Algèbre Linéaire 2

Bernard Le Stum
Université de Rennes 1

Version du 2 décembre 2008

Table des matières

0 Les droites du plan

La notion de *ligne droite* est la base de l'algèbre linéaire. On dit plus simplement *droite* et on précise parfois *droite affine*. En fait, sur une droite affine, on peut marquer des points mais aucun ne joue un rôle privilégié *a priori*. Quand on fixe une origine O, on parle de *droite vectorielle* et on écrit parfois $\overrightarrow{D}$. On peut alors considérer le *vecteur* $\overrightarrow{OP}$ joignant O à P. Et réciproquement, à tout vecteur u de $\overrightarrow{D}$ correspond un unique point P, extrémité du vecteur u d'origine O. On écrit parfois $P = O + u$.

[DESSIN]

Si on fixe un vecteur de base e, tout vecteur u de $\overrightarrow{D}$ est un multiple de e, disons qu'on a $u = x \times e$ avec x un réel. On dit que x est la *composante* du vecteur u dans la base e sur la droite $\overrightarrow{D}$. Si $P = O + u$, on dit alors que x est la coordonnée de P dans le repère (O, u).

[DESSIN]

On s'intéresse à certaines *transformations (dites affines) de la droite* comme les *translations* et les *homothéties*. Une telle transformation est dite *vectorielle* ou *linéaire* si elle fixe l'origine. Une translation correspond à une application de la forme $x \mapsto x + \alpha$. Elle n'est pas linéaire si $\alpha \neq 0$. Une homothétie vectorielle correspond à une application $x \mapsto \lambda x$ avec $\lambda \neq 0$.

[DESSIN]

On peut étudier les droites d'un même *plan*. Comme pour les droites dit *plan vectoriel* si on fixe une *origine* O. Les droites vectorielles sont celles qui passent par O. Par un point, il passe une unique droite parallèle à une autre. Il suit que toute droite affine D s'obtient par translation à partir d'une unique droite vectorielle $\overrightarrow{D}$ passant par O, appelée *direction* de D.

[DESSIN]

On voit assez rapidement qu'il faut fixer deux vecteurs de base e_1 et e_2 pour pouvoir atteindre n'importe quel point du plan à partir de O en ajoutant des multiples de ces vecteurs. Tout vecteur u du plan s'écrira donc de manière unique $x_1e_1 + x_2e_2$ et (x_1, y_2) seront les *composantes* de u dans la *base* (e_1, e_2). Comme ci-dessus, on dira aussi que (x_1, y_2) sont les *coordonnées* du point P, extrémité de u dans le *repÃĺre* (O, e_1, e_2).

[DESSIN]

Si on fixe un repÃÍre, une droite D du plan est caractérisée par son (une) *équation*

$$ax + by = \alpha$$

avec $a \neq 0$ ou $b \neq 0$. En d'autre termes, D corresponds à l'ensemble des solutions (x, y) de cette équation. Et sa direction $\overrightarrow{D}$ est donnée par l'équation *homogène* $ax + by = 0$.

[DESSIN]

Intersecter deux droites revient à résoudre un *système*

$$(S) : \left\{ \begin{array}{l} ax + by = \alpha \\ cx + dy = \beta \end{array} \right. .$$

On trouve soit un point, ou rien, ou une droite. L'intersection des directions s'obtient en résolvant le système homogène

$$(\overrightarrow{S}) : \left\{ \begin{array}{l} ax + by = 0 \\ cx + dy = 0 \end{array} \right. .$$

correspondant.

[DESSIN]

On peut s'intéresser aux *transformations (dites affines) du plan* comme les translations, symétries, homothéties, rotations... Quitte à faire une translation, une telle transformation préserve l'origine et est donnée par une *application linéaire*

$$\varphi : (x, y) \mapsto (ax + by, cx + dy).$$

[DESSIN]

Le *noyau* d'une transformation linéaire est formé de tous les points qui sont envoyés sur l'origine. Cet ensemble correspond au *noyau*

$$\ker \varphi := \{(x, y) \in \mathbf{R}^2, \varphi(x, y) = (0, 0)\}$$

de φ. On retrouve l'ensemble des solutions du système homogène $(\overrightarrow{S})$.

On peut aussi regarder la *projection* du plan sur une droite Δ. Si on demande toujours que l'origine soit préservée, celle-ci correspond à une *forme linéaire*

$$f : (x, y) \to ax + by.$$

[DESSIN]

La droite $\overrightarrow{D}$ d'équation $ax + by = 0$ est alors le noyau de la projection.

On remarque aussi que l'application linéaire φ ci-dessus est donnée par un couple de formes linéaires (f, g). Et que les solutions de $(\overrightarrow{S})$ correspondent à l'intersection des droites données par les noyaux des formes linéaires.

On identifie ainsi implicitement la droite avec $\mathbf{R}$, le plan avec $\mathbf{R}^2$ et on fait de même avec les espaces de dimension supérieure. Les *sous-espaces (affines)* correspondent à des ensembles de solutions de systèmes linéaires. Les transformations géométriques correspondent à des applications linéaires, à translation près. Les problèmes géométriques deviennent des questions algébriques.

Mais on peut aller encore plus loin en utilisant la notation matricielle. On définit un vecteur « colonne » ou un vecteur « ligne » comme un tableau

$$\begin{bmatrix} a \\ b \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} a & b \end{bmatrix}.$$

On dit aussi matrice 2×1 et matrice 1×2. De même, une matrice 2×2 est un tableau

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

On identifiera une matrice 1×1 avec le réel a correspondant mais on pourrait aussi écrire $[a]$. On peut ajouter des matrices de même type en additionnant terme à terme, par exemple

$$A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}, B := \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \quad \text{et alors} \quad A + B = \begin{bmatrix} a+a' & b+b' \\ c+c' & d+d' \end{bmatrix}.$$

On peut aussi multiplier une matrice par une constante, par exemple

$$A := \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{et alors} \quad \lambda A = \begin{bmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{bmatrix}.$$

Mais toute la magie tient dans le fait qu'on peut aussi multiplier un vecteur ligne par un vecteur colonne

$$\begin{bmatrix} \alpha & \beta \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \alpha a + \beta b$$

pour obtenir un scalaire. On peut alors étendre cette méthode pour multiplier une matrice 2×2 à droite par une ligne

$$\begin{bmatrix} \alpha & \beta \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \alpha a + \beta c & \alpha b + \beta d \end{bmatrix}$$

ou à gauche par une colonne

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \alpha c + \beta d \\ \gamma c + \delta d \end{bmatrix}$$

et on peut aussi multiplier les matrices 2×2 entre elles...

Avec ces conventions, on voit ainsi que le système (S) ci-dessus peut se réécrire

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

ou encore

$$AX = B \quad \text{avec} \quad A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad X := \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{et} \quad B := \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$

On a ainsi ramené formellement un problème en deux variables en un problème en une variable ! On dit que A est la matrice du système linéaire. On dit aussi que c'est la *matrice de l'application linéaire* φ ci-dessus. Étudier A revient à étudier φ.

1 Vecteurs dans $\mathbf{R}^n$

Nous allons formaliser un peu tout ce qui précède en oubliant l'origine géométrique des objets. Tant qu'à faire, on va aussi travailler en dimension quelconque.

Définition 1.1 *Si $n \in \mathbf{N}$, alors $\mathbf{R}^n$ désigne l'ensemble des n-uples $(x_1, \ldots, x_n)$ de réels. Un tel n-uple s'appelle aussi un* vecteur *de $\mathbf{R}^n$ et on dit que x_i est sa i-ème* composante *(mais on peut aussi voir $(x_1, \ldots, x_n)$ comme un point de $\mathbf{R}^n$ ...).*

Dans $\mathbf{R}^n$, on dispose d'une addition *(terme à terme)*

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

qui fournit une somme *et d'une* multiplication externe *(terme à terme)*

$$\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$$

qui fournit un produit.

Par exemple, $\mathbf{R}^2$ peut s'identifier au plan comme on l'a vu dans l'introduction, $\mathbf{R}^1$ n'est autre que $\mathbf{R}$ lui même que l'on identifie à la droite et $\mathbf{R}^0$ est le singleton $\{0\}$ (disons, par convention) que l'on peut voir comme un point.

En général, on écrira

$$e_i := (0, \ldots, 0, 1, 0, \ldots, 0)$$

avec le 1 à la i-ème place. On dit parfois que $(e_1, \ldots, e_n)$ est la *base canonique de* $\mathbf{R}^n$ (tout ce qui suit est relatif à cette base). Remarquons en particulier que si $u := (x_1, \ldots, x_n)$, on a

$$x_1 e_1 + \cdots + x_n e_n = (x_1, 0, \ldots, 0) + \cdots + (0, \cdots, 0, x_n) = (x_1, \ldots, x_n) = u.$$

Définition 1.2 *Si $m, n \in \mathbf{N}$, une* matrice $n \times m$ *est un tableau à n lignes et m colonnes*
$$A := [a_{ij}] := \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}.$$
On dit vecteur ligne *si $n = 1$ et* vecteur colonne *si $m = 1$. La* matrice de
$$u := (x_1, \ldots, x_n) \in \mathbf{R}^n$$
est le vecteur colonne
$$[u] := \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$
Enfin, on dit matrice carrée *si $m = n$.*

Pour $m = 0$ ou $n = 0$, on considère par convention qu'il y a une seule matrice, la matrice nulle. Pour $m = 2$ et/ou $n = 2$, on retrouve les matrices à 2 lignes et/ou 2 colonnes introduites plus haut. Et pour $m = n = 1$, on peut identifier matrices et réels (et on le fera).

Si $m, n \in \mathbf{N}$, on note $M_{n\times m}(\mathbf{R})$ l'ensemble des matrices $n \times m$. On écrira simplement $M_n(\mathbf{R})$ si $m = n$.

Définition 1.3 *La* somme *de $A := [a_{ij}]$ et $B := [b_{ij}]$ est $A + B := [a_{ij} + b_{ij}]$ et le* produit *de λ par A est $\lambda A := [\lambda a_{ij}]$.*

Ici encore, ce n'est qu'une généralisation de ce qu'on a vu dans la section précédente. D'autre part, la proposition suivante montre que le passage d'un n-uple au vecteur colonne correspondant est compatible avec nos opérations.

Proposition 1.4 *Si $u, v \in \mathbf{R}^n$, on a $[u+v] = [u]+[v]$. Et si $\lambda \in \mathbf{R}$, on a $[\lambda u] = \lambda[u]$.*

Démonstration : Si $u := (x_1, \ldots, x_n), v := (y_1, \ldots, y_n) \in \mathbf{R}^n$, on a par définition $u + v = (x_1 + y_1, \ldots, x_n + y_n)$ et donc
$$[u + v] := \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}.$$
D'autre part, on a
$$[u] = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad [v] = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \text{et donc} \quad [u + v] = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}.$$

De même, si $u := (x_1, \ldots, x_n) \in \mathbf{R}^n$ et $\lambda \in \mathbf{R}$, on $\lambda u = (\lambda x_1, \ldots, \lambda x_n)$ et donc

$$[\lambda u] := \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix} = \lambda [u].$$

$\square$

Définition 1.5 *Une* forme linéaire *sur* $\mathbf{R}^n$ *est application* $f : \mathbf{R}^n \to \mathbf{R}$ *qui satisfait :*

i) Si $u, v \in \mathbf{R}^n$*, on a* $f(u + v) = f(u) + f(v)$.

ii) Si $u \in \mathbf{R}^n$ *et* $\lambda \in \mathbf{R}$*, alors* $f(\lambda u) = \lambda f(u)$

On note parfois $\check{\mathbf{R}}^n$ l'ensemble de toutes les formes linéaires sur $\mathbf{R}^n$.

Proposition 1.6 *Une forme linéaire sur* $\mathbf{R}^n$ *est une application* $f : \mathbf{R}^n \to \mathbf{R}$ *de la forme*

$$(x_1, \ldots, x_n) \mapsto a_1 x_1 + \cdots + a_n x_n$$

avec $a_1, \ldots, a_n \in \mathbf{R}$.

Démonstration : On se donne tout d'abord

$$f : (x_1, \ldots, x_n) \mapsto a_1 x_1 + \cdots + a_n x_n.$$

Si $u := (\alpha_1, \ldots, \alpha_n)$ et $v := (\beta_1, \ldots, \beta_n)$, on a

$$u + v := (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$$

et donc

$$f(u + v) = a_1(\alpha_1 + \beta_1) + \cdots + a_n(\alpha_n + \beta_n)$$
$$= a_1\alpha_1 + \cdots + a_n\alpha_n + a_1\beta_1 + \cdots + a_n\beta_n = f(u) + f(v).$$

Et de même, si $u := (\alpha_1, \ldots, \alpha_n)$ et $\lambda \in \mathbf{R}$, on a

$$f(\lambda u) = f(\lambda\alpha_1, \ldots, \lambda\alpha_n) = a_1\lambda\alpha_1 + \cdots + a_n\lambda\alpha_n$$

$$= \lambda(a_1\alpha_1 + \cdots + a_n\alpha_n) = \lambda f(u).$$

Réciproquement, on se donne une application f satisfaisant ces conditions et on pose pour tout $i = 1, \ldots, n$, $a_i := f(e_i)$. Rappelons que si $u := (x_1, \ldots, x_n)$, on peut aussi écrire $u = x_1 e_1 + \cdots + x_n e_n$ et on voit alors que

$$f(u) = f(x_1 e_1 + \cdots + x_n e_n) = f(x_1 e_1) + \cdots + f(x_n e_n)$$

$$= x_1 f(e_1) + \cdots + x_n f(e_n) = a_1 x_1 + \cdots + a_n x_n.$$

□

La *matrice* de f est le vecteur ligne

$$[f] := \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}.$$

Par exemple, les projections

$$p_i : u := (x_1, \ldots, x_n) \mapsto x_1$$

sont des formes linéaires. La matrice de p_i est le vecteur ligne $\begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}$ avec 1 à la i-ème place.

Rappelons que si $f, g : I \to \mathbf{R}$ sont deux applications d'un ensemble quelconque vers $\mathbf{R}$, leur *somme* est l'application

$$f + g \quad : \quad \begin{aligned} I &\longrightarrow \mathbf{R} \\ a &\longmapsto f(a) + g(a) \end{aligned}$$

et le *produit* de f par $\lambda \in \mathbf{R}$ est l'application

$$\lambda f \quad : \quad \begin{aligned} I &\longrightarrow \mathbf{R} \\ a &\longmapsto \lambda f(a) \end{aligned}$$

En d'autres termes, on a par définition

$$(f + g)(a) = f(a) + g(a) \quad \text{et} \quad (\lambda f)(a) = \lambda f(a).$$

On appliquera ça au cas $I = \mathbf{R}^n$. En particulier, on voit que si f est la forme linéaire

$$u := (x_1, \ldots, x_n) \mapsto f(u) := a_1 x_1 + \cdots + a_n x_n,$$

on a

$$(a_1 p_1 + \cdots a_n p_n)(u) = a_1 p_1(u) + \cdots a_n p_n(u) = a_1 x_1 + \cdots a_n x_n = f(u).$$

Ceci étant vrai pour tout $u \in \mathbf{R}^n$, on a donc $f := a_1 p_1 + \cdots + a_n p_n$. On dit que $(p_1, \ldots, p_n)$ est la *base canonique* de $\check{\mathbf{R}}^n$.

On montre maintenant que les opérations sur les formes linéaires sont compatibles avec les opérations analogues sur les vecteurs lignes.

Proposition 1.7 *Si f et g sont des formes linéaires sur $\mathbf{R}^n$, alors $f + g$ aussi et on a $[f + g] = [f] + [g]$. De même, si $\lambda \in \mathbf{R}$, alors λf est aussi une forme linéaire sur $\mathbf{R}^n$ et $[\lambda f] = \lambda [f]$.*

Démonstration : On se donne donc deux formes linéaires

$$f : (x_1, \ldots, x_n) \mapsto a_1 x_1 + \cdots + a_n x_n \quad \text{et} \quad g : (x_1, \ldots, x_n) \mapsto b_1 x_1 + \cdots + b_n x_n$$

et on a donc, par définition,

$$(f+g)(x_1,\ldots,x_n) = f(x_1,\ldots,x_n) + g(x_1,\ldots,x_n)$$

$$= a_1x_1 + \cdots + a_nx_n + b_1x_1 + \cdots + b_nx_n = (a_1+b_1)x_1 + \cdots + (a_n+b_n)x_n$$

si bien que $f+g$ est linéaire et

$$[f+g] = \begin{bmatrix} a_1+b_1 & \cdots & a_n+b_n \end{bmatrix} = [f]+[g].$$

De même, si

$$f : (x_1,\ldots,x_n) \mapsto a_1x_1 + \cdots + a_nx_n$$

est une forme linéaire et $\lambda \in \mathbf{R}$, on a

$$(\lambda f)(x_1,\ldots,x_n) = \lambda f(x_1,\ldots,x_n)$$

$$= \lambda(a_1x_1 + \cdots + a_nx_n) = \lambda a_1x_1 + \cdots + \lambda a_nx_n$$

si bien que λf est linéaire et

$$[\lambda f] = \begin{bmatrix} \lambda a_1 & \cdots & \lambda a_n \end{bmatrix} = \lambda[f].$$

$\square$

On peut multiplier les matrices entre elles pourvu qu'elles aient les bonnes dimensions.

Définition 1.8 *Soient $A := [a_{ij}]$ une matrice $n \times m$ et $B := [b_{ij}]$ une matrice $m \times l$. Leur produit AB est la matrice $n \times l$ avec $AB = [c_{ij}]$ où*

$$c_{ij} := \sum_{k=1}^{m} a_{ik}b_{kj} := a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}.$$

On a donc le diagramme suivant

$$\begin{array}{cc} \text{(CARRÉ)} & \left.\overbrace{\begin{bmatrix} b_{11} & \cdots & b_{1l} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{ml} \end{bmatrix}}^{l}\right\} m \\ n\left\{\underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}}_{m}\right. & \left.\underbrace{\begin{bmatrix} c_{11} & \cdots & c_{1l} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nl} \end{bmatrix}}_{l}\right\} n \end{array} .$$

ou chaque entrée de la matrice produit s'obtient en multipliant la ligne correspondante par la colonne correspondante.

On définit aussi les *puissances* d'une matrice carrée en posant

$$A^r = \underbrace{A \times \cdots \times A}, \quad r \text{ fois.}$$

La multiplication d'un vecteur ligne par un vecteur colonne correspond à l'action de la forme linéaire sur le n-uple comme on le voit ci-dessous.

Proposition 1.9 *Si f est une forme linéaire sur $\mathbf{R}^n$ et $u \in \mathbf{R}^n$, on a $[f][u] = f(u)$.*

Démonstration : On se donne donc une forme linéaire f de matrice $[f] = [a_1 \cdots a_n]$ sur $\mathbf{R}^n$ et le vecteur $u = (x_1, \ldots, xn) \in \mathbf{R}^n$. Par définition, on a donc $f(u) = a_1x_1 + \cdots + a_nx_n$ et bien sûr aussi

$$[f][u] = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = a_1x_1 + \cdots + a_nx_n.$$

□

Définition 1.10 *Si f une forme linéaire sur $\mathbf{R}^n$, on dit que*

$$(E) \quad f(x_1, \ldots, x_n) = \alpha$$

est une équation linéaire. Le noyau *de f est*

$$\ker f := \{u \in \mathbf{R}^n, \quad f(u) = 0\}.$$

Si f est non nulle et $\alpha \in \mathbf{R}$, on dit que

$$H := \{u \in \mathbf{R}^n, \quad f(u) = \alpha\}$$

est un hyperplan affine *et que $f(x_1, \ldots, x_n) = \alpha$ est une* équation *de H. Lorsque $\alpha = 0$, on dit que H est un* hyperplan vectoriel.

Bien sûr, dans le cas $n = 2$, un hyperplan correspond à une droite du plan. Pour $n = 1$, c'est un point de la droite. Et dans le cas $n = 3$, on trouve un plan dans l'espace donné par son équation $ax + by + cz = \alpha$. En général, si f est donnée par

$$(x_1, \ldots, x_n) \mapsto a_1x_1 + \cdots + a_nx_n,$$

alors H est *l'ensemble des solutions* de l'équation

$$a_1x_1 + \cdots + a_nx_n = \alpha.$$

Définition 1.11 *Si $f_1, \dots, f_k$ sont des formes linéaires sur $\mathbf{R}^n$ et $\alpha_1, \dots, \alpha_k \in \mathbf{R}$, on dit que*

$$(S) : \begin{cases} f_1(x_1, \dots, x_n) & = & \alpha_1 \\ \vdots & & \vdots \\ f_k(x_1, \dots, x_n) & = & \alpha_k \end{cases}$$

est un système linéaire. Il est dit homogène si $\alpha_1 = \dots = \alpha_m = 0$. Une solution de (S) est un vecteur $(x_1, \dots, x_n)$ qui satisfait toutes les équations. Deux systèmes linéaires sont équivalents *s'ils ont même ensemble de solutions.*

Définition 1.12 *Une partie non vide E de $\mathbf{R}^n$ est un* sous-espace vectoriel *si :*

i) $E \neq \emptyset$

ii) si $u, v \in E$, alors $u + v \in E$

iii) si $u \in E$ et $\lambda \in \mathbf{R}$, alors $\lambda u \in E$.

Proposition 1.13 *Une partie E de $\mathbf{R}^n$ est un sous-espace vectoriel si et seulement si c'est l'ensemble des solutions d'un système linéaire homogène.*

En d'autres termes, un espace affine (resp. vectoriel) est une intersection d'hyperplans affines (resp. vectoriels).

Par convention, $\mathbf{R}^n$ est un sous-espace vectoriel de lui même en tant qu'intersection vide. De même, $\{O\}$ est l'intersection de tous les hyperplans vectoriels (il suffit de prendre les hyperplans des coordonnées). Et bien sûr, tout hyperplan affine (resp. vectoriel) est un sous-espace affine (resp. vectoriel).

Nous énonçons dès maintenant la caractérisation usuelle des sous-espaces vectoriels de $\mathbf{R}^n$ mais nous ne démontrerons que plus tard que la condition est suffisante.

Démonstration : Par hypothèse, E est non vide. De plus, si E est l'intersection d'hyperplans vectoriels définis par des formes linéaires non nulles $f_1, \dots, f_k$, et si $u, v \in E$, on a alors, grâce à la proposition 1.5, pour tout $i = 1, \dots, k$,

$$f_i(u + v) = f_i(u) + f_i(v) = 0 + 0 = 0.$$

et donc $u + v \in E$. De même, si $u \in E$ et $\lambda \in \mathbf{R}$, on a pour tout $i = 1, \dots, k$,

$$f_i(\lambda u) = \lambda f_i(u) = \lambda 0 = 0.$$

et donc $\lambda u \in E$.

Pour la réciproque, on renvoie sur 11.1.

□

On dira aussi qu'une partie non vide E de $\mathbf{R}^n$ est un *sous-espace affine* si c'est l'ensemble des solutions d'un système linéaire (pas nécessairement homogène).

En d'autres termes, on voit donc qu'un espace affine (resp. vectoriel) est une intersection d'hyperplans affines (resp. vectoriels).

Bien sûr, $\mathbf{R}^n$ est un sous-espace vectoriel de lui même en tant qu'intersection vide. De même, $\{O\}$ est l'intersection de tous les hyperplans vectoriels (il suffit de prendre les hyperplans des coordonnées). Et bien sûr, tout hyperplan affine (resp. vectoriel) est un sous-espace affine (resp. vectoriel).

2 Applications linéaires de $\mathbf{R}^m$ vers $\mathbf{R}^n$

On généralise maintenant la notion de somme et de produit par un scalaire aux applications d'un ensemble quelconque I vers $\mathbf{R}^n$.

Définition 2.1 *Si $f, g : I \to \mathbf{R}^n$ sont deux applications de I vers $\mathbf{R}^n$, leur* somme *est l'application*

$$f+g \quad : \quad \begin{array}{rcl} I & \longrightarrow & \mathbf{R}^n \\ a & \longmapsto & f(a)+g(a). \end{array}$$

De même, le produit *de f par $\lambda \in \mathbf{R}$ est l'application*

$$\lambda f \quad : \quad \begin{array}{rcl} I & \longrightarrow & \mathbf{R}^n \\ a & \longmapsto & \lambda f(a). \end{array}$$

Définition 2.2 *Si $f : I \to \mathbf{R}^n$ est une application, ses composantes sont les applications $f_i := p_i \circ f : I \to \mathbf{R}$.*

On voit donc que $f(a) = (f_1(a), \cdots, f_n(a))$ lorsque $a \in I$.

Lemme 2.3 *Si $f, g : I \to \mathbf{R}^n$ sont deux applications de I vers $\mathbf{R}^n$, les composantes de $f+g$ sont les sommes $f_i + g_i$ des composantes. Et si $\lambda \in \mathbf{R}$, les composantes de λf sont les produits λf_i des composantes par λ.*

Définition 2.4 *Une application $f : \mathbf{R}^m \to \mathbf{R}^n$ est* linéaire *si :*

i) Si $u, v \in \mathbf{R}^m$, on a $f(u+v) = f(u) + f(v)$.

ii) Si $u \in \mathbf{R}^m$ et $\lambda \in \mathbf{R}$, alors $f(\lambda u) = \lambda f(u)$

Proposition 2.5 *Une application $f : \mathbf{R}^m \to \mathbf{R}^n$ est linéaire si et seulement si ses composantes sont des formes linéaires.*

Démonstration : On note comme d'habitude, pour tout $i = 1, \ldots, n$, f_i la i-ème composante de f. Pour $u, v \in \mathbf{R}^m$, on a

$$f(u+v) = (f_1(u+v), \ldots, f_n(u+v))$$

et

$$\begin{aligned} f(u)+f(v) &= (f_1(u),\ldots,f_n(u))+(f_1(v),\ldots,f_n(v)) \\ &= (f_1(u)+f_1(v),\ldots,f_n(u)+f_n(v)). \end{aligned}$$

De même, si $u \in \mathbf{R}^m$ et $\lambda \in \mathbf{R}$, on a

$$f(\lambda u) = (f_1(\lambda u),\ldots,f_n(\lambda u))$$

et

$$\lambda f(u) = \lambda(f_1(u),\ldots,f_n(u)) = (\lambda f_1(u),\ldots,\lambda f_n(u)).$$

On voit donc que f satisfait les conditions ci-dessus si et seulement si toutes ses composantes les satisfont aussi. L'assertion résulte donc du lemme 1.5.

□

Dans le cas ou $n = 1$, une application linéaire n'est rien d'autre qu'une forme linéaire sur $\mathbf{R}^m$. Pour $m = n = 2$, on retrouve les transformations de plan étudiées dans la précédente section.

Définition 2.6 *Soit $f : \mathbf{R}^m \to \mathbf{R}^n$ une application linéaire. Si pour tout $i = 1,\ldots,n$, on a*

$$[f_i] =: \begin{bmatrix} a_{i1} & \cdots & a_{im} \end{bmatrix},$$

alors la matrice *de f est*

$$A := [f] := \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \in M_{n\times m}(\mathbf{R}).$$

On peut remarquer que la matrice de f s'obtient en superposant les vecteurs lignes des f_i comme ceci :

$$f = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}.$$

Si on note $e_j := (0,\ldots,0,1,0,\ldots,0)$ avec le 1 à la j-ème place, on a

$$f(e_j) = (a_{1j},\ldots,a_{nj})$$

et on peut donc aussi écrire f en alignant les vecteurs colonnes des $f(e_j)$:

$$[f] = \begin{bmatrix} f(e_1) & \cdots & f(e_m) \end{bmatrix}.$$

Remarquons enfin que la matrice de l'identité est la *matrice unité*

$$I_n := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Définition 2.7 *Le* noyau *de* $f : \mathbf{R}^m \to \mathbf{R}^n$ *est* $\ker f = \{u \in \mathbf{R}^m, f(u) = 0\}$.

