UVSQ 2023/2024
Licence de sciences et technologie, santé
LSMA202N (Algèbre linéaire I)

Notes de cours

Table des matières

1 Ensembles, applications

La théorie des ensembles et la logique sont intrinsèquement liées. On ne développe ici aucun des deux aspects de ce même champ de connaissances qui fait, par ailleurs, l'objet de recherches contemporaines à l'intersection des mathématiques et de l'informatique théorique.

1.1 Descriptions des ensembles en mathématiques

En mathématiques, on distingue généralement deux types de description des ensembles :

(i) la description *en extension*, où l'on écrit entre accolades la liste exhaustive de tous les éléments de l'ensemble.

Par exemple, l'ensemble $\{0, \mathbb{R}, -\pi\}$ a trois éléments qui sont le nombre 0, l'ensemble $\mathbb{R}$ et le nombre $-\pi$.

(ii) La description *en compréhension*, où l'on écrit entre accolades une phrase qui décrit tous les éléments de l'ensemble. Cette phrase est le plus souvent écrite à l'aide d'un mélange de symboles mathématiques et de bouts de textes en langage courant.

Par exemple, l'ensemble $\{f,\ f$ est une fonction croissante $\mathbb{R} \to \mathbb{R}$ et $f(1) = -2\}$ est l'ensemble des fonctions croissantes dont l'ensemble de départ et l'ensemble d'arrivée sont $\mathbb{R}$, et qui vérifient $f(1) = -2$. La fonction $u : \mathbb{R} \to \mathbb{R}$ définie par la formule $\forall t \in \mathbb{R}$, $u(t) = -2 + (t-1)^3$ est un élément de cet ensemble — qui contient bien d'autres éléments, par ailleurs.

Exemples

(i) L'ensemble $E = \{0, 1, 2, 3\}$, défini ici en extension, peut se décrire en compréhension par $E = \{t \in \mathbb{Z},\ 0 \leq t \leq 3\}$.

(ii) L'ensemble $F = \left\{(x, y) \in \mathbb{Z}^2,\ x^2 + y^2 \leq 1\right\}$, décrit ici en compréhension, ne contient que 5 élements et peut s'écrire en extension $F = \{(0, 0), (1, 0), (0, 1), (-1, 0), (0, -1)\}$.

A noter

(i) La description en extension est réservée aux ensembles finis.

(ii) Dans une description en extension, la répétition d'un élément et l'ordre d'écriture n'ont pas d'effet. Par exemple, les ensembles $\left\{2 - 5, \sqrt{2}, \mathbb{N}, -1 - 2, \frac{2}{\sqrt{2}}\right\}$, $\left\{-3, \mathbb{N}, \sqrt{2}\right\}$ et $\left\{\mathbb{N}, -3, \sqrt{2}\right\}$ sont égaux.

(iii) A une sorte de mi-chemin entre extension et compréhension, on utilise parfois des pointillés lorsque le contexte est suffisamment peu ambigu. Par exemple, $\mathbb{N}^* = \{1, 2, \dots\}$, $]1, 19] \cap \mathbb{Z} = \{2, 3, \dots, 19\}$, $3\mathbb{Z} = \{\dots, -6, -3, 0, 3, 6, 9, \dots\}$.

(iv) Il existe un unique ensemble qui ne contient aucun élément : c'est l'ensemble vide, noté $\emptyset$ ou encore $\{\}$. Cette assertion gravite autour de l'axiomatique de la théorie des ensembles — selon le système axiomatique que l'on prend, son existence est un axiome, ou se déduit d'autres axiomes. Si E est n'importe quel ensemble, $\emptyset$ en est un sous-ensemble, que l'on appelle parfois sa *partie vide*.

1.2 Quelques liens entre logique et opérations sur les ensembles

Soient E et F deux ensembles.

Intersection et *et*

On note $E \cap F$ l'*intersection* de E et de F qui est l'ensemble des objets qui appartiennent à E *et* à F :

$$\boxed{\forall x,\ \ (x \in E \cap F) \Longleftrightarrow (x \in E \text{ et } x \in F)}$$

Réunion et *ou*

On note $E \cup F$ la *réunion* de E et de F qui est l'ensemble des objets qui appartiennent à E *ou* à F :

$$\boxed{\forall x,\ \ (x \in E \cup F) \Longleftrightarrow (x \in E \text{ ou } x \in F)}$$

A noter : le *ou* mathématique est *toujours* inclusif.

Inclusion et implication
On note $E \subset F$, ou encore indifféremment $E \subseteq F$, l'assertion E *est inclus dans* F qui signifie que tout élément de E est aussi élément de F :

$$\boxed{\big(E \subset F\big) \Longleftrightarrow \big(\forall x,\ \ (x \in E) \Longrightarrow (x \in F)\big)}$$

On dit aussi que E est une *partie* de F, ou encore que E est un *sous-ensemble* de F.

Exercice Montrer que $E \subseteq F$ équivaut à $\forall x$, $x \notin E$ ou $x \in F$.

A noter
Le symbole $\Longrightarrow$, reservé en mathématique à l'implication logique, ne signifie *jamais* "donc".

Egalité et équivalence
Deux ensembles sont égaux lorsqu'ils ont les mêmes éléments. Autrement dit,

$$\boxed{\big(E = F\big) \Longleftrightarrow \big(\forall x,\ \ (x \in E) \Longleftrightarrow (x \in F)\big)}$$

L'égalité entre deux ensembles équivaut ainsi à la fameuse double inclusion : en termes opératoires, pour montrer que $E = F$, on montre très souvent que $E \subset F$ et $F \subset E$.
De la même façon, l'équivalence entre deux assertions équivaut à la double implication : en termes opératoires, pour montrer que $p \Longleftrightarrow q$, on montre très souvent que $p \Longrightarrow q$ et $q \Longrightarrow p$.

Complémentaire, ou différence
Le *complémentaire de* F *dans* E, noté $E \setminus F$, est l'ensemble des éléments de E qui ne sont pas dans F :

$$\boxed{\forall x,\ \ (x \in E \setminus F) \Longleftrightarrow (x \in E \text{ et } x \notin F)}$$

Souvent, on parle de *différence* de E et F en réservant le terme de *complémentaire* au cas où F est une partie de E.

Différence symétrique
La *différence symétrique* de E et F, notée $E\Delta F = (E \cup F) \setminus (E \cap F)$, est le complémentaire de l'intersection dans la réunion de E et F :

$$\boxed{\forall x,\ \ (x \in E\Delta F) \Longleftrightarrow (x \in E \cup F \text{ et } x \notin E \cap F)}$$

Elle correspond au *ou exclusif*, souvent appelé *xor* en informatique (pour *exclusive or*, en anglais). En français, on dit parfois *ou bien*.

Produit cartésien
Le *produit cartésien* $E \times F$ est l'ensemble des *couples* formés d'un élément de E (en abscisse) et d'un élément de F (en ordonnée), dans cet ordre. On note les couples avec des parenthèses :

$$\boxed{E \times F = \{(x, y), x \in E, y \in F\}\,.}$$

Contrairement à la notation ensembliste pour laquelle $\{x, y\} = \{y, x\}$, l'ordre compte dans la notation des couples : en général, $(x, y) \neq (y, x)$. Plus précisément, $(x, y) = (y, x)$ si, et seulement si $x = y$. Plus généralement,

$$\boxed{(x, y) = (z, t) \Longleftrightarrow x = z \text{ et } y = t.}$$

Ensemble des parties
L'ensemble des parties de E est souvent noté $\mathcal{P}(E)$, ou encore 2^E (pour des raisons de cardinal, voir plus bas et en TD). Par exemple, $\emptyset \in \mathcal{P}(E)$ et $E \in \mathcal{P}(E)$, qui est une traduction des inclusions $\emptyset \subseteq E$ et $E \subseteq E$.

La logique mathématique se base toujours sur le principe du *tiers exclus* : une assertion et sa négation ne peuvent pas être vraies toutes les deux, alors qu'au moins une des deux est vraie. Autrement dit, dire qu'une assertion est vraie équivaut à dire que sa négation est fausse. La célébrissime proposition qui suit est un petit aperçu de la difficulté de la théorie des ensembles, dans laquelle, peut-être plus encore qu'ailleurs, toute évidence apparente doit être interrogée.

Proposition *Il n'existe pas d'ensemble de tous les ensembles.*

Preuve. En raisonnant par l'absurde, on suppose que l'ensemble de tous les ensembles existe et on le note $\mathcal{E}$. On note alors $\mathcal{F}$ son sous-ensemble $\mathcal{F} = \{x \in \mathcal{E},\ x \notin x\}$ (ce serait l'ensemble des ensembles qui ne sont pas éléments d'eux-mêmes). Le principe du tiers exclus impose que ou bien $\mathcal{F} \in \mathcal{F}$, ou bien $\mathcal{F} \notin \mathcal{F}$, mais pas les deux à la fois. Si $\mathcal{F} \in \mathcal{F}$, alors, selon la propriété qui définit $\mathcal{F}$ lui-même, $\mathcal{F} \notin \mathcal{F}$ et le principe du tiers exclus est contredit : impossible. Il est donc faux que $\mathcal{F} \in \mathcal{F}$, ce qui signifie que $\mathcal{F} \notin \mathcal{F}$. Mais alors $\mathcal{F}$ ne vérifie pas la propriété qui définit $\mathcal{F}$ en compréhension : il est faux que $\mathcal{F} \notin \mathcal{F}$, ce qui signifie que $\mathcal{F} \in \mathcal{F}$ et le principe du tiers exclus est à nouveau contredit : impossible. On a prouvé la proposition. ■

1.3 Gros symboles et autres abréviations

Soit $(E_i)_{i\in I}$ une famille d'ensembles, indexée par un ensemble non vide I − on ne fait pas d'autre hypothèse sur l'ensemble I des indices : c'est un ensemble quelconque, non vide♫.

Intersection et quantificateur universel $\forall$
L'intersection des E_i, notée $\bigcap\limits_{i\in I} E_i$ est l'ensemble des objets qui sont dans tous les E_i :

$$\boxed{\forall x,\quad \left(x \in \bigcap_{i\in I} E_i\right) \Longleftrightarrow (\forall i \in I,\ x \in E_i)}$$

Exemple : $\bigcap\limits_{x>0}]-x, x[= \{0\}$.

Réunion et quantificateur existentiel $\exists$
La réunion des E_i, notée $\bigcup\limits_{i\in I} E_i$ est l'ensemble des objets qui sont dans au moins un des E_i :

$$\boxed{\forall x,\quad \left(x \in \bigcup_{i\in I} E_i\right) \Longleftrightarrow (\exists i \in I,\ x \in E_i)}$$

Exemple : $\bigcup\limits_{x>0} [x, +\infty[=]0, +\infty[$.

Produits finis, puissances
Si $n \in \mathbb{N}^*$ et si $E_1, \dots, E_n$ sont des ensembles on note $E_1 \times E_2 \times \dots \times E_n = \prod_{i=1}^n E_i$ l'ensemble des n-uplets dont la i^{e} coordonnée est dans E_i, pour tout i :

$$\prod_{i=1}^n E_i = \{(x_1, x_2, \dots x_n)\,,\ \forall i \in \{1, \dots, n\}\,, x_i \in E_i\}\,.$$

Comme pour les couples, dans un n-uplets, l'ordre compte. En particulier, si tous les E_i sont égaux à un même ensemble E, on note

$$E^n = \prod_{i=1}^n E$$

l'ensemble des n-uplets dont toutes les coordonnées sont dans E. Un exemple que l'on utilisera souvent dans ce cours d'algèbre linéaire : si $n \in \mathbb{N}^*$,

$$\boxed{\mathbb{R}^n = \{(x_1, x_2, \dots x_n)\,,\ \forall i \in \{1, \dots, n\}\,, x_i \in \mathbb{R}\}\,.}$$

♫Cette hypothèse de non vacuité est là pour éviter les ennuis : que serait l'intersection d'une famille vide d'ensembles ?

1.4 Applications

Les applications sont le plus souvent enseignées au lycée en adoptant leur seul point de vue *dynamique*, qui est le suivant. Une *application* $f : E \to F$ est définie par : son *ensemble de départ* E, son *ensemble d'arrivée* F, son *graphe* $\{(x, y) \in E \times F,\ y = f(x)\}$. Son graphe est souvent – mais pas toujours – défini par une formule univoque qui décrit $f(x)$ "en fonction" de x, pour tout $x \in E$. Tout élément x de l'ensemble de départ a une unique image — notée $f(x)$ — dans l'ensemble d'arrivée.

Ainsi, deux applications sont égales lorsqu'elles ont le même ensemble de départ, le même ensemble d'arrivée et le même graphe. Autrement dit, $f : E \to F$ et $g : G \to H$ sont égales si, et seulement si $E = G$, $F = H$ et $\forall x \in E,\ f(x) = g(x)$.

Ce point de vue dynamique, qu'il faut absolument garder à l'esprit car c'est un point de vue opératoire irremplaçable, présente l'inconvénient de ne pas *définir* ce qu'*est* une application en termes de théorie des ensembles. Pour remédier à cela, on définit définitivement une application comme *étant* son graphe.

Définition (application)
Si E et F sont deux ensembles, une *application de E dans F* est une partie f du produit cartésien $E \times F$ telle que

$$\forall x \in E,\ \exists! y \in F,\ (x, y) \in f. \tag{1}$$

Lorsque $(x, y) \in f$, on note $y = f(x)$ — bien garder à l'esprit que cette notation est autorisée par l'axiome d'unicité dans la définition d'une application. Autrement dit, $f = \{(x, y) \in E \times F,\ y = f(x)\}$ – apprécier ici le premier signe $=$ à sa juste valeur. On dit sans changer d'habitude que E est *l'ensemble de départ de f* et F son *ensemble d'arrivée*.

Exemples
(i) Si E est un ensemble, l'application $\mathrm{id}_E : E \to E,\ x \mapsto x$ est appelée *application identité* de E. Son graphe est la *diagonale* de $E \times E$, c'est-à-dire $\{(x, x),\ x \in E\}$.

(ii) Ecrire "*soit $f : \mathbb{R} \to \mathbb{R}$ définie par $\forall x \in \mathbb{R},\ f(x) = \frac{x^2+1}{x^2-1}$*" ne définit pas une application $\mathbb{R} \to \mathbb{R}$. En revanche, l'application $g : \mathbb{R} \to \mathbb{R}$, $x \mapsto \frac{x^2+1}{x^2-1}$ si $x \in \mathbb{R} \setminus \{-1, 1\}$, $1 \mapsto 1$, $-1 \mapsto \sqrt{3}$, elle, est bien définie (exercice : dessiner son graphe).

(iii) Soient n et m des entiers naturels. Si E a n éléments et si F a m éléments, il y a exactement m^n applications $E \to F$.

[Compter naïvement : une application $E \to F$ doit attribuer à chacun des n éléments de E une image parmi les m éléments de F.]

Notation
L'ensemble des applications $E \to F$ est souvent noté $\mathcal{F}(E, F)$, ou encore F^E – cette dernière notation se justifie par l'exemple (iii) ci-dessus.

Définition (composée d'applications)
Si $f : E \to F$ et $g : F \to G$ sont des applications, l'*application composée* $g \circ f : E \to G$ est définie par

$$\forall x \in E,\ g \circ f(x) = g\left(f(x)\right).$$

On schématise souvent la situation par les dessins

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[rr, bend right, "g \circ f"'] & F \arrow[r, "g"] & G \end{tikzcd} \qquad \text{ou encore} \qquad \begin{tikzcd} E \arrow[r, "f"] \arrow[dr, "g \circ f"'] & F \arrow[d, "g"] \\ & G \end{tikzcd}$$

Exemple
On note $A = \mathbb{R} \setminus \{1\}$. La formule $f(x) = \frac{x+1}{x-1}$ définit une application $f : A \to A$ qui vérifie $f \circ f = \mathrm{id}_A$ (on dit que f est une *involution*).

Exercice
Si $f : E \to E$ et $g : E \to E$ sont des applications (mêmes ensembles au départ et à l'arrivée), en général, $f \circ g \neq g \circ f$. Prouver cela en exhibant un contre-exemple. Trouver aussi deux telles applications qui *commutent*, c'est-à-dire qui vérifient $f \circ g = g \circ f$.

Proposition ("associativité" de la composition des applications)
Soient $f : E \to F$, $g : F \to G$ et $h : G \to H$ des applications. Alors, $(h \circ g) \circ f = h \circ (g \circ f)$.

