Cours introductif d'algèbre linéaire

Jean-Philippe Monnier et Daniel Naie

Sommaire

3 septembre, 2020

1. Espaces vectoriels

De l'utilité des espaces vectoriels : la structure d'espace vectoriel (on dit aussi structure linéaire) est une structure mathématique fondamentale. Quand il s'agit de modéliser un problème concret en économie, chimie, physique,..., très souvent, on essaie de se ramener à un problème linéaire. Pourquoi ? Tout simplement, parce que dans ce cadre nous disposons de toute une panoplie d'outils permettant de résoudre assez facilement le problème (par exemple, l'algorithme de Gauss pour les systèmes d'équations linéaires).

1.1. Introduction

Un espace vectoriel est un pays où les habitants s'appellent des vecteurs et dans lequel certaines opérations sont permises entre ces vecteurs.

Vous avez certainement déjà rencontré des espaces vectoriels. Par exemple l'espace $\mathbb{R}^3$ composé des triplets $\boldsymbol{u} = (x, y, z)$. Quelles sont les opérations permises avec ces "vecteurs" ? On se souvient que l'on peut

- additionner deux vecteurs : si $\boldsymbol{u} = (x, y, z)$ et $\boldsymbol{v} = (x', y', z')$, alors $\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v} = (x + x', y + y', z + z')$, c'est-à-dire la somme s'obtient en additionnant terme à terme les coordonnées — c'est encore un élément de $\mathbb{R}^3$
- multiplier un vecteur u par un nombre réel λ en multipliant chaque coordonnée par λ : $\lambda\boldsymbol{u} = (\lambda x, \lambda y, \lambda z)$ qui est encore un élément de $\mathbb{R}^3$.

Les même propriétés sont satisfaites par les éléments de $\mathbb{R}^2$, les paires de nombres réels $\boldsymbol{u} = (x, y)$.

Contre-exemple. Prenons en revanche la partie (ou sous-ensemble) A de $\mathbb{R}^2$ défini par $A = \{(x, y) \mid x, y \geq 0\}$. On peut bien additionner deux éléments de A mais en revanche, si on multiplie un élément de A par un nombre strictement négatif, on sort de A. Cet ensemble ne sera pas un espace vectoriel.

Les vecteurs de $\mathbb{R}^2$ peuvent être dessinés dans le plan : le vecteur $\boldsymbol{u} = (1, 2)$ est la flèche issue de l'origine et de sommet le point $(1, 2)$. Voir la figure 1. La somme de deux vecteurs est obtenue par la règle du parallélogramme, c'est-à-dire $\boldsymbol{u}+\boldsymbol{v}$ est la diagonale du parallélogramme construit avec $\boldsymbol{u}$ et $\boldsymbol{v}$ comme côtés.

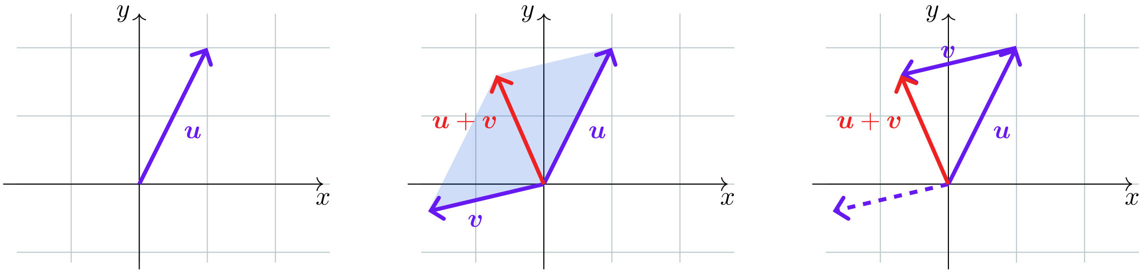

Figure 1: Le vecteur $\boldsymbol{u} = (1, 2)$. La somme des vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$

Remarque. La somme des vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ dans $\mathbb{R}^2$ (ou dans $\mathbb{R}^3$) s'interprète comme le vecteur obtenu *en pensant* le vecteur $\boldsymbol{v}$ comme issu du sommet de $\boldsymbol{u}$; voir le dernier dessin dans la figure 1.

Avec cette interprétation, nous pouvons retrouver les droites dans le plan en utilisant le langage des vecteurs.

Lemme 1.1. *Soient $\boldsymbol{u} \neq \boldsymbol{v} \in \mathbb{R}^2$. L'ensemble des sommets des vecteurs du sous-ensemble $\{(1-\lambda)\boldsymbol{u} + \lambda\boldsymbol{v} \mid \lambda \in \mathbb{R}\}$ est la droite qui passe par les sommets de $\boldsymbol{u}$ et $\boldsymbol{v}$.*

Démonstration. Nous avons $(1-\lambda)\boldsymbol{u}+\lambda\boldsymbol{v} = \boldsymbol{u}+\lambda(\boldsymbol{v}-\boldsymbol{u})$ et les sommets des vecteurs $\boldsymbol{u}+\lambda(\boldsymbol{v}-\boldsymbol{u})$ décrivent la droite passant par le sommet de $\boldsymbol{u}$ et de vecteur directeur $\boldsymbol{v} - \boldsymbol{u} \neq \boldsymbol{0}$. □

Un espace vectoriel sera un ensemble possédant les propriétés mentionnées ci-dessus pour $\mathbb{R}^2$ ou $\mathbb{R}^3$. Ces mêmes propriétés sont satisfaites par les éléments de $\mathbb{R}^n$, où n est un nombre entier ≥ 1.

Nous pouvons utiliser les vecteurs pour démontrer un théorème célèbre dû à Desargues. Ce résultat ne sera pas utilisé par la suite, mais il est introduit ici pour montrer l'utilité des méthodes vectoriels. La preuve pourrait être parcourue sans faire attention au détails lors d'une première lecture du cours.

Théorème 1.2 (Désargues)**.** *Soient $[ABC]$ et $[A'B'C']$ deux triangles tels que*
- *les droites (AA'), (BB') et (CC') s'intersectent en un point*
- *$(AB) \cap (A'B') = \{P\}$, $(BC) \cap (B'C') = \{M\}$, $(AC) \cap (A'C') = \{N\}$.*

Alors les points M, N et P sont alignés.

Démonstration. La figure 2 illustre la configuration du théorème. Nous interprétons le point O comme représentant le vecteur $\boldsymbol{0}$. Les hypothèses du théorème, exprimées en utilisant ces interprétations, deviennent :

- Il existe $\alpha \in \mathbb{R}$ tel que les vecteurs $\overrightarrow{OA}$ et $\overrightarrow{OA'}$ satisfont la relation $\overrightarrow{OA'} = \alpha\overrightarrow{OA}$. De même, $\overrightarrow{OB'} = \beta\overrightarrow{OB}$ et $\overrightarrow{OC'} = \gamma\overrightarrow{OC}$.
- Il existe $\lambda, \lambda' \in \mathbb{R}$ tels que $\overrightarrow{OP} = (1-\lambda)\overrightarrow{OA} + \lambda\overrightarrow{OB} = (1-\lambda')\overrightarrow{OA'} + \lambda'\overrightarrow{OB'}$. De même,

$$\overrightarrow{ON} = (1-\nu)\overrightarrow{OA} + \nu\overrightarrow{OC} = (1-\nu')\overrightarrow{OA'} + \nu'\overrightarrow{OC'}$$
$$\overrightarrow{OM} = (1-\mu)\overrightarrow{OB} + \mu\overrightarrow{OC} = (1-\mu')\overrightarrow{OB'} + \mu'\overrightarrow{OC'}.$$

En utilisant le vecteur $\overrightarrow{OP}$, nous obtenons

$$\overrightarrow{OP} = (1-\lambda)\overrightarrow{OA} + \lambda\overrightarrow{OB} = (1-\lambda')\overrightarrow{OA'} + \lambda'\overrightarrow{OB'} = (1-\lambda')\alpha\overrightarrow{OA} + \lambda'\beta\overrightarrow{OB}.$$

Donc

$$((1-\lambda) - (1-\lambda')\alpha)\overrightarrow{OA} + (\lambda - \lambda'\beta)\overrightarrow{OB} = \mathbf{0}.$$

Il s'ensuit (la diagonale du parallélogramme est nulle si et seulement si les côtés sont nuls) que

$$1-\lambda = (1-\lambda')\alpha \quad \text{et} \quad \lambda = \lambda'\beta.$$

En regardant ces deux relations comme un système en λ et λ', nous arrivons à

$$\lambda = \frac{\beta - \alpha\beta}{\beta - \alpha} \quad \text{et} \quad \lambda' = \frac{1-\alpha}{\beta-\alpha}.$$

Le même raisonnement pour $\overrightarrow{OM}$ et $\overrightarrow{ON}$ implique

$$\mu = \frac{\gamma - \beta\gamma}{\gamma - \beta} \quad \text{et} \quad \mu' = \frac{1-\beta}{\gamma-\beta}$$
$$\nu = \frac{\gamma - \alpha\gamma}{\gamma - \alpha} \quad \text{et} \quad \nu' = \frac{1-\alpha}{\gamma-\alpha}.$$

Pour finir la preuve, il faut comparer les vecteurs $\overrightarrow{MN}$ et $\overrightarrow{MP}$. Nous avons

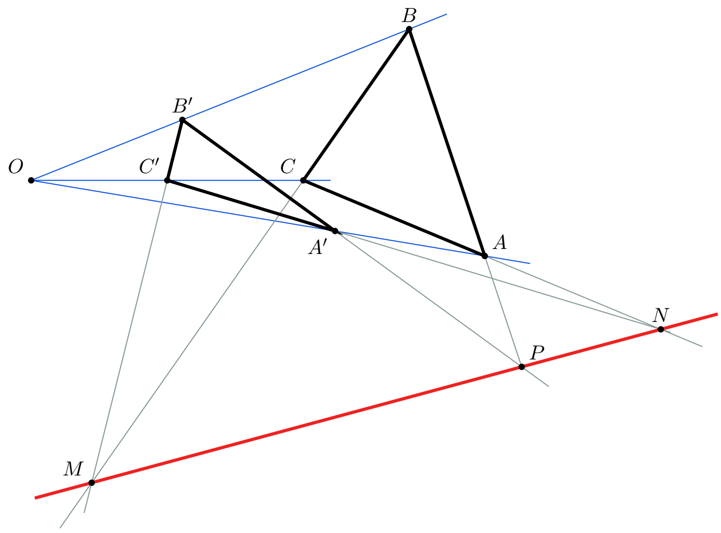

Figure 2: Le théorème de Desargues

$$\begin{aligned}
\overrightarrow{MN} &= \overrightarrow{ON} - \overrightarrow{OM} \\
&= (1-\nu)\overrightarrow{OA} + \nu\overrightarrow{OC} - (1-\mu)\overrightarrow{OB} - \mu\overrightarrow{OC} \\
&= \frac{\alpha\gamma - \alpha}{\gamma - \alpha}\overrightarrow{OA} - \frac{\beta\gamma - \beta}{\gamma - \beta}\overrightarrow{OB} + \frac{\gamma - \alpha\gamma}{\gamma - \alpha}\overrightarrow{OC} - \frac{\gamma - \beta\gamma}{\gamma - \beta}\overrightarrow{OC} \\
&= (\gamma - 1)\left(\frac{\alpha}{\gamma - \alpha}\overrightarrow{OA} - \frac{\beta}{\gamma - \beta}\overrightarrow{OB} + \frac{(\beta-\alpha)\gamma}{(\gamma-\alpha)(\gamma-\beta)}\overrightarrow{OC}\right),
\end{aligned}$$

car

$$\frac{\gamma - \alpha\gamma}{\gamma - \alpha} - \frac{\gamma - \beta\gamma}{\gamma - \beta} = \frac{(\beta-\alpha)\gamma(\gamma-1)}{(\gamma-\alpha)(\gamma-\beta)}.$$

En réorganisant les facteurs,

$$\overrightarrow{MN} = \frac{\gamma - 1}{(\gamma - \alpha)(\beta - \gamma)} \left(\alpha(\beta - \gamma)\,\overrightarrow{OA} + \beta(\gamma - \alpha)\,\overrightarrow{OB} + \gamma(\alpha - \beta)\,\overrightarrow{OC}\right).$$

De la même façon,

$$\overrightarrow{MP} = -\frac{\beta - 1}{(\alpha - \beta)(\beta - \gamma)} \left(\alpha(\beta - \gamma)\,\overrightarrow{OA} + \beta(\gamma - \alpha)\,\overrightarrow{OB} + \gamma(\alpha - \beta)\,\overrightarrow{OC}\right).$$

Donc les points M, N et P sont alignés. □

1.2. Définition d'un espace vectoriel

Pour définir un espace vectoriel, il faut d'abord disposer d'abord d'un ensemble de nombres appelés *scalaires* qui respectent un certain nombre de règles[1]. Cet ensemble sera noté $\mathbb{K}$. Pour nous, $\mathbb{K}$ sera $\mathbb{R}$ ou $\mathbb{C}$ (ou $\mathbb{Q}$).

Définition 1.3. Un $\mathbb{K}$-espace vectoriel est un ensemble E contenant un élément $\mathbf{0} = \mathbf{0}_E$ et muni[2] d'une addition et d'un produit avec les scalaires (de $\mathbb{K}$) satisfaisant les propriétés suivantes : pour tous $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in E$ et pour tous $\lambda, \mu \in \mathbb{K}$,
(i) $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$
(ii) $\boldsymbol{u} + \mathbf{0} = \mathbf{0} + \boldsymbol{u} = \boldsymbol{u}$
(iii) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$
(iv) $\alpha \cdot (\boldsymbol{u} + \boldsymbol{v}) = \alpha \cdot \boldsymbol{u} + \alpha \cdot \boldsymbol{v}$
(v) $(\alpha + \beta) \cdot \boldsymbol{u} = \alpha \cdot \boldsymbol{u} + \beta \cdot \boldsymbol{v}$
(vi) $(\alpha\beta) \cdot \boldsymbol{u} = \alpha \cdot (\beta \cdot u)$
(vii) $1 \cdot \boldsymbol{u} = \boldsymbol{u}$ et $0 \cdot \boldsymbol{u} = \mathbf{0}$.

Lorsqu'il n'y a pas de confusion sur le corps $\mathbb{K}$, nous parlerons simplement d'espace vectoriel plutôt que de $\mathbb{K}$-espace vectoriel. Par la suite, nous écrirons $\lambda\boldsymbol{u}$ pour le produit scalaire de $\boldsymbol{u}$ avec λ.

En général, dans un cours sur les espaces vectoriels, après la définition, il y a beaucoup d'agitation pour justifier que les règles (i) – (vii) font tout ce que nous voulons. Dans le lemme suivant on commence ces vérifications (le premier point dit que $\mathbf{0}$ *est unique* et le deuxième que tout vecteur de E admet un *élément opposé* pour l'addition). Voir aussi l'exercice 1.3 si la curiosité est insoutenable.

Lemme 1.4. *Soit E un espace vectoriel.*
1) *Si $\boldsymbol{u} + \mathbf{0}' = \boldsymbol{u}$ pour tout $\boldsymbol{u} \in E$, alors $\mathbf{0}' = \mathbf{0}$.*
2) *$\boldsymbol{u} + (-1)\boldsymbol{u} = \mathbf{0}$ pour tout $\boldsymbol{u} \in E$.*

Le point 2) justifie la notation $-\boldsymbol{u} = (-1)\boldsymbol{u}$.

[1] $\mathbb{K}$ est un *corps commutatif*

[2] Si la lectrice ou le lecteur sent le besoin d'un degré plus élevé de formalisme, elle peut remplacer la phrase par celle-ci : Un espace vectoriel sur $\mathbb{K}$ est un ensemble E contenant $\mathbf{0}$ et deux applications $A : E \times E \to E$ et $M : \mathbb{K} \times E \to E$ tels que, si on note $\boldsymbol{u} + \boldsymbol{v} = A(\boldsymbol{u}, \boldsymbol{v})$ et $\lambda \cdot \boldsymbol{u} = M(\lambda, \boldsymbol{u})$ pour tous $\boldsymbol{u}, \boldsymbol{v} \in E$ et $\lambda \in \mathbb{K}$, les propriétés suivantes sont satisfaites.

Démonstration. Pour 1), d'après (ii) appliqué à $\mathbf{0}'$, nous avons $\mathbf{0}+\mathbf{0}' = \mathbf{0}'$ et d'après l'hypothèse appliquée à $\mathbf{0}$, $\mathbf{0}+\mathbf{0}' = \mathbf{0}$. Donc
$$\mathbf{0}' = \mathbf{0} + \mathbf{0}' = \mathbf{0}.$$

Pour 2), nous appliquons, dans l'ordre, l'hypothèse, (iiv), (iv) et (iiv), et obtenons
$$\boldsymbol{u} + (-1)\boldsymbol{u} = 1\,\boldsymbol{u} + (-1)\boldsymbol{u} = (1 + (-1))\boldsymbol{u} = 0\boldsymbol{u} = \mathbf{0}.$$
□

Remarque. Il faut bien comprendre qu'un espace vectoriel n'est pas forcément un ensemble d'éléments représentés par des flèches mais avant tout un ensemble sur lequel existent une addition et une multiplication par scalaires. En conséquence, nous ne mettons pas systématiquement une flèche sur les éléments de E car les habitants d'un espace vectoriel peuvent être des fonctions, des matrices, des suites, des polynômes. . .

MÉTHODOLOGIE. Pour montrer qu'un ensemble donné est un espace vectoriel, il faut d'abord savoir quel est le corps de base ($\mathbb{R}$ ou $\mathbb{C}$ ou ...) puis identifier les lois $+$ et $\cdot$; ensuite il y a sept propriétés à vérifier. C'est long. En pratique, on montrera souvent plutôt que c'est un sous-espace vectoriel d'un espace vectoriel bien connu. On les verra bientôt.

1.3. Exemples classiques d'espaces vectoriels

Exemple 1.5. Les espaces $E = \mathbb{R}^n$, avec $n \geq 1$. Lorsque $n = 1$, on a le corps $\mathbb{R}$ lui-même. Pour $n = 2$, $\mathbb{R}^2$ et pour $n = 3$, $\mathbb{R}^3$ que l'on a déjà vu. On peut généraliser :
$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R}, i = 1, \ldots, n\}.$$
L'addition de deux vecteurs et la multiplication par un scalaire se font coordonnée par coordonnée. Explicitement,
$$\begin{aligned}(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)\\ \lambda(x_1, x_2, \ldots, x_n) &= (\lambda x_1, \lambda x_2, \ldots, \lambda x_n).\end{aligned}$$
L'élément nul est $\mathbf{0} = (0, 0, \ldots, 0)$.

Les espaces $\mathbb{R}^n$, $n \geq 1$ sont des $\mathbb{R}$-espaces vectoriels. On introduit de la même façon les espaces $\mathbb{C}^n$ et $\mathbb{Q}^n$. Il est utile de remarquer que $\mathbb{C}^n$ est aussi un $\mathbb{R}$-espace vectoriel —dû au fait que $\mathbb{R}$ est un sous-corps de $\mathbb{C}$, voir l'exemple suivant. Par contre, $\mathbb{R}$ n'est pas un $\mathbb{C}$-espace vectoriel.

Exemple 1.6. Le corps des nombres complexes, $\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}, i = \sqrt{-1}\}$ est un $\mathbb{R}$-espace vectoriel avec les opérations définies par
$$(a + ib) + (c + id) = (a + b) + i(c + d) \quad \text{et} \quad \lambda(a + ib) = \lambda a + i\lambda b.$$
La structure de corps de $\mathbb{C}$ est beaucoup plus riche par rapport à, et *prolonge* celle de $\mathbb{R}$-espace vectoriel. Elle est donnée par la multiplication de deux nombres complexes :
$$(a + ib)(c + id) = (ac - bd) + i(ad + bc).$$
Notez que la multiplication par le scalaire λ est la multiplication en tant que nombres complexes par $\lambda = \lambda + i\,0$.

Exemple 1.7. L'introduction de $\mathbb{R}^n$ se généralise encore plus : soit X un ensemble et $\mathcal{F}(X,\mathbb{R})$ l'ensemble des fonctions définies sur X à valeurs dans $\mathbb{R}$,

$$\mathcal{F}(X,\mathbb{R}) = \{f \mid f : X \to \mathbb{R}\}.$$

Tout comme pour $\mathbb{R}^n$, $\mathcal{F}(X,\mathbb{R})$ devient un $\mathbb{R}$-espace vectoriel en définissant l'addition de deux fonctions (vecteurs) et la multiplication par un scalaire composante par composante. Explicitement, pour l'addition, si $f, g \in \mathcal{F}(X,\mathbb{R})$, alors $f+g$ est la fonction définie pour tout $x \in X$ par

$$(f+g)(x) = f(x) + g(x).$$

Si $N_n = \{1, 2, \ldots, n\}$, alors $\mathcal{F}(N_n,\mathbb{R}) = \mathbb{R}^n$. Donc l'espace précédent est bien une généralisation de $\mathbb{R}^n$. La notation $\mathbb{R}^X$ est parfois utilisée pour désigner $\mathcal{F}(X,\mathbb{R})$ avec cette structure d'espace vectoriel sur $\mathbb{R}$. Qui est $\mathbb{R}^{\mathbb{N}}$? Quels objets mathématiques sont ses éléments ? Un autre cas particulier de l'exemple 1.7 sont les espaces des matrices. Par exemple, si pour $X = \{1,2\} \times \{1,2\}$, $\mathcal{F}(X,\mathbb{C})$ est l'espace $\mathcal{M}_{2\times 2}(\mathbb{C})$ ci-dessus.

Exemple 1.8. Soit $\mathcal{M}_{2\times 2}(\mathbb{C})$, ou $\mathcal{M}_{2,2}(\mathbb{C})$, l'ensemble des matrices carrées de taille 2, à coefficients complexes,

$$\mathcal{M}_{2\times 2}(\mathbb{C}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{C} \right\}.$$

$\mathcal{M}_{2\times 2}(\mathbb{C})$ est un $\mathbb{C}$-espace vectoriel ; l'élément nul est la matrice nulle $0 = 0_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ et les opérations sont définie par

$$M + M' = \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a+a' & b+b' \\ c+c' & d+d' \end{pmatrix}$$

et

$$\lambda M = \lambda \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}.$$

Plus généralement, on a :

Proposition 1.9. *L'ensemble de matrices de taille $m \times n$, $\mathcal{M}_{m\times n}(\mathbb{K})$, $m, n \geq 1$, est un $\mathbb{K}$-espace vectoriel avec la somme et le produit (avec des scalaires) définis composante par composante.*

1.4. Sous-espaces vectoriels

Définition 1.10. Un sous-ensemble F d'un $\mathbb{K}$-espace vectoriel E est un *sous-espace vectoriel* de E si les trois propriétés suivantes sont vérifiées :

(i) $\mathbf{0} \in F$
(ii) pour tous $\boldsymbol{u}, \boldsymbol{v} \in F$, $\boldsymbol{u} + \boldsymbol{v} \in F$
(iii) pour tout $\boldsymbol{u} \in F$ et pour tout $\lambda \in \mathbb{K}$, $\lambda \boldsymbol{u} \in F$.

La condition (i) est une manière convenable de dire que F est non vide.

Proposition 1.11. *Soit E un $\mathbb{K}$-espace vectoriel. Le sous-ensemble $F \subset E$ est un sous-espace vectoriel si et seulement si F, muni de l'addition des vecteurs et de la multiplication par scalaires induites par celles de E, est lui-même un $\mathbb{K}$-espace vectoriel.*

Démonstration. Les deux implications sont des vérifications immédiates basées sur les définitions 1.3 et 1.10. □

Remarque 1.12. En pratique, pour montrer qu'un ensemble F est un espace vectoriel, on cherche un sur-ensemble E de F qui soit un espace vectoriel connu (chercher dans la liste des exemples vus plus haut) et on montre ensuite que F est un sous-espace vectoriel de E. Ceci est en effet beaucoup plus facile —il n'y a que trois propriétés à vérifier —que de montrer directement que F est un espace vectoriel en vérifiant les sept propriétés de la définition 1.3.

Remarque 1.13. Les propriétés (ii) et (iii) de la définition d'un sous-espace vectoriel peuvent être combinées en une seule : pour tous $\boldsymbol{u}, \boldsymbol{v} \in F$ et pour tous $\lambda, \mu \in \mathbb{K}$ on a $\lambda\boldsymbol{u} + \mu\boldsymbol{v} \in F$. (Si on n'éprouve pas de problème lié à l'absence de symétrie, il suffit de vérifier que $\boldsymbol{u} + \mu\boldsymbol{v} \in F$.)

L'expression $\lambda\boldsymbol{u}+\mu\boldsymbol{v}$ est un exemple de *combinaison linéaire* de $\boldsymbol{u}$ et $\boldsymbol{v}$. Plus généralement, dans la définition suivante nous introduisons une notion centrale de l'algèbre linéaire.

Définition 1.14. Soit E un $\mathbb{K}$-espace vectoriel et soit $\mathcal{S} \subset E$ un sous-ensemble. On appelle *combinaison linéaire* (finie) de vecteurs de $\mathcal{S}$ toute expression

$$\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_n\boldsymbol{v}_n$$

avec $n \in \mathbb{N}^*$, $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n \in \mathcal{S}$ et $\alpha_1, \dots, \alpha_n \in \mathbb{K}$.

Par exemple, en utilisant cette notion, $F \subset E$ **est un sous-espace vectoriel si** F **est non-vide et toute combinaison linéaire de vecteurs de** F **appartient à** F.

Exemple. Les sous-ensembles $\{0\}$ et E de E sont toujours des sous-espaces vectoriels de E ; ils sont appelés sous-espaces vectoriels triviaux. Tout sous-espace différent de $\{0\}$ et de E est dit *sous-eapce propre.*

Exemple. Dans $E = \mathbb{R}^2$, la droite D d'équation $y = x$ est un sous-espace vectoriel. On a $D = \mathbb{R}(1,1) = \{(\lambda, \lambda) \mid \lambda \in \mathbb{R}\}$.

Plus généralement, toutes les droites vectorielles $\mathbb{R}\boldsymbol{u}$ avec $\boldsymbol{u} \neq \boldsymbol{0}$ sont des sous-espaces vectoriels. En revanche la réunion des deux droites d'équation $x = 0$ et $y = 0$, n'est pas un sous-espace vectoriel. Sauriez vous dire pourquoi ?

Exemple. Dans $\mathbb{R}^3$, les sous-espace vectoriel non triviaux sont les droites ainsi que les plans vectoriels. On y reviendra dans la section suivante. Pour l'instant, considérons le sous-ensemble

$$F = \{(x, y, z) \mid x + y = 0\} \subset \mathbb{R}^3.$$

Nous obtenons, car la condition définissant F ne porte pas sur z,

$$F = \{(x, y, z) \mid x + y = 0\} = \{(x, -x, z) \mid x, z \in \mathbb{R}\} = \{x(1, -1, 0) + z(0, 0, 1) \mid x, z \in \mathbb{R}\},$$

c'est-à-dire F est le sous-ensemble de toutes les combinaisons linéaires de $(1, -1, 0)$ et $(0, 0, 1)$. Donc F est un sous-espace vectoriel.

Parmi les deux ensembles suivants, quel est celui qui est un sous-espace vectoriel de $\mathbb{R}^3$?

$$P_1 = \{(x,y,z) \mid x+y+z=1\}, \qquad P_2 = \{(x,y,z) \mid x+y+z=0\}.$$

Exemple. En analyse, si I est un intervalle de $\mathbb{R}$, le sous-ensemble $\mathcal{C}(I,\mathbb{R}) \subset \mathcal{F}(I,\mathbb{R})$ des fonctions de I dans $\mathbb{R}$ continues sur I est un sous-espace vectoriel de l'espace des fonctions de I dans $\mathbb{R}$. De même l'ensemble $\mathcal{D}(I,\mathbb{R})$ des fonctions de I dans $\mathbb{R}$ dérivables est un sous-espace vectoriel de $\mathcal{C}(I,\mathbb{R})$. Sauriez-vous montrer cela ?

Exemple. L'ensemble F des suites réelles $u = (u_n)_{n\geq 0}$ vérifiant pour tout $n \geq 0$, $u_{n+2} = u_{n+1} + u_n$ est un sous-espace vectoriel de l'espace de toutes les suites réelles, $\mathbb{R}^{\mathbb{N}}$.

Proposition 1.15. *Toute intersection de sous-espaces vectoriels est un sous-espace vectoriel.*

Démonstration. Montrons d'abord que l'intersection de deux sous-espace vectoriel $F = F_1 \cap F_2$ est un sous-espace vectoriel. Il faut vérifier, d'après la remarque 1.13 que $\mathbf{0} \in F$ et que pour tous $\boldsymbol{u}, \boldsymbol{v} \in F$ et pour tous $\alpha, \beta \in \mathbb{K}$, on a $\alpha\boldsymbol{u} + \beta\boldsymbol{v} \in F$.

- $\mathbf{0} \in F_1$ et $\mathbf{0} \in F_2$ donc $\mathbf{0} \in F_1 \cap F_2$.
- Comme $\boldsymbol{u}, \boldsymbol{v} \in F$, nous avons $\boldsymbol{u}, \boldsymbol{v} \in F_1$ et $\boldsymbol{u}, \boldsymbol{v} \in F_2$. Mais F_1 et F_2 sont des sous-espaces vectoriels, donc $\alpha\boldsymbol{u} + \beta\boldsymbol{v} \in F_1$ et $\alpha\boldsymbol{u} + \beta\boldsymbol{v} \in F_2$, d'où $\alpha\boldsymbol{u} + \beta\boldsymbol{v} \in F$.

Dans le cas général, F s'écrit $F = \bigcap_{i\in I} F_i$ et la démonstration est similaire. □

Remarque. En général, la réunion de deux sous-espaces vectoriels *n'est pas* un sous-espace vectoriel. Par exemple la réunion de deux droites vectorielles quelconques dans $\mathbb{R}^2$ n'est pas un sous-espace vectoriel.

1.5. Exercices

Exercice 1.1. Les sous-ensembles de $\mathbb{R}^2$ formés des couples (x,y) de nombres réels vérifiant les conditions suivantes sont-ils des sous-espaces vectoriels de $\mathbb{R}^2$?

(i) $3x - 2y = 0$	(ii) $xy = 0$	(iii) $x^2 + y^2 = 1$
(iv) $x^2 + y^2 = 0$	(v) $y = 2x^2$	(vi) $x + \lvert y\rvert = 0$
(vii) $x^2 + 4y^2 = 4xy$	(viii) $x \geq 0$	(ix) $x^2 - 4y^2 = 0$.

Exercice 1.2. Dans l'espace vectoriel $\mathbb{R}^n$ on considère les vecteurs $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ et $\boldsymbol{w}'$ tels que $2\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{w}$ et $5\boldsymbol{u} + 3\boldsymbol{v} = \boldsymbol{w}'$. Déterminer $\boldsymbol{w}$ et $\boldsymbol{w}'$ en fonction de $\boldsymbol{u}$ et $\boldsymbol{v}$.

Exercice 1.3. Voici quelques propriétés élémentaires qui découlent de la définition d'un espace vectoriel. Soit E un $\mathbb{K}$-espace vectoriel. Pour tous $\lambda, \mu \in \mathbb{K}$ et tous $\boldsymbol{u}, \boldsymbol{v} \in E$ on a

1) $\lambda\boldsymbol{u} = 0$ si et seulement si $\lambda = 0$ ou $u = 0$
2) $\lambda(\boldsymbol{u} - \boldsymbol{v}) = \lambda\boldsymbol{u} - \lambda\boldsymbol{v}$
3) $(\lambda - \mu)\boldsymbol{u} = \lambda\boldsymbol{u} - \mu\boldsymbol{u}$.

Exercice 1.4. Dans chacun des cas ci-dessous, montrer que $V \subset \mathbb{R}^2$ est un

1) $V = \{(x,0) \mid x \in \mathbb{R}\}$
2) $V = \{(x,-x) \mid x \in \mathbb{R}\}$
3) $V = \{(3x - y, 2x + 5y) \mid x, y \in \mathbb{R}\}$.

Exercice 1.5. Soit m un paramètre réel. Montrer que les triplets réels de la forme (x, y, m) forment un sous-espace vectoriel de $\mathbb{R}^3$ si et seulement si $m = 0$.

Exercice 1.6. Soient a, b, d trois réels. à quelle condition les $(x, y, z) \in \mathbb{R}^3$ tels que

$$ax + by + z + d = 0$$

forment-ils un sous-espace vectoriel de $\mathbb{R}^3$? Donner une interprétation géométrique de cette condition.

Exercice 1.7. On considère le $\mathbb{R}$-espace vectoriel $E = \mathbb{R}$. Quels sont les sous espaces vectoriels de E ? Même question pour $\mathbb{C}$ vu comme $\mathbb{C}$-espace vectoriel. Quels sont les sous-espaces vectoriels de $\mathbb{C}$ vu comme espace vectoriel sur $\mathbb{R}$?

Exercice 1.8. Soit E un $\mathbb{R}$-espace vectoriel. Soient F et G deux sous espaces vectoriels de E. Démontrer que $F \cup G$ est un sous espace vectoriel de E si et seulement si $F \subset G$ ou $G \subset F$.

2. Familles génératrices et familles libres. Bases

Exemple introductif. On se place dans l'espace $\mathbb{R}^3$ et l'on considère les éléments $\boldsymbol{e}_1 = (1, 0, 0)$, $\boldsymbol{e}_2 = (0, 1, 0)$ et $\boldsymbol{e}_3 = (0, 0, 1)$. Soit $\boldsymbol{u} = (x_1, x_2, x_3)$ un vecteur quelconque de $\mathbb{R}^3$. On peut écrire

$$\begin{aligned}\boldsymbol{u} &= (x_1, 0, 0) + (0, x_2, 0) + (0, 0, x_3)\\ &= x_1(1, 0, 0) + x_2(0, 1, 0) + x_3(0, 0, 1)\\ &= x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3.\end{aligned}$$

On dira que $\boldsymbol{u}$ est *combinaison linéaire* de $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ et $\boldsymbol{e}_3$ (voir la définition 1.14). Comme tout vecteur de $\mathbb{R}^3$ s'exprime ainsi (car $\boldsymbol{u}$ est quelconque), on dira que $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ est une *famille génératrice* de $\mathbb{R}^3$. De plus, cette écriture étant unique, on dira que $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ est une *famille libre*. Étant famille libre et génératrice de $\mathbb{R}^3$, on dira que $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ est une *base* de $\mathbb{R}^3$.

L'objet de cette section est de définir ces notions dans un espace vectoriel quelconque E.

2.1. Familles génératrices

Définition 2.1. Soit E un espace vectoriel. On dit que les vecteurs $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p$ engendrent E, ou que la famille $\mathcal{G} = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p\}$ est *génératrice* pour E, si pour tout élément $\boldsymbol{u} \in E$ il existe $\lambda_1, \ldots, \lambda_p \in \mathbb{K}$ tels que

$$\boldsymbol{u} = \lambda_1\boldsymbol{w}_1 + \cdots + \lambda_p\boldsymbol{w}_p.$$

En utilisant la notion de combinaison linéaire introduite dans la définition 1.14, $\mathcal{G}$ est une famille génératrice si tout élément de E est une combinaison linéaire de vecteurs de $\mathcal{G}$. La famille $\mathcal{G}$ peut être infinie ; dans ce cas il faut préciser “tout élément de E est une *combinaison linéaire finie* de vecteurs de $\mathcal{G}$”.

Lemme-Définition 2.2. *Soit E un espace vectoriel et $\mathcal{S}$ une famille d'éléments de E. L'ensemble de toutes les combinaisons linéaires finies de vecteurs de $\mathcal{S}$ est un sous-espace vectoriel de E, appelé le sous-espace engendré par $\mathcal{S}$ et noté* $\mathrm{Vect}(\mathcal{S})$. *De plus,* $\mathrm{Vect}(\mathcal{S})$ *est le plus petit sous-espace vectoriel de E contenant $\mathcal{S}$.*

Démonstration. Nous vérification directement que $\mathrm{Vect}(\mathcal{S})$ est un sous-espace qui contient $\mathcal{S}$. Par exemple, si $\boldsymbol{w} \in \mathcal{S}$, alors la combinaison linéaire $1\,\boldsymbol{w} = \boldsymbol{w} \in \mathrm{Vect}(\mathcal{S})$.