En d'autres termes, le *noyau* de f est l'intersection des noyaux de ses composantes f_i : on a

$$\ker f = \cap_{i=1}^n \ker f_i.$$

Remarquons aussi qu'un sous-espace vectoriel E de $\mathbf{R}^n$ n'est autre que le noyau d'une application linéaire. Il s'agit des n-uples qui sont envoyés par f sur l'origine de $\mathbf{R}^n$.

La compatibilité vue plus haut des opérations sur les matrices avec les formes linéaires s'étend en fait au cas des applications linéaires.

Proposition 2.8 *Soit* $f : \mathbf{R}^m \to \mathbf{R}^n$ *une application linéaire.*

i) Si $u \in \mathbf{R}^m$, *on a* $[f][u] = [f(u)]$.

ii) Si $g : \mathbf{R}^m \to \mathbf{R}^n$ *est une autre application linéaire,* $f + g$ *est linéaire et* $[f + g] = [f] + [g]$.

iii) Si $\lambda \in \mathbf{R}$, *alors* λf *est linéaire et* $[\lambda f] = \lambda[f]$.

iv) Si $g : \mathbf{R}^l \to \mathbf{R}^m$ *est une autre application linéaire, alors* $f \circ g$ *est linéaire et* $[f \circ g] = [f][g]$.

Démonstration : Cela se fait encore en se ramenant aux composantes. Tout d'abord, on a si $u \in \mathbf{R}^m$,

$$[f][u] = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} [u] = \begin{bmatrix} f_1(u) \\ \vdots \\ f_n(u) \end{bmatrix} = [f(u)]$$

Ensuite, si $g : \mathbf{R}^m \to \mathbf{R}^n$ est une autre application linéaire, alors $f + g$ est linéaire car ses composantes $f_i + g_i$ le sont et on a

$$[f + g] = \begin{bmatrix} f_1 + g_1 \\ \vdots \\ f_n + g_n \end{bmatrix} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} + \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix} = [f] + [g]$$

De même, si $\lambda \in \mathbf{R}$, alors λf est linéaire car ses composantes λf_i le sont et on a

$$[\lambda f] = \begin{bmatrix} \lambda f_1 \\ \vdots \\ \lambda f_n \end{bmatrix} = \lambda \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} = \lambda[f]$$

On termine par la composition avec une application $g : \mathbf{R}^l \to \mathbf{R}^m$. On suppose tout d'abord que $n = 1$ si bien que f est une forme linéaire

$$(x_1, \ldots, x_m) \mapsto a_1x_1 + \cdots + a_mx_m.$$

et on a donc pour tout $u \in \mathbf{R}^l$,

$$(f \circ g)(u) = f(g(u)) = a_1g_1(u) + \cdots + a_mg_m(u) = (a_1g_1 + \cdots + a_mg_m)(u).$$

En d'autres termes, $f \circ g$ est une forme linéaire et $f \circ g = a_1 g_1 + \cdots + a_m g_m$. On a donc bien

$$\begin{aligned}[f \circ g] &= [a_1 g_1 + \cdots + a_m g_m] = a_1[g_1] + \cdots + a_m[g_m] \\ &= \begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix} \begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix} = [f][g].\end{aligned}$$

En général, on remarque que les composantes de $f \circ g$ sont

$$p_i \circ (f \circ g) = (p_i \circ f) \circ g = f_i \circ g$$

et il suit que $f \circ g$ sera linéaire et que

$$[f \circ g] = \begin{bmatrix} f_1 \circ g \\ \vdots \\ f_n \circ g \end{bmatrix} = \begin{bmatrix} [f_1][g] \\ \vdots \\ [f_n][g] \end{bmatrix} = \begin{bmatrix} [f_1] \\ \vdots \\ [f_n] \end{bmatrix} [g] = [f][g].$$

□

Corollaire 2.9 *On considère*

i) le système

$$(S) : \left\{ \begin{array}{ccc} a_{11}x_1 + \cdots + a_{1m}x_m & = & b_1 \\ \vdots & & \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m & = & b_n \end{array} \right. ,$$

ii) l'application linéaire

$$f \quad : \quad \begin{array}{ccc} \mathbf{R}^m & \longrightarrow & \mathbf{R}^n \\ (x_1, \ldots, x_m) & \longmapsto & (a_{11}x_1 + \cdots + a_{1m}x_m, \ldots, a_{n1}x_1 + \cdots + a_{nm}x_m) \end{array}$$

ainsi que les vecteurs $u := (x_1, \ldots, x_n)$ *et* $v := (b_1, \ldots, b_n)$.

iii) les matrices

$$A := \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}, \quad B := \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \quad \text{et} \quad X := \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}.$$

Alors, les assertions suivantes sont équivalentes

i) $(x_1, \ldots, x_n)$ *est solution du système* (S)

ii) $f(u) = v$.

iii) $AX = B$

Démonstration : Le deux premières conditions sont clairement équivalentes et il résulte de la proposition précédente qu'elles sont aussi équivalentes à la troisième.

□

Définition 2.10 *Avec les notations du corollaire, on dit que la matrice matrice* $n \times (m+1)$ *obtenue en* concaténant A *et* B*, c'est à dire*

$$A := \begin{bmatrix} a_{11} & \cdots & a_{1m} & b_1 \\ \vdots & & \vdots & \\ a_{n1} & \cdots & a_{nm} & b_n \end{bmatrix},$$

est la matrice du système*.*

Définition 2.11 *Une matrice* $A \in M_n(\mathbf{R})$ *est* inversible *s'il existe une matrice que l'on note* $A^{-1} \in M_n(\mathbf{R})$ *telle que* $AA^{-1} = A^{-1}A = I$*.*

Remarquons que A^{-1}, si elle existe, est unique (exercice !). Mais la matrice

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

par exemple, n'est pas inversible (exercice !), contrairement à

$$A := \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$$

dont l'inverse est

$$A^{-1} = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix} \text{(exercice !).}$$

Proposition 2.12 *Soit* $f : \mathbf{R}^n \to \mathbf{R}^n$ *une application linéaire. Alors,* f *est bijective si et seulement si* $[f]$ *est inversible et dans ce cas,* f^{-1} *est linéaire et* $[f^{-1}] = [f]^{-1}$*.*

Démonstration : Si f est bijective, il existe pour tout $j = 1, \ldots, n$ un unique vecteur $v_j := (b_{1j}, \ldots, b_{nj})$ tel que $f(v_j) = e_i$ et on peut donc considérer la matrice $[b_{ij}]$. Si g est l'application linéaire correspondante, on a $g(e_j) = v_j$ pour tout $j = 1, \ldots, n$. Si $u := (a_1, \ldots, a_n) \in \mathbf{R}^n$, on sait que $u = a_1e_1 + \cdots + a_ne_n$ et il suit que

$$g(u) = g(a_1e_1 + \cdots + a_ne_n)$$

$$= a_1g(e_1) + \cdots + a_ng(e_n) = a_1v_1 + \cdots + a_nv_n$$

si bien que

$$(f \circ g)(u) = f(a_1v_1 + \cdots + a_nv_n)$$

$$= a_1f(v_1) + \cdots + a_nf(v_n) = a_1e_1 + \cdots + a_ne_n = u.$$

On voit donc que g n'est autre que l'application réciproque f^{-1} de f qui et donc bien linéaire, et comme $f^{-1} \circ f = f \circ f^{-1} = \mathrm{Id}_{\mathbf{R}^n}$, on a $[f][f^{-1}] = [f^{-1}][f] = I_n$.

Réciproquement, si $[f]$ est inversible, on considère l'application linéaire g dont la matrice est $[f]^{-1}$. On a alors $[f \circ g] = [f][g] = [f][f]^{-1} = I_n$ et il suit que $f \circ g = \mathrm{Id}_{\mathbf{R}^n}$ et de la même manière, on obtient que $g \circ f = \mathrm{Id}_{\mathbf{R}^n}$. Il suite que g est la réciproque de f et donc que f est bijective.

$\square$

Si A est une matrice inversible et $AX = B$, alors $X = A^{-1}B$. Donc pour résoudre un système carré d'ordre n,

$$(S) : \left\{ \begin{array}{ccc} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n & = & b_n \end{array} \right. ,$$

et si on sait que sa matrice $A := [a_{ij}]$ est inversible et que $A^{-1} = [c_{ij}]$, on aura une unique solution donnée par

$$(S) : \left\{ \begin{array}{ccc} x_1 & = & c_{11}b_1 + \cdots + c_{1n}b_n \\ \vdots & & \vdots \\ x_n & = & c_{n1}b_1 + \cdots + c_{nn}b_n \end{array} \right. .$$

En pratique, ça ne nous avance pas à grand chose car on ne sait pas vraiment inverser des matrices sans résoudre de système *avant*.

Mais par exemple, pour résoudre le système

$$\left\{ \begin{array}{ccc} 5x + 2y & = & 2 \\ 2x + y & = & 1 \end{array} \right. ,$$

qui peut s'écrire $AX = B$, il suffit de calculer

$$X = A^{-1}B = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

La dernière partie de cette section est consacrée à la *méthode du pivot de Gauss* pour résoudre les systèmes linéaires.

Définition 2.13 *Les* opérations élémentaires *sur les lignes d'une matrice sont les suivantes*

i) Échanger deux lignes

ii) Multiplier une ligne par une constante non nulle

iii) Ajouter à une ligne un multiple d'une autre ligne

Le but (ultime) de la méthode de Gauss est de transformer la matrice originale en une matrice dite échelonnée :

$$\begin{bmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & & 0 & * & \cdots & * \\ \vdots & & \vdots & 0 & \cdots & \cdots & 0 & 1 & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & & \vdots & 0 & \ddots & 0 & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & & \vdots & \vdots & & 1 & * & \cdots & * \\ \vdots & & \vdots & \vdots & & & \vdots & \vdots & & 0 & \cdots & \cdots & 0 \\ \vdots & & \vdots & \vdots & & & \vdots & \vdots & & \vdots & & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 & 0 & & 0 & \cdots & \cdots & 0 \end{bmatrix}$$

sur laquelle on peut “lire” les solutions. Par exemple, avec un système de trois équations à 3 inconnues, on essaiera de faire apparaître successivement

$$\begin{bmatrix} 1 & & 0 & \rightarrow & 0 & \alpha \\ \downarrow & & & & \uparrow & \\ 0 & & 1 & & 0 & \beta \\ \downarrow & \nearrow & \downarrow & & \uparrow & \\ 0 & & 0 & \leftarrow & 1 & \gamma \end{bmatrix}.$$

auquel cas, on aura l’unique solution $x = \alpha$, $y = \beta$ et $z = \gamma$.

Enfin, on peut utiliser la méthode du pivot de Gauss pour trouver l’inverse d’une matrice.

Définition 2.14 *Une matrice élémentaire est une matrice obtenue en effectuant une opération élémentaire sur les lignes d’une matrice unité I.*

Par exemple, les matrices 2×2 élémentaires sont les matrices de la forme

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ \lambda & 1 \end{bmatrix}, \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}.$$

Proposition 2.15 *Effectuer une opération élémentaire sur les lignes d’une matrice revient à multiplier à gauche par la matrice élémentaire correspondante.*

Démonstration : C’est un peu laborieux mais allons y. En général, les lignes de la matrice associée à une application linéaire f sont les vecteurs lignes $[f_1], \ldots, [f_n]$ ou $f_1, \ldots, f_n$ désignent les composantes de f. En particulier, comme les composantes de $\mathrm{Id}_{\mathbf{R}^n}$ sont les projections $p_1, \ldots, p_n$, on voit que les lignes de I_n sont les matrices $[p_1], \ldots, [p_n]$. Et comme on a, pour tout $i = 1, \ldots n$, $p_i \circ f = f_i$, on voit que $[p_i][f] = [f_i]$. En d’autres termes, la i-ème ligne d’une matrice s’obtient en multipliant à gauche par la i-ème ligne de la matrice unité.

On considère alors les différents types d’opération élémentaire. Si on échange $[p_i]$ et $[p_j]$, ça va bien échanger $[f_i]$ et $[f_j]$. Et si on multiplie $[p_i]$ par λ, ça aura clairement le même effet sur $[f_i]$. Enfin, si on ajoute $\lambda[p_j]$ à $[p_i]$ et qu’on multiplie ensuite par f, on obtient bien

$$\begin{aligned} ([p_i] + \lambda[p_j])[f] = [p_i + \lambda p_j][f] = [(p_i + \lambda p_j) \circ f] = [p_i \circ f + \lambda p_j \circ f] \\ = [p_i \circ f] + \lambda[p_j \circ f] = [f_i] + \lambda[f_j]. \end{aligned}$$

$\square$

Par exemple, on a

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix},$$

$$\begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \lambda a & \lambda b \\ c & d \end{bmatrix},$$

et

$$\begin{bmatrix} 1 & 0 \\ \lambda & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c + \lambda a & d \end{bmatrix}.$$

Proposition 2.16 *Effectuer une opération élémentaire sur les lignes (de la matrice) d'un système ne change pas l'ensemble des solutions du système.*

Démonstration : Il suffit de montrer que chacune de ces opérations est réversible. Si on a échangé deux lignes, il suffit de les échanger à nouveau. Si on a multiplié une ligne par une constante non nulle λ, il suffit de multiplier celle-si par $1/\lambda$ (d'où l'importance de l'hypothèse $\lambda \neq 0$). Enfin, si on a rajouté à la j-ème ligne le produit par λ de la i-ème, il suffit de retrancher ce même multiple, c'est à dire, ajouter le produit par $-\lambda$ de la i-ème ligne à la j-ème.

□

Proposition 2.17 *Si on applique à I les mêmes opérations élémentaires, et dans le même ordre, que celles qui permettent de transformer A en I, on obtient A^{-1}.*

Démonstration : On suppose donc que l'on obtient I en multipliant à gauche A successivement par des matrices élémentaires $E_1, \ldots, E_r$ et on note B la matrice obtenue en multipliant I successivement par les mêmes matrices dans le même ordre. En d'autres termes, on a

$$E_r \cdots E_2 E_1 A = I \quad \text{et} \quad B := E_r \cdots E_2 E_1 I.$$

On remarque que les matrices élémentaires sont inversibles (leur inverse étant la matrice correspondant à l'opération élémentaire inverse). Il suit que B est inversible et comme on a $BA = I$, A est aussi inversible d'inverse B.

□

En pratique, on applique la méthode du pivot de Gauss à la matrice $[A\ I]$ obtenue en concaténant A et I. Par exemple, pour inverser la matrice

$$A := \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix},$$

on retranche deux fois la première ligne à la seconde dans

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 3 & 0 & 1 \end{bmatrix}$$

pour obtenir

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & -2 & 1 \end{bmatrix}$$

et on ajoute deux fois la seconde à la première avant de multiplier la ligne du bas par -1, ce qui donne

$$\begin{bmatrix} 1 & 0 & -3 & 2 \\ 0 & 1 & 2 & -1 \end{bmatrix}$$

et on a donc

$$A^{-1} = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix}.$$

3 Espaces vectoriels sur un corps

Dans les sections précédentes, on peut remplacer partout $\mathbf{R}$ par un autre corps comme celui des complexes $\mathbf{C}$, des rationnels $\mathbf{Q}$, le corps à deux éléments $\mathbf{F}_2$ ou encore le corps des fractions rationnelles $\mathbf{R}(X)$. Dans la suite, on le désignera par K. Il faut faire attention que des notions comme $a \geq b$ ou $a/2$ n'ont pas toujours une sens dans ce corps K. Par exemple, on ne peut pas dire si i est plus grand que 1 dans $\mathbf{C}$ et on ne peut pas diviser par 2 dans $\mathbf{F}_2$. On ne peut pas non plus extraire de racine en général...

La structure d'espace vectoriel est apparue de manière récurrente depuis le début de ce cours et nous allons la formaliser maintenant.

Définition 3.1 *Un* espace vectoriel *sur K est un ensemble E muni d'une* addition

$$\begin{aligned} E \times E &\longrightarrow E \\ (u,v) &\longmapsto u+v \end{aligned}$$

fournissant une somme *telle que*

i) si $u, v \in E$, alors $u + v = v + u$ (commutativité)

ii) si $u, v, w \in E$, alors $(u + v) + w = u + (v + w)$ (associativité)

iii) il existe $0 \in E$ tel que pour tout $u \in E$, $u + 0 = u$ (zéro)

iv) si $u \in E$, il existe $-u \in E$ tel que $u + (-u) = 0$ (opposé)

et d'une multiplication externe

$$\begin{aligned} K \times E &\longrightarrow E \\ (\lambda,u) &\longmapsto \lambda u \end{aligned}$$

fournissant un produit *telle que*

v) si $u \in E$, alors $1u = u$

vi) si $u \in E$ et $\lambda, \mu \in K$, alors $(\lambda + \mu)u = \lambda u + \mu u$

vii) si $u \in E$ et $\lambda, \mu \in K$, alors $\lambda(\mu u) = (\lambda\mu)u$

viii) si $u, v \in E$ et $\lambda \in K$, alors $\lambda(u + v) = \lambda u + \lambda v$.

Les éléments de E s'appelleront des *vecteurs* et ceux de K des *scalaires*. On écrit parfois 0_E pour préciser qu'il s'agit du zéro de E.

Notations : si $u, v, w \in E$, on écrit $u + v + w := (u + v) + w$; si $u, v \in E$, on écrit $u - v := u + (-v)$; si $u \in E$ et $\lambda, \mu \in K$, on écrit $\lambda\mu u := (\lambda\mu)u$.

Nos premiers exemples viennent de la section précédente.

Tout d'abord, K^n (avec les opérations définies plus haut) est un espace vectoriel comme on le vérifie aisément. Plus généralement, tout sous espace vectoriel de K^n (définition ci-dessus), muni des lois induites, est un espace vectoriel. Cela vaut en particulier pour les hyperplans et donc pour les droites du plan. Mais l'espace $\check{K}^n$ des formes linéaires sur K (définition ci-dessus) est aussi un espace vectoriel pour les opérations habituelles. De même que $M_{n\times m}(K)$ et donc en particulier, les vecteurs lignes ou les vecteurs colonnes de longueur donnée.

Si I est un ensemble quelconque, l'ensemble K^I des applications de I dans K munis des opérations définies plus haut est un espace vectoriel. Si I est un intervalle de $\mathbf{R}$, les fonctions continues à valeurs dans $\mathbf{R}$ définies sur I forment un espace vectoriel $\mathcal{C}(I)$ pour les lois induites. Il en va de même de l'ensemble $\mathcal{C}^1(I)$ des fonctions continûment différentiables lorsque I est ouvert. On peut aussi regarder les fonctions absolument intégrables, etc.

Les familles d'éléments de K indexées par un ensemble I forment aussi un espace vectoriel que l'on identifiera à K^I : il revient au même de se donner la famille $(x_i)_{i\in I}$ ou l'application

$$\begin{aligned} I &\longrightarrow K \\ i &\longmapsto x_i \end{aligned}$$

L'addition et la multiplication se font terme à terme. Comme cas particulier, on trouve l'espace vectoriel $K^{\mathbf{N}}$ des suites indexées par $\mathbf{N}$. On peut aussi regarder les suites convergentes dans $\mathbf{R}$, de même que les suites absolument sommables, etc. On peut aussi considérer les suites de Cauchy dans $\mathbf{Q}$ qui forment un espace vectoriel sur $\mathbf{Q}$.

Les polynômes sur K forment aussi un espace vectoriel $K[X]$.

Il y a cependant des exemples exotiques. On peut munir $\mathbf{R}_{>0}$ de l'addition $u \overrightarrow{+} v = uv$ et de la multiplication externe sur $\mathbf{R}$: $\lambda \overrightarrow{\cdot} u := u^\lambda$. On peut aussi munir l'intervalle $]-\pi/2, \pi/2[$ de

$$u \overrightarrow{+} v := \arctan(\tan u + \tan v)$$

et

$$\lambda \overrightarrow{\cdot} u := \arctan(\lambda \tan u)).$$

Aussi, on peut voir $\mathbf{C}$ comme espace vectoriel sur lui même, sur $\mathbf{R}$ ou sur $\mathbf{Q}$. Enfin, un espace vectoriel sur $\mathbf{F}_2$ est tout simplement un groupe abélien (première partie de la définition d'un espace vectoriel) dans lequel on a toujours $a + a = 0$.

Lemme 3.2 *Soit E un espace vectoriel sur K.*

i) Il existe un unique élément $0 \in E$ satisfaisant la condition donnée.

ii) étant donné un élément $u \in E$ existe un unique élément $-u$ satisfaisant la condition donnée.

iii) Si $u, v \in E$, la condition $u + v = u$ implique $v = 0$.

Démonstration : Notons dès à présent que la troisième assertion implique la première. Notons aussi que si $u, v \in E$, alors

$$v = v + 0 = v + (u + (-u)) = (v + u) + (-u) = (u + v) + (-u).$$

On montre maintenant la seconde assertion : si $u + v = 0$, alors

$$v = 0 + (-u) = (-u) + 0 = -u.$$

Et on montre aussi la troisième : si $u + v = u$, alors

$$v = u + (-u) = 0.$$

□

Proposition 3.3 *Soit E un espace vectoriel sur K.*

i) Si $u \in E$ et $\lambda \in K$, alors

$$\lambda u = 0 \Leftrightarrow (\lambda = 0 \text{ ou } u = 0),$$

ii) si $u \in E$ et $\lambda \in K$, alors $\lambda(-u) = (-\lambda)u = -(\lambda u)$,

iii) si $u, v \in E$ et $\lambda \in K$, alors $\lambda(u - v) = \lambda u - \lambda v$,

iv) si $u \in E$ et $\lambda, \mu \in K$, alors $(\lambda - \mu)u = \lambda u - \mu u$,

v) si $u, v \in E$ et $\lambda \in K^*$ sont tels que $\lambda u = \lambda v$, alors $u = v$

Démonstration : On a

$$0u = (0 + 0)u = 0u + 0u$$

et donc $0u = 0$. On montre de même que $\lambda 0 = 0$. Réciproquement, supposons que $\lambda u = 0$ avec $\lambda \neq 0$. Alors,

$$u = 1u = (\lambda^{-1}\lambda)u = \lambda^{-1}(\lambda u) = \lambda^{-1}0 = 0.$$

D'où la première assertion.

On a

$$\lambda u + \lambda(-u) = \lambda(u + (-u)) = \lambda 0 = 0$$

et donc $\lambda(-u) = -(\lambda u)$. De même, on a

$$\lambda u + (-\lambda)u = (\lambda + (-\lambda))u = 0u = 0$$

et donc $(-\lambda)u = -(\lambda u)$, ce qui nous donne la seconde assertion.

On poursuit. On a

$$\lambda(u-v) = \lambda(u+(-v)) = \lambda u + \lambda(-v) = \lambda u + (-\lambda v) = \lambda u - \lambda v$$

et de la même manière,

$$(\lambda - \mu)u = \lambda u + (-\mu)u = \lambda u + (-\mu u) = \lambda u - \mu u.$$

Enfin, si $\lambda u = \lambda v$, alors

$$\lambda(u-v) = \lambda u - \lambda v = 0$$

et si $\lambda \neq 0$, on a nécessairement $u - v = 0$ et alors $u = v$.

□

On peut maintenant définir la notion de sous-espace vectoriel d'un espace vectoriel quelconque.

Définition 3.4 *Un* sous espace vectoriel *d'un espace vectoriel E est une partie F de E telle que*

i) $F \neq \emptyset$

ii) *si* $u, v \in F$, *alors* $u + v \in F$

iii) *si* $\lambda \in K$ *et* $u \in F$, *alors* $\lambda u \in F$.

Par exemple, tout espace vectoriel E est un sous-espace vectoriel de lui même et la partie réduite au vecteur nul $\{O\}$ est un sous-espace vectoriel de E.

On dispose bien sûr des sous-espaces vectoriels de K^n décrits dans la section précédente, et donc en particulier des droites vectorielles du plan et plus généralement des hyperplans.

Si on regarde nos exemples d'espaces vectoriels, on voit que si I est un intervalle de $\mathbf{R}$, alors $\mathcal{C}^1(I)$ est un sous-espace vectoriel de $\mathcal{C}(I)$ qui est lui même un sous-espace vectoriel de $\mathbf{R}^I$. Ou encore, on voit que les suites convergentes forment un sous-espace vectoriel de $\mathbf{R}^{\mathbf{N}}$. Aussi, les polynômes de degré au plus N forment un sous-espace vectoriel $K[X]_{\leq N}$ de $K[X]$.

Mais il faut être prudent car, dans $\mathbf{R}^2$, la droite d'équation $x + y = 1$ n'est pas un sous-espace vectoriel, le cercle $x^2 + y^2 = 1$ non plus et l'équation $xy = 0$ ne définit pas un sous-espace vectoriel, c'est l'union de deux droites. Attention cependant, les solutions de $x^2 + y^2 = 0$ forment bien un sous-espace vectoriel de $\mathbf{R}^2$. Mais pas dans $\mathbf{C}^2$ (vu comme espace vectoriel sur $\mathbf{C}$).

Notons enfin que sur $\mathbf{F}_2$, la troisième condition est toujours satisfaite si les deux premières le sont.

Proposition 3.5 *Soit E un espace vectoriel sur K.*

i) Si F est un sous-espace vectoriel d'un espace vectoriel E, alors l'addition et la multiplication par les scalaires font de F un espace vectoriel sur K.

ii) Un sous ensemble F de E est un sous espace vectoriel si et seulement si

(a) $0 \in F$

(b) lorsque $u, v \in F$ et $\lambda, \mu \in K$, on a $\lambda u + \mu v \in F$.

Démonstration : Pour la première assertion, toutes les propriété sont héritées de E à part les deux qui font intervenir la notion d'existence. Plus précisément, il suffit de montrer que le vecteur nul est dans F et que l'opposé d'un élément de F est toujours dans F. Or, si $u \in F$, on a bien $-u = (-1)u \in F$ et il suit, comme F est non vide, que l'on a aussi $0 = u + (-u) \in F$.

Montrons maintenant la seconde assertion. On vient de voir que si F est un sous-espace vectoriel de E, la condition a) est toujours satisfaite et on montre facilement que la condition b) l'est aussi : si $u, v \in F$ et $\lambda, \mu \in K$, on a $\lambda u \in F$ et $\mu v \in F$ si bien que $\lambda u + \mu v \in F$. Réciproquement, si ces conditions sont satisfaites, on voit que F est non vide et les deux autres propriétés s'obtiennent à partir de la condition b) en faisant tout d'abord $\lambda = \mu = 1$ pour la première puis $\mu = 0$ pour l'autre.

□

Attention cependant car $\mathbf{R}_{>0}$ n'est pas un sous-espace vectoriel de $\mathbf{R}$ bien que, comme on l'a vu, $\mathbf{R}_{>0}$ est un espace vectoriel pour des lois bien choisies.

4 Applications linéaires

Nous pouvons maintenant considérer des applications linéaires entre différents espaces vectoriels (sur un même corps K) :

Définition 4.1 *Une application $f : E \to F$ entre deux espaces vectoriels sur K est* linéaire *si elle satisfait :*

i) Si $u, v \in E$, on a $f(u + v) = f(u) + f(v)$.

ii) Si $u \in E$ et $\lambda \in \mathbf{R}$, alors $f(\lambda u) = \lambda f(u)$

Leur ensemble se note $L(E, F)$.