Preuve. Il suffit d'écrire l'image d'un élément arbitraire de E par les deux applications $(h \circ g) \circ f$ et $h \circ (g \circ f)$. Dans les deux cas, on trouve $h\left(g\left(f(x)\right)\right)$. ■

A noter
Dans les conditions de la proposition, on note $h \circ g \circ f$, sans parenthèse, pour désigner l'application $(h \circ g) \circ f = h \circ (g \circ f)$. C'est grâce à cette proposition que cette notation sans parenthèse a du sens.

1.5 Applications : opérations sur les parties

Définitions (image directe, image inverse, fibre)
Soient $f : E \to F$ une application, $A \subseteq E$ et $B \subseteq F$.
(i) L'*image (directe)* de A par f est l'ensemble des images par f des éléments de A :

$$\boxed{f(A) = \{f(x),\ x \in A\} = \{y \in F,\ \exists x \in A,\ y = f(x)\}.}$$

En particulier, l'*image de f* est $\mathrm{im}(f) = f(E) = \{y \in F,\ \exists x \in E,\ y = f(x)\} = \{f(x),\ x \in E\}$.
(ii) L'*image réciproque* (ou *inverse*) de B par f est l'ensemble des antécédents par f des éléments de B (gare à la notation !) :

$$\boxed{f^{-1}(B) = \{x \in E,\ f(x) \in B\}.}$$

(iii) Si $b \in F$, la *fibre de b*, que l'on note (dangereusement) $f^{-1}(b)$ est l'ensemble des antécédents de b par f :

$$\boxed{f^{-1}(b) = f^{-1}(\{b\}) = \{x \in E,\ f(x) = b\}.}$$

Exemples
(i) Soit $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$. Alors, $f([-2,-1]) = [1,4]$, $f^{-1}([1,4]) = [-2,-1] \cup [1,2]$, $f^{-1}(1) = \{-1,1\}$ et $f^{-1}(0) = \{0\}$.
(ii) Soit $p : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto x$. Alors, pour tout $x \in \mathbb{R}$, $p^{-1}(x) = \{(x,y),\ y \in \mathbb{R}\} = \{x\} \times \mathbb{R}$. Par ailleurs, si $D = \left\{(x,y) \in \mathbb{R}^2,\ x^2 + y^2 = 1\right\}$, alors $p(D) = [-1,1]$.

Proposition (actions d'une application sur les sous-ensembles)
Soient $f = E \to F$ une application, S et T des parties de E, U et V des parties de F.
(i) Si $U \subseteq V$, alors $f^{-1}(U) \subseteq f^{-1}(V)$
(ii) Si $S \subseteq T$, alors $f(S) \subseteq f(T)$
(iii) $f^{-1}(U) \cap f^{-1}(V) = f^{-1}(U \cap V)$
(iv) $f^{-1}(U) \cup f^{-1}(V) = f^{-1}(U \cup V)$
(v) $f^{-1}(U) \setminus f^{-1}(V) = f^{-1}(U \setminus V)$
(vi) $f(S \cap T) \subseteq f(S) \cap f(T)$
(vii) $f(S \cup T) = f(S) \cup f(T)$
(viii) $f(U) \setminus f(V) \subseteq f(U \setminus V)$

Preuve. Exercice. ■

Exercice
Trouver des contre-exemples qui prouvent qu'il n'y a en général pas d'équivalence dans *(i)* ni dans *(ii)*, qu'il n'y a en général pas d'égalité dans *(vi)* ni dans *(viii)*.

Exercice
Reprendre les assertions *(iii)*,*(iv)*,*(vi)* et *(vii)* ci-dessus en remplaçant S et T par une famille de parties de E et en remplaçant U et V par une famille de parties de F. Autrement dit, avec des notations évidentes, comparer pour l'inclusion :

(a) $f^{-1}\left(\bigcap_{i\in I} U_i\right)$ et $\bigcap_{i\in I} f^{-1}(U_i)$;

(b) $f^{-1}\left(\bigcup_{i\in I} U_i\right)$ et $\bigcup_{i\in I} f^{-1}(U_i)$;

(c) $f\left(\bigcup_{i\in I} S_i\right)$ et $\bigcup_{i\in I} f(S_i)$;

(d) $f\left(\bigcap_{i\in I} S_i\right)$ et $\bigcap_{i\in I} f(S_i)$.

1.6 Injection, surjection, bijection

Définitions (injection, surjection, bijection)
Soit $f : X \to Y$ une application.

(i) f est *injective* lorsque $\forall x, x' \in X,\ f(x) = f(x') \Longrightarrow x = x'$.

(ii) f est *surjective* lorsque $\forall y \in Y,\ \exists x \in X,\ y = f(x)$.

(iii) f est *bijective* lorsque f est injective et surjective.

Paraphrases : autres façons de dire l'injectivité

(i) f est injective si, et seulement si tout élément de l'ensemble d'arrivée a *au plus* un antécédent dans l'ensemble de départ.

(ii) Par contraposition à la définition, f est injective si, et seulement si $\forall x, x' \in X,\ x \neq x' \Longrightarrow f(x) \neq f(x')$.

(iii) f est injective si, et seulement si pour tout $y \in Y$, l'équation $f(x) = y$ a *au plus* une solution x dans X.

Paraphrases : autres façons de dire la surjectivité

(i) f est surjective si, et seulement si tout élément de l'ensemble d'arrivée a *au moins* un antécédent dans l'ensemble de départ.

(ii) En termes plus compacts, f est surjective si, et seulement si $f(X) = Y$, ou encore si, et seulement si $f(X) \supseteq Y$.

(iii) f est surjective si, et seulement si pour tout $y \in Y$, l'équation $f(x) = y$ a *au moins* une solution x dans X.

Paraphrases : premières autres façons de dire la bijectivité

(i) f est bijective si, et seulement si tout élément de l'ensemble d'arrivée a *un unique* un antécédent dans l'ensemble de départ.

(ii) f est bijective si, et seulement si pour tout $y \in Y$, l'équation $f(x) = y$ a *une unique* une solution x dans X.

Exemples prototypiques

(i) Si $A \subseteq X$, l'application $A \to X$, $a \mapsto a$ est injective (prototype d'injection).

(ii) Si $Y \neq \emptyset$, alors $X \times Y \to X$, $(x, y) \mapsto x$ est surjective (prototype de surjection).

(iii) Si X est un ensemble, $\mathrm{id}_X : X \to X$, $x \mapsto x$ est bijective (prototype de bijection).

Exercice
Si $f : X \to Y$ est une application, alors l'application $\widetilde{f} : X \to \mathrm{im}(f)$, $x \mapsto f(x)$ est bien définie et est surjective. Si en outre f est injective, alors $\widetilde{f}$ est une bijection entre X et $f(X)$.

Proposition (bijectivité par existence d'une réciproque)
Une application $f : X \to Y$ est bijective si, et seulement s'il existe $g : Y \to X$ telle que $f \circ g = \mathrm{id}_Y$ et $g \circ f = \mathrm{id}_X$. Dans ces conditions, g est unique et est aussi bijective.

Preuve. (i) On suppose d'abord que g existe. Soient $x, x' \in X$ tels que $f(x) = f(x')$. Alors, $g(f(x)) = g(f(x'))$, ce qui s'écrit encore $x = x'$: l'application f est injective. Pour montrer que f est surjective, soit $y \in Y$, et soit $x = g(y)$ l'image de y par g. Alors, $f(x) = f \circ g(y) = y$, ce qui montre que x est un antécédent de y par f.
(ii) Réciproquement, on suppose que f est bijective et on montre l'existence de g. On donne ici deux versions d'une démonstration.
D'abord, le point de vue dynamique. Puisque f est bijective, pour tout $y \in Y$, il existe un unique $x \in X$ tel que $y = f(x)$. La combinaison de l'existence et de l'unicité autorise, pour chaque $y \in Y$, à nommer ledit $x \in X$

sous la forme $x = g(y)$, ce qui définit une application $g : Y \to X$. Enfin, avec ces notations, $g(f(x)) = g(y) = x$ et $f(g(y)) = f(x) = y$, ce qui montre les égalités $f \circ g = \mathrm{id}_Y$ et $g \circ f = \mathrm{id}_X$.
Ensuite, le point de vue des graphes. On note $\mathcal{G}_f = \{(x, y) \in X \times Y,\ y = f(x)\}$ le graphe♫ de f. Dire que f est une application signifie que $\forall x \in X,\ \exists ! y \in Y,\ (x, y) \in \mathcal{G}_f$. La bijectivité de f sur son graphe se traduit par : $\forall y \in Y,\ \exists ! x \in X,\ f(x) = y$. En d'autres mots, l'ensemble $\{(y, x) \in Y \times X,\ y = f(x)\}$ est le graphe d'une application que l'on note g. Pour tout $(x, y) \in X \times Y$, on a alors $f(x) = y \iff x = g(y)$ et les égalités de composition en résultent.
(iii) Unicité de g. On suppose que $h : Y \to X$ est aussi une application qui vérifie $f \circ h = \mathrm{id}_Y$ et $h \circ f = \mathrm{id}_X$. Soit $y \in Y$. Alors, $h(y) = h(f(g(y))) = g(y)$, la première égalité venant de $f \circ g = \mathrm{id}_Y$, la seconde de $h \circ f = \mathrm{id}_X$. Cela montre que $g = h$. A noter : on a utilisé l'"associativité" de la composition des applications.
(iv) Bijectivité de g : les relations $f \circ g = \mathrm{id}_Y$ et $g \circ f = \mathrm{id}_X$ prouvent, grâce à (i), que g est bijective. ■

Notation et définition
Dans les conditions de la proposition précédente, l'application g est notée f^{-1}. C'est l'*application réciproque* ou encore l'*inverse* de f. Avec ces notations, lorsque $f : X \to Y$ est bijective, pour tout $(x, y) \in X \times Y$,

$$\boxed{f(x) = y \iff x = f^{-1}(y)}$$

A noter
On peut dégager trois grandes classes de méthodes (voisines, malgré tout) pour montrer qu'une application $f : X \to Y$ est bijective.

(i) On montre séparément que f est injective et surjective.

(ii) On montre que pour tout $y \in Y$, l'équation $f(x) = y$ a une unique solution dans X.

(iii) On exhibe la réciproque de f en trouvant une application $g : Y \to X$ telle que $f \circ g = \mathrm{id}_Y$ et $g \circ f = \mathrm{id}_X$.

Proposition (composée de bijections)
Si $f : X \to Y$ et $g : Y \to Z$ sont des bijections, alors $g \circ f : X \to Z$ est également bijective, et $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Preuve. Grâce à l'"associativité" de la composition, les égalités $(g \circ f) \circ \left(f^{-1} \circ g^{-1}\right) = g \circ g^{-1} = \mathrm{id}_Z$ et $\left(f^{-1} \circ g^{-1}\right) \circ (g \circ f) = f^{-1} \circ f = \mathrm{id}_X$ montrent la proposition. ■

Exercices

(i) La composée de deux injections est une injection. La composée de deux surjections est une surjection.

(ii) Soient $i : X \to Y$ et $s : Y \to Z$ des applications. Si $s \circ i$ est injective, alors i est injective. Si $s \circ i$ est surjective, alors s est surjective.

(iii) Si f est bijective, $\left(f^{-1}\right)^{-1} = f$.

(iv) Soit E un ensemble. On note $\mathcal{P}(E)$ l'ensemble des parties de E et $\{0, 1\}^E$ l'ensemble des applications de E dans l'ensemble à deux éléments $\{0, 1\}$. Si $A \in \mathcal{P}(E)$, on note $\mathbf{1}_A$ l'application $E \to \{0, 1\}$ définie par : $\forall x \in E$, $\mathbf{1}_A(x) = 1$ si, et seulement si $x \in A$ (cette application est souvent appelée fonction *indicatrice* de la partie A). Montrer que l'application

$$\begin{array}{ccc} \mathcal{P}(E) & \longrightarrow & \{0,1\}^E \\ A & \longmapsto & \mathbf{1}_A \end{array}$$

est une bijection.

(v) Si f est bijective, alors $f^{-1}(B)$ est (aussi) l'image directe de B par l'application f^{-1}.

Attention à cette notation dangereuse, source de multiples erreurs estudiantines. Alors que $f^{-1}(B)$, vue comme image inverse de B par l'application f, est *toujours* définie, $f^{-1}(B)$, vue comme image directe de B par l'application réciproque f^{-1}, n'a de sens que lorsque f est une bijection.

1.7 Tout petits éléments sur les cardinaux

L'idée qui se cache derrière la notion de *cardinal*, que l'on ne définit pas ici, est celle du "nombre" d'éléments d'un ensemble.

♫Bien entendu, $f = \mathcal{G}_f$, selon la définition d'une application dite en 1.4. Plus bas, encore $g = \mathcal{G}_g$.

Définition (ensembles équipotents)
Si E et F sont deux ensembles, on dit que E est *équipotent* à F lorsqu'il existe une bijection $E \to F$.

Exercice
Soient E, F et G des ensembles. Montrer que :

(i) E est équipotent à E ;

(ii) si E est équipotent à F, alors F est équipotent à E ;

(iii) si E est équipotent à F et si F est équipotent à G, alors E est équipotent à G.

Définition (avoir le même cardinal)
Lorsque deux ensembles E et F sont équipotents, ont dit qu'ils *ont le même cardinal*.

Définitions (ensemble infini, ensemble fini)
Un ensemble E est *infini* lorsqu'il est équipotent à une de ses parties strictes – c'est-à-dire à une partie de E qui soit différente de E. Un ensemble *fini* est un ensemble non infini (sans blague !).

Exercice
(i) Montrer que $\mathbb{N}$ et $\mathbb{R}$ sont infinis.

(ii) Montrer que $]0,1[$, $\mathbb{R}_{>0} = \mathbb{R}_+^* =]0,+\infty[$ et $\mathbb{R}$ sont équipotents.

Proposition (applications entre deux ensembles finis de même cardinal)
Soient E et F deux ensembles finis de même cardinal, et $f : E \to F$ une application. Alors, il y a équivalence entre :

(i) f est injective ;

(ii) f est surjective ;

(iii) f est bijective.

PREUVE. Il suffit de montrer que *(i)* et *(ii)* sont équivalentes.
Si f est injective, elle définit une bijection de E sur $f(E)$. Ainsi, $f(E)$ est une partie de F dont le cardinal égale celui de F. Puisque F est fini, cela entraîne que $F = f(E)$: l'application f est surjective.
Inversement, on suppose que f est surjective. Soit $y \in F$. Puisque f est surjective, soit $x \in f^{-1}(y)$. On considère le sous-ensemble E' de E défini par

$$E' = \left(E \setminus f^{-1}(y)\right) \cup \{x\}\,.$$

On note $i : E' \to E$ la restriction de id_E à E', ce qui signifie que $i(x) = x$ pour tout $x \in E'$. On définit aussi l'application $j : E \to E'$ par la disjonction des cas suivante : pour tout $t \in E$, si $f(t) \neq y$ alors $j(t) = t$ et si $f(t) = y$ alors $j(t) = x$. L'ensemble d'arrivée de j ainsi définie est bien inclus dans E' et, par ailleurs, $i \circ j = \mathrm{id}_E$ et $j \circ i = \mathrm{id}_{E'}$. Ainsi, i est une bijection de E sur E', qui ont donc le même cardinal fini. Comme E' est une partie de E, cela entraîne que $E = E'$, et donc que $f^{-1}(y) = \{x\}$. On a montré que y a un unique antécédent. Comme cela est vrai pour tout $y \in F$, on a montré que f est bijective. ∎

Définition (ensemble dénombrable)
Un ensemble est dit *dénombrable* lorsqu'il est équipotent à l'ensemble $\mathbb{N}$ des entiers naturels.

Exemples
(i) $\mathbb{Z}$, $\mathbb{N}^2$ et $\mathbb{Q}$ sont dénombrables, ainsi que toute union dénombrable d'ensembles dénombrables.

(ii) L'ensemble $\mathbb{R}$ des nombres réels n'est pas dénombrable (argument de la diagonale de Cantor).

Les preuves de (i) et (ii), tout en étant accessibles, ne sont pas évidentes mais très classiques. On en trouvera dans la littérature ou sur internet, à la condition impérative de faire preuve d'esprit critique à leur lecture.

