ENS Paris-Saclay
Préparation à l'agrégation externe de sciences sociales

Éléments d'algèbre linéaire

Pierre Montagnon, Simon Coste et Rafael Grompone von Gioi

L'objet de l'algèbre est l'étude de *structures* algébriques, c'est-à-dire d'ensembles auxquels on confère une certaine structure en les munissant d'opérations.

L'objet principal de l'algèbre linéaire est l'étude de la structure d'espace vectoriel : cette notion a été créée pour modéliser la notion d'espace, en s'inspirant de l'espace qui nous entoure, $\mathbb{R}^3$. La première partie du cours s'intéresse aux propriétés générales des espaces vectoriels ainsi définis. On poursuit par l'étude des applications qui conservent la structure d'espace vectoriel, les applications linéaires. On peut s'intéresser à ces applications sous plusieurs points de vue : celui des fonctions, celui des systèmes linéaires, ou celui des espaces auxquels elles s'appliquent. Ces trois aspects font de l'algèbre linéaire un outil fondamental des mathématiques modernes, dans la mesure où elle [1] apporte à la fois un éclairage théorique sur des structures mathématiques et des moyens extrêmement pratiques et efficaces pour résoudre des problèmes simples.

Pour toute question ou remarque sur ce cours, vous pouvez contacter les auteurs aux adresses suivantes :

simon@scoste.fr
pierre.montagnon@ens-cachan.fr

1. Notez que le mot « algèbre » est féminin, contrairement à une croyance répandue jusque dans les milieux universitaires.

Table des matières

1 Espaces vectoriels et applications linéaires.

1.1 Espaces vectoriels, sous-espaces vectoriels

Les espaces vectoriels constituent le concept central du programme d'algèbre. Leur définition, que nous verrons bientôt, peut sembler longue ou tortueuse, mais l'idée qu'il y a derrière est très simple :

Un espace vectoriel est un ensemble dont on peut additionner les éléments entre eux et les multiplier par des nombres réels.

Plusieurs ensembles bien connus correspondent à cette définition : $\mathbb{R}$ bien entendu, mais aussi l'ensemble $\mathscr{F}(\mathbb{R},\mathbb{R})$ des fonctions définies sur $\mathbb{R}$ et à valeurs réelles, l'ensemble $\mathbb{R}[X]$ des polynômes introduit dans le chapitre précédent, $\mathbb{R}^2$, $\mathbb{R}^3$...

Par analogie avec les espaces $\mathbb{R}^2$ et $\mathbb{R}^3$, les éléments d'un espace vectoriel [2] s'appellent des *vecteurs*. Ainsi, le nombre 2 est un vecteur de l'espace $\mathbb{R}$, le polynôme $1+X^7$ un vecteur de l'espace des polynômes à coefficients réels $\mathbb{R}[X]$ et la fonction exp un vecteur de $\mathscr{F}(\mathbb{R},\mathbb{R})$.

La définition d'un espace vectoriel n'impose pas que l'on puisse multiplier ou diviser des vecteurs entre eux : parfois, la multiplication des vecteurs n'est pas définie [3], comme c'est le cas dans $\mathbb{R}^2$!

Ce que l'on appelle « additionner des vecteurs » correspond en fait à une « loi de composition interne [4] » : si l'on me donne deux vecteurs x et y, je suis capable de les additionner suivant cette loi pour former un nouveau vecteur appelé $x+y$. On s'attend à ce que l'addition soit une opération très facile à réaliser et qu'elle présente des propriétés identiques à celles de l'addition réelle. En particulier, on souhaite qu'elle vérifie les propriétés suivantes :

- Elle doit admettre un élément « neutre », c'est-à-dire qu'il doit exister un élément tel que l'additionner à n'importe quel autre vecteur ne change rien. C'est le *zéro* de l'addition.
- Elle ne doit pas dépendre de l'ordre dans lequel on l'effectue. Autrement dit, $x+y$ doit être le même vecteur que $y+x$.
- Elle ne doit pas dépendre de l'ordre dans lequel on effectue d'éventuelles additions successives : autrement dit, $(x+y)+z$ (on additionne d'abord x et y, puis on rajoute z au résultat) doit être égal à $x+(y+z)$ (on additionne d'abord y et z, puis on ajoute x au résultat).

2. Le terme « espace vectoriel » se réfère bien entendu aux origines géométriques de cette notion, même si la présentation que nous faisons de l'algèbre linéaire n'aborde directement aucune question de nature géométrique.

3. Les espaces dans lesquels on a le droit d'additionner les éléments, de multiplier par des réels et de multiplier des éléments entre eux sont appelés des *algèbres* et ne sont pas au programme.

4. Ce mot savant ne cache pas une réalité très sophistiquée : il s'agit tout simplement d'une fonction notée + qui transforme un couple d'éléments de l'espace E en leur somme, c'est-à-dire d'une fonction

$$+ : \begin{cases} E \times E \to E \\ (x,y) \mapsto x+y \end{cases}$$

Ces trois propiétés sont appelées respectivement *existence d'un élément neutre, commutativité* et *associativité*.

Ensuite, il faut pouvoir *multiplier un vecteur par un nombre réel* : par exemple, multiplier un vecteur par 17. Cela s'appelle une « loi de composition externe [5] », parce que l'on multiplie un vecteur par quelque chose qui n'est pas un vecteur (qui est « externe » à l'espace vectoriel) : un réel, aussi appelé « scalaire ». Cette opération doit là encore obéir à des règles simples pour qu'il soit possible de réaliser des calculs aussi facilement que dans $\mathbb{R}$:

- Par exemple, multiplier par λ puis par μ doit revenir à multiplier par $(\mu \times \lambda)$.
- Il doit y avoir une forme de « compatibilité » entre l'addition et la multiplication externe. Par exemple, il faut pouvoir développer par rapport aux vecteurs, mais aussi par rapport aux scalaires :

$$\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y \quad \text{et} \quad (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot y$$

- Il faut s'assurer que $1 \cdot x = x$ pour tout vecteur x.

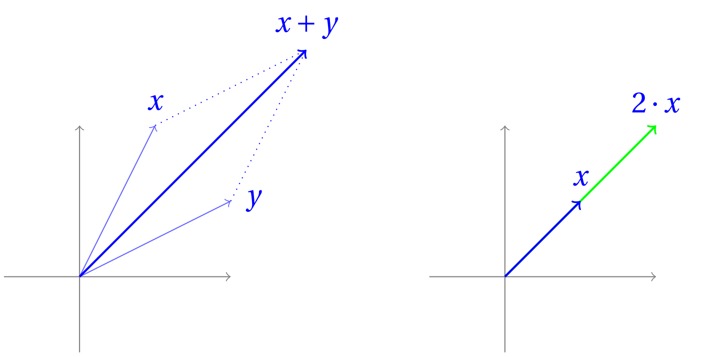
Ces diverses exigences conduisent enfin à la définition d'un espace vectoriel :

Définition 1.1 (Espace vectoriel)**.** — *Soit* E *un ensemble. On dit que* E *est un espace vectoriel sur* $\mathbb{R}$ *(ou plus simplement que* E *est un espace vectoriel) lorsque* E *est muni d'une loi de composition interne notée* $+$ *telle que :*

- *il existe un élément neutre, noté* 0*, tel que pour tout* $x \in \mathrm{E}$, $x + 0 = 0 + x = x$ *(élément neutre)*
- *pour tout* $x \in \mathrm{E}$ *et tout* $y \in \mathrm{E}$, $x + y = y + x$ *(commutativité)*
- *pour tout* $x \in \mathrm{E}, y \in \mathrm{E}, z \in \mathrm{E}$, $(x + y) + z = x + (y + z)$ *(associativité)*

et d'une loi de composition externe notée $\cdot$ *telle que pour tous* $x \in \mathrm{E}, y \in \mathrm{E}$ *et pour tous scalaires* $\lambda, \mu \in \mathbb{R}$*, on ait :*

$$\begin{array}{cc} \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y & (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x \\ (\lambda\mu) \cdot x = \lambda \cdot (\mu \cdot x) & 1 \cdot x = x \end{array}$$



Addition vectorielle   Multiplication externe

FIGURE 1 – Opérations vectorielles.

5. Qui correspond cette fois à une fonction

$$\cdot : \begin{cases} \mathbb{R} \times \mathrm{E} \to \mathrm{E} \\ (\lambda, x) \mapsto \lambda \cdot x \end{cases}$$

La loi de composition externe est notée avec un $\cdot$: par exemple, on écrit $\lambda \cdot x$. Dans la suite du cours, nous omettrons parfois le symbole $\cdot$ et écrirons simplement λx sans risque de créer d'ambiguïté. L'élément neutre pour l'addition d'un espace vectoriel E est parfois noté 0_E pour le distinguer du 0 réel. Par exemple, $0_{\mathbb{R}^3} = (0, 0, 0)$.

Exercice 1.2. Vérifier que $\mathscr{F}(\mathbb{R}, \mathbb{R})$ est un espace vectoriel et déterminer son élément neutre.

EXEMPLE 1.3 (fondamental)**.** $\mathbb{R}^n$, muni de l'addition et de la multiplication par un scalaire, est un espace vectoriel. C'est d'ailleurs l'espace de référence au programme de l'agrégation.

EXEMPLE 1.4. L'ensemble des polynômes à coefficients réels $\mathbb{R}[X]$ est un espace vectoriel [6].

REMARQUE 1.5 (Un peu d'histoire)**.** Les mathématiciens mettaient déjà en œuvre des techniques vectorielles dans le plan ou dans l'espace dans l'Antiquité. Il fallut pourtant attendre le XIXème siècle pour que des mathématiciens comme le britannique Arthur Cayley cherchent à rendre le cadre théorique de ces manipulations vectorielles à la fois plus rigoureux et plus général : l'objectif était de généraliser les notions de *plan* ($\mathbb{R}^2$) et d'*espace* ($\mathbb{R}^3$), pour manipuler des « espaces » possédant 4 dimensions, ou plus.

FIGURE 2 – Hermann G. Grassmann (1807 - 1877), mathématicien allemand.

Un des pionniers dans l'axiomatisation de l'algèbre linéaire fut l'allemand Hermann G. Grassmann [7], mais il fut ignoré de son vivant. Ce fut Giuseppe Peano [8] qui donna la définition définitive des espaces vectoriels, celle qui est exposée dans ce cours, vers la fin des années 1880.

DÉFINITION 1.6 (Sous-espace vectoriel)**. —** *Soit* E *un espace vectoriel et* F *une partie de* E*. On dit que* F *est un sous-espace vectoriel de* E *lorsque :*

6. Lorsque cela n'est pas précisé, les lois de composition choisies sont celles qui définissent habituellement la somme et le produit par un scalaire sur les espaces considérés. Ici, il s'agit par exemple des opérations définies dans le chapitre sur les polynômes.

7. Grassmann (1807 - 1877), qui était professeur de lycée. On lui doit des approches tout à fait nouvelles sur les espaces vectoriels : il fait peut-être partie des premiers mathématiciens à avoir dégagé le concept abstrait d'espace vectoriel, indépendamment des cas particuliers de $\mathbb{R}^2$ et $\mathbb{R}^3$. Desservi par une mauvaise position universitaire et peu rigoureux dans ses écrits, Grassmann fut ignoré de ses contemporains en tant que mathématicien et connut le succès... comme traducteur du *Rig-Véda*, un texte religieux de l'Inde antique écrit en Sanskrit.

8. Mathématicien italien (1858 - 1932).

- 0 *appartient à* F
- F *est stable par addition, c'est-à-dire que pour tout* $x \in F$, $y \in F$, *on a* $x + y \in F$
- F *est stable par multiplication externe, c'est-à-dire que pour tout scalaire* λ *et tout élément* $x \in F$, *on a* $\lambda \cdot x \in F$.

PROPOSITION 1.7. — *Soit* F *un sous-espace vectoriel d'un espace vectoriel. Alors* F *est également un espace vectoriel.*

REMARQUE 1.8. C'est cette proposition que l'on utilisera le plus souvent en pratique pour montrer qu'un ensemble donné est un espace vectoriel : il suffit en effet de montrer que c'est un sous-espace vectoriel d'un espace vectoriel connu, ce qui permet de ne vérifier que les trois propriétés de la définition d'un sous-espace vectoriel au lieu de revenir à la définition un peu pénible d'un espace vectoriel. Le gain de temps réalisé s'explique par le fait que les sous-espaces vectoriels « héritent » des opérations définies sur les espaces vectoriels dans lesquels ils se situent et qu'il n'est donc pas nécessaire de vérifier les propriétés calculatoires de ces opérations une nouvelle fois.

La méthode suivante est celle utilisée en pratique pour établir qu'un ensemble donné est un (sous-)espace vectoriel.

MÉTHODE 1.9. Pour montrer qu'un ensemble F est un sous-espace vectoriel d'un espace E, on vérifie que :
- F est une partie de E
- 0 appartient à F
- F est stable par combinaison linéaire, c'est-à-dire que l'on a :

$$\forall x, y \in F, \quad \forall \lambda \in \mathbb{R}, \quad x + \lambda y \in F$$

REMARQUE 1.10. Il arrive que les problèmes de concours commencent par vous faire démontrer qu'une partie F d'un espace vectoriel E est un sous-espace vectoriel. Dans ce cas, pensez bien à mentionner que F contient 0 et qu'il est inclus dans E.

Exercice 1.11. Soit F l'ensemble des points du plan de la forme $\begin{pmatrix} x \\ 0 \end{pmatrix}$ pour $x \in \mathbb{R}$. Montrer que F est un sous-espace vectoriel de $\mathbb{R}^2$.

Exercice 1.12. Soit F l'ensemble des points du plan de coordonnées $\begin{pmatrix} x \\ 1 \end{pmatrix}$ pour $x \in \mathbb{R}$. F est-il un sous-espace vectoriel de $\mathbb{R}^2$?

Exercice 1.13. Soit F l'ensemble :

$$F = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x + y + 2z = 0 \right\}$$

Montrer que F est un sous-espace vectoriel de $\mathbb{R}^3$.

EXEMPLE 1.14. On note $\mathscr{C}(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$. C'est un espace vectoriel.

Exercice 1.15. On note F l'ensemble des fonctions f continues de $\mathbb{R}$ dans $\mathbb{R}$ telles que $f(1) = 0$. Montrer que F est un sous-espace vectoriel de $\mathscr{C}(\mathbb{R}, \mathbb{R})$.

Exercice 1.16. [Équations différentielles] Soit $\mathscr{C}^1(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions continûment dérivables de $\mathbb{R}$ dans $\mathbb{R}$ (c'est-à-dire de classe $\mathscr{C}^1$ sur $\mathbb{R}$).

1. Montrer que c'est un espace vectoriel.
2. On considère l'ensemble $\mathscr{S}$ des fonctions dérivables y telles que $y' = ay + b$, où a et b sont des réels.
   (a) Montrer que $\mathscr{S} \subset \mathscr{C}^1$.
   (b) Montrer que $\mathscr{S}$ est un sous-espace vectoriel de $\mathscr{C}^1$.

PROPOSITION 1.17. — *Si* F *et* G *sont deux sous-espaces vectoriels d'un espace vectoriel* E*, alors* $F \cap G$ *est un sous-espace vectoriel de* E*. De façon plus générale, une intersection quelconque (éventuellement infinie) de sous-espaces vectoriels de* E *est un sous-espace vectoriel de* E*.*

DÉFINITION 1.18 (Combinaison linéaire)**. —** *Soit* E *un espace vectoriel et soient* $x_1, ..., x_n \in E$ *des vecteurs (avec* $n \geq 1$*). Une* combinaison linéaire *des* x_i *est un vecteur* $x \in E$ *s'écrivant sous la forme :*

$$x = \alpha_1 x_1 + ... + \alpha_n x_n = \sum_{k=1}^{n} \alpha_i x_i$$

où $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$*.*

DÉFINITION 1.19 (Sous-espace vectoriel engendré par une partie)**. —** *Soit* E *un espace vectoriel et* A *une partie de* E*. Le* sous-espace vectoriel de E engendré par A *est le plus petit sous-espace vectoriel de* E *(au sens de l'inclusion) contenant* A*, c'est-à-dire l'intersection de tous les sous-espaces vectoriels de* E *contenant* A*. On le note* vect(A)*. Si* $A = \{x_1, \ldots, x_n\}$ *(avec* $n \in \mathbb{N}^*$*) est une partie finie de* E*, on note simplement* $\mathrm{vect}(x_1, \ldots, x_n) = \mathrm{vect}(A)$*, que l'on appelle aussi* sous-espace vectoriel de E engendré par $x_1, \ldots, x_n$.

Exercice 1.20. Soit E un espace vectoriel. Quel est le sous-espace vectoriel de E engendré par $\{0\}$? Et par E ?

Exercice 1.21. On se place dans $\mathbb{R}^2$. Quel est le sous-espace vectoriel engendré par le vecteur $\begin{pmatrix}1\\2\end{pmatrix}$? Et par la famille $\left\{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right\}$?

Le théorème suivant est très important, puisqu'il donne une description exacte et explicite du sous-espace vectoriel engendré par une partie.

THÉORÈME 1.22. — *Soit* E *un espace vectoriel et* A *une partie de* E*. Alors,* vect(A) *est l'ensemble des combinaisons linéaires d'éléments de* A*.*

Exercice 1.23. Quel est le sous-espace de $\mathbb{R}[X]$ engendré par les polynômes $P = 1$ et $Q = X$?

1.2 Familles libres, familles liées, familles génératrices.

DÉFINITION 1.24 (Famille liée). — *Soit* $(x_i)_{1\leqslant i\leqslant n}$ *une famille de vecteurs d'un espace vectoriel* E. *On dit que cette famille est* liée *s'il existe des scalaires* $\alpha_1,...,\alpha_n$ non tous nuls *tels que*

$$\alpha_1 x_1 + ... + \alpha_n x_n = 0$$

Exercice 1.25. Quelle est la différence entre « non tous nuls » et « tous non nuls » ?

EXEMPLE 1.26. Soit x un vecteur d'un espace vectoriel E. La famille $\{x,x\}$ est liée : en effet,

$$1\cdot x + (-1)\cdot x = x - x = 0$$

De manière générale, si une famille de vecteurs contient deux fois (ou plus) le même vecteur, elle est liée.

REMARQUE 1.27. Dire que la famille $\{x_i\}_{i\in[\![1,n]\!]}$ est liée revient à dire qu'il existe un $i\in[\![1,n]\!]$ et des scalaires (cette fois-ci éventuellement tous nuls) $\lambda_1,\dots,\lambda_n$ tels que :

$$x_i = \sum_{\substack{k=1\\k\neq i}}^{n} \lambda_k x_k$$

Autrement dit, une famille est *liée* si l'un de ses éléments s'exprime comme combinaison linéaire des autres.

EXEMPLE 1.28. Si E est un espace vectoriel et si $n\in\mathbb{N}^*$, toute famille $(x_1,\dots,x_n)\in E^n$ contenant l'élément neutre 0 est liée puisque 0 s'exprime comme combinaison linéaire (à coefficients nuls) des autres vecteurs de la famille.

Exercice 1.29. Montrer que la famille

$$\left\{\begin{pmatrix}1\\2\end{pmatrix},\begin{pmatrix}2\\-1\end{pmatrix},\begin{pmatrix}0\\3\end{pmatrix}\right\}$$

est une famille liée de $\mathbb{R}^2$:

DÉFINITION 1.30 (Famille libre). — *Soit* $(x_i)_{i\in[\![1,n]\!]}$ *une famille de vecteurs d'un espace vectoriel* E. *On dit qu'elle est* libre *lorsqu'elle n'est pas liée. Autrement dit, elle est libre si et seulement si pour tout* $(\alpha_1,...,\alpha_n)\in\mathbb{R}^n$*, on a*

$$\left(\sum_{i=1}^{n} \alpha_i x_i = 0\right) \Longrightarrow \left(\alpha_1 = \alpha_2 = ... = \alpha_n = 0\right)$$

Exercice 1.31. Soient f_1, f_2 et f_3 les trois fonctions définies sur $\mathbb{R}$ par :

$$\forall x\in\mathbb{R},\quad f_1(x) = \sin x \quad f_2(x) = \sin(2x) \quad f_3(x) = \sin(3x)$$

Montrer que (f_1, f_2, f_3) est une famille libre de $\mathscr{F}(\mathbb{R},\mathbb{R})$.

Il importe de comprendre parfaitement cette définition ! La deuxième formulation mérite d'ailleurs qu'on s'y attarde : il s'agit de la négation de la proposition qui constitue la définition 1.24. Insistons : une famille est libre si et seulement si la seule combinaison linéaire nulle des éléments de cette famille est la combinaison dans laquelle tous les scalaires sont nuls.

EXEMPLE 1.32. Soient $n \in \mathbb{N}^*$ et $P_1, \ldots, P_n$ des polynômes à coefficients réels. Quitte à les renuméroter, on peut supposer que leurs degrés sont croissants. Si en plus on a :

$$0 \leq \deg(P_1) < \deg(P_2) < \ldots < \deg(P_n)$$

alors on dit que la famille $\{P_1, \ldots, P_n\}$ est *échelonnée en degré*. Cette famille de polynômes est alors une famille libre de $\mathbb{R}[X]$.

Exercice 1.33. Démontrer le résultat de l'exemple précédent.

De manière plus générale, on a le très utile théorème suivant :

THÉORÈME 1.34 (des familles échelonnées)**. —** *Soit* E *un espace vectoriel de dimension finie non nulle n. Soit* $r \in [\![1, n]\!]$ *et* $F_1, \ldots, F_r$ *des sous-espaces vectoriels de* E *tels que* $F_i \subsetneq F_{i+1}$ *pour tout* $i \in [\![1, r-1]\!]$*. On pose* $F_0 = \{0\}$*. Si* $(x_1, \ldots, x_r) \in E$ *vérifient :*

$$\forall i \in [\![1, r]\!], \quad x_i \in F_i \setminus F_{i-1}$$

alors $(x_1, x_2, \ldots, x_r)$ *est libre.*

Dans le cadre du théorème précédent, la famille $(x_1, \ldots, x_n)$ est dite *échelonnée*.

DÉFINITION 1.35 (Famille génératrice)**. —** *Soit* E *un espace vectoriel et* $x_1, \ldots, x_n$ *des vecteurs de* E *(avec* $n \in \mathbb{N}^*$*). La famille* $\{x_1, ..., x_n\}$ *est dite* génératrice de E *(ou simplement* génératrice *s'il n'y a pas d'ambiguïté) si tout élément de* E *peut s'écrire comme une combinaison linéaire des* x_i*.*

Dans le cadre de la proposition précédente, $\{x_1, ..., x_n\}$ est génératrice si et seulement si E est égale au sous-espace vectoriel de E engendré par cette famille, c'est-à-dire $E = \mathrm{vect}(x_1, \ldots, x_n)$.

EXEMPLE 1.36. On se place dans $\mathbb{R}^2$ et on pose :

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{et} \quad e_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Alors, la famille $\{e_1, e_2\}$ est génératrice. En effet, tout vecteur $\begin{pmatrix} x \\ y \end{pmatrix}$ s'écrit sous la forme

$$(x - y)e_1 + ye_2$$

Exercice 1.37. On se place dans $\mathbb{R}^n$ avec $n \geq 1$. On pose :

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Montrer que la famille $\{e_1, ..., e_n\}$ est génératrice (c'est la généralisation de l'exemple précédent).

Exemple 1.38. La famille $\{1, X, X^2, X^3\}$ de $\mathbb{R}_3[X]$ est une famille génératrice.

Définition 1.39 (Base)**.** — *Soient* E *un espace vectoriel,* $n \in \mathbb{N}^*$ *et* $x_1, ..., x_n$ *des vecteurs de* E*. On dit que la famille* $\{x_1, \ldots, x_n\}$ *est une* base de E *lorsqu'elle est à la fois libre et génératrice.*

Proposition 1.40. — *Soit* E *un espace vectoriel et soit* $n \in \mathbb{N}^*$*. La famille* $\{e_1, \ldots, e_n\}$ *de vecteurs de* E *est une base de* E *si et seulement si pour tout vecteur* $x \in$ E *il existe un* ***unique*** n*-uplet de réels* $(\alpha_1, ..., \alpha_n)$ *tels que*

$$x = \alpha_1 e_1 + \ldots + \alpha_n e_n$$

On attire l'attention du lecteur sur le fait que les scalaires α_i sont uniques, contrairement à ceux qui interviennent dans la définition d'une famille génératrice. Cette propriété est extrêmement importante.

Exercice 1.41. Montrer que la famille définie dans l'exercice 1.37 est une base de $\mathbb{R}^n$. On l'appelle *base canonique de* $\mathbb{R}^n$.

Théorème 1.42. — *Soit* $n \in \mathbb{N}$ *et soient* $P_0, ..., P_n$ *des polynômes. On suppose que pour tout* $i \in [\![0, n]\!]$ *le degré de* P_i *est exactement* i*. Alors, la famille* $\{P_0, ..., P_n\}$ *est une base de* $\mathbb{R}_n[X]$.

Exercice 1.43. Démontrer le théorème 1.42.

Remarque 1.44. Si E est un espace vectoriel, on considère par convention[9] que la famille vide $\varnothing$ est une base du sous-espace vectoriel trivial $\{0\}$.

1.3 Dimension d'un espace vectoriel

Définition 1.45 (Dimension finie)**.** — *Soit* E *un espace vectoriel. On dit qu'il est* de dimension finie *s'il admet une famille génératrice finie.*

Il existe des espaces vectoriels qui ne sont pas de dimension finie, comme par exemple $\mathbb{R}[X]$ ou $\mathscr{F}(\mathbb{R}, \mathbb{R})$.

Théorème 1.46. — *Soit* E *un espace vectoriel de dimension finie. Il existe un unique* $n \in \mathbb{N}$ *tel que toute famille de vecteurs de* E *comprenant* $n+1$ *éléments ou plus est liée, c'est-à-dire tel qu'une famille de vecteurs de* E *ne peut être libre que si elle comporte moins de* n *éléments.*

Attention, ce n'est évidemment pas parce qu'une famille comporte moins de n éléments qu'elle est libre !

Définition 1.47 (Dimension)**.** — *L'entier* n *dont on a déterminé l'existence et l'unicité dans le théorème précédent s'appelle* dimension *de* E *et est noté* dim(E).

Exemple 1.48. La dimension de l'espace vectoriel trivial $\{0\}$ est nulle.

9. ... ou parce que l'on considère que la somme d'un nombre nul d'éléments vaut 0 !

PROPOSITION 1.49. — *Si* F *est un sous-espace vectoriel d'un espace vectoriel* E *de dimension finie, alors* F *est de dimension finie et* $\dim(F) \le \dim(E)$.

THÉORÈME 1.50. — *Toutes les bases d'un espace vectoriel de dimension finie $n \in \mathbb{N}$ ont même cardinal n.*

EXEMPLE 1.51. Le sous-espace vectoriel de $\mathbb{R}^3$ défini par :

$$F = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x + y + z = 0 \right\}$$

admet $\left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$ pour base et est donc de dimension 2.

REMARQUE 1.52. Dans l'exemple 1.51, F est un plan de $\mathbb{R}^3$. Par analogie avec le cas des sous-espaces vectoriels de $\mathbb{R}^2$ et de $\mathbb{R}^3$, on appelle *droites* les espaces vectoriels de dimension 1 et *plans* les espaces vectoriels de dimension 2.

THÉORÈME 1.53. — *Pour tout $n \in \mathbb{N}^*$, la dimension de $\mathbb{R}^n$ est n.*

THÉORÈME 1.54. — *Pour tout $n \in \mathbb{N}$, la dimension de $\mathbb{R}_n[X]$ est $n+1$ et la famille* $\{1, X, \dots, X^n\}$ *forme une base de* $\mathbb{R}_n[X]$ *appelée* base canonique de $\mathbb{R}_n[X]$.

On peut penser à une base d'un espace vectoriel E comme à une collection minimale des « briques élémentaires » dont on peut se servir pour construire les éléments de E par combinaisons linéaires.

Exercice 1.55. Un espace vectoriel de dimension finie non nulle admet-il une unique base ? Une famille libre est-elle nécessairement génératrice ? Toutes les familles génératrices sont-elles libres ?