Si $F = E$, on dit que f est un endomorphisme *de E (ou un* opérateur *sur E) et on note $L(E)$ leur ensemble. Si $F = K$, on dit que f est une* forme linéaire *sur E et on note $\check{E}$ leur ensemble, aussi appelé* dual *de E. Une application linéaire bijective est un* isomorphisme*. Un endomorphisme qui est aussi un isomorphisme est un* automorphisme*. Leur ensemble se note $GL(E)$.*

En général, on dispose toujours de l'identité

$$Id_E \quad : \quad \begin{array}{l} E \longrightarrow E \\ x \longmapsto x \end{array}$$

qui est un automorphisme et de l'application nulle

$$0_{E,F} \quad : \quad \begin{array}{rcl} E & \longrightarrow & F \\ x & \longmapsto & 0 \end{array}$$

qui est l'unique application linéaire constante.

On définit une *homothétie* comme une application de E dans lui même de la forme $u \longmapsto ku$ avec $k \neq 0, 1$. C'est un automorphisme. Aussi, si F est un sous-espace vectoriel de E, l'application d'inclusion $F \hookrightarrow E$ est linéaire.

Un endomorphisme f de E est une *symétrie* si $f \circ f = \mathrm{Id}$ et une *projection* si $f \circ f = f$. Dans le plan, on peut considérer par exemple la symétrie par rapport à l'origine ou a une droite vectorielle, ou la projection sur une droite vectorielle par rapport à une autre.

Nous avons vu que les applications linéaires $f : K^m \to K^n$ sont les applications de la forme

$$(x_1, \ldots, x_m) \mapsto (a_{11}x_1 + \cdots + a_{1m}x_m, \ldots, a_{n1}x_1 + \cdots + a_{nm}x_m).$$

En d'autre terme, on a une bijection

$$\begin{array}{rcl} L(K^m, K^n) & \xrightarrow{\simeq} & M_{n\times m}(K) \\ f & \longmapsto & [f] \end{array}$$

Bien sûr, cela s'applique en particulier aux formes linéaires.

Comme cas particulier, on peut regarder les rotations

$$(x, y) \mapsto (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta).$$

de $\mathbf{R}^2$. Si on identifie comme d'habitude $\mathbf{R}^2$ et $\mathbf{C}$, une rotation n'est autre que l'homothétie $z \mapsto e^{i\theta}z$ de $\mathbf{C}$ vu comme espace vectoriel sur lui même.

Mais les applications de $\mathbf{R}^2$ dans lui même :

$$(x, y) \mapsto (x + 1, y - 1) \quad \text{et} \quad (x, y) \mapsto (x^2, y^2),$$

par exemple, ne sont pas linéaires. Par contre, l'application

$$\ln : (\mathbf{R}_{>0}, \overrightarrow{+}, \overrightarrow{\cdot}) \to (\mathbf{R}, +, .)$$

est non seulement linéaire mais c'est un isomorphisme !

Si I est un ensemble et $x \in I$, on peut regarder l'évaluation en x,

$$\begin{array}{rcl} K^I & \longrightarrow & K \\ \varphi & \longmapsto & \varphi(x) \end{array} .$$

C'est une forme linéaire. Par restriction, c'est aussi le cas des applications d'évaluations $\mathcal{C}(I) \to \mathbf{R}$ ou $\mathcal{C}^1(I) \to \mathbf{R}$ si I est un intervalle (ouvert) de $\mathbf{R}$.

Alternativement, si on voit une application sur I comme une famille indexée par I, cela revient à envoyer une famille sur sa i-ème composante. Par exemple, on peut considérer la forme linéaire sur $K^{\mathbf{N}}$ qui envoie la suite $(x_n)_{n\in\mathbf{N}}$ sur x_0, ou sur un autre terme... Plus subtilement, on peut considérer la forme linéaire qui envoie une suite convergente sur sa limite. Ou même l'application qui envoie une suite de Cauchy dans $\mathbf{Q}$ sur sa limite dans $\mathbf{R}$ qui est donc $\mathbf{Q}$-linéaire.

Deux exemples fondamentaux viennent de l'analyse, il s'agit de la dérivation

$$\begin{array}{rcl} d : \mathcal{C}^1(I) & \longrightarrow & \mathcal{C}(I) \\ f & \longmapsto & f' \end{array}$$

qui est une application linéaire et de l'intégration

$$\begin{array}{rcl} \mathcal{C}([a,b]) & \longrightarrow & \mathbf{R} \\ f & \longmapsto & \int_a^b f(x)dx \end{array}$$

qui est une forme linéaire. Si on fixe $x_0 \in I$, on peut aussi considérer la forme linéaire

$$\begin{array}{rcl} d : \mathcal{C}^1(I) & \longrightarrow & \mathbf{R} \\ f & \longmapsto & f'(x_0) \end{array}$$

ou l'application linéaire

$$\begin{array}{rcl} \mathcal{C}(I) & \longrightarrow & \mathcal{C}^1(I) \\ f & \longmapsto & (x \to \int_{x_0}^x f(t)dt) \end{array} .$$

Pour conclure, remarquons que si on voit $\mathbf{C}$ comme un espace vectoriel sur $\mathbf{R}$, alors la conjugaison complexe

$$\begin{array}{rcl} \mathbf{C} & \xrightarrow{\sigma} & \mathbf{C} \\ z & \longmapsto & \overline{z} \end{array}$$

est un automorphisme (c'est une symétrie). Mais cette application n'est plus linéaire si on voit $\mathbf{C}$ comme espace vectoriel sur lui même. Enfin, pour qu'une application linéaire entre espaces vectoriels sur $\mathbf{F}_2$ soit linéaire, il suffit qu'elle satisfasse la première condition.

Proposition 4.2 *Soit $f : E \to F$ une application linéaire. On a alors $f(0) = 0$ et, si $u \in E$, $f(-u) = -f(u)$.*

Démonstration : On a

$$f(0) + f(0) = f(0 + 0) = f(0)$$

et donc $f(0) = 0$.

De même, si $u \in E$, on a

$$f(u) + f(-u) = f(u + (-u)) = f(0) = 0$$

et donc $f(-u) = -f(u)$.

□

Proposition 4.3 *Une application $f : E \to F$ entre deux espaces vectoriels sur K est linéaire si et seulement si pour tous $u, v \in E, \lambda, \mu \in K$, on a*

$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v).$$

Démonstration : Supposons que $f : E \to F$ est linéaire et soient $u, v \in E, \lambda, \mu \in K$. On a alors

$$f(\lambda u + \mu v) = f(\lambda u) + f(\mu v) = \lambda f(u) + \mu f(v).$$

Inversement, si cette condition est satisfaite, on voit que f est linéaire en prenant d'une part $\lambda = \mu = 1$, et d'autre part $\mu = 0$.

□

Proposition 4.4
i) Si $f : E \to F$ et $g : F \to G$ sont deux applications linéaires, alors gof est aussi linéaire.
ii) Si $f : E \to F$ est un isomorphisme entre deux espaces vectoriels, alors f^{-1} est aussi un isomorphisme.

Démonstration : La première assertion se vérifie aisément. On se donne $u, v \in E, \lambda, \mu \in K$, et on vérifie que

$$(g \circ f)(\lambda u + \mu v) = g(f(\lambda u + \mu v)) = g(\lambda f(u) + \mu f(v))$$
$$= \lambda g(f(u)) + \mu g(f(v))) = \lambda(g \circ f)(u) + \mu(g \circ f)(v).$$

La seconde n'est pas beaucoup plus difficile. On sait déjà que f^{-1} est bijective (par définition). On se donne $u, v \in F, \lambda, \mu \in K$, et on veut montrer que

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

Comme f est bijective, il suffit de montrer que

$$\lambda u + \mu v = f(\lambda f^{-1}(u) + \mu f^{-1}(v))$$

et un rapide calcul du second membre nous donne bien

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

□

Proposition 4.5 *Soit $f : E \to F$ une application linéaire. Alors,* *i) Si G est un sous-espace vectoriel de E, alors*

$$f(G) := \{v \in F, \exists u \in G, f(u) = v\}$$

est un sous-espace vectoriel de F.

ii) Si G est un sous-espace vectoriel de F, alors

$$f^{-1}(G) := \{u \in E, f(u) \in G\}$$

est un sous-espace vectoriel de E.

Démonstration : On se donne tout d'abord un sous-espace G de E ainsi que $v, v' \in f(G)$ et $\lambda, \lambda' \in K$. Par définition, on peut trouver $u, u' \in G$ tels que $f(u) = v$ et $f(u') = v'$. Comme G est un sous-espace vectoriel de E, on a donc $\lambda u + \lambda' u' \in G$ et il suit que

$$\lambda v + \lambda' v' = \lambda f(u) + \lambda' f(u') = f(\lambda u + \lambda' u') \in f(G).$$

On remarque aussi que $0 = f(0) \in f(G)$ car $0 \in G$.

Seconde assertion maintenant : on suppose que G est un sous-espace vectoriel de F et on se donne $u, u' \in f^{-1}(G)$ ainsi que $\lambda, \lambda' \in K$. On a donc $f(u), f(u') \in G$ et il suit que

$$f(\lambda u + \lambda' u') = \lambda f(u) + \lambda' f(u') \in G$$

si bien que $\lambda u + \lambda' u' \in f^{-1}(G)$. Là encore, comme $f(0) = 0 \in G$, on a bien $0 \in f^{-1}(G)$.

$\square$

Définition 4.6 *Si $f : E \to F$ est une application linéaire, son* image *est*

$$\operatorname{Im} f := f(E) = \{v \in F, \exists u \in E, f(u) = v\}$$

et son noyau *est*

$$\ker f := f^{-1}(\{0\}) := \{u \in E, f(u) = 0\}.$$

Par exemple, le noyau de l'identité (ou même d'une homothétie) est réduit à $\{0\}$ et son image est l'espace tout entier. Par contre, c'est le noyau de l'application nulle $0_{E,F} : E \to F$ qui est E tout entier et son image qui est réduite à $\{0\}$.

Bien sûr, on a vu que le noyau de l'application linéaire

$$(x_1, \ldots, x_m) \mapsto (a_{11}x_1 + \cdots + a_{1m}x_m, \ldots, a_{n1}x_1 + \cdots + a_{nm}x_m).$$

de K^m vers K^n est l'ensemble des solutions du système homogène

$$(S) : \left\{ \begin{array}{rcl} a_{11}x_1 + \cdots + a_{1m}x_m & = & 0 \\ \vdots & & \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m & = & 0 \end{array} \right. .$$

Le noyau de la dérivation $d : \mathcal{C}^1(I) \to \mathcal{C}(I)$ est l'ensemble des applications constantes et son image est $\mathcal{C}(I)$ tout entier. Enfin, le noyau de l'application

$$\begin{aligned} \mathcal{C}(I) &\longrightarrow \mathcal{C}^1(I) \\ f &\longmapsto (x \mapsto \textstyle\int_{x_0}^{x} f(t)dt) \end{aligned} \quad .$$

est nul et son image est formé des applications continûment dérivables qui s'annulent en x_0. Ce dernier espace n'est d'ailleurs rien d'autre que le noyau de l'évaluation en x_0.

Proposition 4.7 *Soit $f : E \to F$ une application linéaire. Alors,*

i) $\ker f$ *est un sous-espace vectoriel de* E *et* $\mathrm{im} f$ *est un sous-espace vectoriel de* F*.*

ii) f *est injective si et seulement si* $\ker f = 0$ *et* f *est surjective si et seulement si* $\mathrm{Im}\, f = F$*.*

Démonstration : La première assertion est un cas particulier de la proposition précédente. Et il est clair que, par définition, f est surjective si et seulement si $\mathrm{im}(f) = F$. Maintenant, si $u, v \in E$ satisfont, $f(u) = f(v)$, alors

$$f(u - v) = f(u) - f(v) = 0$$

et il suit que $u - v \in \ker f$. On voit donc que si $\ker f = 0$, alors $u - v = 0$ et donc $u = v$. Cela montre que f est injective. Réciproquement, si $u \in \ker f$, alors $f(u) = 0 = f(0)$ et on voit alors que si f est injective, on a bien $u = 0$.

□

Définition 4.8 *Le noyau d'une forme linéaire non nulle est un* hyperplan*.*

Nous aurons besoin plus tard du résultat suivant :

Lemme 4.9 *Soient $f : E \to F$ et $g : F \to G$ deux applications linéaires. On a* $\ker(f) \subset \ker(g \circ f)$ *avec égalité si Si g est injective. On a* $\mathrm{im}(g \circ f) \subset \mathrm{im}(g)$ *avec égalité si f est surjective.*

Démonstration : Tout d'abord, si $u \in \ker f$, on a

$$(g \circ f)(u) = g(f(u)) = g(0) = 0$$

et donc $u \in \ker(g \circ f)$. Réciproquement, si $u \in \ker(g \circ f)$, on a $0 = (g \circ f)(u) = g(f(u))$ et donc $f(u) \in \ker g$. On voit donc que si g est injective, $f(u) = 0$ et donc $u \in \ker f$.

Montrons maintenant l'autre assertion. Si $w \in \mathrm{im}(g \circ f)$, alors il existe $u \in E$ tel que $(g \circ f)(u) = w$ et on a donc $g(f(u)) = w$ si bien que $w \in \mathrm{im}(g)$. Réciproquement, si $w \in \mathrm{im}(g)$ on peut trouver $v \in F$ tel que $g(v) = w$. Si f est surjective, on peut aussi trouver $u \in E$ tel que $f(u) = v$. On a donc $(g \circ f)(u) = g(f(u)) = g(v) = w$ et $w \in \mathrm{im}(g \circ f)$.

□

Proposition 4.10 *Soit E un espace vectoriel sur K.*

i) Si I est un ensemble, l'ensemble E^I des applications de I dans E est un espace vectoriel si on le munit de

$$(f+g)(x) = f(x) + g(x) \text{ et } (\lambda f)(x) = \lambda f(x).$$

ii) Si F est un autre espace vectoriel, l'ensemble $L(E,F)$ est un sous-espace vectoriel de F^E.

Démonstration : On vérifie aisément la première assertion. En effet, toutes les propriétés se vérifient terme à terme. Par exemple, le zéro est l'application nulle.

Nous allons être plus sérieux pour la seconde assertion. Tout d'abord, $L(E,F)$ est non vide car l'application nulle est linéaire.

Il faut montrer ensuite que si f et g sont linéaires, alors $f+g$ aussi. On se donne donc $u, v \in E$ et $\lambda, \mu \in K$ et on calcule

$$(f+g)(\lambda u + \mu v) = f(\lambda u + \mu v) + g(\lambda u + \mu v)$$

$$= \lambda f(u) + \mu f(v) + \lambda g(u) + \mu g(v) = \lambda f(u) + \lambda g(u) + \mu f(v) + \mu g(v)$$
$$= \lambda(f(u) + g(u)) + \mu(f(v) + g(v)) = \lambda(f+g)(u)) + \mu(f+g)(v)).$$

Vérifions finalement que si f est linéaire, alors pour tout $\lambda \in K$, λf est linéaire. Il vaut mieux procéder en deux étapes pour éviter les confusions. On doit montrer que l'on a toujours

$$(\lambda f)(u+v) = (\lambda f)(u) + (\lambda f)(v)$$

ainsi que

$$(\lambda f)(\mu u) = \mu(\lambda f)(u).$$

Or par définition, on a

$$(\lambda f)(u+v) = \lambda f(u+v) = \lambda(f(u) + f(v)) =$$

$$\lambda f(u) + \lambda f(v) = (\lambda f)(u) + (\lambda f)(v).$$

Et de même pour l'autre égalité :

$$(\lambda f)(\mu u) = \lambda f(\mu u) = \lambda \mu f(u) = \mu \lambda f(u) = \mu(\lambda f)(u).$$

□

Proposition 4.11 *Soit E un espace vectoriel sur K.*

i) L'espace $L(E)$ est une algèbre pour l'addition, la multiplication externe et la composition : cela signifie que c'est un espace vectoriel et que, de plus, on a

(a) Si $f, g \in L(E)$, alors $g \circ f \in L(E)$

(b) Si $f, g, h \in L(E)$, alors $h \circ (g \circ f) = (h \circ g) \circ f$

(c) Si $f \in L(E)$, alors $f \circ \mathrm{Id}_E = \mathrm{Id}_E \circ f = f$

(d) Si $f, g, h \in L(E)$, alors $(f+g)\circ h = f\circ h + g\circ h$ et $f\circ(g+h) = f\circ g + f\circ h$

(e) Si $f, g \in L(E)$ et $\lambda \in K$, alors $\lambda(f \circ g) = (\lambda f)\circ g = f \circ (\lambda g)$.

ii) L'ensemble $GL(E)$ est un groupe pour la composition : cela signifie que

(a) Si $f, g \in GL(E)$, alors $g \circ f \in GL(E)$

(b) Si $f, g, h \in GL(E)$, alors $h \circ (g \circ f) = (h \circ g) \circ f$.

(c) Si $f \in GL(E)$, alors $f \circ \mathrm{Id}_E = \mathrm{Id}_E \circ f = f$

(d) Si $f \in GL(E)$, alors $f^{-1} \in GL(E)$.

Démonstration : On s'attaque à la première assertion. On sait déjà que $L(E)$ est un espace vectoriel. On va passer les autres propriétés dans l'ordre. On a déjà vu la première qui dit que la composée de deux applications linéaires est linéaire. La seconde et la troisième propriété sont déjà vraies pour des applications quelconques. Les deux dernières nécessitent une vérification. On se donne $u \in E$ et on vérifie que

$$[(f+g)\circ h](u) = (f+g)(h(u)) = f(h(u)) + g(h(u))$$

$$= (f\circ h)(u) + (g \circ h)(u) = [f\circ h + g\circ h](u),$$

puis que

$$[f\circ(g+h)](u) = f((g+h)(u)) = f(g(u)+h(u)) = f(g(u)) + f(h(u))$$

$$= (f\circ g)(u) + (f\circ h)(u) = [f\circ g + f\circ h](u),$$

et enfin que

$$[\lambda(f\circ g)](u) = \lambda(f\circ g)(u) = \lambda f(g(u))$$
$$= (\lambda f)(g(u)) = [(\lambda f)\circ g](u)$$

d'une part mais aussi

$$[\lambda(f\circ g)](u) = f(\lambda g(u)) = f((\lambda g)(u)) = [f\circ(\lambda g)](u).$$

La seconde assertion n'apporte rien de nouveau puisque l'on sait que la composée de deux bijections est une bijection et que l'on a déjà vu que la réciproque d'une application linéaire bijective est aussi linéaire (et bijective).

□

Corollaire 4.12 *i) Si $n, m \in \mathbf{N}$, alors $M_{n\times m}(K)$ est un espace vectoriel sur K pour l'addition et la multiplication externe.*

ii) *Si $n \in \mathbf{N}$, alors $M_n(K)$ est une algèbre pour l'addition, la multiplication externe et la multiplication interne. Cela signifie que c'est un espace vectoriel et que, de plus, on a*

   (a) *Si $A, B, C \in M_n(K)$, alors $A(BC) = (AB)C$*

   (b) *Si $A \in M_n(K)$, alors $AI = IA = A$*

   (c) *Si $A, B, C \in M_n(K)$, alors $(A+B)C = AC+BC$ et $A(B+C) = AB+AC$*

   (d) *Si $A, B \in M_n(K)$ et $\lambda \in K$, alors $\lambda(AB) = (\lambda A)B = A(\lambda B)$.*

iii) *L'ensemble $GL_n(K)$ des matrices carrées inversibles d'ordre n est un groupe pour la multiplication. Cela signifie que*

   (a) *Si $A, B \in GL_n(K)$, alors $AB \in GL_n(K)$*

   (b) *Si $A, B, C \in GL_n(K)$, alors $A(BC) = (AB)C$*

   (c) *Si $A \in GL_n(K)$, alors $AI = IA = A$*

   (d) *Si $A \in GL_n(K)$, alors $AA^{-1} = I$ et $A^{-1}A = I$.*

Démonstration : Tout cela s'obtient par transport de structure via les bijections

$$L(K^m, K^n) \simeq M_{n\times m}(K), \quad L(K^n) \simeq M_n(K) \quad \text{et} \quad GL(K^n) \simeq GL_n(K)$$

qui sont compatibles avec l'addition, la multiplication externe et la multiplication interne (appelée composition pour les applications).

□

Il suit que l'on peut calculer formellement dans $M_n(K)$ mais il faut se méfier car l'identité remarquable

$$(A+B)^2 = A^2 + 2AB + B^2$$

n'est pas toujours valide par exemple. Mais par contre, on a toujours

$$(I + A)^k = \sum_{i=0}^{N} \binom{k}{i} A^i$$

si bien que pour calculer

$$\begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^7,$$

on écrit notre matrice sous la forme $I + A$ avec

$$A := \begin{bmatrix} 0 & 0 & 5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

et comme $A^2 = 0$, toutes ses puissances seront nulles, et on en déduit que notre matrice vaut

$$I + 7A := \begin{bmatrix} 1 & 0 & 35 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

5 Sommes directes

Proposition 5.1 *Si E_1 et E_2 sont deux espaces vectoriels, les lois sur $E_1 \times E_2$ définies par*

$$(u_1, u_2) + (v_1, v_2) := (u_1 + u_2, v_1 + v_2)$$

et

$$\lambda(u_1, u_2) := (\lambda u_1, \lambda u_2)$$

font de $E_1 \times E_2$ un espace vectoriel et les projections

$$\begin{aligned} E_1 \times E_2 &\longrightarrow E \\ (u_1, u_2) &\longmapsto u_1 \end{aligned}$$

et

$$\begin{aligned} E \times F &\longrightarrow F \\ (u_1, u_2) &\longmapsto u_2 \end{aligned}$$

sont des applications linéaires.

Démonstration : Toutes les propriétés d'un espace vectoriel se vérifient composante par composante et on voit en particulier que le zéro de $E \times F$ est $(0, 0)$ et que l'opposé d'un vecteur (u_1, u_2) sera le vecteur $(-u_1, -u_2)$.

De même, on voit facilement que les projections sont des applications linéaires. Pour la forme, on peut traiter le cas de la première que l'on appelera p_1. On se donne donc $(u_1, u_2), (v_1, v_2) \in E \times F$ ainsi que $\lambda, \mu \in K$. On calcule d'une part

$$\lambda p_1(u_1, u_2) + \mu p_1(v_1, v_2) = \lambda u_1 + \mu v_1$$

et d'autre part

$$p_1(\lambda(u_1, u_2) + \mu(v_1, v_2) = p_1(\lambda u_1 + \mu v_1, \lambda u_2 + \mu v_2) = \lambda u_1 + \mu v_1.$$

□

Définition 5.2 *On dit alors que l'espace vectoriel $E_1 \times E_2$ est le produit des espaces vectoriels E_1 et E_2.*

On définit plus généralement le produit

$$E_1 \times E_2 \times \cdots \times E_n$$

d'un nombre quelconque d'espaces vectoriels. Comme cas particulier, on retrouve

$$K^n := \underbrace{K \times \cdots \times K}_{n}.$$

Proposition 5.3 *Soient E_1 et E_2 deux sous-espace vectoriel d'un espace vectoriel E. Alors,*

i) L'ensemble

$$E_1 + E_2 = \{u_1 + u_2, \quad u_1 \in E_1 \text{ et } u_2 \in E_2\}$$

est le plus petit sous-espace vectoriel de E contenant E_1 et E_2.

ii) L'ensemble $E_1 \cap E_2$ est le plus grand sous-espace vectoriel de E contenu dans E_1 et E_2.

iii) L'application

$$\begin{array}{rrcl} \Phi : & E_1 \times E_2 & \longrightarrow & E \\ & (u_1, u_2) & \longmapsto & u_1 + u_2 \end{array}$$

est linéaire et son image est $E_1 + E_2$.

iv) L'application

$$\begin{array}{rcl} E_1 \cap E_2 & \longrightarrow & E_1 \times E_2 \\ u & \longmapsto & (u, -u) \end{array}$$

induit un isomorphisme entre $E_1 \cap E_2$ et $\ker \Phi$.

Démonstration : On vérifie d'abord la troisième assertion. La linéarité de l'application Φ se vérifie aisément : si $(u_1, u_2), (v_1, v_2) \in E_1 \times E_2$, et $\lambda, \mu \in K$, on a

$$\Phi(\lambda(u_1, u_2) + \mu(v_1, v_2)) = \Phi(\lambda u_1 + \mu v_1, \lambda u_2 + \mu v_2) = \lambda u_1 + \mu v_1 + \lambda u_2 + \mu v_2.$$

Et d'autre part,

$$\lambda\Phi(u_1, u_2) + \mu\Phi(v_1, v_2) = \lambda(u_1 + u_2) + \mu(v_1 + v_2) = \lambda u_1 + \lambda u_2 + \mu v_1 + \mu v_2.$$

Et l'image de Φ est par définition $E_1 + E_2$.

En particulier, on voit que que $E_1 + E_2$ est un sous-espace vectoriel de E. De plus, il contient bien E_1 et E_2. D'autre part, tout sous-espace vectoriel F contenant E_1 et E_2 contient aussi $E_1 + E_2$: si $u_1 \in E_1$ et $u_2 \in E_2$, alors u_1 et u_2 sont tous les deux dans F et comme F est un sous-espace vectoriel, on a $u_1 + u_2 \in F$. La première assertion est ainsi démontrée.

De même, pour la seconde assertion, on vérifie que $E_1 \cap E_2$ est un sous-espace vectoriel de E : Bien sûr, $O \in E_1 \cap E_2$ et si $u, v \in E_1 \cap E_2$ et $\lambda, \mu \in K$, alors, $\lambda u + \mu v \in E_1$ car E_1 est un sous-espace vectoriel de E et, de même, $\lambda u + \mu v \in E_1$ car E_2 si bien que $\lambda u + \mu v \in E_1 \cap E_2$. Et cette intersection est, par définition, la plus grande partie de E contenue dans E_1 et E_2.

Finalement, l'application

$$\begin{array}{rcl} E_1 \cap E_2 & \xrightarrow{\iota} & E_1 \times E_2 \\ u & \longmapsto & (u, -u) \end{array}$$

est linéaire : Si $u, v \in E_1 \cap E_2$ et $\lambda, \mu \in K$, alors, $\iota(\lambda u + \mu v) = (\lambda u + \mu v, -\lambda u - \mu v) = \lambda(u, -u) + \mu(v, -v)$. Elle est injective car son noyau est nul : si $(u, -u) = (0, 0)$, alors, en particulier, $u = 0$. De plus, son image est formée des couples (u, v) avec $u \in E_1, v \in E_2$ et $u + v = 0$, c'est à dire $\ker \Phi$.

□

Par exemple, on peut considérer le plan E_1 donné par $x_1 = x_4 = 0$ et le plan E_2 donné par $x_1 + x_4 = x_2 + x_3 = 0$ dans K^4. Leur intersection est la droite donnée par $x_1 = x_2 + x_3 = x_4 = 0$ dirigée par le vecteur $(0, 1, -1, 0)$ et leur somme est l'hyperplan d'équation $x_1 + x_4 = 0$.

En général, on peut encore, on peut considérer des sommes

$$E_1 + E_2 + \cdots + E_n$$

ou des intersections

$$E_1 \cap E_2 \cap \cdots \cap E_n$$

de plus de deux sous-espaces (et même d'une infinité !). Nous n'en aurons probablement pas besoin.