Pour montrer que $\mathrm{Vect}(\mathcal{S})$ est le plus petit, il faut montrer que si F est un sous-espace contenant $\mathcal{S}$, alors $\mathrm{Vect}(\mathcal{S}) \subset F$. Soit $\boldsymbol{w} = \lambda_1\boldsymbol{w}_1 + \cdots + \lambda_p\boldsymbol{w}_p \in \mathrm{Vect}(\mathcal{S})$. Nous voulons voir que $\boldsymbol{w} \in F$. Comme $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p \in \mathcal{S} \subset F$, d'après la remarque 1.13 appliquée à $\boldsymbol{w}_1$ et $\boldsymbol{w}_2$,

$$\lambda_1\boldsymbol{w}_1 + \lambda_2\boldsymbol{w}_2 \in F.$$

Puis, en l'appliquant à $\lambda_1\boldsymbol{w}_1 + \lambda_2\boldsymbol{w}_2$ et $\boldsymbol{w}_3$,

$$(\lambda_1\boldsymbol{w}_1 + \lambda_2\boldsymbol{w}_2) + \lambda_3\boldsymbol{w}_3 = \lambda_1\boldsymbol{w}_1 + \lambda_2\boldsymbol{w}_2 + \lambda_3\boldsymbol{w}_3 \in F.$$

Nous continuons et après $p-1$ étapes, nous obtenons $\boldsymbol{w} \in F$. □

Exemple. Nous avons vu que $\mathbb{R}^3$ est engendré par $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Plus généralement, la famille $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ engendre $\mathbb{K}^n$, où $\boldsymbol{e}_k$ est le vecteur dont toutes les composantes sont nulles sauf la k-ième qui est égale à 1.

Exemple. Le $\mathbb{R}$-espace vectoriel $\mathbb{C}$ est engendré par 1 et i. Mais $\mathbb{C}$, vu comme $\mathbb{C}$-espace vectoriel, peut être engendré seulement par 1 (ou par $4\sqrt{3} - 19i$).

Exemple. Nous considérons $\mathbb{R}^3$ est le sous-espace vectoriel $P = \mathrm{Vect}(\boldsymbol{e}_1, \boldsymbol{e}_2)$. Nous remarquons que

$$P = \{(x, y, 0) \mid x, y \in \mathbb{R}\},$$

qu'il est défini par l'équation $z = 0$ et qu'il ressemble à l'espace vectoriel $\mathbb{R}^2$.

Exemple. Nous considérons $\mathcal{M}_2(\mathbb{R})$, l'espace vectoriel des matrices carrées de taille 2×2 à coefficients réels et les matrices

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Alors $\mathrm{Vect}(A, S) = \mathrm{Vect}(E_{12}, E_{21})$. Pour vérifier cette égalité il suffit de monter que $A, S \in \mathrm{Vect}(E_{12}, E_{21})$ et que $E_{12}, E_{21} \in \mathrm{Vect}(A, S)$.

Exemple 2.3 (plans vectoriels)**.** Dans $\mathbb{R}^3$ on considère le sous-ensemble $H \subset \mathbb{R}^3$ défini par $H = \{(x, y, z) \mid 2x - y - 3z = 0\}$, c'est-à-dire l'ensemble des solutions de l'*équation linéaire homogène*[3] $2x - y - 3z = 0$. Alors

$$\begin{aligned} H &= \{(x, 2x - 3z, z) \mid x, z \in \mathbb{R}\} \\ &= \{(x, 2x, 0) + (0, -3z, z) \mid x, z \in \mathbb{R}\} \\ &= \{x(1, 2, 0) + z(0, -3, 1) \mid x, z \in \mathbb{R}\} \\ &= \mathrm{Vect}\left((1, 2, 0), (0, -3, 1)\right), \end{aligned}$$

[3] Une équation homogène est une équation linéaire dont le terme constant est nul.

c'est-à-dire H est un sous-espace vectoriel de $\mathbb{R}^3$ engendré par deux vecteurs. On dit que H est un plan vectoriel —nous verrons plus loin que ces vecteurs sont linéairement indépendants.

De manière générale, une équation linéaire homogène non-triviale dans $\mathbb{K}^n$ définit un sous-espace vectoriel — l'ensemble des solutions —qui sera appelé hyperplan vectoriel.

Proposition 2.4. *Soit E un espace vectoriel.*

1) *Si $\mathcal{A}$ et $\mathcal{A}$ sont deux familles d'éléments de E telles que $\mathcal{A} \subset \mathcal{A}'$ alors* $\mathrm{Vect}(\mathcal{A}) \subset \mathrm{Vect}(\mathcal{A}')$.
2) *Soit $p \in \mathbb{N}$, $p \geq 1$, et soient $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p \in E$. Si $\lambda \in \mathbb{K}^*$, alors* $\mathrm{Vect}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p) = \mathrm{Vect}(\lambda \boldsymbol{v}_1, \ldots, \boldsymbol{v}_p)$.
3) *Si $p \geq 2$, alors*

$$\mathrm{Vect}\Big(\boldsymbol{v}_1 + \sum_{1 \leq k \leq p} \lambda_k \boldsymbol{v}_k, \ldots, \boldsymbol{v}_p\Big) = \mathrm{Vect}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p).$$

Démonstration. Le premier point découle du fait que $Vect(\mathcal{A})$ est le plus petit sous-espace contenant $\mathcal{A}$ et $\mathcal{A} \subset \mathcal{A}' \subset \mathrm{Vect}(\mathcal{A}')$. Les deux autres sont des vérifications immédiates. □

2.2. Familles libres

Définition 2.5. Soit E un espace vectoriel.

(i) Les vecteurs $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p \in E$ sont dits *liés* ou *linéairement dépendants* s'il existe des scalaires $\lambda_1, \ldots, \lambda_p$ non tous nuls, tels que

$$\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_p \boldsymbol{v}_p = \mathbf{0}.$$

(ii) Les vecteurs $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p \in E$ sont dit *libres* ou *linéairement indépendants* s'il ne sont pas liés.

Une famille $\mathcal{F} \subset E$ est dite libre ou liée si les vecteurs qui la compose sont libres ou liés respectivement. Plus correctement, la famille $\mathcal{F}$ est liée s'il existe une combinaison linéaire avec des vecteurs de $\mathcal{F}$ qui soit nulle. La famille est liée si elle n'est pas libre.

Remarque 2.6. Soit $\mathcal{A} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ une famille de vecteurs. Pour montrer que la famille $\mathcal{A}$ est libre (ou que les vecteurs sont linéairement indépendants), on se donne des scalaires $\lambda_1, \ldots, \lambda_n$ tels que $\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_n \boldsymbol{v}_n = \mathbf{0}$ et on cherche à montrer que $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$.

Pour montrer que la famille $\mathcal{A}$ est liée, on cherche une liaison du type $\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_n \boldsymbol{v}_n = \mathbf{0}$ avec $\lambda_1, \ldots, \lambda_n$ non tous nuls.

Quelques exemples qui illustre la remarque précédente :

1) Dans $\mathbb{R}^3$, pour $\boldsymbol{u}_1 = (1,1,0)$, $\boldsymbol{u}_2 = (-1,1,1)$, $\boldsymbol{u}_3 = (0,2,1)$, On a $\boldsymbol{u}_1 + \boldsymbol{u}_2 - \boldsymbol{u}_3 = \mathbf{0}$. Donc ces vecteurs sont liés.

2) Dans $\mathbb{R}^3$, si $\boldsymbol{v}_1 = (1,1,1)$, $\boldsymbol{v}_2 = (1,0,1)$ et $\boldsymbol{v}_3 = (1,1,0)$, la famille $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ est-elle libre ?

Soient $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tels que $\lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \lambda_3 \boldsymbol{v}_3 = \mathbf{0}$. Alors

$$(\lambda_1 + \lambda_2 + \lambda_3, \lambda_1 + \lambda_3, \lambda_1 + \lambda_2) = (0,0,0)$$

et donc les λ_i satisfont le système d'équations linéaires

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ \lambda_1 + \lambda_3 = 0 \\ \lambda_1 + \lambda_2 = 0 \end{cases}$$

que l'on résout par l'élimination (ou la méthode du pivot)

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 0 & (L_1) \\ \quad - \lambda_2 \qquad = 0 & (L_2 - L_1) \\ \qquad\quad - \lambda_3 = 0 & (L_3 - L_1) \end{cases}$$

donc $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

Lemme 2.7. *Soit E un espace vectoriel.*

1) *Toute famille contenant le vecteur nul est liée.*
2) *Une famille $\{\boldsymbol{v}\}$ (c'est-à-dire à un élément) est libre si et seulement si $\boldsymbol{v}$ est non nul.*
3) *Une famille de deux vecteurs $\{\boldsymbol{u}, \boldsymbol{v}\}$ est liée si et seulement si ils sont colinéaires (c'est-à-dire s'il existe $\lambda \in \mathbb{K}$ tel que, ou bien $\boldsymbol{u} = \lambda \boldsymbol{v}$ ou bien $\boldsymbol{v} = \lambda \boldsymbol{u}$).*

Démonstration. L'affirmation 1) s'ensuit du $1 \cdot \mathbf{0} + 0 \cdot \boldsymbol{v}_2 + \cdots + 0 \cdot \boldsymbol{v}_p = \mathbf{0}$.

Pour 2), si la famille est libre, alors d'après 1), $\boldsymbol{v} \neq \mathbf{0}$. Réciproquement, nous argumentons par contraposée Supposons qu'il existe $\lambda \in \mathbb{K}$, $\lambda \neq 0$ tel que $\lambda \boldsymbol{v} = \mathbf{0}$ (c'est-à-dire supposons que la famille est liée). Alors, d'après (vi) et (vii) de la définition 1.3 d'un espace vectoriel,

$$\mathbf{0} = \lambda^{-1}\mathbf{0} = \lambda^{-1}(\lambda \boldsymbol{v}) = (\lambda^{-1}\lambda)\boldsymbol{v} = 1\,\boldsymbol{v} = \boldsymbol{v}.$$

$a \neq 0$, forcément $\lambda = 0$ donc la famille est libre.

Pour 3), si $\boldsymbol{u}$ et $\boldsymbol{v}$ sont linéairement dépendants, alors il existe $\alpha, \beta \in \mathbb{K}$, non tous nuls, tels que

$$\alpha \boldsymbol{u} + \beta \boldsymbol{v} = \mathbf{0}.$$

En supposant $\alpha \neq 0$, nous avons $\boldsymbol{u} = (-\beta/\alpha)\boldsymbol{v}$. La réciproque est évidente. □

Attention. Si la famille $\{\boldsymbol{u}, \boldsymbol{v}\}$ est liée, alors ce n'est pas toujours vrai qu'il existe $\lambda \in \mathbb{K}$ tel que $\boldsymbol{u} = \lambda \boldsymbol{v}$; cela est faux si $\boldsymbol{u} \neq \mathbf{0}$ et $\boldsymbol{v} = \mathbf{0}$. Dans ce cas, c'est $\boldsymbol{v}$ qui s'exprime à l'aide de $\boldsymbol{u}$, $\boldsymbol{v} = 0\,\boldsymbol{u}$.

La proposition suivante est très pratique.

Proposition 2.8. *Soit E un $\mathbb{K}$-espace vectoriel. Les conditions suivantes sont équivalentes.*

1) *Les vecteurs $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ sont liés.*
2) *Il existe un indice $k \in \{1, \ldots, n\}$ tel que $\boldsymbol{v}_k$ soit combinaison linéaire des autres $\boldsymbol{v}_i$, c'est-à-dire il existe $\mu_1, \ldots, \mu_{k-1}, \mu_{k+1}, \ldots, \mu_n \in \mathbb{K}$ tel que*

$$\boldsymbol{v}_k = \mu_1 \boldsymbol{v}_1 + \cdots + \mu_{k-1}\boldsymbol{v}_{k-1} + \mu_{k+1}\boldsymbol{v}_{k+1} + \cdots + \mu_n \boldsymbol{v}_n.$$

Démonstration. 1) $\Rightarrow$ 2) : Si $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ est liée, il existe des scalaires non tous nuls $\lambda_1, \ldots, \lambda_n$ tels que $\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_n \boldsymbol{v}_n = \mathbf{0}$. L'un au moins des λ_i est non nul, par exemple λ_k. Alors

$$\lambda_k \boldsymbol{v}_k = -\sum_{\substack{i=1 \\ i \neq k}}^{n} \lambda_i \boldsymbol{v}_i,$$

et
$$\boldsymbol{v}_k = -\sum_{\substack{i=1\\i\neq k}}^{n} \frac{\lambda_i}{\lambda_k}\,\boldsymbol{v}_i = \sum_{\substack{i=1\\i\neq k}}^{n} \Big(-\frac{\lambda_i}{\lambda_k}\Big)\,\boldsymbol{v}_i = \sum_{\substack{i=1\\i\neq k}}^{n} \mu_i \boldsymbol{v}_i,$$
où $\mu_i = -\frac{\lambda_i}{\lambda_k}$.

Réciproquement, si
$$\boldsymbol{v}_k = \sum_{\substack{i=1\\i\neq k}}^{n} \mu_i \boldsymbol{v}_i,$$
alors $\mu_k \boldsymbol{v}_k - \sum_{\substack{i=1\\i\neq k}}^{n} \mu_i \boldsymbol{v}_i = \boldsymbol{0}$, avec $mu_k = 1$, donc la famille est liée. □

2.3. Bases

Définition. On dit qu'un espace vectoriel E est de *dimension finie* s'il admet une famille génératrice finie.

Définition 2.9. Soit E un espace vectoriel. Si les vecteurs $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ engendrent E et sont linéairement indépendants, on dit qu'ils forment une *base* de E, noté $(\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n)$.

Une base $\mathcal{B} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n)$ est un n-uplet d'éléments de E. La famille $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ est à la fois une famille libre et génératrice de E.

Attention. La notion de famille vue précédemment ne tient pas compte de l'ordre des éléments qui la composent ; en revanche, une base est ordonnée. La base (e_1, e_2) n'est pas la même base que (e_2, e_1). Si l'on tient compte maintenant de l'ordre, c'est que l'on va décrire un vecteur à l'aide de ses coordonnées dans la base considérée. Si l'on change l'ordre de la base, les coordonnées vont changer aussi.

Exemple. Les vecteurs $\boldsymbol{e}_1 = (1, 0, 0)$, $\boldsymbol{e}_2 = (0, 1, 0)$, $\boldsymbol{e}_3 = (0, 0, 1)$ forment une base de $\mathbb{R}^3$ appelée *base canonique*.

En effet, on a vu que tout vecteur de $\mathbb{R}^3$ s'écrit comme combinaison linéaire de $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ donc la famille est génératrice. De plus, la famille est libre car $\lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2 + \lambda_3\boldsymbol{e}_3 = \boldsymbol{0}$ implique $(\lambda_1, \lambda_2, \lambda_3) = (0, 0, 0)$. Donc $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ est bien une base de $\mathbb{R}^3$.

Exemple. Dans la figure 3 les coordonnées (x, y) correspondent à la base canonique $(\boldsymbol{e}_x, \boldsymbol{e}_y)$ et les coordonnées (ξ, η) correspondent à la base $(\boldsymbol{e}_\xi, \boldsymbol{e}_\eta)$ donnée par
$$\boldsymbol{e}_\xi = \frac{3}{2}\,\boldsymbol{e}_x - \frac{1}{2}\,\boldsymbol{e}_y \qquad \text{et} \qquad \boldsymbol{e}_\eta = \frac{1}{4}\,\boldsymbol{e}_x + \boldsymbol{e}_y.$$
Remarquer qu'on a décrit les vecteurs de la nouvelle base $(\boldsymbol{e}_\xi, \boldsymbol{e}_\eta)$ en utilisant les vecteurs de la base canonique —on a exprimé chaque vecteur de la nouvelle base dans la base canonique. Évidemment, on peut utliser la nouvelle base pour exprimer les vecteurs $\boldsymbol{e}_x$ et $\boldsymbol{e}_y$ dans celle-ci. Explicitement on a
$$\boldsymbol{e}_x = \frac{4}{13}\,(2\boldsymbol{e}_\xi + \boldsymbol{e}_\eta) = \frac{8}{13}\,\boldsymbol{e}_\xi + \frac{4}{13}\,\boldsymbol{e}_\eta \quad \text{et} \quad \boldsymbol{e}_y = \frac{4}{13}\,(-\frac{1}{2}\,\boldsymbol{e}_\xi + 3\boldsymbol{e}_\eta) = -\frac{2}{13}\,\boldsymbol{e}_\xi + \frac{12}{13}\,\boldsymbol{e}_\eta.$$

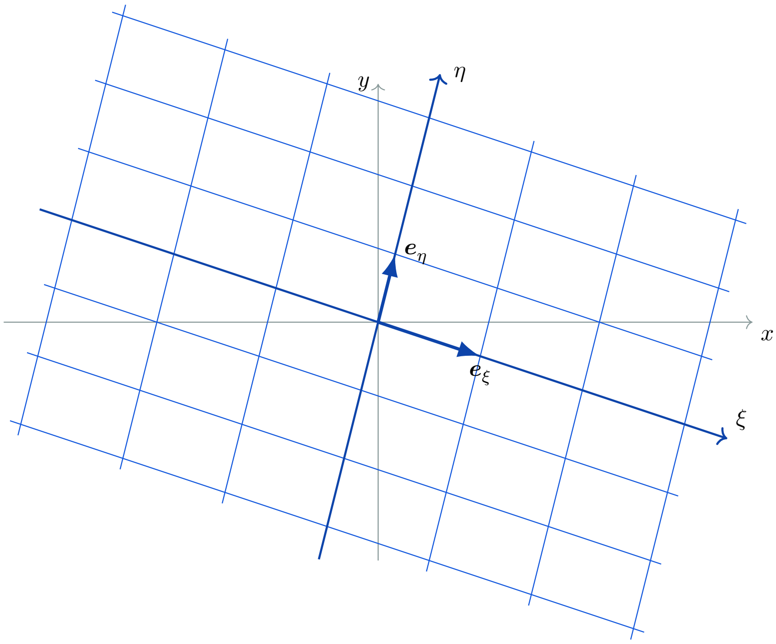

Figure 3: Les vecteurs $\boldsymbol{e}_\xi$ et $\boldsymbol{e}_\eta$ forment une base de $\mathbb{R}^2$. Tout vecteur de $\mathbb{R}^2$ a une écriture unique en utilisant les coordonnées (ξ, η) associées à cette nouvelle base.

Voici un résultat fondamental[4] qui justifie l'intérêt des bases dans la théorie des espaces vectoriels. Ce résultat généralise la remarque de l'exemple précédent.

Proposition 2.10. *Soit E un espace vectoriel de dimension finie. Les vecteurs $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ forment une base de E si et seulement si tout élément $\boldsymbol{v} \in E$ s'écrit de manière unique comme combinaison linéaire des $\boldsymbol{e}_i$, c'est-à-dire il existe $x_1, \dots, x_n \in \mathbb{K}$ uniques tels que*

$$\boldsymbol{v} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n.$$

Les scalaires $x_1, \dots, x_n$ s'appellent les coordonnées de $\boldsymbol{v}$ dans la base $\mathcal{B} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$; elles sont uniques. Pour x_i, la i-ième coordonnée de $\boldsymbol{v}$ dans $\mathcal{B}$, la notion correcte serait $x_i(\boldsymbol{v})$, pour indiquer que la coordonnée dépend du vecteur.

Démonstration. On suppose que $\mathcal{B} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ est une base et on considère un vecteur $\boldsymbol{v} \in E$. L'existence des scalaires $x_1, \dots, x_n$ tels que $\boldsymbol{v} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$ provient simplement du fait que $\mathcal{B}$ est génératrice de E. Pour montrer l'unicité, supposons l'existence de deux décompositions

$$\boldsymbol{v} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n = \mu_1\boldsymbol{e}_1 + \cdots + \mu_n\boldsymbol{e}_n.$$

On a alors

$$(x_1 - \mu_1)\boldsymbol{e}_1 + \cdots + (x_n - \mu_n)\boldsymbol{e}_n = \boldsymbol{0},$$

et puisque la famille est libre, on a

$$x_1 - \mu_1 = x_2 - \mu_2 = \ldots = x_n - \mu_n = 0.$$

[4]La proposition dit qu'une base décrit de manière unique tout vecteur de l'espace considéré. C'est cette description, par des coordonnées, qui nous permet d'effectuer des calculs dans un problème faisant appel aux espaces vectoriels. Dans la section suivante, nous verrons que toute base d'un espace vectoriel E a le même nombre d'éléments. Ce nombre sera nommé la *dimension* de E.

c'est-à-dire $x_1 = \mu_1$, $x_2 = \mu_2$, …, $x_n = \mu_n$. D'où l'unicité.

Pour la partie "seulement si", nous devons justifier que si tout vecteur de E est une combinaison linéaire unique des vecteurs $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$, alors les vecteurs engendrent E et sont libres. Pour voir qu'il sont libres, considérons $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tels que $\sum_i \lambda_i \boldsymbol{e}_i = \mathbf{0}$. Comme

$$\sum_i \lambda_i \boldsymbol{e}_i = \mathbf{0} = \sum_i 0\, \boldsymbol{e}_i$$

et la combinaison linéaire pour $\mathbf{0}$ est unique, $\lambda_i = 0$ pour tout i. □

Remarque 2.11 (Notation)**.** Si E est un espace vectoriel de dimension finie et $\mathcal{B}$ est une base, alors on notera $C_{\mathcal{B}}(\boldsymbol{v})$ la matrice colonne des coordonnées de $\boldsymbol{v}$ dans la base $\mathcal{B}$,

$$C_{\mathcal{B}}(\boldsymbol{v}) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Cette matrice colonne, qu'on appellera aussi vecteur colonne, peut être vue comme un élément de $\mathbb{K}^n$; **fixer une base permet d'identifier l'espace vectoriel E à $\mathbb{K}^n$**.

Apparait ici un problème trivial mais potentiellement perturbant. Des branches différentes des mathématiques se sont développées indépendamment et ont leurs propres notations pour les éléments de $\mathbb{R}^n$ ou $\mathbb{C}^n$. En algèbre linéaire, quand on étudie les bases ou les systèmes d'équations linéaires par exemple, il est naturel d'utiliser l'écriture en colonne des vecteurs ; mais en analyse ou en géométrie analytique, souvent, on préfère l'écriture en ligne. Il est sage de respecter les habitudes de chaque domaine.

Pour nous, **les coordonnées par rapport à une base fixée seront toujours pensées, sinon écrites, comme formant un vecteur colonne, ou matrice colonne**. Par exemple, dans $\mathbb{R}^3$, nous utiliserons les notations

$$(1, 2, 3) \quad \text{et} \quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

pour désigner un même vecteur, $\boldsymbol{v}$. La première écriture et celle *canonique* pour un élément de $\mathbb{R}^3$, l'ensemble des triplets réels. La deuxième sous-entend l'existence d'une base fixée, ici la base canonique, par rapport à laquelle 1, 2 et 3 sont les coordonnées de $\boldsymbol{v}$.

Si nous considérons la base $\mathcal{B} = ((1, 1, 0), (0, 1, 0), (0, 0, 1))$, alors les coordonnées de $\boldsymbol{v} = (1, 2, 3)$ par rapport à $\mathcal{B}$ sont

$$C_{\mathcal{B}}(\boldsymbol{v}) = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

Exemple. Dans $\mathbb{R}^3$ on considère les vecteurs $\boldsymbol{v}_1 = (1, 1, 0)$, $\boldsymbol{v}_2 = (0, 1, 1)$, $\boldsymbol{v}_3 = (1, 0, 1)$ et $\boldsymbol{v}_4 = (1, 1, 1)$. On remarque que ces quatre vecteurs engendrent $\mathbb{R}^3$ car la base canonique vérifie

$$\boldsymbol{e}_1 = \boldsymbol{v}_4 - \boldsymbol{v}_2, \quad \boldsymbol{e}_2 = \boldsymbol{v}_4 - \boldsymbol{v}_3 \quad \text{et} \quad \boldsymbol{e}_3 = \boldsymbol{v}_4 - \boldsymbol{v}_1,$$

donc $\boldsymbol{e}_i \in \operatorname{Vect}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_4)$ et par conséquent

$$\mathbb{R}^3 = \operatorname{Vect}(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3) \subset \operatorname{Vect}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_4).$$

Les $\boldsymbol{v}_i$ ne sont pas linéairement indépendants car

$$\boldsymbol{v}_3 = \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3 = -\boldsymbol{v}_1 - \boldsymbol{v}_2 - \boldsymbol{v}_3 + 3\boldsymbol{v}_4,$$

d'où

$$\boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3 - 2\boldsymbol{v}_4 = \boldsymbol{0}.$$

Les vecteurs $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ et $\boldsymbol{v}_3$ forment une base. Pouvez-vous le démontrer ?

Exemple 2.12. L'espace-temps en cinématique s'identifie, pour un observateur $\mathfrak{R}$ à $\mathbb{R}^4$ avec la base $(\boldsymbol{e}_t, \boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$. Le sous-espace vectoriel $\mathrm{Vect}(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ est l'espace physique de l'observateur au moment $t = 0$. Les événements sont des vecteurs (ou points) dans $\mathbb{R}^4$. Le mouvement rectiligne et uniforme d'un point matériel qui passe par l'origine au moment $t = 0$ est décrit pour $\mathfrak{R}$ par une droite vectorielle engendrée par $(1, v_x, v_y, v_z)$, où $\overrightarrow{v} = (v_x, v_y, v_z) \in \mathbb{R}^3$ est le vecteur vitesse qui caractérise le mouvement. Si $\mathfrak{G}$ est un autre observateur qui se déplace de manière rectiligne et uniforme par rapport à $\mathfrak{R}$ avec la vitesse $\overrightarrow{u} \in \mathbb{R}^3$, comment s'expriment les vecteurs $(\boldsymbol{e}_\tau, \boldsymbol{e}_\xi, \boldsymbol{e}_\eta, \boldsymbol{e}_\zeta)$ à travers lesquels il décrit l'espace-temps aux vecteurs de la base de $\mathfrak{R}$? Pour le moment, pour pouvoir donner une réponse, nous allons considérer l'espace physique comme étant de dimension 1, décrit par $\boldsymbol{e}_z$ pour $\mathfrak{R}$ et par $\boldsymbol{e}_\zeta$ pour $\mathfrak{G}$. En particulier, le vecteur vitesse $\overrightarrow{u}$ devient un nombre réel u ! On a

$$\begin{aligned} \boldsymbol{e}_\tau &= \boldsymbol{e}_t + u\boldsymbol{e}_z \\ \boldsymbol{e}_\zeta &= \boldsymbol{e}_z. \end{aligned}$$

Pour justifier ces formules il faut énoncer les hypothèses qui sous-tendent cette situation physique :

1) $\mathfrak{R}$ et $\mathfrak{G}$ se dépalcent avec vitesse u un par rapport à l'autre
2) $\mathfrak{R}$ et $\mathfrak{G}$ se rencontrent en l'événement O de $\mathbb{R}^2$ qu'ils définissent tous les deux comme ayant lieu à $t = \tau = 0$ —ils utilisent cet événement pour synchroniser leurs horloges
3) le temps est le même pour $\mathfrak{R}$ et $\mathfrak{G}$, c'est-à-dire $t = \tau$
4) les axes de l'esapce des deux observateurs pointent dans la même direction

On a $\boldsymbol{e}_\tau = a\boldsymbol{e}_t + b\boldsymbol{e}_z$ et $\boldsymbol{e}_\zeta = c\boldsymbol{e}_t + d\boldsymbol{e}_z$ et il faut déterminer les coefficients a, b, c et d. En passant en coordonnées, on obtient, pour un vecteur quelconque,

$$t\boldsymbol{e}_t + z\boldsymbol{e}_z = \boldsymbol{v} = \tau\boldsymbol{e}_\tau + \zeta\boldsymbol{e}_\zeta = \tau(a\boldsymbol{e}_t + b\boldsymbol{e}_z) + \zeta(c\boldsymbol{e}_t + d\boldsymbol{e}_z) = (a\tau + c\zeta)\boldsymbol{e}_t + (b\tau + d\zeta)\boldsymbol{e}_z.$$

Donc

$$\begin{aligned} t &= a\tau + c\zeta \\ z &= b\tau + d\zeta. \end{aligned}$$

D'après 3), $a = 1$ et $c = 0$. D'après 1), 2) et 4), $b = u$ et $d = 1$.

Proposition 2.13. *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie qui n'est pas trivial.*

1) *Si $\boldsymbol{u}$ engendre E, alors $(\boldsymbol{u})$ est une base de E.*
2) *Si les vecteurs $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ engendrent E mais ne forment pas une base de E, alors il existe un vecteur parmi eux tel que les $n-1$ autres vecteurs engendrent E.*
3) *Toute famille génératrice finie de E contient une base de E. (Parfois on dit qu'à partir de toute famille génératrice de E on peut extraire une base.)*

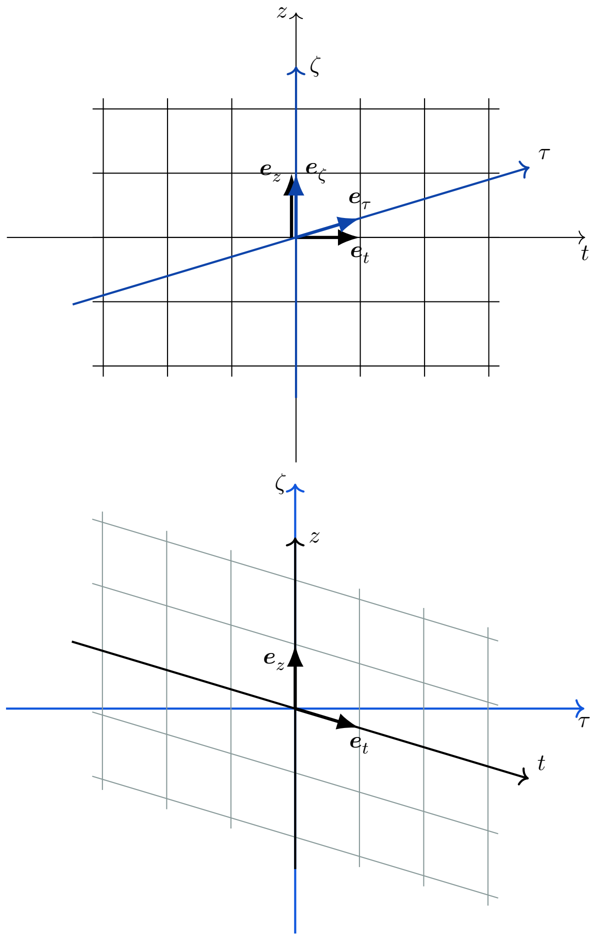

Figure 4: À gauche l'esapce-temps vu par $\mathfrak{K}$, c'est-à-dire sa base est la base canonique ; à droite l'espace-temps vu par $\mathfrak{G}$. La grille de gauche est transformée par le changement de base, c'est-à-dire de point de vue. Elle est représentée à droite dans la base de $\mathfrak{G}$.

Démonstration. Le premier point découle du Lemme 2.7, 2).

Pour démontrer le point 2), on suppose que $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ engendrent E mais qu'ils ne forment pas une base. Alors les vecteurs sont linéairement dépendants, et d'après la proposition 2.8, on peut supposer, en renumérotant les vecteurs, qu'il existe $\alpha_1, \ldots, \alpha_{n-1}$ tels que

$$\boldsymbol{u}_n = \alpha_1 \boldsymbol{u}_1 + \cdots + \alpha_{n-1} \boldsymbol{u}_{n-1}.$$

Nous affirmons que les vecteurs $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-1}$ engendrent E. Soit $\boldsymbol{v} \in E$. D'après l'hypothèse, il existe $\lambda_1, \ldots, \lambda_n$ tels que $\boldsymbol{v} = \sum_{i=1}^n \lambda_i \boldsymbol{u}_i$. En utilisant la relation précédente pour $\boldsymbol{u}_n$, on a

$$\boldsymbol{v} = \sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i = \sum_{i=1}^{n-1} \lambda_i \boldsymbol{u}_i + \sum_{k=1}^{n-1} \alpha_k \boldsymbol{u}_k = \sum_{i=1}^{n-1} (\lambda_i + \alpha_i) \boldsymbol{u}_i.$$

Pour 3), on applique le point précédent à plusieurs reprises. □

Corollaire 2.14. *Tout espace vectoriel de dimension finie admet au moins une base.*

Démonstration. Soit E un espace vectoriel de dimension finie. Par définition, il possède une famille génératrice $\mathcal{G}$. Si E n'est pas trivial, alors $\mathcal{G}$ n'est pas vide et on applique la proposition 2.13. Si $E = \{\mathbf{0}\}$, par convention la famille vide est une base. □

Remarque. Pour $E = \{\mathbf{0}\}$, il y a une unique base qui est la famille vide mais cela est conventionnel et n'a pas beaucoup d'intérêt.

2.4. Exercices

Exercice 2.1. Soient $\boldsymbol{u} = (1, 2, 3, 4)$ et $\boldsymbol{v} = (1, -2, 3, -4)$.
1) Existe-t-il $x, y \in \mathbb{R}$ tels que $(x, 1, y, 1) \in \mathrm{Vect}(\boldsymbol{u}, \boldsymbol{v})$?
2) Existe-t-il $x, y \in \mathbb{R}$ tels que $(x, 1, 1, y) \in \mathrm{Vect}(\boldsymbol{u}, \boldsymbol{v})$?

Exercice 2.2. On considère l'espace vectoriel $\mathbb{R}^4$.
1) Déterminer si les vecteurs $\boldsymbol{u} = (1, 3, -1, 4)$, $\boldsymbol{v} = (2, 1, 1, -2)$ et $\boldsymbol{w} = (4, -3, 5, -14)$ sont linéairement indépendants.
2) Déterminer $a, b, c \in \mathbb{R}$ (s'ils existent) tels que

$$a(1, 3, 0, 1) + b(2, -1, 1, 0) + c(3, 1, -1, 1) = (1, 4, -5, 2).$$

Exercice 2.3. Pour chacune des familles suivantes de vecteurs de $\mathbb{R}^3$, dire si elle est
- une famille libre de $\mathbb{R}^3$
- une famille génératrice de $\mathbb{R}^3$
- une base de $\mathbb{R}^3$.

Donner, pour chaque famille, une base de l'espace vectoriel engendré.
1) $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ avec $\boldsymbol{v}_1 = (-1, 1, 1)$, $\boldsymbol{v}_2 = (1, -1, 1)$, $\boldsymbol{v}_3 = (1, 1, -1)$.
2) $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ avec $\boldsymbol{v}_1 = (1, 1, 1)$, $\boldsymbol{v}_2 = (1, 2, 3)$, $\boldsymbol{v}_3 = (3, 2, 1)$.
3) $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ avec $\boldsymbol{v}_1 = (1, 0, 1)$, $\boldsymbol{v}_2 = (3, 0, 3)$.
4) $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$ avec $\boldsymbol{v}_1 = (1, 2, 3)$, $\boldsymbol{v}_2 = (2, 3, 1)$, $\boldsymbol{v}_3 = (3, 2, 1)$, $\boldsymbol{v}_4 = (1, 1, 1)$.

Exercice 2.4. Pour quels $m \in \mathbb{R}$,
1) la famille $\{(1, 3), (1 - m, 9)\}$ est-elle une base de $\mathbb{R}^2$?
2) la famille $\{(1 + m, 1 - m), (1 - m, 1 + m)\}$ est-elle une base de $\mathbb{R}^2$?
3) la famille $\{(1, 0, -2), (1, 3, -1), (3, 3, m^2 + 6m)\}$ est-elle une base de $\mathbb{R}^3$?