PROPOSITION 1.56. — *Soient* E *un espace vectoriel ainsi que* F *et* G *deux sous-espaces vectoriels de dimension finie de* E. *Si* $F \subset G$ *et si* $\dim(F) \geqslant \dim(G)$, *alors* $F = G$ *et leurs dimensions sont égales.*

La proposition précédente est très utile car elle permet de montrer que des espaces vectoriels sont égaux simplement en montrant une inclusion et en montrant l'égalité des (ou une inégalité sur les) dimensions.

Concluons cette partie avec un théorème d'une importance cruciale en algèbre linéaire :

THÉORÈME 1.57 (de la base incomplète). — *Si* E *est un espace vectoriel de dimension finie non nulle n et si $\mathscr{L} = \{e_1, \dots, e_p\}$ est une famille libre de* E *(avec $p \le n$), alors on peut compléter $\mathscr{L}$ en une base $\mathscr{B} = \{e_1, \dots, e_n\}$ de* E.

En termes moins formels, ce théorème affirme que toute famille libre d'un espace vectoriel de dimension finie peut être complétée en une base.

EXEMPLE 1.58. La famille $\{2, 1+X\}$ est une famille libre de l'espace vectoriel de dimension finie $\mathbb{R}^2[X]$. On peut donc la compléter en une base : par exemple, il suffit de rajouter le vecteur X^2 (pourquoi ?).

Les deux théorèmes suivants sont à la source de la majorité des résultats dans l'algèbre linéaire de dimension finie. Ils sont essentiels à plus d'un titre !

THÉORÈME 1.59. — *Tout espace vectoriel de dimension finie $n \in \mathbb{N}$ possède une base de cardinal n.*

THÉORÈME 1.60. — *Si E est un espace vectoriel de dimension $n \in \mathbb{N}$, une famille libre ou génératrice de E est une base si et seulement son cardinal est n.*

REMARQUE 1.61. Ce théorème est remarquable puisqu'il permet de démontrer qu'une famille de vecteurs d'un espace vectoriel de dimension connue est génératrice... en montrant qu'elle est libre ! Par exemple, la famille

$$\left\{\begin{pmatrix}0\\5\\-2\end{pmatrix}, \begin{pmatrix}1\\-\pi\\0\end{pmatrix}, \begin{pmatrix}0\\\sqrt{2}\\0\end{pmatrix}\right\}$$

est une base de $\mathbb{R}^3$ puisqu'elle est de cardinal 3 et qu'il est facile (exercice !) de montrer qu'elle est libre.

On remarque aussi qu'il est très facile de retrouver le résultat du théorème 1.42 en utilisant la liberté des familles à degrés échelonnés et le fait que $\dim(\mathbb{R}_n[X]) = n+1$ pour tout $n \in \mathbb{N}$. De manière plus générale, il permet de démontrer que toute famille échelonnée d'un espace vectoriel de dimension finie E qui est de cardinal $\dim(E)$ est une base de E.

DÉFINITION 1.62 (Rang d'une famille de vecteurs)**. —** *Soient $n \in \mathbb{N}^*$ et $\{u_1, \dots, u_n\}$ une famille de vecteurs d'un espace vectoriel E. On appelle* rang *de cette famille de vecteurs la dimension* $\dim\big(\mathrm{vect}(u_1, \dots, u_n)\big)$ *de l'espace vectoriel qu'ils engendrent.*

PROPOSITION 1.63. — *Soient $n \in \mathbb{N}^*$ et $\{u_1, \dots, u_n\}$ une famille de vecteurs d'un espace vectoriel E. Le rang de cette famille est le cardinal maximal des sous-familles libres de $\{u_1, \dots, u_n\}$.*

Un cas particulier très utile est le suivant :

EXEMPLE 1.64. Le rang d'une famille échelonnée de vecteurs d'un espace vectoriel de dimension finie est égal à son cardinal.

1.4 Sommes et sommes directes de sous-espaces vectoriels

DÉFINITION 1.65 (Somme de sous-espaces vectoriels)**. —** *Soient F_1 et F_2 deux sous-espaces vectoriels d'un espace vectoriel E. On définit la* somme *de $F_1 + F_2$ comme l'ensemble des vecteurs de la forme $x + y$ où $x \in F_1$ et $y \in F_2$. Autrement dit,*

$$F_1 + F_2 = \{x + y \mid x \in F_1, y \in F_2\}$$

PROPOSITION 1.66. — *La somme de deux sous-espaces vectoriels* F *et* G *d'un espace vectoriel* E *est un sous-espace vectoriel de* E*. Si* F *et* G *sont de dimension finie, alors* $F+G$ *l'est aussi et* $\dim(F+G) \leq \dim(F)+\dim(G)$.

REMARQUE 1.67. Attention, la somme de deux sous-espaces vectoriels n'a rien à voir avec l'union de deux sous-espaces vectoriels ! En général, l'union de deux sous-espaces vectoriels n'est même pas un sous-espace vectoriel, comme le montre l'exercice suivant.

Exercice 1.68. Soient F et G deux sous-espaces vectoriels d'un espace vectoriel E. Montrer que $F \cup G$ est un sous-espace vectoriel de E si et seulement si $F \subset G$ ou $G \subset F$.

DÉFINITION 1.69 (Somme directe). — *Soient* F *et* G *deux sous-espaces vectoriels d'un espace vectoriel* E*. On dit que la somme* $F+G$ *est* directe *lorsque tout vecteur de la somme se décompose* de manière unique *comme somme d'un vecteur de* F *et d'un vecteur de* G.

Lorsque la somme de deux sous-espaces vectoriels est directe, on écrit $F \oplus G$ pour désigner $F+G$, ce qui se lit « somme directe de F et G ».

PROPOSITION 1.70. — *La somme* $F+G$ *de deux sous-espaces vectoriels d'un espace vectoriel* E *est directe si et seulement si pour tout* $x \in F$ *et pour tout* $y \in G$*, on a l'implication*

$$\big(x+y=0\big) \Longrightarrow \big(x=y=0\big)$$

THÉORÈME 1.71. — *La somme* $F+G$ *de deux sous-espaces vectoriels d'un espace vectoriel* E *est directe si et seulement si* $F \cap G = \{0\}$.

EXEMPLE 1.72. On a la relation suivante entre sous-espaces vectoriels de $\mathbb{R}^3$:

$$H = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x, y, z \in \mathbb{R} \text{ et } x+y+z=0 \right\} = \operatorname{vect}\left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right) \oplus \operatorname{vect}\left(\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right)$$

En effet, si $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ vérifie $x+y+z=0$ alors $z=-x-y$ et donc :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ -x-y \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

Comme par ailleurs $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \in H$, on a bien :

$$H = \operatorname{vect}\left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right) + \operatorname{vect}\left(\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right)$$

et il suffit pour conclure de vérifier (exercice : faites-le !) que l'on a bien :

$$\operatorname{vect}\left(\begin{pmatrix}1\\0\\-1\end{pmatrix}\right)\cap\operatorname{vect}\left(\begin{pmatrix}0\\1\\-1\end{pmatrix}\right)=\{0\}$$

Proposition 1.73. — *Soient* F *et* G *deux sous-espaces vectoriels d'un espace vectoriel* E *tels que* F + G *soit de dimension finie. Alors* F *et* G *sont en somme directe si et seulement s'il existe des bases* $\mathscr{B}_F$ *et* $\mathscr{B}_G$ *respectivement de* F *et* G *telles que* $\mathscr{B}_F \cup \mathscr{B}_G$ *soit une base de* F + G*. Si cette condition est vérifiée, toutes les bases de* F *et de* G *vérifient cette propriété*[10]*.*

Proposition 1.74. — *Soient* F *et* G *deux sous-espaces vectoriels d'un espace vectoriel* E *tels que* F + G *soit de dimension finie. Alors* F *et* G *sont en somme directe si et seulement si on a :*

$$\dim\left(F+G\right)=\dim(F)+\dim(G)$$

Définition 1.75 (Sous-espaces vectoriels supplémentaires)**. —** *Soit* E *un espace vectoriel et* F_1, F_2 *deux sous-espaces vectoriels de* E*. Si* $E = F_1 \oplus F_2$*, on dit que* F_1 *est un* supplémentaire *de* F_2 *(on précise parfois « dans* E *») ou que* F_1 *et* F_2 *sont* supplémentaires *(dans* E*).*

Exemple 1.76. Soit E l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et soit $\mathscr{P}$ (respectivement $\mathscr{I}$) l'ensemble des fonctions paires (respectivement impaires). Alors,

$$E=\mathscr{P}\oplus\mathscr{I}$$

En effet, soit f une fonction réelle. Définissons deux fonctions g, h par :

$$\forall x\in\mathbb{R},\quad g(x)=\frac{f(x)+f(-x)}{2}\qquad\text{et}\qquad h(x)=\frac{f(x)-f(-x)}{2}$$

On vérifie facilement que $f = g + h$. La fonction g est paire, et la fonction h est impaire (vérifiez-le en exercice). On a donc bien $\mathscr{P} + \mathscr{I} = E$. Pour montrer que la somme est directe, il ne reste donc plus qu'à vérifier que $\mathscr{P} \cap \mathscr{I} = \{0\}$.

Exercice 1.77. Complétez la dernière étape de la démonstration ci-dessus.

Exercice 1.78. Soit $F_1 = \{(x, y, z) \in \mathbb{R}^3 \mid x - 2y + z = 0\}$ et $F_2 = \{(-\alpha, 2\alpha, \alpha), \alpha \in \mathbb{R}\}$. Montrer que F_1 et F_2 sont des sous-espaces vectoriels supplémentaires de $\mathbb{R}^3$.

Théorème 1.79. — *Tout sous-espace vectoriel d'un espace vectoriel de dimension finie possède un supplémentaire.*

Exercice 1.80. Deux sous-espaces vectoriels supplémentaires sont-ils complémentaires ? Le supplémentaire d'un sous-espace vectoriel d'un espace vectoriel de dimension finie est-il unique ?

10. On dit qu'elles fournissent une base de F + G *par recollement.*

On généralise maintenant la notion de somme directe de deux espaces vectoriels au cas d'un nombre fini quelconque de plusieurs sous-espaces vectoriels. Il n'y a pas de différence fondamentale.

DÉFINITION 1.81 (Somme directe d'un nombre fini de sous-espaces vectoriels)**. —** *Soit $n \in \mathbb{N}^*$ et soient $\mathrm{E}_1, \ldots, \mathrm{E}_n$ des sous-espaces vectoriels d'un espace vectoriel E. On définit la somme des E_i comme l'ensemble des vecteurs s'écrivant sous la forme $x_1 + \ldots + x_n$, où pour tout $i \in [\![1, n]\!]$ on a $x_i \in \mathrm{E}_i$. On la note $\mathrm{E}_1 + \ldots + \mathrm{E}_n$.*

On dit que cette somme est directe si tout vecteur de $\mathrm{E}_1 + \ldots + \mathrm{E}_n$ s'écrit de manière unique *sous la forme $x_1 + \ldots + x_n$, où pour tout $i \in [\![1, n]\!]$ on a $x_i \in \mathrm{E}_i$.*

REMARQUE 1.82. Il arrive souvent que l'on note

$$\bigoplus_{i=1}^{n} \mathrm{E}_i = \mathrm{E}_1 \oplus \mathrm{E}_2 \oplus \ldots \oplus \mathrm{E}_n$$

la somme directe des E_i.

PROPOSITION 1.83. — *Soit $n \in \mathbb{N}^*$ et soient $\mathrm{E}_1, \ldots, \mathrm{E}_n$ des sous-espaces vectoriels d'un espace vectoriel E de dimension finie. La somme $\mathrm{E}_1 + \ldots + \mathrm{E}_n$ est directe si et seulement si on a :*

$$\forall (x_1, \ldots, x_n) \in \mathrm{E}_1 \times \ldots \times \mathrm{E}_n, \big(x_1 + \ldots + x_n = 0\big) \Rightarrow x_1 = \ldots = x_n = 0$$

REMARQUE 1.84. Notez que la condition ci-dessus n'est pas équivalente à la condition $\bigcap_{i=1}^{n} \mathrm{E}_i = \varnothing$, qui ne caractérise pas (si $n \neq 2$) le fait que la somme des E_i soit directe !

THÉORÈME 1.85. — *Soient $n \in \mathbb{N}^*$ et $\mathrm{E}_1, \ldots, \mathrm{E}_n$ des sous-espaces vectoriels d'un espace vectoriel E de dimension finie. Alors,*

$$\mathrm{E} = \bigoplus_{i=1}^{n} \mathrm{E}_i$$

si et seulement si $\mathrm{E} = \mathrm{E}_1 + \ldots + \mathrm{E}_n$ et $\dim(\mathrm{E}_1) + \ldots + \dim(\mathrm{E}_n) = \dim(\mathrm{E})$.

Cette condition équivaut au fait que si $\mathscr{B}_1, \ldots, \mathscr{B}_n$ sont des bases respectivement de $\mathrm{E}_1, \ldots, \mathrm{E}_n$, alors $\mathscr{B} = \bigsqcup_{i=1}^{n} \mathscr{B}_i$ est une base de E (on dit alors que l'on peut obtenir une base de E par recollement *des bases des E_i). Comme dans le cas $n = 2$, elle équivaut aussi au fait qu'il* existe *un tel n-uplet de bases.*

2 Le calcul matriciel.

Dans cette section, nous étudions les *matrices* : ce sont des tableaux de chiffres que l'on peut multiplier, additionner, etc. Elles donnent un moyen très puissant pour effectuer des calculs, représenter certaines applications entre espaces vectoriels ou résoudre des systèmes linéaires. Le grand intérêt des matrices est qu'elles sont faciles à manipuler et que ce sont des objets concrets que l'on visualise facilement. Comme nous le verrons plus loin, elles permettent de synthétiser beaucoup d'informations en très peu d'efforts !

2.1 Les matrices.

Dans tout ce chapitre, on se donne deux entiers strictement positifs m et n.

On identifiera dans les exercices et dans certains énoncés $\mathbb{R}^n$ à $\mathcal{M}_{n,1}$, c'est-à-dire que l'on utilisera indifféremment la notation des vecteurs de $\mathbb{R}^n$ sous forme de n-uplets ou sous forme de vecteurs colonnes.

2.1.1 Définitions, exemples et premières propriétés.

DÉFINITION 2.1 (Matrice)**. —** *Une* matrice de taille $n \times m$ *à coefficients réels* M *est un tableau de nombres réels à n lignes et m colonnes. Le coefficient présent à la ligne i et à la colonne j est souvent noté $a_{i,j}$ (ou $\mathrm{M}_{i,j}$) et la matrice est alors notée :*

$$\mathrm{M} = (a_{i,j})_{\substack{i \in [\![1,n]\!] \\ j \in [\![1,m]\!]}}$$

ou simplement $(a_{i,j})$ s'il n'y a pas d'ambiguïté sur son format. L'ensemble des matrices $n \times m$ à coefficients dans $\mathbb{R}$ est souvent noté $\mathcal{M}_{n,m}(\mathbb{R})$. L'ensemble des matrices carrées de taille *n, c'est-à-dire de format $n \times n$, est souvent noté $\mathcal{M}_n(\mathbb{R})$.*

On utilisera jusqu'à la fin du cours l'expression « matrice » pour désigner une matrice à coefficients réels.

On présente généralement une matrice sous la forme d'un tableau entre parenthèses comme celui-ci :

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$$

DÉFINITION 2.2 (Vecteurs lignes, vecteurs colonnes)**. —** *Si $(a_{i,j})_{\substack{i \in [\![1,n]\!] \\ j \in [\![1,m]\!]}} \in \mathcal{M}_{n,m}(\mathbb{R})$, si $i \in [\![1, n]\!]$ et si $j \in [\![1, m]\!]$, on appelle* i-ème vecteur ligne de A *le vecteur donné par :*

$$\mathrm{L}_i = \begin{pmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,m} \end{pmatrix}$$

et j-ème vecteur colonne de A *le vecteur donné par :*

$$\mathrm{C}_j = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}$$

Il est possible d'additionner des matrices, de les multiplier par un scalaire et même de les multiplier entre elles comme expliqué ci-après.

DÉFINITION 2.3 (Addition de matrices, multiplication externe)**. —** *Soient*

$$\mathrm{A} = (a_{i,j})_{\substack{i \in [\![1,n]\!] \\ j \in [\![1,m]\!]}} \quad \textit{et} \quad \mathrm{B} = (b_{i,j})_{\substack{i \in [\![1,n]\!] \\ j \in [\![1,m]\!]}}$$

deux matrices de même taille $n \times m$*. On définit la matrice* $\mathrm{A}+\mathrm{B}$ *comme la matrice de* $\mathscr{M}_{n,m}(\mathbb{R})$ *dont l'élément d'indices* (i, j) *est* $a_{i,j} + b_{i,j}$*. Si* $\lambda \in \mathbb{R}$*, on définit la matrice* $\lambda \cdot \mathrm{A}$ *comme la matrice de* $\mathscr{M}_{n,m}(\mathbb{R})$ *dont l'élément d'indices* (i, j) *est* $\lambda a_{i,j}$*.*

EXEMPLE 2.4. Voici un exemple de ces deux opérations :

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 3 \\ 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 3 & -1 \end{pmatrix}$$

et

$$5 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 15 \end{pmatrix}$$

PROPOSITION 2.5. — *L'ensemble* $\mathscr{M}_{n,m}(\mathbb{R})$ *muni des lois* $+$ *et* $\cdot$ *définies ci-dessus est un espace vectoriel.*

Exercice 2.6. Montrer que la dimension de $\mathscr{M}_{n,m}(\mathbb{R})$ est $n \times m$. En exhiber une base.

DÉFINITION 2.7 (Matrices particulières)**. —** *On présente ici certaines formes de matrices particulières qu'il est indispensable de connaître.*

1. *On note* $0_{n,m}$ *(ou simplement* 0*) et on appelle* matrice nulle de taille $n \times m$ *l'élément neutre de* $\mathscr{M}_{n,m}$*, c'est-à-dire la matrice de taille* $n \times m$ *dont tous les termes sont des zéros.*
2. *On note* I_n *et on appelle* matrice identité de taille n *la matrice carrée :*

$$\mathrm{I}_n = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

3. *Les matrices carrées de la forme :*

$$\lambda \mathrm{I}_n = \begin{pmatrix} \lambda & 0 & \cdots & \cdots & 0 \\ 0 & \lambda & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \lambda & 0 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

obtenues en multipliant la matrice I_n *par un scalaire* $\lambda \in \mathbb{R}$ *sont appelées des matrices* scalaires.

4. *Les matrices carrées de la forme :*

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \lambda_{n-1} & 0 \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}$$

avec $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ *sont appelées des* matrices diagonales.

5. *Une matrice carrée est dite* triangulaire supérieure *lorsque tous ses termes situés strictement en dessous de sa diagonale*[11] *sont nuls, c'est-à-dire qu'elle est de la forme :*

$$\mathrm{I}_n = \begin{pmatrix} * & * & \cdots & \cdots & * \\ 0 & * & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & * & * \\ 0 & \cdots & \cdots & 0 & * \end{pmatrix}$$

Une matrice carrée est dite triangulaire inférieure *lorsque tous ses termes situés strictement au-dessus de sa diagonale sont nuls, c'est-à-dire qu'elle est de la forme :*

$$\mathrm{I}_n = \begin{pmatrix} * & 0 & \cdots & \cdots & 0 \\ * & * & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & * & 0 \\ * & \cdots & \cdots & * & * \end{pmatrix}$$

Une matrice carrée est simplement dite triangulaire *si elle triangulaire supérieure ou triangulaire inférieure.*

EXEMPLE 2.8. La matrice

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

est triangulaire supérieure.

REMARQUE 2.9. Une manière plus rigoureuse de définir une matrice carrée $\mathrm{A} = (a_{i,j})$ de taille n est triangulaire supérieure consiste à dire que A est triangulaire supérieure

11. Rappelons que si l'on ne précise pas de quelle diagonale il s'agit, c'est de la diagonale principale, c'est-à-dire de celle reliant le coin supérieur gauche de la matrice à son coin inférieur droit, qu'il est question.

si et seulement si pour tout couple d'indices $(i, j) \in [\![1, n]\!]^2$ avec $j < i$, on a $a_{i,j} = 0$. Assurez-vous que vous avez bien compris cette caractérisation, car elle est fréquemment utilisée !

DÉFINITION 2.10 (Tranposée)**. —** *Si* $M \in \mathcal{M}_{n,m}(\mathbb{R})$*, la* matrice transposée de M *(ou plus simplement la* transposée de M*), est la matrice* ${}^tM$ *de* $\mathcal{M}_{m,n}(\mathbb{R})$ *définie par :*

$$\forall i \in [\![1, m]\!], \forall j \in [\![1, n]\!], \quad \left({}^tM\right)_{i,j} = M_{j,i}$$

EXEMPLE 2.11. Si

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 3 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix}$$

alors

$${}^tA = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 0 \end{pmatrix} \quad \text{et} \quad {}^tB = \begin{pmatrix} 3 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix}$$

REMARQUE 2.12. Comme le suggère l'exemple ci-dessus, la transposée d'une matrice carrée triangulaire supérieure est triangulaire inférieure et vice-versa.

2.1.2 Multiplication d'un vecteur colonne par une matrice.

On commence par définir la multiplication d'un vecteur par une matrice. Pour la suite, nous utiliserons fréquemment la notation $(x_i)_{i \in [\![1, m]\!]}$ pour désigner le vecteur colonne

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

DÉFINITION 2.13 (Multiplication d'un vecteur par une matrice)**. —** *Le produit par une matrice* $A = (a_{i,j})_{\substack{i \in [\![1,n]\!] \\ j \in [\![1,m]\!]}}$ *d'un vecteur* $x = (x_i)_{i \in [\![1, m]\!]} \in \mathbb{R}^m$ *est le vecteur* $Ax = (y_i)_{i \in [\![1, n]\!]}$ *de* $\mathbb{R}^n$ *défini par :*

$$\forall i \in [\![1, n]\!], \quad y_i = a_{i,1}x_1 + \ldots + a_{i,m}x_m = \sum_{k=1}^{m} a_{i,k}x_k$$

Dans la définition précédente, le vecteur y s'écrit donc :

$$y = \begin{pmatrix} a_{1,1}x_1 + \ldots + a_{1,m}x_m \\ \vdots \\ a_{n,1}x_1 + \ldots + a_{n,m}x_m \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^m a_{1,k}x_k \\ \vdots \\ \sum_{k=1}^m a_{n,k}x_k \end{pmatrix}$$

Faites bien attention au format : si A possède n lignes et m colonnes, le vecteur x doit avoir m lignes et le vecteur Ax aura n lignes !

PROPOSITION 2.14. — *Soit* $A \in \mathcal{M}_{n,m}(\mathbb{R})$*. Soient* x, y *deux éléments de* $\mathbb{R}^m$ *et* $\lambda \in \mathbb{R}$ *un scalaire. Alors :*

$$A(x + y) = Ax + Ay \quad \textit{et} \quad A(\lambda x) = \lambda(Ax)$$

La proposition précédente montre que la multiplication d'un vecteur par une matrice est « compatible avec les lois de compositions sur $\mathbb{R}^m$ » , c'est-à-dire que l'application $f_A : x \mapsto Ax$ de $\mathbb{R}^m$ dans $\mathbb{R}^n$ est ce que l'on appelle une « application linéaire »; on appelle d'ailleurs cette application *l'application linéaire canoniquement associée à* A. Nous étudierons intensivement les applications linéaires à partir du chapitre suivant.

DÉFINITION 2.15 (Image et noyau d'une matrice)**. —** *Soit* A *une matrice de taille* $n \times m$*. On appelle* image *de* A *l'ensemble :*

$$\mathrm{Im}(A) = \left\{y \in \mathbb{R}^n \mid \exists x \in \mathbb{R}^m \mid y = Ax\right\}$$

On appelle noyau *de* A *l'ensemble :*

$$\ker(A) = \left\{x \in \mathbb{R}^m \mid Ax = 0\right\}$$

PROPOSITION 2.16. — *Le noyau et l'image d'une matrice sont des sous-espaces vectoriels.*

Calculer l'image ou le noyau d'un vecteur n'est pas une mince affaire, et nécessite la résolution d'un système linéaire. Donnons-nous par exemple une matrice $A = (a_{i,j})_{\substack{i\in[\![1,n]\!]\\ j\in[\![1,m]\!]}}$. Son noyau est l'ensemble des $x \in \mathbb{R}^m$ vérifiant :

$$\begin{cases} a_{1,1}x_1 + \ldots + a_{1,m}x_m = 0 \\ \qquad\qquad \cdots \\ a_{n,1}x_1 + \ldots + a_{n,m}x_m = 0 \end{cases}$$

Déterminer le noyau de A demande donc de résoudre un système de n équations à m inconnues. Nous étudierons dans quelques pages une méthode pour résoudre un tel système.

DÉFINITION 2.17 (Rang d'une matrice)**. —** *Soit* A *une matrice* $n \times m$*. Le* rang *de la matrice* A *est le rang de la famille formée par ses vecteurs colonnes. On le note* rg(A).

PROPOSITION 2.18. — *Le rang d'une matrice est la dimension de son image.*

Exercice 2.19. Quel est le rang de la matrice I_n ? Quel est le rang d'une matrice diagonale ?

Exercice 2.20. Calculer le rang de la matrice suivante :

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

THÉORÈME 2.21. — *Le rang de la famille des lignes d'une matrice* $A \in \mathcal{M}_{n,m}(\mathbb{R})$ *est égal au rang de la famille des colonnes de cette même matrice, c'est-à-dire à* rg(A)*. On a donc* $\mathrm{rg}(A) \leqslant \min(m, n)$.

Annonçons tout de suite un théorème fondamental dont nous discuterons plus abondamment dans la section sur le pivot de Gauss :

THÉORÈME 2.22. — *Le rang d'une matrice carrée triangulaire est égal au nombre de coefficients non nuls sur sa diagonale.*

La proposition suivante découle elle aussi d'un résultat que nous verrons plus tard mais trouve tout à fait sa place dans cette section.

PROPOSITION 2.23. — *Une matrice a le même rang que sa transposée.*

En général, calculer le rang d'une matrice n'est pas aisé. On peut facilement s'en sortir en utilisant le théorème 2.21 et le fait que le rang d'une famille échelonnée de vecteurs est facile à déterminer, comme dans l'exemple suivant :

Exercice 2.24. Calculer le rang de la matrice suivante :

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & -8 \\ 0 & 1 & 1 \end{pmatrix}$$

Il existe un algorithme systématique pour calculer le rang de n'importe quelle matrice : il s'agit de la méthode du pivot de Gauss que nous verrons dans la dernière section de ce chapitre.

2.1.3 Multiplication de deux matrices.

Nous allons maintenant le produit de deux matrices entre elles, ce qui constitue une généralisation de la multiplication d'un vecteur par une matrice que nous venons d'étudier.

DÉFINITION 2.25 (Multiplication matricielle)**.** — *Soient l, m, n trois entiers naturels non nuls.*

Soit $\mathrm{A} = (a_{i,j})_{\substack{i \in [\![1,l]\!] \\ j \in [\![1,n]\!]}}$ une matrice de taille $l \times n$ et soit $\mathrm{B} = (b_{i,j})_{\substack{i \in [\![1,n]\!] \\ j \in [\![1,m]\!]}}$ une matrice de format $n \times m$.

La matrice $\mathrm{A} \times \mathrm{B}$ (plus souvent notée AB) est la matrice $\mathrm{B} = (c_{i,j})_{\substack{i \in [\![1,l]\!] \\ j \in [\![1,m]\!]}}$ de taille $l \times m$ dont l'élément d'indices (i, j) est le nombre réel $c_{i,j}$ défini par :

$$c_{i,j} = a_{i,1}b_{1,j} + \dots + a_{i,n}b_{n,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j}$$

Cette opération d'apparence complexe est en réalité très simple à réaliser dès lors que l'on en possède une intuition graphique; la figure 2.1.3 fournit une telle illustration, que le lecteur est invité à étudier attentivement et à mémoriser.