Proposition 5.4 *Soient E_1 et E_2 deux sous-espace vectoriels d'un espace vectoriel E. Alors, les conditions suivantes sont équivalentes :*

i) Tout élément de E s'écrit de manière unique sous la forme $u_1 + u_2$ avec $u \in E_1$ et $v \in E_2$.

ii) On a $\begin{cases} E = E_1 + E_2 \\ E_1 \cap E_2 = 0 \end{cases}$

iii) L'application

$$\Phi : \quad \begin{aligned} E_1 \times E_2 &\longrightarrow E \\ (u, v) &\longmapsto u + v \end{aligned}$$

est un isomorphisme.

Démonstration : On fait une démonstration circulaire.

Il est clair que la première condition implique que

$$E = E_1 + E_2.$$

De plus, sous cette condition, si $u \in E_1 \cap E_2$, on a

$$u + (-u) = 0 + 0$$

et l'unicité implique que $u = 0$.

La seconde condition implique la troisième grâce à la proposition précédente. En effet, on sait que Φ est linéaire, que son noyau est isomorphe à $E_1 \cap E_2$ et donc réduit à 0, et que son image est $E_1 + E_2$, c'est à dire E. L'application est donc linéaire, injective et surjective, c'est à dire un isomorphisme.

Enfin, si la troisième condition est satisfaite, alors Φ est bijective et la première condition en découle immédiatement.

□

Définition 5.5 *On dit alors que E_1 et E_2 sont* supplémentaires *ou que E est* somme directe *de E_1 et E_2 et on écrit $E = E_1 \oplus E_2$.*

Par exemple, si D et Δ sont deux droites (vectorielles) *distinctes* de K^2, on a toujours $K^2 = D \oplus \Delta$. De même, si H est un plan (vectoriel) de K^3 et D une droite (vectorielle) *non contenue dans H*, alors $K^3 = H \oplus D$. Enfin, si $\mathcal{F}$ est l'ensemble des fonctions de $\mathbf{R}$ dans $\mathbf{R}$, $\mathcal{P}$ le sous-espace des fonctions paires et $\mathcal{I}$ celui des fonctions impaires, on a $\mathcal{F} = \mathcal{P} \oplus \mathcal{I}$. On peut aussi considérer les matrices symétriques et antisymétriques.

Proposition 5.6 *Soient E_1 et E_2 deux sous-espace vectoriels supplémentaires d'un espace vectoriel E. Alors, les* projections

$$E \simeq E_1 \times E_2 \to E_1 \quad \text{et} \quad E \simeq E_1 \times E_2 \to E_2$$

sont des applications linéaires surjectives de noyaux respectifs E_2 et E_1.

Démonstration : Les applications sont bien linéaires comme composées d'applications linéaires.

De plus, on peut décrire la première projection p_1 comme suit : si $u \in E$, on peut écrire de manière unique $u = u_1 + u_2$ avec $u_1 \in E_1$ et $u_2 \in E_2$ et alors, $p_1(u) = u_1$. En particulier, on voit que si $p_1(u) = 0$, alors $u_1 = 0$ et donc $u = u_2 \in E_2$. Réciproquement, si $u \in E_2$, on écrit $u = 0 + u$ avec $0 \in E_1$ et $u \in E_2$ et on a donc bien $p_1(u) = 0$ si bien que $u \in \ker p_1$. On montre aussi que p_1 est surjective car $p_1(u) = u$ si $u \in E_1$.

Les assertions relatives à la seconde projection se montrent de la même manière.

□

Dans l'example ci-dessus $K^2 = D \oplus \Delta$, les projections sont respectivement les projections sur une des droites parallèlement à l'autre. Et on a une description analogue pour l'exemple $K^3 = H \oplus D$. Enfin, pour $\mathcal{F} = \mathcal{P} \oplus \mathcal{I}$, les projections sont les applications $\mathcal{F} \to \mathcal{P}$ et $\mathcal{F} \to \mathcal{I}$ qui associent à une fonction ses parties paire et impaire.

6 Systèmes générateurs et libres

Définition 6.1 *Soit E un espace vectoriel sur K. Si $u_1, \ldots, u_n \in E$ et $\lambda_1, \ldots, \lambda_n \in K$, on dit que*

$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n$$

est une combinaison linéaire *de $u_1, \ldots, u_n$ et que*

$$\lambda_1, \ldots, \lambda_n$$

sont les coefficients. *On dit que la combinaison linéaire est* triviale *si tous les coefficients sont nuls.*

Par convention, la combinaison linéaire vide vaut 0. Une combinaison linéaire d'un vecteur isolé est tout simplement un multiple de ce vecteur. Enfin, si $u, v \in E$, une combinaison linéaire de u et v est un vecteur de la forme $\lambda u + \mu v$ avec $\lambda, \mu \in K$.

Comme exemple, on voit que le vecteur $(2, 3, 2) \in K^3$ est combinaison linéaire de $(1, 0, 1)$ et $(0, 1, 0)$. Contrairement au vecteur $(1, 2, 3)$ par exemple.

Les définitions de sous-espaces vectoriels et d'applications linéaires se généralisent à des combinaisons linéaires générales.

Proposition 6.2 *i) Une partie F d'un espace vectoriel E est un sous-espace vectoriel si et seulement si elle est stable par combinaison linéaire : Pour tout $u_1, \ldots, u_n \in F$ et tout $\lambda_1, \ldots, \lambda_n \in K$, on a*

$$\lambda_1 u_1 + \cdots + \lambda_n u_n \in F.$$

ii) Une application $\varphi : E \to F$ est linéaire si et seulement si elle préserve les combinaisons linéaires : Pout tout $u_1, \ldots, u_n \in E$ et tout $\lambda_1, \ldots, \lambda_n \in K$, on a

$$f(\lambda_1 u_1 + \cdots + \lambda_n u_n) = \lambda_1 f(u_1) + \cdots + \lambda_n f(u_n).$$

Démonstration : Les conditions sont clairement suffisantes car il suffit de considérer le cas $n = 2$ (et aussi le cas $n = 0$ pour s'assurer que $0 \in F$ dans la première assertion).

Et on démontre aisément par récurrence sur n que celles-ci sont nécessaires. On sait que $0 \in F$ dans le premier cas et que $f(0) = 0$ dans le second. On suppose la condition satisfaite à l'ordre n. Dans le premier cas, on aura bien

$$\lambda_1 u_1 + \cdots + \lambda_{n+1} u_{n+1} = (\lambda_1 u_1 + \cdots + \lambda_n u_n) + \lambda_{n+1} u_{n+1} \in F.$$

Et dans le second cas

$$f(\lambda_1 u_1 + \cdots + \lambda_{n+1} u_{n+1}) = f((\lambda_1 u_1 + \cdots + \lambda_n u_n) + \lambda_{n+1} u_{n+1})$$

$$= f(\lambda_1 u_1 + \cdots + \lambda_n u_n) + \lambda_{n+1} f(u_{n+1}) = \lambda_1 f(u_1) + \cdots + \lambda_n f(u_n) + \lambda_{n+1} f(u_{n+1}).$$

$\square$

Proposition 6.3 *Soit E un espace vectoriel.*

i) Toute intersection de sous-espaces vectoriels de E est un sous-espace vectoriel de E.

ii) Si $\mathcal{S}$ est une partie de E, il existe un plus petit sous-espace vectoriel F de E contenant $\mathcal{S}$: c'est l'intersection des sous-espace vectoriels contenant $\mathcal{S}$. C'est aussi l'ensemble des combinaisons linéaire d'éléments de $\mathcal{S}$. On le note $\mathrm{Vect}(\mathcal{S})$.

Démonstration : La première assertion se vérifie aisément : si $(E_i)_{i\in I}$ est une famille de sous-espaces vectoriels de d'un espace vectoriel E, on a bien sûr, $0 \in \cap_{i\in I} E_i$ et si $u, v \in \cap_{i\in I} E_i$ et $\lambda, \mu \in K$, on a pour tout $i \in I$, $\lambda u + \mu v \in E_i$ si bien que $\lambda u + \mu v \in \cap_{i\in I} E_i$.

L'existence et la première caractérisation d'un plus petit sous-espace vectoriel contenant une partie $\mathcal{S}$ donnée en résultent formellement : en effet, il suffit de prendre l'intersection de tous les sous-espaces vectoriels qui contiennent $\mathcal{S}$.

Il reste à montrer la dernière assertion. On montre tout d'abord que l'ensemble F des combinaisons linéaires d'éléments de $\mathcal{S}$ est un sous-espace vectoriel : c'est assez pénible à écrire mais il est clair que la somme de deux combinaisons linéaires est encore une combinaison linéaire et qu'il en va de même pour le produit par un scalaire.

Enfin, le sous-espace vectoriel F contient $\mathcal{S}$ et il résulte de la proposition précédente que c'est le plus petit.

□

Définition 6.4 *On dit alors que F est le* sous espace engendré *par $\mathcal{S}$ ou que $\mathcal{S}$ est un* système générateur *de F. Et on écrit donc $F = \mathrm{Vect}(\mathcal{S})$.*

Dans K, toute partie contenant au moins un vecteur non nul est un système générateur. En d'autres termes, seuls $\emptyset$ et $\{0\}$ ne sont pas générateurs.

Dans K^2, un système générateur contient au moins deux éléments non nuls. Mais ce n'est pas suffisant, par exemple $\{(2, -4), (-3, 6)\}$ n'est pas générateur dans $\mathbf{R}^2$. Par contre, $\{(1, 0), (0, 1)\}$ est bien un système générateur de $\mathbf{R}^2$ et plus généralement, la base canonique de K^n est un système générateur.

Les projections sur les axes forment un système générateur de $\check{K}^n$. Aussi, le système $\{1, X, X^2, X^3, \ldots\}$ de $K[X]$ est générateur.

Pour finir, notons que dans l'espace vectoriel $\{0\}$, tout système est générateur (il n'y en a que deux : $\emptyset$ et $\{0\}$) et que tout espace vectoriel E possède au moins un système générateur : E lui même.

Proposition 6.5 *Soit $\mathcal{G}$ un système générateur d'un espace vectoriel E et $\mathcal{S}$ une partie de E contenant $\mathcal{G}$. Alors, $\mathcal{S}$ est aussi un système générateur de E.*

Démonstration : Clair : le plus petit sous-espace vectoriel de E contenant $\mathcal{S}$ contient aussi $\mathcal{G}$. C'est donc E.

□

Proposition 6.6 *Soit $\mathcal{G}$ un système générateur d'un espace vectoriel E et $f : E \to F$ linéaire. Alors $f(\mathcal{G})$ est un système générateur de* $\mathrm{Im} f$*. En particulier, f est surjective si et seulement si $f(\mathcal{G})$ est un système générateur de F.*

Démonstration : On sait que Imf est un sous-espace vectoriel de F qui contient $f(\mathcal{G})$. Réciproquement, tout élément de Imf est combinaison linéaire d'éléments de $f(\mathcal{G})$: un élément v de im$f$ s'écrit par définition $v = f(u)$ avec $u \in E$. Et comme $\mathcal{G}$ est générateur dans E, on peut écrire $u = \lambda_1 u_1 + \cdots + \lambda_n u_n$ avec $u_1, \ldots, u_n \in \mathcal{G}$. Comme f est linéaire, il suit que

$$v = f(\lambda_1 u_1 + \cdots + \lambda_n u_n) = \lambda_1 f(u_1) + \cdots + \lambda_n f(u_n).$$

□

Définition 6.7 *Soit F un sous-espace vectoriel d'un espace vectoriel E. Une* paramétrisation linéaire *de F est une application linéaire $\phi : K^n \to E$ telle que $F :=$ imϕ.*

Par exemple, si H est l'hyperplan d'équation $x+y+z=0$ dans $\mathbf{R}^3$, on peut prendre la paramétrisation

$$\begin{array}{rccc} \phi : & \mathbf{R}^2 & \longrightarrow & \mathbf{R}^3 \\ & (\lambda, \mu) & \longmapsto & (\lambda+\mu, -\lambda, -\mu). \end{array}$$

Notons qu'on peut définir plus généralement, une paramétrisation pas nécessairement linéaire d'une partie F d'un ensemble E comme une application de K^n vers E dont l'image est F : par exemple

$$\begin{array}{ccc} \mathbf{R} & \longrightarrow & \mathbf{R}^2, \\ t & \longmapsto & (\cos t, \sin t) \end{array}$$

est une paramétrisation du cercle S^1.

Proposition 6.8 *Si E est un espace vectoriel sur K et $u_1, \ldots, u_n \in E$, alors l'application*

$$\begin{array}{rccc} \phi : & K^n & \longrightarrow & E \\ & (\lambda_1, \ldots, \lambda_n) & \longmapsto & \lambda_1 u_1 + \cdots + \lambda_n u_n \end{array}$$

est une paramétrisation (linéaire) du sous-espace F engendré par $u_1, \ldots, u_n$.

Démonstration :

Par définition, l'image de ϕ est exactement le sous-espace engendré par $u_1, \ldots, u_n$. Il faut juste s'assurer que ϕ est bien linéaire. En d'autres termes, il faut vérifier que si on se donne $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n, \lambda, \mu \in K$, on a bien

$$\lambda(\lambda_1 u_1 + \cdots + \lambda_n u_n) + \mu(\mu_1 u_1 + \cdots \mu_n u_n)$$

$$= (\lambda\lambda_1 + \mu\mu_1)u_1 + \cdots + (\lambda\lambda_n + \mu\mu_n)u_n.$$

□

Définition 6.9 *Soit E un espace vectoriel sur K. On dit que $u_1, \ldots, u_n \in E$ sont* linéairement dépendants *si il existe une combinaison linéaire non triviale qui s'annule :*

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0.$$

(c'est une relation de dépendance linéaire*). Sinon, on dit que $u_1, \ldots, u_n$ sont* linéairement indépendants.

Par exemple, dans K, deux vecteurs quelconques sont toujours dépendants. En général, un vecteur tout seul est indépendant si et seulement il est non nul. De toutes manières, si l'un des vecteurs $u_1, \ldots, u_n$ est nul, ils sont obligatoirement dépendants.

Dans K^2, deux vecteurs u et v sont dépendants si et seulement si l'un est multiple de l'autre, par exemple $(2, -4)$ et $(-3, 6)$ sont dépendants dans $\mathbf{R}^2$. Par contre, $(1, 0)$ et $(0, 1)$ sont linéairement indépendants et plus généralement, les vecteurs de la base canonique de K^n sont linéairement indépendants.

Notons aussi que les projections sur les axes sont des formes linéaires linéairement indépendantes. Enfin, des polynômes non nuls de degrés distincts sont linéairement indépendants.

On a besoin de généraliser cette notion à des familles.

Définition 6.10 *Une famille $(u_i)_{i \in I}$ d'éléments d'un espace vectoriel E est un* système lié *si il existe $i_1, \ldots, i_n$ distincts dans I et une combinaison linéaire non triviale qui s'annule :*

$$\lambda_1 u_{i_1} + \cdots + \lambda_n u_{i_n} = 0.$$

Sinon, on dit que la famille $(u_i)_{i \in I}$ est libre.

Par exemple, la famille $(1, X, X^2, X^3, \ldots)$ de $K[X]$ est un système libre : en effet, supposons qu'une combinaison linéaire $\lambda_1 X^{i_1} + \cdots \lambda_n X^{i_n}$ avec des exposants distincts s'annule. Et soit i_k le plus grand exposant. Alors, le polynôme $P := \lambda_1 X^{i_1} + \cdots \lambda_n X^{i_n}$ est polynôme nul mais son degré est i_k. Contradiction.

De même, les fonctions $x \mapsto x^\alpha$ pour $\alpha \in \mathbf{R}$ forment une famille libre de $\mathcal{C}(\mathbf{R}_{>0})$. On le montre ainsi : supposons que la fonction $\lambda_1 x^{\alpha_1} + \cdots \lambda_n x^{\alpha_n}$ soit identiquement nulle. On dérive et on multiplie par x pour obtenir que $\lambda_1 \alpha_1 x^{\alpha_1} + \cdots \lambda_n \alpha_n x^{\alpha_n}$ est aussi identiquement nulle. On multiplie la première par α_n et on retranche la seconde pour trouver que $\lambda_1(\alpha_n - \alpha_1)x^{\alpha_1} + \cdots + \lambda_{n-1}(\alpha_n - \alpha_{n-1})x^{\alpha_n}$ est identiquement nulle. On peut ainsi procéder par récurrence sur n.

La plupart des autres exemples importants résultent en fait de la prochaine proposition.

Avant de poursuivre, je souhaite cependant faire la remarque suivante : un système générateur est un ensemble mais un système libre (ou lié) est une famille. Dans un ensemble, les éléments ne sont pas ordonnés mais on ne peut pas les répéter. Dans une famille, c'est le contraire. Étant donné une famille $(u_i)_{i \in I}$, on peut considérer son support $\{u_i, i \in I\}$ qui est un ensemble. De même, tout ensemble $\mathcal{S}$ définit une famille $(u)_{u\in S}$ indexée par lui même. Par exemple, le support de la famille (infinie) $(u_i = (-1)^i)_{i\in\mathbf{N}}$ est l'ensemble

$$S = \{-1, 1\} = \{x \in \mathbf{R}, x^2 = 1\}$$

auquel on associe la famille (finie) $(u_1 := 1, u_{-1} := -1)$ indexée par l'ensemble $\{-1, 1\}$. En pratique, on mélangera allègrement les deux notions mais il faudra faire attention...

Proposition 6.11 *Des vecteurs $u_1, \ldots, u_n$ d'un espace vectoriel E sont dépendants (resp. indépendants) si et seulement si le système $(u_1, \ldots, u_n)$ est lié (resp. libre).*

Démonstration : En effet, dire que le système est lié signifie qu'on peut trouver une combinaison linéaire non triviale d'une *partie* des vecteurs qui s'annule. Il suffit de compléter avec des coefficients nuls pour obtenir une combinaison linéaire non triviale de *tous* les vecteurs qui s'annule.

□

Proposition 6.12 *Soit $\mathcal{L}$ un système libre d'un espace vectoriel E et $\mathcal{S}$ une partie de E contenue dans $\mathcal{L}$. Alors, $\mathcal{S}$ est aussi un système libre de E.*

Démonstration : Clair : toute combinaison linéaire d'éléments de $\mathcal{S}$ est en particulier combinaison linéaire d'éléments de $\mathcal{L}$.

□

Proposition 6.13 *Soit $\mathcal{L}$ un système libre d'un espace vectoriel E et $f : E \to F$ une application linéaire et injective. Alors, $f(\mathcal{L})$ est un système libre de F.*

Démonstration : On écrit $\mathcal{L} =: (u_i)_{i\in I}$ et on remarque que si $f(\mathcal{L})$ est lié, on peut trouver une combinaison linéaire non-triviale nulle

$$\lambda_{i_1} f(u_{i_1}) + \cdots + \lambda_{i_n} f(u_{i_n}) = 0$$

Comme f est linéaire, on a

$$f(\lambda_{i_1} u_{i_1} + \cdots + \lambda_{i_n} u_{i_n}) = 0.$$

Et comme f est injective, on a nécessairement

$$\lambda_{i_1} u_{i_1} + \cdots + \lambda_{i_n} u_{i_n} = 0.$$

□

Proposition 6.14 *Des vecteurs $u_1, \dots, u_n$ d'un espace vectoriel E sont linéairement indépendants si et seulement si l'application*

$$\begin{array}{rcl} K^n & \xrightarrow{\phi} & E \\ (\lambda_1, \dots, \lambda_n) & \longmapsto & \lambda_1 u_1 + \cdots + \lambda_n u_n \end{array}$$

est injective.

Démonstration : Dire que ϕ est injective signifie que son noyau est nul, c'est à dire que lorsque

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0,$$

on a nécessairement

$$\lambda_1 = \cdots = \lambda_n = 0.$$

□

7 Bases

Définition 7.1 *Une* base *d'un espace vectoriel est un système libre et générateur.*

Il faut être prudent car, comme on l'a déjà fait remarquer, un système libre est une famille et un système générateur est un ensemble. En général, on voit plutôt une base comme une famille libre dont le support est un ensemble générateur.

La seule base de l'espace nul est l'ensemble vide. Tout vecteur non nul de K est une base de K. Les vecteurs $(1, 0)$ et $(0, 1)$, forment une base de $\mathbf{R}^2$ mais il en va de même de $(2, 3)$ et $(3, 2)$.

En général, la base canonique $(e_1, \dots, e_n)$ est une base de K^n (on rappelle que e_i est le vecteur dont toutes les composantes sont nulles sauf la i-ème qui vaut 1). Les projections sur les axes forment une base de $\check{K}^n$ et le système $(1, X, X^2, X^3, \dots)$ est une base de $K[X]$. On peut aussi considérer la base $(e_{i,j})$ de $M_{n\times m}(K)$ ou $e_{i,j}$ est la matrice qui vaut zéro partout sauf à l'intersection de la i-ème ligne et de la j-ème colonne ou elle vaut 1. Dans ces trois cas, on parle aussi de *base canonique*.

Proposition 7.2 *Le système $\mathcal{S}$ est une base de E si et seulement si tout élément de E s'écrit de manière unique comme combinaison linéaire d'éléments de $\mathcal{S}$.*

Démonstration : Si $\mathcal{S}$ est une base, c'est un système générateur et donc tout élément de E s'écrit comme combinaison linéaire d'éléments de $\mathcal{S}$. De plus, si on peut écrire

$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n$$

et
$$u = \mu_1 u_1 + \cdots + \mu_n u_n$$
avec $u_1, \ldots, u_n \in \mathcal{S}$ distincts, alors
$$(\lambda_1 - \mu_1)u_1 + \cdots + (\lambda_n - \mu_n)u_n = 0.$$
Le fait que $\mathcal{S}$ est libre entraîne que
$$\lambda_1 - \mu_1 = \cdots = \lambda_n - \mu_n = 0$$
et on a donc
$$\lambda_1 = \mu_1, \ldots, \lambda_n = \mu_n.$$
Réciproquement, notre condition implique trivialement que $\mathcal{S}$ est générateur et il faut s'assurer qu'il est libre. Or, si $u_1, \ldots, u_n \in \mathcal{S}$ sont tels que
$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0$$
alors, comme on peut aussi écrire
$$0u_1 + \cdots + 0u_n = 0,$$
on a nécessairement
$$\lambda_1 = \cdots = \lambda_n = 0.$$

□

Définition 7.3 *Si* $(u_1, \ldots, u_n)$ *est une base de* E *et*
$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n,$$
on dit que $\lambda_1, \ldots, \lambda_n$ *sont les* composantes *de* u.

Par exemple, les composantes du vecteur $(1,4)$ dans la base canonique $((1,0),(0,1))$ de $\mathbf{R}^2$ sont 1 et 4. Mais dans la base $((2,3),(3,2))$, ce sont 2 et -1. En général, les composantes du vecteur $(x_1, \ldots, x_n) \in K^n$ dans la base canonique sont $x_1, \ldots, x_n$. Et les composantes de la forme linéaire
$$(x_1, \ldots, x_n) \mapsto a_1 x_1 + \cdots a_n x_n$$
dans la base canonique de $\check{K}^n$ sont $a_1, \ldots, a_n$. Attention à l'ordre : les composantes du polynôme $a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$ dans la base canonique $(1, X, X^2, \ldots)$ sont $a_0, a_1, a_2, \ldots$.

Proposition 7.4 *Soit* $f : E \to F$ *une application linéaire et* $\mathcal{B}$ *une base de* E. *Alors,* f *est un isomorphisme (resp. surjective, resp. injective) si et seulement si* $f(\mathcal{B})$ *est une base (resp. un système générateur, resp. un système libre) de* F.

Démonstration : C'est une conséquence des résultats analogues sur les systèmes générateurs et libres à part le fait que si $f(\mathcal{B})$ est libre, alors f est injective. Mais ce n'est pas difficile : en effet, supposons que $u \in E$ satisfasse $f(u) = 0$ et écrivons $u = \lambda_1 u_1 + \cdots + \lambda_n u_n$ comme combinaison linéaire d'éléments (distincts) de $\mathcal{B}$. On aura alors

$$\lambda_1 f(u_1) + \cdots + \lambda_n f(u_n) = f(\lambda_1 u_1 + \cdots + \lambda_n u_n) = f(u) = 0$$

et comme $f(u_1), \ldots, f(u_n)$ sont linéairement indépendants, on aura bien $\lambda_1 = \cdots = \lambda_n = 0$ et donc $u = 0$.

□

Il est important dans cette dernière démonstration de regarder l'image de la famille $\mathcal{B}$ et non pas l'image de l'ensemble. Par exemple, avec l'application

$$\begin{array}{ccc} \mathbf{R}^2 & \xrightarrow{f} & \mathbf{R}^2 \\ (x,y) & \longmapsto & (x+y,0) \end{array},$$

l'image de la base canonique est réduite à $\{e_1\}$ qui est libre mais f n'est pas injective !

Proposition 7.5 *Les vecteurs $u_1, \ldots, u_n$ forment une base de E si et seulement si l'application*

$$\begin{array}{ccc} K^n & \xrightarrow{\Phi} & E \\ (\lambda_1, \ldots, \lambda_n) & \longmapsto & \lambda_1 u_1 + \cdots + \lambda_n u_n \end{array}$$

est un isomorphisme.

Démonstration : Résulte des résultats analogues sur les systèmes générateurs et libres.

□

Proposition 7.6 *Soient $(u_1, \ldots, u_n)$ une base de E et $v_1, \ldots, v_n \in F$. Alors, il existe une unique application linéaire $f : E \to F$ telle que*

$$f(u_1) = v_1, \ldots, f(u_n) = v_n.$$

De plus, $v_1, \ldots, v_n$ sont linéairement indépendants (resp. générateurs, resp. forment une base de F) si et seulement si f est injective, (resp. surjective, resp. bijective).

Démonstration : On se donne donc une base $(u_1, \ldots, u_n)$ de E et

$$v_1, \ldots, v_n \in F.$$

Soit $f : E \to F$ une application linéaire telle que

$$f(u_1) = v_1, \ldots, f(u_n) = v_n.$$

Tout $u \in E$ s'écrit de manière unique

$$u =: \lambda_1 u_1 + \cdots + \lambda_n u_n$$

et on a donc

$$f(u) = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

D'où l'unicité. Réciproquement, on peut toujours définir f comme ceci et vérifier qu'elle est bien linéaire : En effet, si on a aussi

$$u' =: \lambda'_1 u_1 + \cdots + \lambda'_n u_n$$

et $\mu, \mu' \in K$, on doit vérifier que

$$f(\mu u + \mu' u') = \mu f(u) + \mu' f(u').$$

Or, on a

$$\mu u + \mu' u' =: (\mu\lambda_1 + \mu'\lambda'_1)u_1 + \cdots + (\mu\lambda_n + \mu'\lambda'_n)u_n$$

et donc

$$f(\mu u + \mu' u') =: (\mu\lambda_1 + \mu'\lambda'_1)v_1 + \cdots + (\mu\lambda_n + \mu'\lambda'_n)v_n.$$

Et c'est bien la même chose que

$$\mu f(u) + \mu' f(u') = \mu(\lambda_1 v_1 + \cdots + \lambda_n v_n) + \mu'(\lambda'_1 v_1 + \cdots + \lambda'_n v_n).$$

Les autres assertions sont simplement un rappel.