1.8 Les attendus du chapitre

Lire tout le chapitre en détail, s'assurer de tout comprendre y compris les exemples, sans laisser de zone d'ombre. Faire, pour soi, une rédaction d'une résolution des exercices énoncés.

Section 1.1
Pour toutes les mathématiques, il est indispensable de savoir comprendre l'écriture d'une ensemble et, inversement, de savoir écrire un ensemble de façon non ambiguë. Cet apprentissage se fait au long cours ; cette section est une sorte de mise au point initiale. S'entraîner à écrire des ensembles avec le formalisme usuel (nombres impairs, entiers multiples de 3, cercle unité dans le plan, nombres irrationnels de l'intervalle $[0,1[$, fonctions impaires, fonctions non croissantes, etc). Inventer, échanger.

Section 1.2
Cette section ne présente aucun objet véritablement nouveau pour les lycéen·ne·s, hormis peut-être la différence symétrique. S'assurer que toutes les connexions entre les opérations ensemblistes ($\cap$, $\cup$, etc) et les opérations logiques (*et, ou*, $\Rightarrow$, etc) sont bien acquises et merveilleusement huilées. La proposition sur la non existence de l'ensemble de tous les ensembles doit être comprise, son résultat doit être retenu comme une veille de l'esprit critique face aux évidences, mais il n'est pas indispensable d'en retenir la preuve.
Se poser la question du mélange des symboles ensemblistes : peut-on écrire autrement $A \cap (B \cup C)$? ou encore $(A \setminus (B \cup C))\,\Delta D$? etc. Inventer des questions, s'efforcer d'y répondre.

Section 1.3
Assimiler les notations qui sont nouvelles. Bien comprendre ce que signifie $\mathbb{R}^n$ lorsque n est un entier naturel. Y a-t-il un rapport entre $(E \times F) \times G$, $E \times (F \times G)$ et $E \times F \times G$? entre $\left(\mathbb{R}^2\right) \times \mathbb{R}$ et $\mathbb{R} \times \left(\mathbb{R}^2\right)$ et $\mathbb{R}^3$? Inventer des questions, s'efforcer d'y répondre.

Section 1.4
S'assurer de bien comprendre ce que signifie que deux applications sont égales, et comment on peut s'y prendre pour montrer une telle égalité en termes opératoires. Prendre conscience qu'une fonction et son graphe sont indiscernables, mais retenir l'aspect opératoire du point de vue dynamique : "une application, c'est un ensemble de départ, un ensemble d'arrivée, une flèche". Savoir sans hésiter la définition de la composée de deux applications. Retenir l'énoncé de l'"associativité" de la composition.

Section 1.5
Avec beaucoup de soin, apprendre la définition de l'image directe et de l'image inverse. Ne pas tenter de retenir par cœur la proposition des actions d'une application sur les sous-ensembles, mais retenir que de tels énoncés existent et savoir les retrouver très vite et sans erreur par un raisonnement élémentaire.

Section 1.6
Avec beaucoup de soin, apprendre la définition d'injection, de surjection, de bijection. Prendre le temps de bien comprendre pourquoi les "paraphrases" sont bien d'autres façons équivalentes de dire les définitions, et retenir ces divers points de vue.
Retenir les exemples prototypiques.
Bien comprendre la preuve de la proposition de bijectivité par existence d'une réciproque. En retenir le point clef – la définition de f^{-1} lorsque l'on sait que f est bijective.
Apprendre avec soin la définition de la réciproque d'une bijection.
Retenir les trois grandes méthodes de preuve d'une bijectivité et s'y référer à chaque fois qu'un raisonnement du cours ou des TD s'y prête.
Bien retenir que l'inverse d'une composée de bijections est la composée de l'inverse, **en changeant l'ordre**.
Etre bien certain·e de bien comprendre la mise en garde sur les deux sens de la notation $f^{-1}(B)$.

Section 1.7
Retenir la définition d'un ensemble infini. Se demander comment montrer que les ensembles familiers ($\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^2$, $\mathbb{N} \times \mathbb{R}$, etc) sont infinis, ont ou non le même cardinal, etc.
Retenir l'énoncé de la proposition sur les applications entre deux ensembles finis de même cardinal, en prenant garde de n'oublier aucune hypothèse. Trouver des exemples d'applications injectives et non bijectives (*resp.* surjectives non bijectives) entre deux ensembles infinis de même cardinal.

2 Espaces vectoriels, indépendance linéaire

2.1 La structure d'espace vectoriel

Définition (espace vectoriel réel)
Soit V un ensemble, muni de deux lois :
– une *loi de composition interne* que l'on appelle *addition* et que l'on note usuellement $+$, c'est-à-dire une application

$$\begin{array}{ccc} V \times V & \longrightarrow & V \\ (v, w) & \longmapsto & v + w \end{array}$$

– une *loi de composition externe* que l'on appelle *multiplication scalaire* et que l'on note usuellement $\cdot$ ou par absence de symbole, c'est-à-dire une application

$$\begin{array}{ccc} \mathbb{R} \times V & \longrightarrow & V \\ (x, w) & \longmapsto & x \cdot w = xw. \end{array}$$

Dans ces conditions, on dit que $(V, +, \cdot)$ a une structure de $\mathbb{R}$-*espace vectoriel*, ou simplement *est* un espace vectoriel sur $\mathbb{R}$ lorsque les axiomes suivants sont vérifiés :

(i) $\forall u, v, w \in V$, $(u + v) + w = u + (v + w)$

(ii) $\forall v, w \in V$, $v + w = w + v$

(iii) il existe $0_V \in V$ tel que $\forall v \in V$, $v + 0_V = v$

(iv) $\forall v \in V$, $\exists w \in V$, $v + w = 0_V$

(v) $\forall x, y \in \mathbb{R}$, $\forall v \in V$, $(x + y)v = xv + yv$

(vi) $\forall x \in \mathbb{R}$, $\forall v, w \in V$, $x(v + w) = xv + xw$

(vii) $\forall x, y \in \mathbb{R}$, $\forall v \in V$, $(xy)v = x(yv)$

(viii) $\forall x \in V$, $1 \cdot v = v$.

A noter
(i) L'axiome (i) autorise à noter $u + v + w$, sans parenthèses, l'élément $u + v + w = (u + v) + w = u + (v + w)$. De la même façon, si n est un entier naturel non nul et si $v_1, \dots, v_n$ sont des vecteurs, tous les parenthésages possibles de la somme

$$v_1 + \cdots + v_n = \sum_{1 \le k \le n} v_k$$

fournissent le même élément de V, ce qui autorise cette notation compacte avec le gros symbole de somme (preuve par récurrence sur n).

(ii) Axiome (iii) : il y a un *unique* tel 0_V.
En effet, si $0'_V$ vérifie également $\forall v \in V$, $v + 0'_V = v$, alors $0'_V + 0_V = 0_V + 0'_V = 0_V$ d'une part (propriété de $0'_V$) et $0'_V + 0_V = 0'_V$ d'autre part (propriété de 0_V) ; donc $0'_V = 0_V$.

(iii) Axiome (iv) : pour tout $v \in V$, il y a un *unique* $w \in V$ tel que $v + w = 0_V$. On le note $w = -v$.
En effet, si w' vérifie également $v + w' = 0_V$, alors $w + v + w' = (w + v) + w' = (v + w) + w' = 0_V + w' = w'$ d'une part, et $w + v + w' = w + (v + w') = w + 0_V = w$ d'autre part. Donc $w = w'$.

Vocabulaire courant
Les éléments d'un espace vectoriel sont appelés *vecteurs*, les nombres réels, par lesquels on peut multiplier les vecteurs, sont les *scalaires*. Le vecteur 0_V, que l'on note souvent simplement 0, est le *vecteur nul*, que l'on appelle parfois simplement *zéro*, au risque de confondre avec le nombre réel 0.
Si $v \in V$, l'unique vecteur w qui vérifie $v + w = 0_V$ est l'*opposé* de v. On le note $-v$ et, comme dans la notation pour les nombres, si $u \in V$, on note $u - v = u + (-v)$.

Quelques règles de calcul
Soit V un espace vectoriel.

(i) $\forall v \in V$, $0 \cdot v = 0_V$.
[En effet, si on note $w = 0 \cdot v$, alors $w + w = (0 + 0) \cdot v = 0 \cdot v = w$. En ajoutant $-w$ aux deux membres de cette égalité, on obtient $w = 0_V$.]

(ii) $\forall x \in \mathbb{R}$, $x \cdot 0_V = 0_V$.

[En effet, $x \cdot 0_V = x\,(0_V + 0_V) = x \cdot 0_V + x \cdot 0_V$. En ajoutant $-x \cdot 0_V$ aux deux membres de cette égalité, on obtient $0_V = x \cdot 0_V$.]

(iii) $\forall v \in V$, $(-1) \cdot v = -v$.

[En effet, $v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 + (-1)) \cdot v = 0 \cdot v = 0_V$.]

(iv) $\forall v \in V$, $v + v = 2 \cdot v$, $v + v + v = 3 \cdot v$. De façon générale, si $n \in \mathbb{N}^*$, $n \cdot v = v + v + \cdots + v$ (n fois).

[Preuve par récurrence sur n).]

Proposition (nullité d'un produit externe)
Soient V un espace vectoriel réel, $v \in V$ et $x \in \mathbb{R}$. Alors, $x \cdot v = 0_V$ si, et seulement si $x = 0$ ou $v = 0_V$.

Preuve. On démontre l'implication de gauche à droite, l'autre découlant des règles de calcul *(i)* et *(ii)* qui précèdent. On suppose que $x \cdot v = 0_V$ et que $x \neq 0$. Alors, le nombre réel $\frac{1}{x}$ existe. On en déduit que $\frac{1}{x}xv = \left(\frac{1}{x}x\right)v = 1 \cdot v = v$ d'une part, et d'autre part, $\frac{1}{x}xv = \frac{1}{x} \cdot (x \cdot v) = \frac{1}{x} \cdot 0_V = 0_V$. Donc $v = 0_V$. ■

Définition (combinaison linéaire)
Soient V un espace vectoriel réel, $n \in \mathbb{N}^*$ et $v_1, \ldots, v_n \in V$. Une *combinaison linéaire* de $v_1, \ldots, v_n$ est un vecteur de V de la forme

$$x_1v_1 + \cdots + x_nv_n = \sum_{1 \leq k \leq n} x_kv_k$$

où $x_1, \ldots, x_n \in \mathbb{R}$.

A noter

(i) En particulier, le vecteur nul est combinaison linéaire de toute famille finie de vecteurs (prendre tous les scalaires nuls).

(ii) Si $(v_i)_{i \in I}$ est une famille quelconque de vecteurs, une combinaison linéaire des v_i est, par définition, une somme de la forme

$$\sum_{i \in I} x_iv_i$$

où $(x_i)_{i \in I}$ est une famille *presque nulle* de nombres réels, ce qui signifie que seuls les x_i d'un sous-ensemble fini de I sont non nuls. Dans ces conditions, la somme est finie et a du sens avec les seuls axiomes d'espace vectoriel.

Exemples fondamentaux d'espaces vectoriels

(i) L'*espace nul*, un peu bébête, constitué d'un seul élément qui est nécessairement son vecteur nul – noter qu'il n'y a qu'une façon de définir $+$ et $\cdot$ sur un singleton ! On le note usuellement $\{0\}$, ou (0), ou encore simplement 0.

(ii) Les *espaces vectoriels* $\mathbb{R}^n$ où $n \in \mathbb{N}^*$.
Les lois sont définies par les formules suivantes : $\forall\,(u_1, \ldots, u_n) \in \mathbb{R}^n$, $\forall\,(v_1, \ldots, v_n) \in \mathbb{R}^n$, $\forall x \in \mathbb{R}$,

$$\begin{cases} (u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots, u_n + v_n) \\ x \cdot (u_1, \ldots, u_n) = (xu_1, \ldots, xu_n) \end{cases}$$

– la somme et la multiplication scalaire se font coordonnée par coordonnée. Exercice : ces lois confèrent à $\mathbb{R}^n$ une structure de $\mathbb{R}$-espace vectoriel.

En particulier, pour $n = 1$, le corps $\mathbb{R}$ des nombres réels est un $\mathbb{R}$-espace vectoriel pour son addition et sa multiplication usuelles.

(iii) Les *espaces de fonctions* $\mathbb{R}^E$
Soient E un ensemble et $\mathbb{R}^E$ l'énorme ensemble des applications $E \to \mathbb{R}$. Les *opérations usuelles* sur $\mathbb{R}^E$ sont les suivantes : $\forall f, g \in \mathbb{R}^E$, $\forall x \in \mathbb{R}$, les fonctions $f + g$ et $x \cdot f$ sont définies par les formules

$$\forall e \in E,\ \ (f + g)\,(e) = f(e) + g(e) \ \text{ et } \ (x \cdot f)\,(e) = xf(e).$$

Exercice : ces lois confèrent à $\mathbb{R}^E$ une structure d'espace vectoriel sur $\mathbb{R}$.
Par exemple, les applications $\mathbb{R} \to \mathbb{R}$ forment un espace vectoriel réel (pour ces lois). C'est également le cas des suites à valeurs réelles, qui sont les applications $\mathbb{N} \to \mathbb{R}$.

(iv) Le corps $\mathbb{C}$ des nombres complexes
L'addition et la multiplication usuelles dans $\mathbb{C}$ confèrent à $\mathbb{C}$ une structure d'espace vectoriel réel.

Exercice : prouver cela.

(v) Les *espaces de matrices* $\mathcal{M}_{p,q}(\mathbb{R})$.
Si p et q sont des entiers naturels non nuls, une *matrice à p lignes et q colonnes et à coefficients réels* est un tableau de nombres du type

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,q} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,q} \\ \vdots & \vdots & & \vdots \\ a_{p,1} & a_{p,2} & \cdots & a_{p,q} \end{pmatrix}$$

où $a_{\ell,c}$ est un nombre réel pour tout $(\ell,c) \in \{1,\ldots,p\} \times \{1,\ldots,q\}$. Si $(\ell,c) \in \{1,\ldots,p\} \times \{1,\ldots,q\}$, le nombre réel $a_{\ell,c}$ est le *coefficient* de la ℓ^{e} ligne et de la c^{e} colonne de A. Deux matrices sont égales si, et seulement si tous leurs coefficients sont égaux. En notation compacte, la matrice ci-dessus est simplement notée

$$\boxed{A = (a_{\ell,c})_{(\ell,c)\in\{1,\ldots p\}\times\{1,\ldots q\}}}$$

ou encore

$$A = (a_{\ell,c})_{1\leq \ell\leq p, 1\leq c\leq q}.$$

Si $p,q \in \mathbb{N}^*$, on note $\boxed{\mathcal{M}_{p,q}(\mathbb{R})}$ l'ensemble des matrices à p lignes et q colonnes et à coefficients réels :

$$\mathcal{M}_{p,q}(\mathbb{R}) = \left\{ \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,q} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,q} \\ \vdots & \vdots & & \vdots \\ a_{p,1} & a_{p,2} & \cdots & a_{p,q} \end{pmatrix},\ \forall (\ell,c) \in \{1,\ldots p\} \times \{1,\ldots q\},\ a_{\ell,c} \in \mathbb{R} \right\}.$$

L'addition et la multiplication scalaire sur $\mathcal{M}_{p,q}(\mathbb{R})$ sont définies par les formules suivantes :
$\forall A = (a_{\ell,c})_{(\ell,c)\in\{1,\ldots p\}\times\{1,\ldots q\}} \in \mathcal{M}_{p,q}(\mathbb{R})$, $\forall B = (b_{\ell,c})_{(\ell,c)\in\{1,\ldots p\}\times\{1,\ldots q\}} \in \mathcal{M}_{p,q}(\mathbb{R})$, $\forall x \in \mathbb{R}$,

$$\boxed{\begin{cases} A + B = (a_{\ell,c} + b_{\ell,c})_{(\ell,c)\in\{1,\ldots p\}\times\{1,\ldots q\}} \\ x \cdot A = (x a_{\ell,c})_{(\ell,c)\in\{1,\ldots p\}\times\{1,\ldots q\}} \end{cases}}$$

Autrement dit, l'addition et la multiplication scalaire se font coefficient par coefficient. Exercice : ces opérations confèrent à $\mathcal{M}_{p,q}(\mathbb{R})$ une structure d'espace vectoriel.