EXEMPLE 2.26. Voici un premier exemple de multiplication matricielle :

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 3 \end{pmatrix} \times \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} 1\times1+1\times0+2\times5 & 1\times1+1\times0+2\times1 \\ 0\times1+2\times0+3\times5 & 0\times1+2\times0+3\times1 \end{pmatrix} = \begin{pmatrix} 11 & 3 \\ 15 & 3 \end{pmatrix}$$

Avec un peu d'entraînement, tout ceci devient très intuitif. Nous attirons cependant l'attention sur le *format* des matrices : les hypothèses sont essentielles. On retiendra donc l'avertissement suivant :

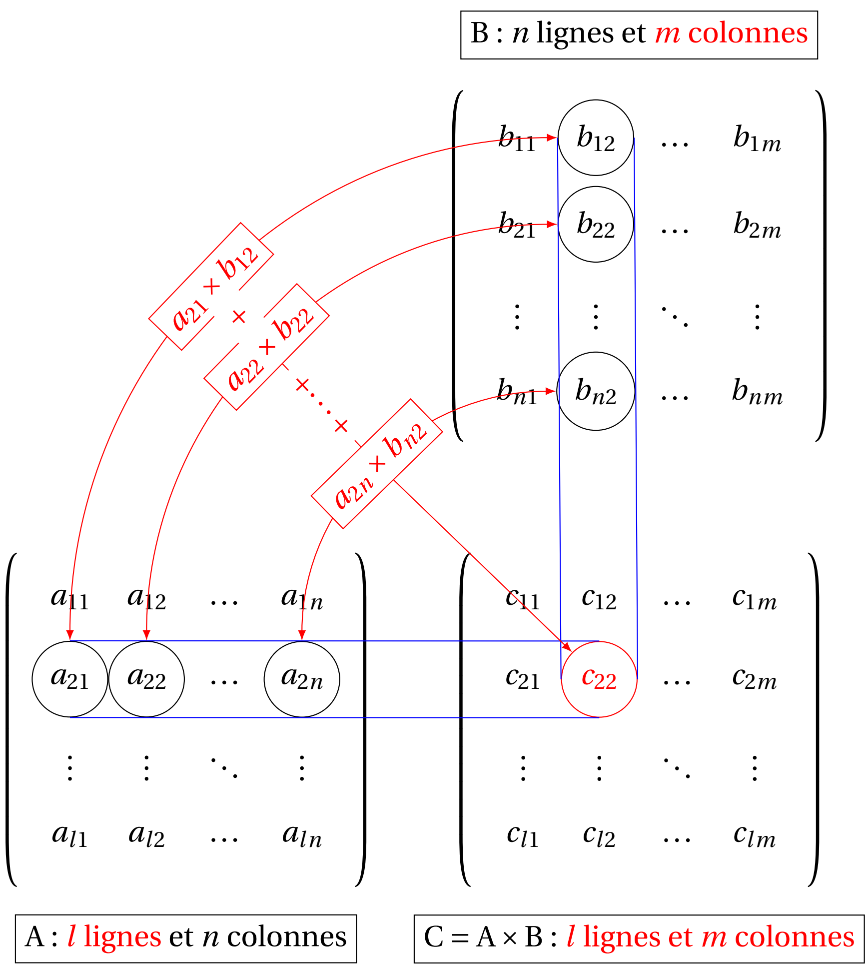


FIGURE 3 – Illustration du calcul pratique d'un produit matriciel

La multiplication $A \times B$ *n'a de sens QUE si le nombre de colonnes de* A *est égal au nombre de lignes de* B.

Il est également recommandé de s'exercer à trouver le format de la matrice d'arrivée. En l'occurrence, la matrice d'arrivée a autant de lignes que A et autant de colonnes que B. Les choses semblent assez claires sur un dessin (voir la figure 4).

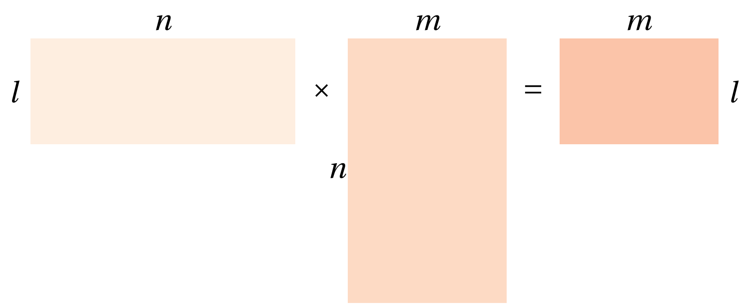


FIGURE 4 – Le format dans la multiplication des matrices.

Bien entendu, lorsque les matrices considérées sont carrées et de même format, tout est plus agréable : les multiplications sont bien définies et le résultat sera encore une matrice carrée de même format.

Exercice 2.27. Soient l, m et n trois entiers naturels non nuls, et deux matrices $A \in \mathcal{M}_{l,m}(\mathbb{R})$ et $B \in \mathcal{M}_{m,n}$. Montrer que :

$$^t(AB) = {}^tB\,{}^tA$$

REMARQUE 2.28. Faisons le lien entre le multiplication de deux matrices et la multiplication d'un vecteur par une matrice présentée au début de cette section. Dans le cadre de la définition précédente, si l'on note $B_1, \dots, B_m$ les vecteurs colonnes qui constituent B alors la matrice $A \times B$ admet pour vecteurs colonnes les vecteurs $AB_1, \dots, AB_n$. En d'autres termes :

$$A \times \begin{pmatrix} B_1 & \dots & B_m \end{pmatrix} = \begin{pmatrix} AB_1 & \dots & AB_m \end{pmatrix}$$

Enfin, on remarque qu'un vecteur de $\mathbb{R}^n$ est une matrice de $\mathcal{M}_{n,1}(\mathbb{R})$ et que les deux définitions du produit d'un vecteur par une matrice coïncident alors.

PROPOSITION 2.29. — *Soit* $n \in \mathbb{N}^*$ *et soit* A *une matrice carrée de taille* $n \times n$. *Alors,*

$$AI_n = I_nA = A$$

PROPOSITION 2.30. — *Si* l, m *et* n *sont trois entiers naturels non nuls, si* $A, A' \in \mathcal{M}_{l,n}$ *et si* $B, B' \in \mathcal{M}_{n,m}$, *alors on a :*

$$A(B + B') = AB + AB' \quad \textit{et} \quad (A + A')B = AB + A'B$$

REMARQUE 2.31. Il est en général **faux** de dire que si A et B sont deux matrices telles que le produit AB et le produit BA existent, on a AB = BA : on dit que le produit matriciel n'est pas commutatif. Par exemple :

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Deux matrices A et B vérifiant AB = BA sont nécessairement carrées et de même taille (exercice : pourquoi ?), et on dit alors qu'elles *commutent.*

Exercice 2.32 (très important)**.** Soient A et B deux matrices non nulles telles que le produit AB existe. Peut-on avoir AB = 0 ? Quelle est la différence avec la multiplication des nombres réels ?

DÉFINITION 2.33 (Puissances itérées)**. —** *Soit* $A \in \mathcal{M}_n(\mathbb{R})$ *une matrice carrée et* k *un entier naturel. On définit* A^k *comme suit :*

$$A^0 = I_n$$
$$\text{Si } k > 0, \quad A^k = \underbrace{A \times \ldots \times A}_{k \text{ fois.}}$$

Exercice 2.34. Calculer les puissances itérées d'une matrice diagonale.

DÉFINITION 2.35 (Nilpotence)**. —** *On dit qu'une matrice carrée* A *est* nilpotente *s'il existe un entier naturel* n *tel que* $A^n = 0$. *On appelle* indice de nilpotence *de* A *le plus petit entier* $n \geq 1$ *vérifiant une telle propriété, c'est-à-dire tel que* $A^{n-1} \neq 0$ *et* $A^n = 0$.

Exercice 2.36. Soit $A = (a_{i,j})_{i,j \in [\![1,n]\!]}$ une matrice carrée dont tous les termes sont strictement positifs, c'est-à-dire $a_{i,j} > 0$ pour tous $i, j \in [\![1, n]\!]$. La matrice A peut-elle être nilpotente ?

PROPOSITION 2.37 (Formule du binôme de Newton)**. —** *Soit* $N \in \mathbb{N}^*$ *et soient* A *et* B *deux matrices de* $\mathcal{M}_n(\mathbb{R})$ ***qui commutent****. Soit n un entier naturel. Alors pour tout* $n \in \mathbb{N}$ *on a :*

$$(A+B)^N = \sum_{k=0}^{N} \binom{N}{k} A^k B^{N-k}$$

REMARQUE 2.38. On insiste lourdement sur l'hypothèse de la proposition précédente : il faut que A et B commutent pour que la formule soit vraie... et elle ne l'est plus en général si A et B ne commutent pas ! Dans le cas N = 2, elle est même tout le temps fausse lorsque la commutativité n'est pas réalisée. En effet, si A et B ne commutent pas, on a

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

et ceci n'est pas égal à $A^2 + 2AB + B^2$ puisque $BA \neq AB$. Il faut donc se méfier : ce piège est tellement classique que tomber dedans peut coûter très cher.

Exercice 2.39. Montrer que l'ensemble des matrices nilpotentes de taille n n'est pas un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ si $n \geq 2$.

PROPOSITION 2.40. — *La multiplication matricielle n'augmente pas le rang, c'est-à-dire que pour toutes matrices* $A \in \mathcal{M}_{m,n}(\mathbb{R})$ *et* $B \in \mathcal{M}_{n,p}(\mathbb{R})$, *on a :*

$$\operatorname{rg}(AB) \leqslant \min\big(\operatorname{rg}(A), \operatorname{rg}(B)\big)$$

2.1.4 Inversion de matrices.

On fixe dans cette section encore entier naturel non nul n. Comme pour les fonctions, il est possible de définir les matrices « inversibles » .

DÉFINITION 2.41 (Inversibilité)**. —** *Soit* A *une matrice carrée de taille* n*. On dit que* A *est* inversible *s'il existe une matrice* B *de même format telle que*

$$AB = BA = I_n$$

B *est alors appelée* inverse de A *et notée* A^{-1}*. L'ensemble des matrices inversibles à coefficients dans* $\mathbb{R}$ *est noté* $GL_n(\mathbb{R})$ *(« groupe linéaire de* $\mathbb{R}^n$ *»).*

PROPOSITION 2.42. — *Soient* A *et* B *deux matrices carrées inversibles de même taille. Alors, la matrice* AB *est encore inversible et on a :*

$$(AB)^{-1} = B^{-1}A^{-1}$$

REMARQUE 2.43. Cette propriété fait de $GL_n(\mathbb{R})$ un *groupe* (d'où son nom), c'est-à-dire un ensemble G muni d'une loi de composition interne $*$ admettant un élément neutre e (c'est-à-dire tel que pour tout $x \in G$ on a $x * e = e * x = x$) et telle que tous les éléments de l'ensemble soient inversibles pour cette loi (c'est-à-dire que pour tout $x \in G$ il existe $y \in G$ tel que $x * y = y * x = e$) : ici, la loi est la multiplication matricielle, l'élément neutre la matrice I_n et l'inverse d'une matrice $A \in GL_n(\mathbb{R})$... son inverse (sic). Cette notion (que nous avons déjà rencontrée sans le dire puisqu'un espace vectoriel est notamment un groupe pour la loi +) n'est pas au programme.

Exercice 2.44. Peut-on dire qu'une matrice est inversible lorsque tous ses termes sont non nuls ?

Exercice 2.45. Montrer que si A est inversible, alors ${}^tA$ aussi. Calculer $({}^tA)^{-1}$.

Exercice 2.46. Une matrice nilpotente (voir définition 2.35) peut-elle être inversible ?

Le théorème suivant sera clair une fois connu le théorème du rang, qui sera présenté dans le chapitre suivant :

Théorème 2.47. — *Soit* $\mathrm{A} \in \mathscr{M}_n(\mathbb{R})$ *une matrice carrée de taille* n*. Les propriétés suivantes sont toutes équivalentes :*

1. A *est inversible.*
2. *Le rang de* A *est égal à* n*.*
3. *Le noyau de* A *est* $\{0\}$*.*
4. *Les colonnes de* A *forment une base de* $\mathbb{R}^n$*.*
5. *Les colonnes de* A *forment une famille libre de* $\mathbb{R}^n$*.*
6. *Les lignes de* A *forment une base de* $\mathbb{R}^n$*.*
7. *Les lignes de* A *forment une famille libre de* $\mathbb{R}^n$*.*
8. *Il existe une matrice* $\mathrm{B} \in \mathscr{M}_n(\mathbb{R})$ *telle que* $\mathrm{BA} = \mathrm{I}_n$*.*
9. *Il existe une matrice* $\mathrm{B} \in \mathscr{M}_n(\mathbb{R})$ *telle que* $\mathrm{AB} = \mathrm{I}_n$*.*
10. *Il existe une matrice* $\mathrm{B} \in \mathscr{M}_n(\mathbb{R})$ *telle que pour tous vecteurs* $\mathrm{X}, \mathrm{Y} \in \mathbb{R}^n$ *:*

$$\mathrm{AX} = \mathrm{Y} \Leftrightarrow \mathrm{X} = \mathrm{BY}$$

Dans ces trois derniers cas, on a $\mathrm{A}^{-1} = \mathrm{B}$*.*

La dernière caractérisation de l'inversibilité donnée par le théorème 2.47 est au fondement d'une méthode infaillible mais parfois fastidieuse pour calculer l'inverse d'une matrice A inversible : elle revient à résoudre le système $\mathrm{AX} = \mathrm{Y}$ de manière à exprimer X en fonction de Y et à faire apparaître une matrice B qui vérifie $\mathrm{X} = \mathrm{BY}$ et qui est, d'après le théorème 2.47, nécessairement l'inverse de A. L'exercice 2.48 ci-dessous peut ainsi être résolu en utilisant le théorème 2.47 ainsi que la méthode du pivot de Gauss pour la résolution de systèmes linéaires. Il peut aussi être traité en utilisant la méthode du pivot de Gauss pour le calcul du rang et le cas échéant en appliquant la méthode d'inversion de Gauss-Jordan. Toutes ces méthodes seront présentées dans la section suivante; pour l'heure, nous vous conseillons de travailler sur cet exercice avec des outils aussi élémentaires que possible pour acquérir l'intuition nécessaire.

Exercice 2.48. Parmi les matrices suivantes, trouver celles qui sont inversibles et celles qui ne le sont pas :

$$\begin{pmatrix} 1 & 5 \\ 2 & 5 \end{pmatrix} \qquad \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \qquad \begin{pmatrix} 4 & 4 \\ 2 & -1 \end{pmatrix} \qquad \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 3 & -1 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 5 \end{pmatrix}$$

Le cas échéant, calculer leur inverse.

On donne maintenant des critères qui permettent d'étudier rapidement l'inversibilité d'une matrice. Attention, certains ne fonctionnent que dans des cas particuliers et n'expriment pas des conditions *nécessaires* d'inversibilité !

Proposition 2.49. — *Une matrice triangulaire est inversible si et seulement si ses coefficients diagonaux sont tous non nuls.*

EXEMPLE 2.50. La matrice

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 5 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

est inversible, la matrice

$$B' = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 5 \\ 0 & 0 & -2 \end{pmatrix}$$

ne l'est pas.

Voici ensuite un critère bien pratique pour étudier l'inversibilité des petites matrices : le déterminant.

THÉORÈME 2.51. — *Soit* A *la matrice de* $\mathcal{M}_2(\mathbb{R})$ *définie par :*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

où a, b, c, d *sont des réels. Alors la matrice* A *est inversible si et seulement* $ad - bc \neq 0$*, et dans ce cas son inverse* A^{-1} *s'écrit sous la forme :*

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

La quantité $ad - bc$ *s'appelle le* déterminant *de la matrice.*

En d'autres termes, le déterminant d'une matrice de taille 2×2 est non nul si et seulement si les deux colonnes (ou les deux lignes) de la matrice sont non proportionnelles.

Il existe un concept de déterminant pour les matrices de taille 3×3, qui sera présenté dans un cours ultérieur.

On pourra à titre d'exercice s'intéresser à nouveau aux matrices de taille 2 de l'exercice 2.48 et retrouver en un clin d'œil les résultats obtenus.

2.2 L'algorithme du pivot de Gauss.

La méthode du pivot de Gauss est une des méthodes les plus utiles de l'algèbre linéaire. Il s'agit d'un algorithme qui n'utilise que des opérations très simples (additions, multiplications) pour modifier des matrices sans changer certaines de leurs caractéristiques. Cet algorithme permet :

- de calculer le rang d'une matrice,
- d'inverser une matrice,
- de résoudre explicitement tous les systèmes linéaires, et par extension de trouver les éléments du noyau d'une matrice et de son image ainsi que de trouver les vecteurs propres associés à une certaine valeur propre [12].

12. Ces deux notions seront abordées dans la dernière section de ce cours.

Il est impératif de savoir reconnaître quand il faut utiliser cet algorithme, puis de savoir l'exécuter rapidement et sans erreurs. Pour cela, il n'y a pas de secret : il suffit de s'entraîner. Le pivot de Gauss se comprend lorsqu'on le pratique!

2.2.1 Le pivot de Gauss.

Soit A une matrice à n lignes et p colonnes. On définit des opérations élémentaires sur A qui « conservent le système associé » :

DÉFINITION 2.52 (Opérations élémentaires)**. —** *Une* opération élémentaire sur la matrice A *est une des trois opérations suivantes :*

1. *addition d'une ligne à une autre ligne.*
2. *multiplication d'une ligne par un scalaire non nul.*
3. *permutation de deux lignes.*

On appellera aussi « opération élémentaire sur A *» le fait d'ajouter à une ligne un multiple non nul d'une autre (ce qui est la combinaison de deux des opérations élémentaires précédemment définies).*

Il est possible de définir de la même manière les opérations élémentaires sur les *colonnes* de la matrice.

REMARQUE 2.53 (importante)**.** Les opérations élémentaires reviennent à multiplier la matrice à gauche par une matrice particulière. Par exemple, si $\lambda \in \mathbb{R}^*$, l'addition de λL_2 à L_1 se traduit en dimension 3 par la multiplication par la matrice suivante :

$$\begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Les matrices des opérations élémentaires sont appelées avec beaucoup d'originalité *matrices des opérations élémentaires.*

Toutes les formes de la méthode du pivot résultent du résultat essentiel suivant.

THÉORÈME 2.54. — *Les opérations élémentaires effectuées sur les lignes d'une matrice ne changent pas l'espace vectoriel engendré par les lignes de cette matrice ; en particulier, elles ne changent pas son rang.*

La méthode du Pivot de Gauss consiste à modifier une matrice à l'aide d'opérations élémentaires, afin de se ramener à une matrice plus simple sur laquelle on pourra lire facilement le rang de la matrice, celui-ci n'étant pas modifié d'après le théorème précédent.

REMARQUE 2.55 (importante)**.** Avant de présenter la méthode du pivot de Gauss, insistons lourdement sur le fait qu'il est **obligatoire** d'expliciter **toutes** les opérations élémentaires que vous effectuez sur la matrice (ou le système linéaire). Vous rendrez compte de ces différentes opérations soit en clair — c'est-à-dire en utilisant le langage courant — soit en utilisant des notations du type $L_i \leftrightarrow L_j$, $L_i \leftarrow \lambda L_i$ ou $L_i \leftarrow L_i + \lambda L_j$, qui signifient respectivement que l'on effectue une interversion de la ligne L_i avec la ligne L_j, que l'on multiplie la ligne L_i par un réel non nul λ ou que l'on ajoute à la ligne L_i un multiple de la ligne L_j. Cela facilite la lecture de votre raisonnement par le correcteur et vous permet de revenir sur vos pas sans avoir à tout recommencer si vous constatez une erreur — ce qui sera sans aucun doute le cas bien souvent, au moins dans les premiers temps !

MÉTHODE 2.56 (Pivot de Gauss et calcul du rang)**.** Soit $A = (a_{i,j}) \in \mathscr{M}_{m,n}(\mathbb{R})$ une matrice. La méthode du pivot de Gauss consiste à utiliser les termes non nuls d'une ligne pour « vider » les lignes suivantes.

1. On se ramène, à l'aide de permutations de lignes, au cas où $a_{1,1} \neq 0$. Si ce n'est pas possible, c'est que la première colonne est nulle : dans ce cas, on l'ignore et on passe directement à la deuxième colonne.
2. On multiplie la première ligne par $\frac{1}{a_{1,1}}$. On obtient alors une matrice de la forme :

$$\begin{pmatrix} 1 & * & \dots & * \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$$

   Le « 1 » en haut s'appelle le *pivot*.
3. En s'aidant des opérations élémentaires, on « nettoie » la première colonne en dessous du pivot : par exemple, pour supprimer le terme $a_{2,1}$ on effectue l'opération élémentaire $L_2 \longleftarrow L_2 - a_{2,1}L_1$, ce qui aboutit à :

$$\begin{pmatrix} 1 & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$$

   Puis on continue jusqu'à obtenir une matrice de la forme suivante :

$$\begin{pmatrix} 1 & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \dots & * \end{pmatrix} = \begin{pmatrix} 1 & * & \dots & * \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

   où la matrice A' est de format $(m-1) \times (n-1)$.
4. On retourne à l'étape 1, mais cette fois en cherchant à agir sur la matrice A' ; notons que les opérations ainsi effectuées ne changent pas la première colonne de A. À l'issue de cette étape, on obtient une matrice de la forme

$$\begin{pmatrix} 1 & * & * & * & \dots & * & * \\ 0 & 1 & * & * & \dots & * & * \\ 0 & 0 & & & & & \\ \vdots & \vdots & & & A' & & \\ & & & & & & \\ 0 & 0 & & & & & \end{pmatrix}$$

où la matrice A'' est de format $(m-2)\times(n-2)$.

5. On réitère cette opération et on s'arrête lorsque le dernier bloc obtenu dans la suite initiée par A, A' et A'' est entièrement nul, ne contient plus qu'une ligne (ce qui peut arriver si $m \leq n$) ou est de taille nulle (ce qui peut arriver si $n \leq m$).

La matrice A ainsi transformée devrait alors ressembler à ce qu'on appelle une matrice « échelonnée [13] », par exemple comme ceci :

$$\begin{pmatrix} 1 & * & * & * & * & * & * & * & * & * \\ 0 & 1 & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Le rang de cette dernière matrice se lit immédiatement : en effet, le rang d'une matrice est le rang de la famille de ses vecteurs lignes (ou de ses vecteurs colonne, les deux étant identiques par le théorème 2.21). Or, la famille des vecteurs ligne de la matrice nouvellement obtenue est une famille échelonnée : le rang de la matrice est donc égal au nombre de vecteurs non nuls de cette famille, c'est-à-dire 4 dans l'exemple ci-dessus.

EXEMPLE 2.57. On désire calculer le rang de la matrice

$$A = \begin{pmatrix} 0 & 6 & 6 \\ 4 & 2 & 0 \\ 1 & 2 & 2 \end{pmatrix}$$

Le coefficient d'indices $(1,1)$ étant nul, on inverse les lignes 1 et 3 (on aurait pu inverser les lignes 1 et 2, mais dans ce cas il aurait fallu multiplier la première ligne par $1/4$ pour avoir un pivot égal à 1, ce qui aurait été plus long).

$$\begin{pmatrix} 1 & 2 & 2 \\ 0 & 6 & 6 \\ 4 & 2 & 0 \end{pmatrix}$$

On « nettoie » ensuite la première colonne : il n'y a que le terme d'indices $(3,1)$ à supprimer, ce que l'on fait par l'opération $L_3 \longleftarrow L_3 - 4L_1$ qui donne

$$\begin{pmatrix} 1 & 2 & 2 \\ 0 & 6 & 6 \\ 0 & -6 & -8 \end{pmatrix}$$

Le terme en $(2,2)$ est non nul. En multipliant la deuxième ligne par $1/6$, on se ramène le pivot à 1 :

$$\begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 1 \\ 0 & -6 & -8 \end{pmatrix}$$

13. Au sens où la famille de ses vecteurs lignes est échelonnée.

Par l'opération $L_3 \leftarrow L_3 + 6L_1$, on aboutit à :

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

Cette matrice est donc de rang 3; par le théorème 2.54, le rang de A est également 3. Nous avons en particulier montré que la matrice A est inversible.

REMARQUE 2.58 (culturelle)**.** Karl Friedrich Gauss (1777 - 1855), prodige allemand, est considéré comme l'un des plus grands mathématiciens de tous les temps. Son nom est attaché à une multitude d'objets, de phénomènes physiques, de théorèmes, d'algorithmes, et il a apporté des contributions absolument majeures à plusieurs branches des mathématiques. Pourtant, une fois n'est pas coutume, Gauss n'est pas l'auteur de l'algorithme qui porte son nom [14] : celui-ci était apparemment connu des mathématiciens chinois du premier siècle, et il était fréquemment utilisé (et enseigné) depuis Newton.

FIGURE 5 – C.F. Gauss, 1777 - 1855.

2.2.2 Inversion des matrices.

La méthode de Gauss sert également à inverser des matrices. Cette méthode est souvent appelée « méthode de Gauss-Jordan [15] » ou « méthode d'élimination de Gauss-Jordan » , mais il n'y a pas de différence conceptuelle avec la méthode de Gauss. La clé de cette section est le théorème suivant, qui est une simple conséquence de la méthode pour calculer le rang [16].

THÉORÈME 2.59. — *Soit* A *une matrice carrée inversible. Par des opérations élémentaires sur les lignes de la matrice, on peut transformer* A *en une matrice triangulaire supérieure avec des* 1 *sur la diagonale.*

14. Il s'agit de la Loi de Stigler : « une découverte scientifique ne porte jamais le nom de son auteur ».
15. Du nom du mathématicien allemand Wilhelm Jordan.
16. On rappelle que A est inversible si et seulement si son rang est n.

MÉTHODE 2.60 (Méthode de Gauss-Jordan)**.** Soit A une matrice carrée de taille n. On désire savoir si elle est inversible, et le cas échéant on veut déterminer son inverse. Pour cela, on calcule le rang de A par le pivot de Gauss, comme dans la section précédente. La matrice est inversible si et seulement si son rang est n, ce qui se traduit par le fait que l'on peut transformer A en une matrice triangulaire supérieure avec des 1 sur la diagonale (c'est le théorème précédent). On se sert ensuite de la diagonale pour « vider » (toujours par des opérations sur les lignes) toutes les colonnes au-dessus des pivots, en utilisant la méthode du pivot « à l'envers » : l'objectif est de se ramener à la matrice I_n.

On rappelle que les opérations élémentaires sur la matrice se traduisent par la multiplication à gauche de la matrice par des matrices élémentaires (voir la remarque 2.53). Notons $P_1, \dots, P_r$ les matrices des opérations élémentaires utilisées pour transformer A en l'identité : le pivot de Gauss (ou de Jordan) est simplement la traduction de l'égalité matricielle

$$P_r P_{r-1} \dots P_1 A = I_n$$

D'après le théorème 2.47, nous avons donc trouvé l'inverse de A : il s'agit de $P_r P_{r-1} \dots P_1$. Toute la difficulté consiste à trouver un moyen de mener à terme l'algorithme en gardant en mémoire les opérations élémentaires P_i. L'exemple suivant propose une méthode très efficace pour le faire.

EXEMPLE 2.61. On désire inverser la matrice suivante :

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

On va modifier la matrice A jusqu'à obtenir l'identité grâce à la méthode du pivot de Gauss, et on va « conserver en mémoire » les opérations que l'on a faites. Pour cela, on adopte une présentation bien pratique : on travaille sur la matrice A_0 définie par

$$A' = \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 2 & 2 & 1 & 0 & 0 & 1 \end{array}\right)$$

En fait, on a simplement accolé à droite de la matrice une matrice identité de taille n. En travaillant sur cette matrice plutôt que sur A, on va appliquer toutes les opérations élémentaires effectuées sur les lignes de A à cette matrice identité, qui va donc garder la trace de ces opérations successives.