□

Théorème 7.7 *(Théorème de la base incomplète) Si $\mathcal{L}$ est un système libre contenu dans un système générateur $\mathcal{G}$, il existe une base $\mathcal{B}$ de E contenue dans $\mathcal{G}$ et contenant $\mathcal{L}$.*

Démonstration : On considère l'ensemble des systèmes libres contenus dans $\mathcal{G}$ qui contiennent $\mathcal{L}$. Et on lui applique le *lemme de Zorn* : comme, trivialement, toute union croissante de systèmes libre est libre, il existe un système libre maximal $\mathcal{B}$ contenu dans $\mathcal{G}$ et contenant $\mathcal{L}$. Montrons que $\mathcal{B}$ est une base. Comme $\mathcal{G}$ est générateur, il suffit de montrer que tout $u \in \mathcal{G}$ est combinaison linéaire d'éléments de $\mathcal{B}$. La maximalité de $\mathcal{B}$ implique que, soit $u \in \mathcal{B}$ auquel cas on a gagné, soit $\mathcal{B} \cup \{u\}$ est lié. Dans ce dernier cas, on peut trouver une combinaison linéaire non triviale nulle

$$\lambda u + \lambda_1 u_1 + \cdots + \lambda_n u_n = 0$$

avec $u_1, \ldots, u_n \in \mathcal{B}$ distincts. Comme $\mathcal{B}$ est libre, on ne peut pas avoir $\lambda = 0$. On voit donc que

$$u = -\frac{\lambda_1}{\lambda}u_1 - \cdots - \frac{\lambda_n}{\lambda}u_n.$$

□

Corollaire 7.8 *i) Tout espace vectoriel possède une base.*

ii) Tout système générateur contient une base.

iii) Tout système libre est contenu dans une base.

Démonstration : C'est immédiat car $\emptyset$ est libre et E est générateur.

□

Considérons par exemple les vecteurs $u := (1, -1, 0)$ et $v := (1, 0, -1)$ de $\mathbf{R}^3$. Ils sont indépendants et on peut donc considérer le système libre $\mathcal{L} := (u, v)$. D'autre part, on sait que (e_1, e_2, e_3) est une base de $\mathbf{R}^3$ et est donc générateur. Il suit que le système $\mathcal{G} := (u, v, e_1, e_2, e_3)$ est générateur et on a $\mathcal{L} \subset \mathcal{G}$. Le théorème nous dit donc qu'il existe une base $\mathcal{B}$ de $\mathbf{R}^3$ avec $\mathcal{L} \subset \mathcal{B} \subset \mathcal{G}$. Ça fait donc un nombre fini de possibilités que l'on peut tester. En fait, on peut prendre $\mathcal{B} := (u, v, e_1)$.

Proposition 7.9 *Tout sous-espace vectoriel de E possède un supplémentaire dans E.*

Démonstration : On se donne donc un sous-espace vectoriel F de E, on choisit une base $\mathcal{C}$ de F et on la complète en une base $\mathcal{B}$ de E. On note $\mathcal{D} := \mathcal{B} \setminus \mathcal{C}$ et G le sous-espace engendré par $\mathcal{D}$. Il faut montrer que $E = F \oplus G$. Or tout $u \in E$ s'écrit de manière unique comme combinaison linéaire d'éléments de $\mathcal{B}$, c'est à dire comme somme d'une combinaison linéaire d'éléments de $\mathcal{C}$ et d'une combinaison linéaire d'éléments de $\mathcal{D}$. Donc tout élément de E s'écrit bien de manière unique comme somme d'un élément de F et d'un élément de G.

□

8 Dimension

Théorème 8.1 *Deux bases d'un même espace vectoriel ont même nombre d'éléments (fini ou infini).*

Démonstration :

On démontre d'abord le lemme suivant

Lemme 8.2 *Soit $(u_1, \ldots, u_n)$ un système générateur d'un espace vectoriel E et $(v_1, \ldots, v_n)$ un système libre de E. Alors $(v_1, \ldots, v_n)$ est une base.*

Démonstration :

On procède par récurrence sur n. Si $n = 0$, alors $E = \{0\}$ et il n'y a rien à faire. En général, on peut écrire, pour chaque $i = 1, \ldots, n$,

$$v_i =: \lambda_{i1} u_1 + \cdots + \lambda_{in} u_n$$

et on pose $v'_i = v_i - \lambda_{in} u_n$. On considère alors le sous espace F de E engendré par $u_1, \ldots, u_{n-1}$. Si le système $(v'_1, \ldots, v'_n)$ était libre, alors $(v'_1, \ldots, v'_{n-1})$ serait aussi libre. Par récurrence, ce serait une base de F et on pourrait écrire v'_n comme combinaison linéaire des autres vecteurs. Contradiction. Il suit que $(v'_1, \ldots, v'_n)$ est lié et on peut donc trouver une combinaison linéaire non-triviale qui s'annule

$$\mu_1 v'_1 + \cdots + \mu_n v'_n = 0.$$

En posant

$$\lambda := \mu_1 \lambda_{1n} + \cdots + \mu_n \lambda_{nn},$$

on obtient

$$\mu_1 v_1 + \cdots + \mu_n v_n = \lambda u_n.$$

Comme $\{v_1, \ldots, v_n\}$ est libre, on a $\lambda \neq 0$ et on peut écrire u_n comme combinaison linéaire de $\{v_1, \ldots, v_n\}$. Par symétrie, il en va de même pour $u_1, \ldots, u_{n-1}$. Et il suit que $\{v_1, \ldots, v_n\}$ est générateur, et donc une base, de E.

□

Pour démontrer le théorème, on peut bien sûr supposer que notre espace vectoriel possède une base finie $(u_1, \ldots, u_n)$. Et on prend n le plus petit possible. La conclusion est alors formelle : si une autre base $\mathcal{B}$ avait au moins $n+1$ éléments, elle contiendrait une famille libre à n éléments. Grâce au lemme, c'est une base avec laquelle on peut écrire une relation linéaire non triviale dans $\mathcal{B}$.

□

Définition 8.3 *Le nombre d'éléments n d'une base d'un espace vectoriel E est la* dimension *de E. On écrit $\dim E = n$. Un espace de dimension 1 (resp. 2) est* une droite *(resp.* un plan*).*

L'espace nul $\{0\}$ est de dimension nulle car sa base est l'ensemble vide $\emptyset$ qui n'a aucun élément. Plus généralement, K^n est de dimension n car sa base canonique possède n éléments et il en va de même de $\check{K}^n$. Et $K[X]$ est de dimension infinie alors que $\dim K[X]_{\leq n} = n + 1$. Enfin, $\mathcal{C}(I)$ est de dimension infinie (en général).

Un hyperplan de K^2 n'est autre qu'une droite, en effet, c'est un ensemble de la forme

$$\{(x, y) \in K^2, \quad ax + by = 0\},$$

avec a, b non tous les deux nuls, dont une base est $(b, -a)$. Et réciproquement, la droite engendrée par un vecteur non nul (a, b) est l'hyperplan d'équation $bx - ay = 0$.

Proposition 8.4 *Si E est un espace vectoriel de dimension n, alors toute base a n éléments, tout système générateur a au moins n éléments et tout système libre a au plus n éléments.*

Démonstration : La première assertion est la définition de la dimension (et résulte en fait du théorème). Les deux autres proviennent du fait que tout système libre est contenu dans une base et que tout système générateur contient une base.

$\square$

Les vecteurs

$$(3,5,7,11),(5,7,11,13),(7,11,13,17),(11,13,17,19),(13,17,19,23)$$

de $\mathbf{R}^4$ sont ils linéairement indépendants ? Non car, dans un espace de dimension 4, 5 vecteurs sont toujours dépendants ! Inutile de calculer à moins d'avoir besoin d'une relation de dépendance linéaire.

Proposition 8.5 *Soit E un espace vectoriel de dimension finie n et $\mathcal{S}$ une partie à n éléments de E. Alors $\mathcal{S}$ est libre si et seulement si $\mathcal{S}$ est une base si et seulement si $\mathcal{S}$ est générateur.*

Démonstration : Si $\mathcal{S}$ est libre, il est contenu dans une base qui a n éléments : il est donc égal à cette base. De même, si $\mathcal{S}$ est générateur, il contient une base à n éléments : il est donc égal à cette base.

$\square$

En pratique, pour montrer que 3 vecteurs de $\mathbf{R}^3$ forment une base, il suffit de montrer qu'ils sont linéairement indépendants *ou* générateurs !

Proposition 8.6 *Si F est un sous-espace vectoriel d'un espace vectoriel E, alors*

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

Si de plus, on a $\dim F = \dim E < \infty$*, alors $E = F$.*

Démonstration : On complète une base de F qui est un système libre de E en une base de E. Bien sûr, si E et F ont même dimension finie, ils vont avoir même base et donc être engendrés par les mêmes éléments. Ils seront donc égaux.

$\square$

Par exemple, on voit que les seuls sous-espaces vectoriels de K sont $\{0\}$ et K car $\dim K = 1$. De même, outre $\{0\}$ et K^2, les seuls sous-espaces vectoriels de K^2 sont les droites vectorielles. Où encore, outre $\{0\}$ et K^3, les seuls sous-espaces vectoriels de K^3 sont les droites et les plans vectoriels.

Montrons par exemple que l'hyperplan H d'équation $x+y+z=0$ dans $\mathbf{R}^3$ a pour base $u := (1,-1,0)$ et $v := (1,0,-1)$. Tout d'abord, u et v ne sont pas multiples l'un de l'autre. Il sont donc indépendants. Donc l'espace F engendré par u et v est de dimension 2. L'hyperplan H ne contient pas $(1,0,0)$ et donc $H \neq \mathbf{R}^3$. Donc, H est de dimension strictement inférieure à 3. D'autre part, comme u et v sont dans H, on a $F \subset H$. On a donc nécessairement $F = H$. Sinon, F aurait une dimension strictement inférieure à H, c'est à dire au plus 1.

Cet argument s'étend pour montrer qu'un hyperplan de $\mathbf{R}^3$ n'est autre qu'un plan.

Proposition 8.7 *Soit $f : E \to F$ une application linéaire. Alors, si f est injective (resp. surjective, resp. bijective), on a*

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

$$\textit{(resp.}\ \dim F \leq \dim E,$$
$$\textit{resp.}\ \dim E = \dim F).$$

Démonstration : Si f est injective, l'image d'une base de E est un système libre de F qui a même nombre d'éléments et qui est contenu dans une base de F. Et si f est surjective, l'image d'une base de E est un système générateur de F qui contient une base de F.

$\square$

L'application linéaire

$$\begin{aligned} \mathbf{R}^4 &\longrightarrow \mathbf{R}^3 \\ (x,y,z,t) &\longmapsto (3x+5y, 7y+11z, 13z+17t) \end{aligned}$$

est elle injective ? Non, parce que $4 > 3$! Et l'application linéaire

$$\begin{aligned} \mathbf{R}^4 &\longrightarrow \mathbf{R}^5 \\ (x,y,z,t) &\longmapsto (3x+5y, 7y+11z, 13z+17t, 19t+23x, 27x+31y) \end{aligned} .$$

Est elle surjective ? Non, parce que $4 < 5$!

Proposition 8.8 *Deux espaces vectoriels de même dimension finie sont isomorphes. En particulier, tout espace vectoriel sur K de dimension finie n est isomorphe à K^n.*

Démonstration : En effet, si on se donne une base de chaque, il existe une unique application linéaire qui envoie la première sur la seconde. Et c'est un isomorphisme.

□

Définition 8.9 *Des vecteurs sont* colinéaires *s'ils appartiennent à une même droite et* coplanaires *s'ils appartiennent à un même plan.*

Par exemple, les vecteurs $(2, -4)$ et $(-3, 6)$ sont colinéaires.

Proposition 8.10 *Deux vecteurs sont linéairement dépendants si et seulement si ils sont colinéaires. De même, trois vecteurs sont linéairement dépendants si et seulement si ils sont coplanaires.*

Démonstration : Plus généralement, n vecteurs sont dépendants si et seulement si l'espace engendré est de dimension strictement inférieure à n et cette dernière condition est équivalente au fait d'appartenir tous à un sous-espace de dimension strictement inférieure à n. On applique ça à $n = 2$ et $n = 3$.

□

Mais attention, trois vecteurs non colinéaires deux à deux peuvent cependant être coplanaires. Prendre par exemple $(1, 0, 0), (0, 1, 0), (1, 1, 0)$ dans $\mathbf{R}^3$.

Proposition 8.11 *Un sous-espace vectoriel H d'un espace vectoriel E est un hyperplan si et seulement si il existe une droite D de E telle que*

$$E = H \oplus D.$$

Et ceci est alors vrai pour toute droite D non contenue dans H.

Démonstration : Si H est un hyperplan, on peut écrire $H = \ker f$ avec $f : E \to K$ linéaire non nulle. Il existe donc $u \in E$ tel que $f(u) \neq 0$ et on notera D la droite dirigée par u. Notons que, réciproquement, si D est une droite non contenue dans H et dirigée par un vecteur u, alors $f(u) \neq 0$.

On a bien sûr $H \cap D = \{0\}$ parce-que c'est un sous-espace vectoriel de D (qui a dimension 1) distinct de D (car u n'est pas dedans). Et on peut facilement voir que tout $v \in E$ s'écrit comme somme d'un élément de H et d'un élément de D :

$$v = (v - \frac{f(v)}{f(u)}u) + \frac{f(v)}{f(u)}u.$$

Réciproquement, si $E = H \oplus D$, alors H est le noyau de l'application composée de la projection sur D et d'un isomorphisme quelconque entre D et K. C'est donc bien un hyperplan.

□

9 Rang

Théorème 9.1 *(Théorème du rang) Si $f : E \to F$ est une application linéaire, alors*

$$\dim E = \dim \ker f + \dim \operatorname{im} f.$$

Démonstration : Si $\dim \ker f = \infty$, l'égalité est claire. Sinon, on choisit une base $(u_1, \ldots, u_t)$ de $\ker f$ et on la prolonge en une suite $u_1, \ldots, u_n$ d'éléments de E. Posons

$$v_{t+1} := f(u_{t+1}), \ldots, v_n := f(u_n).$$

Supposons que $(u_1, \ldots, u_n)$ est libre. Si

$$\lambda_{t+1}v_{t+1} + \cdots + \lambda_n v_n = 0,$$

alors

$$f(\lambda_{t+1}u_{t+1} + \cdots + \lambda_n u_n) = 0$$

et on peut donc écrire

$$\lambda_{t+1}u_{t+1} + \cdots + \lambda_n u_n = \lambda_1 u_1 + \cdots + \lambda_r u_r.$$

Il suit que

$$\lambda_1 = \cdots = \lambda_n = 0$$

et en particulier que $(v_{t+1}, \ldots, v_n)$ est libre.

De même, si on suppose que $(v_{t+1}, \ldots, v_n)$ est libre et si

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0,$$

alors

$$\lambda_{t+1}v_{t+1} + \cdots + \lambda_n v_n = 0$$

si bien que

$$\lambda_{t+1} = \cdots = \lambda_n = 0.$$

On a donc $\lambda_1 u_1 + \cdots + \lambda_r u_r = 0$ et donc aussi

$$\lambda_1 = \cdots = \lambda_t = 0.$$

Il suit que $(u_1, \ldots, u_n)$ est libre.

On conclut alors facilement. Soit n un entier quelconque. Si $\dim E \geq n$, on peut compléter $(u_1, \ldots, u_t)$ en un système libre $(u_1, \ldots, u_n)$ grâce au théorème de la base incomplète. Il suit que $(v_{t+1}, \ldots, v_n)$ est libre et on voit alors que $\dim \operatorname{im} f \geq n - t$. Réciproquement, si $\dim \operatorname{im} f \geq n-t$, on peut y trouver un système libre $(v_{t+1}, \ldots, v_n)$ grâce au théorème de la base incomplète. Pour tout $i = 1, \ldots, n$, il existe $u_i \in E$ tel que $f(u_i) = v_i$ et il suit que $(u_1, \ldots, u_n)$ est nécessairement libre. On voit alors que $\dim E \geq n$. Ceci étant vrai pour tout n, on a bien $\dim \operatorname{im} f = \dim E - t$.

$\square$

La définition ci-dessous justifie la terminologie « Théorème du rang » :

Définition 9.2 *i) Le* rang *d'une partie $\mathcal{S}$ d'un espace vectoriel E est la dimension de l'espace engendré par ce système. On la note* $\mathrm{rg}(\mathcal{S})$.

ii) Le rang *d'une application linéaire f est la dimension de* $\mathrm{Im} f$. *On le note* $\mathrm{rg}(f)$.

iii) Si A est la matrice d'une application linéaire $f : K^m \to K^n$ (dans la base canonique), le rang *de A est le rang de f. On le note* $\mathrm{rg}(A)$.

iv) Le rang *d'un système linéaire*

$$(S) : \left\{ \begin{array}{rcl} a_{11}x_1 + \cdots + a_{1m}x_m & = & b_1 \\ \vdots & & \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m & = & b_n \end{array} \right. ,$$

est le rang de la matrice associée. On le note $\mathrm{rg}(S)$.

Par exemple, le rang de $\mathcal{S} := \{(1,-1,0),(0,1,-1),(-1,0,1)\} \subset \mathbf{R}^3$ est 2 car l'espace engendré est le plan d'équation $x + y + z = 0$.

Les deux premières définitions ne sont pas indépendantes : si $\mathcal{B}$ est une base de E, alors le rang de f est égal au rang de $f(\mathcal{B})$ puisque ce dernier est générateur de $\mathrm{im}(f)$. Par exemple, l'application

$$\begin{array}{ccc} \mathbf{R}^3 & \longrightarrow & \mathbf{R}^3 \\ (x,y,z) & \longmapsto & (x-z, y-x, z-y) \end{array}$$

est de rang 2 car l'image de la base canonique de $\mathbf{R}^3$ est justement l'ensemble $\mathcal{S}$ ci-dessus.

On voit ainsi que le rang d'une matrice est le rang du système formé par ses vecteurs colonnes. Dans notre exemple, on a

$$A = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

Enfin, dans le cas d'un système linéaire homogène de rang r, le théorème du rang nous dit que l'espace des solutions est de dimension $n - r$.

Par exemple,

$$(S) : \left\{ \begin{array}{rcl} x - z & = & 0 \\ y - x & = & 0 \\ z - x & = & 0 \end{array} \right. ,$$

est de rang 2 et l'ensemble des solutions est donc une droite $(3 - 2 = 1)$. C'est bien sûr la droite de direction $(1,1,1)$.

Proposition 9.3 *Soient $f : E \to F$ et $g : F \to G$ deux applications linéaires. Si f est surjective, on a* $\mathrm{rg}(g \circ f) = \mathrm{rg}(g)$ *et si g est injective, on a* $\mathrm{rg}(g \circ f) = \mathrm{rg}(f)$.

Démonstration : On sait que si f est surjective, on a $\mathrm{Im}(g \circ f) = \mathrm{Im}(g)$ et la première assertion en découle. De même, si g est injective, on utilise le théorème du rang et le fait que $\ker(g \circ f) = \ker(f)$.

□

On utilisera ce résultat essentiellement pour des applications linéaires bijectives.

Corollaire 9.4 *Soit $A \in M_{n\times m}(K)$. Si $P \in GL_n(K)$, alors* $\mathrm{rg}(PA) = \mathrm{rg}(A)$. *De même, si $Q \in GL_m(K)$, alors* $\mathrm{rg}(AQ) = \mathrm{rg}(A)$.

Démonstration : On écrit A comme la matrice d'une application f et P comme la matrice d'une application φ (dans la base canonique). Comme P est inversible, on sait que φ est bijective et donc en particulier injective. Il suit que $\mathrm{rg}(\varphi \circ f) = \mathrm{rg} f$ et comme la matrice de la composée $\varphi \circ f$ n'est autre que le produit PA, on a bien $\mathrm{rg}(PA) = \mathrm{rg}(A)$. Et de même pour l'autre égalité.

□

Corollaire 9.5 *Les opérations élémentaires sur les lignes ou sur les colonnes d'une matrice ne changent pas son rang.*

Démonstration : En effet, on sait qu'une faire une opération élémentaire revient à multiplier par une matrice élémentaire, qui est inversible.

□

En fait, le rang de la matrice est le nombre de lignes non nulles après avoir appliqué la méthode du pivot de Gauss.

Le théorème du rang nous permet de démontrer ce que l'on appelle la relation de Grassmann. Mais avant, nous avons besoin d'un lemme.

Lemme 9.6 *Si E et F sont deux espaces vectoriel, alors*

$$\dim E \times F = \dim E + \dim F.$$

En fait, si $(e_i)_{i\in I}$ et $(f_j)_{j\in J}$ sont des bases de E et F respectivement, alors

$$((e_i, 0), (0, f_j))_{i\in I, j\in J}$$

est une base de $E \times F$.

Démonstration : Par définition, tout élément de $E \times F$ s'écrit de manière unique (u, v) avec $u \in E, v \in F$. Et u et v s'écrivent de manière unique

$$u = \lambda_1 e_{i_1} + \cdots + \lambda_k e_{i_k} \quad \text{et} \quad v = \mu_1 f_{i_1} + \cdots + \mu_k f_{i_k}.$$

Il suit que (u, v) s'écrit de manière unique

$$(u, v) = \lambda_1 (e_{i_1}, 0) + \cdots + \lambda_k (e_{i_k}, 0) + \mu_1 (0, f_{i_1}) + \cdots + \mu_k (0, f_{i_k}).$$

$\square$

Cela permet de redémontrer par récurrence sur n que $\dim K^n = n$.

Proposition 9.7 *(Relation de Grassmann) Si E_1 et E_2 sont deux sous-espaces vectoriels de E, alors*

$$\dim E_1 + \dim E_2 = \dim(E_1 + E_2) + \dim(E_1 \cap E_2).$$

Démonstration : On sait que l'application

$$\Phi : \quad \begin{array}{ccc} E_1 \times E_2 & \longrightarrow & E \\ (u_1, u_2) & \longmapsto & u_1 + u_2 \end{array}$$

est linéaire, que son image est $E_1 + E_2$ et que son noyau est isomorphe à $E_1 \cap E_2$. On a donc, grâce au théorème du rang,

$$\dim E_1 \times E_2 = \dim E_1 + E_2 + \dim E_1 \cap E_2.$$

et on utilise la proposition précédente.

$\square$

On peut en déduire par exemple que deux plans de $\mathbf{R}^3$ ont une droite en commun : en effet, si H et H' désignent ces plans, on a

$$\dim(H \cap H') = 2 + 2 - \dim(H + H') \geq 4 - 3 = 1$$

Corollaire 9.8 *Soient E_1 et E_2 deux sous-espaces vectoriels d'un espace vectoriel E de dimension finie. Alors*

$$E = E_1 \oplus E_2 \Leftrightarrow \begin{cases} \dim E = \dim E_1 + \dim E_2 \\ E_1 \cap E_2 = \{0\}. \end{cases}$$

Démonstration : On connaît déjà ce résultat avec la première condition remplacée par $E = E_1 + E_2$. Il suffit donc de montrer que si $E_1 \cap E_2 = \{0\}$, on a

$$E = E_1 + E_2 \Leftrightarrow \dim E = \dim E_1 + \dim E_2.$$

Grâce à la relation de Grassmann, on a $\dim E_1 + \dim E_2 = \dim(E_1 + E_2)$. Et comme $E_1 + E_2$ est un sous-espace vectoriel de E qui est de dimension finie, on a bien

$$E = E_1 + E_2 \Leftrightarrow \dim E = \dim(E_1 + \dim E_2).$$

□

Par exemple, on voit que le supplémentaire d'une droite dans un plan est une autre droite. Et dans l'espace, c'est un plan ne contenant pas cette droite.

Proposition 9.9 *Un sous-espace vectoriel H d'un espace vectoriel E de dimension finie n est un hyperplan si et seulement si* $\dim H = n - 1$.

Démonstration : On sait que H est un hyperplan si et seulement si il existe une droite D avec $E = H \oplus D$. En général, on sait qu'il existe toujours un supplémentaire D pour H et on a alors $\dim D + \dim H = \dim E$. Comme $\dim E = n$, on voir que $\dim D = 1$ si et seulement si $\dim H = n - 1$.

□

Une autre application du théorème du rang est la suivante.

Proposition 9.10 *Soit E un espace vectoriel de dimension finie n et f un endomorphisme de E. Alors f est injective si et seulement si f est surjective si et seulement si f est bijective si et seulement si* $\mathrm{rg}(f) = n$.

Démonstration : Par définition, f est surjective si et seulement si $\mathrm{Im}(f) = E$ et comme $\dim E = n$ est finie, cela signifie que $\mathrm{rg}(f) = \dim \mathrm{Im}(f) = \dim E = n$. Maintenant, grâce au théorème du rang, on a $\mathrm{rg}(f) = n - \dim \ker f$ et on voit que $\mathrm{rg}(f) = n$ si et seulement si $\dim \ker f = 0$, c'est à dire $\ker f = \{0\}$, ce qui signifie bien que f est injective.

□

Corollaire 9.11

i) Une matrice carrée A d'ordre n est inversible si et seulement si $\mathrm{rg}(A) = n$.

ii) Si A et B sont deux matrices carrées d'ordre n telles que $AB = I$, alors A et B sont inversibles et inverses l'une de l'autre.

Démonstration : Pour la première assertion, on écrit A comme matrice d'une application f dans la base canonique et on traduit le résultat précédent.

Pour la seconde assertion, on écrit A et B comme matrices d'applications f et g respectivement si bien que AB est la matrice de $f \circ g$. Notre hypothèse nous dit que $f \circ g = \mathrm{Id}_E$ et implique en particulier que f est surjective. Celle-ci est donc bijective si bien que A est inversible. On en déduit que

$$B = A^{-1}AB = A^{-1}I = A^{-1}$$

comme annoncé. Il suit formellement que B aussi est inversible et que $B^{-1} = A$.

□

10 Matrices d'applications linéaires

Dans cette section, tous les espaces vectoriels sont supposés de dimension finie. On commence par les vecteurs colonnes.

Définition 10.1 *Soit E un espace vectoriel muni d'une base*

$$\mathcal{B} := (u_1, \ldots, u_n).$$

Si $u \in E$ a pour coordonnées $(a_1, \ldots, a_n)$ dans cette base, le vecteur colonne *associé à u est*

$$[u]_{\mathcal{B}} := \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Par exemple, le vecteur colonne associé à $(X-3)^2$ dans la base canonique de $K[X]_{\leq 2}$ est

$$\begin{bmatrix} 9 \\ -6 \\ 1 \end{bmatrix}.$$

Et le vecteur colonne associé à $(2,0)$ dans la base $\{(1,1),(1,-1)\}$ de $\mathbf{R}^2$ est

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

On rappelle qu'il revient au même de se donner une base $\mathcal{B}$ d'un espace vectoriel E ou une paramétrisation (bijective) $\Phi : K^n \to E$: si $(e_1, \ldots, e_n)$ désigne la base canonique de K^n et $\mathcal{B} := (u_1, \ldots, u_n)$, on a tout simplement $\Phi(e_i) = u_i$.

Lemme 10.2 *Si $\Phi : K^n \to E$ est la paramétrisation d'un espace vectoriel E correspondant à une base $\mathcal{B}$ de E, alors pour tout $u \in E$, on a $[u]_{\mathcal{B}} = [\Phi^{-1}(u)]$.*

Démonstration : Par définition, si $(e_1, \ldots, e_n)$ désigne la base canonique de K^n et $\mathcal{B} := (u_1, \ldots, u_n)$ alors, $\Phi(e_i) = u_i$ pour tout i. Donc, si $u := a_1u_1 + \cdots + a_nu_n$, on a bien $\Phi^{-1}(u) = a_1e_1 + \cdots a_ne_n$.

□

On passe maintenant aux applications linéaires.