Exercice 2.5. Soit $\mathcal{B} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ une base de E et soit $\boldsymbol{v} \in E$ dont les coordonnées dans la base $\mathcal{B}$ sont $a_1, \dots, a_n$. Montrer que les vecteurs

$$\boldsymbol{u}_1 + \boldsymbol{v},\ \boldsymbol{u}_2 + \boldsymbol{v},\ \dots,\ \boldsymbol{u}_n + \boldsymbol{v}$$

sont libres si et seulement si $1 + a_1 + \cdots + a_n \neq 0$.

Exercice 2.6. Soit $U \subset \mathbb{R}^3$ défini par $x + y + z = 0$.
1) Montrer que U est un sous-espace vectoriel de $\mathbb{R}^3$.
2) Trouver une base de U et démontrer que c'est une base.
3) Est-ce que toute personne répondant à cette question, produira-t-elle la même base ? Que peut-on conclure concernant une notion de “base naturelle” pour un sous-espace vectoriel ?

Exercice 2.7. On considère le sous-ensemble

$$S \subset \mathcal{M}_2(\mathbb{R}), \quad S = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \,\middle|\, \exists q \in \mathbb{R} \text{ tel que } a + b = c + d = a + c = b + d = k \right\}.$$

1) Montrer que S est une sous-espace vectoriel de $\mathcal{M}_2(\mathbb{R})$.
2) Trouver une base de S.

Exercice 2.8. Soit $(\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n)$ une base de $\mathbb{K}^n$. Existe-t-il des entiers $n \geq 2$ tels que les vecteurs ci-dessous forment des bases de $\mathbb{K}^n$? Si oui, trouver tous les $n \geq 2$ satisfaisant cette propriété.

1) $\boldsymbol{u}_1 - \boldsymbol{u}_2$, $\boldsymbol{u}_2 - \boldsymbol{u}_3, \dots$, $\boldsymbol{u}_{n-1} - \boldsymbol{u}_n$, $\boldsymbol{u}_n - \boldsymbol{u}_1$
2) $\boldsymbol{u}_1 + \boldsymbol{u}_2$, $\boldsymbol{u}_2 + \boldsymbol{u}_3, \dots$, $\boldsymbol{u}_{n-1} + \boldsymbol{u}_n$, $\boldsymbol{u}_n + \boldsymbol{u}_1$

Exercice 2.9. Soit H le sous-espace vectoriel de $\mathbb{R}^3$ engendré par les vecteurs $\boldsymbol{u} = (1, 2, 3)$ et $\boldsymbol{v} = (3, 2, 1)$.

1) Vérifier que $\boldsymbol{u}$ et $\boldsymbol{v}$ ne sont pas colinéaires, en déduire la dimension de H.
2) Montrer que H est aussi engendré par $\boldsymbol{u}' = (1, 1, 1)$ et $\boldsymbol{v}' = (0, 1, 2)$.
3) Montrer que $\mathcal{B} = (\boldsymbol{u}, \boldsymbol{v})$ et $\mathcal{B}' = (\boldsymbol{u}', \boldsymbol{v}')$ sont deux bases de H.
4) Soit $\boldsymbol{w} \in H$ de coordonnées (x, y) dans $\mathcal{B}$ et (x', y') dans $\mathcal{B}'$. Calculer (x, y) en fonction de (x', y'), puis (x', y') en fonction de (x, y).

3. Dimension

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. Une base de E permet de décrire aisément tout vecteur grâce aux coordonnées. Dans la proposition 2.13 on a vu comment fabriquer une base à partir d'une famille génératrice de E —on enlève des vecteurs en conservant le caractère générateur de la famille restante jusqu'à temps qu'elle soit libre.

Nous verrons ici comment fabriquer une base à partir d'une famille libre —on rajoutera des vecteurs en conservant le caractère libre jusqu'à temps que la nouvelle famille soit génératrice de E. Chemin faisant, on établira le résultat principale : Toute base de E a le même nombre d'éléments, c'est-à-dire le même cardinal. Ce nombre est appelé la *dimension de E*.

Lors d'un premier contact avec les espaces vectoriels, la première partie de cette section est difficile ayant un degré d'abstraction élevé. Le lecteur peut retenir les résultats principaux (voir le Théorème 3.2 et ne pas s'attarder sur les détails lors d'une première lecture ; plus tard, après avoir acquis une certaine aisance en algèbre linéaire, il pourrait y revenir.

3.1. Lemme de Steinitz et dimension d'un espace vectoriel

Nous commençons par un résultat concernant les familles libres, similaire à la proposition 2.13.

Lemme 3.1. *Soit E un espace vectoriel et soit $\mathcal{F} = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_p\}$ une famille libre. Alors, ou bien les vecteurs de $\mathcal{F}$ forment une base de E, ou bien il existe $\boldsymbol{w} \in E$ tel que la famille $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_p, \boldsymbol{w}\}$ soit libre.*

Démonstration. Le sous-espace engendré par $\mathcal{F}$, $\mathrm{Vect}(\mathcal{F})$ est, ou bien E tout entier, ou bien un sous-espace propre. Dans le premier cas la famille est une base. Dans le deuxième, il existe $\boldsymbol{w} \in E \setminus \mathrm{Vect}(\mathcal{F})$, $\boldsymbol{w} \neq \boldsymbol{0}$. Montrons que $\boldsymbol{u}_1, \dots, \boldsymbol{u}_p, \boldsymbol{w}$ sont linéairement indépendants.

Soient $\lambda_1, \dots, \lambda_p, \mu \in \mathbb{K}$ tels que

$$\lambda_1 \boldsymbol{u}_1 + \cdot + \lambda_p \boldsymbol{u}_p + \mu \boldsymbol{w} = \boldsymbol{0}.$$

Si $\mu = 0$, alors les $\lambda_i = 0$ car $\mathcal{F}$ est libre. Si $\mu \neq 0$, nous obtenons la contradiction

$$\mathrm{Vect}(\mathcal{F}) \not\ni \boldsymbol{w} = -\frac{\lambda_1}{\mu} \boldsymbol{u}_1 - \cdots - \frac{\lambda_p}{\mu} \boldsymbol{u}_p$$

□

Nous arrivons au résultat technique principal, appelé le lemme d'échange de Steinitz. Ces conséquences importantes sont énumérées dans le théorème 3.2.

Lemme d'échange de Steinitz. *Soient* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$ *vecteurs linéairement indépendants. Soit* $0 \leq k < m$ *et soient* $\boldsymbol{f}_{k+1}$, $\boldsymbol{f}_{k+2}, \dots, \boldsymbol{f}_p$, $k+1 \leq p$, *tels que les vecteurs*

$$\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{f}_{k+1}, \boldsymbol{f}_{k+2}, \dots, \boldsymbol{f}_p$$

engendrent E*. Alors, après renumérotation éventuelle des indices des* $\boldsymbol{f}_j$*, les vecteurs*

$$\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{u}_{k+1}, \boldsymbol{f}_{k+2}, \dots, \boldsymbol{f}_p$$

engendrent E*.*

Démonstration. Comme $\mathrm{Vect}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{f}_{k+1}, \boldsymbol{f}_{k+2}, \dots, \boldsymbol{f}_p) = E$, il existe $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ tels que

$$\boldsymbol{u}_{k=1} = \sum_{j=1}^{k} \lambda_j \boldsymbol{u}_j + \sum_{j=k+1}^{p} \lambda_j \boldsymbol{f}_j.$$

Nous remarquons qu'il existe au moins un scalaire λ_j non nul avec $j \geq k+1$ car les $\boldsymbol{u}_i$ sont linéairement indépendants. En renumérotant les indices, nous pouvons supposer $\lambda_{k+1} \neq 0$. Alors

$$\boldsymbol{f}_{k+1} = \sum_{j=1}^{k} \Big(-\frac{\lambda_j}{\lambda_{k+1}}\Big) \boldsymbol{u}_j + \frac{1}{\lambda_{k+1}} \boldsymbol{u}_{k+1} + \sum_{j=k+2}^{p} \Big(-\frac{\lambda_j}{\lambda_{k+1}}\Big) \boldsymbol{f}_j$$

Pour finir, montrons que les vecteurs $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{k+1}, \boldsymbol{f}_{k+2}, \dots, \boldsymbol{f}_p$ engendrent E. Soit $\boldsymbol{v} \in E$. Par hypothèse, $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{f}_{k+1}, \dots, \boldsymbol{f}_p$ engendrent E. Donc il existe des scalaires α_j tels que

$$\boldsymbol{v} = \sum_{j=1}^{k} \alpha_j \boldsymbol{u}_j + \sum_{j=k+1}^{p} \alpha_j \boldsymbol{f}_j.$$

En utilisant l'expression précédente pour $\boldsymbol{f}_{k+1}$,

$$\begin{aligned}
\boldsymbol{v} &= \sum_{j=1}^{k} \alpha_j \boldsymbol{u}_j + \alpha_{k+1} \boldsymbol{f}_{k+1} + \sum_{j=k+2}^{p} \alpha_j \boldsymbol{f}_j \\
&= \sum_{j=1}^{k} \alpha_j \boldsymbol{u}_j + \alpha_{k+1} \left(\sum_{j=1}^{k} \Big(-\frac{\lambda_j}{\lambda_{k+1}}\Big) \boldsymbol{u}_j + \frac{1}{\lambda_{k+1}} \boldsymbol{u}_{k+1} + \sum_{j=k+2}^{p} \Big(-\frac{\lambda_j}{\lambda_{k+1}}\Big) \boldsymbol{f}_j \right) + \sum_{j=k+2}^{p} \alpha_j \boldsymbol{f}_j \\
&= \sum_{j=1}^{k} \Big(\alpha_j - \frac{\alpha_{k+1}\lambda_j}{\lambda_{k+1}}\Big) \boldsymbol{u}_j + \frac{\alpha_{k+1}}{\lambda_{k+1}} \boldsymbol{u}_{k+1} + \sum_{j=k+2}^{p} \Big(\alpha_j - \frac{\alpha_{k+1}\lambda_j}{\lambda_{k+1}}\Big) \boldsymbol{f}_j.
\end{aligned}$$

□

Une des conséquence du lemme est le fait que toutes les bases de E, espace de dimension finie, ont le même nombre d'élément. Ce nombre est appelé la *dimension* de E.

Théorème 3.2. *Soit* E *un* $\mathbb{K}$*-espace vectoriel de dimension finie.*

(a) *Si* $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_m\}$ *est une famille libre et* $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_p\}$ *est une famille génératrice, alors* $m \leq p$*.*

(b) *Toutes les bases de E ont le même nombre d'éléments appelé la dimension de E et noté* $\dim(E)$ *(ou* $\dim_{\mathbb{K}}(E)$ *pour préciser le corps des scalaires).*
(c) *Toute famille génératrice ayant* $\dim(E)$ *éléments est une base.*
(d) *Toute famille libre ayant* $\dim(E)$ *éléments est une base.*

Démonstration. Pour (a), nous supposons par l'absurde que $p < m$. Alors, en appliquant le lemme de Steinitz p fois, nous obtenons que $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ engendrent E. Donc

$$\boldsymbol{u}_{p+1} \in \operatorname{Vect}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p),$$

ce qui contredit la liberté de la famille $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$.

Pour (b), si $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$ et $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ sont deux bases, alors $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$ est libre et $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ est génératrice. D'après (A), $m \leq n$. En échangeant les rôles, $n \leq m$, donc $m = n$.

Le (c) découle du point précédent et de la proposition 2.13 —de toute famille génératrice on peut extraire une base.

Le (d) découle du (b) et du lemme 3.1 — toute famille libre peut être complétée à une base. □

Exemple 3.3. $\dim(\mathbb{K}^n) = n$, $\dim(\mathcal{M}_2(\mathbb{R}) = 4$, $\dim_{\mathbb{C}}(\mathbb{C}) = 1$, $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.

3.2. Dimensions des sous-espaces vectoriels

Définition. La dimension d'un sous-espace vectoriel F d'un espace vectoriel E est sa dimension en tant qu'espace vectoriel.

Proposition 3.4. *Pour tout sous-espace vectoriel $F \subset E$, on a* $\dim(F) \leq \dim(E)$ *avec égalité si et seulement si $F = E$.*

Démonstration. On suppose $F \neq \{\mathbf{0}\}$. Soit $n = \dim(E) \geq 1$ et soit $\boldsymbol{u}_1 \in F$, $\boldsymbol{u}_1 \neq \mathbf{0}$. La famille $\{\boldsymbol{u}_1\}$ est libre. Ou bien elle est une base de F, ou bien, en utilisant le lemme 3.1 (appliqué à F), elle peut être augmentée à une famille libre ayant un élément de plus. Cette opération peut être itérée au plus $n-1$ fois, car une famille libre dans F est libre dans E et, d'après Théorème 3.2 (D), une telle famille a au plus n éléments. □

Corollaire 3.5. *Si F, F' sont deux sous-espaces vectoriels tels que $F \subset F'$, alors* $\dim(F) \leq \dim(F')$ *avec égalité si et seulement si $F = F'$.*

Démonstration. D'après la proposition 3.4. □

La dimension est un très bon moyen de classifier les sous-espaces vectoriels d'un espace vectoriel. Par exemple, les sous-espaces vectoriels de $\mathbb{R}^3$, classifiés suivant leurs dimensions, sont :

- en dimension 0, $\{\mathbf{0}\}$
- en dimension 1, les droites vectorielles, $D = \operatorname{Vect}\{\boldsymbol{u}\}$ avec $\boldsymbol{u} \neq \mathbf{0}$
- en dimension 2, les plans vectoriels, $\Pi = \operatorname{Vect}\{\boldsymbol{u}, \boldsymbol{v}\}$ avec $\boldsymbol{u}$ et $\boldsymbol{v}$ linéairement indépendants
- en dimension 3, $\mathbb{R}^3$.

Un début de compréhension de la géométrie des sous-espaces vectoriels de E est donné par la formule suivante, due à Grassmann

Proposition 3.6 (dimension de la somme)**.** *Soient E un espace vectoriel de dimension finie et soient F et G deux sous-espaces vectoriels de E. Si on note par $F+G$ le sous-espace* $\mathrm{Vect}(F,G)$*, alors*

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

Démonstration. Nous savons que les sous-espaces F, G, $F+G$, $F \cap G$ sont de dimensions finies. Soit $\mathcal{B} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ une base de $F \cap G$. On complète $\mathcal{B}$ en $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_s)$ base de F et en $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{w}_1, \dots, \boldsymbol{w}_t)$ base de G. Montrons que

$$\mathcal{F} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_s, \boldsymbol{w}_1, \dots, \boldsymbol{w}_t)$$

est une base de $F+G$. La famille $\mathcal{F}$ est clairement génératrice de $F+G$. Il reste à vérifier qu'elle est libre. Si

$$\lambda_1 \boldsymbol{u}_1 + \cdots + \lambda_k \boldsymbol{u}_k + \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_s \boldsymbol{v}_s + \beta_1 \boldsymbol{w}_1 + \cdots + \beta_t \boldsymbol{w}_t = \boldsymbol{0},$$

alors le vecteur

$$\boldsymbol{u} = \lambda_1 \boldsymbol{u}_1 + \cdots + \lambda_k \boldsymbol{u}_k + \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_s \boldsymbol{v}_s = -(\beta_1 \boldsymbol{w}_1 + \cdots + \beta_t \boldsymbol{w}_t)$$

appartient à $F \cap G$. La première égalité implique $\alpha_1 = \cdots = \alpha_s = 0$ car $\mathcal{B}$ est une base de $F \cap G$. Donc

$$\lambda_1 \boldsymbol{u}_1 + \cdots + \lambda_k \boldsymbol{u}_k + \beta_1 \boldsymbol{w}_1 + \cdots + \beta_t \boldsymbol{w}_t = \boldsymbol{0}$$

et, comme $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{w}_1, \dots, \boldsymbol{w}_t)$ est une base de G,

$$\lambda_1 = \cdots = \lambda_k = \beta_1 = \cdots = \beta_t = 0$$

On obtient $\dim(F+G) = k+s+t = (k+s)+(k+t)-k = \dim(F)+\dim(G)-\dim(F\cap G)$. □

Définition 3.7. Un hyperplan d'un espace vectoriel de dimension n est un sous-espace vectoriel de dimension $n-1$.

Nous avons vu dans l'exemple 2.3 qu'une équation linéaire homogène non triviale (à coefficients réels) définit un hyperplan (c'est-à-dire plan) vectoriel dans $\mathbb{R}^3$ ou, plus généralement, dans $\mathbb{R}^n$. De plus, nous savons qu'une intersection de sous-espaces vectoriels est un sous-espace vectoriel, voir la proposition 1.15. Il s'ensuit que l'ensemble des solutions d'un système d'équations linéaires homogène est une intersection (finie) d'hyperplans vectoriels, donc un sous-espace vectoriel.

Question. Peut-on déterminer la dimension du sous-espace des solutions d'un système d'équations linéaires homogène ? La suite du cours fournira les éléments nécessaires pour répondre à cette question.

Exemple. $F = \{(x,y,z,t) \mid x+y+z+t = 0\} \subset \mathbb{R}^4$ et $F' = \{(x,y,z,t) \mid x-2y+z = 0\} \subset \mathbb{R}^4$ sont deux hyperplans de $\mathbb{R}^4$. Líntersection $F \cap G$ représente l'espace des solutions du système homogène

$$\begin{cases} x + y + z + t = 0 \\ x - 2y + z \phantom{{}+t} = 0. \end{cases}$$

Nous exprimons z et t en fonction de x et y ; la deuxième équation donne $z = -x + 2y$ et la première devient $t = -3y$. Donc

$$F \cap G = \{(x, y, -x + 2y, -3y) \mid x, y \in \mathbb{R}\}.$$

Puisque $(x, y, -x + 2y, -3y) = x(1, 0, -1, 0) + y(0, 1, 2, -3)$, $F \cap G$ est le sous-espace deux dimensionnelle, $\mathrm{Vect}((1, 0, -1, 0), (0, 1, 2, -3))$. Quelle est la dimension de $F + G$?

3.3. Application : rang d'une famille de vecteurs

Soit E un espace vectoriel sur $\mathbb{K}$.

Définition 3.8. Soit $\mathcal{F} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ une famille d'éléments d'un espace vectoriel E. On appelle rang de $\mathcal{F}$, noté $\mathrm{rg}(\mathcal{F})$, la dimension de l'espace vectoriel engendré par $\mathcal{F}$, c'est-à-dire $\mathrm{rg}(\mathcal{F}) = \dim(\mathrm{Vect}(\mathcal{F}))$.

Proposition 3.9. *Si $\mathcal{F} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$, alors $\mathrm{rg}(\mathcal{F}) \leq n$ avec égalité si et seulement si $\mathcal{F}$ est une famille libre.*

Démonstration. La première partie résulte de la proposition 2.13 car $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ est une famille génératrice de $\mathrm{Vect}(\mathcal{F})$.

Si $\mathrm{rg}(\mathcal{F}) = n$, alors $\dim(\mathrm{Vect}(\mathcal{F})) = n$ et la famille génératrice $\mathcal{F}$ de n vecteurs de cet espace est alors une base donc forcément libre. Réciproquement, si $\mathcal{F}$ est libre, c'est une base (puisque libre et génératrice) de $\mathrm{Vect}(\mathcal{F})$, donc $\dim(\mathrm{Vect}(\mathcal{F})) = n$. □

On retiendra que le rang d'une famille libre est le cardinal de cette famille. Pour $\{\boldsymbol{u}\}$, le rang vaut 0 si $\boldsymbol{u} = \boldsymbol{0}$ et 1 sinon. Pour $\{\boldsymbol{u}, \boldsymbol{v}\}$, le rang vaut 0 si $\boldsymbol{u} = \boldsymbol{v} = \boldsymbol{0}$, 1 si la famille est liée (avec un vecteur non nul au moins), et 2 si la famille est libre.

Exemple. Quel est le rang de la famille $\mathcal{F}$ de vecteurs de $\mathbb{R}^3$ composée de $\boldsymbol{v}_1 = (1, 2, 3)$, $\boldsymbol{v}_2 = (3, 1, 2)$, $\boldsymbol{v}_3 = (1, 2, 3)$ et $\boldsymbol{v}_4 = (1, 1, 1)$? Au plus 4 à cause du cardinal de $\mathcal{F}$, mais au plus 3 à cause de la dimension de $\mathbb{R}^3$. Pour déterminer $\mathrm{rg}(\mathcal{F}) \leq 3$ il faut faire une étude de dimension. Cet algorithme sera présenté par la suite.

Lemme 3.10. *Soit $\mathcal{F} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ une famille de vecteurs. Les quatre opérations suivantes ne changent pas le sous-espace vectoriel engendré par $\mathcal{F}$, et donc ne modifie pas le rang de la famille.*

(I) *Multiplier un vecteur $\boldsymbol{v}_i$ par $\lambda \neq 0$.*
(II) *Remplacer un vecteur $\boldsymbol{v}_i$ par $\boldsymbol{v}_i+$ (combinaison linéaire des autres vecteurs).*
(III) *Intervertir deux vecteurs de la famille.*
(IV) *Éliminer les vecteurs nuls éventuels de la famille.*

Démonstration. Ce lemme est une reformulation de la proposition 2.8 3). □

Les opérations (I-IV) ci-dessus s'appellent les *opérations élémentaires* pour les familles de vecteurs. En les itérant à partir de la famille $\mathcal{F} = (vv_1, \ldots, \boldsymbol{v}_n)$, on peut arriver à une famille libre $(\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_r)$ avec $r \leq n$ et $\mathrm{Vect}(\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_r) = \mathrm{Vect}(\mathcal{F})$. Par conséquent $\mathrm{rg}(\mathcal{F}) = r$.

En pratique, on utilise la représentation matricielle de la famille $\mathcal{F}$, chaque colonne de la matrice représentant les coordonnées d'un des vecteurs dans une base fixée.

On rappelle la notation introduite dans Remarque-Notation 2.11 : si $\mathcal{B} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ est une base de E, et $\boldsymbol{v} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \ldots + a_n\boldsymbol{e}_n$, alors la matrice $C_{\mathcal{B}}(\boldsymbol{v})$ est la matrice à n lignes et 1 colonne

$$C_{\mathcal{B}}(\boldsymbol{v}) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Elle s'appelle le vecteur colonne de $\boldsymbol{v}$ dans la base $\mathcal{B}$ —le vecteur colonne de $\boldsymbol{v}$ contient en colonne les coordonnées de $\boldsymbol{v}$ dans la base $\mathcal{B}$.

Définition 3.11. Soit $\mathcal{B} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ une base de E et soit $\mathcal{F} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ une famille telle que

$$\begin{aligned} \boldsymbol{v}_1 &= a_{1\,1}\boldsymbol{e}_1 + a_{2\,1}\boldsymbol{e}_2 + \cdots + a_{n\,1}\boldsymbol{e}_n \\ \boldsymbol{v}_2 &= a_{1\,2}\boldsymbol{e}_1 + a_{2\,2}\boldsymbol{e}_2 + \cdots + a_{n\,2}\boldsymbol{e}_n \\ &\vdots \\ \boldsymbol{v}_p &= a_{1\,p}\boldsymbol{e}_1 + a_{2\,p}\boldsymbol{e}_2 + \cdots + a_{n\,p}\boldsymbol{e}_n. \end{aligned}$$

La matrice à n lignes et p colonnes (de taille $n \times p$)

$$A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le p}}$$

s'appelle la matrice de décomposition de la famille $\mathcal{F}$ dans la base $\mathcal{B}$.

Description de l'algorithme pour déterminer le rang d'une famille de vecteurs.. La matrice A contient toutes les informations concernant les vecteurs de la famille $\mathcal{F}$; la i-ième colonne de A contient les coordonnées du vecteur $\boldsymbol{v}_i$ dans la base $\mathcal{B}$, c'est-à-dire le vecteur colonne de $\boldsymbol{v}_i$ dans la base $\mathcal{B}$.

Les opérations élémentaires (i-iii) sur les vecteurs d'une famille se traduisent par des opérations sur les colonnes de la matrice associée à la famille :

(i) multiplier une colonne par un scalaire $\lambda \neq 0$

(ii) ajouter à une colonne une autre colonne multipliée par un scalaire

(iii) permuter des colonnes.

En effectuant ces opérations selon la technique du *pivot de Gauss*, on va arriver à une matrice échelonnée verticalement c'est-à-dire dont les colonnes contiennent des zéros mais de façon ordonnée : *les zéros sont en haut de la colonne et leur nombre croît strictement de gauche à droite.* Il faut remarquer que la taille de la matrice finale est identique à la taille de la matrice initiale.

$$\begin{pmatrix} * & 0 & 0 & \cdots & 0 \\ * & * & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ * & * & & \cdots & \end{pmatrix}$$

Le rang de la famille sera donné par le nombre de colonnes non nulles de la matrice ainsi obtenue.

Proposition 3.12. *Soit $\mathcal{F} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ une famille et A la matrice de décomposition de $\mathcal{F}$ dans la base $\mathcal{B}$. Si, en itérant les opérations* (i-iii) *décrites au-dessus on arrive à une matrice échelonnée verticalement A', la matrice de décomposition dans la base $\mathcal{B}$ de la famille $\mathcal{F}' = \{\boldsymbol{v}'_1, \dots, \boldsymbol{v}'_r, \boldsymbol{0}, \dots, \boldsymbol{0}\}$, alors* $\mathrm{rg}(\mathcal{F}) = r$ *et la famille $\mathcal{F}'$ est une base de* $\mathrm{Vect}(\mathcal{F})$.

Démonstration. La famille $\{\boldsymbol{v}'_1, \dots, \boldsymbol{v}'_r\}$ est libre. Son rang est donc r. $\square$

Exemple. Calculons le rang de la famille $\mathcal{F} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ où $\boldsymbol{v}_1 = (1,1,1)$, $\boldsymbol{v}_2 = (1,2,3)$ et $\boldsymbol{v}_3 = (1, \frac{3}{2}, 2)$. (On rappelle que, par exemple, $\boldsymbol{v}_2 = (1,2,3) = \boldsymbol{e}_1 + 2\boldsymbol{e}_2 + 3\boldsymbol{e}_3$ dans la base canonique de $\mathbb{R}^3$.) *A priori*, le rang de $\mathcal{F}$ est au plus 3 (qui est le cardinal de la famille et la dimension de l'espace).

Dans la base canonique, on obtient par la méthode du pivot sur les colonnes en commençant avec la matrice de décomposition dans la base canonique,

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & \frac{3}{2} \\ 1 & 3 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & \frac{1}{2} \\ 1 & 2 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \end{pmatrix}.$$

On en déduit que le rang de la famille $\mathcal{F}$ est 2 et qu'une base du sous-espace vectoriel engendré est $(\boldsymbol{v}_1, \boldsymbol{w})$ où $\boldsymbol{w} = (0,1,2)$. Pouvez-vous donner une base de $\mathrm{Vect}(\mathcal{F})$ extraite de la famille génératrice $\mathcal{F}$?

Exemple. Calculer le rang de la famille $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$ où $\boldsymbol{v}_1 = (1,2,3)$, $\boldsymbol{v}_2 = (3,2,1)$, $\boldsymbol{v}_3 = (2,3,1)$ et $\boldsymbol{v}_4 = (1,1,1)$. *A priori* le rang est au plus 3 (dimension de l'espace).

En partant avec la matrice de décomposition dans la base canonique, nous avons, par la méthode du pivot de Gauss sur les colonnes,

$$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 2 & 2 & 3 & 1 \\ 3 & 1 & 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -4 & -1 & -1 \\ 3 & -8 & -5 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -4 & 0 & 0 \\ 3 & -8 & 3 & 0 \end{pmatrix}.$$

On en déduit que le rang est 3 et qu'une base du sous-espace vectoriel engendré est $(\boldsymbol{v}_1, \boldsymbol{w}_1 = (0,4,-8), \boldsymbol{w}_2 = (0,0,3))$ ou encore $(\boldsymbol{v}_1, (0,1,2), (0,0,1))$.

3.4. Exercices

Exercice 3.1. Soient $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{w} \in \mathbb{R}^5$. On considère les sous-espaces vectoriels V et W engendrés par $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ et $\boldsymbol{v}_3$, et respectivement par $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$ et $\boldsymbol{w}$. Montrer que, ou bien $\dim(W) = \dim(V)$, ou bien $\dim(W) = \dim(V) + 1$.

Exercice 3.2. Les affirmations suivantes sont-elles vraies ou fausses ? Justifier les réponses.

1) Toute famille génératrice d'un espace vectoriel contient une base de cet espace vectoriel.
2) La dimension d'un espace est le nombre de vecteurs de cet espace.
3) Toute famille contenant une famille liée est liée.
4) La base de $\mathbb{R}^3$ est $(\boldsymbol{e}_1 = (1,0,0), \boldsymbol{e}_2 = (0,1,0), \boldsymbol{e}_3 = (0,0,1))$.
5) Si $E = \mathrm{Vect}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$ et si $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ est libre, alors $\dim(E) = 3$.
6) $\mathrm{Vect}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_p) = \mathrm{Vect}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{p-1})$ si et seulement si $\boldsymbol{u}_p$ est combinaison linéaire de $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_{p-1}$.

7) Soient $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ trois vecteurs d'un espace vectoriel E. On suppose que $\{\boldsymbol{u}, \boldsymbol{v}\}$, $\{\boldsymbol{u}, \boldsymbol{w}\}$ et $\{\boldsymbol{v}, \boldsymbol{w}\}$ sont libres, alors la famille $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ est libre.
8) Soient p vecteurs $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ d'un espace vectoriel E. On suppose qu'aucun de ces vecteurs n'est combinaison linéaire des autres. Alors la famille $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\}$ est libre. (On pourrait commencer par traiter le cas $p = 3$.)

Exercice 3.3. On considère l'espace vectoriel $\mathbb{R}^3$ et $F = \{(x, y, z) \in E \mid 2x - y + z = 0\}$.
1) Montrer que F est un sous-espace vectoriel de E.
2) Trouver une base pour F.
3) Compléter cette base à une base de $\mathbb{R}^3$ en utilisant les vecteurs de la base canonique.

Exercice 3.4. Trouver une base et donner la dimension pour chacun des sous-espaces vectoriels de $\mathbb{R}^4$ suivants.
1) $\{(a + b, a - 3b, a, b) \mid a, b \in \mathbb{R}\}$
2) $\{(a - b, b + c, a, b + 2c) \mid a, b, c \in \mathbb{R}\}$
3) $\{(a, b, c, d) \mid a + b = c + d\}$.

Exercice 3.5. Dans $\mathbb{R}^6$, soient $\boldsymbol{v}_1 = (1, 1, 0, 2, 3, 0)$, $\boldsymbol{v}_2 = (1, 1, 1, 2, 3, 0)$ et $\boldsymbol{v}_3 = (1, 1, 0, 2, 3, 1)$. Vérifier que la famille $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ est libre, puis compléter cette famille en une base de $\mathbb{R}^6$ à l'aide des vecteurs de la base canonique.

Exercice 3.6. Soient V et W deux sous-espaces vectoriels distincts de $\mathbb{R}^3$ de dimension 2. Montrer que $\dim(V \cap W) = 1$. Interpréter géométriquement le résultat.

Exercice 3.7. Soit $V \subset \mathbb{R}^4$ un sous-espace vectoriel de dimension 2. Est-ce que toute base de $\mathbb{R}^4$ s'obtient en rajoutant deux vecteurs convenables à une base de V ? Argumenter la réponse. Formuler et résoudre l'exercice correspondant pour $D \subset \mathbb{R}^3$ de dimension 1.

Exercice 3.8. Donner une base et la dimension de $\mathrm{Vect}((1, 2, 3, 4), (1, 0, 5, 0), (0, 1, -1, 2))$.

Exercice 3.9. Soit E un espace vectoriel et soit $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_4\}$ une famille génératrice de E.
1) Les familles suivantes sont-elles génératrices de E ?

$$\begin{array}{ll} \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4, \mathbf{0}\} & \{\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\} \\ \{\boldsymbol{u}_1, \boldsymbol{u}_2 - 3\boldsymbol{u}_3, \boldsymbol{u}_2 + \boldsymbol{u}_3 + \boldsymbol{u}_4, \boldsymbol{u}_4\} & \{\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{u}_2 + \boldsymbol{u}_3, \boldsymbol{u}_3 + \boldsymbol{u}_4, \boldsymbol{u}_4 + \boldsymbol{u}_1\} \end{array}$$

2) Si $\dim E = n$, quelle inégalité vérifie n ?

Exercice 3.10. Soit E un espace vectoriel et soit $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_4\}$ une famille libre de E.
1) Les familles suivantes sont-elles libres ?

$$\begin{array}{ll} \{\boldsymbol{u}_1, \boldsymbol{u}_2, 0, \boldsymbol{u}_4\} & \{\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{u}_3 + \boldsymbol{u}_4\} \\ \{\boldsymbol{u}_1, \boldsymbol{u}_2 + \boldsymbol{u}_3 + \boldsymbol{u}_4, \boldsymbol{u}_4\} & \{\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{u}_2 + \boldsymbol{u}_3, \boldsymbol{u}_3 + \boldsymbol{u}_4, \boldsymbol{u}_4 + \boldsymbol{u}_1\} \end{array}$$

2) Si $\dim E = n$, quelle inégalité vérifie n ?

Exercice 3.11. Soit $U = \mathrm{Vect}\big((-1, 2, 2, 5, 0), (0, 2, 1, 2, 1), (1, 2, 0, -1, -2)\big)$ et soient V le sous-espace vectoriel engendré par les vecteurs $\boldsymbol{v}_1 = (-2, 0, 2, 6, -2)$, $\boldsymbol{v}_2 = (4, 10, 1, -2, 9)$, $\boldsymbol{v}_3 = (1, 6, 2, 3, 4)$ et $\boldsymbol{v}_4 = (-1, 2, -3, 4, 0)$.

1) Calculer les dimensions des sous-espaces U, V et $U+V$.
2) Construire une base de U et la compléter à une base de $U+V$.
3) Quelle est la dimension de $U \cap V$?
4) Calculer le rang de la famille de vecteurs formée avec $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ et les deux premiers vecteurs de la famille génératrice de U.
5) Donner une base pour $U \cap V$.

4. Équations de sous-espaces vectoriels

4.1. Systèmes d'équations linéaires et méthode du pivot de Gauss (un résumé)

On appelle équation linéaire à $n \geq 1$ variables (ou inconnues) $x_1, \ldots, x_n$ une équation de la forme
$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b,$$
où les coefficients $a_1, \ldots, a_n \in \mathbb{K}$. Lorsque $b = 0$ l'équation est dite homogène.

Définition. Une collection d'équations linéaires $L_1, \ldots, L_m$, $m \geq 1$, qui portent sur les mêmes inconnues
$$(L_\bullet) \quad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 & (L_1) \\ \vdots & \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n = b_m & (L_m) \end{cases}$$
s'appelle un *système d'équations linéaires.* Il est dit sans second membre ou *homogène* si tous les b_i sont nuls.

Le système linéaire $(L_\bullet)$ à n inconnues étant sonné, on note $\mathrm{Sol}(L_\bullet)$ l'ensemble des n-uplets $(x_1, \ldots, x_n) \in \mathbb{K}^n$ vérifiant les équations de $(L_\bullet)$, c'est-à-dire l'*ensemble solution* du système linéaire.