On commence par effectuer l'opération $L_3 \leftarrow L_3 - 2L_1$, qui donne :

$$A_1 = \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & -2 & 1 & -2 & 0 & 1 \end{array}\right)$$

Puis, l'opération $L_3 \leftarrow L_3 + 2L_2$ donne :

$$A_2 = \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & 2 & 1 \end{array}\right)$$

À ce stade, on peut lire que le rang de A est 3, donc nous sommes bien sûrs que la matrice est inversible. Pour trouver son inverse, il faut encore continuer le pivot, mais cette fois dans l'autre sens, afin de transformer la partie gauche en la matrice identité : il ne reste en fait plus qu'à effectuer l'opération $L_1 \leftarrow L_1 - 2L_2$. On obtient :

$$A_3 = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & 2 & 1 \end{array}\right)$$

C'est alors que le miracle se produit : la partie droite de la matrice A_3 est précisément l'inverse de la matrice A. On a donc

$$A^{-1} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ -2 & 2 & 1 \end{pmatrix}$$

Bien sûr, il n'y a rien de miraculeux dans tout ça, puisque la partie droite de la matrice A_1 était l'identité : nous l'avons simplement multipliée par les trois mêmes matrices élémentaires P_1, P_2 et P_3 (que nous n'avons pas eu besoin d'expliciter !) qui ont servi à transformer la matrice A en l'identité, ce qui montre que la partie droite de A_4 est la matrice $P_3P_2P_1I_3 = P_3P_2P_1 = A^{-1}$.

Exercice 2.62. Dire si les matrices suivantes sont inversibles, et le cas échéant calculer leur inverse :

$$A = \begin{pmatrix} 1 & 5 & 1 \\ 0 & 2 & 3 \\ 2 & 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 3 & 2 \\ 1 & 2 & 5 \\ 1 & 1 & 1 \end{pmatrix}$$

2.2.3 Résolution des systèmes linéaires.

La méthode de Gauss permet enfin de résoudre des systèmes linéaires. Un système linéaire en les inconnues $x_1, \ldots, x_p$ est une famille finie d'équations de la forme :

$$\begin{cases} a_{1,1}x_1 + \ldots + a_{1,p}x_p = b_1 \\ a_{2,1}x_1 + \ldots + a_{2,p}x_p = b_2 \\ \qquad \cdots \\ a_{n,1}x_1 + \ldots + a_{n,p}x_p = b_n \end{cases} \tag{1}$$

où $a_{i,j}$ (pour $i \in [\![1, n]\!]$ et $j \in [\![1, p]\!]$) et $b_1, \ldots, b_n$ sont des réels.

Un système linéaire peut s'écrire comme une équation matricielle. En effet, posons $A = (a_{i,j})_{\substack{i \in [\![1,n]\!] \\ j \in [\![1,p]\!]}}$, posons $b = (b_i)_{i \in [\![1,n]\!]}$ et définissons $x = (x_j)_{j \in [\![1,p]\!]}$ comme étant le vecteur des inconnues. Alors le système est équivalent à :

$$Ax = b$$

Définition 2.63 (Second membre et matrices d'un système linéaire)**. —** *On appelle* second membre *du système 1 le vecteur*

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

On appelle matrice *de ce système la matrice*

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,p} \end{pmatrix} \in \mathscr{M}_{n,p}(\mathbb{R})$$

Et matrice complète *de ce système la matrice formée de la concaténation de* A *et de* b*, donnée par :*

$$A_b = \begin{pmatrix} a_{1,1} & \dots & a_{1,p} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,p} & b_n \end{pmatrix} \in \mathscr{M}_{n,p+1}(\mathbb{R})$$

Tous les systèmes linéaires n'admettent pas de solution, loin de là ! Par exemple, le système

$$\begin{cases} x = 1 \\ x = 0 \end{cases}$$

n'admet évidemment pas de solution.

DÉFINITION 2.64 (Système homogène, système compatible)**. —** *Soit* $\mathscr{S}$ *un système linéaire de la forme :*

$$\begin{cases} a_{1,1}x_1 + \dots + a_{1,p}x_p = b_1 \\ a_{2,1}x_1 + \dots + a_{2,p}x_p = b_2 \\ \dots \\ a_{n,1}x_1 + \dots + a_{n,p}x_p = b_n \end{cases}$$

- *Si le second membre est nul (c'est-à-dire si* $b_1 = \dots = b_n = 0$*), le système est dit homogène.*
- *Si* $\mathscr{S}$ *admet une solution, on dit que le système est compatible.*

PROPOSITION 2.65. — *Soit un système linéaire* $\mathscr{S}$ *à* n *équations et* p *inconnues admettant* $Ax = b$ *pour forme matricielle. L'ensemble des solutions de* $\mathscr{S}$ *est :*

- *soit vide,*
- *soit un ensemble de la forme* $x_0 + \ker(A)$*, où* x_0 *est une solution quelconque du système.*

Si $p = n$ *et si la matrice du système est inversible (on dit alors que l'on a affaire à un* système de Cramer*) alors* $\mathscr{S}$ *admet une unique solution donnée par* $x = A^{-1}b$*.*

THÉORÈME 2.66. — *Soit* $\mathscr{S}$ *un système d'équations linéaires de matrice* A *et de second membre* b*. Soit* A_b *la matrice complète du système* $\mathscr{S}$*. Alors toute matrice obtenue à partir d'opérations élémentaires sur* A_b *est la matrice complète d'un système équivalent à* $\mathscr{S}$*, c'est-à-dire d'un système dont les solutions sont exactement celles de* $\mathscr{S}$*.*

Soit $\mathscr{S}$ un système linéaire, d'écriture matricielle $Ax = b$ avec A inversible. Grâce à la méthode du pivot de Gauss, en faisant des opérations élémentaires sur les lignes de A, on peut transformer A en une matrice triangulaire avec des termes diagonaux non nuls. Cela permet ensuite de résoudre le système de proche en proche, en « remontant » le long des équations ; le théorème 2.66 nous assure que ces transformations auront transformé le système de départ en un système strictement équivalent, et donc que les solutions du système triangulaire obtenu seront exactement celles du système initial.

Dans le cas plus général d'une matrice A non nécessairement inversible (et même non nécessairement carrée!), le pivot de Gauss permet de transformer le système en un système échelonné qui rendra aisée la discussion sur l'existence et le nombre des solutions.

EXEMPLE 2.67. On cherche à résoudre le système linéaire suivant :

$$\left\{\begin{array}{cccccc} 2x & +y & +2z & = & 1 \\ 2x & +2y & +z & = & 0 \\ & y & +z & = & 0 \end{array}\right.$$

En effectuant l'opération $L_2 \leftarrow L_2 - L_1$, on élimine l'inconnue x de la deuxième équation. On obtient :

$$\left\{\begin{array}{cccccc} 2x & +y & +2z & = & 1 \\ & y & -z & = & -2 \\ & y & +z & = & 0 \end{array}\right.$$

Puis, en effectuant $L_3 \leftarrow L_3 - L_2$ dans ce nouveau système, on élimine l'inconnue z de la dernière équation :

$$\left\{\begin{array}{cccccc} 2x & +y & +2z & = & 1 \\ & y & -z & = & -2 \\ & & 2z & = & 2 \end{array}\right.$$

Le système est alors triangulaire et se résout en remontant de proche en proche : on obtient $z = 1$, puis $y = -1$, et enfin $x = 0$. L'unique solution du système est donc le vecteur ${}^t(0, -1, 1)$.

REMARQUE 2.68. Notez que dans l'exemple ci-dessus nous nous sommes efforcés de présenter les différentes équations en formant une colonne pour chaque variable. Cette présentation fait clairement apparaître les matrices des différents systèmes et réduit de beaucoup le nombre d'erreurs commises lors des calculs ; nous la recommandons fortement.

EXEMPLE 2.69. On veut maintenant résoudre le système suivant :

$$\left\{\begin{array}{cccccc} x & +y & & -t & = & 0 \\ 2x & & +z & +t & = & 1 \\ x & 2y & -z & -t & = & 2 \end{array}\right.$$

Cette fois, il y a quatre inconnues et le système n'a que trois équations. La matrice du système n'étant pas carrée, elle ne peut pas être inversible. D'après la proposition 2.65, soit il n'y a pas de solutions, soit il y en a une infinité. Nous allons effectuer des opérations élémentaires pour résoudre le système. On commence par éliminer la variable x des deux dernières équations (ce qui revient à « vider » la première colonne de la matrice complète du système sous sa diagonale) en effectuant les opérations élémentaires

$L_2 \leftarrow L_2 - 2L_1$ et $L_3 \leftarrow L_3 - L_1$, ce qui mène au système :

$$\begin{cases} x & +y & & -t & = & 0 \\ & -2y & +z & +3t & = & 1 \\ & y & -z & & = & 2 \end{cases}$$

On se débarrasse ensuite de la variable y dans la dernière équation grâce aux opérations successives $L_2 \leftarrow -\frac{1}{2}L_2$ (qui sert à faire apparaître le pivot de la deuxième colonne) et $L_3 \leftarrow L_3 - L_2$ qui mènent au système :

$$\begin{cases} x & +y & & -t & = & 0 \\ & y & -\frac{1}{2}z & -\frac{3}{2}t & = & -\frac{1}{2} \\ & & -\frac{1}{2}z & +\frac{3}{2}t & = & \frac{5}{2} \end{cases}$$

La matrice du système ainsi obtenu est échelonnée ; l'algorithme du pivot de Gauss s'arrête.

Ici, la variable t va servir de paramètre au système : comme on sait qu'il n'existe pas de solution unique au système, on accepte de laisser la valeur de t libre et d'exprimer les autres variables en fonction de t. En pratique, la dernière équation donne $z = 3t-5$. En injectant ceci dans la deuxième équation, on obtient $y = 3t - 3$. Puis en revenant à la première, on obtient $x = 3 - 2t$. Les solutions du système sont donc tous les vecteurs de la forme :

$$\begin{pmatrix} 3-2t \\ 3t-3 \\ 3t-5 \\ t \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \\ -5 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 3 \\ 3 \\ 1 \end{pmatrix} \qquad \text{(pour } t \in \mathbb{R}\text{)}$$

Cet ensemble de solutions a bien la forme gouvernée par le deuxième cas de la proposition 2.65 : il s'agit de l'ensemble $s_0 + \ker(A)$, où s_0 est la solution particulière ${}^t(3, -3, -5, 0)$ et A est la matrice 3×4 du système, dont le noyau s'écrit :

$$\ker(A) = \operatorname{vect}\left(\begin{pmatrix} -2 \\ 3 \\ 3 \\ 1 \end{pmatrix}\right)$$

REMARQUE 2.70. Notez que nous avons condensé la rédaction de l'exemple ci-dessus en effectuant plusieurs opérations élémentaires successives sur le système en une seule étape. Vous pouvez le faire au tableau en gardant à l'esprit quelques points :

1. Choisissez la ligne qui contient le pivot que vous souhaitez utiliser et n'effectuez que des opérations qui consistent à ajouter un multiple de cette ligne à d'autres lignes. Veillez bien à garder cette ligne de référence inchangée durant l'application de vos opérations successives ! Cela permet par exemple de « vider » une colonne sous la diagonale en une seule étape et n'a aucun risque d'occasionner un surcroît d'erreurs de calcul.
2. S'il vous est nécessaire de multiplier une ligne par un réel non nul pour y faire apparaître un pivot et que vous souhaitez utiliser directement ce pivot, par exemple, pour « vider » une colonne sous la diagonale en une seule étape, écrivez immédiatement dans le système d'arrivée la ligne contenant le pivot, résultat de votre première opération élémentaire et faites vos calculs en vous y référant.

3. En cas de doute, écrivez toutes les étapes. Une erreur coûte bien plus de temps à localiser et à corriger que n'en fait gagner une étape de raisonnement dans laquelle les opérations élémentaires se cumulent.

EXEMPLE 2.71. Soit A la matrice suivante :

$$\begin{pmatrix} 2 & 1 \\ 1 & -5 \end{pmatrix}$$

On cherche à savoir s'il existe $X = \begin{pmatrix} x \\ y \end{pmatrix}$ tel que AX soit le vecteur $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Cela revient à la résolution du système suivant :

$$\begin{cases} 2x & +y & = & 1 \\ x & -5y & = & 0 \end{cases}$$

En effectuant les opérations $L_1 \leftrightarrow L_2$ puis $L_1 \leftarrow L_1 - 2L_2$, on obtient le système suivant :

$$\begin{cases} x & -5y & = & 0 \\ & 11y & = & 1 \end{cases}$$

La deuxième équation équivaut à $y = \frac{1}{11}$. En reportant ceci dans la première équation, on obtient $x = \frac{5}{11}$. L'unique solution du système est donc le vecteur $X = \begin{pmatrix} 5/11 \\ 1/11 \end{pmatrix}$. N'oubliez pas de vérifier le résultat en calculant AX.

Exercice 2.72. Résoudre le système suivant :

$$\begin{cases} x & +2y & -z & = & 1 \\ 3x & & -7z & = & 2 \\ 0 & y & -2z & = & 5 \end{cases}$$

REMARQUE 2.73. On a vu dans la section précédente comment la résolution de système permet d'inverser une matrice. Il est bon de remarquer que dans le cas d'un système de Cramer (c'est-à-dire d'un système de n équations linéaires à n inconnues dont la matrice est inversible) il suffit à l'inverse de connaître l'inverse de la matrice pour résoudre le système. Dans le cas de l'exercice 2.72 par exemple, la matrice du système donnée par :

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 0 & -7 \\ 0 & 1 & -2 \end{pmatrix}$$

admet un inverse donné par :

$$A^{-1} = \frac{1}{16} \begin{pmatrix} 7 & 3 & -14 \\ 6 & -2 & 4 \\ 3 & -1 & -6 \end{pmatrix}$$

que l'on peut par exemple calculer en utilisant la méthode d'élimination de Gauss-Jordan. Il suffit alors d'écrire que la solution du système est donnée par :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1} \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} = \begin{pmatrix} \frac{-57}{16} \\ \frac{11}{8} \\ \frac{-29}{16} \end{pmatrix}$$

Le choix de la méthode à utiliser pour résoudre un système d'équations linéaires vous appartient donc, même s'il est bon de noter que l'inversion de matrice n'est une méthode admissible que... lorsque la matrice du système est inversible!

3 Applications linéaires et changement de base.

Dans ce chapitre, nous présentons le troisième concept central de notre cours d'algèbre linéaire : celui d'application linéaire.

3.1 Les applications linéaires.

Les *applications linéaires* sont des applications qui « conservent la structure linéaire » entre des espaces vectoriels. Autrement dit, elles conservent les additions et les multiplications externes au sens suivant :

DÉFINITION 3.1 (Application linéaire). — *Soient* E *et* F *deux espaces vectoriels et* $f : \mathrm{E} \to \mathrm{F}$ *une application. On dit que* f *est une* application linéaire *lorsque pour tout couple* x, y *d'éléments de* E, *on a*

$$f(x+y) = f(x) + f(y)$$

et lorsque pour tout x *dans* E *et tout scalaire* λ,

$$f(\lambda x) = \lambda f(x)$$

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

On remarquera que si f est linéaire, alors on a nécessairement $f(0) = 0$. En pratique, on utilisera la méthode suivante pour montrer qu'une application est linéaire :

MÉTHODE 3.2. Pour montrer qu'une application $f : \mathrm{E} \to \mathrm{F}$ est linéaire, on vérifie simplement que l'identité suivante est vérifiée :

$$\forall (x, y) \in \mathrm{E}^2, \forall \lambda \in \mathbb{R}, \qquad f(x+\lambda y) = f(x) + f(\lambda y)$$

Exercice 3.3. L'application de $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = 3x$ pour tout $x \in \mathbb{R}$ est-elle linéaire ? Et l'application nulle de $\mathbb{R}$ dans $\mathbb{R}$? Et l'application $g : \mathbb{R} \to \mathbb{R}$ définie par $g(x) = 3x+1$ pour tout $x \in \mathbb{R}$?

Exercice 3.4. On définit une application $\Delta : \mathbb{R}[\mathrm{X}] \to \mathbb{R}[\mathrm{X}]$ par

$$\Delta(\mathrm{P}) = \mathrm{P}'$$

Montrer que Δ est linéaire.

THÉORÈME 3.5. — *Soit* E *un espace vectoriel de dimension finie non nulle* n *et* $\mathscr{B} = \{e_1, \ldots, e_n\}$ *une base de* E. *Soit* F *un espace vectoriel et* $\{u_1, \ldots, u_n\}$ *une famille de vecteurs de* F. *Alors, il existe* une unique application linéaire de E dans F *telle que pour tout* $i \in [\![1, n]\!]$,

$$f(e_i) = u_i$$

REMARQUE 3.6. Le théorème précédent est très important : il dit que toute application linéaire de E dans F est uniquement déterminée par les images des vecteurs de la base de départ, c'est-à-dire par la famille $f(e_1), \ldots, f(e_n)$. Autrement dit, pour connaître parfaitement une application linéaire, il suffit de connaître une base de l'espace de départ et les images des vecteurs de cette base. Cette remarque est au fondement de la *représentation matricielle des applications linéaires* que nous étudierons plus loin.

L'exercice suivant donne une intuition de la démonstration (facile!) du théorème 3.5 :

Exercice 3.7. On se place dans $\mathbb{R}^3$, que l'on munit de sa base canonique (e_1, e_2, e_3). Soit f l'unique application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^3$ telle que $f(e_1) = (2,0,4)$, $f(e_2) = (1,1,0)$ et $f(e_3) = (5,0,0)$. Trouver l'image du vecteur ${}^t(x, y, z) \in \mathbb{R}^3$ par f.

DÉFINITION 3.8 (Inversibilité)**. —** *On dit qu'une application linéaire est inversible si et seulement si elle est bijective.*

DÉFINITION 3.9 (Isomorphisme)**. —** *Si* E *et* F *sont deux espaces vectoriels, on appelle* isomorphisme *de* E *dans* F *toute application linéaire* $f : \mathrm{E} \to \mathrm{F}$ *inversible.*

Remarquez qu'étymologiquement, un *isomorphisme* est un objet qui relie deux objets de même structure. On s'attend donc à ce que deux espaces vectoriels reliés par un isomorphisme se « ressemblent » beaucoup ; c'est le cas, en un sens précisé par le théorème 3.20 ci-dessous.

DÉFINITION 3.10 (Endomorphisme, automorphisme)**. —** *Soit* E *un espace vectoriel. Une application linéaire de* E *dans* E *est appelée* un endomorphisme. *Si cette application est également inversible, alors on dit que c'est un* automorphisme.

On note simplement $\mathscr{L}(\mathrm{E}) = \mathscr{L}(\mathrm{E}, \mathrm{E})$ *l'ensemble des endomorphismes de* E *et* GL(E) *(«* groupe linéaire de E *») l'ensemble de ses automorphismes.*

On définit le noyau et l'image d'une application linéaire de la même manière que pour les matrices.

DÉFINITION 3.11 (Image, noyau)**. —** *Soient* E *et* F *deux espaces vectoriels et* $f : \mathrm{E} \to \mathrm{F}$ *une application linéaire. Le* noyau de f, *noté* $\ker(f)$, *est l'ensemble des* x *tels que* $f(x) = 0$ *:*

$$\ker(f) = \left\{x \in \mathrm{E} \mid f(x) = 0\right\}$$

L'image de f, *notée* $\mathrm{Im}(f)$, *est l'ensemble des* $y \in \mathrm{F}$ *tels qu'il existe* $x \in \mathrm{E}$ *avec* $f(x) = y$ *:*

$$\mathrm{Im}(f) = \left\{y \in \mathrm{F} \mid \exists x \in \mathrm{E}, f(x) = y\right\}$$

On vérifiera également que si A est une matrice et que f_A l'application linéaire canoniquement associé à A (dont on rappelle qu'il s'agit de l'application $f_\mathrm{A} : x \mapsto \mathrm{A}x$ de $\mathbb{R}^m$ dans $\mathbb{R}^n$, où $n \times m$ est le format de la matrice A), alors le noyau de A (au sens des matrices) est identique au noyau de f_A (au sens des applications linéaires), et de même pour leur image.

Exercice 3.12. Soit $\Delta : \mathbb{R} \to \mathbb{R}[\mathrm{X}]$ l'opérateur de dérivation des polynômes. Quel est son noyau ? Quelle est son image ?

EXEMPLE 3.13. Soit E un espace vectoriel. Le noyau de l'application nulle est E et son image est {0}. Le noyau de l'application identité est {0} et son image est E.

EXEMPLE 3.14. Soit f l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par :

$$f\begin{pmatrix} x \\ y \end{pmatrix} = x + y$$

Son noyau est l'ensemble des vecteurs $\begin{pmatrix} x \\ y \end{pmatrix}$ tels que $x + y = 0$. On a donc :

$$\ker(f) = \left\{ \begin{pmatrix} x \\ -x \end{pmatrix} \mid x \in \mathbb{R} \right\}$$

On vérifiera qu'il s'agit du sous-espace vectoriel engendré par le vecteur $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Enfin, l'image de f est $\mathbb{R}$ tout entier. En effet, pour tout $x \in \mathbb{R}$, on a

$$f \begin{pmatrix} x \\ 0 \end{pmatrix} = x + 0 = x$$

PROPOSITION 3.15. — *Le noyau et l'image d'une application linéaire sont des sous-espaces vectoriels respectivement de son ensemble de départ et de son ensemble d'arrivée.*

DÉFINITION 3.16 (Rang d'une application linéaire)**.** — *Si* E *et* F *sont des espaces vectoriels et si* $f : \mathrm{E} \to \mathrm{F}$ *est une application linéaire telle que* $\mathrm{Im}(f)$ *est de dimension finie, alors on appelle* rang de f *et on note* $\mathrm{rg}(f)$ *la dimension de* $\mathrm{Im}(f)$*.*

PROPOSITION 3.17. — *Si* E *et* F *sont des espaces vectoriels et si* E *est de dimension finie* $n \in \mathbb{N}^*$ *et admet une base* $\{e_1, \ldots, e_n\}$*, alors pour toute application linéaire* $f \in \mathscr{L}(\mathrm{E}, \mathrm{F})$ *on a :*

$$\mathrm{Im}(f) = \mathrm{vect}\big(f(e_1), \ldots, f(e_n)\big)$$

EXEMPLE 3.18. Si $n \in \mathbb{N}^*$, l'application linéaire

$$\Delta_0 : \begin{cases} \mathbb{R}_n[\mathrm{X}] \to \mathbb{R}_{n-1}[\mathrm{X}] \\ \mathrm{P} \mapsto \mathrm{P}(\mathrm{X}) - \mathrm{P}(\mathrm{X}-1) \end{cases}$$

dont on vérifiera à titre d'exercice qu'elle est bien définie (et en particulier qu'elle est bien à valeurs dans $\mathbb{R}_{n-1}[\mathrm{X}]$) a pour image le sous-espace vectoriel de $\mathbb{R}_{n-1}[\mathrm{X}]$ engendré par la famille $\{\Delta_0(1), \Delta_0(\mathrm{X}), \ldots, \Delta_0(\mathrm{X}^n)\}$, c'est-à-dire le sous-espace vectoriel de $\mathbb{R}_{n-1}[\mathrm{X}]$ engendré par la famille $\{\Delta_0(\mathrm{X}), \ldots, \Delta_0(\mathrm{X}^n)\}$ (puisque $\Delta_0(1) = 0$) dont on montre (autre exercice !) qu'elle est à degrés échelonnés. Comme elle est de cardinal n, c'est une base de $\mathbb{R}_{n-1}[\mathrm{X}]$ et Δ_0 est surjective.

THÉORÈME 3.19. — *Soit* f *une application linéaire entre deux espaces vectoriels. Alors,* f *est injective si et seulement si* $\ker(f) = \{0\}$*.*

THÉORÈME 3.20. — *Si* E *et* F *sont deux espaces vectoriels, si* E *est de dimension finie non nulle* n *et admet* $\{e_1, \ldots, e_n\}$ *pour base et si* $f \in \mathscr{L}(\mathrm{E}, \mathrm{F})$*, alors* f *est un isomorphisme si et seulement si* $\{f(e_1), \ldots, f(e_n)\}$ *est une base de* F*. Si cette condition est vérifiée, on en déduit que* F *est de dimension finie égale à* n*.*

DÉFINITION 3.21 (Itérées d'une application linéaire)**.** — *On définit également les itérées d'une application linéaire* $f : \mathrm{E} \to \mathrm{F}$ *comme suit : on pose* $f^0 = \mathrm{Id}$ *et pour tout* $n > 0$ *on pose* $f^n = \underbrace{f \circ \ldots \circ f}_{n \text{ fois}}$*.*

Remarquez que si la notation utilisée rappelle la notation des puissances (et on verra dans la dernière section de ce chapitre que ce n'est pas un hasard!), on ne considère pas dans la définition ci-dessus un *produit* d'applications linéaires, ce qui n'a aucun sens en général, mais bien leur *composée*.

Exercice 3.22. Soit A un endomorphisme d'un espace vectoriel E tel que $f^2 = 0$. Montrer que $\mathrm{Im}(f) \subset \ker(f)$. La réciproque est-elle vraie?

On termine cette section par l'étude d'un type particulier d'applications linéaires : les formes linéaires.

DÉFINITION 3.23. — *Soit* E *un espace vectoriel. Une* forme linéaire sur E *est une application linéaire de* E *dans* $\mathbb{R}$.

Exercice 3.24. Quel est le rang d'une forme linéaire?

DÉFINITION 3.25. — *Soit* E *un espace vectoriel. Le noyau d'une forme linéaire* non nulle *sur* E *s'appelle un* hyperplan.

THÉORÈME 3.26. — *Soit* E *un espace vectoriel de dimension finie n. Alors, tout sous-espace vectoriel de dimension $n-1$ est un hyperplan, et réciproquement tout hyperplan est de dimension $n-1$.*

Exercice 3.27. Décrire tous les hyperplans de $\mathbb{R}$ et de $\mathbb{R}^2$.

PROPOSITION 3.28. — *Tout supplémentaire d'un hyperplan est une droite vectorielle.*

3.2 Théorème du rang.

Le théorème suivant est d'une importance capitale.

THÉORÈME 3.29 (Théorème du rang)**. —** *Soit* E *un espace vectoriel de dimension finie $n \in \mathbb{N}$. Soit* F *un espace vectoriel quelconque et soit $f : E \to F$ une application linéaire. Alors,*

$$n = \dim \ker(f) + \mathrm{rg}(f)$$

Noter que n est la dimension de l'espace de départ de f et non celle de son espace d'arrivée! C'est une confusion fréquente. Une manière de mémoriser la formule est sans doute d'en apprendre la démonstration, à la fois simple et subtile. Nous l'esquissons dans la remarque suivante.

REMARQUE 3.30. La démonstration de ce théorème utilise des techniques intéressantes qui aident à répondre à certaines questions d'oral. On reprend les notations de l'énoncé du théorème. Alors $\ker(f)$ est un sous-espace vectoriel de E : il admet donc un supplémentaire $\tilde{E}$, c'est-à-dire que $E = \tilde{E} \oplus \ker(f)$. Considérons maintenant la restriction de f à $\tilde{E}$, notée $\tilde{f}$. On a $\ker(\tilde{f}) = \ker(f) \cap \tilde{E}$ (pourquoi?). Or, comme $\tilde{E}$ et $\ker(f)$ sont en somme directe, cette intersection est réduite à $\{0\}$: donc $\ker(\tilde{f}) = \{0\}$, et $\tilde{f}$ est injective. En particulier, $\tilde{f}$ est un isomorphisme de $\tilde{E}$ vers $\mathrm{Im}(f)$ (pourquoi?), donc $\dim(\tilde{E}) = \dim \mathrm{Im}(f)$ d'après le théorème 3.20. Comme $\dim(\tilde{E}) = n - \dim \ker(f)$, on en déduit le résultat.