Vecteurs-ligne
L'espace $\mathcal{M}_{1,q}(\mathbb{R})$ des matrices à une ligne et q colonnes est aussi l'espace $\mathbb{R}^q$. Pour désigner les éléments de $\mathcal{M}_{1,q}(\mathbb{R})$, on utilise indifféremment les vocables q-uplets, vecteurs-ligne de dimension q, matrices-ligne à q colonnes.

Vecteurs-colonnes
L'espace $\mathcal{M}_{p,1}(\mathbb{R})$ est l'espace des matrices à p lignes et une colonne. On utilise indifféremment les vocables vecteurs-colonne de dimension p ou matrices-colonne à p lignes pour désigner les éléments de $\mathcal{M}_{p,1}(\mathbb{R})$.

Proposition (produit d'espaces vectoriels réels)
Soient V et W deux $\mathbb{R}$-espaces vectoriels dont on note identiquement $+$ et $\cdot$ les lois. Les opérations définies par : $\forall (v,w) \in V \times W,\ \forall (v',w') \in V \times W,\ \forall x \in \mathbb{R}$,

$$\begin{cases} (v,w) + (v',w') = (v+v', w+w') \\ x \cdot (v,w) = (xv, xw) \end{cases}$$

confèrent à $V \times W$ une structure de $\mathbb{R}$-espace vectoriel.

Preuve. Exercice ■

Définition (espace produit)
Dans les conditions de la proposition précédente, l'espace vectoriel $V \times W$ est l'*espace produit* ou encore la *somme directe externe* des espaces V et W.

2.2 Sous-espaces vectoriels

Définition (sous-espace vectoriel)
Soient V un espace vectoriel réel et W un sous-ensemble de V. On dit que W est un *sous-espace vectoriel* de V lorsque :

(i) $0_V \in W$

(ii) W est *stable pour la somme* : $\forall v, w \in W$, $v + w \in W$

(iii) W est *stable pour la multiplication scalaire* : $\forall v \in W$, $\forall x \in \mathbb{R}$, $x \cdot v \in W$.

Remarque
Un sous-espace vectoriel de V est une partie de V qui soit encore un espace vectoriel pour l'addition et la multiplication scalaire de V. Exercice : écrire une preuve de cela.

Proposition (stabilité des sous-espaces par combinaisons linéaires)
Soient V un espace vectoriel réel et W un sous-ensemble non vide de V. Alors, W est un sous-espace vectoriel de V si, et seulement si W est stable par combinaisons linéaires*, ce qui signifie que pour tous $n \in \mathbb{N}^*$, $v_1, \dots, v_n \in W$ et $x_1, \dots, x_n \in \mathbb{R}$,*

$$x_1 v_1 + \cdots + x_n v_n \in W.$$

Preuve. Si W est stable par combinaisons linéaires, il contient 0_V et est stable pour la somme et la multiplication scalaire qui sont des combinaisons linéaires particulières ; donc W est un sous-espace vectoriel de V. Inversement, si W est stable pour $+$ et $\cdot$, soient $v_1, \dots, v_n \in W$ et $x_1, \dots, x_n \in \mathbb{R}$. Alors, tous les $x_k v_k$ sont dans W, et donc leur somme aussi, ce qui prouve que la combinaison linéaire $x_1 v_1 + \cdots + x_n v_n$ est encore dans W — pour écrire une preuve complète, procéder par récurrence sur n. ■

Exemples

(i) Si V est un $\mathbb{R}$-espace vectoriel, $\{0_V\}$ et V en sont deux sous-espaces vectoriels.

(ii) Si V est un $\mathbb{R}$-espace vectoriel et si $v \in \mathbb{V}$, l'ensemble $\{xv,\ x \in \mathbb{R}\}$ des vecteurs *colinéaires à v* est un sous-espace vectoriel de V. Si $v \neq 0_V$, ce sous-espace est la *droite engendrée par v*.

(iii) Plus généralement, si V est un $\mathbb{R}$-espace vectoriel et si $v_1, \dots, v_n \in V$, l'ensemble des combinaisons linéaires de $v_1, \dots, v_n$ est un sous-espace vectoriel de V, que l'on notera en abrégé

$$\sum_{k=1}^{n} \mathbb{R} v_k = \mathbb{R} v_1 + \cdots + \mathbb{R} v_n.$$

(iv) L'ensemble des applications continues $\mathbb{R} \to \mathbb{R}$ est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{R}}$. C'est vrai aussi de l'ensemble des applications $\mathbb{R} \to \mathbb{R}$ dérivables, mais pas de l'ensemble des applications $\mathbb{R} \to \mathbb{R}$ croissantes.

Proposition (intersection de sous-espaces)
Soit $(V_i)_{i \in I}$ une famille non vide de sous-espaces vectoriels d'un espace vectoriel réel V. Alors, l'intersection $\bigcap_{i \in I} V_i$ est encore un sous-espace vectoriel de V.

Preuve. On note $\mathcal{I}$ l'intersection des V_i, $i \in I$. Puisque 0_V est dans tous les V_i, il est dans $\mathcal{I}$. Soient $v, w \in \mathcal{I}$ et $x \in \mathbb{R}$. Pour chaque $i \in I$, $v, w \in V_i$ et donc $v + w \in V_i$ et $xv \in V_i$ puisque V_i est un sous-espace. Comme cela est vrai pour tous les i, on a montré que $v + w \in \mathcal{I}$ et que $xv \in \mathcal{I}$. ■

Définition (sous-espace vectoriel engendré par une partie)
Soient V un espace vectoriel réel et E un sous-ensemble de V. Le *sous-espace vectoriel engendré par E* est l'intersection de tous les sous-espaces vectoriels de V qui contiennent E. On le note $\mathrm{Vect}(E)$ ou $\mathrm{Vect}\, E$.

Remarques

(i) Avec ces notations, la proposition sur l'intersection de sous-espaces assure que $\mathrm{Vect}\, E$ est un sous-espace vectoriel de V, ce qui rend l'appellation légitime.

(ii) Si V est un espace vectoriel réel, le sous-espace vectoriel engendré par l'ensemble vide est le sous-espace nul : $\mathrm{Vect}\,(\emptyset) = \{0_V\}$.

Proposition (minimalité des sous-espaces engendrés)
Soient V un espace vectoriel réel et E un sous-ensemble de V. Alors, $\operatorname{Vect} E$ *est le plus petit sous-espace vectoriel de V qui contient E, au sens de l'inclusion. Autrement dit, pour tout sous-espace vectoriel W de V,*

$$\boxed{E \subseteq W \Longrightarrow \operatorname{Vect}(E) \subseteq W}$$

Preuve. Si W est un sous-espace de V qui contient E, il appartient à l'ensemble des sous-espaces dont on prend l'intersection pour définir $\operatorname{Vect} E$. Donc il contient $\operatorname{Vect} E$. ∎

Proposition (caractérisation des sous-espaces engendrés)
Soient V un espace vectoriel réel et E une partie non vide de V. Alors, $\operatorname{Vect} E$ *est l'ensemble des combinaisons linéaires des éléments de E.*

Autrement dit, $\operatorname{Vect} E = \left\{ \sum_{1 \leq k \leq n} x_k v_k,\ n \in \mathbb{N}^*,\ v_1, \ldots, v_n \in E,\ x_1, \ldots x_n \in \mathbb{R} \right\}$.

Preuve. On montre d'abord que l'ensemble $\mathcal{E}$ des combinaisons linéaires des éléments de E est un sous-espace vectoriel de V (facile), qui contient E (facile aussi). Donc $\mathcal{E}$ contient $\operatorname{Vect} E$, par minimalité des sous-espaces engendrés. Inversement, puisque $\operatorname{Vect} E$ contient E et est stable par combinaisons linéaires (c'est un sous-espace), il contient les combinaisons linéaires des éléments de E ; donc $\operatorname{Vect} E$ contient $\mathcal{E}$. ∎

Exemples

(i) Dans n'importe quel espace vectoriel réel, $\operatorname{Vect}\{v\} = \{xv,\ x \in \mathbb{R}\} = \mathbb{R}v$.

(ii) (a) Soit $P = \left\{(x, y, z) \in \mathbb{R}^3,\ x - y + 2z = 0\right\}$. Alors, P est un sous-espace vectoriel de $\mathbb{R}^3$ (exercice).

(b) Dans $\mathbb{R}^3$, soient $v = (1, 1, 0)$ et $w = (-2, 0, 1)$ et P le sous-espace de l'exemple (a). Alors, $v \in P$ et $w \in P$ — il suffit de faire le calcul. On en déduit que $\operatorname{Vect}\{v, w\} \subseteq P$.

(c) Avec les notations précédentes, soit $(x, y, z) \in P$. Alors, $x = y - 2z$ et on peut écrire

$$(x, y, z) = (y - 2z, y, z) = y(1, 1, 0) + z(-2, 0, 1) = yv + zw$$

ce qui montre que $(x, y, z) \in \operatorname{Vect}\{v, w\}$: on a montré que $P \subseteq \operatorname{Vect}\{v, w\}$.

(d) Conclusion : on a montré que $P = \operatorname{Vect}\{v, w\}$, ce qui s'écrit encore sous les formes suivantes, dont il importe de prendre l'habitude en passant couramment de l'une à l'autre.

$$\begin{aligned} P &= \left\{(x, y, z) \in \mathbb{R}^3,\ x - y + 2z = 0\right\} \\ &= \operatorname{Vect}\{(1, 1, 0), (-2, 0, 1)\} \\ &= \{s(1, 1, 0) + t(-2, 0, 1),\ s, t \in \mathbb{R}\} \\ &= \mathbb{R}(1, 1, 0) + \mathbb{R}(-2, 0, 1). \end{aligned}$$

(iii) Dans $\mathcal{M}_{2,1}(\mathbb{R})$, soient $a = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $c = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ et $v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Alors, $v = 2a + b + 0 \cdot c$, mais aussi $v = 3a + 0 \cdot b + c$. Ainsi, on a écrit v comme deux combinaisons linéaires différentes de a, b et c : en général, une écriture comme combinaison linéaire n'est pas unique.

(iv) Dans $\mathcal{M}_{3,1}(\mathbb{R})$, toute combinaison linéaire de $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ et $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ a une troisième coordonné nulle. Par exemple, $\begin{pmatrix} \sqrt{3} \\ -1 \\ 1 \end{pmatrix}$ n'est pas dans le sous-espace $\operatorname{Vect}\{e_1, e_2\}$.

(v) Pour tout $t \in \mathbb{R}$, $\cos\left(t + \frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}(\cos t - \sin t)$. Cette formule trigonométrique montre que dans l'espace vectoriel des applications $\mathbb{R} \to \mathbb{R}$, l'application $t \mapsto \cos\left(t + \frac{\pi}{4}\right)$ est dans le sous-espace engendré par les fonctions cosinus et sinus.

Proposition (somme de sous-espaces)
Soient V un espace vectoriel réel, W_1 et W_2 des sous-espaces vectoriels de V. Alors

$$\mathrm{Vect}\,(W_1 \cup W_2) = W_1 + W_2$$

où

$$W_1 + W_2 = \{w_1 + w_2,\ w_1 \in W_1,\ w_2 \in W_2\}\,.$$

Preuve. Puisque W_1 et W_2 sont des sous-espaces, la somme $W_1 + W_2$ est aussi l'ensemble des combinaisons linéaires des éléments de $W_1 \cup W_2$. ■

A noter
De la même façon, si $W_1, \ldots, W_n$ sont des sous-espaces vectoriels d'un même espace vectoriel réel V, alors

$$\mathrm{Vect}\left(\bigcup_{k=1}^{n} W_k\right) = \sum_{k=1}^{n} W_k = \left\{\sum_{k=1}^{n} w_k,\ \forall k \in \{1, \ldots, n\},\ w_k \in W_k\right\}.$$

Preuve : exercice.

2.3 Familles génératrices, familles libres

Définition (famille génératrice)
Soient V un espace vectoriel réel, W un sous-espace vectoriel de V et E un sous-ensemble de V. On dit que les vecteurs de E forment une *famille génératrice de W* — ou encore une *partie génératrice de W* — lorsque $W = \mathrm{Vect}(E)$. Lorsque $\mathrm{Vect}(E) = V$ (tout entier), on dit simplement que les vecteurs de E forment une *famille génératrice* (tout court, sans préciser de quoi).

Exemples

(i) Si V est un espace vectoriel réel, l'ensemble de tous les vecteurs de V forme une famille génératrice de V (!).

(ii) Dans $\mathbb{R}^3$, on note $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ et $e_3 = (0,0,1)$.
Comme dans l'exemple (iii) ci-dessus, $\{e_1, e_2\}$ n'est pas une famille génératrice (de $\mathbb{R}^3$) puisque tout combinaison linéaire de e_1 et e_2 a une troisième coordonnée nulle. En revanche, pour tout $(x,y,z) \in \mathbb{R}^3$, la formule

$$(x,y,z) = x(1,0,0) + y(0,1,0) + z(0,0,1)$$

montre que $\{e_1, e_2, e_3\}$ est une famille génératrice (de $\mathbb{R}^3$).

(iii) Si $\{v_1, \ldots, v_n\}$ est une famille génératrice, alors pour tout v, la famille $\{v, v_1, \ldots, v_n\}$ est encore génératrice. Plus généralement, *toute sur-famille d'une famille génératrice est encore génératrice.*

(iv) Le $\mathbb{R}$-espace vectoriel $\mathbb{C}$ est engendré par $\{1, i\}$, autre façon de dire que tout nombre complexe s'écrit sous la forme $a + ib$ où a et b sont des réels.

(v) Lorsqu'un espace vectoriel admet une partie génératrice finie, on dit qu'il est *finiment engendré.*

Trouver une famille génératrice est une manière de décrire un sous-espace vectoriel. Par souci d'économie — et d'unicité de l'écriture, dans le cas des espaces vectoriels, on le verra —, on est amené à rechercher une famille génératrice "la plus petite possible". La suite du chapitre apporte une réponse complète à cette démarche.

Définition (relation linéaire entre des vecteurs)
Soient V un espace vectoriel réel, $n \in \mathbb{N}^*$ et $v_1, \ldots, v_n \in V$. Une *relation linéaire entre les v_k* est un n-uplet $(x_1, \ldots, x_n) \in \mathbb{R}^n$ tel que $x_1 v_1 + \cdots + x_n v_n = 0_V$. Lorsqu'une relation linéaire $(x_1, \ldots, x_n)$ est non nulle, c'est-à-dire lorsque $(x_1, \ldots, x_n) \neq (0, \ldots, 0)$, on dit que $(x_1, \ldots, x_n)$ est une *relation linéaire non triviale entre les v_k.*

A noter
Avec les mêmes notations, le n-uplet $(0, \ldots, 0)$ est toujours une relation linéaire entre les vecteurs $v_1, \ldots, v_n$. On l'appelle *relation linéaire triviale.*

Définition (dépendance et indépendance linéaire)
Soient V un espace vectoriel réel, $n \in \mathbb{N}^*$ et $v_1, \ldots, v_n \in V$. On dit que les vecteurs $v_1, \ldots, v_n \in V$ sont *linéairement dépendants*, ou encore qu'ils forment une *famille liée* lorsqu'il existe une relation linéaire non triviale entre eux. Sinon, on dit qu'ils sont *linéairement indépendants*, ou encore qu'ils forment une *famille libre*.

Autrement dit,

$$v_1, \ldots, v_n \text{ sont linéairement dépendants} \iff \exists\,(x_1, \ldots, x_n) \in \mathbb{R}^n \setminus \{(0, \ldots, 0)\}\,,\ \sum_{k=1}^{n} x_k v_k = 0_V$$

et

$$v_1, \ldots, v_n \text{ sont linéairement indépendants} \Leftrightarrow \forall\,(x_1, \ldots, x_n) \in \mathbb{R}^n, \left(\sum_{k=1}^{n} x_k v_k = 0_V \Rightarrow (x_1, \ldots, x_n) = (0, \ldots, 0)\right)$$

A noter

(i) Par convention, parce que cela nous sera utile, on dira que $\emptyset$ est une famille libre.

(ii) Plus généralement, si $(v_i)_{i \in I}$ est une famille quelconque de vecteurs de V, une relation linéaire entre les v_i est, par définition, une famille *presque nulle* $(x_i)_{i \in I}$ de réels tels que

$$\sum_{i \in I} x_i v_i = 0_V.$$

Noter que cette somme est finie puisque seuls les x_i d'un sous-ensemble fini de I sont non nuls. Par définition toujours, la famille $(v_i)_{i \in I}$ est libre lorsque toute relation linéaire entre les v_i est triviale — la relation linéaire triviale indexée par I est la famille dont tous les termes sont nuls. Toujours par définition, la famille $(v_i)_{i \in I}$ est *liée* lorsqu'elle n'est pas libre.