Lemme 4.1. *Si $(L_\bullet)$ est un système linéaire **homogène** à n inconnues, alors $\mathrm{Sol}(L_\bullet) \subset \mathbb{K}^n$ est un sous-eapce vectoriel.*

Démonstration. Une vérification immédiate de la définition d'un sous-espace vectoriel. Voir aussi la discussion précédant la question de la page 23. □

Proposition 4.2. *Soit $(L_\bullet)$ un système linéaire à n inconnues et soit $(L^h_\bullet)$ le système homogène associé. Si $\mathrm{Sol}(L_\bullet) \neq \emptyset$ et si $\boldsymbol{p}^* = (x_1^*, \ldots, x_n^*) \in \mathrm{Sol}(L_\bullet) \subset \mathbb{K}^n$ est une solution particulière, alors*
$$\mathrm{Sol}(L_\bullet) = \boldsymbol{u}^* + \mathrm{Sol}(L^h_\bullet).$$

Démonstration. Il faut démontrer la double inclusion. Par exemple, pour montrer "$\subset$", soit $\boldsymbol{q} = (q_1, \ldots, q_n)$ une autre solution de $(L_\bullet)$. Nous considérons la différence $\boldsymbol{q} - \boldsymbol{p}^*$ et nous évaluons en $\boldsymbol{q} - \boldsymbol{p}^*$ le membre de gauche de chaque équation L_i, $1 \leq i \leq m$. Nous obtenons

$$\begin{aligned} a_{i,1}(q_1 - x_1^*) + \cdots + a_{i,n}(q_n - x_n^*) &= (a_{i,1}q_1 + \cdots + a_{i,n}q_n) - (a_{i,1}x_1^* + \cdots + a_{i,n}x_n^*) \\ &= b_i - b_i = 0. \end{aligned}$$

Donc $\mathrm{Sol}(L_\bullet) \subset \boldsymbol{u}^* + \mathrm{Sol}(L_\bullet^h)$. L'argument pour l'autre inclusion est similaire □

Définition 4.3. Deux systèmes linéaires à n inconnues sont dits *équivalents* s'ils ont le même ensemble solution.

Lemme 4.4. *Soit $L_\bullet$ un système de $m \geq 2$ équations linéaires $L_1, \ldots, L_m$. Si $(L'_\bullet)$ est le système obtenu à partir de $(L_\bullet)$ en effectuant une des trois opérations suivantes,*
- (i) *multiplier une équation L_i par $\lambda \neq 0$*
- (ii) *remplacer une équation L_i par $L_i + \mu L_j$, avec $\mu \in \mathbb{K}$ et $j \neq i$*
- (iii) *permuter deux équations,*

alors $(L_\bullet)$ et $(L'_\bullet)$ sont équivalents.

Démonstration. Vérification directe. (Il faudrait comparer les trois règles avec celles qui apparaissent dans § 3.3 et permettent d'échelonner une matrice en utilisant ses colonnes.) □

Définition 4.5. Soit $(L_\bullet)$ un système de m équations linéaires. Pour tout $1 \leq i \leq m$, soit, s'il existe, j_i l'indice du premier coefficient non nul[5] de l'équation L_i. On dit que $(L_\bullet)$ est *échelonné* s'il existe un entier $r \leq m$ tel que
- (1) la suite $j_1, \ldots, j_r$ est strictement croissante (en particulier j_i existe pour tout $1 \leq i \leq r$)
- (2) $a_{i,j} = 0$ pour tous $r < i \leq m$ et $1 \leq j \leq n$.

Si $(L_\bullet)$ est échelonné, l'entier r s'appelle le rang du système. Notation : $\mathrm{rg}(L_\bullet)$.

Exemple. Pour $m = 4$, $n = 6$, le système ci-dessous est échelonné de rang 4. (Remarquer que le système est facile à résoudre. L'espace des solutions s'exprime en fonction des inconnues x_2 et x_4 dont les indices n'apparaissent pas dans la suite $(j_i)_{1 \leq i \leq 4}$.)

$$\begin{cases} 5x_1 - x_2 - x_3 - 2x_4 + x_6 = b_1 \\ 3x_3 + 2x_5 = b_2 \\ - x_5 + x_6 = b_3 \\ 5x_6 = b_4 \end{cases}$$

Théorème 4.6 (Pivot de Gauss)**.** *Tout système d'équations linéaires est équivalent à un système échelonné.*

Esquisse de preuve. Soit $(L_\bullet)$ un système d'équations linéaires avec m équations et n inconnues. Nous allons appliquer les trois opérations introduites dans le lemme 4.4, opérations qui transforment le système en un système équivalent

Soit $(j_i)_{1 \leq i \leq m}$ la suite des indices des premiers coefficients non nuls des équations L_1, $L_2, \ldots, L_m$. En permutant des équations si nécessaire (l'opération (iii) du lemme), on peut supposer que j_1 est le plus grand élément de la suite (j_i). Nous appliquons le pivot de Gauss

[5] L'équation L_i est de la forme $a_{i,j_i}x_{j_i} + \cdots + a_{i,n}x_n = b_i$ avec $a_{i,j_i} \neq 0$.

en utilisant la première équation (les opérations (i) et (ii) du lemme) et arrivons à un système $(L'_\bullet)$ pour lequel la suite (j'_i) satisfait

$$j'_1 = j_1 > j'_i$$

pour tout $i > 1$.

Par la suite, nous appliquons le même procédé en considérant les équations $L'_2, L'_3, \dots, L'_m$. □

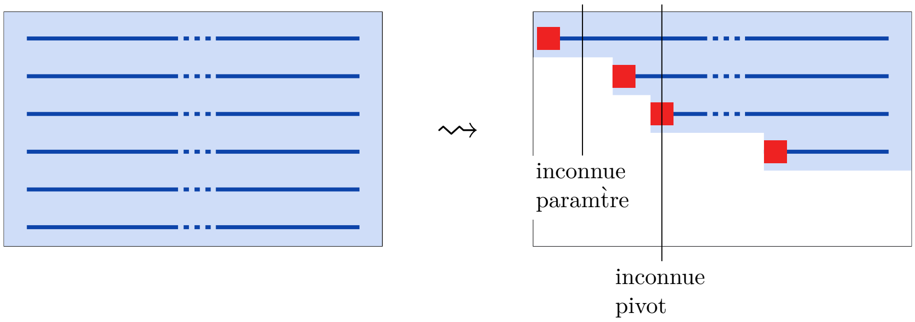

Figure 5: Échelonner un système linéaire. Les inconnues pivot correspondent aux éléments pivot représentés pour chaque équation du système échelonné Les autres inconnues deviennent des paramètres lors de la résolution du système.

Dans ce qui suit, à partir d'un système d'équations linéaires homogène donné, nous verrons comment trouver une base et la dimension du sous-espace vectoriel des solutions. Puis, dans § 4.3, partant d'un sous-espace vectoriel de $\mathbb{K}^n$ décrit par une base (ou une famille génératrice), nous construirons un système homogène d'équations linéaires le définissant. Chemin faisant, nous obtiendrons une preuve du théorème ci-dessous (réciproque du lemme 4.1).

Théorème 4.7. *Tout sous-espace vectoriel d'un espace vectoriel de dimension finie peut être défini par un système d'équations linéaires.*

4.2. Résolution des systèmes d'équations linéaires homogène

Le but de cette partie c'est de déterminer précisément (dimension et base) le sous-espace vectoriel décrit par un système d'équations linéaires homogène, c'est-à-dire de mettre en pratique le théorème 4.6. Par conséquent, nous utiliserons essentiellement la méthode du pivot de Gauss.

Prenons, pour commencer, un exemple situé dans $\mathbb{R}^5$. Soit $F \subset \mathbb{R}^5$ le sous-espace solution du système homogène

$$\begin{cases} x+y \phantom{{}+3z} + t + u = 0 \\ x+y+z+2t+2u = 0 \\ x+y+3z+4t+4u = 0. \end{cases}$$

On cherche sa dimension ainsi qu'une base. Pour cela on échelonne le système en suivant l'algorithme de la preuve du théorème 4.6 ; on obtient successivement,

$$\begin{cases} x+y \phantom{{}+3z} + t + u = 0 & (L'_1 = L_1) \\ \phantom{x+y+{}} z + t + u = 0 & (L'_2 = L_2 - L_1) \\ \phantom{x+y+{}} 3z + 3t + 3u = 0 & (L'_3 = L_3 - L_1) \end{cases}$$

et

$$\begin{cases} x + y \quad + \quad t + \quad u = 0 & (L_1') \\ \qquad\quad z + \quad t + \quad u = 0 & (L_2') \\ \qquad\qquad 0t + 0u = 0 & (L_3' - 3L_2'). \end{cases}$$

On remarque que le rang du système est 2. On voit qu'il se ramène à un système à deux équations et que l'on peut exprimer les deux inconnues x et z (*ayant été utilisées pour le pivot*) en fonction des trois autres y, t, u. Ces inconnues y, t, u deviennent des paramètres. On a

$$\begin{cases} x = -y - t - u \\ z = -t - u \end{cases}$$

et donc

$$F = \{(-y - t - u, y, -t - u, t, u) \mid y, t, u \in \mathbb{R}\}.$$

Pour trouver une base, on sépare les paramètres dans l'expression du vecteur général de F :

$$\begin{aligned}(-y - t - u, y, -t - u, t, u) &= (-y, y, 0, 0, 0) + (-t, 0, -t, t, 0) + (-u, 0, -u, 0, u) \\ &= y(-1, 1, 0, 0, 0) + t(-1, 0, -1, 1, 0) + u(-1, 0, -1, 0, 1).\end{aligned}$$

Donc $(x, y, z, t, u) \in F$ si et seulement si

$$\begin{pmatrix} x \\ y \\ z \\ t \\ u \end{pmatrix} = y \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + u \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} = y\boldsymbol{u}_1 + t\boldsymbol{u}_2 + u\boldsymbol{u}_3.$$

Cela prouve que les vecteurs $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ et $\boldsymbol{u}_3$ forment une famille génératrice de F. De plus, ils forment une famille libre : en effet si $\lambda_1\boldsymbol{u}_1 + \lambda_2\boldsymbol{u}_2 + \lambda_3\boldsymbol{u}_3 = \boldsymbol{0}$, alors, en inspectant les deuxième, quatrième et cinquième coordonnées du vecteur de gauche, on obtient $\lambda_i = 0$ pour tout i.

Remarque. $\dim(F) = \dim(\mathbb{R}^5) - 2 = 3$, où 2 est le nombre d'équations obtenues après échelonnement, c'est-à-dire le rang du système.

Cette remarque est vrai de manière générale ; la formule correspondante s'appelle *la formule du rang* et est une conséquence du fait que tout système (homogène) peut être échelonné (Théorème 4.6).

Théorème 4.8 (formule du rang)**.** *Soit* $(L_\bullet)$ *un système d'équations linéaires dans* $\mathbb{K}^n$*. Alors* $\dim(\mathrm{Sol}(L_\bullet)) = n - \mathrm{rg}(L'_\bullet)$*, où* $(L'_\bullet)$ *est un système échelonné équivalent à* $(L_\bullet)$*.*

Démonstration. La preuve reprend les idées de l'exemple précédent. Elle peut sembler technique, car dans le cas général, il faut expliciter les nombreux indices qui apparaissent lors de la construction de la base du l'espace solution.

Soit m le nombre d'équations du système considéré et soit F l'espace solution associé. D'après le théorème du pivot de Gauss, nous pouvons construire un système homogène échelonné $(L'_\bullet)$ équivalent à $(L_\bullet)$. Soit $r = \mathrm{rg}(L'_\bullet)$. Comme $(L'_\bullet)$ est un système homogène, ses équations $L'_{r+1}, \ldots, L'_m$ sont triviales. Nous les retirons du système ; cette opération ne modifie pas l'espace solution F. **Par la suite, les équations triviales seront ignorées**.

Soit $(j'_i)_{1\leq i\leq r}$ la suite des pivots, c'est-à-dire la suite des indices des premiers coefficients non nuls des équations L'_i. Pour déterminer le sous-espace vectoriel F, nous utilisons les équations de $(L'_\bullet)$ dans l'ordre inverse, L'_r, L'_{r-1},... pour exprimer les inconnues $x_{j'_i}$ en fonction des autres inconnues. Nous obtenons ainsi une description de l'élément général de F en fonction de $n-r$ paramètres.

Affirmation. $\dim(F) = n - r$.

Pour justifier ceci, soit

$$K = \{k_1, \dots, k_{n-r}\} = \{1, \dots, n\} \smallsetminus \{j'_1, \dots, j'_r\}$$

C'est l'ensemble des indices des inconnues qui n'ont pas été utilisées comme pivot, c'est-à-dire des paramètres. Les équations L'_r et L'_{r-1} deviennent

$$x_{j'_r} + \sum_{k\in K} \alpha_{j'_r,k} x_k = 0 \quad \text{et} \quad x_{j'_{r-1}} + \alpha_{j'_{r-1},j'_r} x_{j'_r} + \sum_{k\in K} \alpha_{j'_{r-1},k} x_k = 0$$

car $j'_{r-1} < j'_r$. On obtient, en notant $\mu = \alpha_{j'_{r-1},j'_r}$,

$$\begin{aligned} x_{j'_r} &= -\sum_{k\in K} \alpha_{j'_r,k} x_k \\ x_{j'_{r-1}} &= -\mu x_{j'_r} - \sum_{k\in K} \alpha_{j'_{r-1},k} x_k = \sum_{k\in K} (\mu\alpha_{j'_r,k} - \alpha_{j'_{r-1},k}) x_k \end{aligned}$$

En poursuivant les calculs avec les autres équations du système, nous arrivons à $\boldsymbol{u} \in F$ si et seulement si

$$x_j = \begin{cases} \text{combinaison linéaire des } x_k,\ k \in K & , \quad \text{si } j \notin K \\ x_k, & \quad \text{si } j \in K. \end{cases}$$

Donc il existe $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-r}$ vecteurs de F tels que

- $\boldsymbol{u} = x_{k_1}\boldsymbol{u}_1 + \cdots + x_{k_{n-r}}\boldsymbol{u}_{n-r}$
- La coordonnée k_i des vecteurs $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-r}$ est nulle sauf pour $\boldsymbol{u}_i$.

C'est cette dernière propriété qui nous dit que les vecteurs $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-r}$ sont linéairement indépendants. Donc $F = \mathrm{Vect}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-r})$ et $\dim(F) = n - r$. □

Corollaire-Définition 4.9. *Soit $(L_\bullet)$ un système d'équations linéaires homogènes à n inconnues. Tout système échelonné équivalent à $(L_\bullet)$ a le même rang $n - \dim(\mathrm{Sol}(L_\bullet))$ appelé le rang du système linéaire.*

Corollaire. *Soit $(L_\bullet)$ un système de m équations linéaires homogènes à n inconnues. Alors* $\dim(\mathrm{Sol}(L_\bullet)) \leq n - m$

Démonstration. Comme $r = \mathrm{rg}(L_\bullet) \leq m$, alors $\dim(\mathrm{Sol}(L_\bullet)) = n - r \geq n - m$. □

Définition 4.10. Les équations linéaires homogènes $L_1, \dots, L_m$ sont dites linéairement indépendantes si $\mathrm{rg}(L_\bullet) = m$.

Dire que les équations sont linéairement indépendantes est équivalent à dire qu'après avoir échelonné le système associé, on n'a pas perdu d'équations. Dans ce cas, on a $\dim(\mathrm{Sol}(L_\bullet)) = n - m$.

On retiendra que $n-k$ équations linéairement indépendantes déterminent un sous-espace vectoriel de dimension k. En particulier, un sous-espace vectoriel donné par une seule équation (non nulle) est de dimension $\dim E - 1 = n-1$ c'est-à-dire un hyperplan.

Dans § 4.3 nous verrons qu'un sous-espace vectoriel de dimension k est déterminé par $n-k$ équations linéaires homogènes linéairement indépendantes.

Exemple. Prenons un système de trois équations homogènes dans $\mathbb{R}^3$ et notons F l'espace solution. *A priori*, $\dim(F) \leq 3$ car on est dans $\mathbb{R}^3$) et $\dim(F) \geq 3-3=0$ car F est défini par 3 équations. Cela ne donne finalement aucun renseignement précis. Regardons le cas particulier suivant :

$$\begin{cases} 2x - y + z = 0 \\ x + y + z = 0 \\ 3x \phantom{{}-y} + 2z = 0 \end{cases}$$

Appliquons la méthode du pivot. Successivement le système devient

$$\begin{cases} x + y + z = 0 & (L_1' = L_2) \\ - 3y - z = 0 & (L_2' = L_1 - 2L_2) \\ - 3y - z = 0 & (L_3' = L_3 - 3L_2) \end{cases}$$

et

$$\begin{cases} x + y + \phantom{\frac{1}{3}}z = 0 \\ \phantom{x+{}} y + \frac{1}{3}z = 0. \end{cases}$$

Dans ce cas, le système est de rang 3 et $\dim(F) = 3-2 = 1$. Le système définit une droite. En prenant l'inconnue z comme paramètre, le système peut s'écrire, en remontant à partir de la dernière équation,

$$\begin{cases} x = -\frac{2}{3}z \\ y = -\frac{1}{3}z \\ z = z \end{cases} \iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \begin{pmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ 1 \end{pmatrix}.$$

Cela montre que le vecteur $\boldsymbol{u} = \left(-\frac{2}{3}, -\frac{1}{3}, 1\right)$ est une base de F.

4.3. Équations linéaires d'un sous-espace vectoriel donné par une famille génératrice

Supposons que l'on connaisse un sous-espace vectoriel F simplement par une de ses familles génératrices, c'est-à-dire $F = \mathrm{Vect}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$. On veut décrire F à l'aide d'un système d'équations linéaires. Voyons la méthode sur un exemple. Cet algorithme représente l'idée de la preuve du théorème 4.7.

Soit $F = \mathrm{Vect}\{\boldsymbol{v}_1, \boldsymbol{v}_2\} \subset \mathbb{R}^4$ avec $\boldsymbol{v}_1 = (1,1,-1,1)$ et $\boldsymbol{v}_2 = (2,0,1,0)$. Soit $\boldsymbol{u} = (x,y,z,t) \in \mathbb{R}^4$. Alors

$$\boldsymbol{u} \in F \qquad \text{si et seulement si} \qquad \exists \lambda, \mu \in \mathbb{R} \quad \text{tels que} \quad \boldsymbol{u} = \lambda \boldsymbol{v}_1 + \mu \boldsymbol{v}_2.$$

En interprétant cette dernière identité composante par composante, on a $\boldsymbol{u} \in F$ si et seulement s'il existe $\lambda, \mu \in \mathbb{R}$ tels que

$$\begin{cases} x & = \lambda + 2\mu \\ \quad y & = \lambda \\ \qquad z & = -\lambda + \mu \\ \qquad\quad t & = \lambda \end{cases} \quad \text{c'est-à-dire} \quad \begin{cases} x \quad - \lambda - 2\mu = 0 \\ \quad y \quad - \lambda \qquad = 0 \\ \qquad z \ + \lambda - \ \mu = 0 \\ \qquad\quad t - \lambda \qquad = 0 \end{cases}$$

La technique consiste à appliquer le pivot de Gauss avec λ et μ comme *pivots* à ce système de quatre équations avec six inconnues. Les équations dans lesquelles λ et μ ont été éliminées formeront un système d'équations pour F.

Nous avons

$$\begin{cases} \qquad\quad t - \lambda \qquad = 0 \\ \qquad z \ + \lambda - \ \mu = 0 \\ \quad y \quad - \lambda \qquad = 0 \\ x \quad - \lambda - 2\mu = 0 \end{cases} \rightsquigarrow \begin{cases} \qquad\quad t - \lambda \qquad = 0 & (L_1) \\ \qquad z + t \qquad - \ \mu = 0 & (L_2 + L_1) \\ \quad y \ - t \qquad\qquad = 0 & (L_3 - L_1) \\ x \quad - t \qquad - 2\mu = 0 & (L_4 - L_1) \end{cases}$$

et finalement, en utilisant μ comme pivot,

$$\begin{cases} \qquad\qquad t - \lambda \qquad = 0 \\ \qquad z + \ t \qquad - \mu = 0 \\ \quad y \qquad - \ t \qquad = 0 \\ x \ - 2z - 3t \qquad = 0 \end{cases}$$

On garde les équations ne contenant plus de λ et μ. On a donc

$$F = \left\{ (x, y, z, t) \in \mathbb{R}^4 \;\middle|\; \begin{cases} x \ - 2z - 3t = 0 \\ \quad y \qquad - \ t = 0 \end{cases} \right\}.$$

Esquisse de preuve du théorème 4.7. Soit $F = \mathrm{Vect}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_p\} \subset E$, où E est un $\mathbb{K}$-espace vectoriel de dimension n. Soit $\mathcal{B} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ une base de E. Alors

$$\begin{aligned} \boldsymbol{v}_1 &= a_{1,1}\boldsymbol{e}_1 + \cdots + a_{n,1}\boldsymbol{e}_n \\ &\ \vdots \\ \boldsymbol{v}_p &= a_{1,p}\boldsymbol{e}_1 + \cdots + a_{n,p}\boldsymbol{e}_n \end{aligned}$$

et le rang de la famille de vecteurs colonnes

$$C_{\mathcal{B}}(\boldsymbol{v}_1) = \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{n,1} \end{pmatrix}, \dots, \quad C_{\mathcal{B}}(\boldsymbol{v}_{p-1}) = \begin{pmatrix} a_{1,p-1} \\ \vdots \\ a_{n,p-1} \end{pmatrix}, \quad C_{\mathcal{B}}(\boldsymbol{v}_p) = \begin{pmatrix} a_{1,p} \\ \vdots \\ a_{n,p} \end{pmatrix} \tag{4.1}$$

est la dimension de F. Nous supposons, pour simplifier l'argument, que $\dim(F) = p$, c'est-à-dire les vecteurs $\boldsymbol{v}_j$ sont linéairement indépendants. *En particulier* $p \leq n$.

Comme dans l'exemple ci-dessus, $\boldsymbol{v} \in F$ si et seulement s'il existe $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ tels que

$$\boldsymbol{v} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n = \lambda_1\boldsymbol{v}_1 + \cdots + \lambda_p\boldsymbol{v}_p$$

c'est-à-dire

$$\begin{cases} x_1 = \lambda_1 a_{1,1} + \cdots + \lambda_p a_{1,p} \\ \quad\vdots \\ x_n = \lambda_1 a_{n,1} + \cdots + \lambda_p a_{n,p}. \end{cases}$$

En réarrangeant les termes, nous arrivons au système

$$(L_\bullet) \quad \left\{ \begin{array}{llllllll} a_{1,1}\lambda_1 & +\cdots+ & a_{1,p}\lambda_p & - \; x_1 & & & = & 0 \\ a_{2,1}\lambda_1 & +\cdots+ & a_{2,p}\lambda_p & & - \; x_2 & & = & 0 \\ \quad\vdots & & & & & & & \\ a_{n,1}\lambda_1 & +\cdots+ & a_{n,p}\lambda_p & & & - \; x_n & = & 0. \end{array} \right.$$

Le système $(L_\bullet)$ est à $p+n$ inconnues avec n équations. Son espace solution $\mathrm{Sol}(L_\bullet)$ vit dans $\mathbb{R}^{p+n}$ et est de dimension p —il est échelonné dans les n variables x_i, donc son rang est n.

On recommence l'échelonnement du système par la méthode de Gauss en utilisant comme variables pivot les λ_k. Ceci peut être fait car le rang de la famille des vecteurs (4.1) est $p \leq n$. Après p opérations (nous ne comptons pas les permutations d'équations), on arrive à un système du type décrit dans le figure 6.

$(L'_\bullet)$

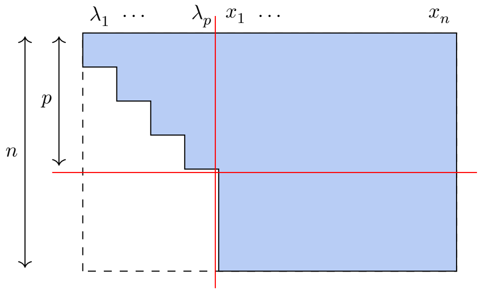

Figure 6: Le système $(L_\bullet)$ est transformé en $(L'_\bullet)$ par la méthode de Gauss qui utilise les inconnues λ_k comme pivot. Dans cette représentation graphique, $n = 7$ et $p = 4$. Les $n-p$ équations en bas à droite représentent un système d'équations pour le sous-espace F initial.

Pour finir la preuve, il reste à montrer que les $n-p$ équations en bas à droite —équations ne faisant pas apparaître les λ_k —représentent un système d'équations pour le sous-espace F dans les coordonnées $x_1, \ldots, x_n$. Explicitement, en notant par $(\Lambda_\bullet)$ ce système d'équations linéaires homogènes (à n inconnues), il faut monter

$$\mathrm{Sol}(\Lambda_\bullet) = \mathrm{Vect}\left(C_{\mathcal{B}}(\boldsymbol{v}_1), \ldots, C_{\mathcal{B}}(\boldsymbol{v}_p)\right) \subset \mathbb{R}^n.$$

Cette double inclusion sera discutée en TD. $\square$

4.4. Le rang d'une famille de vecteurs colonnes et le rang d'un système d'équations linéaires

Nous avons rencontré deux notions de rang. D'abord, nous avons introduit le rang d'une famille de vecteurs dans la définition 3.8 comme étant la dimension du sous-espace engendré par la famille. Du point de vue pratique, la proposition 3.12 nous dit que le rang peut être calculé en échelonnant (en travaillant en colonnes) la matrice de decomposition de la famille dans une base fixée. Par la suite, dans le corollaire-définition 4.9, nous avons introduit la notion de rang d'un système d'équations linéaires homogènes et avons vu qu'il peut être calculé en échelonnant le système (en travaillant avec les équations, c'est-à-dire “en ligne”).

Il faut comprendre que ces deux notions de rang sont calculées en utilisant un même objet d'algèbre linéaire, une *matrice.* Dans le premier cas, la matrice de décomposition dans une base fixée nous permet de calculer le rang par le pivor de Gauss. Dans le deuxième cas, la matrice n'a pas encore été mise en évidence : c'est la *matrice des coefficients du système.* Les opérations de la méthode du pivot que nous avons utilisées pour échelonner le système, peuvent être vues comme opérations sur les lignes de la matrice des coefficients. Par exemple, le système

$$\begin{cases} x - y - z - 3s + t = 0 \\ 2x - y + z + t = 0 \\ 2y + 2s + t = 0 \\ x - 2y + 2z + s - t = 0 \end{cases}$$

devient, en utlisant succéssivement x (la première équation) comme pivot et puis y (la deuxème équation)

$$\begin{cases} x - y - z - 3s + t = 0 \\ y + 3z + 6s - t = 0 \\ 2y + 2s + t = 0 \\ -y + 3z + 4s - 2t = 0 \end{cases} \rightsquigarrow \begin{cases} x - y - z - 3s + t = 0 \\ y + 3z + 6s - t = 0 \\ -6z - 10s + 3t = 0 \\ 6z + 10s - 3t = 0 \end{cases}$$

c'est-à-dire il est équivalent au système échelonné ci-dessous.

$$\begin{cases} x - y - z - 3s + t = 0 \\ y + 3z + 6s - t = 0 \\ -6z - 10s + 3t = 0 \end{cases}$$

En particulier, son rang est 3. Maintenant, toutes les opérations effectuées lors de l'échelonnement du système, peuvent être lues comme *opérations sur les lignes des matrices des coefficients* —on interprète les coefficients de manière détachée, tout en sachant que chaque colonne correspond à une inconnue :

$$\begin{pmatrix} 1 & -1 & -1 & -3 & 1 \\ 2 & -1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 2 & 1 \\ 1 & -2 & 2 & 1 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 & -1 & -3 & 1 \\ 0 & 1 & 3 & 6 & -1 \\ 0 & 2 & 0 & 2 & 1 \\ 0 & -1 & 3 & 4 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 & -1 & -3 & 1 \\ 0 & 1 & 3 & 6 & -1 \\ 0 & 0 & -6 & -10 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

La matrice initiale (celle de gauche) décrit aussi cinq vecteurs colonnes de $\mathbb{R}^4$ dans la base canonique. Si nous echelonnons cette matrice, en travaillant en colonne cette fois ci, nous obtenons le rang de la famille formée par

$$\begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ -1 \\ 2 \\ -2 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} -3 \\ 0 \\ 2 \\ 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}.$$

Le résultat est 3 de nouveau —il faut faire ce calcul. Le résultat général suggéré par cet exemple est le suivant :

Proposition-Définition 4.11. *Soit A une matrice de taille $n \times p$. Soient $\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_p \in \mathbb{R}^n$ les p vecteurs colonnes de A et soient $\boldsymbol{L}_1, \boldsymbol{L}_2 \ldots, \boldsymbol{L}_n \in \mathbb{R}^p$ les n vecteurs lignes de A. Alors*

$$\dim(\mathrm{Vect}(\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_p)) = \dim(\mathrm{Vect}(\boldsymbol{L}_1, \boldsymbol{L}_2 \ldots, \boldsymbol{L}_n)) =: \mathrm{rg}(A)$$

et ce nombre est appelé le rang de A.

Démonstration. Soit $r = \dim(\mathrm{Vect}(\boldsymbol{L}_1, \ldots, \boldsymbol{L}_p))$, c'est-à-dire en utlisant les opérations (par "ligne") introduites dans le lemme 4.4, la matrice A est échelonnée en une matrice A' avec

$$A' = \begin{pmatrix} 1 & a'_{1,2} & \ldots & a'_{1,r} & a'_{1,r+1} & \ldots & a'_{1,p} \\ 0 & 1 & \ldots & a'_{2,r} & a'_{2,r+1} & \ldots & a'_{2,p} \\ \vdots & & & & & & \\ 0 & 0 & \ldots & 1 & a'_{r,r+1} & \ldots & a'_{r,p} \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & & & & & & \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Nous affirmons que les vecteurs colonnes de A qui correspondent aux r pivots es lignes obtenus dans A' sont linéairement indépendants. Il s'ensuit que

$$\dim(\mathrm{Vect}(\boldsymbol{L}_1, \boldsymbol{L}_2 \ldots, \boldsymbol{L}_n)) \leq \dim(\mathrm{Vect}(\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_p)).$$

Pour justifier l'affirmation, on suppose que les r pivots des lignes apparaissent sur les premières colonnes. On considère une combinaison linéaire quelconque nule formée avec les r premières colonnes de A, c'est-à-dire on considère α_1, α_2, ..., α_r des scalaires tels que

$$\alpha_1 \boldsymbol{C}_1 + \alpha_2 \boldsymbol{C}_2 + \cdots + \alpha_r \boldsymbol{C}_r = 0_{n \times 1}.$$

Mais cette relation linéaire n'est pas modifiée par les opérations efféctuées pour passer de A à A'. Donc

$$\alpha_1 \boldsymbol{C}'_1 + \alpha_2 \boldsymbol{C}'_2 + \cdots + \alpha_r \boldsymbol{C}'_r = 0_{n \times 1},$$

où on a noté par $\boldsymbol{C}'_j$ la j-ème colonne de A'. Donc, en utlisant la forme des $\boldsymbol{C}'_j$, on arrive à

$$\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0.$$

En faisant le même raisonnement mais en échangeant les rôles des lignes et des colonnes, on arrive à

$$\dim(\mathrm{Vect}(\boldsymbol{L}_1, \boldsymbol{L}_2 \ldots, \boldsymbol{L}_n)) \geq \dim(\mathrm{Vect}(\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_p)),$$

d'où le résultat. □

4.5. Exercices

Exercice 4.1. Déterminer une base et la dimension du sous-espace vectoriel V de $\mathbb{R}^4$ formé des (x, y, z, t) tels que

$$\begin{cases} 9x + 13y + 30z + 19t = 0 \\ 8x + 9y + 19z + 22t = 0 \\ x - y - 4z + 7t = 0 \end{cases}$$

Exercice 4.2. Soient $U, V \subset \mathbb{R}^4$ définis par

$$\begin{cases} x + y - 2z + \ \ t = 0 \\ -x + y + \ \ z - 3t = 0 \end{cases} \quad \text{et} \quad \begin{cases} x - 2y + z + 2t = 0 \\ \qquad y + z - 3t = 0. \end{cases}$$

1) Expliquer pourquoi U et V sont des sous-espaces vectoriels.

2) Déterminer une base de $U \cap V$ et puis étendre la base à une base de U et par la suite, à une de $U + V$.

Exercice 4.3. Dans $\mathbb{R}^4$, soient $F = \{(a+b, a-b, a, 0),\ a, b \in \mathbb{R}\}$ et $G = \{(y, x, y, x-y),\ x, y \in \mathbb{R}\}$.

1) Trouver des bases de $F, G, F \cap G, F + G$.

2) Donner des équations de $F, G, F \cap G, F + G$ en nombre minimal dans la base canonique de $\mathbb{R}^4$.

Exercice 4.4. Dans $\mathbb{R}^3$, soient $\boldsymbol{u}_1 = (1, 6, -4)$, $\boldsymbol{u}_2 = (3, 2, 4)$, $\boldsymbol{u}_3 = (1, -1, 4)$, $\boldsymbol{u}_4 = (1, 3, -2)$. On pose $F = \mathrm{Vect}(\boldsymbol{u}_1, \boldsymbol{u}_2)$, $G = \mathrm{Vect}(\boldsymbol{u}_3, \boldsymbol{u}_4)$.

1) Déterminer une base puis des équations en nombre minimal dans la base canonique de $\mathbb{R}^3$ pour F et G.

2) Déterminer une base de $F \cap G$.

3) Que peut-on dire de $F + G$?

Exercice 4.5. Soit $F \subset \mathbb{R}^5$ le sous-espace vectoriel engendré par les vecteurs $\boldsymbol{v}_1 = (-2, 0, 2, 6, -2)$, $\boldsymbol{v}_2 = (4, 10, 1, -2, 9)$, $\boldsymbol{v}_3 = (1, 6, 2, 3, 4)$ et $\boldsymbol{v}_4 = (-1, 2, -3, 4, 0)$.

1) Trouver une base et la dimension de F.

2) Trouver un système d'équations homogènes dont l'espace solution soit F.

3) Soit H l'espace des solutions de $x_4 = 0$. Trouver un système d'équations homogènes à quatre inconnues, de rang maximal, dont l'espace solution soit le sous-espace vectoriel $F \cap H \subset H$.

5. Applications linéaires

Le décor est constitué de deux espaces vectoriels E et F sur un même corps $\mathbb{K}$. On souhaite pouvoir passer de E à F tout en conservant la structure d'espace vectoriel. On va considérer pour cela des applications respectant les lois d'un espace vectoriel : l'addition et la multiplication par scalaires. La question à laquelle on essaiera de répondre est : Quelles sont les applications linéaires (c'est-à-dire celles qui commutent avec les combinaisons linéaires) entre E et F ?

5.1. Définition et premières propriétés

On va considérer des applications d'un espace vectoriel dans un autre qui respectent la structure d'espace vectoriel. Cela permettra, entre autres, de pouvoir transporter des résultats vrais dans l'un vers l'autre.

Définition. Soient E et F deux $\mathbb{K}$-espaces vectoriels. On dit que l'application $f : E \to F$ est linéaire si
(i) pour tous $\boldsymbol{u}, \boldsymbol{v} \in E$ on a $f(\boldsymbol{u}+\boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v})$
(ii) pour tout $\lambda \in \mathbb{K}$ et pour tout $\boldsymbol{u} \in E$, on a $f(\lambda\boldsymbol{u}) = \lambda f(\boldsymbol{u})$.

Les deux conditions (i) et (ii) de la définition peuvent être réunies en une seule : pour tous $\lambda, \mu \in \mathbb{K}$ et pour tous $u, v \in E$ on a

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

Une application linéaire est donc une application telle que l'image de toute combinaison linéaire est la combinaison linéaire correspondante des images.

Lemme 5.1. *Si $f : E \to F$ est linéaire, alors*
1) $f(\mathbf{0}_E) = \mathbf{0}_F$
2) *pour tout $\boldsymbol{u} \in E$ on a $f(-\boldsymbol{u}) = -f(\boldsymbol{u})$*
3) *pour tous $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in E$ et pour tous $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ on a $f\left(\sum_{i=1}^n \lambda_i \boldsymbol{u}_i\right) = \sum_{i=1}^n \lambda_i f(\boldsymbol{u}_i)$.*

Démonstration. 1) est la propriété (i) de la définition avec $\boldsymbol{v} = -\boldsymbol{u}$. 2) est la propriété (ii) de la définition avec $\lambda = -1$. Le dernier point se démontre par récurrence. □

Exemples. On indique, pour chaque exemple, le domaine, le codomaine et la correspondance.
— $\mathbb{R} \to \mathbb{R}$ définie par $x \mapsto 2x$ est linéaire
— $\mathbb{R}^2 \to \mathbb{R}^2$ définie par $(x, y) \mapsto (x+y, x-y)$ est linéaire
— $\mathbb{R}^3 \to \mathbb{R}$ définie par $(x, y, z) \mapsto x$ est linéaire
— $\mathbb{R} \to \mathbb{R}$ définie par $x \mapsto x^2$ n'est pas linéaire (on prend $\lambda = 2$ dans la propriété (ii) de la définition)
— $E \to E$ définie par $x \mapsto \lambda x$, avec $\lambda \in \mathbb{K}$ fixé, est linéaire (*homothétie de rapport* λ)
— $D : \mathbb{R}[X] \to \mathbb{R}[X]$ définie par $D(P) = P'$ est linéaire
— $\mathcal{C}([a, b], \mathbb{R}) \to \mathbb{R}$ définie par $f \mapsto \int_a^b f(t)\, dt$ est linéaire.