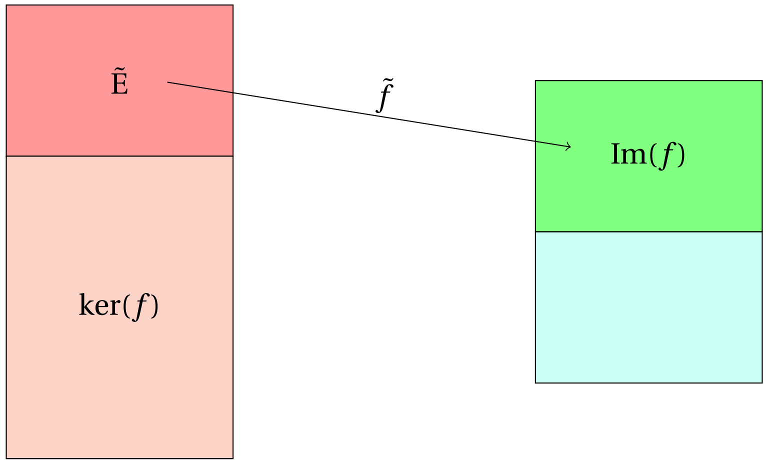


FIGURE 6 – Une illustration de la démonstration du théorème du rang.

THÉORÈME 3.31 (Formule de Grassmann)**. —** *Soient* F, G *deux sous-espaces vectoriels d'un espace vectoriel* E *de dimension finie. Alors*

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

On insiste sur le fait que l'espace ambiant doit être de dimension finie : cette formule n'a pas de sens si les dimensions considérées sont infinies ! La formule de Grassmann donne aussi un critère bien utile pour montrer que des sous-espaces sont en somme directe, en regardant seulement leurs dimensions.

REMARQUE 3.32. La formule de Grassmann se réécrit $\dim(F+G)=\dim(F)+\dim(G)-\dim(F\cap G)$. Cette formule est l'exact analogue de la formule donnant le cardinal de l'union de deux ensembles finis [17] : en effet, si A et B sont deux sous-ensembles finis d'un même ensemble, alors

$$\mathrm{card}(A\cup B)=\mathrm{card}(A)+\mathrm{card}(B)-\mathrm{card}(A\cap B)$$

Notez que la formule de Grassmann permet de retrouver très rapidement la proposition 1.74.

Le théorème suivant est extrêmement important. Il faut absolument comprendre que ses conclusions ne sont vraies que lorsque l'espace considéré est de dimension finie !

THÉORÈME 3.33. — *Soit* E *un espace vectoriel de dimension finie et soit f un endomorphisme de* E*. Alors il est équivalent de dire que f est bijectif, injectif ou surjectif.*

On retiendra la formulation lapidaire suivante :

E **de dimension finie** et $f : E \to E$. Alors :
f injectif $\Leftrightarrow$ f surjectif $\Leftrightarrow$ f bijectif

17. Ce n'est pas un hasard puisque la formule de Grassmann peut se démontrer en utilisant des réunions de bases bien choisies et en examinant leurs cardinaux respectifs.

L'économie de pensée que permet ce théorème est considérable : pour montrer qu'une application est bijective, on doit en général prouver qu'elle est à la fois injective et surjective, ce qui en général n'est pas une mince affaire. Grâce à ce théorème, il suffit de montrer que :

1. L'application considérée est linéaire, ce qui est en général très simple.
2. Elle est soit injective, soit surjective : il suffit de choisir le fait le plus simple à établir.

Exercice 3.34. Démontrer le théorème 3.33.

Notons que ce théorème admet une généralisation au cas d'une application linéaire $f : \mathrm{E} \to \mathrm{F}$ entre deux espaces vectoriels **de même dimension finie**, mais qu'il est **faux** en-dehors de ce cas : par exemple, une application linéaire $f : \mathbb{R}^2 \to \mathbb{R}^3$ ne saurait être surjective (exercice : montrez-le grâce au théorème du rang!) mais peut tout à fait être injective. De même, il n'est peut-être pas superflu de répéter que ce théorème n'est pas vrai si E est de dimension infinie, comme on le voit en considérant l'application de dérivation des polynômes étudiée dans l'exercice 3.12, qui est un endomorphisme surjectif mais non injectif de $\mathbb{R}[\mathrm{X}]$.

EXEMPLE 3.35 (Polynômes interpolateurs de Lagrange)**.** Nous donnons maintenant une illustration impressionnante des résultats précédents. Soit n un entier naturel. On se donne $n+1$ réels $x_0, \dots, x_n$ tous distincts et $n+1$ réels quelconques $y_0, \dots, y_{n+1}$. Nous inspirant d'un problème soulevé en leur temps par les contemporains de Joseph-Louis Lagrange [18], nous posons la question suivante :

Existe-t-il un polynôme $\mathrm{P} \in \mathbb{R}_n[\mathrm{X}]$ *tel que pour tout* $i \in [\![0, n]\!]$ *on ait* $\mathrm{P}(x_i) = y_i$ *?*

On pourrait construire un tel polynôme à la main [19] : il suffit de bricoler de tels polynômes et d'en trouver une expression explicite. Mais les résultats précédents d'algèbre linéaire permettent de démontrer l'existence d'un tel polynôme presque sans effort, le tout avec élégance et raffinement. Considérons donc l'application $\varphi : \mathbb{R}_n[\mathrm{X}] \to \mathbb{R}^{n+1}$, définie par :

$$\forall \mathrm{P} \in \mathbb{R}_n[\mathrm{X}], \quad \varphi(\mathrm{P}) = \big(\mathrm{P}(x_0), \dots, \mathrm{P}(x_n)\big)$$

On vérifie sans difficulté qu'il s'agit d'une application linéaire. Son noyau est l'ensemble des polynômes $\mathrm{P} \in \mathbb{R}_n[\mathrm{X}]$, tels que $\mathrm{P}(x_0) = \dots = \mathrm{P}(x_n) = 0$; mais un tel polynôme P a $n+1$ racines *distinctes* (on rappelle que les x_i sont distincts) et est de degré au plus n : il est donc nécessairement nul. On en déduit que $\ker(\varphi) = \{0\}$. L'application φ est donc injective. Comme les deux espaces $\mathbb{R}_n[\mathrm{X}]$ et $\mathbb{R}^{n+1}$ ont la même dimension $n+1$, l'application φ est un isomorphisme (théorème 3.33). **Elle est donc surjective** : pour tout $(y_0, \dots, y_n) \in \mathbb{R}^{n+1}$, il existe un polynôme $\mathrm{P} \in \mathbb{R}_n[\mathrm{X}]$ tel que $\varphi(\mathrm{P}) = (y_0, \dots, y_n)$, ce qui est exactement ce que nous cherchons.

En résumé, pour montrer un résultat d'existence, nous nous sommes ramenés à prouver un résultat de surjectivité, et en fait nous n'avons eu qu'à démontrer un résultat beaucoup plus simple d'injectivité.

18. 1736 - 1813. Fierté de la science hexagonale, originaire du Piémont turinois.

19. Et en fait, on vous invite à le faire en exercice. Indication : considérez les polynômes P_i définis pour tout $i \in [\![0, n]\!]$ par $\mathrm{P}_i(\mathrm{X}) = \prod_{k \neq i} \frac{\mathrm{X} - x_k}{x_i - x_k}$.

3.3 Les liens entre applications linéaires et matrices.

Dans toute cette section, on se donne $n, m \in \mathbb{N}^*$.

Soit $\mathscr{B} = (e_1, ..., e_n)$ une base d'un espace vectoriel E de dimension n, et F un espace vectoriel de dimension m. Soit $f : \mathrm{E} \to \mathrm{F}$ une application linéaire. En vertu du théorème 3.5, pour connaître f, il suffit de connaître l'image de $e_1, ..., e_n$. Notons $u_i = f(e_i)$ pour tout $i \in [\![1, n]\!]$.

Fixons maintenant une base $\mathscr{B}' = \{e'_1, ..., e'_m\}$ de F. En vertu du théorème 1.40, chaque vecteur de F s'exprime de manière unique comme une combinaison linéaire des e'_i. Ainsi, pour tout $i \in [\![1, n]\!]$, il existe un unique n-uplet de réels $(a_{1,i}, ..., a_{m,i})$ tels que

$$u_i = a_{1,i}e'_1 + ... + a_{m,i}e'_m$$

Définissons alors la matrice $\mathrm{A} \in \mathscr{M}_{m,n}$ dont le terme d'indices (i, j) est $a_{i,j}$. Cette matrice est appelée *matrice de f dans les bases $\mathscr{B}$ et $\mathscr{B}'$*. On note alors

$$\mathrm{A} = \mathrm{mat}_{\mathscr{B}}^{\mathscr{B}'}(f)$$

A est donc la matrice dont les colonnes sont les vecteurs coordonnées des vecteurs $f(e_i)$ sur la base $\mathscr{B}'$; on écrit parfois les vecteurs de $\mathscr{B}'$ en colonne à droite de la matrice et les vecteurs $f(e_i)$ en ligne au-dessus d'elle pour souligner cette décomposition, comme ceci :

$$\mathrm{mat}_{\mathscr{B}}^{\mathscr{B}'}(f) = \begin{array}{c} \begin{array}{ccc} f(e_1) & \dots & f(e_n) \end{array} \\ \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} \end{array} \begin{array}{c} \\ e'_1 \\ \\ e'_m \end{array}$$

EXEMPLE 3.36. On se place dans le plan $\mathbb{R}^2$ et on considère l'application identité, définie par $\mathrm{Id}(x) = x$ pour tout $x \in \mathbb{R}^2$. Choisissons une base $\mathscr{B} = \{e_1, e_2\}$, par exemple la base canonique $e_1 = (1, 0)$ et $e_2 = (0, 1)$. Alors $\mathrm{Id}(e_1) = e_1$ et $\mathrm{Id}(e_2) = e_2$, d'où :

$$\mathrm{mat}_{\mathscr{B}}^{\mathscr{B}}(\mathrm{Id}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I}_2$$

EXEMPLE 3.37. On munit $\mathbb{R}^3$ de la base canonique $\mathscr{B}$ et on munit $\mathbb{R}^2$ de la base canonique $\mathscr{B}'$. Soit f l'application linéaire définie par :

$$\forall x, y, z \in \mathbb{R}, \quad f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y \\ x + y + z \end{pmatrix}$$

La matrice de f dans les bases $\mathscr{B}$ et $\mathscr{B}'$ est de taille 2×3. Pour tout $i \in [\![1, 2]\!]$, sa i-ème colonne est formée des coordonnées de $f(e_i)$ dans la base (e'_1, e'_2). Il suffit donc de calculer $f(e_1), f(e_2)$ et $f(e_3)$. On obtient :

$$\mathrm{mat}_{\mathscr{B}}^{\mathscr{B}'}(f) = \begin{array}{c} \begin{array}{ccc} f(e_1) & f(e_2) & f(e_3) \end{array} \\ \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \end{array} \begin{array}{c} \\ e'_1 \\ e'_2 \end{array}$$

Exercice 3.38. Trouver la matrice (dans les bases canoniques) de l'application $f : \mathbb{R}^4 \to \mathbb{R}^7$ définie par

$$\forall x,y,z,t \in \mathbb{R}^4, \quad f\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x+y \\ y-t \\ 5t \\ 3t+8z \\ 0 \\ x+y+5z-3t \\ 2x-y \end{pmatrix}$$

THÉORÈME 3.39. — *Si* E *est un espace vectoriel de dimension* n *et si* F *est un espace vectoriel de dimension* m *et si* $\mathscr{B}$ *est une base de* E *et* $\mathscr{B}'$ *une base de* F*, alors l'application :*

$$\Phi : \begin{cases} \mathscr{L}(\mathrm{E},\mathrm{F}) \to \mathscr{M}_{m,n}(\mathbb{R}) \\ f \mapsto \mathrm{mat}_{\mathscr{B}}^{\mathscr{B}'}(f) \end{cases}$$

est un isomorphisme. En particulier, on a $\dim(\mathscr{L}(\mathrm{E},\mathrm{F})) = nm$ *et* $\mathscr{L}(\mathrm{E}) = n^2$.

Nous laissons au lecteur le soin de transposer au cadre des applications linéaires les résultats sur les matrices étudiés à la section 2.1 et au cadre des matrices les résultats sur les applications linéaires étudiés aux sections 3.1 et 3.2. En particulier, nous l'invitons à regarder en détail le théorème 2.47 à la lumière du théorème du rang.

REMARQUE 3.40 (d'une certaine importance)**.** Nous y reviendrons lourdement par la suite, mais retenons tout de suite qu'une application linéaire ne possède pas une unique matrice : il y autant de matrices que de choix de bases au départ et à l'arrivée. Cela implique notamment qu'une même application linéaire peut être représentée par des matrices de formes très différentes les unes des autres selon le choix des bases dans lesquelles on la représente. L'objet de la dernière partie de ce cours sera justement de trouver des bases dans lesquelles les matrices d'un endomorphisme donné ont la forme la plus simple possible.

Exercice 3.41. On munit $\mathbb{R}^2$ de la base canonique $\mathscr{B}$ et de la base $\mathscr{B}' = (e'_1, e'_2)$ définie par $e'_1 = (4,0), e'_2 = (1,1)$. On définit l'application $f : \mathbb{R}^2 \to \mathbb{R}^2$ par :

$$\forall x,y \in \mathbb{R}, \quad f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 2y \end{pmatrix}$$

Trouver la matrice de l'application f :
- Dans la base $\mathscr{B}$ au départ et à l'arrivée.
- Dans la base $\mathscr{B}$ au départ et $\mathscr{B}'$ à l'arrivée.
- Dans la base $\mathscr{B}'$ au départ et à l'arrivée.
- Dans la base $\mathscr{B}'$ au départ et $\mathscr{B}$ à l'arrivée.

PROPOSITION 3.42. — *Si* E, F *et* G *sont trois espaces vectoriels de dimension finie dotés de bases respectives* $\mathscr{B}, \mathscr{B}'$ *et* $\mathscr{B}''$ *et si* $f : \mathrm{E} \to \mathrm{F}$ *et* $g : \mathrm{F} \to \mathrm{G}$ *sont deux applications linéaires, alors :*

$$\mathrm{mat}_{\mathscr{B}}^{\mathscr{B}''}(g \circ f) = \mathrm{mat}_{\mathscr{B}'}^{\mathscr{B}''}(g)\mathrm{mat}_{\mathscr{B}}^{\mathscr{B}'}(f)$$

REMARQUE 3.43. En d'autres termes, la multiplication matricielle traduit la composition des applications linéaires, ce qui justifie sa définition plutôt alambiquée. L'exemple suivant montre que cela permet entre autres de calculer très facilement des composées d'applications linéaires sans cela plutôt fastidieuses à déterminer.

EXEMPLE 3.44. Considérons les applications $f, g : \mathbb{R}^3 \to \mathbb{R}^3$ définies par :

$$f : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 2x+y-z \\ 3x+2y+z \\ x+z \end{pmatrix} \quad \text{et} \quad g : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 5x+3z \\ 2x-7y \\ x+y+z \end{pmatrix}$$

Alors l'application $g \circ f : \mathbb{R}^3 \to \mathbb{R}^3$ admet pour matrice dans la base canonique $\mathscr{B}$ de $\mathbb{R}^3$ le produit

$$\mathrm{mat}_{\mathscr{B}}^{\mathscr{B}}(g)\mathrm{mat}_{\mathscr{B}}^{\mathscr{B}}(f) = \begin{pmatrix} 2 & 1 & -1 \\ 3 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 & 0 & 3 \\ 2 & -7 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 11 & -8 & 5 \\ 20 & -13 & 10 \\ 6 & 1 & 4 \end{pmatrix}$$

et s'écrit donc :

$$g \circ f : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 11x-8y+5z \\ 20x-13y+10z \\ 6x+y+4z \end{pmatrix}$$

Nous nous intéressons maintenant au changement de base : si l'on dispose d'une application linéaire $f : \mathrm{E} \to \mathrm{F}$, comment seront modifiées les matrices de f si l'on change les bases des espaces vectoriels E et F ?

DÉFINITION 3.45 (Matrices équivalentes)**. —** *Soient* $\mathrm{A}, \mathrm{B} \in \mathscr{M}_{n,m}(\mathbb{R})$*. On dit que* A *est* équivalente à B *(ou que* A et B sont équivalentes*) s'il existe deux matrices inversibles* $\mathrm{P} \in \mathrm{GL}_n(\mathbb{R})$ *et* $\mathrm{Q} \in \mathrm{GL}_m(\mathbb{R})$ *telles que* $\mathrm{A} = \mathrm{PBQ}^{-1}$.

Nous étudierons plus précisément la notion de changement de base par la suite, mais nous pouvons déjà révéler que dans le cadre de la définition précédente, A et B sont équivalentes si et seulement si elles représentent une même application linéaire $f : \mathbb{R}^m \to \mathbb{R}^n$ dans deux couples de bases différents [20]. On verra que Q est alors la matrice de passage de l'ancienne à la nouvelle base de $\mathbb{R}^m$ et que P est la matrice de passage de l'ancienne à la nouvelle base de $\mathbb{R}^n$.

PROPOSITION 3.46. — *Deux matrices de même taille sont équivalentes si et seulement si elles ont le même rang.*

C'est de cette proposition que l'on déduit qu'une matrice et sa transposée ont même rang. Nous vous laissons chercher pourquoi [21] !

DÉFINITION 3.47 (Matrices semblables)**. —** *Soient* $\mathrm{A}, \mathrm{B} \in \mathscr{M}_n(\mathbb{R})$*. On dit que* A *est* semblable à B *(ou que* A et B sont semblables*) s'il existe une matrice inversible* $\mathrm{P} \in \mathrm{GL}_n(\mathbb{R})$ *telle que* $\mathrm{A} = \mathrm{PBP}^{-1}$.

---

20. C'est-à-dire qu'il existe deux bases $\mathscr{B}_m$ et $\mathscr{B}'_m$ de $\mathbb{R}^m$ et deux bases $\mathscr{B}_n$ et $\mathscr{B}'_n$ de $\mathbb{R}^n$ telles que $\mathrm{A} = \mathrm{mat}_{\mathscr{B}_m}^{\mathscr{B}_n}(f)$ et $\mathrm{B} = \mathrm{mat}_{\mathscr{B}'_m} = (f)^{\mathscr{B}'_n}$.

21. Indice : si pour tout $r \in [\![1, \min(m,n)]\!]$ on définit la matrice J_r comme la matrice dont tous les coefficients sont nuls à l'exception des coefficients d'indice (i,i) pour $i \in [\![1, r]\!]$ qui valent 1, alors le rang de J_r et celui de ${}^t\mathrm{J}_r$ sont faciles à calculer.

On verra plus loin que deux matrices sont semblables si elles représentent un même endomorphisme de $\mathbb{R}^n$ dans deux bases différentes [22] : la matrice P est alors la matrice de passage entre ces deux bases. Il peut alors être intéressant de classer toutes les matrices en regroupant toutes les matrices qui sont semblables entre elles et de chercher parmi une classe de similitude donnée quelles sont les matrices dont la forme est la plus simple. C'est la théorie de la *réduction des endomorphismes*, dont nous aurons un aperçu dans le dernier chapitre.

Exercice 3.48. Soient $A, B, C \in \mathcal{M}_n(\mathbb{R})$. Montrer que :

1. Si A est semblable à B alors B est semblable à A.
2. Si A est semblable à B et si B est semblable à C, alors A est semblable à C.
3. A est semblable à elle-même.

PROPOSITION 3.49. — *Si* A *est semblable à* B *et si* P *est une matrice inversible telle que* $A = PBP^{-1}$ *, alors pour tout entier naturel n on a :*

$$A^n = PB^nP^{-1}$$

La démonstration de cette proposition est simple : si $A = PDP^{-1}$, alors :

$$A^n = \underbrace{(PDP^{-1}) \times \ldots \times (PDP^{-1})}_{n \text{ fois}}$$

En utilisant l'associativité du produit matriciel, on a donc

$$A^n = PD(P^{-1}P)D(P^{-1}P)D \times \ldots \times D(P^{-1}P)DP^{-1}$$

Les termes en PP^{-1} se simplifient, et on obtient $A^n = PD^nP^{-1}$.

Cette proposition est très utile : souvent, pour calculer les puissances d'une matrice A, on cherche si elle est semblable à une matrice B qui est très simple (par exemple, une matrice diagonale) et pour laquelle les puissances sont faciles à calculer. Au lieu de calculer (difficilement) les puissances de A, il suffit de calculer celles de B qui sont plus faciles à trouver, et retourner dans la base de départ.

3.4 Changement de base.

Soient $\mathscr{B} = (e_i)_{i \in [\![1,n]\!]}$ et $\mathscr{B}' = (e'_i)_{i \in [\![1,n]\!]}$ deux bases d'un même espace vectoriel de dimension finie $n \in \mathbb{N}^*$. Supposons que nous ayons un vecteur x et que nous connaissions ses coordonnées dans la base $\mathscr{B}$, disons $x = a_1 e_1 + \ldots + a_n e_n$. Comment trouver ses coordonnées dans la base $\mathscr{B}'$?

Il faut pour cela connaître les coordonnées des vecteurs de la première base $\mathscr{B}$ dans la deuxième base $\mathbb{B}'$. Supposons par exemple que

$$e_i = \lambda_{1,i} e'_1 + \ldots + \lambda_{n,i} e'_n$$

22. C'est-à-dire que si A et B sont deux matrices semblables de $\mathcal{M}_n(\mathbb{R})$, il existe $f : \mathbb{R}^n \to \mathbb{R}^n$ linéaire et qu'il existe deux bases $\mathscr{B}$ et $\mathscr{B}'$ de $\mathbb{R}^n$ telles que $A = \mathrm{mat}_{\mathscr{B}}^{\mathscr{B}}(f)$ et $B = \mathrm{mat}_{\mathscr{B}'}^{\mathscr{B}'}(f)$.

Alors on peut écrire :

$$\begin{aligned} x &= a_1 e_1 + \ldots + a_n e_n \\ &= a_1\big(\lambda_{1,1} e'_1 + \ldots + \lambda_{n,1} e'_n\big) + \ldots + a_n\big(\lambda_{1,n} e'_1 + \ldots + \lambda_{n,n} e'_n\big) \end{aligned}$$

En regroupant les termes en e'_i, on obtient

$$x = \big(a_1\lambda_{1,1} + \ldots + a_n\lambda_{n,1}\big)e'_1 + \ldots + \big(a_1\lambda_{n,1} + \ldots + a_n\lambda_{n,n}\big)e'_n$$

En vertu de l'unicité de la décomposition d'un vecteur dans la base $\mathscr{B}'$ (c'est le théorème 1.40), nous voyons que pour tout $i \in [\![1, n]\!]$, la i-ème coordonnée de x dans cette base est :

$$a_1\lambda_{i,1} + \ldots + a_n\lambda_{i,n}$$

On reconnaît alors le i-ème terme de la multiplication matrice-vecteur Λx, où Λ est la matrice $(\lambda_{i,j})_{i,j\in[\![1,n]\!]}$. Cette matrice est précisément appelée la *matrice de passage de la base $\mathscr{B}'$ à la base $\mathscr{B}$*.

DÉFINITION 3.50 (Matrice de passage)**. —** *La* matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$*, notée* $C(\mathscr{B},\mathscr{B}')$*, est la matrice dont la i-ème colonne donne les coordonnées du vecteur e'_i dans la base $\mathscr{B}$.*

REMARQUE 3.51 (source de vertiges)**.** La nomenclature est parfois mystérieuse. Ce que l'on appelle « matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{B}'$ » est la matrice qui sert à exprimer les vecteurs de $\mathscr{B}'$ dans la base $\mathscr{B}$. Autrement dit, si l'on dispose du vecteur X des coordonnées de x dans la base $\mathscr{B}$ et qu'on veut celles de x dans la base $\mathscr{B}'$, il faudra multiplier X par... la matrice de passage de $\mathscr{B}'$ à $\mathscr{B}$. On retiendra donc l'adage suivant :

Soit X *le vecteur des coordonnées de x dans la base $\mathscr{B}$, et soit* X′ *les coordonnées x dans la base $\mathscr{B}'$. Si* P *est la matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{B}'$, alors*

$$X = PX'$$

REMARQUE 3.52. La notation $C(\mathscr{B},\mathscr{B}')$ n'est pas standard et il est nécessaire de préciser ce qu'elle désigne lorsque vous l'utilisez (ou plus généralement lorsque vous utilisez n'importe quelle notation pour une matrice de passage).

EXEMPLE 3.53. On souhaite écrire le polynôme $X + 3X^2$ dans la base (exercice : pourquoi en est-ce une ?) $\mathscr{B}' = \{2, X-1, X^2+1\}$ de $\mathbb{R}_2[X]$. La matrice de passage de la base canonique $\mathscr{B}$ de $\mathbb{R}_2[X]$ dans $\mathscr{B}'$ est :

$$C(\mathscr{B},\mathscr{B}') = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Si l'on note Y le vecteur des coordonnées de $X + 3X^2$ dans $\mathscr{B}'$, on sait que :

$$\begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} = C(\mathscr{B},\mathscr{B}')Y$$

On détermine donc Y en inversant $\mathrm{C}(\mathscr{B},\mathscr{B}')$, ce qui donne :

$$\mathrm{Y} = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}$$

On vérifie pour finir qu'on a effectivement $\mathrm{X} + 3\mathrm{X}^2 = -2 + (\mathrm{X}-1) + 3(\mathrm{X}^2+1)$.

PROPOSITION 3.54. — *La matrice* $\mathrm{C}(\mathscr{B},\mathscr{B}')$ *est la matrice de l'application identité*

$$\mathrm{Id} : \mathrm{E} \to \mathrm{E}$$

dans les bases $\mathscr{B}'$ *(au départ) et* $\mathscr{B}$ *(à l'arrivée).*

PROPOSITION 3.55. — *La matrice de passage* $\mathrm{C}(\mathscr{B},\mathscr{B}')$ *est inversible et :*

$$\big(\mathrm{C}(\mathscr{B},\mathscr{B}')\big)^{-1} = \mathrm{C}(\mathscr{B}',\mathscr{B})$$

PROPOSITION 3.56. — *Toute matrice inversible* $\mathrm{A} \in \mathrm{GL}_n(\mathbb{R})$ *est une matrice de passage entre la base canonique et la base de* $\mathbb{R}^n$ *formée des vecteurs colonnes de* A.

Soit E un espace vectoriel de dimension n et F un espace vectoriel de dimension m. On fixe une base $\mathscr{B}_\mathrm{E}$ de E et une base $\mathscr{B}_\mathrm{F}$ de F. Soit $u : \mathrm{E} \to \mathrm{F}$ une application linéaire et soit A sa matrice dans les bases choisies. Que se passe-t-il lorsqu'on change de base au départ et à l'arrivée ?

Donnons-nous deux autres bases $\mathscr{B}'_\mathrm{E} = \{e'_1, \dots, e'_n\}$ et $\mathscr{B}'_\mathrm{F} = \{f'_1, \dots, f'_m\}$, de (respectivement) E et F. Notons $\tilde{\mathrm{A}}$ la matrice de u dans ces nouvelles bases. Il s'agit de la matrice de taille $m \times n$ dont la i-ème colonne donne les coordonnées de l'image du vecteur e'_i par u. Le théorème suivant donne la relation qui existe entre A et $\tilde{\mathrm{A}}$:

THÉORÈME 3.57 (Formule du changement de base)**. —** *On reprend les notations ci-dessus. Alors,*

$$\tilde{\mathrm{A}} = \mathrm{C}(\mathscr{B}'_\mathrm{F}, \mathscr{B}_\mathrm{F}) \times \mathrm{A} \times \mathrm{C}(\mathscr{B}_\mathrm{E}, \mathscr{B}'_\mathrm{E})$$

Autrement dit,

$$\mathrm{mat}_{\mathscr{B}'_\mathrm{E}}^{\mathscr{B}'_\mathrm{F}}(u) = \mathrm{C}(\mathscr{B}'_\mathrm{F}, \mathscr{B}_\mathrm{F}) \times \mathrm{mat}_{\mathscr{B}_\mathrm{E}}^{\mathscr{B}_\mathrm{F}}(u) \times \mathrm{C}(\mathscr{B}_\mathrm{E}, \mathscr{B}'_\mathrm{E})$$

Certains visualisent mieux cette formule lorsqu'elle est écrite sur un diagramme :

$$\begin{tikzcd}
(\mathrm{E}, \mathscr{B}_\mathrm{E}) \arrow[r, "\mathrm{A}"] & (\mathrm{F}, \mathscr{B}_\mathrm{F}) \arrow[d, "\mathrm{C}(\mathscr{B}'_\mathrm{F}, \mathscr{B}_\mathrm{F})"] \\
(\mathrm{E}, \mathscr{B}'_\mathrm{E}) \arrow[u, "\mathrm{C}(\mathscr{B}_\mathrm{E}, \mathscr{B}'_\mathrm{E})"] \arrow[r, "\tilde{\mathrm{A}}"] & (\mathrm{F}, \mathscr{B}'_\mathrm{F})
\end{tikzcd}$$

FIGURE 7 – Diagramme illustrant la formule de changement de base. La matrice $\tilde{\mathrm{A}}$ dans les nouvelles bases s'exprime en fonction de la matrice dans les anciennes bases via la formule $\tilde{\mathrm{A}} = \mathrm{C}(\mathscr{B}'_\mathrm{F}, \mathscr{B}_\mathrm{F})\mathrm{A}\mathrm{C}(\mathscr{B}_\mathrm{E}, \mathscr{B}'_\mathrm{E})$.