Proposition (unicité de l'écriture d'une combinaison linéaire entre vecteurs indépendants)
Soient V un espace vectoriel réel, $n \in \mathbb{N}^*$ et $\{v_1, \ldots, v_n\}$ une famille libre de vecteurs de V. Alors, tout vecteur de $\mathrm{Vect}\,\{v_1, \ldots, v_n\}$ *s'écrit* de manière unique *sous la forme d'une combinaison linéaire de $v_1, \ldots, v_n$.*

Autrement dit, avec les notations de la proposition, si $\{v_1, \ldots, v_n\}$ est une famille libre, alors pour tous $x_1, \ldots, x_n, x'_1, \ldots, x'_n \in \mathbb{R}$,

$$\left(\sum_{k=1}^{n} x_k v_k = \sum_{k=1}^{n} x'_k v_k\right) \Longrightarrow \left(\forall k \in \{1, \ldots, n\}\,,\ x_k = x'_k\right).$$

Preuve. Si $\sum_k x_k v_k = \sum_k x'_k v_k$, alors $\sum_k (x_k - x'_k)\, v_k$ est une relation linéaire entre les v_k — on comprendra que cette façon de parler, communément admise et à laquelle il faut s'habituer, signifie que $(x_k - x'_k)_{1 \leq k \leq n}$ est une relation linéaire entre les v_k. Elle est triviale puisque les v_k sont linéairement indépendants. ∎

Exemples

(i) Si v est un vecteur de V, la famille $\{v\}$ est libre si, et seulement si $v \neq 0_V$.

(ii) Toute famille contenant le vecteur nul est liée.

(iii) Soient $u, v \in V$. Si $u = 0_V$ ou $v = 0_V$, alors la famille $\{u, v\}$ est liée. Si en revanche $u \neq 0_V$ et $v \neq 0_V$, la famille $\{u, v\}$ est liée si, et seulement s'il existe $x \in \mathbb{R} \setminus \{0\}$ tel que $u = xv$, c'est-à-dire si, et seulement si $\mathrm{Vect}\,\{u\} = \mathrm{Vect}\,\{v\}$.

(iv) Soient $u_1, \ldots, u_n \in V$. Si $v \in \mathrm{Vect}\,\{u_1, \ldots, u_n\}$, alors la famille $\{v, u_1, \ldots, u_n\}$ est liée.
En effet, écrire v comme combinaison linéaire des u_k fournit une relation linéaire non triviale entre $v, u_1, \ldots, u_n$ dans laquelle le coefficient de v est 1.

(v) Si $n \geq 1$ et si $\{v_1, \dots, v_n\}$ est une famille libre, alors la famille $\{v_1, \dots, v_{n-1}\}$ est encore libre. Plus généralement, *toute sous-famille d'une famille libre est encore libre.*

(vi) Dans l'espace vectoriel des applications $\mathbb{R} \to \mathbb{R}$, on note cos la fonction cosinus, sin la fonction sinus, 1 la fonction constante égale à 1 et 0 la fonction constante égale à 0 (c'est le neutre de $\mathbb{R}^{\mathbb{R}}$). Alors, les vecteurs cos, sin et 1 sont linéairement indépendants.

En effet, supposons que $a, b, c \in \mathbb{R}$ vérifient $a \cdot \cos + b \cdot \sin + c \cdot 1 = 0$. Alors, pour tout $x \in \mathbb{R}$, $a \cos x + b \sin x + c = 0$. Cette égalité est vraie pour $x = 0$, ce qui fournit la relation $a + c = 0$. En écrivant encore l'égalité pour $x = \pi$ et $x = \frac{\pi}{2}$, on obtient également $-a + c = 0$ et $b + c = 0$. On résout aisément ce système linéaire en (a, b, c), qui admet $(0, 0, 0)$ pour unique solution.

(vii) [*Où l'on voit que démontrer une indépendance linéaire de vecteurs revient à résoudre un système linéaire homogène.*]
Dans $\mathcal{M}_{3,1}(\mathbb{R})$, on note

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{et} \quad v_2' = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

On cherche à savoir si les familles $\{v_1, v_2, v_3\}$ et $\{v_1, v_2', v_3\}$ sont libres ou liées.

On commence par la première. On suppose que $(x_1, x_2, x_3) \in \mathbb{R}^3$ vérifient $x_1 v_1 + x_2 v_2 + x_3 v_3 = 0$ (on a noté 0 le vecteur nul de $\mathcal{M}_{3,1}(\mathbb{R})$). Cela équivaut au système linéaire homogène

$$\begin{cases} x_1 + x_3 = 0 \\ 2x_1 + x_2 = 0 \\ 3x_1 + 2x_2 + x_3 = 0 \end{cases}$$

que l'on résout par substitutions (écrire par exemple x_3 et x_2 en fonction de x_1 à l'aide des deux premières équations et reporter dans la troisième) : son unique solution est le triplet $(x_1, x_2, x_3) = (0, 0, 0)$. Autrement dit, la famille (v_1, v_2, v_3) est libre.

Ensuite, la seconde. On suppose que $(x_1, x_2, x_3) \in \mathbb{R}^3$ vérifient $x_1 v_1 + x_2 v_2' + x_3 v_3 = 0$, ce qui équivaut au système linéaire homogène

$$\begin{cases} x_1 + x_3 = 0 \\ 2x_1 + x_2 = 0 \\ 3x_1 + x_2 + x_3 = 0. \end{cases}$$

On le résout : l'ensemble des solutions du système est l'ensemble des triplets (x_1, x_2, x_3) tels que $x_2 - 2x_3 = 0$ et $x_1 + x_3 = 0$. Autrement dit, $x_1 v_1 + x_2 v_2 + x_3 v_3 = 0$ si, et seulement si $(x_1, x_2, x_3) = (-x_3, 2x_3, x_3)$ est proportionnel à $(-1, 2, 1)$. Par exemple, $-v_1 + 2v_2 + v_3 = 0$ fournit une relation linéaire non triviale entre v_1, v_2' et v_3 : la famille (v_1, v_2', v_3) est liée.

Lemme ("trop" de vecteurs sont forcément liés)
Soient $p, q \in \mathbb{N}^*$, V un espace vectoriel réel, $u_1, \dots, u_p \in V$, $v_1, \dots v_q \in V$. On suppose que
(i) $v_1, \dots v_q \in \mathrm{Vect}\{u_1, \dots, u_p\}$
(ii) $q \geq p + 1$.
Alors, la famille $\{v_1, \dots v_q\}$ est liée.

Preuve. On procède par récurrence sur p. Il suffit de prouver le résultat lorsque $q = p+1$, ce que l'on suppose. Si $p = 1$, les vecteurs v_1 et v_2 sont colinéaires à u_1. Soient donc $x_1, x_2 \in \mathbb{R}$ tels que $v_1 = x_1 u_1$ et $v_2 = x_2 u_1$. Si $x_1 = 0$, alors $v_1 = 0_V$ et la famille $\{v_1, v_2\}$ est liée. Si au contraire $x_1 \neq 0$, alors $u_1 = \frac{1}{x_1} v_1$ et $v_2 - \frac{x_2}{x_1} v_1 = 0_V$, ce qui montre que la famille $\{v_1, v_2\}$ est encore liée.
On suppose $p \geq 2$ et on note $W = \mathrm{Vect}\{u_1, \dots, u_{p-1}\}$. En écrivant les v_k comme combinaisons linéaires de $u_1, \dots, u_p$ et en isolant u_p, on obtient $x_1, \dots, x_{p+1} \in \mathbb{R}$ et $w_1, \dots, w_{p+1} \in W$ tels que

$$\begin{cases} v_1 - x_1 u_p = w_1 \\ v_2 - x_2 u_p = w_2 \\ \qquad \vdots \\ v_p - x_p u_p = w_p \\ v_{p+1} - x_{p+1} u_p = w_{p+1}. \end{cases}$$

Si tous les x_k sont nuls, alors $v_1, \dots, v_{p+1} \in W$ et l'hypothèse de récurrence assure que $\{v_1, \dots, v_{p+1}\}$ est liée. Sinon, quitte à ré-ordonner les v_k, on suppose que $x_{p+1} \neq 0$. La dernière ligne du système s'écrit alors $u_p = \frac{1}{x_{p+1}}(v_{p+1} - w_{p+1})$. En reportant dans les p premières égalités du système, on obtient $w'_1, \dots, w'_p \in W$ tels que

$$\begin{cases} v_1 - \frac{x_1}{x_{p+1}} v_{p+1} = w'_1 \\ v_2 - \frac{x_2}{x_{p+1}} v_{p+1} = w'_2 \\ \qquad \vdots \\ v_p - \frac{x_q}{x_{p+1}} v_{p+1} = w'_p. \end{cases}$$

Alors, puisque W admet une famille génératrice de $p-1$ vecteurs (au plus), l'hypothèse de récurrence assure qu'il existe $(y_1, \dots, y_p) \in \mathbb{R}^p$, $(y_1, \dots, y_p) \neq (0, \dots, 0)$, tel que

$$\sum_{k=1}^{p} y_k \left(v_k - \frac{x_k}{x_{p+1}} v_{p+1} \right) = 0_V.$$

Cela fournit une relation linéaire non triviale entre les vecteurs $v_1, \dots, v_{p+1}$. ■

Corollaire (une famille libre a moins d'éléments qu'une famille génératrice)
Soit V un espace vectoriel réel finiment engendré et $\ell, g \in \mathbb{N}$. On suppose que $\{u_1, \dots u_\ell\}$ est une famille libre et que $\{v_1, \dots v_g\}$ est une famille génératrice (de V). Alors, $\ell \leq g$.

Preuve. C'est une conséquence directe du lemme. ■

Exemple
Dans $\mathcal{M}_{2,1}(\mathbb{R})$, tout vecteur est combinaison linéaire de $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et de $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Le lemme assure que trois vecteurs de $\mathcal{M}_{2,1}(\mathbb{R})$ sont donc toujours linéairement dépendants. Par exemple, les vecteurs $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et de $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ sont liés (il est très facile de trouver entre eux une relation linéaire non triviale, par ailleurs), même s'ils sont *deux à deux* linéairement indépendants.

2.4 Bases d'un espace vectoriel

Définition (base)
Soient V un espace vectoriel réel, $n \geq 1$ et $v_1, \dots, v_n \in V$. On dit que le n-uplet $(v_1, \dots, v_n) \in V^n$ est une *base de V* lorsque la famille $\{v_1, \dots, v_n\}$ est à la fois libre et génératrice.

A noter
Avec les notations de la définition, une base est un n-uplet. En particulier, l'ordre d'écriture des vecteurs d'une base compte. Par exemple, si (v, w) est une base d'un espace vectoriel, alors (w, v) en est une autre, distincte de la première (exercice).

Proposition (écriture des vecteurs dans une base)
Soient V un espace vectoriel réel, $n \geq 1$ et $(v_1, \dots, v_n) \in V^n$. Alors, les assertions suivantes sont équivalentes :

(i) $(v_1, \dots, v_n)$ une base de V ;

(ii) tout vecteur $v \in V$ s'écrit, de manière unique, sous la forme

$$v = \sum_{k=1}^{n} x_k v_k$$

où $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Preuve. La famille $\{v_1, \dots, v_n\}$ est génératrice si, et seulement si tout vecteur de V est combinaison linéaire des v_k. Elle est libre si, et seulement si l'écriture est unique. ■

Définition (coordonnées d'un vecteur dans une base)
Dans la situation de la proposition précédente, le n-uplet de réels $(x_1, \ldots, x_n)$ est le n-uplet des *coordonnées de v*.

A noter
(i) Une paraphrase importante de la proposition précédente est la suivante (on reviendra dessus) : si $(v_1, \ldots, v_n)$ une base de V, l'application

$$\begin{array}{ccc} \mathbb{R}^n & \longrightarrow & V \\ (x_1, \ldots, x_n) & \longmapsto & \displaystyle\sum_{k=1}^{n} x_k v_k \end{array}$$

est une bijection de $\mathbb{R}^n$ sur V.

(ii) Plus généralement, une famille quelconque $(v_i)_{i \in I}$ est une base de V lorsqu'elle est à la fois libre et génératrice. Tout vecteur de V s'écrit alors encore, de façon unique, comme combinaison linéaire des v_i.

A ce stade, la question de l'existence de bases reste entière. On y répond complètement dans le cas des espaces vectoriels finiment engendrés, même si cette dernière hypothèse peut être levée en toute généralité. Le théorème qui suit est important dans le sens où il fonde la théorie de la dimension, et parce qu'il s'avère extrêmement opératoire.

Théorème (de la base incomplète)
Soient V un espace vectoriel réel, $\ell \in \mathbb{N}^*$, $v_1, \ldots, v_\ell \in V$ et G une partie génératrice finie de V. On suppose que

(i) la famille $\{v_1, \ldots, v_\ell\}$ est libre

(ii) $\{v_1, \ldots, v_\ell\} \subseteq G$.

Alors, il existe un entier $d \geq \ell$ et $v_{\ell+1}, \ldots, v_d \in G$ tels que $(v_1, \ldots, v_d)$ soit une base de V.

Le slogan : *dans un espace vectoriel, toute famille libre se complète en une base.*

Preuve. L'ensemble des parties de G qui contiennent $\{v_1, \ldots, v_\ell\}$ et qui sont formées de vecteurs linéairement indépendants n'est pas vide puisque $\{v_1, \ldots, v_\ell\}$ est un de ses éléments. En outre, toutes ces parties ont un cardinal fini. Il en existe donc une famille libre de G contenant $\{v_1, \ldots, v_\ell\}$, dont le nombre d'éléments est maximum. On en choisit une comme cela et on la nomme $\{v_1, \ldots, v_d\}$, $d \geq \ell$. Il s'agit de montrer que $(v_1, \ldots, v_d)$ est une base de V. Par définition, la famille $\{v_1, \ldots, v_d\}$ est libre. Il reste dont à montrer qu'elle est génératrice. Puisque tout élément de V est combinaison linéaire d'éléments de G, il suffit de montrer que tout élément de $G \setminus \{v_1, \ldots, v_d\}$ est dans $\operatorname{Vect}\{v_1, \ldots, v_d\}$. Soit donc $v \in G \setminus \{v_1, \ldots, v_d\}$. Par minimalité de d, la famille $\{v, v_1, \ldots, v_d\}$, qui contient $d+1$ éléments, est nécessairement liée. Soit donc $xv + x_1v_1 + \cdots + x_dv_d = 0_V$ une relation linéaire non triviale. Puisque la famille $\{v_1, \ldots, v_d\}$ est libre, $x \neq 0$. Donc $v = \frac{-1}{x}(x_1v_1 + \cdots + x_dv_d) \in \operatorname{Vect}\{v_1, \ldots, v_d\}$. ∎

Corollaire (existence de bases)
Tout espace vectoriel réel finiment engendré et non nul admet au moins une base.

Preuve. On note G un système générateur fini. Puisque V est non nul, G contient au moins un vecteur non nul v_1. La famille $\{v_1\}$ est alors libre et on applique le théorème de la base incomplète avec $\ell = 1$. ∎

A noter

(i) Le théorème de la base incomplète est encore vrai dans les espaces vectoriels qui ne sont pas finiment engendrés, à condition d'admettre ce que l'on appelle l'*axiome du choix*. Il prend alors la forme suivante : si L est une famille libre et si G est une famille génératrice qui contient L, alors il existe B telle que $L \subseteq B \subseteq G$ et telle que les éléments de B forment une base. De la même façon, conséquemment, tout espace vectoriel non nul admet au moins une base.

(ii) On aurait pu faire un autre choix de présentation de l'existence de bases en démontrant qu'une base est une *famille génératrice minimale* (on ne peut plus enlever de vecteurs à une telle famille sans en perdre le caractère générateur), ou encore une *famille libre maximale* (on ne peut plus ajouter de vecteurs à une telle famille sans en perdre le caractère libre). Ces caractérisations des bases peuvent être démontrées à titre d'exercice et retenues comme des résultats du cours.