Il existe un lien fondamental entre applications linéaires et bases.

Proposition 5.2. *Une application linéaire est entièrement déterminée par l'image des vecteurs d'une base du domaine de définition.*

Démonstration. En effet, si E est un espace vectoriel ayant pour base $\mathcal{B} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ et si $f : E \to F$ est linéaire, alors pour tout $\boldsymbol{u} \in E$, il existe $a_1, \dots, a_n \in \mathbb{K}$ tels que

$$\boldsymbol{u} = a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n.$$

Alors

$$f(\boldsymbol{u}) = f(a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n) = a_1 f(\boldsymbol{e}_1) + \cdots + a_n f(\boldsymbol{e}_n).$$

Donc la connaissance des $f(\boldsymbol{e}_1), \dots, f(\boldsymbol{e}_n)$ conduit à la connaissance de l'image de tout vecteur de E. □

Une application linéaire $f : E \to \mathbb{K}$, donc à valeurs dans $\mathbb{K}$ – le corps de base de son domaine de définition –, est appelée *forme linéaire*. Par la suite, le but sera d'écrire explicitement les formes linéaires définies sur E de dimension finie.

Lemme 5.3. *Si $\mathcal{B} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ est une base du $\mathbb{K}$-espace vectoriel E, alors les applications coordonnées associées à $\mathcal{B}$, $x_1, \dots, x_n : E \to \mathbb{K}$ définies par*

$$x_i(\boldsymbol{u}) = x_i(a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n) = a_i \quad \textit{pour tout} \quad 1 \le i \le n,$$

sont linéaires.

Démonstration. En découle des propriétés des bases. □

En utilisant le lemme précédent, on décrit toutes les formes linéaires sur E.

Proposition 5.4. *Soit E un $\mathbb{K}$-espace vectoriel et soit $\mathcal{B} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ une base de E. Alors toute forme linéaire de E est une combinaison linéaire des formes linéaires coordonnées associées à $\mathcal{B}$.*

Démonstration. Soit $f : E \to \mathbb{K}$ une forme linéaire. Si $\boldsymbol{u} \in E$, alors $\boldsymbol{u} = x_1(\boldsymbol{u})\boldsymbol{e}_1 + \cdots + x_n(\boldsymbol{u})\boldsymbol{e}_n$. D'après la proposition 5.2,

$$\begin{aligned} f(\boldsymbol{u}) &= f(x_1(\boldsymbol{u})\boldsymbol{e}_1 + \cdots + x_n(\boldsymbol{u})\boldsymbol{e}_n) \\ &= x_1(\boldsymbol{u})f(\boldsymbol{e}_1) + \cdots + x_n(\boldsymbol{u})f(\boldsymbol{e}_n) \end{aligned}$$

c'est-à-dire

$$f = f(\boldsymbol{e}_1)\, x_1 + \cdots + f(\boldsymbol{e}_n)\, x_n.$$

Réciproquement, il est clair qu'une application de cette forme est bien une forme linéaire. □

Exemple. Une forme linéaire sur $\mathbb{R}^3$ est de la forme $f : \mathbb{R}^3 \to \mathbb{R}, (x, y, z) \mapsto ax + by + cz$. De même, l'application

$$I : P = a + bX + cX^2 \mapsto \int_0^1 P(t)\, dt$$

est une forme linéaire sur $\mathbb{R}_{\le 2}[X]$. On a

$$\begin{aligned} I(P) &= \int_0^1 (a + bt + ct^2)\, dt \\ &= a\int_0^1 dt + b\int_0^1 t\, dt + c\int_0^1 t^2\, dt \\ &= aI(1) + bI(X) + cI(X^2) \\ &= a + \frac{1}{2}b + \frac{1}{3}c. \end{aligned}$$

Dans l'espace des polynômes de degré ≤ 2, les coefficients a, b et c représentent les formes coordonnées associées à la base canonique $(1, X, X^2)$.

Corollaire 5.5. *Soit $f : V \to \mathbb{K}^p$ une application et soient $y_1, \dots, y_p$ les applications coordonnées de $\mathbb{K}^p$ associées à la base canonique. Alors f est linéaire si et seulement si $y_j \circ f$ est linéaire pour tout $1 \le j \le p$ (c'est-à-dire les composantes de f sont linéaires).*

Démonstration. Un sens est clair car la composition de deux applications linéaires est une application linéaire (voir 5.11).

Supposons que $f_j = y_j \circ f$ est linéaire pour pour tout j. Alors, pour tous $\boldsymbol{u}, \boldsymbol{v} \in E$ et pour tous $\alpha, \beta \in \mathbb{K}$, on a

$$\begin{aligned} f(\alpha \boldsymbol{u} + \beta \boldsymbol{v}) &= \big(f_1(\alpha \boldsymbol{u} + \beta \boldsymbol{v}), \ldots, f_p(\alpha \boldsymbol{u} + \beta \boldsymbol{v})\big) \\ &= \big(\alpha f_1(\boldsymbol{u}) + \beta f_1(\boldsymbol{v}), \ldots, \alpha f_p(\boldsymbol{u}) + \beta f_p(\boldsymbol{v})\big) \\ &= \alpha\big(f_1(\boldsymbol{u}), \ldots, f_p(\boldsymbol{u})\big) + \beta\big(f_1(\boldsymbol{v}), \ldots, f_p(\boldsymbol{v})\big) \\ &= \alpha f(\boldsymbol{u}) + \beta f(\boldsymbol{v}). \end{aligned}$$

□

5.2. Noyau et image d'une application linéaire

Proposition 5.6. *Soit $f : E \to F$ une application linéaire.*

(i) *Pour tout sous-espace vectoriel E' de E, l'image de E', $f(E') = \{f(\boldsymbol{v}) \mid \boldsymbol{v} \in E'\}$ est un sous-espace vectoriel de F.*

(ii) *Pour tout sous-espace vectoriel F' de F, l'image réciproque de F', $f^{-1}(F') = \{\boldsymbol{v} \in E \mid f(\boldsymbol{v}) \in F'\}$ est un sous-espace vectoriel de E.*

Démonstration. A vérifier en exercice. □

En particulier, suite à ce résultat, si $f : E \to F$ est une application linéaire,

$$\mathrm{im}(f) = f(E) = \{f(\boldsymbol{v}) \in E \mid \boldsymbol{v} \in E\}$$

est un sous-espace vectoriel de F. Il existe aussi un sous-espace naturel de E associé à f.

Définition. Soit $f : E \to F$ une application linéaire. On appelle noyau de f noté $\ker(f)$ l'image réciproque de $\{\mathbf{0}_F\}$ (qui est un sous-espace vectoriel de E).

On a le diagramme suivant correspondant à une application linéaire.

$$\ker(f) \subset E \xrightarrow{f} \mathrm{im}(f) \subset F$$

On rappelle qu'en général, une fonction $\varphi : X \to Y$ est dite

- *injective* si pour tout $y \in Y$, $\mathrm{card}(f^{-1}(\{y\}) \geq 1$
- *surjective* si pour tout $y \in Y$, $\mathrm{card}(f^{-1}(\{y\}) \leq 1$
- *bijective* si pour tout $y \in Y$, $\mathrm{card}(f^{-1}(\{y\}) = 1$, c'est-à-dire f est injective et surjective.

Par exemple, si X est un ensemble d'étudiants et si $Y = \{0, 1, \ldots, 20\}$ ensemble des notes, alors l'application φ, note à un examen donné, n'est en général pas injective car deux étudiants différents peuvent très bien avoir la même note; en général elle n'est pas surjective non plus, car il arrive fréquemment que certaines notes ne soient pas attribuées. Un exemple de fonction injective est celle qui à tout véhicule attribue sa plaque minéralogique. Par ailleurs, il est clair que φ est surjective si et seulement si $\mathrm{im}(\varphi) = Y$.

L'injectivité et la surjectivité se caractérisent facilement dans le cas d'applications linéaires.

Lemme 5.7. *Si $f : E \to F$ une application linéaire, alors f est injective si et seulement si $\ker(f) = \{\mathbf{0}_E\}$.*

Démonstration. On a toujours $f(\mathbf{0}_E) = \mathbf{0}_F$, donc $\{\mathbf{0}_E\} \subset \ker(f) = f^{-1}(\mathbf{0}_F)$. Alors, si f est injective, on a $\ker(f) = \{\mathbf{0}_E\}$. Réciproquement, si $\ker(f) = \{\mathbf{0}_E\}$, comme pour deux vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ ayant la même image,

$$f(\boldsymbol{u} - \boldsymbol{v}) = f(\boldsymbol{u}) - f(\boldsymbol{v}) = \mathbf{0}_F,$$

on conclut que $\boldsymbol{u} - \boldsymbol{v} = \mathbf{0}_E$. □

Remarque (méthodes de calcul)**.** Pour déterminer le noyau d'une application linéaire f entre deux espaces vectoriel de dimension fini, on résout l'équation $f(\boldsymbol{v}) = 0$. Cela se traduit par la résolution d'un système linéaire (après avoir choisi une base de E).

Puisque $\mathrm{im}(f)$ est un sous-espace vectoriel vectoriel de F, $\mathrm{im}(f) = F$ équivaut à $\dim(\mathrm{im}(f)) = \dim F$, donc il suffit de connaître la dimension de $\mathrm{im}(f)$ pour savoir si elle est surjective : on considère les vecteurs images d'une base de E et on calcul la dimension du sous-espace engendré.

Cette dernière remarque suggère la définition suivante qui permettra de conclure que f est surjective si et seulement si $\mathrm{rg}(f) = \dim(F)$.

Définition. Si $f : E \to F$ est une application linéaire, on appelle *rang* de f, noté $\mathrm{rg}(f)$, la dimension de l'image de f.

Comment trouver le rang d'une application lináire ? À l'aide d'une famille génératrice. Explicitement, si $\mathcal{B} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ est une base de E, alors

$$\mathrm{im}(f) = \mathrm{Vect}\{f(\boldsymbol{e}_1), f(\boldsymbol{e}_2), \dots, f(\boldsymbol{e}_n)\}.$$

Donc le rang de f est le rang de la famille de vecteurs $\{f(\boldsymbol{e}_1), f(\boldsymbol{e}_2), \dots, f(\boldsymbol{e}_n)\}$.

Exemple. On considère l'application f de $\mathbb{R}^3$ dans $\mathbb{R}^3$ définie par

$$(x, y, z) \longmapsto [f](x', y', z') = (x + z, y + z, x + z).$$

On veut savoir si f est injective, surjective, bijective ? On recherche donc le noyau et l'image de f. Pour le noyau on a

$$(x, y, z) \in \ker(f) \quad \text{si et seulement si} \quad f(x, y, z) = (0, 0, 0) \quad \text{c'est-à-dire} \quad \begin{cases} x \quad + z = 0 \\ \quad y + z = 0 \\ x \quad + z = 0 \end{cases}$$

système qui admet la solution générale $(x, y, z) = z(-1, -1, 1)$, $z \in \mathbb{R}$. On voit donc que le noyau est une droite vectorielle engendrée par le vecteur $\boldsymbol{u} = (-1, -1, 1)$.

Pour l'image on calcule $f(\boldsymbol{e}_1) = (1, 0, 1)$, $f(\boldsymbol{e}_2) = (0, 1, 0)$ et $f(\boldsymbol{e}_3) = (1, 1, 1)$. Puis on cherche le rang de cette famille ; on a

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \quad \rightsquigarrow \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \rightsquigarrow \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Donc le rang de f vaut 2.

On en déduit que f n'est ni injective, ni surjective et donc non bijective.

Remarque. On a $\dim(\ker(f)) + \dim(\mathrm{im}(f)) = 1 + 2 = 3 = \dim E$. Cette relation n'est pas un hasard et on verra bientôt pourquoi.

Exemple. Soit $\varphi : \mathbb{R}_2[X] \to \mathbb{R}_2[X]$ définie par $P \overset{\varphi}{\mapsto} XP'$. Pour déterminer le noyau, on a

$$P \in \ker(\varphi) \quad \text{ssi} \quad XP' = 0 \quad \text{ssi} \quad P' = 0 \quad \text{ssi} \quad P = \text{cste}\,.$$

Donc $\ker(\varphi) = \mathbb{R}$ et $\dim(\ker(\varphi)) = 1$. Pour l'image, on a

$$\mathrm{im}(f) = \mathrm{Vect}(f(1), f(X), f(X^2)) = \mathrm{Vect}(0, X, 2X^2) = \mathrm{Vect}(X, X^2).$$

En particulier, $\dim(\mathrm{im}(\varphi)) = 2$. À nouveau, $\dim(\ker(\varphi)) + \dim(\mathrm{im}(\varphi)) = 1 + 2 = \dim \mathbb{R}_2[X]$.

5.3. Théorème du rang

Ce résultat très important lie la dimension du noyau et celle de l'image, relation auquelle faisaient référence les deux exemples précédents.

Définition 5.8. Soit $\varphi : E \to F$ une application linéaire avec F de dimension fini. On appelle le rang de φ, noté $\mathrm{rank}(\varphi)$, la dimension de l'image de φ.

Théorème 5.9 (du rang)**.** *Soit $f : E \to F$ une application linéaire. Si E est de dimension finie, alors*

$$\dim(\ker(f)) + rank(\varphi) = \dim(E).$$

Démonstration. Soit $(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_p)$ une base de $\ker(f)$; on complète cette base (via le théorème de la base incomplète) en une base $(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_p, \boldsymbol{e}_{p+1}, \dots, \boldsymbol{e}_n)$ de E. On sait que l'on a

$$\mathrm{im}(f) = \mathrm{Vect}(f(\boldsymbol{e}_1), \dots, f(\boldsymbol{e}_n)) = \mathrm{Vect}(f(\boldsymbol{e}_{p+1}), \dots, f(\boldsymbol{e}_n)),$$

car les premiers vecteurs image sont nuls. Il suffit de montrer que $\{f(\boldsymbol{e}_{p+1}), \dots, f(\boldsymbol{e}_n)\}$ est une famille libre. Supposons qu'il existe $\lambda_{p+1}, \dots, \lambda_n \in \mathbb{R}$ tels que

$$\lambda_{p+1} f(\boldsymbol{e}_{p+1}) + \cdots + \lambda_n f(\boldsymbol{e}_n) = 0.$$

Alors par linéarité, on a $f(\lambda_{p+1}\boldsymbol{e}_{p+1} + \cdots + \lambda_n \boldsymbol{e}_n) = 0$, donc

$$\lambda_{p+1}\boldsymbol{e}_{p+1} + \cdots + \lambda_n \boldsymbol{e}_n \in \ker(f),$$

c'est-à-dire il existe $\lambda_1, \dots, \lambda_p$ tels que

$$\lambda_{p+1}\boldsymbol{e}_{p+1} + \cdots + \lambda_n \boldsymbol{e}_n = \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_p \boldsymbol{e}_p.$$

Or la famille $(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n)$ est libre, donc tous les λ_i sont nuls. Il s'ensuit que la famille $\{f(\boldsymbol{e}_{p+1}), \dots, f(\boldsymbol{e}_n)\}$ est une base de $\mathrm{im}(f)$, ce qui entraîne $\dim(\mathrm{im}(f)) = n - p$. Finalement,

$$\dim(\ker(f)) + \dim(\mathrm{im}(f)) = p + (n - p) = n = \dim E.$$

On obtient la formule du théorème, car d'après la définition, $\mathrm{rank}(f) = \dim(\mathrm{im}(f))$. □

Corollaire 5.10. *Si $f : E \to F$ est une application linéaire avec $\dim F = \dim E < +\infty$, alors f est injective si et seulement si f est surjective (donc si et seulement si f est bijective).*

Démonstration. D'après le théorème du rang,

$$\dim(\mathrm{im}(f)) = \dim(E) - \dim(\ker(f)) = \dim(F) - \dim(\ker(f)).$$

Comme f est injective si et seulement si $\ker(f) = \{\mathbf{0}\}$, c'est-à-dire si et seulement si $\dim(\ker(f)) = 0$, l'équivalence s'ensuit. (On rappelle que $\dim(\mathrm{im}(f)) = \dim(F)$ si et seulement si $\mathrm{im}(f) = F$, car $\mathrm{im}(f) \subset F$ est un sous-espace vectoriel.) □

Remarque. On retiendra que dans le cas où les espaces de départ et d'arrivée ont même dimension *finie*, il suffit de montrer l'injectivité pour montrer la bijectivité. La dimension finie est nécessaire dans le corollaire précédent ; par exemple, l'application linéaire $f : \mathbb{R}[X] \to \mathbb{R}[X]$ définie par $f(P) = XP$ est injective mais non surjective.

Voici d'autres conséquences immédiates du théorème 5.9 (du rang). En raisonnant comme dans le corollaire précédent,

si $f : E \to F$ est un isomorphisme, alors $\dim(E) = \dim(F)$
si $f : E \to F$ est injective, alors $\dim(E) \leq \dim(F)$
si $f : E \to F$ est surjective, alors $\dim(E) \geq \dim(F)$.

Vocabulaire et notations.

– L'ensemble des applications linéaires de E dans F est noté $\mathcal{L}(E, F)$ ou $\mathrm{Hom}(E, F)$.
– Une application linéaire bijective $E \to F$ s'appelle un *isomorphisme*. On dit alors que les espaces vectoriels E et F sont isomorphes.
– Une application linéaire d'un espace vectoriel dans lui-même s'appelle un *endomorphisme*. L'ensemble des endomorphismes de E est noté $\mathcal{L}(E)$ ou $\mathrm{End}(E)$.
– Un endomorphisme bijectif s'appelle un *automorphisme*.

5.4. Composition d'applications linéaires

On peut composer des applications linéaires si les espaces vectoriels se correspondent bien. On a le résultat facile à établir suivant :

Proposition 5.11. *Soient E, F, et G trois espaces vectoriels et soient $f : E \to F$ et $g : F \to G$ deux applications linéaires. Alors la fonction composée $g \circ f : E \to G$ est une application linéaire.*

Pour $f : E \to \mathbb{R}^p$ nous avons parlé des composantes de f, les formes linéaires $f_1, \ldots, f_p$. Chaque f_j est la composée de f avec la fonction coordonnée $y_j : \mathbb{R}^p \to \mathbb{R}$, c'est-à-dire

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[rr, bend right, "f_j"'] & \mathbb{R}^p \arrow[r, "y_j"] & \mathbb{R}. \end{tikzcd}$$

Un autre exemple est la composition de D avec lui-même, où $D : \mathbb{R}_n[X] \to \mathbb{R}_n[X]$ est l'opérateur de dérivation, $P \mapsto P'$. Alors $D^2 = D \circ D : P \mapsto P''$ est l'opérateur de $\mathbb{R}_n[X]$ dans $\mathbb{R}_n[X]$ nomé l'*opérateur de dérivation d'ordre deux* qui est donc linéaire.

5.5. Quelques applications linéaires classiques

Dans cette section on discutera quelques exemples "géométriques" d'applications linéaires, les homothéties, projections et symétries. Par la suite, E sera un espace vectoriel de dimension finie. (Cette hypothèse de finitude simplifiera beaucoup les arguments, mais elle n'est pas nécessaire pour introduire ces exemples.)

Définition (Homothétie)**.** Soit k un scalaire fixé. L'endomorphisme de E qui à $\boldsymbol{v}$ associe $k\boldsymbol{v}$ est appelé homothétie de rapport k.

On remarque que l'homothétie de rapport k est bijective si et seulement si $k \neq 0$. De plus, la composition de deux homothéties est encore une homothétie.

Avant d'introduire les projections et les symétries, nous devons discuter les sommes directes de sous-espaces vectoriels. Soient F et G deux sous-espaces vectoriels de E. On rappelle que le sous-espace vectoriel somme $F + G$ est défini par

$$F + G = \{\boldsymbol{v} + \boldsymbol{w} \mid \boldsymbol{v} \in F, \boldsymbol{w} \in \mathbb{G}\}.$$

On dit que la somme de F et G est *directe* si $F \cap G = \{\mathbf{0}\}$. Dans ce cas, on note la somme par $F \oplus G$ on dit que F et G sont en somme directe. Cette notion permet d'assurer l'unicité de la décomposition de tout vecteur de $F + G$. Plus précisément, on a :

Proposition 5.12. *Deux sous-espaces vectoriels F et G sont en somme directe, si et seulement si tout vecteur de $F + G$ se décompose de manière unique en la somme d'un vecteur de F et d'un vecteur de G.*

Démonstration. Soit $\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w} \in F + G$, avec $\boldsymbol{v} \in F$ et $\boldsymbol{w} \in G$. Si $\boldsymbol{u} = \boldsymbol{v}' + \boldsymbol{w}'$ est une autre décomposition de $\boldsymbol{u}$, alors

$$F \ni \boldsymbol{v} - \boldsymbol{v}' = \boldsymbol{w} - \boldsymbol{w}' \in G,$$

c'est-à-dire $\boldsymbol{v} - \boldsymbol{v}' \in F \cap G$. Par conséquent, si $F \cap G = \{\mathbf{0}\}$, alors l'écriture est unique, et si $F \cap G \neq \{\mathbf{0}\}$, alors l'écriture n'est pas unique, d'où l'équivalence. □

Définition. On dit que deux sous-espaces vectoriels F et G de E sont *supplémentaires* lorsque $F \cap G = \{\mathbf{0}\}$ et $F + G = E$, autrement dit lorsque $F \oplus G = E$.

On dira aussi que E est la somme directe de F et G. Dans la figure 7 on représente $\mathbb{R}^3$ comme la somme directe d'une droite vectorielle et d'un plan vectoriel.

Exemple. Deux droites vectorielles distinctes de $\mathbb{R}^2$ sont supplémentaires —pensez aux droites d'équation $x = 0$ et $y = 0$ respectivement. Tout vecteur de $\mathbb{R}^2$ admet une et une seule décomposition suivant les deux droites.

Proposition 5.13. *Si F et G sont des sous-espaces vectoriels de E qui est de dimension finie, alors F et G sont supplémentaires si et seulement si $F \cap G = \{\mathbf{0}\}$ et $\dim(F) + \dim(G) = \dim(E)$.*

Démonstration. Il suffit de démontrer que si $F \cap G = \{\mathbf{0}\}$ et $\dim(F) + \dim(G) = \dim(E)$, alors F et G sont supplémentaires. Donc, d'après la définition, il faut montrer que $F + G = E$ (car on sait que $F \cap G = \{\mathbf{0}\}$). Le sous-espace $F + G$ est de dimension $\dim(F) + \dim(G)$ car $F \cap G = \{\mathbf{0}\}$, d'où l'égalité. □

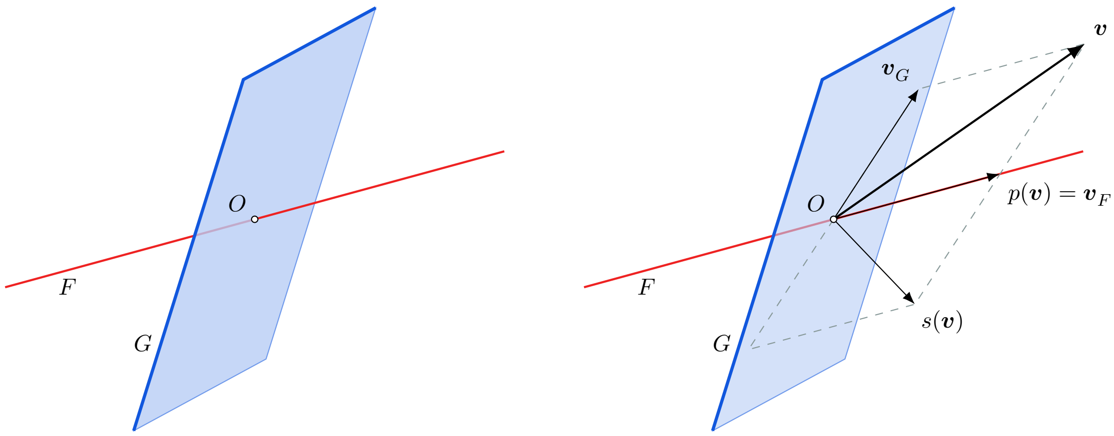

Figure 7: F est une droite vectorielle et G un plan vectoriel. On a $\mathbb{R}^3 = F \oplus G$. À droite, on représente la décomposition de $\boldsymbol{v}$ comme somme de $\boldsymbol{v}_F$ et $\boldsymbol{v}_G$ ainsi que les images de $\boldsymbol{v}$ par la projection sur F et la symétrie par rapport à F.

On peut maintenant introduire les projections et les symétries.

Définition. Soient $F, G \subset E$ deux sous-espaces vectoriels supplémentaires.

- La projection de E sur F parallèlement à G est l'endomorphisme $p = p_{F,G} : E \to E$ défini par
$$\boldsymbol{v} \mapsto \boldsymbol{v}_F,$$
où $\boldsymbol{v} = \boldsymbol{v}_F + \boldsymbol{v}_G$ est la décomposition unique de $\boldsymbol{v}$ dans la somme directe $F \oplus G = E$.
- La symétrie de E par rapport à F et parallèlement à G est l'endomorphisme $s = s_{F,G} : E \to E$ défini par
$$\boldsymbol{v} \mapsto \boldsymbol{v}_F - \boldsymbol{v}_G.$$

Proposition 5.14. *Soient $F, G \subset E$ deux sous-espaces vectoriels supplémentaires. Les endomorphismes $p = p_{F,G}$ et $s = s_{F,G}$ satisfont les propriétés suivantes :*

1) $p \circ p = p$, $\ker(p) = G$, $\operatorname{im}(p) = F$
2) $s \circ s = \mathrm{Id}_E$ *(c'est-à-dire s est involutive ; en particulier s est bijective et $s^{-1} = s$)*
3) $s = 2p - \mathrm{Id}_E$.

Démonstration. Tout au long de cette preuve on écrira, pour tout $\boldsymbol{v} \in E$, $\boldsymbol{v} = \boldsymbol{v}_F + \boldsymbol{v}_G$ pour la décomposition unique de $\boldsymbol{v}$, avec $\boldsymbol{v} \in F$ et $\boldsymbol{v}_G \in G$.

D'après les définitions de p et s on a pour tout $\boldsymbol{v} \in E$
$$(p \circ p)(\boldsymbol{v}) = p(p(\boldsymbol{v})) = p(\boldsymbol{v}_F) = \boldsymbol{v}_F = p(\boldsymbol{v})$$
et
$$(s \circ s)(\boldsymbol{v}) = s(s(\boldsymbol{v})) = s(\boldsymbol{v}_F - \boldsymbol{v}_G) = \boldsymbol{v}_F - (-\boldsymbol{v}_G) = \boldsymbol{v}_F + \boldsymbol{v}_G = \boldsymbol{v}.$$
À partir de ces identités les deux premiers points sont évidents. Pour voir le troisième, on calcule
$$s(\boldsymbol{v}) = \boldsymbol{v}_F - \boldsymbol{v}_G = \boldsymbol{v}_F - (\boldsymbol{v} - \boldsymbol{v}_F) = 2\boldsymbol{v}_F - \boldsymbol{v} = 2p(\boldsymbol{v}) - \boldsymbol{v}.$$
□

On notera que p n'est pas bijective sauf dans le cas très particulier où $F = E$, c'est-à-dire quand $p = Id_E$. Par contre, la symétrie est toujours bijective et son inverse est elle-même.

Exemple. On prend $E = \mathbb{R}^3$, $\mathcal{B} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ la base canonique et les sous-espaces supplémentaires $F = \mathrm{Vect}(\boldsymbol{e}_1, \boldsymbol{e}_2)$ et $G = \mathrm{Vect}(\boldsymbol{e}_3)$. Alors la projection et symétrie correspondantes sont

$$p : (x, y, z) \mapsto (x, y, 0) \qquad \text{et} \qquad s : (x, y, z) \mapsto (x, y, -z).$$

6. Matrices et applications linéaires

6.1. Matrices à coefficients dans un corps $\mathbb{K}$

Nous avons déjà manipulé des matrices dans l'exemple 1.8 et la proposition 1.9, puis lors de l'étude des sous-espaces vectoriels de $\mathbb{R}^n$ et de la résolution des systèmes linéaires. Dans ce chapitre nous allons associer une matrice à une application linéaire après avoir fixé des base pour le domaine et le codomaine de l'application. Les matrices associées aux applications linéaires représentent l'équivalent des matrices colonnes des coefficients associées aux vecteurs d'un espace vectoriel dès qu'une base de celui-ci est fixée.

On rappelle d'abord la notion de matrice ainsi que les opérations naturelles définies à l'aide des matrices.

Définition. Une matrice A de taille $n \times p$ à coefficients dans un corps $\mathbb{K}$ est un ensemble d'éléments[6] $\left(a_{i,j}\right)_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ de $\mathbb{K}$ que l'on représente par le tableau

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix}.$$

Les $a_{i,j}$ sont appelés les coefficients de la matrice A ayant n lignes et p colonnes. Les matrices

$$L_i(A) = \begin{pmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,p} \end{pmatrix} \quad \text{et} \quad C_j(A) = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{j,n} \end{pmatrix}$$

sont appelées la i-ème ligne de A et, respectivement, la j-ème colonne de A. Une matrice de taille $1 \times p$ est appelée une matrice ligne et une matrice de taille $n \times 1$ une matrice colonne.

On note $\mathcal{M}_{n\times p}(\mathbb{K})$ l'ensemble de toutes les matrice de taille $n \times p$ à coefficients dans $\mathbb{K}$. Lorsque $p = n$, on dit que la matrice est carrée d'ordre n. Cet ensemble a une structure naturelle d'espace vectoriel sur $\mathbb{K}$ à travers l'addition et la multiplication par un scalaire, opérations définies composante par composante. Explicitement, si $A = \left(a_{i,j}\right)_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$, $B = \left(b_{i,j}\right)_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ sont des des matrices de $\mathcal{M}_{n\times p}(\mathbb{K})$ et $\lambda \in \mathbb{K}$, alors

$$A + B = \left(a_{i,j} + b_{i,j}\right)_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} \in \mathcal{M}_{n\times p}(\mathbb{K})$$

[6] Les éléments $a_{i,j}$ apparaissent dans la littérature sous d'autres notations : par exemple $a_{i\,j}$ ou a^i_j. Dans cette dernière notation, l'indice de la ligne est en exposant !

et
$$\lambda A = \left(\lambda a_{i,j}\right)_{\substack{1\leq i\leq n\\1\leq j\leq p}} \in \mathcal{M}_{n\times p}(\mathbb{K}).$$

Exemple. Si $A = \left(\begin{smallmatrix}1 & 0\\ -2 & 1\end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix}2 & 3\\ 4 & -12\end{smallmatrix}\right)$, alors $A+B = \left(\begin{smallmatrix}3 & 3\\ 2 & -11\end{smallmatrix}\right)$ et $(-4)A = \left(\begin{smallmatrix}-4 & 0\\ 8 & -4\end{smallmatrix}\right)$.

Pour introduire et justifier le produit des matrices, nous allons considérer un système linéaire ayant n équations et p inconnues :
$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,p}x_p = b_1\\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,p}x_p = b_2\\ \qquad\qquad\qquad\qquad\vdots\\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,p}x_p = b_n. \end{cases}$$

En utilisant les matrices $A = (a_{i,j})$ de taille $b \times p$, $X = \begin{pmatrix} x_1\\ \vdots\\ x_p\end{pmatrix}$ de taille $p \times 1$ et $B = \begin{pmatrix} b_1\\ \vdots\\ b_n\end{pmatrix}$ de taille $n \times 1$, le système s'écrit sous forme matricielle comme
$$AX = B.$$

Cette écriture définit le produit matricielle AB d'une matrice quelconque A avec une matrice colonne X ayant le nombre de lignes égale au nombre de colonnes de A. Il fait sens de poser la définition suivante.

Définition 6.1. Soient $A = \left(a_{i,j}\right)_{\substack{1\leq i\leq m\\1\leq j\leq n}} \in \mathcal{M}_{m\times n}(\mathbb{K})$ et $B = \left(b_{k,l}\right)_{\substack{1\leq k\leq n\\1\leq l\leq p}} \in \mathcal{M}_{n\times p}(\mathbb{K})$, on définit le produit AB par $AB = \left(c_{i,l}\right)_{\substack{1\leq i\leq m\\1\leq l\leq m}} \in \mathcal{M}_{m\times p}(\mathbb{K})$ où
$$c_{i,l} = \sum_{j=1}^{n} a_{i,j}b_{j,l}.$$

Voyons de suite un exemple. On considère le produit d'une matrice de taille 3×2 avec une de taille 2×4 —le résultat est une matrice de taille 3×4 :
$$\begin{pmatrix} 2 & 1\\ 3 & -3\\ 1 & 4\end{pmatrix}\begin{pmatrix} 1 & 3 & 0 & -1\\ 3 & 3 & 1 & 0\end{pmatrix} = \begin{pmatrix} 5 & 9 & 1 & -2\\ -6 & 0 & -3 & -3\\ 13 & 15 & 4 & -1\end{pmatrix}.$$

En pratique pour calculer l'élément $c_{i,l}$ du produit $AB = (c_{i,l})$, on fait le produit (scalaire) de la ligne i de A par la colonne l de B.

Remarque. Le produit de matrices ne commute pas en général. Par exemple
$$\begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix} 0 & 1\\ 0 & 0\end{pmatrix} = \begin{pmatrix} 0 & 0\\ 0 & 1\end{pmatrix} \qquad \text{et} \qquad \begin{pmatrix} 0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix} = \begin{pmatrix} 1 & 0\\ 0 & 0\end{pmatrix}.$$

De plus, on peut avoir $AB = 0$ sans que $A = 0$ ou $B = 0$. Par exemple
$$\begin{pmatrix} 0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix} 0 & 1\\ 0 & 0\end{pmatrix} = \begin{pmatrix} 0 & 0\\ 0 & 0\end{pmatrix}.$$

Une telle matrice carrée A, telle que $A^q = 0$ pour un certain $q \in \mathbb{N}^*$ est dite *nilpotente.*

6.2. Représentation des applications linéaires par des matrices

Définition 6.2. Soient E et F deux espaces vectoriels de dimensions finies et soient $\mathcal{E} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ une base de E et $\mathcal{F} = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_p)$ une base de F. Si $\varphi : E \mapsto F$ est une application linéaire, on appelle la *matrice de φ dans les bases $\mathcal{E}$ et $\mathcal{F}$* la matrice de taille $p \times n$ dont la i-ème colonne représente les coordonnées de $\varphi(\boldsymbol{e}_i)$ dans la base $\mathcal{F}$, c'est-à-dire la i-ème colonne est égale à $C_{\mathcal{F}}(\varphi(\boldsymbol{e}_i))$.