Définition 10.3 *Soient E et F deux espaces vectoriels munis de bases*

$$\mathcal{B} := (u_1, \ldots, u_m)$$

et $\mathcal{C}$ respectivement. Alors, la matrice d'une application linéaire *$f : E \to F$ est*

$$[f]^{\mathcal{C}}_{\mathcal{B}} := [[f(u_1)]_{\mathcal{C}} \cdots [f(u_m)]_{\mathcal{C}}].$$

Lorsque $E = F$ et $\mathcal{B} = \mathcal{C}$, on écrit tout simplement $[f]_{\mathcal{B}}$.

Par exemple, dans la base canonique, la matrice de l'application linéaire

$$\begin{array}{rcl} \mathbf{R}[X]_{\leq 2} & \longrightarrow & R[X]_{\leq 2} \\ P & \longmapsto & P' \end{array}$$

est

$$A := \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Le lemme suivant permet de ramener toutes les questions sur les matrices d'applications linéaires à des espaces de la forme K^n.

Lemme 10.4 *Soit $f : E \to F$ une application linéaire, $\Phi : K^m \to E$ la paramétrisation associée à une base $\mathcal{B}$ de E et $\Psi : K^n \to F$ la paramétrisation associée à une base $\mathcal{C}$ de F. On a alors*

$$[f]^{\mathcal{C}}_{\mathcal{B}} = [\Psi^{-1} \circ f \circ \Phi].$$

Démonstration : Si $(e_1, \ldots, e_m)$ désigne la base canonique de K^m et $\mathcal{B} := (u_1, \ldots, u_m)$, il s'agit de montrer que pour tout i, on a

$$[(\Psi^{-1} \circ f \circ \Phi)(e_i)] = [f(u_i)]_{\mathcal{C}},$$

ce qui résulte immédiatement du lemme.

□

Proposition 10.5

Soient E et F deux espaces vectoriels de dimension m et n, respectivement, munis de bases $\mathcal{B}$ et $\mathcal{C}$, respectivement. Alors, l'application

$$\begin{aligned} L(E,F) &\longrightarrow M_{n\times m}(K) \\ f &\longmapsto [f]_{\mathcal{B}}^{\mathcal{C}} \end{aligned}$$

est un isomorphisme d'espaces vectoriels.

Démonstration : Si $\Phi : K^m \to E$ est la paramétrisation associée à $\mathcal{B}$ et $\Psi : K^n \to F$ la paramétrisation associée à $\mathcal{C}$, notre application se décompose en deux isomorphismes

$$\begin{array}{ccccc} L(E,F) & \xrightarrow{\simeq} & L(K^m,K^n) & \xrightarrow{\simeq} & M_{n\times m}(K). \\ f & \longmapsto & \Psi^{-1}\circ f\circ\Phi & & \\ & & g & \longmapsto & [g] \end{array}$$

En fait, on sait déjà que la seconde application est un isomorphisme. Comme la première est clairement bijective, d'inverse $g \mapsto \Psi\circ g\circ\Phi^{-1}$, il faut juste s'assurer qu'elle est linéaire, c'est à dire l'identité

$$\Psi^{-1}\circ(\lambda f+\mu g)\circ\Phi = \lambda(\Psi^{-1}\circ f\circ\Phi)+\mu(\Psi^{-1}\circ g\circ\Phi).$$

□

Corollaire 10.6 *Si E et F sont deux espaces vectoriels, on a*

$$\dim L(E,F) = \dim E\times\dim F.$$

Démonstration :

□

En particulier, $\dim \check{E} = \dim E$ et $\dim L(E) = (\dim E)^2$.

Proposition 10.7 *Si $\mathcal{B}$ et $\mathcal{C}$ sont des bases d'espaces vectoriels E et F respectivement, si $u \in E$ et si $f : E \to F$ est linéaire, on a*

$$[f(u)]_{\mathcal{C}} = [f]_{\mathcal{B}}^{\mathcal{C}}[u]_{\mathcal{B}}.$$

Si $\mathcal{B}$, $\mathcal{C}$ et $\mathcal{D}$ sont des bases de E, F et G respectivement, et si $f : E \to F$ et $g : F \to G$ sont linéaires, on a

$$[g\circ f]_{\mathcal{B}}^{\mathcal{D}} = [g]_{\mathcal{C}}^{\mathcal{D}}[f]_{\mathcal{B}}^{\mathcal{C}}.$$

Démonstration : À nouveau, on se ramène au cas de la base canonique en utilisant des paramétrisations. Avec nos notations habituelles, la première assertion se ramènera donc à vérifier que

$$[\Psi^{-1}(f(u))] = [\Psi^{-1} \circ f \circ \Phi][\Phi^{-1}(u)],$$

et comme nous connaissons déjà le résultat pour la base canonique, on est réduit à

$$\Psi^{-1}(f(u)) = (\Psi^{-1} \circ f \circ \Phi)(\Phi^{-1}(u)).$$

La seconde assertion se démontre de la même manière en faisant intervenir trois paramétrisations.

□

Corollaire 10.8 *Soit E un espace vectoriel de dimension n muni d'une base $\mathcal{B}$. Alors,*

i) L'application

$$\begin{aligned} L(E) &\longrightarrow M_n(K) \\ f &\longmapsto [f]_{\mathcal{B}} \end{aligned}$$

est un isomorphisme d'algèbres : C'est un isomorphisme d'espaces vectoriels et on a en plus, pour $f, g \in L(E)$,

$$[g \circ f]_{\mathcal{B}} = [g]_{\mathcal{B}}[f]_{\mathcal{B}}.$$

ii) L'application

$$\begin{aligned} GL(E) &\longrightarrow GL_n(K) \\ f &\longmapsto [f]_{\mathcal{B}} \end{aligned}$$

est un isomorphisme de groupes : C'est une bijection et on a en plus, pour $f, g \in L(E)$,

$$[g \circ f]_{\mathcal{B}} = [g]_{\mathcal{B}}[f]_{\mathcal{B}}.$$

Démonstration : Cela résulte des résultats analogues pour les bases canoniques en passant par les paramétrisations.

□

Définition 10.9 *Soit E un espace vectoriel muni de deux bases $\mathcal{B}$ et $\mathcal{B}'$. La* matrice de passage *de $\mathcal{B}$ à $\mathcal{B}'$ est $[\mathrm{Id}_E]^{\mathcal{B}}_{\mathcal{B}'}$.*

Attention à l'ordre : on munit l'espace de *départ* de la « nouvelle » base $\mathcal{B}'$ et l'espace d'*arrivée* de l' « ancienne » base $\mathcal{B}$. En pratique, on écrit les vecteurs de la nouvelle base dans l'ancienne.

Pour donner un exemple, il faut deux bases du même espace vectoriel. Par exemple, on peut prendre pour $\mathcal{B}$ la base canonique de $\mathbf{R}^2$ et $\mathcal{B}' := \{(1,1),(1,-1)\}$. La matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est alors

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

Mais la matrice de passage dans l'autre sens, de $\mathcal{B}'$ à $\mathcal{B}$ est

$$P^{-1} := \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

Ceci est général et on a le résultat suivant :

Proposition 10.10 *i) Une matrice de passage est inversible. Plus précisément, l'inverse de la matrice de passage d'une base $\mathcal{B}$ à une base $\mathcal{B}'$ est la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$.*

ii) Soit E un espace vectoriel muni de deux bases $\mathcal{B}$ et $\mathcal{B}'$ et F un espace vectoriel muni de deux bases $\mathcal{C}$ et $\mathcal{C}'$. On note P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ et Q la matrice de passage de $\mathcal{C}$ à $\mathcal{C}'$. Soit $f : E \to F$ une application linéaire, A sa matrice dans les bases $\mathcal{B}$ et $\mathcal{C}$ et A' sa matrice dans les bases $\mathcal{B}'$ et $\mathcal{C}'$. On a alors,

$$A' = Q^{-1}AP.$$

Démonstration : La première assertion est immédiate (vérifier). Pour la seconde, on remarque que

$$f := \mathrm{Id}_F \circ f \circ \mathrm{Id}_E$$

et donc que

$$[f]_{\mathcal{B}'}^{\mathcal{C}'} := [\mathrm{Id}_F]_{\mathcal{C}}^{\mathcal{C}'}[F]_{\mathcal{B}}^{\mathcal{C}}[\mathrm{Id}_E]_{\mathcal{B}'}^{\mathcal{B}}.$$

□

Définition 10.11 *Si A et A' sont deux matrices $n \times m$ telles qu'ils existe des matrices inversibles P et Q avec $A' = Q^{-1}AP$, on dit que A et A' sont* équivalentes.

Si A et A' sont deux matrices carrées d'ordre n telles qu'il existe une matrice inversible P avec $A' = P^{-1}AP$, on dit qu'elles sont semblables.

On peut montrer que deux matrices sont équivalentes si et seulement si elles ont même rang (dans un sens c'est clair, non). Par contre le fait d'être semblable est une propriété très forte.

Regardons par exemple l'application linéaire

$$\begin{aligned} \mathbf{R}^2 &\xrightarrow{f} \mathbf{R}^2 \\ (x,y) &\longmapsto (x+2y, 2x+y) \end{aligned}$$

Sa matrice dans la base canonique est

$$A := \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

Il suit que sa matrice dans la base $\mathcal{B}' := \{(1,1),(1,-1)\}$ est

$$A' = P^{-1}AP = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}.$$

C'est pratique pour calculer les puissances de A :

$$A^n = PA'^nP^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3^n & 0 \\ 0 & (-1)^n \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{3^n+(-1)^n}{2} & \frac{3^n-(-1)^n}{2} \\ \frac{3^n-(-1)^n}{2} & \frac{3^n+(-1)^n}{2} \end{bmatrix}.$$

Application : la double suite récurrente

$$\begin{cases} u_{n+1} & = & u_n + 2v_n \\ v_{n+1} & = & 2u_n + v_n \end{cases}$$

avec $u_0 = 3$ et $v_0 = 2$? On écrit

$$\begin{bmatrix} u_{n+1} \\ v_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} u_n \\ v_n \end{bmatrix}$$

ou encore

$$X_{n+1} = AX_n \quad \text{avec} \quad X_n := \begin{bmatrix} u_n \\ v_n \end{bmatrix} \quad \text{et} \quad X_0 := \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

On a donc

$$X_n := A^nX_0 = \begin{bmatrix} \frac{3^n+(-1)^n}{2} & \frac{3^n-(-1)^n}{2} \\ \frac{3^n-(-1)^n}{2} & \frac{3^n+(-1)^n}{2} \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{5.3^n+(-1)^n}{2} \\ \frac{5.3^n-(-1)^n}{2} \end{bmatrix}$$

si bien que

$$u_n = \frac{5.3^n + (-1)^n}{2} \quad \text{et} \quad v_n = \frac{5.3^n - (-1)^n}{2}.$$

11 Dualité

Proposition 11.1 *Soit F un sous-espace vectoriel de dimension m d'un espace vectoriel E de dimension finie n et $r := n - m$. Alors, il existe $f_1, \ldots, f_r \in \check{E}$ linéairement indépendantes telles que*

$$F = \{u \in E, \quad f_1(u) = \cdots = f_r(u) = 0\}$$

Démonstration : Soit G un supplémentaire de F dans E et $p : E \to G$ la projection. Soit $\phi : K^{n-m} \simeq G$ une paramétrisation bijective (donné par une base de G). On regarde la composée $q := \phi^{-1} \circ p$ qui est donnée par r formes linéaires $f_1, \ldots, f_r$. Comme F est le noyau de q, on a bien l'égalité annoncée. Le fait que $f_1, \ldots, f_r$ sont linéairement indépendantes résulte du fait que q est surjectif grace au corollaire 11.13 que nous démontrerons plus bas.

□

En particulier, on voit qu'un sous-espace vectoriel de K^n est toujours l'ensemble des solutions d'un système homogène.

Proposition 11.2 *Si $\mathcal{B} := (u_1, \ldots, u_n)$ est une base d'un espace vectoriel E, alors*

$$\check{\mathcal{B}} := (\check{u}_1, \ldots, \check{u}_n),$$

où $\check{u}_i$ est l'unique forme linéaire telle que

$$\check{u}_i(u_j) = \delta_{ij} = \begin{cases} 1 & \text{si} \quad i = j \\ 0 & \text{sinon} \end{cases},$$

est une base de $\check{E}$. De plus, on a toujours

$$f = f(u_1)\check{u}_1 + \cdots + f(u_n)\check{u}_n$$

et

$$u = \check{u}_1(u)u_1 + \cdots + \check{u}_n(u)u_n.$$

Démonstration : Il est clair que notre famille est libre : si $\lambda_1, \ldots, \lambda_n \in K$ satisfont $\lambda_1\check{u}_1 + \cdots + \lambda_n\check{u}_n = 0$ et qu'on applique cette égalité au vecteur u_i, on trouve que $\lambda_i = 0$. On montre ensuite que si $f \in \check{E}$, alors

$$f = f(u_1)\check{u}_1 + \cdots + f(u_n)\check{u}_n.$$

Pour cela, il suffit d'appliquer chaque membre à u_i pour $i = 1, \ldots, n$. On en déduit que $(\check{u}_1, \ldots, \check{u}_n)$ est une base. Pour démontrer la seconde formule, on écrit

$$u = \lambda_1 u_1 + \cdots \lambda_n u_n$$

et on calcule $\check{u}_i(u) = \lambda_i$.

□

Définition 11.3 *Avec les notations de la proposition, on dit que $\check{\mathcal{B}}$ est la* base duale *de $\mathcal{B}$.*

Par exemple, la base duale de la base canonique $(e_1, \ldots, e_n)$ de K^n est la base canonique $(p_1, \ldots, p_n)$ de $\check{K}^n$ (formée des projections sur les axes). En général, on retrouve que $\dim \check{E} = \dim E$.

Proposition 11.4 *Soit E un espace vectoriel sur K. Si $u \in E$, l'application*

$$\begin{tikzcd}[row sep=small] \check{E} \arrow[r, "\varphi_u"] & K \\ f \arrow[r, mapsto] & f(u) \end{tikzcd}$$

est une forme linéaire sur $\check{E}$. L'application induite

$$\begin{tikzcd}[row sep=small] E \arrow[r] & \check{\check{E}} \\ u \arrow[r, mapsto] & \varphi_u \end{tikzcd}$$

est linéaire. Si de plus, $\dim E < \infty$*, c'est un isomorphisme.*

Démonstration : Les deux première assertions se vérifient aisément : la première résulte de l'identité

$$(\lambda f + \mu g)(u) = \lambda f(u) + \mu g(u)$$

et la seconde de l'identité

$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v).$$

Si de plus, $\dim E < \infty$, on peut choisir une base $(u_1, \ldots, u_n)$ de E. Comme nos espaces ont même dimension finie, il suffit pour conclure de montrer que notre application est injective. Si u est dans le noyau, alors pour tout $i = 1, \ldots n$, on a $\check{u}_i(u) = \varphi_u(\check{u}_i) = 0$. Il suit que

$$u = \check{u}_1(u)u_1 + \cdots + \check{u}_n(u)u_n = 0.$$

□

Définition 11.5 *Si $f : E \to F$ est une application linéaire, l'*application duale *est*

$$\begin{tikzcd}[row sep=small] \check{f} : \check{F} \arrow[r] & \check{E} \\ g \arrow[r] & g \circ f. \end{tikzcd}$$

Proposition 11.6 *Si $f : E \to F$ est une application linéaire alors $\check{f}$ aussi.*

Démonstration : On a toujours

$$\check{f}(\lambda g + \mu h) = (\lambda g + \mu h) \circ f = \lambda g \circ f + \mu h \circ f = \lambda \check{f}(g) + \mu \check{f}(h)$$

si bien que $\check{f}$ est bien linéaire.

□

Définition 11.7 *La* transposée *de la matrice*

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$$

est la matrice

$$^tA := \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1m} & \cdots & a_{nm} \end{bmatrix}.$$

Proposition 11.8 *Si E et F sont de dimension finie, avec bases $\mathcal{B}$ et $\mathcal{C}$ respectivement, et si $A := [f]^{\mathcal{C}}_{\mathcal{B}}$, alors $[\check{f}]^{\check{\mathcal{B}}}_{\check{\mathcal{C}}} = {}^tA$.*

Démonstration : Si on écrit $\mathcal{B} = (u_1, \ldots, u_n)$ et $\mathcal{C} = (v_1, \ldots, v_m)$, on a

$$\check{f}(\check{v}_i)(u_j) = (\check{v}_i \circ f)(u_j) = \check{v}_i(f(u_j))$$

et l'assertion en résulte : la j-ème composante de $\check{f}(\check{v}_i)$ est identique à la i-ème composante de $f(u_j)$.

□

Proposition 11.9 *Si E et F sont deux espaces vectoriels, l'application*

$$\begin{aligned} L(E,F) &\longrightarrow L(\check{F}, \check{E}) \\ f &\longmapsto \check{f} \end{aligned}$$

est linéaire. C'est un isomorphisme si E et F sont de dimension finie.

Démonstration : On vérifie facilement que l'on a toujours

$$(\lambda_1\check{f}_1 + \lambda_2\check{f}_2) = (\lambda_1 f_1 + \lambda_2 f_2)\check{}.$$

Pour la seconde assertion, on peut fixer des bases et on voit immédiatement que si ${}^tA = 0$, alors $A = 0$. Il suit que le noyau de l'application est nul, donc que celle ci est injective et elle est donc bijective car les deux espaces ont même dimension.

□

Proposition 11.10 *Si $f : E \to F$ est une application linéaire, alors* $\mathrm{rg}(\check{f}) = \mathrm{rg}(f)$.

Démonstration : On note r le rang de f et on se donne une base $u_{r+1}, \ldots u_n$ de $\ker f$ que l'on complète en une base $u_1, \ldots, u_n$ de E. Il suit du théorème du rang que les vecteurs $v_1 := f(u_1), \ldots, v_r := f(u_r)$ sont linéairement indépendants et on complète en une base $v_1, \dots, v_m$ de F. Nous savons que $\mathrm{Im}(\check{f})$ est engendré par les formes linéaires $\check{f}(\check{v}_j)$ pour tout $j = 1, \dots m$. Nous allons les calculer.

Pour tout $i = 1, \dots, n$, on a $\check{f}(\check{v}_j)(u_i) = \check{v}_j(f(u_i))$ qui est nul si $i > r$ car alors $f(u_i) = 0$ et égal à $\check{v}_j(v_i)$ si $i \leq r$, c'est à dire 0 si $i \neq j$ et 1 sinon. Pour résumer, on a $\check{f}(\check{v}_j) = \check{u}_j$ si $j \leq r$ et 0 sinon. On voit donc que $\mathrm{Im}(\check{f})$ est engendré par $\check{u}_1, \dots, \check{u}_r$ et a donc même dimension r que $\mathrm{Im}(f)$.

□

Corollaire 11.11 *Si $A \in M_{n\times m}(K)$, on a* $\mathrm{Rg}({}^tA) = \mathrm{Rg}(A)$.

En particulier, on voit que le rang d'une matrice est égal au rang du système des ses vecteurs lignes.

Corollaire 11.12 *Soit $f : E \to F$ une application linéaire entre espaces vectoriels de dimension finie. Alors, f est injective (resp. surjective, resp. bijective) si et seulement si $\check{f}$ est surjective (resp. injective, resp. bijective).*

Démonstration : En effet, on sait que f est injective si et seulement si $\dim E = \mathrm{Rg}(f)$, ce qui est équivalent à $\dim \check{E} = \mathrm{Rg}(\check{f})$ qui veut dire que $\check{f}$ est surjective car $\check{E}$ est l'espace d'arrivée de $\check{f}$. On démontre de même la seconde assertion et la troisième en découle.

□

Nous obtenons maintenant le résultat dont nous avions besoin pour conclure la démonstration de la proposition 11.1.

Corollaire 11.13 *Soient E un espace vectoriel et $f_1, \dots, f_r \in \check{E}$. Alors, $f_1, \dots, f_r$ sont linéairement indépendants si et seulement si l'application*

$$\begin{array}{rcl} E & \xrightarrow{f} & K^r \\ u & \longmapsto & (f_1(u), \dots f_r(u)) \end{array}$$

est surjective. Ils sont générateurs si et seulement si l'application est injective. Enfin, c'est une base de $\check{E}$ si et seulement si l'application est bijective.

Démonstration : Par définition, si p_i désigne la projection sur le i-ème facteur, alors $\check{f}(p_i) = f_i$. Comme les projections forment une base de $\check{K}^r$, on voit que $f_1, \dots, f_r$ sont linéairement indépendants (resp. générateurs, resp. une base) si et seulement si $\check{f}$ est injective (resp. surjective, resp. bijective) et on vient de voir que c'est équivalent à f surjective (resp. injective, resp. bijective).

□

Proposition 11.14 *Soit E un espace vectoriel, $\mathcal{S}$ un ensemble de formes linéaires et*

$$F := \{u \in E, \quad \forall f \in \mathcal{S}, f(u) = 0\}.$$

Alors, $\dim F + \mathrm{Rg}(\mathcal{S}) = \dim E$.

Démonstration : On remarque tout d'abord que cet énoncé ne dépend que du sous-espace $\mathrm{Vect}(\mathcal{S}) \subset \check{E}$. On peut supposer qu'il est de dimension finie car sinon, E sera de dimension infinie. On peut donc supposer que $\mathcal{S} = (f_1, \dots, f_r)$ avec $f_1, \dots, f_r$ linéairement indépendantes. Et on applique le théorème du rang à l'application $f : E \to K^r$ correspondante, qui est surjective par le corollaire 11.13.

□

Proposition 11.15 *i) Si $f : E \to F$ et $g : F \to G$ sont deux applications linéaires, alors, $(g \circ f)\check{} = \check{f} \circ \check{g}$.*

ii) Si $A \in M_{n\times m}(K)$ et $B \in M_{m\times l}(K)$ alors ${}^t(AB) = {}^tB{}^tA$.

Démonstration : La seconde assertion résulte immédiatement de la première et celle ci se vérifie aisément : si $h \in \check{G}$, on a

$$(g \circ f)\check{}\,(h) = h \circ g \circ f$$

et

$$(\check{f} \circ \check{g})(h) = \check{f}(\check{g}(h)) = \check{g}(h) \circ f = h \circ g \circ f.$$

□

On finit avec une définition.

Définition 11.16 *Une matrice $A \in M_n(K)$ est* symétrique *(resp.* antisymétrique*) si elle satisfait ${}^tA = A$ (resp. ${}^tA = -A$).*

On peut montrer que les matrices symétriques (resp. antisymétriques) forment un sous-espace vectoriel $SM_n(K)$ (resp. $AM_n(K)$ et que si $1 + 1 \neq 0$ dans K, alors

$$M_n(K) = SM_n(K) \oplus AM_n(K).$$

12 Déterminants

On rappelle que le *groupe symétrique* S_n est l'ensemble de toutes les *permutations* σ de l'ensemble $\{1, \ldots, n\}$ (bijections de l'ensemble sur lui même). On note une permutation sous forme d'une matrice à deux lignes

$$\begin{bmatrix} 1 & \ldots & i & \cdots & n \\ \sigma(1) & \ldots & \sigma(i) & \cdots & \sigma(n) \end{bmatrix}$$

Une permutation qui échange simplement deux entiers distincts i est j est une *transposition* et on la note $(i \quad j)$. Plus généralement, une *permutation circulaire* (ou *cycle*) de longueur k envoie i_1 sur i_2 puis i_2 sur i_3 et ainsi de suite jusqu'à i_k qui est envoyé sur i_1. On la note $(i_1 \quad \cdots \quad i_k)$. Toute permutation peut s'obtenir en composant des transpositions. Toute permutation est composée d'un certains nombre de cycles disjoints.

La *signature* $\epsilon(\sigma)$ d'une transposition $(i \quad j)$ est -1 par définition. Plus généralement, la signature d'un cycle de longueur k est $(-1)^{k-1}$. En fait, on définit la signature d'une permutation quelconque σ comme $(-1)^i$ ou i est le nombre d'*inversions* de la permutation (chaque fois que $\sigma(j) < \sigma(i)$ alors que $i < j$). En fait ϵ est l'unique application de S_n sur $\{\pm 1\}$ qui satisfait $\epsilon(\sigma\tau) = \epsilon(\sigma)\epsilon(\tau)$ (pour $n > 1$).

Nous aurons aussi besoin de considérer le *groupe alterné* A_n de toutes les permutations *paires*, c'est à dire, dont la signature est 1. Remarquons que toutes les permutations impaires, dont la signature est -1, s'obtiennent en multipliant une transposition fixée τ par une permutation paire. On résume ça avec la formule $S_n = A_n \coprod \tau A_n$.

On passe maintenant à la définition du déterminant d'une matrice :

Définition 12.1 *Le* déterminant *de*

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

est

$$\det A := \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} := \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Par exemple, on a

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

et pour les matrices 3×3, on dispose de la *règle de Sarrus* (attention, ne fonctionne pas avec les matrices 4×4) :

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

Par exemple,

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

Enfin, notons aussi que si A est triangulaire (supérieure ou inférieure), alors $\det A = a_{11} \cdots a_{nn}$ car tous les autres termes ont au moins un facteur nul !

Proposition 12.2 *Si $A \in M_n(K)$, alors $\det {}^tA = \det A$.*

Cela signifie que toute propriété du déterminant relative aux colonnes sera aussi valable pour les lignes.

Démonstration : Pour $\sigma \in S_n$ fixé, on a $a_{\sigma(1)1} \cdots a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)}$ et $\epsilon(\sigma^{-1}) = \epsilon(\sigma)$ car leur produit est $\epsilon(\mathrm{Id})$ qui vaut 1. On a donc

$$\det {}^tA = \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} = \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma^{-1}) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)}$$

et en remplaçant σ^{-1} par σ, on trouve bien $\det A$.

□

On va maintenant développer la théorie des formes alternées.

Définition 12.3 *Soient $E_1, \ldots, E_n, F$ des espaces vectoriels. Une application*

$$\omega : E_1 \times \ldots \times E_n \to F$$

qui est linéaire en chaque variable est dite multilinéaire *ou plus précisément n-*linéaire*.*

Par exemple, les applications

$$\begin{aligned} K^2 \times K^2 &\longrightarrow K \\ ((a,b),(c,d)) &\longrightarrow ad - bc \end{aligned}$$

ou

$$\begin{aligned} K^2 \times K^2 &\longrightarrow K \\ ((a,b),(c,d)) &\longrightarrow ac + bd \end{aligned}$$

sont bilinéaires.

Définition 12.4 *Soit E un espace vectoriel. Une* forme n-linéaire *sur E est une application multilinéaire*

$$\omega : E^n := E \times E \times \ldots \times E \to K.$$

On dit que ω est alternée *si on a*

$$\omega(u_1, \ldots, u_n) = 0$$

chaque fois qu'il existe $i \neq j$ tel que $u_i = u_j$.

Les deux exemples ci-dessus sont des formes bilinéaires sur K^2 mais seule la première est alternée.

Proposition 12.5 *Si une forme multilinéaire ω est alternée et si $\sigma \in \mathcal{S}_n$, on a*

$$\omega(u_{\sigma(1)}, \ldots, u_{\sigma(n)}) = \epsilon(\sigma)\omega(u_1, \ldots, u_n)$$

Et la réciproque est vraie si $1 + 1 \neq 0$ dans K.

L'hypothèse est bien sûr satisfaite si $K = \mathbf{R}$ mais pas is $K = \mathbf{F}_2$.

Démonstration : Comme tout permutation est un produit de transpositions et que la signature est multiplicative, on peut supposer que σ est la transposition de i et de j. On a alors

$$\omega(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_n) + \omega(u_1, \ldots, u_j, \ldots, u_i, \ldots, u_n)$$

$$= \omega(u_1, \ldots, u_i + u_j, \ldots, u_i + u_j, \ldots, u_n) = 0.$$

Pour la réciproque, on utilise la même transposition. Si $u_i = u_j$, alors σ ne bouge pas les vecteurs et on en déduit que

$$\omega(u_1, \ldots, u_n) = -\omega(u_1, \ldots, u_n)$$

qui est donc nul si $-1 \neq 1$.