Exemples

(i) Dans $\mathbb{R}^n$, les vecteurs

$$(1,0,0,\ldots,0),(0,1,0,\ldots 0),\ldots,(0,0,\ldots,0,1)$$

forment, dans cet ordre, une base que l'on nomme *base canonique* de $\mathbb{R}^n$.

(ii) De façon analogue, les vecteurs

$$\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\\\vdots\\0\end{pmatrix},\ldots,\begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix}$$

forment, dans cet ordre, la *base canonique* de $\mathcal{M}_{n,1}(\mathbb{R})$.

(iii) De façon plus générale, si $p,q\in\mathbb{N}^*$, on appelle *base canonique* de $\mathcal{M}_{p,q}(\mathbb{R})$ la famille des matrices

$$E_{\ell,c}=(\delta_{i,c}\delta_{\ell,j})_{1\leq i\leq p,1\leq j\leq q}$$

où δ désigne le symbole de Kronecker (défini par $\delta_{i,j}=0$ si $i\neq j$ et $\delta_{i,j}=1$ si $i=j$). Autrement dit, $E_{\ell,c}$ est la matrice dont tous les coefficients sont nuls, à l'exception du coefficient de la $\ell^{ième}$ ligne et de la $c^{ième}$ colonne, qui vaut 1.

[Pour le choix de l'ordre dans la base canonique de $\mathcal{M}_{p,q}(\mathbb{R})$, on prend souvent l'ordre alphabétique sur $\{1,\ldots,p\}\times\{1,\ldots,q\}$.]

(iv) Un résultat d'analyse élémentaire assure que si $f:\mathbb{R}\to\mathbb{R}$ vérifie $f''+f=0$, alors il existe un unique couple (a,b) tel que, pour tout $t\in\mathbb{R}$, $f(t)=a\cos t+b\sin t$. Autrement dit, le couple $(\cos,\sin)$ est une base de l'espace vectoriel des solutions de l'équation différentielle $y''+y=0$.

[Exercice : montrer directement, sans utiliser ce résultat, que l'ensemble des fonctions $y:\mathbb{R}\to\mathbb{R}$ qui vérifient $y''+y=0$ (qui est une équation différentielle linéaire homogène) est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{R}}$.]

2.5 Dimension d'un espace vectoriel finiment engendré

Théorème (invariance du cardinal des bases)

Dans un espace vectoriel réel finiment engendré, toutes les bases ont le même nombre d'éléments.

Preuve. Si $(u_1,\ldots,u_m)$ et $(v_1,\ldots,v_n)$ sont des bases, alors la famille $\{u_1,\ldots,u_m\}$ est libre et la famille $\{v_1,\ldots,v_n\}$ est génératrice ; donc $m\leq n$, d'après le théorème "une famille libre a moins d'éléments qu'une famille génératrice". Par symétrie, on a aussi $n\leq m$. ∎

Définition (dimension d'un espace vectoriel finiment engendré)

Si V est un espace vectoriel réel finiment engendré, le cardinal commun à toutes ses bases est la *dimension de* V. On la note $\dim_{\mathbb{R}}(V)$, ou simplement $\dim V$ ou $\dim(V)$.

A noter

(i) Dans un espace vectoriel quelconque, toutes les bases ont aussi le même cardinal. C'est ce cardinal commun qu'on appelle alors la dimension de l'espace. Un espace vectoriel finiment engendré est aussi appelé *espace vectoriel de dimension finie.*

(ii) Puisque l'espace nul n'a pas de base, on dit qu'il est de dimension 0. Cette convention est souvent commode dans les énoncés que l'on rencontrera.

Exemples

(i) Pour tout $n\in\mathbb{N}^*$, les espaces $\mathbb{R}^n$ et $\mathcal{M}_{n,1}(\mathbb{R})$ sont de dimension n (on en a exhibé des bases).

(ii) Si $p,q\in\mathbb{N}^*$, l'espace vectoriel $\mathcal{M}_{p,q}(\mathbb{R})$ est de dimension pq (compter le nombre de matrices dans la base canonique). En particulier, l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$ est de dimension n^2.

(iii) Tout nombre complexe s'écrit, de manière unique, sous la forme $a+ib$ où $a,b\in\mathbb{R}$. Autrement dit, $(1,i)$ est une base du $\mathbb{R}$-espace vectoriel $\mathbb{C}$. Ainsi, $\dim_{\mathbb{R}}\mathbb{C}=2$.

(iv) L'ensemble des solutions de l'équation différentielle linéaire d'ordre 2 homogène $y''+y=0$ est un sous-espace vectoriel de dimension 2 de $\mathbb{R}^{\mathbb{R}}$.

Proposition (caractérisation des bases en dimension finie)
Soient V un espace vectoriel réel de dimension finie $d \geq 1$, et $v_1, \ldots, v_d \in V$. Les assertions suivantes sont équivalentes.

(i) $\{v_1, \ldots, v_d\}$ *est libre*

(ii) $\{v_1, \ldots, v_d\}$ *est génératrice*

(iii) $(v_1, \ldots, v_d)$ *est une base.*

Preuve. On suppose que *(i)* est vraie. Par le théorème de la base incomplète, on peut compléter $\{v_1, \ldots, v_d\}$ en une base de V. Comme V est de dimension d, une telle base a d éléments. Donc $(v_1, \ldots, v_d)$ est (déjà) une base. On suppose que *(ii)* est vraie. Tous les vecteurs v_k ne sont pas nuls, puisque $\dim V \geq 1$. Soit donc $m \in \{1, \ldots d\}$ tel que $v_m \neq 0$. Par le théorème de la base incomplète, on peut compléter la famille libre $\{v_m\}$ par des v_k pour obtenir une base de V. Comme V est de dimension d, une telle base a d éléments. Donc on a besoin de tous les v_k pour obtenir une base : $(v_1, \ldots, v_d)$ est une base. ∎

Proposition (dimension d'un sous-espace)
Soient V un espace vectoriel réel de dimension finie, et W un sous-espace vectoriel de V. Alors, W est aussi de dimension finie, et $\dim W \leq \dim V$.

Preuve. C'est le théorème de la base incomplète (compléter une base de W en une base de V). ∎

Proposition (caractérisation de l'égalité de deux sous-espaces)
Soient V un espace vectoriel réel, W_1 et W_2 deux sous-espaces vectoriels de V de dimension finie. Alors,

$$\left.\begin{array}{c} W_1 \subseteq W_2 \\ \text{et} \\ \dim W_1 = \dim W_2 \end{array}\right\} \Longrightarrow W_1 = W_2.$$

Preuve. On complète une base de W_1 en une base de W_2. Comme les dimensions sont égales, la base de W_1 est aussi une base de W_2. ∎

2.6 Somme directe de sous-espaces vectoriels

Définition (somme directe de sous-espaces)
Soient V un espace vectoriel, W_1 et W_2 deux sous-espaces vectoriels de V. On dit que V est *somme directe* de W_1 et W_2 lorsque tout vecteur de V s'écrit, de manière unique, comme somme d'un vecteur de W_1 et d'un vecteur de W_2. On note alors

$$V = W_1 \oplus W_2$$

et on dit, dans ces conditions, que W_1 et W_2 sont des sous-espaces *supplémentaires* de V.
Plus généralement, si $W_1, \ldots, W_n$ sont des sous-espaces de V, on dit que V est *somme directe* de $W_1, \ldots, W_n$ lorsque tout vecteur de V s'écrit, de manière unique, comme somme de vecteurs des W_k, $1 \leq k \leq n$. On note encore

$$V = W_1 \oplus \cdots \oplus W_n = \bigoplus_{k=1}^{n} W_k.$$

A noter
Encore plus généralement, si $(W_i)_{i \in I}$ est une famille non vide de sous-espaces de V, on dit que V est la somme directe des W_i lorsque tout vecteur de V s'écrit, de façon unique, comme une somme

$$\sum_{j \in J} w_j$$

où J est une partie finie de I et où, $w_j \in W_j$ pour tout $j \in J$.

Proposition (somme directe et dimension)
Soient V un espace vectoriel de dimension finie et $W_1, \ldots, W_n$ des sous-espaces de V. Alors,

$$V = \bigoplus_{k=1}^{n} W_k \Longrightarrow \dim V = \sum_{k=1}^{n} \dim (W_k).$$

Preuve. Prendre des bases des W_k et les mettre bout à bout : on obtient une base de V. Cela suffit à démontrer le résultat. ■

Proposition (caractérisation des sommes directes de deux sous-espaces)
Soient V un espace vectoriel de dimension finie, W_1 et W_2 des sous-espaces de V. Les assertions suivantes sont équivalentes :
(i) $V = W_1 \oplus W_2$
(ii) $V = W_1 + W_2$ *et* $W_1 \cap W_2 = \{0\}$
(iii) $\dim(W_1) + \dim(W_2) = \dim(V)$ *et* $W_1 \cap W_2 = \{0\}$

Preuve. (i)⇒(ii) est évident : la somme est garantie par l'existence de la décomposition sous la forme $w_1 + w_2$, l'intersection l'est par l'unicité de cette écriture. (ii)⇒(iii) est aussi évident grâce à la proposition précédente. (iii)⇒(i) On suppose que $\dim(W_1) + \dim(W_2) = \dim(V)$ et $W_1 \cap W_2 = \{0\}$. On prend une base $(w_1, \ldots, w_{d_1})$ de W_1 et une base $(w_{d_1+1}, \ldots, w_{d_1+d_2})$ de W_2. Alors, $\dim V = d_1 + d_2$ et la famille $\{w_1, \ldots, w_{d_1+d_2}\}$ est libre à cause de l'hypothèse sur l'intersection. On en déduit que $(w_1, \ldots, w_{d_1+d_2})$ est une base de V, ce qui permet de conclure. ■

2.7 Les attendus du chapitre

Le lire en détail, s'assurer de tout comprendre y compris les exemples, sans laisser de zone d'ombre. Faire, pour soi, une rédaction d'une résolution des exercices énoncés.

Section 2.1
Apprendre soigneusement les huit axiomes d'un espace vectoriel et les considérer commes des règles du jeu exclusives des opérations dans un espace vectoriel. Comprendre en quoi l'associativité de l'addition (axiome (i)) permet de ne pas écrire les parenthèses d'une somme, ou d'une combinaison linéaire.
Assimiler les règles de calcul et la proposition de nullité d'un produit externe ; elles ont l'air évidente et un peu bébêtes, mais les travailler jusqu'à se rendre compte que ce n'est pas le cas et qu'elles sont bien des conséquences des axiomes.

Apprendre la définition d'une combinaison linéaire.

S'approprier les exemples d'espaces vectoriels ($\mathbb{R}^n$, matrices, applications $E \to \mathbb{R}$, suites de nombres réels). Les avoir *toujours* à l'esprit comme exemples fondamentaux, sur lesquels on peut faire fonctionner *tous* les énoncés abstraits pour leur donner du sens dès la première approche.

Apprendre les définitions de la somme de deux matrices de mêmes dimensions et du produit d'une matrice par un nombre.

Apprendre la définition d'un espace vectoriel produit.

Section 2.2
Apprendre la définition d'un sous-espace vectoriel, et sa caractérisation par stabilité par combinaisons linéaires. Bien comprendre les premiers exemples, retenir ce que l'on appelle une droite vectorielle.

La notion de sous-espace vectoriel engendré par une partie est à la fois délicate et fondamentale. S'assurer de bien en avoir appris la définition, et de bien comprendre et retenir les deux autres aspects opératoires : la proposition de minimalité et la caractérisation. Lire les exemples, bien les comprendre, sans aucune zone d'ombre.

Apprendre ce qu'est la somme de deux sous-espaces vectoriels, retenir que c'est le sous-espace engendré par la leur réunion.

Section 2.3
Apprendre la définition d'une famille génératrice, la relier à la notion de sous-espace engendré. Comprendre et retenir qu'une sur-famille d'une famille génératrice est encore génératrice.

Apprendre avec beaucoup de soin ce que signifient la dépendance et l'indépendance linéaires. Retenir minutieusement les deux encadrés est fondamental.

Bien comprendre en quoi l'indépendance linéaire est synonyme d'unicité d'écriture comme combinaison linéaire. Comprendre et retenir qu'une sous-famille d'une famille libre est encore libre (et qu'une sur-famille d'une famille liée est encore liée).

A la lecture des exemples, bien comprendre en quoi tester une indépendance linéaire ou chercher à savoir si une famille est génératrice revient à résoudre des systèmes linéaires.

Bien comprendre et retenir le lemme *"trop" de vecteurs sont forcément liés.* Lire et comprendre la preuve avec soin, savoir en décrire les grandes lignes.

Section 2.4
Apprendre la définition d'une base, notion fondamentale du chapitre.

Apprendre ce que sont les *coordonnées* d'une vecteur dans une base, en s'assurant de bien avoir compris en quoi le caractère générateur et l'indépendance linéaire sont nécessaires à leurs définitions.

Retenir l'énoncé et les grandes lignes de la preuve du théorème de la base incomplète. Ce résultat est très opératoire et sert souvent.

Apprendre ce que sont les bases canoniques des espaces de matrices, en particulier des espaces de vecteurs-ligne ($\mathcal{M}_{1,n}(\mathbb{R}) = \mathbb{R}^n$) et de vecteurs-colonne ($\mathcal{M}_{n,1}(\mathbb{R})$).

Section 2.5
Apprendre avec soin ce qu'est la dimension d'un espace vectoriel, en sachant relier cette définition au théorème d'invariance du cardinal des bases.

Retenir la dimension des espaces de matrices et savoir argumenter leur calcul.

Apprendre et retenir les deux dernières propositions avec leurs preuves. La dernière, notamment, est extrêmement opératoire et très pratique.

Section 2.6
Apprendre et retenir la définition de la somme directe de sous-espaces, en ne laissant aucune zone d'ombre dans le cas de deux sous-espaces. Ne pas négliger cette fin de chapitre : la notion est fondamentale ; elle est en particulier sans cesse utilisée dans l'étude de la réduction des endomorphismes (ou des matrices carrées).

Savoir utiliser les caractérisations d'une somme directe, en mesurant à quel point les considérations sur les dimensions simplifient le mode opératoire.

3 Applications linéaires

Les applications linéaires sont celles qui préservent la structure entre deux espaces vectoriels.

3.1 Définitions, généralités, premiers exemples

Définition (application linéaire)
Soient V et W deux espaces vectoriels réels, et $f : V \to W$ une application. On dit que f est *linéaire* lorsque
(i) $\forall v, w \in V,\ f(v+w) = f(v) + f(w)$;
(ii) $\forall x \in \mathbb{R}, \forall v \in V,\ f(x \cdot v) = x \cdot f(v)$.

Une application linéaire $V \to W$ est aussi appelée *homomorphisme (ou morphisme) d'espaces vectoriels*. Lorsque $V = W$, on parle d'*endomorphisme*. L'ensemble des applications linéaires $V \to W$ et l'ensemble des endomorphismes de V sont respectivement notés

$$\mathrm{Hom}(V, W) \quad \text{et} \quad \mathrm{End}(V).$$

Proposition (les applications linéaires sont celles qui préservent les combinaisons linéaires)
Soient V et W deux espaces vectoriels réels, et $f : V \to W$ une application. Alors, f est linéaire si, et seulement si pour tout $n \in \mathbb{N}^*$, pour tous $v_1, \ldots, v_n \in V$, pour tous $x_1, \ldots, x_n \in \mathbb{R}$,

$$f\left(\sum_{k=1}^{n} x_k v_k\right) = \sum_{k=1}^{n} x_k f\left(v_k\right).$$

Preuve. ($\Leftarrow$) est immédiat : la somme et la multiplication scalaire sont des combinaisons linéaires particulières. ($\Rightarrow$) se fait par récurrence sur n. ■

A noter
Si $f : V \to W$ est linéaire, alors $f\left(0_V\right) = 0_W$.
[En effet, $f\left(0_V\right) = f\left(0_V + 0_V\right) = f\left(0_V\right) + f\left(0_V\right)$. En ajoutant $-f\left(0_V\right)$ aux deux membres de l'égalité, il vient $0_W = f\left(0_V\right)$.]