Par la suite, on utilisera la notation $\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)$ pour désigner la matrice de l'application linéaire φ dans les bases fixées $\mathcal{E}$ pour le domaine et $\mathcal{F}$ pour le codomaine. Si, pour tout $1 \leq i \leq n$,

$$\varphi(\boldsymbol{e}_i) = a_{1,i}\boldsymbol{f}_1 + a_{2,i}\boldsymbol{f}_2 + \cdots + a_{p,i}\boldsymbol{f}_p \tag{6.1}$$

alors

$$\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{p,1} & a_{p,2} & \cdots & a_{p,n} \end{pmatrix}.$$

Quand les bases sont sous-entendues, ou l'accent ne portera pas sur les bases, on dira “soit A la matrice de φ dans les bases $\mathcal{E}$ et $\mathcal{F}$...”. De point de vue graphique, on essayera le plus souvant, de faire la distinction entre l'application linéaire et la matrice associée :

$$E \xrightarrow{\quad\varphi\quad} F \qquad \text{et} \qquad E, \mathcal{E} \xrightarrow{\quad A \quad} F, \mathcal{F}$$

Remarque. Une application linéaire donnée existe indépendamment des choix des bases pour le domaine ou le codomaine. Par contre, la représentation matricielle de l'application linéaire dépend fortement des bases choisies. Nous avons déjà rencontré ce phénomène quand nous avons vus que la matrice des coefficients d'un vecteur dépend de la base choisie.

Si φ est un endomorphisme $\varphi : E \to E$ et si $\mathcal{F} = \mathcal{E}$, c'est-à-dire si on fixe la même base $\mathcal{E}$ au départ et a l'arrivée, on appelle la matrice $\mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(\varphi)$ la matrice de φ dans la base $\mathcal{E}$.

Exemple. Soit $D : \mathbb{R}_3[X] \to \mathbb{R}_3[X]$ l'application linéaire définie par $P \mapsto P'$. On fixe la base canonique $\mathcal{X} = (1, X, X^2, X^3)$ pour les deux espaces vectoriels —de départ et d'arrivée. Alors la matrice de D dans la base $\mathcal{X}$ est

$$\mathrm{Mat}_{\mathcal{X}}^{\mathcal{X}}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

car $D(1) = 0$, $D(X) = 1$, $D(X^2) = 2X$ et $D(X^3) = 3X^2$. Mais si pour le domaine de définition on choisit la base $\mathcal{B} = (1, (1-X), (1-X)^2, (1-X)^3)$ et pour le codomaine $\mathcal{X}$, alors

$$\mathrm{Mat}_{\mathcal{X}}^{\mathcal{B}}(D) = \begin{pmatrix} 0 & -1 & -2 & -3 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Exemple. Soit E un espace vectoriel de dimension finie et soit $\mathcal{B} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ une base de E. La matrice de l'application identité $Id_E : \boldsymbol{v} \mapsto \boldsymbol{v}$ dans la base $\mathcal{B}$ est la matrice identité, c'est-à-dire

$$\mathrm{Mat}_{\mathcal{B}}^{\mathcal{B}}(Id_E) = I_n = \begin{pmatrix} 1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & 1 \end{pmatrix}$$

puisque pour tout $1 \leq i \leq n$, on a $Id_E(\boldsymbol{e}_i) = \boldsymbol{e}_i$.

Exemple. Soient $E = F \oplus G$, c'est-à-dire F et G sont deux sous-espaces supplémentaires de E, et soit $\mathcal{B}$ une base de E dont les premiers vecteurs forment une base de F et les suivants une base de G. Si $p_{F,G}$ et $s_{F,G}$ sont la projection sur F parallèlement à G et, respectivement la symétrie par rapport à F parallèlement à G, alors

$$\mathrm{Mat}_{\mathcal{B}}^{\mathcal{B}}(p_{F,G}) = \begin{pmatrix} I_f & \\ & 0_g \end{pmatrix} \quad \text{et} \quad \mathrm{Mat}_{\mathcal{B}}^{\mathcal{B}}(s_{F,G}) = \begin{pmatrix} I_f & \\ & -I_g \end{pmatrix},$$

où $f = \dim(F)$ et $g = \dim(G)$. Remarquons la relation matricielle

$$2\begin{pmatrix} I_f & \\ & 0_g \end{pmatrix} - I_{f+g} = \begin{pmatrix} I_f & \\ & -I_g \end{pmatrix}$$

qui traduit matriciellement la relation $2p_{F,G} - Id_E = s_{F,G}$ établie dans la proposition 5.14 3).

Le choix de la place des indices de $a_{k,i}$ dans l'expression de $\varphi(\boldsymbol{e}_i)$ dans la base $\mathcal{F}$, voir (6.1), est justifiée par le résultat suivant et par le comportement des matrices associées lors de la composition des applications linéaires.

Matrice associée et coordonnées

Proposition 6.3. *Soit $\varphi : E \to F$ une application linéaire et soient $\mathcal{E} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ une base de E et $\mathcal{F} = (\boldsymbol{f}_1, \dots, \boldsymbol{f}_p)$ une base de F. Pour tout $\boldsymbol{v} \in E$,*

$$C_{\mathcal{F}}(\varphi(\boldsymbol{v})) = \mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)\, C_{\mathcal{E}}(\boldsymbol{v}),$$

où $C_{\mathcal{E}}(\boldsymbol{v})$ est la matrice colonne des coefficients de $\boldsymbol{v}$ dans la base $\mathcal{E}$.

Démonstration. On suppose que $n = \dim(E)$ et que $p = \dim(F)$. Soit $\boldsymbol{v} \in E$ et soient

$$\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}} = (a_{i,j})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}, \quad X = C_{\mathcal{E}}(\boldsymbol{v}) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad Y = C_{\mathcal{F}}(\varphi(\boldsymbol{v})) = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}.$$

Alors

$$\sum_{k=1}^{p} y_k \boldsymbol{f}_k = \varphi(\boldsymbol{v}) = \varphi\Big(\sum_{j=1}^{n} x_j \boldsymbol{e}_j\Big) = \sum_{j=1}^{n} x_j \varphi(\boldsymbol{e}_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{n} a_{i,j} \boldsymbol{f}_i = \sum_{i=1}^{n} \sum_{j=1}^{n} x_j a_{i,j} \boldsymbol{f}_i.$$

Donc

$$\sum_{k=1}^{p} y_k \boldsymbol{f}_k = \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} x_j a_{i,j}\Big) \boldsymbol{f}_i.$$

Comme les $\boldsymbol{f}_k$ forment une base de F, on conclut que

$$y_k = \sum_{j=1}^{n} x_j a_{i,j} = \sum_{j=1}^{n} a_{i,j} x_j$$

pour tout $1 \leq k \leq p$. Donc $Y = (a_{i,j})X$. $\square$

Remarque. On a choisit la notation $\mathrm{Mat}^{\mathcal{E}}_{\mathcal{F}}(\varphi)$ pour la matrice de l'endomorphisme pour avoir dans la relation de la proposition la règle mnémotechnique

$$C_{\mathcal{F}}(\varphi(\boldsymbol{v})) = \mathrm{Mat}^{\cancel{\mathcal{E}}}_{\mathcal{F}}(\varphi)\, C_{\cancel{\mathcal{E}}}(\boldsymbol{v}).$$

On retrouvera cette règle surtout dans le theoreme 6.7.

Exemple. Pour la dérivée formelle considérée dans un des exemples précédents, $D : \mathbb{R}_3[X] \to \mathbb{R}_3[X]$ avec $D : P \mapsto P'$, si on prend $P = a_0 + a_1 X + a_2 X^2 + a_3 X^3$, alors, en appliquant la proposition 6.3, les coefficients de $D(P)$ sont donnés par le produit

$$\mathrm{Mat}^{\mathcal{X}}_{\mathcal{X}}(D) \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ 0 \end{pmatrix}.$$

Ceci signifie que $P' = D(P) = a_1 + 2a_2 X + 3a_3 X^3$, ce qui est correct.

Exemple. Soit $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ l'application linéaire définie par $(x, y, z) \mapsto (x + y + 3z, x - y + z, 2y)$, où (x, y, z) sont les coordonnées correspondant à la base canonique $\mathcal{E} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$. Alors, d'après la proposition 6.3,

$$\mathrm{Mat}^{\mathcal{B}}_{\mathcal{B}}(\varphi) = \begin{pmatrix} 1 & 1 & 3 \\ 1 & -1 & 1 \\ 0 & 2 & 0 \end{pmatrix}.$$

On peut vérifier la matrice en calculant les images des vecteurs de la base :

$$\begin{aligned} \varphi(\boldsymbol{e}_1) &= (1, 1, 0) = \boldsymbol{e}_1 + \boldsymbol{e}_2 \\ \varphi(\boldsymbol{e}_2) &= (1, -1, 2) = \boldsymbol{e}_1 - \boldsymbol{e}_2 + 2\boldsymbol{e}_3 \\ \varphi(\boldsymbol{e}_3) &= (3, 1, 0) = 3\boldsymbol{e}_1 + \boldsymbol{e}_2. \end{aligned}$$

Donc si une application linéaire est décrite en utilisant *les coordonnées* (associées à certaines bases), alors les coefficients qui apparaissent dans cette définition explicite en coordonnées sont *les lignes de la matrice.*

Pour détailler dans le cas général ce que nous venons de remarquer dans l'exemple préc'edent, soient $A = \mathrm{Mat}^{\mathcal{E}}_{\mathcal{F}}(\varphi) = (a_{i,j})$, $X = C_{\mathcal{E}}(\boldsymbol{v}) = (x_j)$ et $Y = C_{\mathcal{F}}(\varphi(\boldsymbol{v})) = (y_i)$ —avec les deux dernières des matrices colonnes. Alors

$$y_i = a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n$$

pour tout $1 \leq i \leq p$, c'est-à-dire $Y = AX$. Symboliquement, en interprétant les bases comme des "matrices" ligne, la relation (6.1) peut s'écrire comme $\varphi(\mathcal{E}) = \mathcal{F}\, A$. Donc on a les identités

$$Y = AX \qquad \text{et} \qquad \varphi(\mathcal{E}) = \mathcal{F}\, A.$$

Les coordonnées et les bases doivent aller dans des sens opposés.

Matrices associées et composition des applications linéaires

Proposition 6.4. *Soient E, F et G trois espaces vectoriels de bases respectives $\mathcal{E}$, $\mathcal{F}$ et $\mathcal{G}$. Si $\varphi \in \mathcal{L}(E,F)$ et $\psi \in \mathcal{L}(F,G)$, alors*

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{E}}(\varphi \circ \varphi) = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{F}}(\psi)\ \mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi).$$

On dit que la composition des applications linéaires se traduit par le produit des matrices correspondantes. Les diagrammes représentant l'énoncé sont les suivants.

$$\begin{tikzcd} E \arrow[r, "\varphi"] \arrow[rr, bend right, "\varphi\circ\varphi"'] & F \arrow[r, "\psi"] & G \end{tikzcd}$$

$$\begin{tikzcd} E, \mathcal{E} \arrow[r, "\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)"] \arrow[rr, bend right, "\substack{\mathrm{Mat}_{\mathcal{G}}^{\mathcal{E}}(\varphi\circ\varphi) \\ \mathrm{Mat}_{\mathcal{G}}^{\mathcal{F}}(\psi)\ \mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)}"'] & F, \mathcal{F} \arrow[r, "\mathrm{Mat}_{\mathcal{G}}^{\mathcal{F}}(\psi)"] & G, \mathcal{G} \end{tikzcd}$$

Démonstration. Les diagrammes suggèrent une justification de l'égalité de l'énoncé en utlisant les coordonnées et la proposition 6.3 ; si $\boldsymbol{u} \in E$ est un vecteur quelconque, alors

$$C_{\mathcal{G}}(\psi \circ \varphi(\boldsymbol{u})) = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{E}}(\psi \circ \varphi)\, C_{\mathcal{E}}(\boldsymbol{u})$$

et en même temps

$$C_{\mathcal{G}}(\psi \circ \varphi(\boldsymbol{u})) = C_{\mathcal{G}}(\psi(\varphi(\boldsymbol{u}))) = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{F}}(\psi)\, C_{\mathcal{F}}(\varphi(\boldsymbol{u})) = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{F}}(\psi)\ \mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)\, C_{\mathcal{E}}(\boldsymbol{u}).$$

Donc

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{E}}(\psi \circ \varphi)\, C_{\mathcal{E}}(\boldsymbol{u}) = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{F}}(\psi)\ \mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)\, C_{\mathcal{E}}(\boldsymbol{u}),$$

c'est-à-dire

$$\left[\mathrm{Mat}_{\mathcal{G}}^{\mathcal{E}}(\psi \circ \varphi) - \mathrm{Mat}_{\mathcal{G}}^{\mathcal{F}}(\psi)\ \mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)\right] C_{\mathcal{E}}(\boldsymbol{u}) = 0.$$

Comme $\boldsymbol{u}$ est quelconque, il s'ensuit que le premier facteur est nul. □

Le résultat principal de cette section est le suivant ; il nous dit qu'il existe un choix de bases pour le domaine et le codomaine d'une application linéaire tel que la matrice associée soit très simple. Plusieures questions vont se poser naturellement à la suite de ce résultat.

Théorème 6.5. *Soit $\varphi : E \to F$ une application linéaire avec $n = \dim(E)$ et $p = \dim(F)$. Si $r = \mathrm{rank}(\varphi) \overset{déf}{=} \dim(\mathrm{im}(\varphi))$, alors il existe $\mathcal{E}$ une base de E et $\mathcal{F}$ une base de F telle que*

$$\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi) = \begin{pmatrix} I_r & \cdots \\ \cdots & 0_{p-r,n-r} \end{pmatrix}.$$

Démonstration. Soit $(f_1, \ldots, f_r)$ une base de $\mathrm{im}(\varphi) \subset F$. On prolonge cette famille libre de F à une base de F, $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_r, \boldsymbol{f}_{r+1}, \ldots, \boldsymbol{f}_p)$. Pour tout $1 \leq j \leq r$, on choisit $\boldsymbol{e}_j \in E$ tel que

$$\varphi(\boldsymbol{e}_j) = \boldsymbol{f}_j \in \mathrm{im}(\varphi).$$

La famille $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r\}$ ainsi construite est libre — sinon une combinaison nulle non triviale serait envoyée par φ en une combinaison des $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_r$ nulle et non triviale.

Soit $(\boldsymbol{e}_{r+1}, \ldots, \boldsymbol{e}_n)$ une base de $\ker(\varphi) \subset E$. Pour démontrer le théorème il suffit de montrer que $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ est une base de E, c'est-à-dire de montrer qu'elle est libre et engendre E.

Pour voir que la famille est libre, soient $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tels que $\sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i = \mathbf{0}$. Alors

$$\alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_r \boldsymbol{e}_r = -\alpha_{r+1} \boldsymbol{e}_{r+1} - \cdots - \alpha_n \boldsymbol{e}_n$$

et, en appliquant φ, on obtient

$$\alpha_1 \boldsymbol{f}_1 + \cdots + \alpha_r \boldsymbol{f}_r = \varphi(\alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_r \boldsymbol{e}_r) = \varphi(-\alpha_{r+1} \boldsymbol{e}_{r+1} - \cdots - \alpha_n \boldsymbol{e}_n) = \mathbf{0}.$$

Donc $\alpha_1 = \cdots = \alpha_r = 0$. Il s'ensuit que $\mathbf{0} = -\alpha_{r+1} \boldsymbol{e}_{r+1} - \cdots - \alpha_n \boldsymbol{e}_n$ et donc que tous les α_i sont nuls.

Pour voir que la famille est génératrice, soit $\boldsymbol{v} \in E$. Alors, il existe $\beta_1, \ldots, \beta_r \in \mathbb{K}$ tels que

$$\varphi(\boldsymbol{v}) = \beta_1 \boldsymbol{f}_1 + \cdots + \beta_r \boldsymbol{f}_r = \beta_1 \varphi(\boldsymbol{e}_1) + \cdots + \beta_r \varphi(\boldsymbol{e}_r),$$

d'après le choix des $\boldsymbol{e}_i$, $i \leq r$. Alors $\boldsymbol{v} - \sum_{i=1}^{r} \beta_i \boldsymbol{e}_i \in \ker(\varphi)$. □

Exemple (l'homothétie)**.** Si $h : E \to E$ est une homothétie de rapport k, alors pour tout vecteur $\boldsymbol{v} \in E$ on a $h(\boldsymbol{v}) = k\boldsymbol{v}$ et en particulier pour les vecteurs d'une base. Donc, quelle que soit la base $\mathcal{B}$, on a $\mathrm{Mat}_{\mathcal{B}}^{\mathcal{B}}(h) = kI_n$, où $n = \dim(E)$.

D'abord, on notera que la matrice est kI_n pour toute base de E —ceci est faux en général, la matrice d'un endomorphisme variant avec la base chosie. Par la suite, en appliquant le théorème 6.5, on obtient la matrice I_n si les bases choisies pour le domaine et le codomaine sont différentes. Par exemple, si $\mathcal{B} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ et $\mathcal{B}' = (k\boldsymbol{e}_1, \ldots, k\boldsymbol{e}_n)$, alors

$$\mathrm{Mat}_{\mathcal{B}'}^{\mathcal{B}}(h) = I_n.$$

Cet exemple montre que **pour un endomorphisme**, le fait d'avoir des choix indépendants pour les deux bases permet d'obtenir une matrice très simple (voir le théorème 6.5), mais, qu'en même temps, toute signification géométrique de l'endomorphisme est perdue. Plusieures questions se posent naturellement :

- Si $\varphi : E \to E$ est un endomorphisme avec $\dim(E)$ finie, existe-t-il une base $\mathcal{E}$ telle que $\mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(\varphi)$ soit très simple ? Que voudrait dire simple dans ce contexte ?
- Comment est modifiée la matrice d'une application linéaire quand on change une ou les deux bases auxquelles cette matrice est associée ?

La réponse à la deuxième question sera donnée par la suite. La réponse à la première quesiton occupera le dernier chapitre de notre cours. Le fait d'avoir moins de degrés de liberté dans le choix des bases nous suggère que la matrice recherchée ne va plus avoir la simplicité qui apparait dans le théorème 6.5, mais elle recélera plus d'informations géométriques.

6.3. Changement de base pour les applications linéaires

Soit E un espace vectoriel de dimension finie et soient $\mathcal{E}$ et $\mathcal{E}'$ deux bases de E. En exprimant les vectuers de la base $\mathcal{E}'$ en fonction de la base $\mathcal{E}$ on obtient une matrice carrée nomée la *matrice de passage de $\mathcal{E}$ vers $\mathcal{E}'$*. C'est la matrice de décomposition de la famille $\mathcal{E}'$ dans la base $\mathcal{E}$, c'est-à-dire la matrice dont les colonnes sont les coordonnées des vecteurs de la nouvelle base $\mathcal{E}'$ dans l'ancienne base $\mathcal{E}$. Il s'ensuit que cette matrice

$$(p_{i,j}) = \begin{pmatrix} C_{\mathcal{E}}(\boldsymbol{e}'_1) & C_{\mathcal{E}}(\boldsymbol{e}'_2) & \cdots & C_{\mathcal{E}}(\boldsymbol{e}'_n) \end{pmatrix}$$

est la matrice de l'application Id_E dans les bases $\mathcal{E}'$ et $\mathcal{E}$. L'ordre des bases ici est très important ; on a

$$Id_E(\boldsymbol{e}'_j) = \boldsymbol{e}'_j = \sum_{i=1}^{n} p_{i,j}\boldsymbol{e}_i.$$

Définition 6.6. Si $\mathcal{E}$ et $\mathcal{E}'$ sont deux bases de E, espace vectoriel de dimension n, alors la matrice de passage de $\mathcal{E}$ vers $\mathcal{E}'$, ou encore matrice de changement de base de $\mathcal{E}$ vers $\mathcal{E}'$ est

$$\mathrm{Pass}_{\mathcal{E}}^{\mathcal{E}'} = \mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}'}(Id_E) \in M_n(\mathbb{K}).$$

Si $X = C_{\mathcal{E}}(\boldsymbol{v})$ et $X' = C_{\mathcal{E}'}(\boldsymbol{v})$ sont les matrices colonnes des coordonnées d'un vecteur quelconque dans les bases $\mathcal{E}$ et respectivement $\mathcal{E}'$, alors, d'après la proposition 6.3,

$$X = \mathrm{Pass}_{\mathcal{E}}^{\mathcal{E}'}\, X'.$$

On peut maintenant donner la formule du changement de bases pour les applications liéaires.

Théorème 6.7. *Soient E, F deux espaces vectoriels de dimensions finies et soit $\varphi : E \to F$ une application linéaire. Si $\mathcal{E}$ et $\mathcal{E}'$ sont deux bases de E, et $\mathcal{F}$ et $\mathcal{F}'$ deux bases de F, alors*

$$\mathrm{Mat}_{\mathcal{F}'}^{\mathcal{E}'}(\varphi) = \mathrm{Pass}_{\mathcal{F}'}^{\mathcal{F}}\, \mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)\, \mathrm{Pass}_{\mathcal{E}}^{\mathcal{E}'}.$$

Les diagrammes correspondants (pour les applications et pour les matrices) sont les suivants.

$$\begin{tikzcd} E \arrow[r, "\varphi"] & F \arrow[d, "\mathrm{Id}_F"] \\ E \arrow[u, "\mathrm{Id}_E"] \arrow[r, "\varphi"] & F \end{tikzcd} \qquad \begin{tikzcd} E, \mathcal{E} \arrow[r, "\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)"] & F, \mathcal{F} \arrow[d, "\mathrm{Pass}_{\mathcal{F}'}^{\mathcal{F}}"] \\ E, \mathcal{E}' \arrow[u, "\mathrm{Pass}_{\mathcal{E}}^{\mathcal{E}'}"] \arrow[r, "\mathrm{Mat}_{\mathcal{F}'}^{\mathcal{E}'}(\varphi)"] & F, \mathcal{F}' \end{tikzcd}$$

Démonstration. La preuve est un corrolaire de la proposition 6.4 et de la définition de la matrice de passage en utlisant le deuxième diagramme ci-dessus. □

Avant de poursuivre avec l'étude des matrices associées à une application linéaire, il faut comprendre plus en détail les matrices de passage. Par exemple, si $\varphi : E \to E$ est un endomorphisme et $\mathcal{E}$ et $\mathcal{E}'$ deux bases de E, alors, d'après le théorème 6.7, on a

$$M' = \mathrm{Pass}_{\mathcal{E}'}^{\mathcal{E}}\, M\, \mathrm{Pass}_{\mathcal{E}}^{\mathcal{E}'},$$

où M est la matrice de φ dans la base $\mathcal{E}$ et M' celle dans la base $\mathcal{E}'$. Les matrices de passages qui apparaissent vérifient la relation

$$\mathrm{Pass}_{\mathcal{E}}^{\mathcal{E}'}\, \mathrm{Pass}_{\mathcal{E}'}^{\mathcal{E}} = I_n \tag{$\sharp$}$$

car (pour la deuxième égalité on utilise la proposition 6.4)

$$\mathrm{Pass}_{\mathcal{E}}^{\mathcal{E}'}\, \mathrm{Pass}_{\mathcal{E}'}^{\mathcal{E}} = \mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}'}(Id_E)\, \mathrm{Mat}_{\mathcal{E}'}^{\mathcal{E}}(Id_E) = \mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(Id_E \circ Id_E) = \mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(Id_E) = I_n.$$

Les matrices vérifiant la relation ($\sharp$) sont dites *inversibles*.

6.4. Matrices inversibles

Définition. Une matrice A carrée de $\mathcal{M}_n(\mathbb{K})$ est dite *inversible* lorsqu'il existe une matrice carrée A' telle que $AA' = A'A = I_n$. Dans ce cas la matrice A' s'appelle l'inverse de A et se note A^{-1}.

Exemple. Si $A = \left(\begin{smallmatrix} 0 & 1 \\ 2 & 0 \end{smallmatrix}\right)$, alors A est inversible et $A^{-1} = \begin{pmatrix} 0 & 1/2 \\ 1 & 0 \end{pmatrix}$ car $AA' = I_2 = A'A$.

Précédemment, on a vu que les matrices de passage sont inversibles. En outre, il existe un lien important entre matrices inversibles et endomorphismes.

Théorème 6.8. *Soit $f : E \to E$ un endomorphisme d'un espace vectoriel muni d'une base $\mathcal{E}$ et soit $M = \mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(f)$ la matrice de f dans la base $\mathcal{E}$. Les affirmations suivantes sont vraies :*

(i) *si f est bijective alors l'application inverse f^{-1} est linéaire*

(ii) *f est bijective si et seulement si M est inversible*

(iii) *si f est bijective alors $\mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(f^{-1}) = M^{-1}$.*

Démonstration. Pour (i), comme f est bijective, notons f^{-1} l'application inverse (ou réciproque). Soient $\boldsymbol{u}, \boldsymbol{v} \in E$ et soient $\alpha, \beta \in \mathbb{K}$. On a

$$f(\alpha f^{-1}(\boldsymbol{u}) + \beta f^{-1}(\boldsymbol{v})) = \alpha\, f(f^{-1}\boldsymbol{u}) + \beta\, f(f^{-1}\boldsymbol{v}) = \alpha\boldsymbol{u} + \beta\boldsymbol{v}$$

et en appliquant f^{-1} on arrive à

$$\alpha f^{-1}(\boldsymbol{u}) + \beta f^{-1}(\boldsymbol{v}) = f^{-1}(\alpha\boldsymbol{u} + \beta\boldsymbol{v}).$$

Pour (ii), si f est bijective, alors on note par M' la matrice de f^{-1} dans la base $\mathcal{E}$; puisque l'on a $f \circ f^{-1} = f^{-1} \circ f = Id_E$, d'après la proposition 6.4, en passant aux matrices, cela devient $MM' = M'M = I_n$, c'est-à-dire M est inversible. Réciproquement, si M est inversible, alors il existe la matrice M^{-1} telle que

$$MM^{-1} = M^{-1}M = I_n.$$

Soit g l'application linéaire pour laquelle $\mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(g) = M^{-1}$. D'après la proposition 6.4,

$$\mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(f \circ g) = \mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(f)\, \mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(g) = M\, M^{-1} = I_n.$$

Donc $f \circ g = Id_E$ et simmilairement $g \circ f = Id_V$, c'est-à-dire f est inversible et $g = f^{-1}$.

Le point (iii) résulte immédiatement du (ii). □

Calcul de l'inverse d'une matrice. Il existe plusieures méthodes pour le calcul explicite de l'inverse d'une matrice. En voici une première méthode algorithmique basé sur la méthode du pivot de Gauss. Soit A une matrice carrée inversible. Sachant que $AX = Y$ si et seulement si $X = A^{-1}Y$, on inverse le système $AX = Y$ pour obtenir X en fonction de Y et, ainsi, on en déduit A^{-1}.

Exemple. On veut inverser la matrice $A = \left(\begin{smallmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{smallmatrix}\right)$. On résout par la méthode de Gauss le système $AX = Y$. On obteint successivement,

$$\begin{cases} \phantom{x_1 +{}} x_2 + x_3 = y_1 \\ x_1 \phantom{{}+x_2} + x_3 = y_2 \\ x_1 + x_2 \phantom{{}+x_3} = y_3 \end{cases} \qquad \begin{cases} x_1 \phantom{{}+x_2} + x_3 = y_2 \\ x_2 - x_3 = -y_2 + y_3 \\ x_2 + x_3 = y_1 \end{cases} \qquad \begin{cases} x_1 \phantom{{}+x_2} + x_3 = y_2 \\ x_2 - x_3 = -y_2 + y_3 \\ 2x_3 = y_1 + y_2 - y_3 \end{cases}$$

c'est-à-dire

$$\begin{cases} x_1 \quad = \frac{1}{2}(-y_1 + y_2 + y_3) \\ \quad x_2 \quad = \frac{1}{2}(y_1 - y_2 + y_3) \\ \qquad x_3 = \frac{1}{2}(y_1 + y_2 - y_3). \end{cases}$$

On en déduit

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

Ayant introduit la notion de matrice inversible (pour les matrices carrées), revenons maintenant au théorème 6.7 quand l'application linéaire considérée est un endomorphisme.

Corollaire 6.9. *Soit $f : E \to E$ un endomorphisme de l'espace vectoriel de dimension finie E. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E. Si M_f et M'_f sont les matrices de f dans les bases $\mathcal{E}$ et $\mathbf{e}'$ respectivement, et si P est la matrice de passage de $\mathcal{E}$ vers $\mathcal{E}'$, alors*

$$M'_f = P^{-1} M_f P.$$

Démonstration. D'après le théorème 6.7,

$$\mathrm{Mat}_{\mathcal{E}'}^{\mathcal{E}'}(f) = \mathrm{Pass}_{\mathcal{E}'}^{\mathcal{E}}\ \mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(f)\ \mathrm{Pass}_{\mathcal{E}}^{\mathcal{E}'}.$$

En utilisant les hypothèses, $M_f = \mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}}(f)$, $M'_f = \mathrm{Mat}_{\mathcal{E}'}^{\mathcal{E}'}(f)$, $P = \mathrm{Pass}_{\mathcal{E}}^{\mathcal{E}'}$ et $\mathrm{Pass}_{\mathcal{E}'}^{\mathcal{E}} = (\mathrm{Pass}_{\mathcal{E}}^{\mathcal{E}'})^{-1} = P^{-1}$. □

On sait maintenant comment change la matrice d'un endomorphisme quand on passe d'une base à une autre. Cette relation entre les différentes matrices appelle la définition suivante.

Définition 6.10 (matrices semblables)**.** Les matrices carrées $A, B \in \mathcal{M}_n(\mathbb{K})$ sont dites *semblables* s'il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Ainsi, pour un endomorphisme fixé f, les matrices associées dans différentes bases sont toutes semblables ; réciproquement, si deux matrices sont semblables et une est la matrice de f, alors l'autre l'est aussi.

Remarque 6.11. Parmi les fonctions définies sur $\mathcal{M}_n(\mathbb{K})$ (à valeurs dans $\mathbb{K}$), celles dont la valeur ne varie pas quand on remplace une matrice par une matrice semblable remplissent un rôle spécial : à travers elles il est possible de construire des *invariants* des endomorphismes. Si $\alpha : \mathcal{M}_N(\mathbb{K}) \to \mathbb{K}$ est une telle fonction, alors en définissant

$$\alpha(f) = \alpha(M_f),$$

on obtient un résultat qui dpend seulement de f et non pas de la base dans laquelle la matrice M_f est écrite.

Un des invariants importants d'un endomorphisme est la trace.

Exemple 6.12. (trace) Si $f : E \to E$ est un endomorphisme, la trace de f est définie par

$$\operatorname{tr}(f) = \operatorname{tr}(M_f) = \sum_{i=1}^{n} a_{i,i}$$

où $M_f = (a_{i,i})$ est la matrice de f dans une base $\mathcal{E}$. En particulier, la trace d'une matrice carrée A est la somme des éléments de la diagonale de A. Pour démontrer que la trace est constante sur les matrices semblables, on établit un résultat plus général : *si* $A \in \mathcal{M}_{n\times p}(\mathbb{K})$ *et* $B \in \mathcal{M}_{p\times n}(\mathbb{K})$*, alors*

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

En effet,

$$\operatorname{tr}(AB) = \sum_{i=1}^{n}\left(\sum_{j=1}^{p} a_{i,j}b_{j,i}\right) = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{i,j}b_{j,i} = \sum_{j=1}^{p}\sum_{i=1}^{n} b_{j,i}a_{i,j} = \sum_{j=1}^{p}\left(\sum_{i=1}^{n} b_{j,i}a_{i,j}\right) = \operatorname{tr}(BA).$$

Alors

$$\operatorname{tr}(P^{-1}M_fP) = \operatorname{tr}\left(P^{-1}(M_fP)\right) = \operatorname{tr}\left((M_fP)P^{-1}\right) = \operatorname{tr}(M_fPP^{-1}) = \operatorname{tr}(M_f).$$

6.5. Opérations dans $\mathcal{L}(E, F)$ et traduction en terme de matrices

Soient E et F deux espaces vectoriels sur $\mathbb{K}$. Pour $\varphi, \psi \in \mathcal{L}(E, F)$ et pour $\lambda \in \mathbb{K}$ on définit $\lambda\varphi$ et $\varphi + \psi$ par

$$(\lambda\varphi)(\boldsymbol{v}) = \lambda\varphi(\boldsymbol{v}) \qquad \text{et} \qquad (\varphi + \psi)(\boldsymbol{v}) = \varphi(\boldsymbol{v}) + \psi(\boldsymbol{v})$$

pour tout $\boldsymbol{v} \in E$. On vérifie aisément que $\lambda\varphi$ et $\varphi + \psi$ sont des applications linéaires, donc $\mathcal{L}(E, F)$ est un $\mathbb{K}$-espace vectoriel.

Par la suite on veut voir que si E et F sont de dimensions finies, n et p respectivement, la structure de cet espace vectoriel est celle de $\mathcal{M}_{p\times n}(\mathbb{K})$ dès que des bases sont fixées pour E et F.

Pour commencer, on rappelle que $\mathcal{M}_{p\times n}(\mathbb{K})$, l'ensemble des matrices de taille fixée à coefficients dans $\mathbb{K}$, est un espace vectoriel sur $\mathbb{K}$ (avec les opérations définies coefficient par coefficient. Si, pour tous $1 \le i \le p$ et $1 \le j \le n$, on définit la *matrice élémentaire* $E_{i,j}$ comme étant la matrice dont tous les coefficients sont nuls sauf celui d'indices (i, j) qui vaut 1, alors $(E_{1,1}, E_{1,2}, \dots, E_{p,n})$ est une base de $\mathcal{M}_{p\times n}(\mathbb{K})$. En particulier

$$\dim(\mathcal{M}_{p\times n}(\mathbb{K})) = np.$$

Pour finir, il suffit de voir que les opérations sur les matrices correspondent à des opérations sur les applications linéaires. Soient $\mathcal{E}$ et $\mathcal{F}$ des bases de E et F. Pour tout $\varphi \in \mathcal{L}(E, F)$, on a introduit la matrice $\operatorname{Mat}^{\mathcal{E}}_{\mathcal{F}}(\varphi) \in \mathcal{M}_{p\times n}(\mathbb{K})$. On a ainsi défini la fonction

$$\operatorname{Mat}^{\mathcal{E}}_{\mathcal{F}} : \mathcal{L}(E, F) \longrightarrow \mathcal{M}_{p\times n}(\mathbb{K}).$$

Théorème 6.13. *La fonction* $\operatorname{Mat}^{\mathcal{E}}_{\mathcal{F}} : \mathcal{L}(E, F) \to \mathcal{M}_{p\times n}(\mathbb{K})$ *est une application linéaire bijective.*

Démonstration. Soient φ et ψ deux applications linéaires. On veut montrer que

$$\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi+\psi)=\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)+\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\psi).$$

La j-ème colonne de $\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi+\psi)$ est la matrice colonne $C_{\mathcal{F}}((\varphi+\psi)(\boldsymbol{e}_j))$. Mais, d'après la définition de $\varphi+\psi$,

$$C_{\mathcal{F}}((\varphi+\psi)(\boldsymbol{e}_j))=C_{\mathcal{F}}(\varphi(\boldsymbol{e}_j)+\psi(\boldsymbol{e}_j))=C_{\mathcal{F}}(\varphi(\boldsymbol{e}_j))+C_{\mathcal{F}}(\psi(\boldsymbol{e}_j)),$$

d'où le résultat. On montre de la même manière que $\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\lambda\varphi)=\lambda\,\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)$.

Comme on sait qu'une application linéaire est entièrement déterminée par l'image d'une base, dès qu'une base est fixée au départ et à l'arrivée, la matrice de l'application détermine complètement l'application. Donc l'isomorphisme entre $\mathcal{L}(E,F)$ et $\mathcal{M}_{p\times n}(\mathbb{K})$. □

Exemple. On considère l'endomorphisme $\varphi:\mathbb{R}_2[X]\to\mathbb{R}_2[X]$ défini par $P\mapsto P'+2P$. C'est la somme de deux endomorphsimes ; le premier est la dérivation et le deuxième la homothétie de rapport 2 qui ont pour matrices associées dans la base canonique $(1,X,X^2)$,

$$A=\begin{pmatrix}0&1&0\\0&0&2\\0&0&0\end{pmatrix}\quad\text{et}\quad B=\begin{pmatrix}2&0&0\\0&2&0\\0&0&2\end{pmatrix}$$

respectivement. Par conséquent, la matrice de la somme est

$$A+B=\begin{pmatrix}2&1&0\\0&2&2\\0&0&2\end{pmatrix}.$$

Ceci signifie que la matrice colonne des coefficients de l'image du polynôme $P=a_0+a_1X+a_2X^2$ par φ est

$$\begin{pmatrix}2&1&0\\0&2&2\\0&0&2\end{pmatrix}\begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix}=\begin{pmatrix}2a_0+a_1\\2a_1+2a_2\\2a_2\end{pmatrix}.$$

c'est-à-dire que $\varphi(P)=2a_0+a_1+(2a_1+2a_2)X+2a_2X^2$, formule que l'on peut vérifier aisément en calculant directement $\varphi(P)$.