Les théorèmes de changement de base sont d'une grande importance, mais les formules ne sont ni intuitives ni évidentes à mémoriser, notamment en raison des notations difficiles. Il est conseillé de s'exercer à retrouver tout seul ces résultats : c'est le seul moyen d'acquérir les réflexes nécessaires.

Évoquons pour finir deux résultats que nous avons annoncés plus haut :

THÉORÈME 3.58. — *Soient* $A, B \in \mathcal{M}_{n,m}$. *Il existe* $P \in GL_n(\mathbb{R})$ *et* $Q \in GL_m(\mathbb{R})$ *telles que* $A = PBQ^{-1}$ *si et seulement s'il existe* $f \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$, *deux bases* $\mathcal{B}_m$ *et* $\mathcal{B}'_m$ *de* $\mathbb{R}^m$ *et deux bases* $\mathcal{B}_n$ *et* $\mathcal{B}'_n$ *de* $\mathbb{R}^n$ *telles que* $A = \mathrm{mat}_{\mathcal{B}_m}^{\mathcal{B}_n}(f)$ *et* $B = \mathrm{mat}_{\mathcal{B}'_m}^{\mathcal{B}'_n}(f)$. *Dans ce cas, on a* $P = C(\mathcal{B}_n, \mathcal{B}'_n)$ *et* $Q = C(\mathcal{B}_m, \mathcal{B}'_m)$.

En d'autres termes et comme nous l'avons déjà annoncé, deux matrices sont équivalentes si et seulement si elles représentent la même application linéaire dans deux couples de bases différents.

THÉORÈME 3.59. — *Soient* $A, B \in \mathcal{M}_n(\mathbb{R})$. *Il existe* $P \in GL_n(\mathbb{R})$ *telle que* $A = PBP^{-1}$ *si et seulement s'il existe* $f : \mathbb{R}^n \to \mathbb{R}^n$ *linéaire et existe deux bases* $\mathcal{B}$ *et* $\mathcal{B}'$ *de* $\mathbb{R}^n$ *telles que* $A = \mathrm{mat}_{\mathcal{B}}^{\mathcal{B}}(f)$ *et* $B = \mathrm{mat}_{\mathcal{B}'}^{\mathcal{B}'}(f)$. *Dans ce cas, on a* $P = C(\mathcal{B}, \mathcal{B}')$.

En d'autres termes et comme nous l'avons déjà annoncé, deux matrices sont semblables si et seulement si elles représentent le même endomorphisme dans deux bases différentes.

4 Réduction des endomorphismes.

Soit E un espace vectoriel et f un endomorphisme de E. Nous voulons étudier comment f agit sur E : quel est son noyau ? Quelle est son image ? Comment se comporte f sur tel ou tel sous-espace vectoriel ? La réduction des endomorphismes a pour objectif de réduire l'étude de l'action de f sur E tout entier à l'étude de l'action de f sur des sous-espaces vectoriels plus petits et bien choisis. Il s'agit d'une théorie très vaste que nous n'allons qu'effleurer.

4.1 Sous-espaces stables et polynômes annulateurs.

DÉFINITION 4.1 (Sous-espace stable par un endomorphisme)**. —** *On dit qu'un sous-espace vectoriel* F *de* E *est* f-stable*, ou* stable par f*, lorsque* $f(\mathrm{F}) \subset \mathrm{F}$.

Autrement dit, un sous-espace est stable par F si $f(x)$ « reste dans F » lorsque x appartient à F.

DÉFINITION 4.2 (Endomorphisme induit sur un sous-espace stable)**. —** *Si* F *est stable par* f*, alors* f *induit un endomorphisme de* F*, noté* f_{F}*, défini par* $f_{\mathrm{F}}(x) = f(x)$ *pour tout* $x \in \mathrm{F}$.

REMARQUE 4.3. L'endomorphisme f_{F} n'est pas la simple restriction $f_{|\mathrm{F}}$ de f à F : son ensemble d'arrivée est F et non E. Cette remarque d'apparence tatillonne a des conséquences évidentes sur la représentation matricielle de f_{F}...

LEMME 4.4. — *Soient* f *et* g *deux endomorphismes de* E *tels que* $f \circ g = g \circ f$ *(on dit qu'ils* commutent*). Alors,* $\ker(f)$ *et* $\mathrm{im}(f)$ *sont stables par* g.

DÉFINITION 4.5 (Polynômes d'endomorphismes et de matrices)**. —** *Soit* $\mathrm{N} \in \mathbb{N}$ *et soit* $\mathrm{P} \in \mathbb{R}[\mathrm{X}]$ *un polynôme s'écrivant* $\mathrm{P} = a_{\mathrm{N}}\mathrm{X}^{\mathrm{N}} + \ldots + a_0$*. On définit l'endomorphisme* $\mathrm{P}(f)$ *par :*

$$\mathrm{P}(f) : x \mapsto a_{\mathrm{N}} f^{\mathrm{N}}(x) + \ldots + a_1 f(x) + a_0 x$$

De même, si $\mathrm{M} \in \mathscr{M}_n(\mathbb{R})$ *avec* $n \in \mathbb{N}^*$*, on définit la matrice* $\mathrm{P}(\mathrm{M}) \in \mathscr{M}_n(\mathbb{R})$ *par :*

$$\mathrm{P}(\mathrm{M}) = a_{\mathrm{N}}\mathrm{M}^{\mathrm{N}} + \ldots + a_1\mathrm{M} + a_0\mathrm{I}_n$$

On notera que si $\mathrm{P} = 1$, $\mathrm{P}(f)$ **n'est pas l'application constante égale à** 1 (qui ne serait pas linéaire) mais bien l'application identité sur E ! De même, si $\mathrm{P} = 1$ et si $\mathrm{M} \in \mathscr{M}_n(\mathbb{R})$ avec $n \in \mathbb{N}^*$ alors $\mathrm{P}(\mathrm{M}) = \mathrm{I}_n$.

EXEMPLE 4.6. Soit $\mathrm{P}(\mathrm{X}) = 3\mathrm{X}^2 + 2$ et soit f un endomorphisme. Alors,

$$\mathrm{P}(f) = 3f^2 + 2\mathrm{Id}_{\mathrm{E}}$$

Autrement dit, pour tout $x \in \mathrm{E}$ on a $\mathrm{P}(f)(x) = 3f^2(x) + 2x$.

PROPOSITION 4.7. — *Soient* P, Q *deux polynômes et* f *un endomorphisme. Alors,* $\mathrm{P}(f)$ *et* $\mathrm{Q}(f)$ *commutent et* $(\mathrm{PQ})(f) = \mathrm{P}(f) \circ \mathrm{Q}(f)$.

DÉFINITION 4.8. — *Soit f un endomorphisme de* E*. On dit qu'un polynôme* P *est annulateur de f si* $P(f) = 0$*. De même, si* A *est une matrice, on dit que* P *est annulateur de* A *si* $P(A) = 0$.

THÉORÈME 4.9. — *Tout endomorphisme de* E *admet un polynôme annulateur non nul. Toute matrice carrée admet un polynôme annulateur non nul.*

La démonstration du théorème 4.9 est édifiante : elle est symptomatique de l'utilisation de la dimension finie en algèbre.

Exercice 4.10. Démontrer la proposition précédente.

REMARQUE 4.11. La proposition 4.9 est fausse si E est de dimension infinie. Par exemple, on pourra vérifier que l'endomorphisme de dérivation sur $\mathbb{R}[X]$ n'admet pas de polynôme annulateur.

Un polynôme annulateur est un outil très efficace pour étudier les endomorphismes ou les matrices. Dans ce cours, nous n'irons pas trop loin dans cette direction, mais nous verrons tout de même comment un polynôme annulateur permet de trouver les valeurs propres d'un endomorphisme.

4.2 Éléments propres et diagonalisation.

Dans toute la suite, E est un espace vectoriel de dimension finie $n \in \mathbb{N}^*$ sur $\mathbb{R}$ et f est un endomorphisme de E.

DÉFINITION 4.12 (Éléments propres). — *Une* valeur propre *de f est un scalaire $\lambda \in \mathbb{R}$ pour lequel il existe un vecteur $x \in E$* ***non nul*** *tel que $f(x) = \lambda x$. Un vecteur non nul $x \in E$ tel que $f(x) = \lambda x$ est appelé* vecteur propre associé à la valeur propre λ. *Les* valeurs propres *et les* vecteurs propres associés*) d'une matrice carrée sont définies respectivement comme les valeurs propres et les vecteurs propres associés de l'endomorphisme canoniquement associé à cette matrice.*

Les résultats ci-dessous sont énoncés pour les endomorphismes ; on laisse le soin au lecteur de les adapter au cas des matrices carrées.

EXEMPLE 4.13. L'endomorphisme f est injectif si et seulement si 0 n'est pas valeur propre de f.

De façon plus générale, voici une caractérisation des valeurs propres qu'il est très important d'avoir en tête :

THÉORÈME 4.14. — *Le scalaire $\lambda \in \mathbb{R}$ est valeur propre de f si et seulement si l'endomorphisme $f - \lambda \mathrm{Id}$ est non injectif, ou si et seulement s'il n'est pas bijectif, ou si et seulement s'il n'est pas surjectif, ces trois conditions étant équivalentes.*

EXEMPLE 4.15. Les valeurs propres d'une matrice triangulaire (supérieure ou inférieure) sont ses termes diagonaux. Par exemple, les valeurs propres de la matrice

$$\begin{pmatrix} 1 & 0 & 5 \\ 0 & 5 & 2 \\ 0 & 0 & -2 \end{pmatrix}$$

sont $1, 5$ et -2. Cette technique est très efficace et fréquemment utilisée dans les problèmes : si une matrice se révèle être triangulaire, on connaît automatiquement ses valeurs propres !

THÉORÈME 4.16. — *Soit* A *une matrice et* P *un polynôme annulateur de* A*. Alors, toutes les valeurs propres de* A *sont des racines de* P*.*

REMARQUE 4.17. Attention : nous n'avons pas dit que toutes les racines d'un polynôme annulateur sont valeurs propres de la matrice ! Par exemple, le polynôme $(\mathrm{X}-1)(\mathrm{X}-2)\mathrm{X}^3$ est annulateur de la matrice identité... mais 0 et 2 ne sont certainement pas des valeurs propres de Id !

Le théorème 4.16 est très utile et nous vous invitons à l'utiliser sans retenue (dans le bon sens toutefois, voir la remarque ci-dessus !) ; toutefois, l'énoncé exige parfois sa démonstration, qu'il faut donc savoir reproduire sans hésitation.

Exercice 4.18. Démontrer le théorème 4.16.

DÉFINITION 4.19 (Sous-espace propre)**. —** *Soit* $\lambda \in \mathbb{R}$*. Le sous-espace vectoriel*

$$\ker(f - \lambda \mathrm{Id})$$

est appelé sous-espace propre associé à la valeur propre λ*, et on le note* $\mathrm{E}_\lambda(f)$*. Il contient l'élément nul et tous les éventuels vecteurs propres associés à la valeur* λ*. Si* λ *n'est pas valeur propre de* f*,* $\mathrm{E}_\lambda(f) = \{0\}$*.*

PROPOSITION 4.20. — *Les sous-espaces propres de* f *sont stables par* f*.*

DÉFINITION 4.21 (Spectre)**. —** *Le* spectre de f *est l'ensemble de toutes ses valeurs propres. On le note* $\mathrm{Sp}(f)$*.*

THÉORÈME 4.22. — *Soit* λ, μ *deux valeurs propres distinctes de* f*. Alors, les sous-espaces propres* $\mathrm{E}_\lambda(f)$ *et* $\mathrm{E}_\mu(f)$ *sont en somme directe. De manière générale, si* $\lambda_1, \dots, \lambda_n$ *sont les valeurs propres de* f*, alors les sous-espaces* $\mathrm{E}_{\lambda_i}(f)$ *sont tous en somme directe.*

Autrement dit, deux vecteurs propres associés à deux valeurs propres distinctes forment une famille libre. En fait, par le même raisonnement, on montre que toute famille de vecteurs propres associés à des valeurs propres distinctes est une famille libre. On en déduit le fait suivant à l'aide du théorème 1.85 :

PROPOSITION 4.23. — *On rappelle que* $\dim(\mathrm{E}) = n$*. Alors* f *admet au plus* n *valeurs propres distinctes.*

On peut trouver les éléments propres d'une matrice de taille 2 en utilisant le pivot de Gauss et la proposition 4.14. Plus généralement, on peut trouver les éléments propres d'une matrice carrée en en calculant le polynôme caractéristique (voir le point de cours sur les déterminants).

EXEMPLE 4.24. On cherche les valeurs propres de la matrice suivante :

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

On cherche donc $\lambda \in \mathbb{R}$ tel que $A - \lambda I_2$ soit non inversible. Pour cela, on considère $\lambda \in \mathbb{R}$ et on effectue un pivot de Gauss sur la matrice $A - \lambda I_2$ pour déterminer son rang en fonction de λ. En effectuant successivement les opérations $L_1 \leftrightarrow L_2$ (pour ne pas avoir à diviser par $1 - \lambda$ dont on ne sait pas s'il est non nul!), $L_1 \leftarrow -L_1$ et $L_2 \leftarrow L_2 - (1-\lambda)L_1$, on trouve :

$$\mathrm{rg}\,(A - \lambda I_2) = \mathrm{rg}\begin{pmatrix} 1-\lambda & 2 \\ -1 & 1-\lambda \end{pmatrix} = \mathrm{rg}\begin{pmatrix} -1 & 1-\lambda \\ 1-\lambda & 2 \end{pmatrix} = \mathrm{rg}\begin{pmatrix} 1 & \lambda-1 \\ 1-\lambda & 2 \end{pmatrix} = \mathrm{rg}\begin{pmatrix} 1 & \lambda-1 \\ 0 & 3-2\lambda+\lambda^2 \end{pmatrix}$$

λ est donc valeur propre de A si et seulement si $3 - 2\lambda + \lambda^2 = 0$, mais aucun réel ne vérifie cette équation. A n'admet donc aucune valeur propre.

On aurait aussi pu utiliser le critère d'inversibilité des matrices carrées de taille 2 pour faire apparaître le polynôme $3 - 2\lambda + \lambda^2$ sans avoir à utiliser le pivot de Gauss.

Une méthode similaire à celle développée dans l'exercice précédent fonctionne dans le cas des matrices carrées de taille $p \in \mathbb{N}^*$, mais elle fait apparaître des polynômes de degré p dont on ne sait pas en général déterminer les racines, sauf s'il existe des racines évidentes qui permettent de se ramener à des polynômes de plus bas degré par factorisation.

De manière générale, cette méthode n'est à mettre en œuvre qu'après avoir recherché de manière intelligente d'éventuelles valeurs propres évidentes de la matrice proposée. Illustrons notre propos :

EXEMPLE 4.25 (important)**.** Par exemple, $n = 3$ et si l'énoncé invite à étudier un endomorphisme f dont la matrice dans une base $\mathscr{B} = \{e_1, e_2, e_3\}$ de E est :

$$A = \begin{pmatrix} 3 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 5 & -3 \end{pmatrix}$$

on dispose de trois valeurs propres évidentes distinctes, qui sont les trois seules valeurs propres de f d'après la proposition 4.23. En effet, il apparaît que $f(e_1) = 3e_1$ et $f(e_3) = -3e_3$, ce qui montre que 3 et -3 sont valeurs propres (et on connaît même des vecteurs propres associés!). De plus, la somme des coefficients de chaque ligne de A vaut 2; cela signifie (retenez-le!) que A admet 2 pour valeur propre et qu'un vecteur propre associé est le vecteur ${}^t(1,1,1)$, ce que nous vous invitons à vérifier.

Voici une autre technique pour trouver les valeurs propres d'un endomorphisme ou d'une matrice sans avoir à utiliser le pivot de Gauss. Elle résulte du fait qu'une matrice et sa transposée ont même rang :

PROPOSITION 4.26. — *Une matrice carrée et sa transposée ont les mêmes valeurs propres.*

EXEMPLE 4.27. D'après l'exemple 4.25, la matrice :

$$ {}^t\mathrm{A} = \begin{pmatrix} 3 & 0 & 0 \\ -1 & 2 & 5 \\ 0 & 0 & -3 \end{pmatrix} $$

admet pour valeurs propres distinctes -3, 2 et 3. Bien entendu, les vecteurs propres associés ne sont pas les mêmes que dans le cas de A!

De la même façon, on voit que la matrice :

$$ \mathrm{B} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & -1 \end{pmatrix} $$

admet pour valeurs propres évidentes -1 et 1 (il suffit de considérer sa transposée et de suivre le même raisonnement que dans l'exemple 4.25) et est de rang 2 (en effet, sa deuxième ligne est somme des deux autres) donc admet aussi 0 pour valeur propre.

Une fois déterminées ces valeurs propres, il ne reste plus qu'à chercher les vecteurs propres associés en résolvant grâce à la méthode du pivot les différents systèmes d'équations linéaires issus des équations $\mathrm{BX} = -\mathrm{X}$, $\mathrm{BX} = 0$ et $\mathrm{BX} = \mathrm{X}$.

Exercice 4.28. Montrer que deux matrices carrées semblables ont les mêmes valeurs propres. Que dire des vecteurs propres associés?

DÉFINITION 4.29 (Diagonalisabilité)**. —** *f dit* diagonalisable *si et seulement s'il existe une base de* E *dans laquelle la matrice de f est diagonale.*

Une matrice $\mathrm{A} \in \mathscr{M}_n(\mathbb{R})$ *est dite* diagonalisable *si et seulement si elle est semblable à une matrice diagonale, c'est-à-dire s'il existe une matrice inversible* $\mathrm{P} \in \mathrm{GL}_n(\mathbb{R})$ *et une matrice diagonale* $\mathrm{D} \in \mathscr{M}_n(\mathbb{R})$ *telles que* $\mathrm{PDP}^{-1} = \mathrm{A}$.

Un endomorphisme est donc diagonalisable si et seulement si sa matrice (dans n'importe quelle base) est diagonalisable.

EXEMPLE 4.30. Une matrice diagonale est diagonalisable.

D'après l'exercice 4.37, si A est une matrice diagonalisable et si D est une matrice diagonale semblable à A alors les coefficients diagonaux de D sont les valeurs propres de A, une valeur propre $\lambda \in \mathbb{R}$ étant répétée sur la diagonale de D un nombre de fois égal à $\dim(\mathrm{E}_\lambda(\mathrm{A}))$.

Nous allons maintenant énoncer des caractérisations très importantes de la diagonalisabilité, qu'il faut impérativement connaître et savoir appliquer.

THÉORÈME 4.31. — *Les propositions suivantes sont équivalentes :*

1. *f est diagonalisable*
2. E *admet une base constituée de vecteurs propres de f*
3. E *est somme directe des sous-espaces propres de f*
4. E *est somme des sous-espaces propres de f*
5. E *est somme de certains sous-espaces propres de f*
6. $\sum_{\lambda\in\mathrm{Sp}(f)} \dim\big(\mathrm{E}_\lambda(f)\big) = n$

EXEMPLE 4.32. La matrice $A \in \mathcal{M}_3(\mathbb{R})$ dont tous les coefficients valent 1, c'est-à-dire la matrice :

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

est diagonalisable. En effet, le rang de A est égal à 1 puisqu'elle est non nulle et que toutes ses colonnes sont colinéaires deux à deux, ce qui implique que la dimension de son noyau est égale à 2 par le théorème du rang, et donc que 0 est valeur propre de A et que l'on a $\dim(E_0(A)) = \dim(\ker(A)) = 2$. Comme la somme des coefficients de chaque ligne de A vaut 3, on a :

$$A \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}$$

donc 3 est valeur propre de A. On a alors :

$$n = \dim(E) \geq \dim(E_0(A)) + \dim(E_3(A)) \geq 2 + 1 = 3$$

d'où l'on déduit les faits suivants :

1. A est diagonalisable
2. A admet 0 et 3 pour seules valeurs propres
3. $\dim(E_3(A)) = 1$, et donc :

$$E_3(A) = \operatorname{vect}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right)$$

On va maintenant *diagonaliser* A, c'est-à-dire que l'on va exhiber la matrice de passage P de la base canonique de $\mathbb{R}^3$ vers une base de vecteurs propres de A et écrire l'équation de diagonalisation $A = PDP^{-1}$, où D est une matrice diagonale.

En remarquant qu'une base de $E_0(A) = \ker(A)$ est donnée par :

$$\left\{\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right\}$$

on obtient une base de $\mathbb{R}^3$ composée de vecteurs propres de A par recollement des bases de $E_0(A)$ et de $E_n(A)$ ainsi exhibées. La matrice de passage associée est :

$$P = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}$$

Son inverse se calcule aisément par l'une des méthodes exposées dans le chapitre consacré au calcul matriciel, et on trouve :

$$P^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{pmatrix}$$

On écrit finalement l'*équation de diagonalisation* de A en construisant la matrice diagonale D dont les coefficients diagonaux sont, dans l'ordre, les valeurs propres correspondant aux différents vecteurs propres qui constituent P et en écrivant l'équation $A = PDP^{-1}$:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ 1/3 & 1/3 & -2/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

THÉORÈME 4.33. — *Si f admet n valeurs propres distinctes, alors f est diagonalisable.*

EXEMPLE 4.34. La matrice A étudiée dans l'exemple 4.25 ainsi que sa transposée sont diagonalisables.

REMARQUE 4.35. La réciproque de cette proposition est évidemment fausse : il existe des matrices diagonalisables qui n'admettent qu'une seule valeur propre, comme par exemple la matrice identité. Un autre exemple est fourni par la matrice de taille 3 et dont tous les coefficients valent 1, qui a été étudiée dans l'exemple 4.32.

Toutes les matrices ne sont pas diagonalisables, comme le montre l'exercice suivant :

Exercice 4.36. Montrer que la matrice

$$A = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

n'est pas diagonalisable.

Exercice 4.37. Supposons que $f^2 = \mathrm{Id}$. Montrer que :

$$E = \ker(f - \mathrm{Id}) \oplus \ker(f + \mathrm{Id})$$

Cet endomorphisme est-il diagonalisable ?

Exercice 4.38. Que dire d'un endomorphisme nilpotent et diagonalisable ?

4.3 Applications

4.3.1 Calcul des puissances itérées d'une matrice.

Supposons que l'on dispose d'une matrice carrée $A \in \mathscr{M}_n(\mathbb{R})$ et que l'on veuille calculer sa puissance k-ème pour un certain $k \in \mathbb{N}$. La méthode naïve qui consiste à calculer A^2, puis A^3... et ce jusqu'à A^k permet évidemment d'obtenir le résultat, mais le temps nécessaire pour réaliser ces calculs à la main est prohibitif si k n'est pas très petit. Si la matrice A est diagonalisable, les choses deviennent beaucoup plus simples : en effet, supposons qu'il existe $P \in GL_n(\mathbb{R})$ et Δ une matrice diagonale telles que $A = P\Delta P^{-1}$. Alors on a $A = P\Delta^k P^{-1}$ d'après la proposition 3.49.

Or la matrice Δ est diagonale. Elle s'écrit donc sous la forme :

$$\Delta = \begin{pmatrix} a_1 & 0 & \cdots & \cdots & 0 \\ 0 & a_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & a_n \end{pmatrix}$$

Il est alors immédiat que

$$\Delta^k = \begin{pmatrix} a_1^k & 0 & \cdots & \cdots & 0 \\ 0 & a_2^k & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & a_n^k \end{pmatrix}$$

Tout ce que l'on doit faire pour calculer A^k se résume donc à deux choses :

- Diagonaliser A, c'est-à-dire trouver P et D, calculer P^{-1} et écrire explicitement l'équation de diagonalisation $A = PDP^{-1}$.
- Calculer Δ^k.

Cette méthode est parfois très efficace; évidemment, elle suppose que A soit diagonalisable, ce qui n'est pas le cas de toutes les matrices.

4.3.2 Suites récurrentes linéaires.

La diagonalisation permet d'étudier facilement les suites récurrentes linéaires. Plutôt que de faire un long discours théorique sur le sujet, nous allons donner un exemple simple utilisant la méthode de la sous-section précédente. Soient a, b et c trois nombres réels. On définit trois suites $(u_n)_{n\in\mathbb{N}}, (v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ par : $u_0 = a, v_0 = b, w_0 = c$, et par la formule de récurrence suivante :

$$\forall n \in \mathbb{N}, \quad \left\{ \begin{array}{ccccc} u_{n+1} & = & 3u_n & +v_n & -w_n \\ v_{n+1} & = & & v_n & -w_n \\ w_{n+1} & = & & & 2w_n \end{array} \right.$$

Définissons une matrice A de taille 3×3 et une suite de vecteurs $(X_n)_{n\in\mathbb{N}}$ par

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix} \qquad \text{et} \qquad \forall n \in \mathbb{N}, \quad X_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$$

Les termes de la suite $(X_n)_{n\in\mathbb{N}}$ sont donc liés par la relation de récurrence :

$$\forall n \in \mathbb{N}, \quad X_{n+1} = AX_n$$

En particulier, une récurrence facile (écrivez-la!) montre que $X_n = A^n X_0$ pour tout $n \in \mathbb{N}$. Supposons que l'on veuille trouver u_{31}. Pour cela, on peut utiliser la relation $X_{31} = A^{31}X_0$ et calculer la puissance trente-et-unième de A, mais comme expliqué dans la sous-section précédente, cela risque d'être très lourd ! Si l'on peut diagonaliser A, tout sera plus simple. Ici, c'est le cas : on constate rapidement que son spectre est $\{1,2,3\}$, et que des vecteurs propres associés à $1,2,3$ sont respectivement :

$$x_1 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \quad \text{et} \quad x_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

De sorte que la matrice de passage est

$$\mathrm{P} = \begin{pmatrix} 1 & 1 & 2 \\ 0 & -2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

On calcule son inverse (par la méthode du pivot de Gauss) :

$$\mathrm{P}^{-1} = \begin{pmatrix} 1 & 1/2 & -3/2 \\ 0 & -1/2 & -1/2 \\ 0 & 0 & 1 \end{pmatrix}$$

Et finalement, on obtient l'équation de diagonalisation de A :

$$\mathrm{A} = \begin{pmatrix} 1 & 1 & 2 \\ 0 & -2 & -1 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \times \begin{pmatrix} 1 & 1/2 & -3/2 \\ 0 & -1/2 & -1/2 \\ 0 & 0 & 1 \end{pmatrix}$$

ainsi que les itérées successives de A^k :

$$\forall k \in \mathbb{N}, \quad \mathrm{A}^k = \begin{pmatrix} 1 & 1 & 2 \\ 0 & -2 & -1 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 3^k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^k \end{pmatrix} \times \begin{pmatrix} 1 & 1/2 & -3/2 \\ 0 & -1/2 & -1/2 \\ 0 & 0 & 1 \end{pmatrix}$$

Ainsi, on obtient rapidement $u_{31} = 3^3 1 a + \frac{3^{31}-1}{2} b + \left(2^{32} - \frac{3^{32}+1}{2}\right) c$, ce qui aurait été très fastidieux sans diagonaliser A ! Cette méthode vaut le coup d'être comprise et maîtrisée car on la retrouve régulièrement à l'oral.