□

Proposition 12.6 *Soient E un espace vectoriel de dimension n et $f_1, \ldots, f_n \in \check{E}$. Alors, l'application*

$$\begin{array}{rcl} E^n & \longrightarrow & K \\ (u_1, \ldots, u_n) & \longmapsto & |f_i(u_j)| \end{array}$$

est multilinéaire alternée.

Démonstration : On vérifie d'abord que l'application est bien multilinéaire. On a fixé $f_1, \ldots, f_n$ et un entier $0 \leq j_0 \leq n$. On se donne $(u_1, \ldots, u_n)$ et $(v_1, \ldots, v_n)$ satisfaisant $u_j = v_j$ pour $j \neq j_0$ et on pose $w_j = u_j = v_j$ pour $j \neq j_0$ et $w_{j_0} = \lambda u_{j_0} + \mu v_{j_0}$ avec $\lambda, \mu \in K$. Il faut montrer que

$$|f_i(w_j)| = \lambda |f_i(u_j)| + \mu |f_i(v_j)|,$$

ou encore

$$\sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) f_1(w_{\sigma(1)}) \cdots f_n(w_{\sigma(n)})$$

$$= \lambda \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) f_1(u_{\sigma(1)}) \cdots f_n(u_{\sigma(n)}) + \mu \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) f_1(v_{\sigma(1)}) \cdots f_n(v_{\sigma(n)}).$$

Par linéarité, il suffit de vérifier ça pour chaque σ et il suffit de regarder le facteur avec $\sigma(i) = j_0$ car tous les autres sont identiques. On est donc réduit à montrer que $f_i(w_{j_0}) = \lambda f_i(u_{j_0}) + \mu f_i(v_{j_0})$ et cela résulte donc du fait que f_i est linéaire.

On montre maintenant que c'est une forme alternée. On suppose donc que $u_i = u_j$ avec $1 \leq i < j \leq n$ et on écrit $S_n = A_n \coprod \tau A_n$ ou τ est la transposition qui échange i et j. On voit que notre déterminant est égal à

$$\sum_{\sigma \in A_n} \epsilon(\sigma) f_1(u_{\sigma(1)}) \dots f_n(u_{\sigma(n)}) + \sum_{\sigma \in A_n} \epsilon(\tau\sigma) f_1(u_{\tau\sigma(1)}) \dots f_n(u_{\tau\sigma(n)})$$

et on veut montrer que c'est nul. Il suffit alors de remarquer que l'on a toujours $u_{\tau(k)} = u_k$ et que $\epsilon(\tau\sigma) = \epsilon(\tau)\epsilon(\sigma) = -\epsilon(\sigma)$.

$\square$

Corollaire 12.7 *Si $u_1, \dots, u_n \in E$, l'application*

$$\begin{aligned} \check{E}^n &\longrightarrow K \\ (f_1, \dots, f_n) &\longmapsto |f_i(u_j)| \end{aligned}$$

est multilinéaire alternée.

Démonstration : C'est tout simplement l'assertion analogue sur les lignes et on sait que si A est une matrice, on a $\det {}^tA = \det A$.

$\square$

Proposition 12.8 *i) Échanger deux colonnes d'un déterminant le transforme en son opposé.*

ii) Ajouter à une colonne d'un déterminant un multiple d'une autre colonne ne change pas le déterminant

iii) Multiplier une colonne d'un déterminant par un constante multiplie le déterminant par la même constante.

On a les résultats analogues pour les lignes.

Démonstration : C'est une conséquence immédiate du fait que le déterminant est multilinéaire et alterné en les colonnes. Et aussi que le déterminant ne change pas quand on conisdère la transposée.

$\square$

On peut donc calculer un déterminant par la méthode de Gauss qui permet de se ramener à un matrice triangulaire. Par exemple

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

Définition 12.9 *Soit E un espace vectoriel et $\mathcal{B} := (u_1, \ldots, u_n)$ une base de E. Le déterminant associé à $\mathcal{B}$ est l'application multilinéaire alternée associé à la base duale de $\mathcal{B}$:*

$$\begin{array}{ccc} E^n & \xrightarrow{\det_{\mathcal{B}}} & K \\ (v_1, \ldots, v_n) & \longmapsto & |\check{u}_i(v_j)|. \end{array}$$

En d'autres termes, si pour $i = 1, \ldots, n$, on a

$$v_i := \lambda_{1i} u_1 + \cdots + \lambda_{ni} u_n \in E,$$

alors $\check{u}_i(v_j) = \lambda_{ij}$ et il suit que

$$\det_{\mathcal{B}}(v_1, \ldots, v_n) = |\lambda_{ij}|.$$

Remarquons en particulier que

$$\det_{\mathcal{B}}(u_1, \ldots, u_n) = |I_n| = 1.$$

On peut aussi calculer les déterminants par récurrence sur l'ordre en utilisant le résultat suivant :

Proposition 12.10 *Soit $\mathcal{B} = (u_1, \ldots, u_n)$ une base de E et $v_1, \ldots, v_n \in E$. S'il existe i_0, j_0 tels que $v_{j_0} = u_{i_0}$, alors*

$$\det_{\mathcal{B}}(v_1, \ldots, v_n) = (-1)^{i_0+j_0} \det_{\mathcal{B}'}(v'_1, \ldots, \widehat{v'_{j_0}}, \ldots, v'_n)$$

avec $\mathcal{B}' = (u_1, \ldots, \widehat{u_{i_0}}, \ldots, u_n)$ et $v'_j = v_j - \check{u}_{i_0}(v_j) u_{i_0}$.

Cela veut dire que si on a une colonne composée uniquement de zéros sauf sur une ligne ou il y a un 1, on peut supprimer cette ligne et cette colonne, quite à prendre l'opposé.

Démonstration : Quitte à permuter les u_i avec le cycle $(i_0, \ldots, n)$, on peut supposer que $i_0 = n$. De même, comme le determinant est alterné, on peut supposer que $j_0 = n$: il suffit de faire agir le cycle $(j_0, \ldots, n)$ sur les indices. Et on doit donc démontrer que

$$\det_{\mathcal{B}}(v_1, \ldots, v_{n-1}, u_n) = \det_{\mathcal{B}'}(v'_1, \ldots, v'_{n-1})$$

avec $\mathcal{B}' = (u_1, \ldots, u_{n-1})$ et $v'_j = v_j - \check{u}_n(v_j) u_n$. On écrit pour tout $k = 1, \ldots, n$,

$$v_k := \lambda_{1k} u_1 + \cdots + \lambda_{nk} u_n \in E$$

si bien que, pour tout $k = 1, \ldots, n-1$,

$$v'_k := \lambda_{1k} u_1 + \cdots + \lambda_{n-1k} u_{n-1}.$$

Comme $v_n = u_n$, on a $\lambda_{nn} = 1$ et $\lambda_{ln} = 0$ pour $l \neq n$ et il suit que

$$\det_{\mathcal{B}}(v_1, \ldots, v_n) = \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) \lambda_{1\sigma(1)} \cdots \lambda_{n\sigma(n)}$$

$$= \sum_{\sigma \in \mathcal{S}_n, \sigma(n)=n} \epsilon(\sigma)\lambda_{1\sigma(1)} \cdots \lambda_{n-1\sigma(n-1)}.$$

$$= \sum_{\sigma \in \mathcal{S}_{n-1}} \epsilon(\sigma)\lambda_{1\sigma(1)} \cdots \lambda_{n-1\sigma(n-1)} = \det_{\mathcal{B}'}(v'_1, \ldots, v'_{n-1}).$$

□

Appliquons cette méthode. Tout d'abord, par linéarité, on a

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

Et il suit que

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

Proposition 12.11 *Soit $\mathcal{B} := (u_1, \ldots, u_n)$ une base de E et $v_1, \ldots, v_n \in E$. On a alors pour toute forme n-linéaire alternée sur E,*

$$\omega(v_1, \ldots, v_n) = \det_{\mathcal{B}}(v_1, \ldots, v_n)\omega(u_1, \ldots, u_n).$$

Démonstration :

On écrit pour tout $i = 1, \ldots, n$,

$$v_i := \lambda_{i1}u_1 + \cdots + \lambda_{in}u_n.$$

Par multilinéarité, on a

$$\omega(v_1, \ldots, v_n) = \sum_{i_1, \ldots, i_n \in \{1, \ldots, n\}} \lambda_{1i_1} \cdots \lambda_{ni_n}\omega(u_{i_1}, \ldots, u_{in}).$$

Comme ω est alternée, alors $\omega(u_{i_1}, \ldots, u_{in}) = 0$ chaque fois que deux indices sont identiques et on peut réécrire cette formule

$$\omega(v_1, \ldots, v_n) = \sum_{\sigma \in S_n} \lambda_{1\sigma(1)} \cdots \lambda_{n\sigma(n)}\omega(u_{\sigma(1)}, \ldots, u_{\sigma(n)})$$

$$= \sum_{\sigma \in S_n} \lambda_{1\sigma(1)} \cdots \lambda_{n\sigma(n)}\epsilon(\sigma)\omega(u_1, \ldots, u_n).$$

□

Notons en particulier que si $\omega(u_1, \ldots, u_n) = 0$, alors nécessairement $\omega = 0$.

Proposition 12.12 *Soit $\mathcal{B}$ une base de E et $v_1, \ldots, v_n \in E$. Alors, les conditions suivantes sont équivalentes :*

i) $\det_{\mathcal{B}}(v_1, \ldots, v_n) \neq 0$

ii) $v_1, \ldots, v_n$ *forment une base $\mathcal{C}$ de E*

On a alors

$$\det_{\mathcal{B}}(\mathcal{C}) \det_{\mathcal{C}}(\mathcal{B}) = 1.$$

Démonstration : Si les vecteurs sont liés, on peut écrire l'un d'entre eux, disons v_n pour simplifier, en fonction des autres :

$$v_n = \lambda_1 v_1 + \cdots + \lambda_{n-1} v_{n-1}.$$

On a alors,

$$\det_{\mathcal{B}}(v_1, \ldots, v_n) = \lambda_1 \det_{\mathcal{B}}(v_1, \ldots, v_{n-1}, v_1) + \cdots + \lambda_{n-1} \det_{\mathcal{B}}(v_1, \ldots, v_{n-1}, v_{n-1}) = 0$$

car $\det_{\mathcal{B}}$ est multilinéaire et alternée.

Inversement, si $v_1, \ldots, v_n$ forment une base $\mathcal{C}$ de E on applique la proposition précédente à $\omega = \det_{\mathcal{C}}$, ce qui donne

$$1 = \det_{\mathcal{B}}(v_1, \ldots, v_n) \det_{\mathcal{C}}(u_1, \ldots, u_n)$$

comme annoncé, et en particulier, le fait que $\det_{\mathcal{B}}(v_1, \ldots, v_n) \neq 0$.

□

Proposition 12.13 *Si E est un espace vectoriel de dimension $n > 0$, les formes n-linéaire alternées sur E forment un sous-espace vectoriel D de dimension 1 de l'espace de toutes les applications $E^n \to K$.*

Démonstration : On vérifie aisément que D est bien stable sous l'addition et par multiplication par un scalaire. Ensuite, on fixe une base $\mathcal{B} := (u_1, \ldots, u_n)$. On sait que $\det_{\mathcal{B}}$ est une forme linéaire alternée non nulle. Il suit que D est un sous-espace vectoriel de dimension ≥ 1.

On considère maintenant l'application d'évaluation en $\mathcal{B}$

$$\begin{aligned} D &\longrightarrow K \\ \omega &\longmapsto \omega(u_1, \ldots, u_n) \end{aligned}$$

qui est une forme linéaire. Son noyau est nul comme on l'a vu à la suite de la proposition 12.12 et elle est donc injective. Il suit que $\dim D \leq 1$ et on a fini.

□

En particulier, on voit que si $\mathcal{B}$ est une base de E, alors $\det_{\mathcal{B}}$ est une base de D.

Proposition 12.14 *Soit E un espace vectoriel de dimension n et $f \in L(E)$. Alors, il existe un unique* $\det f \in K$ *tel que, pour toute forme n-linéaire alternÃľe ω sur E et pour tout $v_1, \ldots, v_n \in E$, on ait*

$$\omega(f(v_1), \ldots, f(v_n)) = (\det f) \times \omega(v_1, \ldots, v_n).$$

En fait, si $\mathcal{B} = (u_1, \ldots, u_n)$ est une base de E, on a

$$\det f = \det_{\mathcal{B}}(f(u_1), \ldots, f(u_n)).$$

Démonstration : Si ω est une forme n-linéaire alternée, alors l'application

$$(v_1, \ldots, v_n) \mapsto \omega(f(v_1), \ldots, f(v_n))$$

est aussi n-linéaire alternée. C'est immédiat. Comme celles-ci forment un espace de dimension 1, si $\omega \neq 0$, il existe un unique $\lambda \in K$ tel que pour tout $v_1, \ldots, v_n \in E$, on ait

$$\omega(f(v_1), \ldots, f(v_n)) = \lambda \times \omega(v_1, \ldots, v_n).$$

Bien sûr, λ depend *a priori* de ω mais tout autre forme n-linéaire alternée est un multiple de ω et on retrouve donc le même λ que l'on note $\det f$.

La seconde assertion est un cas particulier de la première avec $\omega = \det_{\mathcal{B}}$ et $v_i = u_i$ pour tout i.

□

Définition 12.15 *On dit alors que* $\det f$ *est le* déterminant *de f.*

Notez que celui-ci ne dépend pas du choix d'une base !

Proposition 12.16 *Soit E un espace vectoriel de dimension finie et $f, g \in L(E)$. On a alors*

$$\det(g \circ f) = (\det g)(\det f).$$

De plus, f est bijective si et seulement si $\det f \neq 0$.

Démonstration : La première assertion est immédiate : si $\dim E = n$, ω est une forme n-linéaire alternée et $v_1, \ldots, v_n \in E$, on a

$$\omega((g \circ f)(v_1), \ldots, (g \circ f)(v_n)) = (\det g) \times \omega(f(v_1), \ldots, f(v_n))$$

$$= (\det g)(\det f) \times \omega(v_1, \ldots, v_n).$$

La seconde résulte de la proposition 12.12.

□

Proposition 12.17 *Soit E un espace vectoriel de dimension finie et $\mathcal{B}$ une base de E. On a alors*

$$\det[f]_{\mathcal{B}} = \det f.$$

Démonstration : C'est la formule de la proposition 12.14.

□

Corollaire 12.18 *Si $A, B \in M_n(K)$, alors $\det(AB) = (\det A)(\det B)$. De plus, A est inversible si et seulement si $\det A \neq 0$.*

Démonstration : Traduction à partir des applications linéaires.

□

Par exemple, on a

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

13 Diagonalisation

On suppose traditionnellement que les espaces vectoriels sont de *dimension finie* dans ce qui suit.

Définition 13.1 *Soit f un endomorphisme d'un espace vectoriel E. Alors, $\lambda \in K$ est une* valeur propre *pour f s'il existe $u \neq 0$ tel que $f(u) = \lambda u$. On dit alors que u est un* vecteur propre *pour f.*

Exemples :

i) L'application nulle n'a qu'une valeur propre : 0.
ii) L'identité n'a qu'une valeur propre : 1.
iii) Une projection a (au plus) deux valeurs propre : 0 et 1.
iv) Une symétrie a (au plus) deux valeur propres : 1 et -1.
v) L'homothétie de rapport λ a une seule valeur propre : λ.
vi) La rotation d'angle θ n'a pas de valeur propre réelle sauf si θ est l'angle plat (mais une valeur propre complexe $e^{i\theta}$).
vii) l'application linéaire

$$\begin{array}{ccc} \mathbf{R}^2 & \xrightarrow{f} & \mathbf{R}^2 \\ (x, y) & \longmapsto & (x+2y, 2x+y) \end{array}$$

satisfait $f(1,1) = (3,3) = 3(1,1)$ et $f(1,-1) = (-1,1) = -(1,-1)$ si bien que 3 et -1 sont des valeurs propres.

Proposition 13.2 *Soit f un endomorphisme d'un espace vectoriel E, $\lambda \in K$ et*

$$E_\lambda := \ker(\lambda \mathrm{Id}_E - f).$$

Alors, $E_\lambda \neq \{0\}$ si et seulement si λ est une valeur propre de f et alors, les vecteurs propres associés à λ sont les vecteurs non nuls de E_λ.

Démonstration : En effet, dire que $u \in E_\lambda$ signifie que $(\lambda \mathrm{Id}_E - f)(u) = 0$ ou encore que $\lambda u = f(u)$.

□

Définition 13.3 *Si λ est une valeur propre pour f, on dit que E_λ est un* sous-espace propre *pour f.*

Par exemple, si p est la projection sur F de direction G, le sous-espace propre associé à 0 est G et celui associé à 1 est F. Les sous-espaces propres de l'application $(x, y) \mapsto (x+2y, 2x+y)$ sont les droites d'équations $x - y = 0$ et $x + y = 0$.

Proposition 13.4 *Soit f un endomorphisme d'un espace vectoriel E. Alors f est bijectif si et seulement si 0 n'est pas valeur propre pour f. Dans, ce cas, si λ est une valeur propre pour f, alors λ^{-1} est une valeur propre pour f^{-1} avec même sous-espace propre.*

Démonstration : Dire que 0 est valeur propre pour f signifie qu'il existe $u \neq 0$ tel que $f(u) = 0$ c'est à dire que $\ker f = 0$, ce qui signifie que f n'est pas bijective puisque $\dim E < \infty$.

Si f est bijective et u est un vecteur propre de f pour une valeur propre λ, on a $f(u) = \lambda u$ si bien que $u = f^{-1}(\lambda u)$ et comme f^{-1} est linéaire, $u = \lambda f^{-1}(u)$ et donc finalement $\lambda^{-1}u = f^{-1}(u)$.

□

Proposition 13.5 *Soit f un endomorphisme d'un espace vectoriel de dimension finie E et λ une valeur propre de f. Si $k \in \mathbf{N}$, alors λ^k est valeur propre de f^k avec même sous-espace propre. Cela est toujours valide pour $k < 0$ si f est bijective.*

Démonstration : On sait déjà que $1 = \lambda^0$ est valeur propre pour $\mathrm{Id}_E = f^0$. On montre ensuite l'assertion par récurrence sur $k > 0$. On suppose donc que λ est valeur propre de f et que $f^k(u) = \lambda^k u$. On en déduit que

$$f^{k+1}(u) = f^k(f(u)) = f^k(\lambda u) = \lambda f^k(u) = \lambda\lambda^k u = \lambda^{k+1}u.$$

Enfin, si f est bijective, on sait que λ^{-1} est valeur propre pour f^{-1} avec même sous-espace propre que λ. D'autre part, on a $\lambda^{-k} = (\lambda^{-1})^k$ et $f^{-k} = (f^{-1})^k$.

□

C'est comme ça qu'on détermine les valeurs propres d'une symétrie ($s^2 = \mathrm{Id}$) ou d'une projection ($p^2 = p$) par exemple.

Définition 13.6 *Si f est un endomorphisme d'un espace vectoriel E, alors*

$$P := \det(X\mathrm{Id}_E - f) \in K[X]$$

est le polynôme caractéristique *de f.*

Pour l'endomorphisme $(x, y) \mapsto (x + 2y, 2x + y)$, par exemple, on trouve

$$\begin{vmatrix} X-1 & -2 \\ -2 & X-1 \end{vmatrix} = (X-1)^2 - 4 = (X-3)(X+1).$$

Le célèbre *théorème de Cayley-Hamilton* dit que l'on a toujours $P(f) = 0$. Dans notre exemple, on aura donc $f^2 - 2f + 3\mathrm{Id} = 0$. Nous n'aurons pas le temps de discuter ce théorème.

Proposition 13.7 *Soient f un endomorphisme d'un espace vectoriel E et P son polynôme caractéristique. Alors, $\lambda \in K$ est valeur propre de f si et seulement si λ est racine de P.*

Démonstration : En effet, dire que λ est racine de P signifie que $\det(\lambda \mathrm{Id}_E - f) = 0$, ce qui signifie que $\lambda \mathrm{Id}_E - f$ n'est pas bijective ou encore que son noyau E_λ n'est pas nul.

□

Définition 13.8 *Un endomorphisme f d'un espace vectoriel E est* diagonalisable *s'il existe une base formée de vecteurs propres.*

Théorème 13.9 *Soit f un endomorphisme d'un espace vectoriel de dimension finie E, $\lambda_1, \ldots, \lambda_k$ ses valeurs propres et $E_1, \ldots, E_k$ les sous-espaces propres. Alors,* $\dim E_1 + \cdots + \dim E_k \leq \dim E$ *avec égalité si et seulement si f est diagonalisable. C'est le cas en particulier si* $\dim E = n$ *et qu'il y a n valeurs propres distinctes.*

Lemme 13.10 *Si $F_i := E_1 + \cdots + E_i$, alors $F_{i+1} = F_i \oplus E_{i+1}$. De plus, F_i possède une base formée de vecteurs propres.*

Démonstration du lemme : Pour la première assertion, il suffit bien sûr de montrer que $F_i \cap E_{i+1} = 0$ et on va en fait montrer que $F_j \cap E_{i+1} = 0$ pour $j \leq i$. On procède par récurrence sur $i + j$. Si $u \in E_{i+1}$, on a par définition $f(u) = \lambda_{i+1} u$ mais si on a aussi $u \in F_j$, on peut écrire $u = u_1 + \cdots + u_j$ avec $u_1 \in E_1, \ldots, u_j \in E_j$ et on a donc aussi

$$f(u) = \lambda_1 u_1 + \cdots + \lambda_j u_j.$$

On en déduit que

$$((\lambda_{i+1} - \lambda_1)u_1 + \cdots + (\lambda_{i+1} - \lambda_{j-1})u_{j-1}) + (\lambda_{i+1} - \lambda_j)u_j = 0.$$

Par récurrence, on voit que les deux termes sont nuls et en particulier que $(\lambda_{i+1} - \lambda_j)u_j = 0$, ce qui montre que $u_j = 0$ et donc que $u \in F_{j-1}$ et on a fini (par récurrence).

La seconde assertion se démontre grace à la première par récurrence sur i puisqu'une base de F_{i+1} est composé d'une base de F_i et d'une base de E_{i+1}.

□

Démonstration du théorème :

Il résulte du lemme que $\dim E_1 + \cdots + \dim E_k = \dim F_k \leq \dim E$ avec égalité si et seulement si $F_k = E$. Et dans ce dernier cas, E aura une base formée de vecteurs propres. Réciproquement, si $\mathcal{B}$ est une base de E formée de vecteurs propres, alors $\mathcal{B} \subset \cup E_i \subset F_k$ et c'est donc aussi une base de F_k qui est donc nécessairement égal à E.

Bien sûr, si on a n valeurs propres distinctes, alors $k = n$ et comme chaque $E_i \neq 0$, on a $\dim F_n \geq n$ et donc $E = F_n$.

□

Remarquons que la dernière condition n'est pas nécessaire car l'identité à une seule valeur propre, quelle que soit la dimension de E mais est pourtant diagonalisable, et même diagonale ! Le phénomène analogue se produit plus généralement avec les projections ou les symétries par exemple.

Corollaire 13.11 *Si le polynôme caractéristique de f a toutes ses racines dans K et que celles-ci sont distinctes, alors f est diagonalisable.*

Démonstration : On a vu que ce sont justement les valeurs propres de f.

□

Remarquons que la première condition est toujours satisfaite si $K = \mathbf{C}$.

Définition 13.12 *Le* polynôme caractéristique *de $A \in M_n(K)$ est $P := \det(XI_n - A) \in K[X]$. Les* valeurs propres *de A sont les racines de P dans K. La matrice A est* diagonalisable *si elle est semblable à une matrice diagonale.*

Proposition 13.13 *Si P est le polynôme caractéristique de $A \in M_n(K)$, alors*

$$P = X^n - tX^{n-1} + \cdots + (-1)^n\delta$$

ou $\delta = \det A$ et $t = \mathrm{tr}A$ est la somme des éléments diagonaux.

Démonstration : Clairement, on a $P(0) = \det(-A) = (-1)^n \det A$. Pour la trace, c'est un peu plus compliqué et il faut regarder ce qui apparaît dans le coefficient de X^{n-1} quand on calcule le déterminant.

□

Notons en particulier que $\det A$ est égal au produit des valeurs propres et que $\mathrm{tr}A$ est égal à la somme (si on les prend avec leur multiplicité et dans un corps suffisement grand)

Proposition 13.14 *Soit A la matrice d'un endomorphisme f d'un espace vectoriel E dans une base $\mathcal{B}$. Alors, le polynôme caractéristique de A est identique au polynôme caractéristique de f, les valeurs propres de A sont identiques aux valeurs propres de F et enfin, A est diagonalisable si et seulement si f est diagonalisable.*

Démonstration : La première assertion résulte du fait que $XI_n - A$ est la matrice de $X\mathrm{Id}_E - f$ dans la base $\mathcal{B}$. La seconde en découle. Enfin, dire que f est diagonalisable signifie qu'il existe une base formée de vecteurs propres, c'est à dire, une matrice de passage P telle que $A' := P^{-1}AP$ soit diagonale.

□

Quelques exemples pour clore ce chapitre : on considère les matrices suivantes de $M_2(\mathbf{R})$:

$$A := \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad B := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad C := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} et \quad \text{et} \quad D := \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

i) La matrice A a pour polynôme caractéristique $X^2 - 2X - 3$ et donc deux valeurs propres distinctes 3 et -1. Elle est diagonalisable.
ii) La matrice B a pour polynôme caractéristique $X^2 + 1$ et n'a donc aucune valeur propre. Elle n'est pas diagonalisable.
iii) La matrice C a pour polynôme caractéristique $X^2 - 2X + 1$ et a donc 1 pour unique valeur propre. Et le sous espace propre est l'axe des x ! Elle n'est donc pas diagonalisable.
iv) La matrice D a pour polynôme caractéristique $X^2 + 2X + 1$ et a donc -1 pour unique valeur propre. Mais cette fois ci, elles est diagonalisable (et même diagonale).

14 Produit scalaire

Définition 14.1 *La matrice d'une forme bilinéaire $\varphi : E \times E \to K$ dans une base $\mathcal{B} := (u_1, \ldots, u_n)$ de l'espace vectoriel E est*

$$[\varphi]_{\mathcal{B}} := [\varphi(u_i, u_j)].$$

Par exemple, l'application

$$\begin{array}{ccc} \mathbf{R}^2 \times \mathbf{R}^2 & \longrightarrow & \mathbf{R} \\ ((a,b),(c,d)) & \longrightarrow & ac + bd \end{array}$$

a pour matrice $I := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ dans la base canonique. Et le déterminant

$$\begin{array}{ccc} \mathbf{R}^2 \times \mathbf{R}^2 & \longrightarrow & \mathbf{R} \\ ((a,b),(c,d)) & \longrightarrow & ad - bc \end{array}$$

a pour matrice $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

Proposition 14.2 *Si φ est une forme bilinéaire sur E et $\mathcal{B}$ une base E, on a pour tout $u, v \in E$,*

$$\varphi(u, v) = {}^t[u]_{\mathcal{B}}[\varphi]_{\mathcal{B}}[v]_{\mathcal{B}}.$$

Démonstration : Pour alléger les notations, on omet le nom de la base dans les indices. On peut écrire

$$[u] = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} \quad \text{et} \quad [v] = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}.$$

Par bilinéarité, il suit que dans la base $\mathcal{B} = (u_1, \ldots, u_n)$, on a

$$\varphi(u, v) = \varphi(\lambda_1 u_1 + \cdots + \lambda_n u_n, \mu_1 u_1 + \cdots + \mu_n u_n)$$

n'est autre que la somme de tous les $\lambda_i \mu_j \varphi(u_i, u_j)$ pour $i, j = 1, \ldots, n$. D'autre part, si on pose $b_{ij} := \varphi(u_i, u_j)$, on voit que ${}^t[u][\varphi]$ est le vecteur ligne dont la j-ème composante est $\lambda_1 b_{1j} + \cdots + \lambda_n b_{nj}$. Pour obtenir${}^t[u][\varphi]t[v]$, on multiplie par μ_j et on somme sur j. On trouve bien la somme de tous les $\lambda_i b_{ij} \mu_j$.