Par ailleurs, dans la proposition 6.4, on a vu que la composition des applications linéaires correspond à la multiplication des matrices de ces applications linéaires dès que des bases sont fixées pour chaque espace vectoriel concerné De plus, dans le thérème 6.8, on a vu que pour un endomorphsime bijectif, l'endomorphisme inverse correspond a la matrice inverse de l'endomorphisme. Tous ces résultats se formalisent en

Théorème–Définition 6.14. *L'espace vectoriel $\mathcal{M}_n(\mathbb{K})$ avec la multiplication des matrices est une $\mathbb{K}$-algèbre, c'est-à-dire*
- *l'application $\mathbb{K}\to\mathcal{M}_n(\mathbb{K})$ définie par $\lambda\mapsto\lambda I_n$ commute avec la somme et la multiplication*
- *pour toutes $A,B,C\in\mathcal{M}_n(\mathbb{K})$, on a $A(BC)=A(BC)$*
- *pour toutes $A,B,C\in\mathcal{M}_n(\mathbb{K})$, on a $A(B+C)=AB+AC$.*

De plus, si E est un $\mathbb{K}$-espace vectoriel de dimension n, alors $\mathcal{L}(E)$ est une $\mathbb{K}$-algèbre et l'isomorphisme linéaire $\mathrm{Mat}_{\mathcal{E}}^{\mathcal{E}} : \mathcal{L}(E) \to \mathcal{M}_n(\mathbb{K})$ *est un isomorphisme de $\mathbb{K}$-algèbres*[7]*.*

Démonstration. Il reste à établir l'associativité du produit des matrices et la distributivité du produit par rapport à la somme. Le coefficient (i, j) de $A(BC)$ est donné par

$$\sum_{k=1}^{n} a_{i,k}\Big(\sum_{l=1}^{n} b_{k,l}c_{l,j}\Big) = \sum_{k=1}^{n}\sum_{l=1}^{n} a_{i,k}b_{k,l}c_{l,j} = \sum_{l=1}^{n}\sum_{k=1}^{n} a_{i,k}b_{k,l}c_{l,j} = \sum_{l=1}^{n}\Big(\sum_{k=1}^{n} a_{i,k}b_{k,l}\Big)c_{l,j}$$

donc $A(BC) = (AB)C$. Le même argument établit la distributivité. $\square$

Exemple. On considère l'endomorphisme $\psi : \mathbb{R}_2[X] \to \mathbb{R}_2[X]$ défini par $P \mapsto P''$. On a $\psi = D \circ D = D^2$. La matrice M_ψ de ψ est égale au carré de la matrice de M_D de D.

$$M_\psi = M_D^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Ceci est bien correct puisque $\psi(a_0 + a_1X + a_2X^2) = 2a_2$.

En revenant à l'isomorphisme du théorème 6.13 $\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}} : \mathcal{L}(E, F) \to \mathcal{M}_{p\times n}(\mathbb{K})$, il faut remarquer que les deux notions de rang que nouus avons définies, à savoir le rang d'une application linéaire (voir la définition 5.8) et le rang d'une matrice(voir la proposition-définition 4.11), se correspondent.

Corollaire 6.15 (les rangs)**.** *Soit $\varphi : E \to F$ une application linéaire avec E et F de dimensions finies. Si $\mathcal{E}$ et $\mathcal{F}$ sont des bases de E et F respectivement, alors*

$$\mathrm{rank}(\varphi) = \mathrm{rank}(\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)).$$

Démonstration. Le rang de φ est le rang de l'image de φ, c'est-à-dire le rang de la famille formée par les vecteurs

$$\varphi(\boldsymbol{e}_1), \varphi(\boldsymbol{e}_2), \dots, \varphi(\boldsymbol{e}_n),$$

où $\mathcal{E} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n)$. Maintenant les colonnes de la matrice $\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)$ sont les matrices colonnes des coefficients des vectueurs $\varphi(\boldsymbol{e}_j)$, $1 \le j \le n$. On arrive à

$$\begin{aligned}
\mathrm{rank}(\varphi) &\overset{\text{déf}}{=} \dim(\mathrm{im}(\varphi)) = \dim(\mathrm{Vect}(\varphi(\boldsymbol{e}_1), \varphi(\boldsymbol{e}_2), \dots, \varphi(\boldsymbol{e}_n))) \\
&= \dim(\mathrm{Vect}(C_{\mathcal{F}}(\varphi(\boldsymbol{e}_1)), C_{\mathcal{F}}(\varphi(\boldsymbol{e}_2)), \dots, C_{\mathcal{F}}(\varphi(\boldsymbol{e}_n)))) \\
&= \dim\{\text{l'espace des colonnes de } \mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)\} \\
&\overset{\text{déf}}{=} \mathrm{rank}(\mathrm{Mat}_{\mathcal{F}}^{\mathcal{E}}(\varphi)).
\end{aligned}$$

$\square$

[7] Pour la structure de $\mathbb{K}$-algèbre de $\mathcal{L}(E)$, l'application $\mathbb{K} \to \mathcal{L}(E)$ est définie en enlvoyant λ en l'homothétie de rapport λ.

Exemple. Calculer le rang de l'application linéaire $\varphi : \mathbb{R}^3 \to \mathbb{R}^4$ définit par

$$(x, y, z) \mapsto (x + y + z, x - y + z, 2y, x - z).$$

On a vu que sa matrice dans les bases canoniques est $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & -1 \end{pmatrix}$ qui devient $\begin{pmatrix} 1 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 2 & 0 \\ 1 & -1 & -2 \end{pmatrix}$ après échelonnement. Donc le rang de φ est 3. (On peut trouver le rang en calculant la dimension du noyau et en utilisant le théorème du rang.)

6.6. Transposition

On présente une opération qui permet, pour une matrice, de transformer les lignes en colonnes et inversement.

Définition. Soit $A = (a_{i,j})_{\substack{1 \le i \le n \\ 1 \le j \le p}} \in \mathcal{M}_{n \times p}(\mathbb{K})$. On appelle *transposée* de A, ou la *matrice transposée* de A, la matrice $A^T = (\alpha_{j,i})_{\substack{1 \le j \le p \\ 1 \le i \le n}} \in \mathcal{M}_{p \times n}(\mathbb{K})$ définie par

$$\alpha_{j,i} = a_{i,j}$$

pour tous $1 \le j \le p$ et $1 \le i \le n$.

La correspondence $A \mapsto A^T$ est une application $\cdot^T : \mathcal{M}_{n \times p}(\mathbb{K}) \to \mathcal{M}_{p \times n}(\mathbb{K})$. Un premier exemple, très simple, aide à comprendre cette notion. Si $A = \left(\begin{smallmatrix} 2 & 1 & 2 \\ 3 & 2 & 4 \end{smallmatrix}\right)$ alors $A^T = \left(\begin{smallmatrix} 2 & 3 \\ 1 & 2 \\ 2 & 4 \end{smallmatrix}\right)$.

En ce qui concerne les propriétés de la transposée, on a le résultat suivant. Les points (ii) et (iii) disent que l'applcation $\cdot^T$ est linéaire, et le point (i) implique que $\cdot^T$ est un isomorphisme.

Proposition 6.16. *Pour toutes matrices A, B et tout scalaire α, on a :*

(i) $(A^T)^T = A$
(ii) $(A + B)^T = A^T + B^T$
(iii) $(\alpha A)^T = \alpha\, A^T$
(iv) $(AB)^T = B^T A^T$ *(attention au renversement)*
(v) *si A est inversible, alors $(A^T)^{-1} = (A^{-1})^T$.*

Démonstration. Les trois premières propriétés sont immédiates. Pour (iv), soit $C = AB$, ou $A \in \mathcal{M}_{n,p}$ et $B \in \mathcal{M}_{p,q}$. On pose $C = (c_{i,k})_{\substack{1 \le i \le n \\ 1 \le k \le q}} \in \mathcal{M}_{n,q}$. Alors

$$c_{i.k} = \sum_{j=1}^{p} a_{i,j} b_{j,k}$$

et, d'après la définition, $C^T = (\gamma_{k,i})_{\substack{1 \le k \le q \\ 1 \le i \le n}}$ avec $\gamma_{k,i}\beta_{k,j} = c_{i,k}$ (de même pour A^T et B^T). On a

$$\gamma_{k,i} = c_{i,k} = \sum_{j=1}^{p} a_{i,j} b_{j,k} = \sum_{j=1}^{p} \alpha_{j,i} \beta_{k,j} = \sum_{j=1}^{p} \beta_{k,j} \alpha_{j,i},$$

donc $C^T = B^T A^T$.

Pour (v), en utilisant (iv), on obtient

$$(A^{-1})^T A^T = (A\, A^{-1})^T = (I_n)^T = I_n.$$

Donc l'inverse de A^T est $(A^{-1})^T$. $\square$

7. Déterminants

7.1. Introduction

On commence par l'exemple de l'aire des parallélogrammes dans le plan. On considère l'espace vectoriel $\mathbb{R}^2$ muni de la base canonique $(\boldsymbol{e}_x, \boldsymbol{e}_y)$. À tout couple de vecteurs $(\boldsymbol{v}, \boldsymbol{w})$ on associe le parallélogramme $\square_{\boldsymbol{v},\boldsymbol{w}}$. En fixant la superficie du rectangle $\square_{\boldsymbol{e}_x,\boldsymbol{e}_y}$ comme l'unité (orienté) d'aire, il s'ensuit que pour tout couple $(\boldsymbol{v}, \boldsymbol{w})$ on peut calculer l'aire (orientée) du parallélogramme $\square_{\boldsymbol{v},\boldsymbol{w}}$. En notant par $\mu(\boldsymbol{v}, \boldsymbol{w})$ cette aire, on vient de construire l'application

$$\mu : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}, \qquad \mu(\boldsymbol{u}, \boldsymbol{v}) = \text{aire}(\square_{\boldsymbol{v},\boldsymbol{w}}).$$

Plusieurs questions se présentent naturellement : Quelles sont les propriétés de μ ? Comment calculer $\mu(\boldsymbol{v}, \boldsymbol{w})$ quand les coordonnées des vecteurs sont connues dans la base considérée ? Peut-on généraliser cette construction dans $\mathbb{R}^n$, voir dans un espace vectoriel de dimension finie sur $\mathbb{K}$?

Affirmation. μ est linéaire en chaque variable et antisymétrique, c'est-à-dire pour tous vecteurs $\boldsymbol{u}$ et $\boldsymbol{v}$ on a $\mu(\boldsymbol{u}, \boldsymbol{v}) = -\mu(\boldsymbol{v}, \boldsymbol{u})$.

La justification de cette affirmation se fait en plusieurs étapes en montrant successivement les identités suivantes :

(i) $\mu(\boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w}) = \mu(\boldsymbol{u}, \boldsymbol{w}) + \mu(\boldsymbol{v}, \boldsymbol{w})$ et $\mu(\boldsymbol{u}, \boldsymbol{v} + \boldsymbol{w}) = \mu(\boldsymbol{u}, \boldsymbol{v}) + \mu(\boldsymbol{u}, \boldsymbol{w})$
(ii) $\mu(-\boldsymbol{u}, \boldsymbol{v}) = -\mu(\boldsymbol{u}, \boldsymbol{v})$
(iii) $\mu(n\boldsymbol{u}, \boldsymbol{v}) = n\mu(\boldsymbol{u}, \boldsymbol{v})$ pour tout $n \in \mathbb{N}$
(iv) $\mu(q\boldsymbol{u}, \boldsymbol{v}) = q\mu(\boldsymbol{u}, \boldsymbol{v})$ pour tout $q \in \mathbb{Q}$ et donc, par continuité[8], $\mu(\lambda\boldsymbol{u}, \boldsymbol{v}) = \lambda\,\mu(\boldsymbol{u}, \boldsymbol{v})$ pour tout $\lambda \in \mathbb{R}$.
(v) $\mu(\boldsymbol{u}, \boldsymbol{v}) = -\mu(\boldsymbol{v}, \boldsymbol{u})$.

On indique brièvement le chemin. pour commencer, on inspecte la figure 8 et on en déduit l'additivité de μ dans la première variable et par un raisonnement analogue, l'additivité dans la deuxième variable. Par la suite, comme $0 = \mu(\boldsymbol{0}, \boldsymbol{v}) = \mu(\boldsymbol{u} - \boldsymbol{u}, \boldsymbol{v})$, on obtient (ii). En utilisant de nouveau l'additivité, on montre que μ commute avec la multiplication par scalaire dans chacune des variables. Par exemple, pour tous $n \in \mathbb{N}$, on a

$$\mu(n\boldsymbol{u}, \boldsymbol{v}) = \mu(\boldsymbol{u}, \boldsymbol{v}) + \cdots + \mu(\boldsymbol{u}, \boldsymbol{v} = n\mu(\boldsymbol{u}, \boldsymbol{v}).$$

En remplaçant $\boldsymbol{u}$ par $\frac{1}{n}\,\boldsymbol{u}$, on a $\mu(\frac{1}{n}\,\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{n}\,\mu(\boldsymbol{u}, \boldsymbol{v})$. Donc on obtient la commutativité avec la multiplication par tout scalaire rationnel et donc, par continuité, par tout scalaire réel. En fin, pour (v), on développe

$$\begin{aligned}
0 &= \mu(\boldsymbol{u} + \boldsymbol{v}, \boldsymbol{u} + \boldsymbol{v}) \\
&= \mu(\boldsymbol{u}, \boldsymbol{u} + \boldsymbol{v}) + \mu(\boldsymbol{v}, \boldsymbol{u} + \boldsymbol{v}) \\
&= \mu(\boldsymbol{u}, \boldsymbol{u}) + \mu(\boldsymbol{u}, \boldsymbol{v}) + \mu(\boldsymbol{v}, \boldsymbol{u}) + \mu(\boldsymbol{v}, \boldsymbol{v}) \\
&= \mu(\boldsymbol{u}, \boldsymbol{v}) + \mu(\boldsymbol{v}, \boldsymbol{u}).
\end{aligned}$$

Affirmation. Si $\boldsymbol{u} = (a, b)$ et $\boldsymbol{v} = (c, d)$, alors $\mu(\boldsymbol{u}, \boldsymbol{v}) = ad - bc$.

[8] Le côté naturel de la continuité. . .

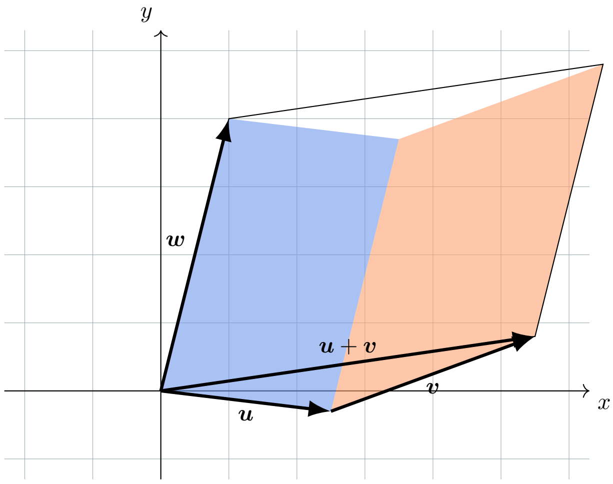

Figure 8: L'additivité du déterminant en chaque variable : l'aire du parallélogramme engendré par les vecteurs $\boldsymbol{u}+\boldsymbol{v}$ et $\boldsymbol{w}$ et égale à la somme des aires des parallélogrammes engendrés par $\boldsymbol{u}$ et $\boldsymbol{w}$ et, respectivement, par $\boldsymbol{v}$ et $\boldsymbol{w}$.

Ceci découle des linéarités établies précédemment. On a

$$\begin{aligned}
\mu(\boldsymbol{u},\boldsymbol{v}) &= \mu(a\boldsymbol{e}_x + b\boldsymbol{e}_y, c\boldsymbol{e}_x + d\boldsymbol{e}_y)\\
&= a\mu(\boldsymbol{e}_x, c\boldsymbol{e}_x + d\boldsymbol{e}_y) + b\mu(\boldsymbol{e}_y, c\boldsymbol{e}_x + d\boldsymbol{e}_y)\\
&= ac\,\mu(\boldsymbol{e}_x,\boldsymbol{e}_x) + ad\,\mu(\boldsymbol{e}_x,\boldsymbol{e}_y) + bc\,\mu(\boldsymbol{e}_y,\boldsymbol{e}_x) + bd\,\mu(\boldsymbol{e}_y,\boldsymbol{e}_y)\\
&= ad - bc.
\end{aligned}$$

La notion d'aire algébrique introduite ci-dessus, se généralise en plusieurs dimensions (le volume algébrique des parallélépipèdes dans $\mathbb{R}^3$, par exemple) ou pour tout espace vectoriel de dimension finie E sur $\mathbb{K}$ sous de nom de *déterminant*.

7.2. Généralités

Définition (application multi-linéaire alternée)**.** Soit E un espace vectoriel sur $\mathbb{K}$ et soit $f : E^n = E \times \ldots \times E \to \mathbb{K}$ une application.

- ⋄ f est dite multi-linéaire lorsqu'elle est linéaire par rapport à chaque variable
- ⋄ f est dite alternée lorsque pour tous $1 \le i < j \le n$ et pour tout $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) \in E^n$, on a

$$f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_j, \ldots, \boldsymbol{v}_n) = -f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_n).$$

Il est facile de voir que la deuxième condition est équivalente à demander que pour tous $1 \le i < j \le n$ et tout $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) \in E^n$ avec $\boldsymbol{v}_i = \boldsymbol{v}_j$, on a $f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = 0$. Par exemple, l'application aire algébrique $\mu : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ introduite précédemment est bilinéaire alternée.

Définition (permutation)**.** Une fonction bijective $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ est appelée une *permutation*. L'ensemble de toutes les permutations de $\{1, \ldots, n\}$ est noté $\mathfrak{S}_n$.

Définition (signature d'une permutation)**.** Soit $\sigma \in \mathfrak{S}_n$ une permutation. On dit que la paire $\{i, j\}$, avec $1 \le i < j \le n$, est en *inversion* pour σ si $\sigma(i) > \sigma(j)$. La signature de σ, notée $\varepsilon(\sigma)$, est l'entier $(-1)^{N(\sigma)}$ où $N(\sigma)$ est le nombre de paires en inversion pour σ.

Théorème–Définition 7.1. *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base. Il existe une unique forme multi-linéaire alternée sur E^n, appelée déterminant dans la base $\mathcal{B}$ et notée $\det_{\mathcal{B}} : E^n \to \mathbb{K}$ telle que $\det_{\mathcal{B}}(e_1, \dots, e_n) = 1$.*

De plus, tout autre forme multi-linéaire alternée sur E^n est un multiple de celle-ci.

Démonstration. On développe la preuve dans le cas $n = 3$. C'est un calcul qui aboutira à une formule explicite. Le cas général se traite de la même manière ; on indiquera la formule du déterminant pour n général à la fin.

Soient, pour tout $1 \le i \le n$, la matrice colonne des coefficients de $\boldsymbol{v}_i$,

$$C_{\mathcal{B}}(\boldsymbol{v}_i) = \begin{pmatrix} a_{1,i} \\ \vdots \\ a_{n,i} \end{pmatrix}.$$

Alors, pour $n = 3$, comme f est une forme multi-linéaire alternée, on obtient

$$\begin{aligned}
&f(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3) \\
&\quad = f(a_{1,1}\boldsymbol{e}_1 + a_{2,1}\boldsymbol{e}_2 + a_{3,1}\boldsymbol{e}_3, a_{1,2}\boldsymbol{e}_1 + a_{2,2}\boldsymbol{e}_2 + a_{3,2}\boldsymbol{e}_3, a_{1,3}\boldsymbol{e}_1 + a_{2,3}\boldsymbol{e}_2 + a_{3,3}\boldsymbol{e}_3) \\
&\quad = a_{1,1}a_{2,2}a_{3,3}\, f(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3) + a_{1,1}a_{3,2}a_{2,3}\, f(\boldsymbol{e}_1, \boldsymbol{e}_3, \boldsymbol{e}_2) + a_{2,1}a_{1,2}a_{3,3}\, f(\boldsymbol{e}_2, \boldsymbol{e}_1, \boldsymbol{e}_3) \\
&\qquad + a_{2,1}a_{3,2}a_{1,3}\, f(\boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_1) + a_{3,1}a_{2,2}a_{1,3}\, f(\boldsymbol{e}_3, \boldsymbol{e}_2, \boldsymbol{e}_1) + a_{3,1}a_{1,2}a_{2,3}\, f(\boldsymbol{e}_3, \boldsymbol{e}_1, \boldsymbol{e}_2) \\
&\quad = \sum_{\sigma \in \mathfrak{S}_3} a_{1,\sigma(1)}a_{2,\sigma(2)}a_{3,\sigma(3)}\, f(\boldsymbol{e}_{\sigma(1)}, \boldsymbol{e}_{\sigma(2)}, \boldsymbol{e}_{\sigma(3)}) \\
&\quad = \sum_{\sigma \in \mathfrak{S}_3} \varepsilon(\sigma) a_{1,\sigma(1)}a_{2,\sigma(2)}a_{3,\sigma(3)}\, f(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3) \\
&\quad = \sum_{\sigma \in \mathfrak{S}_3} \varepsilon(\sigma) a_{1,\sigma(1)}a_{2,\sigma(2)}a_{3,\sigma(3)}.
\end{aligned}$$

Donc en fixant la valeur de l'application multi-linéaire alternée sur la base $\mathcal{B}$, on fixe sa valeur sur tout n-uplet de vecteurs de E.

Il est évident maintenant qu'en général, si on utilise les mêmes notations, on obtient

$$f(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}. \tag{7.1}$$

Le pas difficile est franchie dans la quatrième égalité dans la suite d'égalités ci-dessus. □

La preuve du théorème suggère, d'une part, un premier exemple, et de l'autre, des techniques de calcul du déterminant. Nous allons étudier ces techniques dans la section suivante. En ce qui concerne l'exemple, on considère $E = \mathbb{K}^2$ —ou plus généralement $E = \mathbb{K}^n$ —muni de la base canonique $\mathcal{B}$. Pour une famille (ordonnée) de deux vecteurs dans $\mathbb{K}^2$, on a

$$\det_{\mathcal{B}}\left(\begin{pmatrix} a_{1,1} \\ a_{2,1} \end{pmatrix}, \begin{pmatrix} a_{1,2} \\ a_{2,2} \end{pmatrix} \right) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

Pour une famille (ordonnée) de trois vecteurs dans $\mathbb{K}^3$, on obtient

$$\begin{aligned}
\det_{\mathcal{B}}\left(\begin{pmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \end{pmatrix}, \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ a_{3,2} \end{pmatrix}, \begin{pmatrix} a_{1,3} \\ a_{2,3} \\ a_{3,3} \end{pmatrix} \right) &= a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} \\
&\quad - a_{1,2}a_{2,1}a_{3,3} - a_{1,1}a_{2,3}a_{3,2} - a_{1,3}a_{2,2}a_{3,1}.
\end{aligned} \tag{7.2}$$

Étant donnée une matrice *carrée* $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$, les colonne de la matrice peuvent être interprétées comme une famille ordonnée de vecteurs de $\mathbb{K}^n$. On introduit la définition suivante :

Définition. Soit $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$. Le déterminant de A, noté $\det(A)$ ou $|A|$, est le déterminant des vecteurs colonnes de A dans la base canonique de $\mathbb{K}^n$.

Par conséquent, la formule précédente pour le calcul du déterminant dans la base canonique de $\mathbb{K}^3$ d'une famille de trois vecteurs et aussi la formule pour calculer le déterminant d'une matrice 3×3.

Remarque 7.2. Dans la formule du déterminant établie dans (7.2), il y a trois termes apparaissant avec le signe $+$ et trois apparaissant avec le signe $-$. Chaque terme, indépendamment du signe, est le produit de trois éléments de la matrice ; on peut penser chaque terme dans la formule du déterminant comme un triangle dont les sommets se trouvent sur certaines positions de la matrice. La figure 9 représente ces six triangles (isocèles). Il est possible d'utiliser

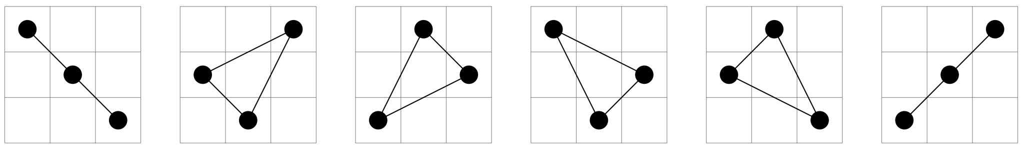

Figure 9: Les six termes du déterminant d'une matrice 3×3 représentés comme triangles isocèles sur les positions de la matrices. Le quatrième triangle correspond au terme $-a_{1,1}a_{2,3}a_{3,2}$.

ces schémas comme moyen mnémotechnique pour retrouver la formule du déterminant.

Comme le déterminant est une application multi-linéaire alternée, en découlent les premières propriétés importantes.

Corollaire 7.3. *Soit E un $\mathbb{K}$-espace vectoriel de dimension n, soit $\mathcal{B}$ une base de E et soit $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ une famille ordonnée.*

(i) *Si un des vecteurs $\boldsymbol{v}_i$ est nul, alors $\det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = 0$.*

(ii) *Si $\boldsymbol{v}_i = \boldsymbol{v}_j$ pour $i \neq j$, alors $\det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = 0$.*

(iii) *Pour tout i et tous $\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_n \in \mathbb{K}$,*

$$\det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \sum_{j=1}^{i-1} \lambda_j \boldsymbol{v}_j + \boldsymbol{v}_i + \sum_{j=i+1}^{n} \lambda_j \boldsymbol{v}_j, \ldots, \boldsymbol{v}_n).$$

(Le déterminant de la famille ne change pas si on ajoute à un des vecteurs une combinaison linéaire des autres vecteurs.)

Démonstration. La première affirmation est une conséquence de la linéarité ou par calcul directe en utilisant la formule (7.1). Pour (ii) on utilise le fait que $\det_{\mathcal{B}}$ est alternée.

Pour établir (iii) on considère le cas $i = n$. On a

$$\begin{aligned}
\det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n + \sum_{j=1}^{n-1} \lambda_j \boldsymbol{v}_j) &= \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) + \sum_{j=1}^{n-1} \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-1}, \lambda_j \boldsymbol{v}_j) \\
&= \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) + \sum_{j=1}^{n-1} \lambda_j \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-1}, \boldsymbol{v}_j) \\
&= \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n).
\end{aligned}$$

□

La proposition ci-après est aussi une conséquence (mais plus subtile) du fait que le déterminant est multi-linéaire et alterné. En pensant à l'exemple de l'aire dans le plan, cette proposition répond à la question suivante : comment peut-on comparer les aires d'un même parallélogramme qui sont mesurées en utilisant des bases différentes (c'est-à-dire des "unités de mesure" différentes) ?

Proposition 7.4. *Soient $\mathcal{B} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ et $\mathcal{B}' = (\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n)$ deux bases de E. Alors*

(i) *pour tous vecteurs $\boldsymbol{v}_i \in E$, $1 \le i \le n$, on a $\det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$*

(ii) *$\det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(\mathcal{B}) = 1$.*

Démonstration. Soit $\delta = \det_{\mathcal{B}}(\mathcal{B}')$ et soit $g : E^n \to \mathbb{K}$ l'application définie par

$$E^n \ni (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) \mapsto [g]\delta \det_{\mathcal{B}'}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$$

L'application g est une forme n-linéaire alternée sur E telle que

$$g(\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n) = g(\mathcal{B}') = \delta \cdot 1 = \delta.$$

De même, $\det_{\mathcal{B}}$ est une forme n-linéaire sur E telle que $\det_{\mathcal{B}}(\mathcal{B}') = \delta$. D'après le théorème 7.1, $g = \det_{\mathcal{B}}$ et donc pour tout $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) \in E^n$, on a (i), c'est-à-dire

$$\det_{\mathcal{B}}(\mathcal{B}')\ \det_{\mathcal{B}'}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = g(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n).$$

Pour obtenir (ii), on utilise la famille $\mathcal{B}$ dans (i). On obtient $1 = \det_{\mathcal{B}}(\mathcal{B}) = \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(\mathcal{B})$. □

Si E est de dimension n, alors le déterminant (dans n'importe quelle base de E) permet de décider si une famille de n vecteurs est libre ou liée. (Attention, il faut que le nombre de vecteurs soit égal à la dimension de l'espace.)

Proposition 7.5. *Soit $\mathcal{B}$ une base quelconque de l'espace vectoriel n-dimensionnel E. Si $\mathcal{F}$ est une famille de n vecteurs de E, alors $\mathcal{F}$ est libre si et seulement si $\det_{\mathcal{B}}(\mathcal{F}) \neq 0$.*

Démonstration. Dans le sens direct, si $\mathcal{F}$ est libre, alors $\mathcal{F}$ est une base de E (voir le théorème 3.2, (c)). En utilisant la proposition 7.4, $\det_{\mathcal{B}}(\mathcal{F}) \det_{\mathcal{F}}(\mathcal{B}) = 1$ et donc $\det_{\mathcal{B}}(\mathcal{F}) \neq 0$.

Pour justifier l'affirmation réciproque, on suppose $\mathcal{F}$ liée ; quitte a renommer les vecteurs, on peut supposer que $\boldsymbol{v}_n = \sum_{i=1}^{n-1} \lambda_i \boldsymbol{v}_i$. Alors, en appliquant le point (iii) du corollaire 7.3,

$$\det_{\mathcal{B}}(\mathcal{F}) = \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n - \sum_{i=1}^{n-1} \lambda_i \boldsymbol{v}_i) = \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-1}, \boldsymbol{0}) = 0.$$

□

7.3. Calcul pratique des déterminants

Pour calculer un déterminant de n vecteurs $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ dans la base $\mathcal{B}$ de E on utilise le calcul matriciel, c'est-à-dire la technique et la formule apparaissant dans la preuve du théorème 7.1. Explicitement, on écrit la matrice de décomposition de cette famille dans la base $\mathcal{B}$ et on calcule le déterminant de cette matrice.

Avant de regarder en détails ces règles de calcul, il faut comprendre conceptuellement les objets manipulés lors de cette opération et les mettre en liaison avec les considérations précédentes concernant les applications linéaires.

- Fixer une base $\mathcal{B}$ du $\mathbb{K}$-espace vectoriel n-dimensionnel E, revient à identifier tout vecteur de E à vecteur colonne de $\mathbb{K}^n$. Symboliquement, on considère l'application linéaire bijective (l'isomorphisme) $C_{\mathcal{B}} : E \to \mathbb{K}^n$.
- Identifier une famille ordonnée de n vecteurs de E à la matrice de décomposition de cette famille dans la base $\mathcal{B}$ revient à considérer l'application "produit"
$$A_{\mathcal{B}} : E \times \cdots \times E \longrightarrow \mathcal{M}_{n,n}(\mathbb{K})$$
définie par
$$A_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = \begin{pmatrix} C_{\mathcal{B}}(\boldsymbol{v}_1) & \cdots & C_{\mathcal{B}}(\boldsymbol{v}_n) \end{pmatrix} \in \mathcal{M}_{n,n}(\mathbb{K}).$$
Cette application est linéaire sur $\mathbb{K}$.
- Les déterminants de la matrice et de la famille sont reliés par la formule
$$\det(A_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)) = \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n).$$
Ceci est équivalent à la commutativité du diagramme ci-dessus.
$$\begin{tikzcd} E \times \cdots \times E \arrow[rr, "A_{\mathcal{B}}"] \arrow[dr, "\det_{\mathcal{B}}"'] & & \mathcal{M}_{n,n} \arrow[dl, "\det"] \\ & \mathbb{K} & \end{tikzcd}$$
- Est-ce que le déterminant est constant sur les matrices semblables ? En d'autres termes, peut-on définir le déterminant d'un endomorphisme ?
On rappelle qu'à tout endomorphisme $\varphi \in \mathrm{End}(E)$, on a associé une matrice carrée $\mathrm{Mat}_{\mathcal{B}}^{\mathcal{B}}(\varphi)$ qui dépend de la base $\mathcal{B}$. En changeant la base, la matrice est remplacée par une matrice semblable. Est-ce que les valeurs de la fonction composée
$$\mathrm{End}(E) \xrightarrow{\mathrm{Mat}_{\mathcal{B}}^{\mathcal{B}}} \mathcal{M}_n(\mathbb{K}) \xrightarrow{\det} \mathbb{K}$$
sont-elles indépendantes de la base $\mathcal{B}$? La réponse sera positive et on la justifiera dans la section 7.4.

Développement suivant une ligne ou une colonne

On a vu après la preuve du théorème 7.1 comment calculer le déterminant d'une matrice 2×2 ou 3×3. Par exemple pour $A = (a_{i,j})$ une matrice de taille 3×3,
$$\det(A) = a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{1,2}a_{2,1}a_{3,3} - a_{1,1}a_{3,2}a_{2,3} - a_{1,3}a_{2,2}a_{3,1}.$$
La figure 9 suggérait une méthode mnémotechnique pour retenir cette formule ; une autre méthode serait la suivante : on recopie les deux premières lignes de la matrice en-dessous de la matrice ; on additionne le produit des trois diagonales principales puis on retranche le produit des trois autres diagonales. Par exemple,
$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 \\ 3 & -1 & 2 \\ -2 & 3 & 1 \end{vmatrix} &= 1 \cdot (-1) \cdot 1 + 3 \cdot 3 \cdot 3 + (-2) \cdot 2 \cdot 2 - (-2) \cdot (-1) \cdot 3 - 1 \cdot 3 \cdot 2 - 3 \cdot 2 \cdot 1 \\ &= -1 + 27 - 8 - 6 - 6 - 6 \\ &= 0. \end{aligned}$$

Pour calculer des déterminants d'ordre plus élevé, on pourra effectuer des développements suivant une ligne ou une colonne de la matrice pour se ramener à des déterminants d'ordre plus petit.

Définition. Soit A une matrice carrée. Soit $A_{i,j}$ la matrice obtenue en supprimant la ligne i et la colonne j de A. On appelle *mineur* de place (i,j) de A le déterminant $\det(A_{i,j})$. On appelle *cofacteur* de place (i,j) de A l'expression $(-1)^{i+j}\det(A_{i,j})$.

Remarque. En pratique, on ne calcule pas le signe $(-1)^{i+j}$ apparaissant dans la formule du cofacteur ; on applique l'alternance des signes sachant que l'élément place $(1,1)$ est muni du signe $+$ et chaque déplacement sur l'horizontale ou la verticale introduit un changement de signe. Graphiquement on obtient le tableau des signes suivant.

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

On remarque que les éléments sur la diagonale ont toujours le signe $+$.

Proposition 7.6. *Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$. Pour tout $1 \leq i \leq n$, la formule de développement du déterminant de A suivant la ligne i,*

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

est vérifiée.

La formule de développement du déterminant suivant la colonne j est vraie aussi. Elle s'obtient de la formule suivant la ligne par transposition. On y reviendra après la démonstration de la proposition 7.6.