4.3.3 Probabilités

Les méthodes que nous venons de présenter trouvent une application directe dans de nombreux problèmes de probabilités, que ce soit dans l'étude des graphes probabilistes (ou chaînes de Markov) ou dans l'explicitation du terme général de suites liées entre elles par des relations de récurrence linéaire provenant de la formule des probabilités totales. Découvrez le cours de probabilités pour en savoir plus !

5 Correction des exercices.

Solution de l'exercice 1.2. Notons $F = \mathscr{F}(\mathbb{R}, \mathbb{R})$. L'addition et la multiplication externe de F sont définies par les relations connues depuis longtemps :

$$\forall f, g \in F,\ \forall \lambda \in \mathbb{R},\ \forall x \in \mathbb{R}, \quad (f+g)(x) = f(x) + g(x) \quad \text{et} \quad (\lambda f)(x) = \lambda f(x)$$

On vérifie immédiatement que les deux lois de composition ainsi définies vérifient différentes propriétés de la définition 1.1 en revenant aux opérations sur des réels. Par exemple, on vérifie l'associativité de la multiplication en remarquant que si $f, g, h \in F$, alors pour tout $x \in \mathbb{R}$ on a :

$$\begin{aligned}\big(f + (g+h)\big)(x) &= f(x) + (g+h)(x) = f(x) + g(x) + h(x)\\ &= (f+g)(x) = h(x) = \big((f+g)+h\big)(x)\end{aligned}$$

et donc $f + (g+h) = (f+h) + g$.

L'élément neutre 0_F de F est la fonction nulle, définie par $0_F(x) = 0$ pour tout $x \in \mathbb{R}$.

Solution de l'exercice 1.11. On a bien $0_{\mathbb{R}^2} = \begin{pmatrix}0\\0\end{pmatrix} \in F \subset \mathbb{R}^2$. De plus, si $x, y, \lambda \in \mathbb{R}$ alors :

$$\begin{pmatrix}x\\0\end{pmatrix} + \begin{pmatrix}y\\0\end{pmatrix} = \begin{pmatrix}x+y\\0\end{pmatrix} \in F$$

et

$$\lambda \begin{pmatrix}x\\0\end{pmatrix} = \begin{pmatrix}\lambda x\\0\end{pmatrix} \in F$$

F est donc bien un sous-espace vectoriel de $\mathbb{R}^2$.

Solution de l'exercice 1.12. $(0,0)$ n'est pas dans F donc F n'est pas un sous-espace vectoriel de $\mathbb{R}^2$.

Solution de l'exercice 1.13. On a bien :

$$0_{\mathbb{R}^3} = \begin{pmatrix}0\\0\\0\end{pmatrix} \in F$$

puisque $0 + 0 + 2 \times 0 = 0$. De plus, si $x, y, z, x', y', z' \in R$, si $\lambda \in \mathbb{R}$ et si on a $x + y + 2z = x' + y' + 2z' = 0$ alors :

$$\begin{pmatrix}x\\y\\z\end{pmatrix} + \begin{pmatrix}x'\\y'\\z'\end{pmatrix} = \begin{pmatrix}x+x'\\y+y'\\z+z'\end{pmatrix} \in F$$

puisque $x + x' + y + y' + 2(z+z') = (x+y+2z) + (x'+y'+2z') = 0 + 0 = 0$, et

$$\lambda \begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}\lambda x\\\lambda y\\\lambda z\end{pmatrix} \in F$$

puisque $\lambda x + \lambda y + 2\lambda z = \lambda(x+y+2z) = \lambda \times 0 = 0$.

Solution de l'exercice 1.15. On a bien $O_{\mathscr{C}(\mathbb{R},\mathbb{R})} \in F$, et si $f, g \in F$ et $\lambda \in \mathbb{R}$ on a $(f+g)(1) = f(1) + g(1) = 0 + 0 = 0$ donc $f + g \in F$ et $(\lambda f)(1) = \lambda f(1) = \lambda \times 0 = 0$ donc $\lambda f \in F$. F est donc un sous-espace vectoriel de $\mathscr{C}(\mathbb{R}, \mathbb{R})$.

Notez qu'il n'est pas nécessaire de mentionner le fait que $f + g$ et λf sont des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ puisque l'on sait déjà que $\mathscr{C}(\mathbb{R}, \mathbb{R})$ est un espace vectoriel !

Solution de l'exercice 1.16. C'est un exercice important ; la rédaction précise de la solution est fortement conseillée au lecteur.

1. La rédaction est la même que plus haut et utilise le théorème d'analyse qui affirme que l'ensemble des fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$ est stable par addition et multiplication externe (en fait, c'est aussi un espace vectoriel !), de même que $\mathscr{C}(\mathbb{R}, \mathbb{R})$.
2. (a) Si $y \in \mathscr{S}$, y est dérivable et $y' = ay + b$ est continue puisque $ay + b$ est continue [23] (car dérivable !), donc $y \in \mathscr{C}^1(\mathbb{R}, \mathbb{R})$.
   (b) La linéarité de la dérivation et un argument de stabilité permettent de conclure.

Solution de l'exercice 1.20. $\{0\}$ est un sous-espace vectoriel contenant $\{0\}$ donc contient le sous-espace vectoriel vect(0), mais ce dernier contient nécessairement l'élément neutre 0, donc $\text{vect}(0) = \{0\}$. De même, on a $E \subset \text{vect}(E) \subset E$ donc $\text{vect}(E) = E$.

Solution de l'exercice 1.21. $\text{vect}\left(\begin{pmatrix}1\\2\end{pmatrix}\right)$ étant stable par multiplication externe, il contient nécessairement les multiples de $\begin{pmatrix}1\\2\end{pmatrix}$, d'où :

$$\left\{\begin{pmatrix}x\\2x\end{pmatrix}, x \in \mathrm{R}\right\} \subset \text{vect}\left(\begin{pmatrix}1\\2\end{pmatrix}\right)$$

L'espace de gauche étant un sous-espace vectoriel de E qui contient $\begin{pmatrix}1\\2\end{pmatrix}$, on a réciproquement :

$$\text{vect}\left(\begin{pmatrix}1\\2\end{pmatrix}\right) \subset \left\{\begin{pmatrix}x\\2x\end{pmatrix}, x \in \mathrm{R}\right\}$$

d'où l'égalité :

$$\text{vect}\left(\begin{pmatrix}1\\2\end{pmatrix}\right) = \left\{\begin{pmatrix}x\\2x\end{pmatrix}, x \in \mathrm{R}\right\}$$

On remarque ensuite que tous les éléments de $\mathbb{R}^2$ s'écrivent comme combinaison linéaire des vecteurs $\begin{pmatrix}0\\1\end{pmatrix}$ et $\begin{pmatrix}1\\0\end{pmatrix}$, ce qui indique que $\mathbb{R}^2$ tout entier est contenu dans tout sous-espace vectoriel de $\mathbb{R}^2$ contenant ces deux vecteurs, et donc que le sous-espace vectoriel de $\mathbb{R}^2$ engendré par ces vecteurs est $\mathbb{R}^2$ lui-même.

Solution de l'exercice 1.23. Le théorème 1.22 permet directement d'écrire que ce sous-espace est $\{aX + b \mid a, b \in \mathbb{R}\} = \mathbb{R}_1[X]$.

23. Ce raisonnement permet en fait de montrer par récurrence que y est indéfiniment dérivable sur $\mathbb{R}$.

Solution de l'exercice 1.25. Des réels sont dits non tous nuls s'ils ne sont pas tous nuls, c'est-à-dire si au moins l'un d'entre eux est non nul. Ils sont dits tous non nuls... s'ils sont tous non nuls, c'est-à-dire qu'aucun d'entre eux n'est nul.

Solution de l'exercice 1.29. En notant respectivement e_1, e_2 et e_3 les trois vecteurs qui composent la famille dans leur ordre d'apparition dans l'énoncé, on vérifie que $2e_1 - e_2 = e_3$, c'est-à-dire que $2e_1 - e_2 - e_3 = 0$, donc la famille est bien liée.

Solution de l'exercice 1.31. Soient a, b, c trois réels tels que $af_1 + bf_2 + cf_3 = 0$. Pour montrer que la famille est libre, il faut montrer que $a = b = c = 0$. Pour cela, l'idée est de tester l'égalité sur valeurs particulières de x. Par exemple, lorsque $x = \pi/2$, on obtient $a\sin(\pi/2) + b\sin(\pi) + c\sin(3\pi/2) = 0$, ce qui donne $a - c = 0$ donc $a = c$. Puis, en prenant $x = \frac{2\pi}{3}$, on obtient $a/2 - b/2 = 0$, donc $a = b$ et donc $a = b = c$. Nous avons donc à ce stade l'égalité suivante, valable pour tout x :

$$a\big(\sin(x) + \sin(2x) + \sin(3x)\big) = 0$$

Il suffit alors de choisir un x tel que $\sin(x) + \sin(2x) + \sin(3x) \neq 0$, par exemple $\pi/4$: on aura forcément $a = 0$, donc $a = b = c = 0$ et la famille est libre.

Solution de l'exercice 1.33. Soient $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que $\sum_{k=1}^{n} \lambda_k P_k = 0$. On remarque que $\sum_{k=1}^{n-1} \lambda_k P_k$ est une somme de polynômes de degré strictement inférieur à $\deg(P_n)$ et donc est elle-même un polynôme de degré strictement inférieur à $\deg(P_n)$. Mais on a :

$$\lambda_n P_n + \sum_{k=1}^{n-1} \lambda_k P_k = 0$$

donc le terme de plus haut degré de $\lambda_n P_n + \sum_{k=1}^{n-1} \lambda_k P_k$ est nul. Or ce terme est égal au produit de λ_n par le terme de plus haut de degré de P_n, qui est non nul puisque $\deg(P_n) \geq 0$: on a donc $\lambda_n = 0$. On peut donc écrire $\sum_{k=1}^{n-1} \lambda_k P_k = 0$ et réitérer le raisonnement [24] pour trouver :

$$\lambda_1 = \ldots = \lambda_n = 0$$

ce qui montre bien que la famille $\{P_1, \ldots, P_n\}$ est libre.

Solution de l'exercice 1.37. Il suffit de remarquer que pour tout $(x_1, \ldots, x_n) \in \mathbb{R}^n$ on a :

$$\big(x_1 \cdots x_n\big) = x_1 e_1 + \ldots + x_n e_n = \sum_{k=1}^{n} x_k e_k$$

Solution de l'exercice 1.41. On a vu dans l'exercice 1.37 que la famille était génératrice. Il reste à montrer qu'elle est libre; soient donc $\lambda_1, \ldots, \lambda_n \in R$ tels que $\sum_{k=1} \lambda_k e_k = 0$. Cette relation se réécrit :

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

soit $\lambda_1 = \ldots = \lambda_n = 0$, ce qui permet de conclure que $\{e_1, \ldots, e_n\}$ est bien libre et est donc une base de $\mathbb{R}^n$.

24. Exercice : écrivez précisément la récurrence qui se cache derrière ce raisonnement itéré !

Solution de l'exercice 1.43. Il suffit d'utiliser le résultat de l'exercice 1.33 pour voir que $\{P_0,\dots,P_n\}$ est libre. Pour montrer qu'elle est génératrice, on remarque tout d'abord que pour tout $r \in [\![0, n]\!]$ on a $\mathbb{R}_r[X] = \mathrm{vect}(1,\dots,X^r)$.

On démontre ensuite par récurrence forte les propositions définies par :

$$\forall r \in [\![0, n]\!], \quad \mathscr{P}_r : \text{« } 1, X, \dots, X^r \in \mathrm{vect}(P_0,\dots,P_r) \text{ »}$$

ce qui permettra de conclure puisque $\mathscr{P}_n$ donnera alors :

$$\mathbb{R}_n[X] = \mathrm{vect}(1,\dots,X^n) \subset \mathrm{vect}(P_0,\dots,P_n) \subset \mathbb{R}_n[X]$$

d'où l'égalité attendue.

$\mathscr{P}_0$ est vraie puisque P_0 est de degré 0 donc est égal à une constante non nulle a, ce qui permet d'écrire $1 = \frac{P_0}{a} \in \mathrm{vect}(P_0)$.

Soit $r \in [\![0, n-1]\!]$; on suppose $\mathscr{P}_0,\dots,\mathscr{P}_r$ vraies et on note $a_{r+1} \neq 0$ le terme de plus haut degré de P_{r+1}. Alors $P_{r+1} - a_{r+1}X^{r+1}$ est de degré strictement inférieur à $r+1$ et appartient alors à $\mathrm{vect}(1, X,\dots,X^r)$ et donc à $\mathrm{vect}(P_0,\dots,P_r)$ d'après les hypothèses de récurrence tenues pour vraies. Il existe donc $Q \in \mathrm{vect}(P_0,\dots,P_r)$ tel que $P_{r+1} - a_{r+1}X^{r+1} = Q$. Alors on a, puisque $a_{r+1} \neq 0$:

$$X^{r+1} = \frac{1}{a_{r+1}}\big(P_{r+1} - Q\big) \in \mathrm{vect}(P_0,\dots,P_{r+1})$$

ce qui montre que $\mathscr{P}_{r+1}$ est vraie et permet de conclure par le principe de récurrence.

La famille $\{P_1,\dots,P_n\}$ est donc bien génératrice de $\mathbb{R}_n[X]$; puisqu'elle est libre, c'en est une base.

Solution de l'exercice 1.55. La réponse à toutes ces questions est « non ». On pourra pour s'en convaincre considérer les deux bases de $\mathbb{R}$ fournies par $\{1\}$ et $\{2\}$, la famille libre mais non génératrice $\{1, X^2\}$ de $\mathbb{R}_2[X]$ et la famille génératrice non libre $\{1, 2\}$ de $\mathbb{R}$.

Solution de l'exercice 1.68. Si par exemple $F \subset G$ (le cas symétrique étant similaire), on a bien entendu $F \cup G = G$ qui est un sous-espace vectoriel de E. Si $F \not\subset G$ et $G \not\subset F$, soient $x \in G \setminus F$ et $y \in F \setminus G$. Alors $x, y \in F \cup G$ mais $x + y$ n'appartient pas à F (sinon on aurait $x = (x + y) - y \in F$ comme différence d'éléments de F) ni à G (sinon on aurait $y = (x + y) - x \in G$ comme différence d'éléments de G), ce qui montre que $x + y \notin F \cup G$ et donc que $F \cup G$ n'est pas stable par addition ; ce n'est donc pas un sous-espace vectoriel de E.

Solution de l'exercice 1.77. Si $f : \mathbb{R} \to \mathbb{R}$ est à la fois paire et impaire, alors pour tout $x \in \mathbb{R}$ on a :

$$f(x) = f(-x) = -f(x)$$

donc $f(x) = 0$, ce qui montre que la fonction nulle est la seule fonction paire et impaire de $\mathscr{F}(\mathbb{R}, \mathbb{R})$ et clôt la démonstration.

Solution de l'exercice 1.78. Pour montrer que ces deux espaces sont supplémentaires dans $\mathbb{R}^3$, il faut montrer deux choses : qu'ils sont en somme directe, puis que leur somme est égale à $\mathbb{R}^3$.

Pour montrer qu'ils sont en somme directe, c'est très simple : il suffit de prouver que $F_1 \cap F_2 = \{0\}$. Soit donc $X = (x, y, z)$ un élément de $F_1 \cap F_2$. Comme $X \in F_2$, il existe

$\alpha \in \mathbb{R}$ tel que $x = -\alpha$, $y = 2\alpha$ et $z = \alpha$. Mais d'autre part, $X \in F_1$, donc il doit vérifier $x-2y+z = 0$. Compte tenu de ce qui vient d'être dit, cela équivaut à $-\alpha-2\times(-2\alpha)+\alpha = 0$, soit $4\alpha = 0$ et $\alpha = 0$. Alors $X = 0$ et les deux sous-espaces sont bien en somme directe.

Nous devons maintenant montrer que cette somme directe est égale à $\mathbb{R}^3$ tout entier. Une méthode efficace consiste à utiliser le critère sur les dimensions : si l'on réussit à montrer que $\dim(F_1)+\dim(F_2) = 3$, alors on aura bien $F_1 \oplus F_2 = \mathbb{R}^3$. Il est évident que $\dim(F_2) = 1$, parce que F_2 est l'espace vectoriel engendré par le vecteur non nul $(-1,2,1)$.

Il nous reste donc à montrer que $\dim(F_1) = 2$. Pour cela, observons que F_1 est le noyau de la forme linéaire ϕ définie par

$$\phi\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = x-2y+z$$

Il est bien connu que le noyau d'une forme linéaire non nulle sur un espace vectoriel de dimension n est un sous-espace vectoriel de dimension $n-1$ (c'est précisément la conclusion du théorème du rang) : on a alors $\dim(F_1) = \dim\ker(\phi) = 2$, ce qui termine la démonstration.

On retiendra de l'exercice précédent que pour déterminer la dimension d'un sous-espace vectoriel, il peut parfois être utile d'exprimer ce sous-espace vectoriel comme le noyau d'une application linéaire dont on connaît facilement le rang (par exemple une forme linéaire, qui est de rang 1 si elle n'est pas nulle).

Solution de l'exercice 1.80. Un sous-espace vectoriel d'un espace vectoriel E contient toujours l'élément O_E donc deux sous-espaces vectoriels de E, quels qu'ils soient, ne sauraient être complémentaires (puisqu'ils ne sont pas disjoints!). On démontre facilement (faites-le!) que l'unique supplémentaire de E dans E est $\{0\}$ et que l'unique supplémentaire de $\{0\}$ dans E est E. Cependant, si E est de dimension finie et si F est un sous-espace vectoriel non trivial de E (c'est-à-dire si $F \neq \{0\}$ et $F \neq E$), alors F admet un supplémentaire G que l'on peut écrire $G = \mathrm{vect}(e_1,\dots,e_r)$ avec $r \in \mathbb{N}$ et $e_1,\dots,e_r \in G$ (puisque E et donc G est de dimension finie). Si l'on choisit $x \in F$ tel que $x \neq 0$ (ce qui est possible puisque $F \neq \{0\}$) alors le sous-espace vectoriel $G' = \mathrm{vect}(e_1+x, e_2,\dots,e_r)$ de E est un supplémentaire de F (exercice : démontrez-le!) distinct de G : en effet, il contient e_1+x que G ne contient pas puisque $x \notin G$.

Solution de l'exercice 2.6. Il suffit de vérifier que la famille :

$$\left\{\begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & & & \vdots \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & & \vdots \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & \dots & \dots & \dots & 0 \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ \vdots & & & 0 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}\right\}$$

composée des $m\times n$ matrices $E_{i,j}$ (où pour tout i $[\![1, m]\!]$ et tout $j \in [\![1, n]\!]$ on définit $E_{i,j}$ comme la matrice dont tous les coefficients sont nuls à l'exception du coefficient d'indices (i, j) qui vaut 1) est libre et génératrice dans $\mathscr{M}_{n,m}(\mathbb{R})$. Cette base est appelée *base canonique de* $\mathscr{M}_{n,m}(\mathbb{R})$, qui est donc de dimension $m \times n$.

Solution de l'exercice 2.19. Le rang de la matrice identité est n : en effet, les vecteurs colonnes de I_n forment une famille échelonnée de n vecteurs de $\mathbb{R}^n$.

Plus généralement, le rang d'une matrice diagonale est égal à son nombre de coefficients non nuls. En effet, si $A \in \mathscr{M}_n(\mathbb{R})$ est une matrice diagonale on démontre facilement que la famille formée par les colonnes non nulles de A est libre (par exemple parce qu'il est possible de la compléter[25] en une famille échelonnée en remplaçant les colonnes nulles par des colonnes remplies de 0 à l'exception d'un 1 au niveau de la diagonale), et il est clair qu'il s'agit de la famille libre maximale que l'on peut construire à partir des vecteurs colonnes de A puisque les autres colonnes de A sont nulles.

Solution de l'exercice 2.20. La première colonne de la matrice forme une famille libre (d'un seul élément). La deuxième colonne est nulle donc il n'existe pas de famille libre contenant cette colonne. Par conséquent, le cardinal maximal d'une famille libre constituée par des vecteurs colonnes de A est 1, d'où $\operatorname{rg}(A) = 1$.

Solution de l'exercice 2.24. Après permutation, les trois colonnes de la matrice forment une famille échelonnée de $\mathbb{R}^3$ donc sont libres. Le rang de la matrice est donc égal à 3.

Solution de l'exercice 2.27. On note $a_{i,j}$ le coefficient générique de A et $b_{i,j}$ celui de B, et on définit $c_{i,j}$ comme le coefficient générique de AB. On remarque tout d'abord que la matrice ${}^t(AB)$ est de format $n \times l$ et que c'est aussi le cas de la matrice ${}^tB\,{}^tA$. On écrit ensuite :

$$\forall i \in [\![1, n]\!], \forall j \in [\![1, l]\!], \quad \left({}^tB\,{}^tA\right)_{i,j} = \sum_{k=1}^{m} \left({}^tB\right)_{i,k} \left({}^tA\right)_{k,j} = \sum_{k=1}^{m} b_{k,i} a_{j,k} = \sum_{k=1}^{m} a_{j,k} b_{k,i} = c_{j,i}$$

ce qui établit l'égalité attendue.

Solution de l'exercice 2.32. C'est possible ! On a par exemple :

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

La différence fondamentale de la multiplication matricielle avec la multiplication des réels, outre sa non-commutativité, est le fait que $AB = 0$ n'entraîne pas forcément $A = 0$ ou $B = 0$.

Solution de l'exercice 2.34. Soient $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. On démontre par une récurrence facile (exercice : écrivez-la !) que pour tout $k \in \mathbb{N}$ on a :

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \lambda_{n-1} & 0 \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2^k & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \lambda_{n-1}^k & 0 \\ 0 & \cdots & \cdots & 0 & \lambda_n^k \end{pmatrix}$$

25. Ce raisonnement s'appuie sur le fait qu'une sous-famille d'une famille libre est libre. Sauriez-vous le démontrer ?

Solution de l'exercice 2.36. La définition du produit matriciel permet de montrer que le produit de deux matrices dont les coefficients sont tous strictement positifs est une matrice qui possède elle-même des coefficients tous strictement positifs. Une récurrence facile (exercice : écrivez-la!) permet alors de montrer que les puissances successives de A ont toutes des coefficients tous strictement positifs, et donc qu'aucune puissance de A n'est nulle. A ne peut donc pas être nilpotente!

Solution de l'exercice 2.39. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ sont toutes deux nilpotentes (en effet, leur carré est nul), mais leur somme $\mathrm{S} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ne l'est pas puisque son carré est I_2 et que ses puissances successives sont données par :

$$\forall k \in \mathbb{N}, \quad \mathrm{S}^k = \begin{cases} \mathrm{I}_2 \text{ si } k \text{ est pair} \\ \mathrm{S} \text{ si } k \text{ est impair} \end{cases}$$

L'ensemble des matrices nilpotentes de taille 2 n'est donc pas stable par somme; on généralise le contre-exemple précédent à n'importe quelle taille $n \geq 2$ en complétant les matrices de taille 2 ci-dessus par des lignes et des colonnes nulles.

Par contre, l'ensemble des matrices nilpotentes de taille n est bien stable par multiplication externe pour tout $n \in \mathbb{N}$; on dit qu'il s'agit d'un *cône* de $\mathcal{M}_n(\mathbb{R})$.

Enfin, l'ensemble des matrices nilpotentes de taille 1 est l'ensemble $\{(0)\}$ qui, lui, est bien un sous-espace vectoriel.

Solution de l'exercice 2.44. Ce n'est pas automatique. Par exemple, la matrice $\mathrm{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ est de rang 1 donc ne peut être inversible puisque s'il existait $\mathrm{B} \in \mathcal{M}_2(\mathbb{R})$ tel que $\mathrm{AB} = \mathrm{I}_2$ on aurait $\mathrm{rg}(\mathrm{I}_2) = 2 \leq \min(\mathrm{rg}(\mathrm{A}), \mathrm{rg}(\mathrm{B})) \leq 1$ d'après la proposition 2.40, ce qui est absurde. On ne peut pas non plus dire qu'une matrice dont tous les coefficients sont non nuls est non inversible comme en témoigne l'exemple de la matrice $(1) \in \mathcal{M}_1(\mathbb{R})$ qui est sa propre inverse.

Solution de l'exercice 2.45. Supposons A inversible. Si n désigne la taille de A, il suffit d'écrire

$${}^t\mathrm{A}\,{}^t\left(\mathrm{A}^{-1}\right) = {}^t\left(\mathrm{A}^{-1}\mathrm{A}\right) = {}^t\mathrm{I}_n = \mathrm{I}_n$$

et

$${}^t\left(\mathrm{A}^{-1}\right){}^t\mathrm{A} = {}^t\left(\mathrm{A}\mathrm{A}^{-1}\right) = {}^t\mathrm{I}_n = \mathrm{I}_n$$

pour voir que ${}^t\mathrm{A}$ est elle aussi inversible et que son inverse est ${}^t\left(\mathrm{A}^{-1}\right)$.

Solution de l'exercice 2.46. Si A est nilpotente alors il existe $k \in \mathbb{N}$ tel que $\mathrm{A}^k = 0$. Soit k' le plus petit entier positif tel que $\mathrm{A}^{k'} = 0$; nécessairement $k' \geq 1$ (puisque A^0 est la matrice identité). Supposons A inversible. On a alors $\mathrm{A}^{-1}\mathrm{A}^{k'} = \mathrm{A}^{-1}\mathrm{A}\mathrm{A}^{k'-1} = \mathrm{A}^{k'-1} \neq 0$ par définition de k' mais aussi $\mathrm{A}^{-1}\mathrm{A}^{k'} = \mathrm{A}^{-1} \times 0 = 0$, ce qui est absurde. A n'est donc pas inversible.

Solution de l'exercice 2.48. Il est d'ores et déjà bon de remarquer qu'une matrice carrée A de taille 2 est inversible si et seulement si ses vecteurs colonnes ne sont pas colinéaires, c'est-à-dire, en notant C_1 et C_2 lesdits vecteurs, qu'il n'existe pas de $\lambda \in \mathbb{R}$ tel que $C_1 = \lambda C_2$ ou $C_2 = \lambda C_1$. En effet, cela signifie que la famille (C_1, C_2) n'est pas liée (voir la remarque 1.27), c'est-à-dire qu'elle est libre et donc que le rang de A est égal à 2. On verra un peu plus loin dans le cours un critère calculatoire très pratique pour déterminer si une matrice carrée de taille 2 est inversible et qui repose directement sur cette observation.

En règle générale, le fait que les colonnes d'une matrice soient colinéaires deux à deux est équivalent au fait que son rang est égal à 0 (cas où toutes les colonnes sont nulles) ou à 1 (cas où la matrice est non nulle mais où toutes ses colonnes sont proportionnelles à l'une de ses colonnes non nulles). Lire directement sur une matrice carrée de taille 3 si elle est inversible est déjà beaucoup plus ardu puisqu'il s'agit de déterminer si ses trois vecteurs colonnes sont *coplanaires*!

Les observations que nous venons de faire se transposent (!) directement au cas des lignes en considérant la transposée de la matrice.

Intéressons-nous maintenant aux matrices proposées par l'énoncé ; à l'exception de la dernière, on discutera de leur inversibilité sans utiliser le pivot de Gauss, et on inversera celles pour lesquelles cela est possible en utilisant la méthode du pivot. Nous vous conseillons donc de limiter votre lecture aux preuves d'inversibilité et de revenir à cet exercice une fois lue la section sur le pivot de Gauss !

La première matrice est inversible puisque ses deux colonnes ne sont pas colinéaires ; détaillons pour la première et dernière fois cet argument. si l'on avait $C_1 = \lambda C_2$ avec $\lambda \in \mathbb{R}$, on devrait avoir $\lambda = \frac{1}{5}$ (on le voit en considérant les coefficients de la première ligne), ce qui n'est pas compatible avec les coefficients de la deuxième ligne. Inversement, si l'on avait $C_2 = \lambda C_1$ avec $\lambda \in \mathbb{R}$, on devrait avoir $\lambda = 5$, ce qui n'est à nouveau pas compatible avec les coefficients de la deuxième ligne.