$\square$

Définition 14.3 *Soit E un espace vectoriel sur $\mathbf{R}$. Une forme bilinéaire*

$$(u, v) \longmapsto \langle u, v \rangle$$

sur E est symétrique *si pour tout $u, v \in E$, on a*

$$\langle v, u \rangle = \langle u, v \rangle.$$

Et bien sûr, linéaire par rapport à la seconde variable signifie que l'on a toujours

$$\langle u, \lambda_1 v_1 + \lambda_2 v_2 \rangle = \lambda_1 \langle u, v_1 \rangle + \lambda_2 \langle u, v_2 \rangle.$$

Par symétrie, elle sera alors automatiquement linéaire aussi par rapport à la première variable.

On voit que le premier exemple ci-dessus est symétrique mais pas le second. En fait, une application bilinéaire est symétrique si et seulement si sa matrice dans une base quelconque est symétrique.

Définition 14.4 *Une forme bilinéaire symétrique $\langle -, - \rangle$ sur un espace vectoriel E est* définie positive *ou encore, un* produit scalaire*, si on a la propriété*

$$\langle u, u \rangle > 0 \Leftrightarrow u \neq 0.$$

L'exemple ci-dessus est bien un produit scalaire car on a bien $a^2 + b^2 > 0$ si et seulement si $(a, b) \neq (0, 0)$.

On admettra le théorème suivant qui est bien pratique :

Théorème 14.5 *Une forme bilinéaire symétrique est un produit scalaire si et seulement si les valeurs propres de sa matrice dans une base quelconque sont strictement positives.*

Démonstration : Admis.

□

Définition 14.6 *Soit E un espace vectoriel sur $\mathbf{R}$. Une* norme *sur E est une application $\| - \| : E \to \mathbf{R}$ telle que*

i) Si $u \in E$, alors $\|u\| > 0 \Leftrightarrow u \neq 0$

ii) Si $u, v \in E$ alors $\|u + v\| \leq \|u\| + \|v\|$

iii) si $\lambda \in \mathbf{R}$ et $u \in E$, alors $\|\lambda u\| = |\lambda| \|u\|$

Sur $\mathbf{R}^2$, on a par exemple les normes $\|(x, y)\|_1 = |x| + |y|$ ou $\|(x, y)\|_2 = \sqrt{x^2 + y^2}$ ou encore $\|(x, y)\|_\infty = \max(|x|, |y|)$ comme on peut aisément le vérifier.

Proposition 14.7 *Soit $\langle -, - \rangle$ un produit scalaire. Alors,*

i) La fonction

$$\| - \| : E \to \mathbf{R}, u \longmapsto \sqrt{\langle u, u \rangle}$$

est une norme sur E et on a toujours

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

ii) (Inégalité de Cauchy-Schwartz) On a toujours

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

Démonstration : On montre d'abord l'inégalité de Cauchy-Schwartz. On peut supposer $u \neq 0$. On regarde le polynôme

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

Celui-ci est toujours positif (ou nul). Son discriminant est donc négatif (ou nul) et on a donc

$$\langle u, v\rangle^2 - \|u\|^2\|v\|^2 \leq 0.$$

Il suit que

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

On vérifie ensuite que l'on a bien une norme. La première et la dernière condition résultent directement des définitions. La seconde est conséquence de l'inégalité de Cauchy-Schwartz : on a d'une part

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

et d'autre part

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

Enfin, pour la formule, il suffit de développer le membre de droite.

□

L'négalité de Cauchy-Schwartz est très importante car elle permet de définir l'*angle* entre deux vecteurs non nuls u et v comme l'unique $\theta \in [-\pi/2, \pi/2]$ tel que

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\|v\|}.$$

Définition 14.8 *Un* espace vectoriel euclidien *est un espace vectoriel de dimension finie sur* $\mathbf{R}$ *muni d'un produit scalaire* $\langle -, -\rangle$.

Sur $\mathbf{R}^n$, le produit scalaire standard est

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle = x_1y_1 + \cdots + x_ny_n$$

et la norme associée est

$$\|(x_1, \ldots, x_n)\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Définition 14.9 *Soit E un espace vectoriel euclidien. Deux vecteurs u et v sont* orthogonaux *si $\langle u, v\rangle = 0$ et on écrit alors $u \perp v$.*

Par exemple, dans $\mathbf{R}^2$ avec le produit scalaire habituel, on a donc $(1,0)\perp(0,1)$.

Définition 14.10 *Soit E un espace vectoriel euclidien.*

i) Deux parties $\mathcal{S}$ et $\mathcal{T}$ de E sont orthogonales *si pour tout $u \in \mathcal{S}$ et $v \in \mathcal{T}$, on a $u \perp v$. On écrira alors $\mathcal{S} \perp \mathcal{T}$*

ii) La partie orthogonale *à une partie $\mathcal{S}$ de E est*

$$\mathcal{S}^\perp := \{u \in E, \quad u \perp \mathcal{S}\}.$$

On voit donc que $\mathcal{S} \perp \mathcal{T}$ si et seulement si $\mathcal{T} \subset \mathcal{S}^\perp$.

Par exemple, dans $\mathbf{R}^2$, l'axe des x est l'orthogonal à l'axe des y.

Proposition 14.11 *Soit E un espace vectoriel euclidien. Alors,*

i) Si $\mathcal{S}$ est une partie de E, on a $\mathcal{S} \subset \mathcal{S}^{\perp\perp}$.

ii) Si $\mathcal{S} \subset \mathcal{T} \subset E$, alors $\mathcal{T}^\perp \subset \mathcal{S}^\perp \subset E$.

Démonstration : Si $u \in \mathcal{S}$ et $v \in \mathcal{S}^\perp$, alors $u \perp v$ et par symétrie $v \perp u$. Il suit que $u \in \mathcal{S}^{\perp\perp}$ et on a donc bien $\mathcal{S} \subset \mathcal{S}^{\perp\perp}$. Supposons maintenant que $\mathcal{S} \subset \mathcal{T}$ et donnons nous $u \in \mathcal{T}^\perp$ et $v \in S$. On a bien évidemment $v \in T$ et donc $u \perp v$, ce qui montre que $u \in \mathcal{S}^\perp$.

□

Lemme 14.12 *Soit E un espace vectoriel euclidien. Si $u \in E$, l'application*

$$\begin{array}{ccc} E & \xrightarrow{\langle u,-\rangle} & \mathbf{R} \\ v & \longmapsto & \langle u,v\rangle \end{array}$$

est une forme linéaire. De plus, l'application

$$\begin{array}{ccc} E & \longrightarrow & \check{E} \\ u & \longmapsto & \langle u,-\rangle \end{array}$$

est un isomorphisme d'espace vectoriel.

Démonstration : La première assertion exprime simplement la linéarité par rapport à la première variable. Et le fait que la seconde application est linéaire exprime la linéarité par rapport à l'autre variable. Enfin, comme les espaces ont même dimension, pour montrer qu'on a un isomorphism, il suffit de vérifier que le noyau est trivial. Or si l'application $\langle u,-\rangle$ est nulle, on a en particulier $\|u\|^2 = \langle u,u\rangle = 0$ et donc $u = 0$.

□

Par exemple, dans $\mathbf{R}^n$ avec le produit scalaire standard, on voit que $\langle e_i, u\rangle = p_i(u)$ et donc que l'isomorphisme envoie la base canonique de $\mathbf{R}^n$ sur le base canonique de $\check{\mathbf{R}}^n$. En d'autres termes, elle transforme un vecteur colonne en vecteur ligne.

Proposition 14.13 *Soit E un espace vectoriel euclidien. Si $\mathcal{S} \subset E$, $\mathcal{S}^\perp$ est un sous-espace vectoriel de E et*

$$\dim E = \dim \mathcal{S}^\perp + \mathrm{rg}\mathcal{S}.$$

Démonstration : Sous l'isomorphisme

$$\begin{aligned} E &\xrightarrow{\simeq} \check{E} \\ u &\longmapsto \langle u, -\rangle, \end{aligned}$$

une partie $\mathcal{S}$ de E correspond à une partie $\mathcal{T} \subset \check{E}$ et on a

$$\mathcal{S}^\perp = \{u \in E, \quad \forall f \in \mathcal{T}, f(u) = 0\}.$$

L'assertion résulte donc de la proposition 11.14

$\square$

Proposition 14.14 *Soit E un espace vectoriel euclidien. Alors,*

i) Si F est un sous-espace vectoriel de E, on a

$$E = F \oplus F^\perp \text{ et } F^{\perp^\perp} = F.$$

ii) Si F et G sont deux sous-espace vectoriel de E, alors

$$(F+G)^\perp = F^\perp \cap G^\perp$$

et

$$(F \cap G)^\perp = F^\perp + G^\perp.$$

Démonstration : Si F est un sous-espace vectoriel de E, il est clair que $F \cap F^\perp = 0$. De plus, comme on vient de le voir,

$$\dim E = \dim F^\perp + \dim F.$$

On a donc bien

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

On voit aussi que $\dim F^{\perp^\perp} = \dim F$ et comme $F \subset F^{\perp^\perp}$, on a nécessairement $F^{\perp^\perp} = F$.

Montrons maintenant la seconde assertion. On a $F, G \subset F+G$ et donc $(F+G)^\perp \subset F^\perp, G^\perp$ par renversement d'inclusion. Si H est un sous-espace vectoriel de E tel que $H \subset F^\perp, G^\perp$, alors $F, G \subset H^\perp$ et donc $F + G \subset H^\perp$ si bien que $H \subset (F+G)^\perp$. La dernière égalité se démontre exactement de la même manière : On a $F^\perp, G^\perp \subset (F \cap G)^\perp$ par renversement d'inclusion. D'autre part, si H est un sous-espace vectoriel de E tel que $F^\perp, G^\perp \subset H$, alors $H^\perp \subset F, G$ et donc $H^\perp \subset F \cap G$ si bien que $(F \cap G)^\perp \subset H$.

□

Définition 14.15 *Si F est un sous-espace vectoriel d'un espace euclidien E, la* projection orthogonale *sur F est la projection associée à la somme directe $E = F \oplus F^{\perp}$.*

En d'autre termes, si $u \in E$ il existe un unique $v \in F$ tel que $(v-u)\perp F$: c'est la projection orthogonale de u sur F. En fait, il suffit de vérifier que $(v-u)\perp u_i$ pour une base (u_i) de F.

Par exemple, pour trouver le projetÃĪ orthogonal du vecteur $(1,5)$ sur la droite Δ d'équation $y = x$ dans $\mathbf{R}^2$, il faut résoudre simultanément

$$y = x \quad \text{et} \quad \langle (x-1, y-5), (1,1)\rangle = 0$$

car $(1,1)$ est une base de Δ. On a donc $x = y = 3$ et on trouve le vecteur $(3,3)$.

Définition 14.16 *Un système $\mathcal{S}$ d'éléments non-nuls de E est* orthogonal *si pour tout $u \neq v$ dans $\mathcal{S}$, on a $u\perp v$. Il est* orthonormal *si de plus, pour tout $u \in \mathcal{S}$, on a $\|u\| = 1$.*

Par exemple, le système $((1,0),(0,1))$ est orthonormal dans $\mathbf{R}^2$ muni du produit scalaire standard contrairement à $((1,1),(1,-1))$ qui est cependant orthogonal.

Proposition 14.17 *Une base $\mathcal{B} := (u_1, \ldots, u_n)$ est orthonormale si et seulement si pour tout $i = 1, \ldots, n$, on a $\check{u}_i = \langle u_i, -\rangle$.*

Démonstration : Pour que deux applications linéaires coincident, il suffit qu'elles prennent les mêmes valeurs sur tous les vecteurs d'une base. Ici, cela se traduit par $\check{u}_i(u_j) = \langle u_i, u_j\rangle$ pour tout $i, j = 1, \ldots, n$, c'est à dire $\langle u_i, u_j\rangle = 0$ pour $i \neq j$ et $\langle u_i, u_i\rangle = 1$ pour tout i. Où encore, $u_i \perp u_j$ pour $i \neq j$ et $\|u_i\| = 1$ pour tout i.

□

Proposition 14.18 *Soit F un sous-espace vectoriel d'un espace vectoriel euclidien E et $(u_1, \ldots, u_m)$ une base orthonormale de F. Alors, la projection orthogonale sur F est donnée par*

$$u \mapsto \langle u, u_1\rangle u_1 + \cdots + \langle u, u_m\rangle u_m.$$

Démonstration : On note pour simplifier $\lambda_i := \langle u, u_i\rangle$ et on pose

$$v := \lambda_1 u_1 + \cdots + \lambda_m u_m.$$

Comme $(u_1, \ldots, u_m)$ une famille orthonormale, on voit que, pour tout $i = 1, \ldots, m$, on a

$$\langle v, u_i\rangle = \langle \lambda_1 u_1 + \cdots + \lambda_m u_m, u_i\rangle = \lambda_1\langle u_1, u_i\rangle + \cdots + \lambda_m\langle u_m, u_i\rangle = \lambda_i$$

On a donc, si on pose $w = u - v$,

$$\langle w, u_i \rangle = \langle u - v, u_i \rangle = \langle u, u_i \rangle - \langle u, u_i \rangle = \lambda_i - \lambda_i = 0$$

si bien que $w \perp F$. Il suit que $u = v + w$ avec $v \in F$ et $w \in F^\perp$. Cela veut bien dire que v est la projection orthogonale de u sur F.

□

Proposition 14.19 *Soit E un espace vectoriel euclidien. Alors,*

i) Tout système orthogonal est libre.

ii) Tout système orthogonal (resp. orthonormal) est contenu dans une base orthogonale (resp. orthonormale).

iii) L'espace E possède une base orthonormale.

Démonstration : Soit $\mathcal{S}$ un système orthogonal. Si on a une combinaison linéaire nulle d'éléments de $\mathcal{S}$,

$$\lambda_1 u_1 + \cdots + \lambda_r u_r = 0,$$

alors pour tout $i = 1, \dots r$, on a

$$\langle u_i, \lambda_1 u_1 + \cdots + \lambda_r u_r \rangle = 0$$

et donc $\lambda_i \|u_i\|^2 = 0$, ce qui donne $\lambda_i = 0$ car $u_i \neq 0$. Il suit que $\mathcal{S}$ est libre.

Montrons ensuite par récurrence sur la dimension de E que tout système orthogonal (resp. orthonormal) *non-vide* $\mathcal{S}$ est contenu dans une base orthogonale (resp. orthonormale). Soit $F = \mathcal{S}$. On sait que $E = F \oplus F^\perp$. Comme $\mathcal{S}$ est libre, c'est une base de F et comme $\mathrm{Rg}(\mathcal{S}) > 0$, on a $\dim F^\perp < \dim E$. Par récurrence, il possède une base orthonormale qui nous permet de compléter $\mathcal{S}$. C'est clairement une base orthonormale, et même orthonormale si $\mathcal{S}$ l'est.

Pour finir, il suffit de montrer que si $E \neq 0$, il existe des systèmes orthonormaux non vides : on peut trouver un vecteur $u \neq 0$ et on prend le système réduit au vecteur $\frac{u}{\|u\|}$.

□

La variante constructive de cette proposition est le procédé d'orthonormalisation de Gram-Schmidt : On part d'une base $(u_1, \dots, u_m)$ de E et on pose

$$v_1 := u_1, \quad v_2 = u_2 - \frac{\langle u_2, u_1 \rangle}{\|u_1\|^2} u_1, \quad v_3 = u_3 - \frac{\langle u_3, u_1 \rangle}{\|u_1\|^2} u_1 - \frac{\langle u_3, u_2 \rangle}{\|u_2\|^2} u_2, \quad \dots$$

pour obtenir une base orthogonale puis on normalise :

$$w_1 := \frac{v_1}{\|v_1\|}, \quad w_2 := \frac{v_2}{\|v_2\|}, \quad \dots .$$

Par exemple, si on veut une base orthonormale du plan d'équation $x + y + z = 0$, on part de la base formée de $(1, -1, 0)$ et $(1, 0, -1)$ et on remplace le second vecteur par

$$(1, 0, -1) - \frac{\langle (1, -1, 0), (1, 0, -1)\rangle}{\|(1, -1, 0)\|^2}(1, -1, 0) = (1, 0, -1) - \frac{1}{2}(1, -1, 0) = (\frac{1}{2}, \frac{1}{2}, -1)$$

pour obtenir une base orthogonale puis on normalise en divisant respectivement par

$$\|v_1\| = \sqrt{2} \quad \text{et} \quad \|v_1\| = \sqrt{\frac{3}{2}}$$

pour trouver

$$(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0) \quad \text{et} \quad (\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{3}).$$

Proposition 14.20 *Soit $(u_1, \ldots, u_n)$ une base orthogonale d'un espace affine euclidien. Alors,*

i) Les coordonnées de $v \in E$ sont

$$\frac{\langle v, u_1\rangle}{\|u_1\|^2}, \quad \ldots, \quad \frac{\langle v, u_n\rangle}{\|u_n\|^2}.$$

ii) Si

$$v = \lambda_1 u_1 + \cdots + \lambda_n u_n$$

et

$$w = \mu_1 u_1 + \cdots + \mu_n u_n,$$

alors

$$\langle w, z\rangle = \lambda_1\mu_1\|u_1\|^2 + \cdots + \lambda_n\mu_n\|u_n\|^2.$$

Démonstration : Un rapide calcul nous donne la seconde formule et on en déduit la première, car alors

$$\langle v, u_i\rangle = \lambda_i\|u_i\|^2.$$

□

15 Quelques exercices

Exercice 1 *Soient F et G deux sous-espaces vectoriels d'un espace vectoriel E. Montrer que si $F \cup G$ est un sous-espace vectoriel de E, alors $F \subset G$ ou $G \subset F$.*

Exercice 2 *Soient F, G et H trois sous-espaces vectoriels d'un espace vectoriel E avec $F \subset G$. Montrer que*

$$F + (H \cap G) = (F + H) \cap G.$$

Exercice 3 *Soient F, G et H trois sous-espaces vectoriels d'un espace vectoriel E. A-t-on*

$$(F + H) \cap G = (F \cap G) + (H \cap G)?$$

Exercice 4 *Soit $f : E \to F$ une application linéaire. Montrer que l'application $G \to f(G)$ est une bijection de l'ensemble des sous-espaces vectoriels de E qui contiennent $\ker f$ sur l'ensemble des sous-espaces vectoriels de $\mathrm{Im} f$.*

Exercice 5 *On rappelle qu'un endomorphisme p de E est une* projection *si $p^2 := p \circ p = p$. Montrer que p est un projecteur si et seulement si $q := \mathrm{Id}_E - p$ en est un. Montrer qu'alors, $\ker p = \mathrm{Im} q$, $\mathrm{Im} p = \ker q$ et $E = \ker p \oplus \mathrm{Im} p$.*

Exercice 6 *Soient p et q deux projecteurs de E. Montrer que $\mathrm{Im} p \subset \mathrm{Im} q$ si et seulement si $q \circ p = p$.*

Exercice 7 *($1 + 1 \neq 0$) On rappelle qu'un endomorphisme s de E est une* symétrie *si $s^2 = \mathrm{Id}_E$. Montre que s est une symétrie si et seulement si $p = \frac{s+\mathrm{Id}_E}{2}$ est un projecteur. En déduire qu'alors*

$$E = \ker(s - Id_E) \oplus \ker(s + Id_E).$$

Exercice 8 *Montrer qu'une homothétie est une application linéaire distincte de l'identité qui laisse les droites invariantes.*

Exercice 9 *Montrer que les sous-espaces vectoriels de $\mathbf{R}^4$ engendrés par $u := (2, 3, -1, 0)$ et $v := (-3, 1, 0, 2)$, d'une part et $u' := (-5, 9, -2, 6)$ et $v' := (5, 2, -1, -2)$, d'autre part, sont identiques*

Exercice 10 *Soit E l'espace vectoriel des applications $[0, \pi] \to \mathbf{R}$. Montrer que*

$$\mathcal{S} := (1, sinx, sin2x, sin3x, \ldots)$$

est un système libre de E.

Exercice 11 *Montrer que l'application*

$$M_n(\mathbf{R}) \to M_n(\mathbf{R}), A \mapsto {}^tA$$

est une symétrie vectorielle.

Exercice 12 *On rappelle qu'une matrice* symétrique *(resp.* antisymétrique*) est une matrice A telle que*

$${}^tA = A \text{ (resp. } {}^tA = -A).$$

Leur ensemble se note $SM_n(\mathbf{R})$ (resp. $AM_n(\mathbf{R})$). Montrer que $SM_n(\mathbf{R})$ et $AM_n(\mathbf{R})$ sont des sous-espaces vectoriels de $M_n(\mathbf{R})$ et que

$$M_n(\mathbf{R}) = SM_n(\mathbf{R}) \oplus AM_n(\mathbf{R}).$$

Exercice 13 *On note C (resp. C_k) l'ensemble des carrés magiques 3-3 (resp. de trace k), c'est à dire, les $A \in M_3(\mathbf{R})$ telles que les sommes des éléments des lignes, des colonnes et des diagonales soient toutes égales (resp. égales à k).*

i) Montrer que C_k est non-vide et que si $M, N \in C_k$, alors $M - N \in C_0$.

ii) Montrer que

$$C_0 = SC_0 \oplus AC_0$$

avec

$$SC_0 = C_0 \cap SM_3(\mathbf{R}) \text{ et } AC_0 = C_0 \cap AM_3(\mathbf{R}).$$

iii) Déterminer une base de SC_0 puis une base de AC_0. En déduire une base de C_0 puis une base de C.

iv) Trouver un carré magique de trace 27 dont toutes les entrées sont distinctes.

Exercice 14 *Soit E l'espace des polynômes de degré au plus 3 sur $\mathbf{R}$. Soient f_1, f_2, f_3 et f_4 les formes linéaires qui envoient P sur $P(0), P(1), P'(0)$ et $P'(1)$ respectivement. Montrer que c'est une base de $\check{E}$. Déterminer la base de E dont c'est la base duale.*

Exercice 15 *Soit $A \in M_n(\mathbf{R})$ telle que $A^2+A+I = 0$. Montrer que A est inversible.*

Exercice 16 *Soit $N \in M_n(\mathbf{R})$ telle que $N^r = 0$. Montrer que $A := I + N$ est inversible.*

Exercice 17 *Calculer A^n pour*

$$A := \begin{bmatrix} 1 & -1 \\ 1 & -3 \end{bmatrix}.$$

(On pourra chercher une relation linéaire entre les premières puissances de A, puis chercher le reste dans la division de X^n par le polynôme correspondant)

Exercice 18 *Même question avec*

$$A := \begin{bmatrix} -15 & 10 & 8 \\ -8 & 6 & 4 \\ -24 & 15 & 13 \end{bmatrix}.$$

Exercice 19 *Résoudre le système*

$$\begin{cases} u_{n+1} &= & u_n - v_n \\ v_{n+1} &= & u_n + 3v_n \end{cases}$$

avec conditions initiales $u_0 = 1$ *et* $v_0 = 2$.

Exercice 20 *Soit* $A \in M_n(\mathbf{R})$ *telle que* $A^2 + A + I = 0$. *Calculer* $\det A$. *Même question avec* $A^2 - A + I = 0$.

Exercice 21 *Montrer que si* n *est impair, il n'existe pas de* $A \in M_n(\mathbf{R})$ *telle que* $A^2 + I = 0$. *Même question avec* $A^2 - \sqrt{2}A + I = 0$.

Exercice 22 *Montrer que si* $A \in M_n(\mathbf{C})$, *alors* $\det \bar{A} = \overline{\det A}$. *En déduire que si* $A, B \in M_n(\mathbf{R})$ *satisfont* $AB = BA$, *alors* $\det(A^2 + B^2) \geq 0$.

Exercice 23 *Montrer que si* n *est impair et* $A \in M_n(\mathbf{R})$ *antisymétrique, alors* $\det A = 0$.

Exercice 24 *Calculer pour tout* $n \geq 2$

$$\Delta_n := \begin{vmatrix} 0 & \cdots & \cdots & 0 & a_n \\ x & \ddots & & \vdots & a_{n-1} \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & x & a_1 \end{vmatrix}$$

puis

$$\Gamma_n := \begin{vmatrix} x & 0 & \cdots & 0 & a_n \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & x & a_1 \\ a_n & \cdots & \cdots & a_1 & x \end{vmatrix}.$$

Exercice 25 *Calculer*

$$\Delta_n := \begin{vmatrix} 1+x^2 & -x & 0 & \cdots & 0 \\ -x & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -x \\ 0 & \cdots & 0 & -x & 1+x^2 \end{vmatrix}$$

Exercice 26 *Montrer que si* $P \in \mathbf{R}[X]$ *est tel que* $\deg P < n$, *et*

$$A = \begin{bmatrix} P(x) & \cdots & P(x+n) \\ \vdots & & \vdots \\ P(x+n) & \cdots & P(x+n+m) \end{bmatrix},$$

alors $\det A = 0$. *Calculer* $\det B$ *avec*

$$B = \begin{bmatrix} 0 & \cdots & n^2 \\ \vdots & & \vdots \\ n^2 & \cdots & (n+m)^2 \end{bmatrix}.$$

Exercice 27 *Soient* $\lambda_1, \ldots, \lambda_n$ *et* $a \neq b$ *fixés. Calculer*

$$\Delta := \begin{vmatrix} \lambda_1 & b & \cdots & b \\ a & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ a & \cdots & a & \lambda_n \end{vmatrix}.$$

On pourra montrer que

$$\Delta(X) := \begin{vmatrix} \lambda_1 + X & b+X & \cdots & b+X \\ a+X & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & b+X \\ a+X & \cdots & a+X & \lambda_n + X \end{vmatrix}$$

est un polynôme de degré au plus 1, *calculer* $\Delta(-a)$ *et* $\Delta(-b)$ *et en déduire* $\Delta = \Delta(0)$.

Exercice 28 *On considère, pour* $a \in \mathbf{R}$, *l'application linéaire*

$$f_a : \mathbf{R}[X]_{\leq 3} \to \mathbf{R}[X]_{\leq 3}, P \mapsto P(X+a)$$

ainsi que sa matrice M_a *dans la base canonique. Montrer que* M_a *est inversible et calculer* M_a^n *pour* $n \in \mathbf{Z}$. *On pourra d'abord remarquer que pour* $a, b \in \mathbf{R}$, *on a* $f_{a+b} = f_a \circ f_b$.

Exercice 29 *Montrer que f est un projecteur si et seulement si f est diagonalisable et ses valeurs propres $\in \{0, 1\}$.*

Exercice 30 *Calculer le polynôme caractéristique de*

$$\begin{bmatrix} 1 & -2 & -2 \\ 0 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Cette matrice est-elle diagonalisable ?

Index

Références