Esquisse de preuve. On considère le développement suivant la linge n. Le cas général suit en utilisant la multi-linéarité alternée du déterminant. D'après la preuve du théorème 7.1,

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}.$$

L'idée est de voir l'ensemble des permutations comme une réunion disjointe de n sous-ensemble indexés par l'image de n. Le premier, $\mathfrak{S}_n^1$, est celui composé des permutations σ pour lesquelles $\sigma(n) = 1$. Le deuxième, $\mathfrak{S}_n^2$, contient les permutations qui vérifient $\sigma(n) = 2$, et ainsi de suite. Chacun de ces sous-ensembles est en bijection avec l'ensemble des permutations de $n-1$ objets. Le cœur de l'argument est formé par l'étude des signatures des permutations apparaissant à travers ce développement. En admettant cette étude, on obtient

$$\begin{aligned} \det(A) &= \sum_{\sigma \in \mathfrak{S}_n^1} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n-1,\sigma(n-1)} a_{n,1} + \cdots + \sum_{\sigma \in \mathfrak{S}_n^n} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n-1,\sigma(n-1)} a_{n,n} \\ &= a_{n,1} \sum_{\gamma \in \mathfrak{S}_{n-1}} (-1)^{n-1} \varepsilon(\gamma) a_{1,\gamma(1)} \cdots a_{n-1,\gamma(n-1)} + \cdots \\ &\quad + a_{n,n} \sum_{\gamma \in \mathfrak{S}_{n-1}} (-1)^{n-n} \varepsilon(\gamma) a_{1,\gamma(1)} \cdots a_{n-1,\gamma(n-1)} \end{aligned}$$

d'où le résultat, car $(-1)^{n-k} = (-1)^{n+k}$. $\square$

Remarque. Lorsqu'on choisit de développer suivant une ligne (ou une colonne), on a tout intérêt à choisir une ligne (ou colonne) comportant beaucoup de zéros. Dans l'exemple suivant, on développe suivant la deuxième ligne et par la suite, suivant la dernière linge (pour les déterminants 3×3).

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 3 & 1 & 1 \\ 1 & 0 & 4 & 0 \end{vmatrix} = (-1) \cdot 1 \cdot \begin{vmatrix} 1 & 1 & 1 \\ 3 & 1 & 1 \\ 0 & 4 & 0 \end{vmatrix} + 0 \cdot |*| + (-1) \cdot 1 \cdot \begin{vmatrix} 1 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 0 & 0 \end{vmatrix} + 0 \cdot |*|$$

$$= (-1) \cdot (-1) \cdot 4 \cdot \begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} = 4(1-3) - (1-3) = -6.$$

Proposition 7.7. *Si $A \in \mathcal{M}_n(\mathbb{K})$, alors* $\det(A^T) = \det(A)$.

Démonstration. Si $A^T = (b_{i,j})$ avec $b_{i,j} = a_{j,i}$, alors

$$\det(A^T) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) b_{1,\sigma(1)} b_{2,\sigma(2)} \cdots b_{n,\sigma(n)} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}.$$

Comme chaque σ est une bijection, en réordonnant les facteurs, on a

$$a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n} = a_{1,\sigma^{-1}(1)} a_{2,\sigma^{-1}(2)} \cdots a_{n,\sigma^{-1}(n)}.$$

De plus $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$, donc

$$\det(A^T) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma^{-1}) a_{1,\sigma^{-1}(1)} a_{2,\sigma^{-1}(2)} \cdots a_{n,\sigma^{-1}(n)}$$

et en changeant l'indice de sommation, $\det(A^T) = \det(A)$. □

En mettant ensemble le déterminant de la transposée et la formule de développement suivant la ligne i, on obtient la formule de développement du déterminant suivant la colonne i :

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

Les formules de développement suivant une ligne ou une colonne permettent de calculer le déterminant d'une matrice triangulaire. Une matrice carrée est dite *triangulaire supérieure* (ou inférieure) si tous les éléments au-dessous (ou au-dessus) de la diagonale sont nuls.

Corollaire 7.8 (matrice triangulaire)**.** *Si une matrice carrée est triangulaire, alors son déterminant est égal au produit des éléments diagonaux.*

Démonstration. Supposons que la matrice $A = (a_{i,j})$ soit triangulaire inférieure. On développe suivant la première ligne, cela donne $a_{1,1}$ multiplié par le déterminant de $A_{1,1}$ qui est aussi une matrice triangulaire inférieure. On itère le processus. Lorsqu'il s'arrête, on a bien obtenu le produit des éléments diagonaux. □

En particulier, si A est diagonale, c'est-à-dire $A = \mathrm{Diag}(a_1, a_2, \ldots, a_n)$, alors son déterminant est égal au produit des éléments diagonaux, c'est-à-dire $\det(A) = a_1 \cdot a_2 \cdots a_n$.

Remarque 7.9. Dans les algorithmes de calcul du déterminant d'une matrice carrée, on se ramène à des situations triangulaires. Pour ce faire, on utilise la méthode du pivot de Gauss, c'est-à-dire on utilise des opérations élémentaires sur les lignes et les colonnes des matrices successives apparaissant lors du calcul.

Les opérations élémentaires dont il est question dans la remarque précédente, sont les opérations énumérées dans le corollaire 7.3 —établies pour le déterminant d'une famille de vecteurs dans une base quelconque.

Opérations élémentaires sur les lignes et les colonnes

Ci-dessous, on rassemble et expriment en termes matriciels les règles à respecter lors du calcul du déterminant d'une matrice. Ces règles sont utilisées pour faire apparaî tre de multiples zéros dans la matrice en réalisant des opérations élémentaires sur les lignes ou les colonnes.

- si une colonne est nulle, le déterminant est nul
- si une colonne est combinaison linéaire des autres, le déterminant est nul
- on ne change pas la valeur d'un déterminant si l'on ajoute à une colonne une combinaison linéaire des autres
- si l'on multiplie une colonne par un scalaire λ, le déterminant est multiplié par λ
- si on échange deux colonnes, le déterminant change de signe
- les propriétés précédentes sont aussi valables pour les lignes.

Par exemple, pour le calcul suivant, on rajoute un multiple de la quatrième colonne à toutes des autres tels que $a_{1,i} = 0$ pour $i \leq 3$.

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

Pour la troisième égalité, on a ajouté la deuxième colonne à la première et puis on a échangé les deux colonnes.

Exemple 7.10 (déterminant de Vandermond)**.** Si $a_1, a_2, \ldots, a_n \in \mathbb{K}$, alors

$$V(a_1, \ldots, a_n) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

Pour voir ceci, on effectue les opérations suivantes (on note L_i la ligne i) :

$$L_n \rightsquigarrow L_n - a_1 L_{n-1} \quad L_{n-1} \rightsquigarrow L_{n-1} - a_1 L_{n-2} \quad \cdots \quad L_2 \rightsquigarrow L_2 - a_1 L_1.$$

Le déterminant $V(a_1, \ldots, a_n)$ devient

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & a_2 - a_1 & \cdots & a_n - a_1 \\ 0 & a_2^2 - a_1 a_2 & \cdots & a_n^2 - a_1 a_n \\ \vdots & \vdots & & \vdots \\ 0 & a_2^{n-1} - a_1 a_2^{n-2} & \cdots & a_n^{n-1} - a_1 a_n^{n-2} \end{vmatrix} = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & a_2 - a_1 & \cdots & a_n - a_1 \\ 0 & a_2(a_2 - a_1) & \cdots & a_n(a_n - a_1) \\ \vdots & \vdots & & \vdots \\ 0 & a_2^{n-2}(a_2 - a_1) & \cdots & a_n^{n-2}(a_n - a_1) \end{vmatrix}.$$

En développant suivant la première colonne et en sortant le facgteur commun de cahcune des autres colonnes, on a

$$V(a_1, \ldots, a_n) = (a_2 - a_1) \cdots (a_n - a_1) V(a_2, \ldots, a_n).$$

Le résultat s'ensuit par récurrence.

7.4. Déterminant d'un endomorphisme

On revient à la question concernant le comportement de la composition (voir page 66)

$$\mathrm{End}(E) \xrightarrow{\mathrm{Mat}_{\mathcal{B}}^{\mathcal{B}}} \mathcal{M}_n(\mathbb{K}) \xrightarrow{\det} \mathbb{K}$$

si la base $\mathcal{B}$ est remplacée par une autre base $\mathcal{B}'$. Pour répondre à cette question, il faut relier cette composition à une application multi-linéaire alternée. L'application linéaire $\mathrm{Mat}_{\mathcal{B}}^{\mathcal{B}}$ associe à un endomorphisme φ la matrice de décomposition de la famille ordonnée $(\varphi(\boldsymbol{e}_1), \ldots, \varphi(\boldsymbol{e}_n))$ dans la base $\mathcal{B} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$. L'application multi-linéaire alternée recherchée apparaît naturellement si on regarde la matrice de décomposition associée à toute famille de vecteurs $(\varphi(\boldsymbol{v}_1), \ldots, \varphi(\boldsymbol{v}_n))$, c'est-à-dire en considérant le diagramme suivant.

$$\begin{tikzcd}
E^n \arrow[r, "\varphi\times\cdots\times\varphi"] \arrow[d, "\det_{\mathcal{B}}"'] & E^n \arrow[r, "A_{\mathcal{B}}"] \arrow[d, "\det_{\mathcal{B}}"] & \mathcal{M}_{n,n} \arrow[dl, "\det"] \\
\mathbb{K} \arrow[r, "c_{\mathcal{B}}(\varphi)"] & \mathbb{K} &
\end{tikzcd} \tag{7.3}$$

Pour une famille de vecteurs $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$, on a (interprétation sur le carré du diagramme ; le triangle sera utilisé dans la démonstration suivante)

$$\det_{\mathcal{B}}(\varphi(\boldsymbol{v}_1,), \ldots, \varphi(\boldsymbol{v}_n)) = c_{\mathcal{B}}(\varphi)\ \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n),$$

où $c_{\mathcal{B}}(\varphi)$ est une constante qui dépend de l'endomorphisme considéré ainsi que de la base $\mathcal{B}$. Cette constante est produite par le théorème-définition 7.1 car $\det_{\mathcal{B}} \circ (\varphi \times \cdots \times \varphi)$ est une application n-linéaire alternée. Le théorème suivant montre en fait que la constante $c_{\mathcal{B}}(\varphi)$ est indépendante de la base $\mathcal{B}$ et par la suite qu'elle est égale au déterminant de $\mathrm{Mat}_{\mathcal{B}}^{\boldsymbol{b}}(\varphi)$.

Théorème 7.11. *Soit E un espace vectoriel de dimension $n \geq 1$ et soit $\varphi \in \mathcal{L}(E)$. Il existe une unique constante $c(\varphi) \in \mathbb{K}$ telle que pour toute base $\mathcal{B} = (e_1, \ldots, e_n)$ de E et pour toute famille ordonnée $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) \in E^n$ on ait*

$$\det_{\mathcal{B}}(\varphi(\boldsymbol{v}_1), \ldots, \varphi(\boldsymbol{v}_n)) = c(\varphi) \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n).$$

De plus, $c(\varphi) = \det(\mathrm{Mat}_{\mathcal{B}}^{\mathcal{B}}(\varphi))$.

Démonstration. On a vu dans le diagramme 7.3 qu'il existe une constante $c_{\mathcal{B}}(\varphi)$ telle que

$$\det_{\mathcal{B}}(\varphi(\boldsymbol{v}_1), \ldots, \varphi(\boldsymbol{v}_n)) = c_{\mathcal{B}}(\varphi) \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n).$$

Pour justifier qu'elle est indépendante de la base $\mathcal{B}$, soit $\mathcal{B}' = (e_1', \ldots, e_n')$ une autre base de E. Alors

$$\begin{aligned}
\det_{\mathcal{B}'}(\varphi(\boldsymbol{v}_1), \ldots, \varphi(\boldsymbol{v}_n)) = {} & \det_{\mathcal{B}'}(\mathcal{B})\ \det_{\mathcal{B}}(\varphi(\boldsymbol{v}_1), \ldots, \varphi(\boldsymbol{v}_n)) && \text{proposition 7.4, (i)} \\
& \det_{\mathcal{B}'}(\mathcal{B})\, c_{\mathcal{B}}(\varphi)\ \det_{\mathcal{B}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) && \\
& c_{\mathcal{B}}(\varphi)\ \det_{\mathcal{B}'}(\mathcal{B})\ \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) && \text{proposition 7.4, (i)} \\
& c_{\mathcal{B}}(\varphi)\ \det_{\mathcal{B}'}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) && \text{proposition 7.4, (ii)}
\end{aligned}$$

donc $c_{\mathcal{B}'}(\varphi) = c_{\mathcal{B}}(\varphi)$, c'est-à-dire la constante est indépendante de la base choisie. Pour identifier cette constante (et donc obtenir son unicité aussi), on utilise le triangle commutatif du diagramme 7.3 ; en prenant $\boldsymbol{v}_i = \boldsymbol{e}_i$ pour tout i, on a

$$c(\varphi) = \det_{\mathcal{B}}(\varphi(\boldsymbol{e}_1), \ldots, \varphi(\boldsymbol{e}_n)) = \det(A_{\mathcal{B}}(\varphi(\boldsymbol{e}_1), \ldots, \varphi(\boldsymbol{e}_n))) = \det(\mathrm{Mat}_{\mathcal{B}}^{\mathcal{B}}(\varphi)).$$

□

On a démontré dans le théorème précédent que la constante $c(\varphi)$ est le déterminant de toute matrice associée à l'endomorphisme φ. Cet élément de $\mathbb{K}$ s'appelle le *déterminant de* φ et se note $\det(\varphi)$.

Exemple. Pour $Id_E : E \to E$ on a $\det(Id_E) = \det(I_n) = 1$.

Remarque. Le déterminant d'un endomorphisme est donc une notion intrinsèque indépendante de la base choisie contrairement au déterminant d'une famille de vecteurs, notion qui dépend de la base choisie. En particulier si A est la matrice de φ dans la base $\mathcal{B}$, A' est la matrice de φ dans la base $\mathcal{B}'$ et P est la matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$, alors

$$A' = P^{-1} A\, P \qquad \text{et} \qquad \det(A') = \det(P^{-1} A\, P) = \det(A).$$

Les deux propositions suivantes établissent les propriétés importantes du déterminant par rapport à la composition des endomorphismes. Elles sont codées dans l'affirmation

$$\det : \mathcal{L}(E) \longrightarrow \mathbb{K} \quad \text{est un morphisme de groupes.}$$

Proposition 7.12. *Si* $\varphi, \psi \in \mathcal{L}(E)$, *alors* $\det(\varphi \circ \psi) = \det \varphi \, \det \psi$.

Démonstration. Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. D'après les définition et le théorème 7.11,

$$\begin{aligned} \det(\varphi \circ \psi) &= \det_{\mathcal{B}}(\varphi \circ \psi(e_1), \ldots, \varphi \circ \psi(e_n)) \\ &= \det_{\mathcal{B}}(\varphi(\psi(e_1)), \ldots, \varphi(\psi(e_n))) \\ &= \det(\varphi) \, \det_{\mathcal{B}}(\psi(e_1), \ldots, \psi(e_n)) \\ &= \det(\varphi) \, \det(\psi). \end{aligned}$$

□

Proposition 7.13. *Soit E un espace vectoriel de dimension $n \geq 1$ et soit $\varphi \in \mathcal{L}(E)$. L'endomorphisme φ est bijectif (c'est-à-dire est un automorphisme) si et seulement si* $\det(\varphi) \neq 0$.

De plus, si φ est un automorphisme, alors $\det(\varphi^{-1}) = 1/\det(\varphi)$.

Démonstration. L'application φ est un automorphisme de E si, et seulement si, la famille $(\varphi(e_1), \ldots, \varphi(e_n))$ est libre. D'après la proposition 7.5, la famille est libre si, et seulement si, $\det_{\mathcal{B}}(\varphi(e_1), \ldots, \varphi(e_n)) \neq 0$. Mais

$$\det_{\mathcal{B}}(\varphi(e_1), \ldots, \varphi(e_n)) = \det(\varphi),$$

donc φ est un automorphisme si et seulement si $\det(\varphi) \neq 0$.

Si φ est un automorphisme, on a $\varphi \circ \varphi^{-1} = Id_E$. En appliquant le proposition 7.12,

$$\det(\varphi) \, \det(\varphi^{-1}) = \det(Id_E) = 1.$$

□

Exemple. L'application $f : \mathbb{R}^3 \to \mathbb{R}^3$, $(x, y, z) \mapsto (-x + y + z, x - y + z, x + y - z)$ est un automorphisme car la matrice de f dans la base canonique est

$$\begin{vmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{vmatrix}$$

et son déterminant est égal à $4 \neq 0$.

À travers l'isomorphisme de $\mathbb{K}$-espaces vectoriels $\mathcal{L}(E) \to \mathcal{M}_n(\mathbb{K})$, isomorphisme qui commute avec la composition des endomorphismes et le produit des matrices, les propriétés de l'application déterminant pour les endomorphismes sont transportées pour les matrices.

Corollaire 7.14. *Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ et soit $\lambda \in \mathbb{K}$. Alors*
(i) $\det(\lambda A) = \lambda^n \det(A)$ *(attention à l'exposant)*
(ii) $\det(AB) = \det(A) \det(B)$
(iii) *A est inversible si et seulement si* $\det(A) \neq 0$
(iv) *si A est inversible, alors* $\det(A^{-1}) = \frac{1}{\det(A)}$.

Démonstration. Le premier point est établi par récurrence en utilisant la formule de décomposition du déterminant. Pour les autres points, il suffit de considérer les endomorphismes φ et ψ de $\mathbb{K}^n$ qui ont pour matrices respectives A et B dans la base canonique de $\mathbb{K}^n$ et d'appliquer les propositions précédentes. $\square$

8. Matrices carrées inversibles et déterminants

8.1. Formules pour les coefficients de A^{-1} quand A est inversible

Dans l'exemple de la page 55, on a introduit une méthode algorithmique pour calculer A^{-1}, où A est une matrice inversible ; la méthode consiste à inverser le système $Ax = y$ associé à A. Dans cette section on présente les formules théorique pour les coefficients de A.

On rappelle que pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ et pour tous $1 \leq i, j \leq n$, on note par $A_{i,j}$ la matrice obtenue en supprimant la ligne i et la colonne j de A. Le cofacteur de place (i, j), ou de $a_{i,j}$, est l'expression $(-1)^{i+j} \det(A_{i,j})$.

Définition. Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$. La matrice $\operatorname{co}(A) = \big((-1)^{i+j} \det(A_{i,j})\big) \in \mathcal{M}_n(\mathbb{K})$ est appelée la *comatrice* de A.

Lemme 8.1. *Soit $A \in \mathrm{Mat}_n(\mathbb{K})$ et soient $j \neq k$. Alors la somme des produits des coefficient de la colonne j par les cofacteurs des coefficients de la colonne k est nulle, c'est-à-dire*

$$\sum_{i=1}^{n} (-1)^{i+k} a_{i,j} \det(A_{i,k}) = 0.$$

Démonstration. Soit A' la matrice obtenue en remplaçant dans A la k-ème colonne par la j-ème colonne ($k \neq j$). On a $\det(A') = 0$. Or les cofacteurs des coefficients de la k-ème colonne de A' ne dépendent pas de ces coefficients et par conséquent, ce sont aussi les cofacteurs des coefficients de la k-ème colonne de A. On obtient la formule

$$0 = \det(A') = \sum_{i=1}^{n} (-1)^{i+k} a_{i,j} \det(A_{i,k})$$

en développant $\det(A')$ suivant la k-ème colonne. □

Par transposition, le même énoncé est vrai pour les lignes de A, c'est-à-dire si $i \neq k$, alors

$$\sum_{j=1}^{n} (-1)^{j+k} a_{i,j} \det(A_{k,j}) = 0.$$

Proposition 8.2. *Si $A \in \mathcal{M}_n(\mathbb{K})$, alors $A \cdot \mathrm{co}(A^T) = \mathrm{co}(A^T) \cdot A = \det(A)\, I_n$, où $\mathrm{co}(A^T)$ est la comatrice de la transposée de A. De plus, si A est inversible, alors*

$$A^{-1} = \frac{1}{\det(A)}\ \mathrm{co}(A^T). \tag{8.1}$$

Démonstration. On calcule les coefficients de la matrice produit $M = (m_{i,k}) = A \cdot \mathrm{co}(A^T)$. Si on pose $\mathrm{co}(A^T) = (b_{j,k})$, alors $b_{j,k} = (-1)^{j+k} \det((A^T)_{j,k}) = (-1)^{j+k} \det(A_{k,j})$ et donc

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

D'après la formule du lemme 8.1 (lue pour les lignes) et la formule de décomposition du déterminant d'après une ligne, il s'ensuit que

$$m_{i,k} = \begin{cases} \det(A) & \text{si } i = k \\ 0 & \text{si } i \neq k. \end{cases}$$

Donc $A \cdot \mathrm{co}(A^T) = \det(A)\, I_n$. On prouve de même que $\mathrm{co}(A^T) \cdot A = \det(A)\, I_n$.

Pour finir, on note que la formule pour A^{-1} devient évidente quand A est inversible. □

Exemple. Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible, alors $A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ d'après la formule (8.1), car $ad - bc \neq 0$. Maintenant si $A = (a_{i,j})$ est une matrice de taille 3×3 inversible, alors, d'après cette même formule,

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} & -\begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} & \begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{2,2} & a_{2,3} \end{vmatrix} \\ -\begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} & \begin{vmatrix} a_{1,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} & -\begin{vmatrix} a_{1,1} & a_{1,3} \\ a_{2,1} & a_{2,3} \end{vmatrix} \\ \begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} & -\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} & \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} \end{pmatrix}.$$

Du point de vue pratique, la formule (8.1) n'est pas très efficace car il faut calculer, en général, le déterminant de A puis n^2 déterminants d'ordre $n - 1$. On pourra l'utiliser de manière satisfaisante dans les cas $n = 2$ ou 3 comme indiqué dans l'exemple précédent.

Par la suite, on présente une troisième méthode d'inversion, celle de Gauss-Jordan. Elle est de fait la plus efficace et est utilisée dans les algorithmes d'algèbre linéaire pour le calcul de l'inverse. Cette méthode consiste à résoudre simultanément par la méthode de Gauss les systèmes linéaires

$$AX = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad AX = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad AX = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Pour cela on concatène les matrices A et I_n, puis on met en œuvre la méthode du pivot sur les lignes (entières) jusqu'à obtenir à gauche la matrice identité. Il ne faut pas se contenter de faire apparaître des zéros sous la diagonale mais aussi au-dessus de celle-ci. Ce que l'on obtient à la fin du processus à droite de I_n est A^{-1}.

Par exemple, pour calculer l'inverse de la matrice

$$A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix},$$

on applique l'algorithme de Gauss-Jordan et on obtient successivement.

$$\left(\begin{array}{ccc:ccc} -1 & 1 & 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc:ccc} 1 & -1 & -1 & -1 & 0 & 0 \\ 1 & -1 & 1 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc:ccc} 1 & -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{ccc:ccc} 1 & -1 & -1 & -1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 & 1 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc:ccc} 1 & -1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{ccc:ccc} 1 & 0 & -1 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right)$$

On en déduit $A^{-1} = \frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

Calcul du rang d'une matrice. Soit $A = (a_{i,j}) \in \mathcal{M}_{p,n}(\mathbb{K})$ et soit k un entier tel que $1 \leq k \leq \inf(p, n)$. On appelle *déterminant d'ordre k extrait de A*, le déterminant de toute matrice carrée obtenue en supprimant dans A, $p-k$ lignes et $n-k$ colonnes, c'est-à-dire tout déterminant de la forme

$$\begin{vmatrix} a_{i_1,j_1} & \cdots & a_{i_1,j_k} \\ \vdots & & \vdots \\ a_{i_k,j_1} & \cdots & a_{i_k,j_k} \end{vmatrix},$$

où $1 \leq i_1 < i_2 < \cdots < i_k \leq p$ et $1 \leq j_1 < j_2 < \cdots < j_k \leq n$. Par exemple, si

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{pmatrix}.$$

Alors

$$\begin{vmatrix} a_{1,1} & a_{1,4} \\ a_{3,1} & a_{3,4} \end{vmatrix} \qquad \text{et} \qquad \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,4} \end{vmatrix}$$

sont des déterminants d'ordre 2 et, respectivement, 3 extraits de A.

Théorème 8.3. *Soit $A = (a_{i,j}) \in \mathcal{M}_{p,n}(\mathbb{K})$ et soit ρ le plus grand entier $k \leq \inf(p,n)$ tel qu'il existe un déterminant d'ordre k extrait de A non nul. Alors* $\mathrm{rg}(A) = \rho$.

Démonstration. Si $A = 0$, c'est évident. Soit $A = (a_{i,j}) \in \mathcal{M}_{p,n}(\mathbb{K})$ une matrice non nulle et soit $r = \mathrm{rg}(A)$. On a $1 \leq r \leq \inf(p,n)$ et $1 \leq \rho \leq \inf(p,n)$. On note par la suite $c_1, \dots, c_n$ les colonnes de A.

On affirme que $r \geq \rho$. En effet, soit

$$A_\rho = \begin{pmatrix} a_{i_1,j_1} & \cdots & a_{i_1,j_\rho} \\ \vdots & & \vdots \\ a_{i_\rho,j_1} & \cdots & a_{i_\rho,j_\rho} \end{pmatrix}$$

une matrice carrée d'ordre ρ —obtenue en supprimant $p - \rho$ lignes et $n - \rho$ colonnes de A —telle que $\Delta_\rho = \det(A_\rho) \neq 0$. Il s'ensuit que les colonnes de A_ρ sont libres, donc que les colonnes $c_{j_1}, \dots, c_{j_\rho}$ de A sont libres, c'est-à-dire que $r = \mathrm{rg}(A) \geq \rho$.

On affirme que $\rho \geq r$. Pour justifier cette inégalité, soit $(c_{j_1}, \dots, c_{j_r})$ une famille ordonnée libre extraite de la famille ordonnée $(c_1, \dots, c_n)$, où $1 \leq j_1 < j_2 < \cdots < j_r \leq n$. On considère la matrice $C = \begin{pmatrix} c_{j_1} & \cdots & c_{j_r} \end{pmatrix} \in \mathcal{M}_{p,r}(\mathbb{K})$. Comme $\mathrm{rg}(C) = r$, on peut extraire r lignes linéairement indépendantes de la famille des lignes de C. Si A_r est la matrice carrée d'ordre r extraite de C, et donc de A. formée avec ces lignes, alors $\det(A_r) \neq 0$. Par conséquent $\rho \geq r$. $\square$

En pratique, pour calculer le rang d'une matrice, l'ancienne méthode basée sur l'algorithme du pivot de Gauss est meilleure, c'est-à-dire plus performante (d'ailleurs comme pour le calcul de l'inverse d'une matrice).

8.2. Systèmes de Cramer

Soit n un entier ≥ 1. On considère un système linéaire à coefficients dans $\mathbb{K}$ ayant n inconnues x_1, x_2,…,x_n et n équations

$$(S) \quad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,n}x_n = b_1 \\ \quad \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \ldots + a_{n,n}x_n = b_n. \end{cases}$$

Il peut se mettre sous la forme matricielle $AX = B$ où $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{K})$ est une matrice carrée, et $X = (x_i)_{1 \leq i \leq n} \in \mathcal{M}_{n,1}(\mathbb{K})$ et $B = (b_i)_{1 \leq i \leq n} \in \mathcal{M}_{n,1}(\mathbb{K})$ sont des matrices colonne. Le déterminant $\det(A)$ s'appelle le déterminant du système.

Le système (S) est dit de Cramer si A est inversible autrement dit si $\det(A) \neq 0$. Dans ce cas, on peut multiplier par A^{-1} à gauche et on obtient la solution unique $X^* = A^{-1}B$. Les formule explicites pour les coefficients de A^{-1} nous permettent de donner des formules utilisant les déterminants pour la solution X^*.

Proposition 8.4 (Cramer)**.** *Si A est inversible, alors le système (de Cramer) $AX = B$ admet une solution unique donnée par les formules de Cramer,*

$$x_k^* = \frac{\Delta_k}{\det(A)}, \quad \textit{où} \quad \Delta_k = \begin{vmatrix} a_{1,1} & \dots & a_{1,k-1} & b_1 & a_{1,k+1} & \dots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,k-1} & b_n & a_{n,k+1} & \dots & a_{n,n} \end{vmatrix},$$

pour tout $k = 1, \dots, n$. (Le déterminant Δ_k correspond à la matrice obtenue en remplaçant dans A la k-ème colonne par la colonne B.)

Démonstration. Soient $C_1, \dots, C_n$ les colonnes de A et soit $\mathcal{B}$ la base canonique de $\mathbb{K}^n$. On a $\det(A) = \det_{\mathcal{B}}(C_1, \dots, C_n)$.

Puisque X^* est la solution du système, on a $AX^* = B$, c'est-à-dire

$$\sum_{i=1}^{n} C_i x_i^* = B.$$

Alors, pour tout $1 \leq k \leq n$,

$$\begin{aligned} \Delta_k &= \det_{\mathcal{B}}(C_1, \dots, C_{k-1}, B, C_{k+1}, \dots, C_n) \\ &= \sum_{i=1}^{n} x_i^* \det_{\mathcal{B}}(C_1, \dots, C_{k-1}, C_i, C_{k+1}, \dots, C_n) \\ &= x_i^* \det(A). \end{aligned}$$

□

Remarque. Dans le cas d'un système linéaire homogène de Cramer, c'est-à-dire quand A est inversible et $B = 0$, on a

$$AX = 0 \quad \text{si et seulement si} \quad X = 0.$$

(En effet, on sait qu'il y a une unique solution, de plus la solution nulle est évidente donc c'est la bonne.)

Exemple. On considère le système

$$\begin{cases} -x + y + z = 1 \\ x - y + z = 1 \\ x + y - z = 1. \end{cases}$$

On a $A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$ et on a déjà vu que $\det(A) = 4$. Donc la solution a les composantes

$$x_1 = \frac{1}{4} \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{vmatrix} = 1, \quad x_2 = \frac{1}{4} \begin{vmatrix} -1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \end{vmatrix} = 1, \quad x_3 = \frac{1}{4} \begin{vmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 1.$$

En pratique, dès que $n \geq 4$, il faut remarquer que les formules de Cramer ne sont pas efficaces car il faut calculer un déterminant d'ordre n ainsi que n déterminants d'ordre $n - 1$.

A. Déterminant de Sylvester

Soit $F = a_m X^m + \cdots + a_1 X + a_0$ un polynôme à coefficients dans $\mathbb{C}$ de degré m. Pour tout $d \geq 0$, on associe à F l'application linéaire

$$\mu_d = \mu_d^F : \mathbb{C}[X]_{d-1} \longrightarrow \mathbb{C}[X]_{d+n-1}, \qquad P \mapsto FP.$$

Dans les bases $(X^{d-1}, \ldots, X, 1)$ et, respectivement, $(X^{d+n-1}, \ldots, X, 1)$, la matrice de μ_d est égale à

$$M_d = M_d^F = \begin{pmatrix} a_n & 0 & 0 & \cdots & 0 & 0 \\ a_{n-1} & a_n & 0 & \cdots & 0 & 0 \\ a_{n-2} & a_{n-1} & a_n & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_0 & a_1 \\ 0 & 0 & 0 & \cdots & 0 & a_0 \end{pmatrix} \in \mathcal{M}_{d+n,n}(\mathbb{C}).$$

Il est évident que $\mathrm{rg}(M_d) = d$.

Soient $F, G \in \mathbb{C}[X]$ deux polynômes de degrés m et n respectivemnt. On définit l'application linéaire $\varphi : \mathbb{C}[X]_{n-1} \times \mathbb{C}[X]_{m-1} \to \mathbb{C}[X]_{m+n-1}$ associée à ces deux polynômes, définie par

$$(P, Q) \mapsto [\varphi] PF + QG.$$

Il est utile de remarquer que

$$\varphi((P, Q)) = \mu_{n-1}^F(P) + \mu_{n-1}^G(Q).$$

Le domaine et le codomaine sont des sous-espaces vectorield de dimension $m + n$. La matrice carrée de φ dans les bases $\big((X^{n-1}, 0), \ldots, (X, 0), (1, 0), (0, X^{m-1}), \ldots, (0, X), (0, 1)\big)$ et $(X^{m+n-1}, \ldots, X, 1)$ est

$$S(F, G) = \begin{pmatrix} M_{n-1}^F & M_{m-1}^G \end{pmatrix} \in \mathcal{M}_{m+n,m+n}(\mathbb{C})$$

et est appelée la matrice de Sylvester de F et G.

Théorème A.1. *Soient $F, G \in \mathbb{C}[X]$ de degrés m et n respectivement. Les polynômes F et G ont une racine commune si et seulement si l'application de Sylvester φ n'est pas un isomorphisme, c'est-à-dire si et seulement si*

$$\det(S(F, G)) = 0.$$

Démonstration. Si F et G ont une racine commune $\alpha \in \mathbb{C}$, alors $F = (X - \alpha)F_0$ et $G = (X - \alpha)G_0$. Il s'ensuit que

$$\varphi(G_0, -F_0) = G_0 F - F_0 G = (X - \alpha)(G_0 F_0 - F_0 G_0) = 0.$$

Si F et G sont premier entre eux, alors en utilisant la relation de Bezout, il existe U et V de degrés $\leq n - 1$ et, respectivement, $\leq m - 1$ tels que $UF + VG = 1$. En utilisant la division euclidienne, on montre que $(XU, XV) \mapsto X$ et ainsi de suite, c'est-à-dire que φ est un isomorphisme. $\square$

Définition A.2. L'expression $r(F,G) = \det(S(F,G))$ s'appelle le *résultant des polynômes F et G*.

Exemple. Si $F = aX^2 + bX + c$ et $G = F'$, alors

$$r(F,F') = \begin{vmatrix} a & 2a & 0 \\ b & b & 2a \\ c & 0 & b \end{vmatrix} = ab^2 + 4a^2c - 2ab^2 = a(4ac - b^2) = -a\varDelta$$

où $\varDelta$ est le discriminant de la forme quadratique F. On retrouve le résultat bien connu suivant : l'équation $aX^2 + bX + c = 0$, $a \neq 0$, admet une racine double si et seulement si $\varDelta = b^2 - 4ac = 0$.

Exemple. Si $F = aX^3 + bX^2 + cX + d$ et $G = F'$, alors

$$r(F,F') = \begin{vmatrix} a & & 3a & & \\ b & a & 2b & 3a & \\ c & b & c & 2b & 3a \\ d & c & & c & 2b \\ & d & & & c \end{vmatrix} = a \begin{vmatrix} a & 2b & 3a & \\ b & c & 2b & 3a \\ c & & c & 2b \\ d & & & c \end{vmatrix} + 3a \begin{vmatrix} b & a & 3a & \\ c & b & 2b & 3a \\ d & c & c & 2b \\ & d & & c \end{vmatrix}.$$

Le premier déterminant 4×4, en développant d'après la première ligne, est égale à

$$\begin{aligned} D_1 &= a \begin{vmatrix} c & 2b & 3a \\ & c & 2b \\ & & c \end{vmatrix} - 2b \begin{vmatrix} b & 2b & 3a \\ c & c & 2b \\ d & & c \end{vmatrix} + 3a \begin{vmatrix} b & c & 3a \\ c & & 2b \\ d & & c \end{vmatrix} \\ &= ac^3 - 2b(bc^2 + 4b^2d - 3acd - 2bc^2) + 3a(2bcd - c^3) \\ &= 2(6abcd - ac^3 - 4b^3d + b^2c^2). \end{aligned}$$

Pour le deuxième on obtient

$$\begin{aligned} D_2 &= b \begin{vmatrix} b & 2b & 3a \\ c & c & 2b \\ d & & c \end{vmatrix} - a \begin{vmatrix} c & 2b & 3a \\ d & c & 2b \\ & & c \end{vmatrix} + 3a \begin{vmatrix} c & b & 3a \\ d & c & 2b \\ & d & c \end{vmatrix} \\ &= b(bc^2 + 4b^2d - 3acd - 2bc^2) - a(c^3 - 2bcd) + 3a(c^3 + 3ad^2 - 2bcd - bcd) \\ &= -b^2c^2 + 4b^3d - 3abcd - ac^3 + 2abcd + 3ac^3 + 9a^2d^2 - 9abcd \\ &= 9a^2d^2 - 10abcd + 2ac^3 + 4b^3d - b^2c^2. \end{aligned}$$

Donc

$$r(F,G) = a(27a^2d^2 - 18abcd + 4ac^3 + 4b^3d - b^2c^2).$$