Notons A cette première matrice et cherchons son inverse en utilisant la méthode du pivot. Considérons les deux vecteurs de $\mathbb{R}^2$ donnés par :

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

et supposons que $Y = AX$. Cela signifie que :

$$\begin{cases} x_1 & +5x_2 & = & y_1 \\ 2x_1 & +5x_2 & = & y_2 \end{cases}$$

ce qui équivaut par l'opération élémentaire $L_2 \leftarrow L_2 - L_1$ au système :

$$\begin{cases} x_1 & +5x_2 & = & y_1 \\ x_1 & & = & y_2 - y_1 \end{cases}$$

que l'on résout de proche en proche et dont on trouve qu'il équivaut à :

$$\begin{cases} x_1 & = & -y_1 & +y_2 \\ x_2 & = & \frac{2}{5}y_1 & -\frac{1}{5}y_2 \end{cases}$$

soit $X = A'Y$ où

$$A' = \begin{pmatrix} -1 & 1 \\ \frac{2}{5} & -\frac{1}{5} \end{pmatrix}$$

ce qui montre que $A^{-1} = A'$.

La deuxième matrice n'est pas inversible puisque $C_1 = 0 \times C_2$. Comme elle n'est pas nulle, elle est de rang 1.

La troisième matrice est aussi de rang 1 — et donc non inversible — puisqu'elle n'est pas nulle et vérifie $C_2 = 2C_1$.

La quatrième matrice est inversible car ses deux colonnes ne sont pas colinéaires. On calcule son inverse par la méthode du pivot comme précédemment et on trouve que :

$$\begin{pmatrix} 4 & 4 \\ 2 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/12 & 1/3 \\ 1/6 & -1/3 \end{pmatrix}$$

La cinquième matrice est de rang 1 — et donc non inversible — puisqu'elle n'est pas nulle et vérifie $C_2 = -C_1$.

La sixième matrice, quant à elle, est de rang au moins 2 puisque ses colonnes ne sont pas colinéaires deux à deux. Nous la noterons C. On montre qu'elle est inversible en résolvant directement le système qui permet de trouver son inverse. Considérons les deux vecteurs de $\mathbb{R}^3$ donnés par :

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

et supposons que $Y = CX$. Cela signifie que :

$$\left\{ \begin{array}{rrrcl} 3x_1 & -x_2 & +2x_3 & = & y_1 \\ & x_2 & & = & y_2 \\ x_1 & & +5x_3 & = & y_3 \end{array} \right.$$

ce qui équivaut, en substituant x_2 à y_2 en utilisant L_2, à :

$$\left\{ \begin{array}{rrrcl} 3x_1 & & +2x_3 & = & y_1 + y_2 \\ & x_2 & & = & y_2 \\ x_1 & & +5x_3 & = & y_3 \end{array} \right.$$

Ce système est équivalent via l'opération élémentaire $L_1 \leftarrow L_1 - 3L_3$ au système suivant :

$$\left\{ \begin{array}{rrrcl} & & -13x_3 & = & y_1 + y_2 - 3y_3 \\ & x_2 & & = & y_2 \\ x_1 & & +5x_3 & = & y_3 \end{array} \right.$$

c'est-à-dire, en résolvant de proche en proche :

$$\left\{ \begin{array}{rcrrr} x_1 & = & \frac{5}{13}y_1 & +\frac{5}{13}y_2 & -\frac{2}{13}y_3 \\ x_2 & = & & y_2 & \\ x_3 & = & -\frac{1}{13}y_1 & +\frac{1}{13}y_2 & +\frac{3}{13}y_3 \end{array} \right.$$

soit enfin $Y = C'X$ avec

$$C' = \frac{1}{13} \begin{pmatrix} 5 & 5 & -2 \\ 0 & 13 & 0 \\ -1 & 1 & 3 \end{pmatrix}$$

ce qui montre que C est inversible et que $C^{-1} = C'$.

Solution de l'exercice 2.62. La méthode proposée ici est différente de celle mise en œuvre dans l'exercice 2.48 puisqu'il s'agit d'une transformation de Gauss-Jordan similaire à celle exposée dans l'exemple 2.61.

Lorsque l'on met cette méthode en pratique, on cherche d'abord à déterminer si une matrice est de rang maximal avant de tenter de l'inverser. Toutefois, il peut être prudent de raisonner d'emblée avec une double matrice comme dans l'exemple 2.61 pour ne pas perdre la mémoire des opérations effectuées dans le pivot pour transformer la matrice en une matrice triangulaire supérieure.

Quelle que soit la manière de raisonner que vous adoptez (que ce soit par équivalence de matrices, par la résolution d'un système linéaire comme dans la solution de l'exercice 2.48 ou par la méthode présentée ici), il est *obligatoire* d'écrire en détail les opérations effectuées (voir la remarque 2.55).

Intéressons-nous à la première matrice donnée par l'énoncé. On la transforme comme dans l'exemple 2.61 en la matrice :

$$A_0 = \left(\begin{array}{ccc|ccc} 1 & 5 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

à laquelle on applique la transformation $L_3 \leftarrow L_3 - 2L_1$ pour obtenir la matrice :

$$A_1 = \left(\begin{array}{ccc|ccc} 1 & 5 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 0 & -9 & -2 & -2 & 0 & 1 \end{array}\right)$$

On divise alors L_2 par 2 pour faire apparaître un pivot unitaire, ce qui donne la matrice :

$$A_2 = \left(\begin{array}{ccc|ccc} 1 & 5 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3/2 & 0 & 1/2 & 0 \\ 0 & -9 & -2 & -2 & 0 & 1 \end{array}\right)$$

On applique alors à A_2 l'opération $L_3 \leftarrow L_3 + 9L_2$, ce qui donne :

$$A_3 = \left(\begin{array}{ccc|ccc} 1 & 5 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3/2 & 0 & 1/2 & 0 \\ 0 & 0 & 23/2 & -2 & 9/2 & 1 \end{array}\right)$$

On applique enfin à A_3 l'opération $L_3 \leftarrow \frac{2}{23}L_3$ et on obtient :

$$A_4 = \left(\begin{array}{ccc|ccc} 1 & 5 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & -4/23 & 9/23 & 2/23 \end{array}\right)$$

A est donc bien inversible puisqu'il est possible de la transformer par des opérations élémentaires en une matrice triangulaire inversible. On poursuit donc la transformation de A en I_3 en appliquant à A_4 l'opération $L_1 \leftarrow L_1 - 5L_2$, qui donne :

$$A_5 = \left(\begin{array}{ccc|ccc} 1 & 0 & -13/2 & 1 & -5/2 & 0 \\ 0 & 1 & 3/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & -4/23 & 9/23 & 2/23 \end{array}\right)$$

On applique ensuite à A_5 l'opération $L_1 \leftarrow L_1 + \frac{13}{2}L_3$ et on obtient :

$$A_6 = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -3/23 & 1/23 & 13/23 \\ 0 & 1 & 3/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & -4/23 & 9/23 & 2/23 \end{array}\right)$$

On applique enfin à A_6 l'opération $L_2 \leftarrow L_2 - \frac{3}{2}L_3$, ce qui donne la matrice :

$$A_7 = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -3/23 & 1/23 & 13/23 \\ 0 & 1 & 0 & 6/23 & -2/23 & -3/23 \\ 0 & 0 & 1 & -4/23 & 9/23 & 2/23 \end{array}\right)$$

On en déduit la valeur de A^{-1} :

$$A^{-1} = \frac{1}{23}\begin{pmatrix} -3 & 1 & 13 \\ 6 & -2 & -3 \\ -4 & 9 & 2 \end{pmatrix}$$

On obtient par un raisonnement exactement similaire (en utilisant cette fois des permutations de lignes pour mettre les pivots sur la diagonale des matrices successives) que les matrices B et C sont inversibles et que :

$$B^{-1} = \begin{pmatrix} -1 & 1 & 0 \\ 1/2 & -1 & 1/2 \\ 1 & 0 & 0 \end{pmatrix}$$

et

$$C^{-1} = \frac{1}{10}\begin{pmatrix} -3 & 4 & -1 \\ -1 & -2 & 3 \\ 11 & 2 & -3 \end{pmatrix}$$

Solution de l'exercice 2.72. En appliquant à

$$\left\{\begin{array}{rrrcr} x & +2y & -z & = & 1 \\ 3x & & -7z & = & 2 \\ & y & -2z & = & 5 \end{array}\right.$$

l'opération élémentaire $L_2 \leftarrow L_2 - 3L_1$, on obtient le système :

$$\left\{\begin{array}{rrrcr} x & +2y & -z & = & 1 \\ & -6y & -4z & = & -1 \\ & y & -2z & = & 5 \end{array}\right.$$

En appliquant à ce dernier système l'opération élémentaire $L_2 \leftrightarrow L_3$ (pour faire apparaître un pivot sur la deuxième ligne sans avoir à la diviser par -6), on obtient :

$$\left\{\begin{array}{rrrcr} x & +2y & -z & = & 1 \\ & y & -2z & = & 5 \\ & -6y & -4z & = & -1 \end{array}\right.$$

ce qui donne, une fois réalisée l'opération élémentaire $L_3 \leftarrow L_3 + 6L_2$:

$$\left\{\begin{array}{rrrcr} x & +2y & -z & = & 1 \\ & y & -2z & = & 5 \\ & & -16z & = & 29 \end{array}\right.$$

d'où, en remontant de proche en proche, la solution :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{-57}{16} \\ \frac{11}{8} \\ \frac{-29}{16} \end{pmatrix}$$

Solution de l'exercice 3.3. L'application $x \mapsto 3x$ est évidemment linéaire, de même pour l'application nulle (il suffit de revenir à la définition !). En revanche, l'application $g : x \mapsto 3x + 1$ n'est pas linéaire, par exemple parce que $g(0) = 1 \neq 0$.

Solution de l'exercice 3.4. Si $\lambda \in \mathbb{R}$ et si $P = \sum_{k=0}^{n} a_k X^k$ et $Q = \sum_{k=0}^{n} b_k X^k$ avec $n \in \mathbb{N}$ et $a_0, \dots, a_n, b_0, \dots, b_n \in \mathbb{R}$ alors :

$$\begin{aligned} \Delta(\lambda P + Q) = \Delta\left(\sum_{k=0}^{n} (\lambda a_k + b_k) X^k\right) &= \sum_{k=1}^{n} k(\lambda a_k + b_k) X^{k-1} \\ &= \lambda \sum_{k=1}^{n} k a_k X^{k-1} + \sum_{k=1}^{n} k b_k X^{k-1} = \lambda\Delta(P) + \Delta(Q) \end{aligned}$$

d'où la linéarité de Δ.

Solution de l'exercice 3.7. Il suffit d'écrire que pour si $x, y, z \in \mathbb{R}$ alors :

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = f(xe_1 + ye_2 + ze_3) = xf(e_1) + yf(e_2) + zf(e_3)$$

d'où l'on déduit

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x\begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix} + y\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2x + y + 5z \\ y \\ 4x \end{pmatrix}$$

Solution de l'exercice 3.12. La dérivée d'un polynôme est nulle si et seulement s'il est constant : en effet, la dérivée d'un polynôme nul est constant et la dérivée d'un polynôme non constant de degré $n \in \mathbb{N}^*$ est un polynôme de degré $n - 1 > -\infty$ donc non nul. On a donc $\ker(\Delta) = \mathbb{R}_0[X]$. De plus, tout polynôme est le polynôme dérivée d'un autre polynôme (en d'autres termes, l'application Δ est surjective) : en effet, si $P = \sum_{k=0}^{n} a_k X^k$ avec $n > 0$ alors $P = Q'$ avec $Q = \sum_{k=0}^{n} \frac{a_k}{k+1} X^{k+1}$. On a donc $\mathrm{Im}(D elta) = \mathbb{R}[X]$.

Solution de l'exercice 3.22. Soit $y \in \mathrm{Im}(f)$. Alors il existe $x \in E$ tel que $y = f(x)$; mais cela permet d'écrire que $f(y) = f\big(f(x)\big) = f^2(x) = 0$ et donc que $y \in \ker(f)$, ce qui montre que $\mathrm{Im}(f) \subset \ker(f)$. La réciproque est vraie : si $\mathrm{Im}(f) \subset \ker(f)$, on a pour tout $x \in E$ la relation $f(x) \in \mathrm{Im}(f)$ d'où l'on tire $f(x) \in \ker(f)$ et donc $f\big(f(x)\big) = 0$, c'est-à-dire $f^2(x) = 0$. Ceci étant vrai pour tout $x \in E$, on a bien $f^2 = 0$.

Solution de l'exercice 3.24. L'image d'une forme linéaire est un sous-espace vectoriel de $\mathbb{R}$; il est donc de dimension 0 (dans le cas où la forme est nulle) ou 1 (dans le cas où la forme est non nulle). En d'autres termes, une forme linéaire est nulle ou surjective.

Solution de l'exercice 3.27. Le noyau d'une forme linéaire non nulle de $\mathbb{R}$ dans $\mathbb{R}$ est de dimension $\{0\}$: c'est alors nécessairement $\{0\}$, qui est donc le seul hyperplan de $\mathbb{R}$.

Les hyperplans de $\mathbb{R}^2$ sont les noyaux des formes linéaires non nulles de $\mathbb{R}^2$. Or on montre que ces formes linéaires sont exactement les applications $f : \mathbb{R}^2 \to \mathbb{R}$ définies par :

$$f : \begin{pmatrix} x \\ y \end{pmatrix} = ax + by$$

avec $a, b \in \mathbb{R}$ et $(a, b) \neq (0, 0)$. En effet, il suffit de prendre $a = f(e_1)$ et $b = f(e_2)$, où $\{e_1, e_2\}$ est la base canonique de $\mathbb{R}^2$, pour vérifier que toute forme linéaire non nulle est de cette forme, et réciproquement il est clair qu'une telle application est une forme linéaire non nulle.

Les hyperplans de $\mathbb{R}^2$ sont donc exactement les ensembles de la forme :

$$\mathrm{H}_{a,b} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid ax + by = 0 \right\}$$

avec $a, b \in \mathbb{R}$ et $(a, b) \neq (0, 0)$.

Il n'est pas difficile de voir que ces ensembles sont en réalité les droites du plan $\mathbb{R}^2$ passant par 0 !

De manière plus générale, on montre (exactement de la même façon) que si $n \in \mathbb{N}^*$, les hyperplans de $\mathbb{R}^n$ sont exactement les ensembles de la forme :

$$\mathrm{H}_{a_1,\ldots,a_n} = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \mid a_1 x_1 + \ldots + a_n x_n = 0 \right\}$$

avec $a_1, \ldots, a_n \in \mathbb{R}$ non tous nuls.

Solution de l'exercice 3.34. Avec les notations du théorème, on a par le théorème du rang $\dim(\mathrm{E}) = \dim(\ker(f)) + \mathrm{rg}(f)$. On sait que f est injectif si et seulement si $\ker(f) = \{0\}$, c'est-à-dire $\dim(\ker(f)) = 0$, ce qui implique alors que $\mathrm{rg}(f) = \dim(\mathrm{E})$ et donc, puisque $\mathrm{Im}(f) \subset \mathrm{E}$, que $\mathrm{Im}(f) = \mathrm{E}$ et donc que f est surjectif, donc bijectif. De même, f est surjectif si et seulement si $\mathrm{rg}(f) = \dim(\mathrm{E})$, ce qui implique que $\dim(\ker(f)) = 0$ et donc que f est injectif, donc bijectif.

Solution de l'exercice 3.38. Il suffit d'écrire que si $\mathscr{B}_4 = \{e_1, \ldots, e_4\}$ est la base canonique de $\mathbb{R}^4$ alors on a :

$$f(e_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 2 \end{pmatrix}, f(e_2) = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, f(e_3) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 8 \\ 0 \\ 5 \\ 0 \end{pmatrix}, f(e_4) = \begin{pmatrix} 0 \\ -1 \\ 5 \\ 3 \\ 0 \\ -3 \\ 0 \end{pmatrix}$$

Les décompositions de ces vecteurs sur la base canonique $\mathscr{B}_7$ de $\mathbb{R}^7$ sont évidentes : ils sont égaux à leurs vecteurs coordonnées sur cette base. On peut donc écrire :

$$\operatorname{mat}_{\mathscr{B}_4}^{\mathscr{B}_7}(f) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 8 & 3 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 5 & -3 \\ 2 & -1 & 0 & 0 \end{pmatrix}$$

Solution de l'exercice 3.41. On détermine comme précédemment que :

$$\operatorname{mat}_{\mathscr{B}}^{\mathscr{B}}(f) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

En calculant $f(e'_1) = 4f(e_1) = 4e_1$ et $f(e'_2) = f(e_1) + f(e_2) = e_1 + 2e_2$ on obtient :

$$\operatorname{mat}_{\mathscr{B}'}^{\mathscr{B}}(f) = \begin{pmatrix} 4 & 1 \\ 0 & 1 \end{pmatrix}$$

On décompose ensuite dans la base $\mathscr{B}'$ les vecteurs $f(e_1) = e_1 = \frac{1}{4}e'_1$ et $f(e_2) = 2e_2 = 2e'_2 - \frac{1}{2}e'_1$ et on écrit :

$$\operatorname{mat}_{\mathscr{B}}^{\mathscr{B}'}(f) = \begin{pmatrix} 1/4 & -1/2 \\ 0 & 2 \end{pmatrix}$$

Enfin, on décompose dans la base $\mathscr{B}'$ les vecteurs $f(e'_1) = e'_1$ et $f(e'_2) = f(e_1) + f(e_2) = 2e'_2 - \frac{1}{4}e'1$, ce qui permet d'écrire :

$$\operatorname{mat}_{\mathscr{B}'}^{\mathscr{B}'}(f) = \begin{pmatrix} 1 & -1/4 \\ 0 & 2 \end{pmatrix}$$

Remarquez que la méthode canonique pour effectuer les décompositions sur la base $\mathscr{B}'$ consiste à poser un système de deux équations à deux inconnues qui traduit, par exemple, l'équation $f(e_1) = ae'_1 + be'_2$ d'inconnues a et b, mais qu'il est possible, comme souvent, de les réaliser de tête, ce qui fait évidemment gagner un temps précieux.

Solution de l'exercice 3.48. Il suffit de revenir aux définitions.

1. Il existe $P \in GL_n(\mathbb{R})$ telle que $A = PBP^{-1}$. On a donc $B = P^{-1}AP = P^{-1}A\left(P^{-1}\right)^{-1}$; comme $P^{-1} \in GL_n(\mathbb{R})$, B est bien semblable à A. Notez que cette proposition justifie la terminologie « A et B sont semblables » introduite dans la définition de la similitude.
2. Il existe $P, P' \in GL_n(\mathbb{R})$ telles que $A = PBP^{-1}$ et $B = P'CP'^{-1}$. On a donc $A = PP'CP'^{-1}P^{-1} = (PP')C(PP')^{-1}$. Comme $PP' \in GL_n(\mathbb{R})$, A est bien semblable à C.
3. Il suffit de prendre $P = I_n$.

On peut démontrer exactement les mêmes résultats pour la relation d'équivalence.

Solution de l'exercice 4.10. La dimension de $\mathscr{L}(\mathrm{E})$, que l'on notera M, est $(\dim(\mathrm{E}))^2$. La famille d'endomorphismes $\{f^0, f^1, \ldots, f^{\mathrm{M}}\}$ est une famille à $\mathrm{M}+1$ éléments de $\mathscr{L}(\mathrm{E})$ donc est liée d'après la définition même de la dimension (voir la définition 1.47). Il existe donc une combinaison linéaire nulle de ces endomorphismes dont les coefficients ne sont pas tous nuls, c'est-à-dire un polynôme annulateur non nul de f. La proposition relative aux matrices carrées se démontre de la même façon ou bien en utilisant l'endomorphisme canoniquement associé à une matrice carrée pour se ramener au cas que l'on vient de traiter.

Solution de l'exercice ??. Soit P un polynôme annulateur non nul de f *de degré minimal.* Si $\mathrm{P}(0) = 0$, c'est-à-dire si 0 est racine de P, alors P s'écrit XQ avec $\mathrm{Q} \in \mathbb{R}[\mathrm{X}]$, d'où l'on déduit que $0 = \mathrm{P}(f) = f\mathrm{Q}(f)$ et donc, si f est inversible, que $0 = \mathrm{Q}(f)$ (en composant les deux membres de l'égalité $0 = f\mathrm{Q}(f)$ par f^{-1}, ce qui montre que Q est un polynôme annulateur de f non nul (car $\mathrm{P} \neq 0$) et de degré strictement inférieur à celui de P (puisque $\mathrm{XQ} = \mathrm{P}$), ce qui contredit la définition de P. On ne peut donc avoir $\mathrm{P}(0) = 0$ si f est inversible. Réciproquement, si $\mathrm{P}(0) \neq 0$ alors on peut écrire :

$$\mathrm{P} = a_0 + \sum_{k=1}^{\deg(\mathrm{P})} a_k \mathrm{X}^k$$

où $a_0 \in \mathbb{R}^*$ et où $a_1, \ldots, a_{\deg(\mathrm{P})} \in \mathbb{R}$. En évaluant cette relation en f on obtient :

$$0 = \mathrm{P}(f) = a_0 \mathrm{Id} + \sum_{k=1}^{\deg(\mathrm{P})} a_k f^k$$

On a donc, puisque $a_0 \neq 0$:

$$\mathrm{Id} = -\frac{1}{a_0} \sum_{k=1}^{\deg(\mathrm{P})} a_k f^k$$

soit :

$$\mathrm{Id} = \left(-\frac{1}{a_0} \sum_{k=1}^{\deg(\mathrm{P})} a_k f^{k-1}\right) \circ f = f \circ \left(-\frac{1}{a_0} \sum_{k=1}^{\deg(\mathrm{P})} a_k f^{k-1}\right)$$

d'où l'inversibilité de f et l'expression de f^{-1} comme :

$$f^{-1} = -\frac{1}{a_0} \sum_{k=1}^{\deg(\mathrm{P})} a_k f^{k-1}$$

qui est bien un polynôme en f.

Solution de l'exercice 4.16. On démontre tout d'abord par une récurrence simple (que nous vous invitons à écrire) que pour toute valeur propre $\lambda \in \mathbb{R}$ de f et tout vecteur propre $x \in \mathrm{E}$ associé à λ on a :

$$\forall k \in \mathbb{N}, f^k(x) = \lambda^k x$$

Il suffit ensuite d'écrire P sous la forme $\sum_{k=0}^{\deg(\mathrm{P})} a_k \mathrm{X}^k$ où les a_i sont des réels et d'écrire que pour toute valeur propre $\lambda \in \mathbb{R}$ de f et tout vecteur propre $x \in \mathrm{E}$ associé à λ on a :

$$0 = \big(\mathrm{P}(f)\big)(x) = \left(\sum_{k=0}^{\deg(\mathrm{P})} a_k f^k\right)(x) = \sum_{k=0}^{\deg(\mathrm{P})} a_k f^k(x) = \sum_{k=0}^{\deg(\mathrm{P})} a_k \lambda^k x = \mathrm{P}(\lambda) x$$

d'où $\mathrm{P}(\lambda) = 0$ puisque $x \neq 0$.

Solution de l'exercice 4.37. Soient A et B deux matrices semblables de taille $p \in \mathbb{N}^*$ et $P \in GL_p(\mathbb{R})$ telle que $A = PBP^{-1}$. Comme la relation de similitude est symétrique, il suffit de montrer que toute valeur propre de A est valeur propre de B pour obtenir, par symétrie, que les valeurs propres de A sont exactement celles de B. Soit donc $\lambda \in \mathbb{R}$ une valeur propre de A et $x \in \mathbb{R}^p$ un vecteur propre associé. Alors $Ax = \lambda x$ donc $PBP^{-1}x = \lambda x$ soit $BP^{-1}x = P^{-1}x$; mais $P^{-1}x \neq 0$ car P^{-1} est inversible et $x \neq 0$, donc $P^{-1}x$ est vecteur propre de B associé à λ, qui est bien valeur propre de B.

On a par ailleurs montré que si x est un vecteur propre de A associé à une certaine valeur propre alors $P^{-1}x$ est un vecteur propre de B associé à la même valeur propre. On montre par symétrie que si y est un vecteur propre de B associé à une certaine valeur propre alors Py est un vecteur propre de x associé à la même valeur propre, ce qui permet d'écrire que si $\lambda \in \mathrm{Sp}(A) = \mathrm{Sp}(B)$ on a :

$$E_\lambda(B) = \{P^{-1}x \mid x \in E_\lambda(A)\}$$

et

$$E_\lambda(A) = \{Py \mid y \in \mathbb{E}_\lambda(B)\}$$

On constate notamment que les sous-espaces propres de A et B associés à une valeur propre donnée sont isomorphes (l'isomorphisme étant réalisé par la multiplication par P ou par P^{-1}).

Solution de l'exercice 4.36. A ne possède qu'une seule valeur propre, 1. Si elle était diagonalisable, elle devrait donc être semblable à la matrice identité. Ce n'est évidemment pas le cas ! En effet, s'il existait $P \in GL_3(\mathbb{R})$ telle que $A = PI_3P^{-1}$ on aurait $A = PP^{-1} = I_3$, ce qui est faux. A n'est donc pas diagonalisable.

On démontre de la même façon qu'une matrice semblable à une matrice scalaire est nécessairement cette même matrice scalaire et qu'aucune matrice non scalaire qui n'admet qu'une seule valeur propre n'est diagonalisable.

Solution de l'exercice 4.37. On cherche dans un premier temps à montrer que E s'écrit comme la somme de $\ker(f-\mathrm{Id})$ et de $\ker(f+\mathrm{Id})$, c'est-à-dire que tout $x \in E$ peut s'écrire comme la somme d'éléments de ces deux sous-espaces vectoriels. La relation $f^2 = \mathrm{Id}$ se réécrit sous la forme :

$$(f-\mathrm{Id}) \circ (f+\mathrm{Id}) = (f+\mathrm{Id}) \circ (f-\mathrm{Id}) = 0$$

ce qui indique que tout élément de la forme $f(y)+y$ est dans $\ker(f-\mathrm{Id})$ est que tout élément de la forme $f(y)-y$ est dans $\ker(f+\mathrm{Id})$. Il suffit donc d'écrire x comme somme de tels éléments ; pour cela, on écrit :

$$x = \frac{1}{2}\Big[\underbrace{x+f(x)}_{\in\ker(f-\mathrm{Id})} + \underbrace{f(-x)-(-x)}_{\in\ker(f-\mathrm{Id})}\Big]$$

On sait par ailleurs que les sous-espaces propres $\ker(f+\mathrm{Id}) = E_{-1}(f)$ et $\ker(f-\mathrm{Id}) = E_1(f)$ sont en somme directe (et ce même si -1 et/ou 1 ne sont pas valeurs propres de f puisqu'alors le sous-espace propre associé est $\{0\}$), d'où la relation :

$$E = \ker(f-\mathrm{Id}) \oplus \ker(f+\mathrm{Id})$$

E s'écrit alors comme somme directe de sous-espaces propres de f, ce qui implique que f est diagonalisable.

Solution de l'exercice 4.38. Supposons f nilpotent. Alors sa seule valeur propre possible est 0 : en effet, si $\lambda \neq 0$ était valeur propre de f et si $x \in \mathrm{E}$ était un vecteur propre associé, on aurait $f^k(x) = \lambda^k x \neq 0$ pour tout $k \in \mathbb{N}$ par une récurrence facile (écrivez-la !), ce qui est contraire au fait que $f^k = 0$ pour un certain $k \in \mathbb{N}^*$. Si de plus f est diagonalisable, la matrice de f dans n'importe quelle base est semblable à la matrice nulle, donc elle-même nulle, et $f = 0$. Comme l'endomorphisme nul est bien diagonalisable, on en déduit qu'un endomorphisme nilpotent est diagonalisable si et seulement s'il est nul.