Exemples
(i) Si V et W sont n'importe quels espaces vectoriels réels, l'application nulle $V \to W$, $v \mapsto 0_W$ et l'identité $\mathrm{id}_V : V \to V$, $v \mapsto v$ sont linéaires.
(ii) (Projections d'un produit) Si V et W des espaces vectoriels réels, les applications

$$\begin{array}{rccc} p_1 : & V \times W & \longrightarrow & V \\ & (v, w) & \longmapsto & v \end{array} \quad \text{et} \quad \begin{array}{rccc} p_2 : & V \times W & \longrightarrow & V \\ & (v, w) & \longmapsto & w \end{array}$$

sont linéaires (preuve : exercice).
[Cette linéarité provient immédiatement de la définition des lois d'un espace produit, à laquelle elle équivaut.]

D'une manière générale, si $n \in \mathbb{N}^*$ et si $V_1, \ldots, V_n$ sont des espaces vectoriels réels, alors pour tout $j \in \{1, \ldots, n\}$, la j^{e} *projection*

$$\begin{array}{ccc} \prod_{k=1}^{n} V_k & \longrightarrow & V_j \\ (v_1, \ldots, v_n) & \longmapsto & v_j \end{array}$$

est linéaire (preuve : exercice).
(iii) Si $n \in \mathbb{N}^*$ et si $A = \left(a_{j,k}\right)_{1 \le j,k \le n} \in \mathcal{M}_n\left(\mathbb{R}\right)$ est une matrice carrée, la *trace* de A, notée $\mathrm{Tr}(A)$, est la somme de ses éléments diagonaux :

$$\mathrm{Tr}(A) = \sum_{k=1}^{n} a_{k,k}.$$

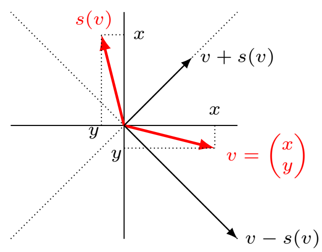
L'application $\mathcal{M}_n\left(\mathbb{R}\right) \to \mathbb{R}$, $A \mapsto \mathrm{Tr}(A)$ est linéaire (preuve : exercice).

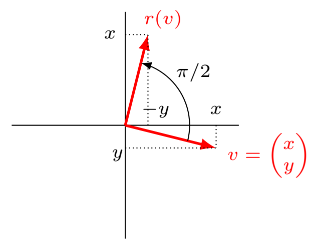
(iv) Les applications $s, r : \mathcal{M}_{2,1}(\mathbb{R}) \to \mathcal{M}_{2,1}(\mathbb{R})$ définies par

$$\forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathcal{M}_{2,1}(\mathbb{R}), \ s\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix} \quad \text{et} \quad r\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}$$

sont linéaires (preuve : exercice).

[Si on munit $\mathcal{M}_{2,1}(\mathbb{R})$ du produit scalaire standard et si on oriente ce plan euclidien par sa base canonique, ces applications vérifient : $\forall v \in \mathcal{M}_{2,1}(\mathbb{R})$, $s(v) + v \in \mathrm{Vect}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $s(v) - v \in \mathrm{Vect}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ d'une part, $\|r(v)\| = \|v\|$ et $\mathrm{mes}\,(\widehat{v, r(v)}) = \pi/2$ d'autre part : s est la *symétrie vectorielle orthogonale* par rapport à la droite $\mathrm{Vect}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et r est la *rotation vectorielle* d'angle $\pi/2$.]





(v) L'application $\mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \mapsto x - 2y + 3z$ est linéaire (exercice). D'une manière générale, si $n \geq 1$ et si $a_1, \ldots, a_n \in \mathbb{R}^n$, alors l'application

$$\begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ (x_1, \ldots, x_n) & \longmapsto & a_1x_1 + \cdots + a_nx_n = \displaystyle\sum_{k=1}^{n} a_kx_k \end{array}$$

est linéaire (preuve : exercice).

(vi) L'application $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^2 + y$ n'est pas linéaire. En effet, par exemple, $f((-1, 0) + (1, 0)) = f(0, 0) = 0$ est différent de $f(-1, 0) + f(1, 0) = 1 + 1 = 2$.

(vii) Si V est un espace vectoriel réel et si $\lambda \in \mathbb{R}^*$, alors l'application

$$\begin{array}{ccc} V & \longrightarrow & V \\ v & \longmapsto & \lambda \cdot v \end{array}$$

est linéaire (preuve : exercice). C'est l'*homothétie* de rapport λ.

Exercice
Soient V et W deux espaces vectoriels réels et $f : V \to W$. On suppose que f est *additive*, c'est-à-dire que $f(u + v) = f(u) + f(v)$ pour tous $u, v \in V$. Montrer que $f(x \cdot v) = x \cdot f(v)$, pour tous $v \in V$ et $x \in \mathbb{Q}$.

Proposition (composée d'applications linéaires)
La composée de deux applications linéaires est encore linéaire.

Preuve. Soient U, V et W des espaces vectoriels réels, $f : U \to V$ et $g : V \to W$ des applications linéaires. Soient $u, v \in U$ et $x \in \mathbb{R}$. Alors, $g \circ f(u+v) = g(f(u+v)) = g(f(u)+f(v)) = g(f(u))+g(f(v)) = g \circ f(u)+g \circ f(v)$ et $g \circ f(x \cdot v) = g(f(x \cdot v)) = g(x \cdot f(v)) = x \cdot g(f(v)) = x \cdot g \circ f(v)$. On a montré que $g \circ f$ est linéaire. ■

Définition (isomorphisme d'espaces vectoriels)
Soient V et W des espaces vectoriels réels et $f : V \to W$. On dit que f est un *isomorphisme (d'espaces vectoriels)* lorsque f est à la fois linéaire et bijective. On dit également que V et W sont *isomorphes* lorsqu'il existe un isomorphisme entre V et W. Lorsque $V = W$, un endomorphisme bijectif de V est appelé *automorphisme* de V. L'ensemble des applications linéaires bijectives $V \to W$ et l'ensemble des automorphismes de V sont respectivement notés

$$\mathrm{Isom}(V, W) \quad \text{et} \quad \mathrm{Aut}(V).$$

C'est une notion importante que l'on retrouve dans l'étude de toutes les structures algébriques. Deux espaces vectoriels isomorphes ont, en tant qu'espaces vectoriels, des structures indiscernables. L'enjeu de la question des isomorphismes est celui de la classification des espaces vectoriels, qui a une réponse très simple comme on le verra plus bas.

Proposition (réciproque d'un isomorphisme)
La réciproque d'un isomorphisme (d'espaces vectoriels) est encore un isomorphisme.

Preuve. Soient V et V' des espaces vectoriels réels et $f \in \operatorname{Isom}(V, V')$. Il suffit de montrer que f^{-1} est linéaire. Soient $u', v' \in V'$ et $x \in \mathbb{R}$. Puisque f est bijective, en notant $u = f^{-1}(u')$ et $v = f^{-1}(v')$, on a $u' = f(u)$ et $v' = f(v)$. Par ailleurs, puisque f est linéaire, $u' + v' = f(u) + f(v) = f(u + v)$ et $xu' = xf(u) = f(xu)$. Alors, $f^{-1}(u' + v') = u + v = f^{-1}(u') + f^{-1}(v')$ et $f^{-1}(xu') = xu = xf^{-1}(u')$. ■

Exemples

(i) L'identité est un automorphisme (!).

(ii) Si V est un espace vectoriel réel et si $\lambda \in \mathbb{R}^*$, alors l'homothétie de rapport λ est un automorphisme de V, dont la réciproque est l'homothétie de rapport $1/\lambda$.

(iii) Si $n \in \mathbb{N}^*$, l'application

$$\begin{array}{ccc} \mathcal{M}_{n,1}(\mathbb{R}) & \longrightarrow & \mathbb{R}^n \\ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} & \longmapsto & (x_1, x_2, \ldots, x_n) \end{array}$$

est un isomorphisme d'espaces vectoriels, dont la réciproque est $(x_1, x_2, \ldots, x_n) \mapsto \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

(iv) Si V et W sont des espaces vectoriels réels, l'application

$$\begin{array}{cccc} \iota : & V \times W & \longrightarrow & W \times V \\ & (v, w) & \longmapsto & (w, v) \end{array}$$

est un isomorphisme dont la réciproque est $W \times V \to V \times W$, $(w, v) \mapsto (v, w)$.

(v) Si $p, q \in \mathbb{N}^*$, l'application

$$\begin{array}{ccc} \mathbb{R}^p \times \mathbb{R}^q & \longrightarrow & \mathbb{R}^{p+q} \\ ((x_1 \ldots, x_p), (y_1, \ldots, y_q)) & \longmapsto & (x_1 \ldots, x_p, y_1 \ldots, y_q) \end{array}$$

est un isomorphisme de $\mathbb{R}$-espaces vectoriels dont la réciproque est

$$\begin{array}{ccc} \mathbb{R}^{p+q} & \longrightarrow & \mathbb{R}^p \times \mathbb{R}^q \\ (x_1 \ldots, x_{p+q}) & \longmapsto & ((x_1 \ldots, x_p), (x_{p+1} \ldots, x_{p+q})) . \end{array}$$

(vi) Retour à l'exemple (iv) de la salve d'exemples précédente. Un calcul élémentaire montre que $s^2 = \mathrm{id}_{\mathcal{M}_{2,1}(\mathbb{R})}$ et que $r^4 = \mathrm{id}_{\mathcal{M}_{2,1}(\mathbb{R})}$ (où l'on noté l'itération de la composition sous forme de puissance). Cela montre que s et r sont des automorphismes de $\mathcal{M}_{2,1}(\mathbb{R})$, dont les réciproques sont respectivement s et r^3.

Proposition (images directes et inverses d'un sous-espace par une application linéaire)
Soient V et W des espaces vectoriels réels et $f : V \to W$ une application linéaire.

(i) Si V' est un sous-espace vectoriel de V, alors $f(V')$ est un sous-espace vectoriel de W.

(ii) Si W' est un sous-espace vectoriel de W, alors $f^{-1}(W')$ est un sous-espace vectoriel de V.

Preuve.

(i) Soient $w, w' \in f(V')$ et $x \in \mathbb{R}$. Soient $v, v' \in V'$ tels que $w = f(v)$ et $w' = f(v')$. Alors, $w + w' = f(v) + f(v') = f(v + v') \in f(V')$ et $xw = xf(v) = f(xv) \in f(V')$, puisque V' est un sous-espace de V.

(ii) Soient $v, v' \in f^{-1}(W')$ et $x \in \mathbb{R}$. Alors, $f(v + v') = f(v) + f(v') \in W'$ et $f(xv) = xf(v) \in W'$ puisque W' est un sous-espace de W. Donc $v + v' \in f^{-1}(W')$ et $xv \in f^{-1}(W')$. ■

Définition (noyau d'une application linéaire)
Soient V et W des espaces vectoriels réels et $f : V \to W$ une application linéaire. Le *noyau* de f est l'ensemble des vecteurs de V qui ont 0_W pour image :

$$\ker(f) = \{v \in V,\ f(v) = 0_W\} = f^{-1}(\{0\}).$$

Un noyau est toujours un sous-space vectoriel de l'espace de départ, c'est une conséquence immédiate de la proposition précédente.

Proposition (noyau et injectivité d'une application linéaire)
Une application linéaire est injective si, et seulement si son noyau est l'espace nul.

Preuve. Soit $f \in \mathrm{Hom}(V, W)$. Si f est injective, puisque $f(0_V) = 0_W$, 0_V est l'unique vecteur de V dont l'image soit nulle. Donc $\ker(f) = \{0_V\}$. Inversement, on suppose que $\ker(f) = \{0_V\}$. Soient $v, v' \in V$ tels que $f(v) = f(v')$. Alors, $f(v - v') = 0_W$, ce qui entraîne que $v - v' \in \ker(f)$ et donc que $v - v' = 0_V$. ■

A noter
Lorsqu'on veut montrer qu'une application linéaire est injective, on calcule *toujours* son noyau : on ne refait pas sans cesse le calcul fait une fois pour toutes dans la preuve de la proposition précédente.

3.2 Bases et applications linéaires

Si $(v_1, \ldots, v_d)$ est une base d'un espace vectoriel réel V et si $f \in \mathrm{Hom}(V, W)$, alors l'image par f d'un vecteur écrit sous la forme

$$v = \sum_{k=1}^{d} x_k v_k$$

où $(x_1, \ldots, x_d) \in \mathbb{R}^d$ sont les coordonnées de v dans la base $(v_1, \ldots, v_d)$, est nécessairement, par linéarité, égale au vecteur

$$f(v) = \sum_{k=1}^{d} x_k f(v_k). \tag{2}$$

Proposition (une application linéaire est déterminée par l'image d'une base)
Soient V et W deux espaces vectoriels réels, $f, g \in \mathrm{Hom}(V, W)$ et $(v_1, \ldots, v_d)$ une base de V. Alors,

$$f = g \iff \forall k \in \{1, \ldots d\},\ f(v_k) = g(v_k).$$

Preuve. Seule l'implication ($\Longleftarrow$) est non triviale. Elle est garantie par la linéarité de f et de g, traduites dans la formule (2). ■

En d'autres termes, *pour que deux applications linéaires soient égales, il suffit qu'elles coïncident sur une base.*

Proposition (définir une application linéaire par l'image d'une base)
Soient V et W deux espaces vectoriels réels, $(v_1, \ldots, v_d)$ une base de V et $w_1, \ldots, w_d \in W$. Alors, il existe une unique application linéaire $f : V \to W$ telle que

$$\forall k \in \{1, \ldots d\},\ f(v_k) = w_k.$$

Preuve. L'unicité de f est garantie par la proposition précédente. En outre, la formule (2) montre que si $v \in V$ se développe en $v = x_1 v_1 + \cdots + x_d v_d$ dans la base $(v_1, \ldots, v_d)$ et si f est l'application linéaire cherchée, alors, nécessairement, $f(v) = x_1 f(v_1) + \cdots + x_d f(v_d) = x_1 w_1 + \cdots + x_d w_d$. On montre successivement que cela définit bien une application f, qu'elle est linéaire et qu'elle envoie chaque v_k sur w_k.
Pour tout $v = x_1 v_1 + \cdots + x_d v_d \in V$, on pose $f(v) = x_1 w_1 + \cdots + x_d w_d$. Cela définit bien une application parce que, d'une part, tout vecteur de V s'écrit comme une telle combinaison linéaire des v_k et, d'autre part, parce que cette combinaison linéaire est unique : les x_k sont uniquement déterminés par v (ce point est le point de la preuve le plus délicat). On montre alors que f ainsi définie, est linéaire.
Soient $v, v' \in V$ et soient $(x_1, \ldots, x_d) \in \mathbb{R}^d$ et $(x'_1, \ldots, x'_d) \in \mathbb{R}^d$ leurs coordonnées respectives dans la base $(v_1, \ldots, v_d)$. Soit aussi $x \in \mathbb{R}$. Alors, $v + v' = (x_1 + x'_1) v_1 + \cdots + (x_d + x'_d) v_d$ et $xv = (xx_1) v_1 + \cdots + (xx_d) v_d$,

ce qui implique que $f(v+v') = (x_1+x_1')\,w_1 + \cdots + (x_d+x_d')\,w_d = (x_1w_1+\cdots+x_dw_d) + (x_1'w_1+\cdots+x_d'w_d) = f(v)+f(v')$ et que $f(xv) = xx_1w_1+\cdots+xx_dw_d = x\,(x_1w_1+\cdots+x_dw_d) = xf(v)$. On a montré que f est linéaire.
Enfin, pour tout k, par construction, $f(v_k) = w_k$. ■

Définition (formes coordonnées dans une base)
Soient V un espace vectoriel réel et $(v_1,\ldots,v_d)$ une base de V. Pour tout $j\in\{1,\ldots,d\}$, l'application linéaire qui envoie v_j sur 1 et tous les autres v_k sur 0 est appelée j^{e} *forme coordonnée dans la base* $(v_1,\ldots,v_d)$. On la note v_j^* ou dx_j.

L'image d'un vecteur par la j^{e} forme coordonnée dans la base $(v_1,\ldots,v_d)$ est la j^{e} coordonnée de ce vecteur dans la base $(v_1,\ldots,v_d)$:

$$v_j^* : \begin{array}{ccc} V & \longrightarrow & \mathbb{R} \\ v=\displaystyle\sum_{k=1}^{d} x_kv_k & \longmapsto & x_j. \end{array}$$

Autrement dit, *le nombre réel $v_j^*(v)$ est la j^{e} coordonnée du vecteur v dans la base $(v_1,\ldots,v_d)$.* A l'aide de ces formes coordonnées, une fois la base $(v_1,\ldots,v_d)$ fixée, la décomposition de tout vecteur v de V s'écrit

$$v=\sum_{k=1}^{d} v_k^*(v)v_k.$$

Proposition (image d'une base par une application linéaire)
Soient V et W des espaces vectoriels réels, $f:V\to W$ une application linéaire et $(v_1,\ldots,v_d)$ une base de V.

(i) f est injective si, et seulement si la famille $\{f(v_1),\ldots,f(v_d)\}$ est libre

(ii) f est surjective si, et seulement si la famille $\{f(v_1),\ldots,f(v_d)\}$ est génératrice

(iii) f est bijective si, et seulement si $(f(v_1),\ldots,f(v_d))$ est une base de W

Preuve. Il suffit de montrer *(i)* et *(ii)*.

(i) On suppose f injective. Soient $x_1,\ldots x_d\in\mathbb{R}$ tels que $x_1f(v_1)+\cdots+x_df(v_d)=0_W$. Alors, comme f est linéaire, $f(x_1v_1+\cdots+x_dv_d)=0_W$. Puisque f est injective, cela impose que $x_1v_1+\cdots+x_dv_d$, ce qui force tous les x_k à être nuls puisque la famille $\{v_1,\ldots,v_d\}$ est libre. On a montré que $\{f(v_1),\ldots,f(v_d)\}$ est libre. Inversement, on suppose que $\{f(v_1),\ldots,f(v_d)\}$ est libre. Soit alors $v\in\ker(f)$. On développe v dans la base $(v_1,\ldots,v_d)$: soient $x_1,\ldots x_d\in\mathbb{R}$ tels que $v=x_1v_1+\cdots+x_dv_d$. En prenant l'image par f, puisque $v\in\ker(f)$, on obtient $0_W=x_1f(v_1)+\cdots+x_df(v_d)$. Mais la famille $\{f(v_1),\ldots,f(v_d)\}$ est libre, et donc les x_k sont nécessairement nuls. Donc $v=0_V$. On a montré que f est injective.

(ii) On suppose f surjective. Soit $w\in W$. Puisque f est surjective, soit $v\in V$ tel que $w=f(v)$. On développe v dans la base $(v_1,\ldots,v_d)$: soient $x_1,\ldots x_d\in\mathbb{R}$ tels que $v=x_1v_1+\cdots+x_dv_d$. Alors, $w=x_1f(v_1)+\cdots+x_df(v_d)\in\operatorname{Vect}\{f(v_1),\ldots,f(v_d)\}$. On a montré que $\{f(v_1),\ldots,f(v_d)\}$ est génératrice. Inversement, on suppose que $\{f(v_1),\ldots,f(v_d)\}$ est une famille génératrice de W. Soit $w\in W$. Puisque $\{f(v_1),\ldots,f(v_d)\}$ est génératrice, soient $x_1,\ldots x_d\in\mathbb{R}$ tels que $w=x_1f(v_1)+\cdots+x_df(v_d)$. Alors, $w=f(x_1v_1+\cdots+x_dv_d)\in\operatorname{im}(f)$. On a montré que f est surjective. ■

Proposition ($\mathbb{R}^d$ est le prototype d'espace vectoriel réel de dimension d)
Soient $d\in\mathbb{N}^*$ et V un espace vectoriel réel de dimension d. Pour toute base $(v_1,\ldots,v_d)$ de V, l'application

$$\begin{array}{ccc} \mathbb{R}^d & \longrightarrow & V \\ (x_1,\ldots,x_d) & \longmapsto & \displaystyle\sum_{k=1}^{d} x_kv_k \end{array}$$

est un isomorphisme d'espaces vectoriels, dont la réciproque est

$$\begin{array}{ccc} V & \longrightarrow & \mathbb{R}^d \\ v & \longmapsto & (v_1^*(v),\ldots,v_1^*(v))\,. \end{array}$$

Preuve. Que ces applications soient des bijections réciproques l'une de l'autre vient directement de la définition des formes coordonnées. Que la première (et donc la seconde aussi) soit linéaire est immédiat. ■

S'ensuit le corollaire suivant qui répond de façon complète à la question de la classification des espaces vectoriels réels de dimensions finies à isomorphisme près.

Proposition (deux espaces vectoriels de même dimension sont isomorphes)
(i) Tout espace vectoriel réel de dimension finie $d \in \mathbb{N}^*$ est isomorphe à $\mathbb{R}^d$.

(ii) Deux espaces vectoriels réels de dimensions finies sont isomorphes si, et seulement s'ils ont la même dimension.

Preuve. *(i)* et le sens ($\Leftarrow$) de *(ii)* viennent directement de la proposition précédente. Pour ($\Rightarrow$), on utilise la proposition sur l'image d'une base par une application linéaire, qui montre que si $f = V \to W$ est un isomorphisme, il envoie une base sur une base, obligeant V et W à avoir des bases qui ont le même nombre d'éléments et donc à avoir la même dimension. ■

A noter
(i) L'hypothèse sur la dimension finie est, à vrai dire, inutile. Modulo l'axiome du choix qui fournit l'existence de bases, *deux espaces vectoriels (réels) sont isomorphes si, et seulement s'ils ont la même dimension.*

(ii) Un espaces vectoriel réel V de dimension finie d est isomorphe à $\mathbb{R}^d$, on l'a montré par le choix d'une base. Changer de base fournit autant d'autres isomorphismes entre V et $\mathbb{R}^d$. Dans le jargon, à cause du caractère impératif du choix d'une base dans cet argumentation, on dit que tout espace vectoriel réel de dimension finie d est *non canoniquement* isomorphe à $\mathbb{R}^d$.

3.3 L'espace vectoriel $\mathrm{Hom}(V, W)$

Si V et W sont des espaces vectoriels réels, on munit l'ensemble $\mathrm{Hom}(V, W)$ des deux lois suivantes : pour tous $f, g \in \mathrm{Hom}(V, W)$ et pour tout $x \in \mathbb{R}$, les applications $f + g$ et $x \cdot f$ sont définies par

$$\forall v \in V,\ (f+g)(v) = f(v) + g(v) \ \text{ et } \ (x \cdot f)(v) = x \cdot f(v).$$

Proposition (l'espace $\mathrm{Hom}(V, W)$)
Pour les deux lois définies ci-dessus, $\mathrm{Hom}(V, W)$ *est un espace vectoriel réel.*

Preuve. Exercice. Le vecteur nul de $\mathrm{Hom}(V, W)$ est l'application nulle $V \to W$, $v \mapsto 0_W$. ■

Proposition (dimension de $\mathrm{Hom}(V, W)$)
Si V et W sont des espaces vectoriels de dimensions finies, alors l'espace vectoriel $\mathrm{Hom}(V, W)$ *est aussi de dimension finie et*

$$\dim_{\mathbb{R}} \mathrm{Hom}(V, W) = \dim_{\mathbb{R}}(V) \times \dim_{\mathbb{R}}(W).$$

Preuve. Soient $(v_1, \ldots, v_d)$ une base de V et $(w_1, \ldots, w_e)$ une base de W, où $d = \dim_{\mathbb{R}}(V)$ et $e = \dim_{\mathbb{R}}(W)$. Pour tout $(j, k) \in \{1, \ldots, d\} \times \{1, \ldots, e\}$, on note $\delta_{j,k}$ l'application linéaire $V \to W$ définie, en vertu de la proposition "définir une application linéaire par l'image d'une base", par :

$$\forall i \in \{1, \ldots, d\},\ \delta_{j,k}(v_i) = \begin{cases} 0_W & si \ \ i \neq j \\ w_k & si \ \ i = j. \end{cases}$$

Autrement dit, pour tout $v = x_1 v_1 + \cdots + x_d v_d \in V$, $\delta_{j,k}(v) = x_j w_k$.
On montre que $\mathcal{F} = \{\delta_{j,k},\ (j, k) \in \{1, \ldots, d\} \times \{1, \ldots, e\}\}$ est une famille libre et génératrice de $\mathrm{Hom}(V, W)$; comme $\mathcal{F}$ a $d \times e$ éléments, cela démontrera la proposition.
(i) $\mathcal{F}$ *est génératrice.* En effet, soit $f \in \mathrm{Hom}(V, W)$. On développe chaque $f(v_j)$ dans la base $(w_1, \ldots, w_e)$ de W : pour tout $(j, k) \in \{1, \ldots, d\} \times \{1, \ldots, e\}$, soit $f_{j,k} \in \mathbb{R}$ tel que $f(v_j) = f_{j,1} w_1 + \cdots + f_{j,e} w_e$. On note alors

$$g = \sum_{1 \leq j \leq d, 1 \leq k \leq e} f_{j,k} \delta_{j,k}.$$

L'application g est dans $\mathrm{Hom}(V,W)$, puisque c'est une combinaison linéaire des $\delta_{j,k}$ qui sont dans $\mathrm{Hom}(V,W)$. En outre, pour tout $i \in \{1,\ldots,d\}$, on a

$$g\left(v_i\right) = \sum_{k=1}^{e} f_{i,k} w_k = f\left(v_i\right).$$

Cela montre que les applications linéaires f et g coïncident sur la base $(v_1,\ldots,v_d)$ de V : elles sont donc égales. Cela montre que f est combinaison linéaire des $\delta_{j,k}$.
(ii) $\mathcal{F}$ *est libre.* Soit $\left(x_{j,k}\right)_{(j,k)\in\{1,\ldots,d\}\times\{1,\ldots,e\}}$ une famille de nombres réels tels que

$$\sum_{1\leq j\leq d, 1\leq k\leq e} x_{j,k}\delta_{j,k} = 0 \tag{3}$$

où l'on a noté 0 le vecteur nul de $\mathrm{Hom}(V,W)$. Soit $i \in \{1,\ldots,d\}$. En prenant l'image de v_i dans l'égalité (3), on obtient $x_{i,1}w_1 + \cdots + x_{i,e}w_e = 0_W$, ce qui impose que tous les $x_{i,k}$, $1 \leq k \leq e$ soient nuls puisque les w_k sont linéairement indépendants. Comme cela est vrai pour tout $i \in \{1,\ldots,d\}$, tous les $x_{i,k}$ sont nuls : on a montré que $\mathcal{F}$ est libre. ■

Définition (forme linéaire, espace dual)
Si V est un espace vectoriel réel, une *forme linéaire sur* V est une application linéaire $V \to \mathbb{R}$. L'espace vectoriel $\mathrm{Hom}(V,\mathbb{R})$ des formes linéaires sur V est appelé *espace dual* de V. On le note souvent V^*.

Proposition (dimension du dual)
Soit V un espace vectoriel de dimension finie. Alors, $\dim V = \dim V^*$.

Preuve. C'est une application directe de la proposition précédente. ■

A noter
Ce résultat, typique de la dimension finie, est faux en dimension infinie. Par exemple, l'espace dual $\mathbb{R}[X]^*$ de l'espace vectoriel $\mathbb{R}[X]$ des polynômes à coefficients réels, qui est de dimension dénombrable, est (isomorphe à) l'espace de toutes les suites à coefficients réels, qui est lui de dimension non dénombrable.

Proposition (formes linéaires sur $\mathbb{R}^d$)
Si $d \in \mathbb{N}^*$, les formes linéaires sur $\mathbb{R}^d$ sont les *applications de la forme*

$$\begin{array}{ccc} \mathbb{R}^d & \longrightarrow & \mathbb{R} \\ (x_1,\ldots,x_d) & \longmapsto & \displaystyle\sum_{k=1}^{d} a_k x_k \end{array}$$

où $(a_1,\ldots,a_d) \in \mathbb{R}^d$.

Preuve. D'une part, ces applications sont linéaires. D'autre part, si on note e_k le k^{e} vecteur de la base canonique de $\mathbb{R}^d$, une forme linéaire $\ell \in \left(\mathbb{R}^d\right)^*$ étant donnée, la formule

$$\ell\left(x_1,\ldots,x_d\right) = \ell\left(\sum_{k=1}^{d} x_k e_k\right) = \sum_{k=1}^{d} \ell\left(e_k\right) x_k$$

montre que ℓ est de la forme requise — avec $a_k = \ell\left(e_k\right)$. ■

Proposition (liens opératoires entre addition, multiplication scalaire et composition)
Soient f,g,h des applications linéaires et $x \in \mathbb{R}$. Les assertions suivantes sont vraies, chaque fois les les espaces de départ et d'arrivée des applications mentionnées donnent du sens aux égalités.
(i) $f \circ (g+h) = f\circ g + f \circ h$
(ii) $(f+g)\circ h = f\circ h + g\circ h$
(iii) $f\circ(x\cdot g) = x\cdot f\circ g$
(iv) $(x\cdot f)\circ g = x\cdot f\circ g$

Preuve. *(i)* Si v est dans l'espace de départ commun à g et h, alors $f \circ (g+h)(v) = (f \circ g + f \circ h)(v)$. Le calcul lui-même est laissé en exercice. Noter que cette propriété vient de la définition de la somme de deux applications *et* de la linéarité de f.
(ii) Si v est dans l'espace de départ de h, alors $(f+g) \circ h(v) = (f \circ h + g \circ h)(v)$. Le calcul lui-même est laissé en exercice. Noter que cette propriété ne vient que de la définition de la somme de deux applications (leur linéarité n'intervient pas).
(iii) Si v est dans l'espace de départ de g, alors $f \circ (x \cdot g)(v) = (x \cdot f \circ g)(v)$. Le calcul lui-même est laissé en exercice. Noter que cette propriété vient de la définition de la multiplication scalaire des applications *et* de la linéarité de f.
(iv) Si v est dans l'espace de départ de g, alors $(x \cdot f) \circ g(v) = (x \cdot f \circ g)(v)$. Le calcul lui-même est laissé en exercice. Noter que cette propriété ne vient que de la définition de la multiplication scalaire des applications (leur linéarité n'intervient pas). ■

A noter (vocabulaire : $\mathbb{R}$-*algèbre*)
Soit V un espace vectoriel réel. Cette dernière proposition appliquée à $\mathrm{End}(V)$, combinée à la structure d'espace vectoriel de $\mathrm{End}(V)$, à l'associativité de la composition et à l'existence de id_V comme élément neutre pour la composition confèrent à $\mathrm{End}(V)$ une structure de $\mathbb{R}$-*algèbre unifère*.

3.4 Le théorème du rang

Définition (rang d'une application linéaire)
Le *rang* d'une application linéaire f est la dimension de son image (qui est, on le rappelle, un sous-espace vectoriel de l'espace d'arrivée). On le note $\mathrm{rg}(f)$.

Le théorème du rang est l'un des théorèmes les plus opératoires de l'algèbre linéaire.

Proposition (théorème du rang)
Soient V et W deux espaces vectoriels réels de dimensions finies, et $f \in \mathrm{Hom}(V, W)$. Alors,

$$\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} \ker(f) + \dim_{\mathbb{R}} \mathrm{im}(f).$$

Preuve. Comme $\mathrm{im}(f)$ est un sous-espace vectoriel de W qui est de dimension finie, soient $r = \mathrm{rg}(f)$ et $(w_1, \ldots, w_r)$ une base de $\mathrm{im}(f)$. Pour chaque $j \in \{1, \ldots, r\}$, soit $v_j \in f^{-1}(w_j)$. Soient aussi $k = \dim \ker(f)$ et $(u_1, \ldots, u_k)$ une base de $\ker(f)$. On montre alors que $(u_1, \ldots, u_k, v_1, \ldots v_r)$ est une base de V, ce qui démontre le théorème.
(i) *La famille $\{u_1, \ldots, u_k, v_1, \ldots v_r\}$ est génératrice.*
En effet, soit $v \in V$. Puisque $(w_1, \ldots, w_r)$ est une base de $\mathrm{im}(f)$, soient $y_1, \ldots, y_r \in \mathbb{R}$ tels que $f(v) = y_1 w_1 + \cdots + y_r w_r$. Alors, $f(v) = f(y_1 v_1 + \cdots + y_r v_r)$, ce qui s'écrit encore $v - (y_1 v_1 + \cdots + y_r v_r) \in \ker(f)$. Sachant que $(u_1, \ldots, u_k)$ est une base de $\ker(f)$, on développe ce vecteur dans cette base : soient $x_1, \ldots, x_k \in \mathbb{R}$ tels que $v - (y_1 v_1 + \cdots + y_r v_r) = x_1 u_1 + \cdots + x_k u_k$. Cette égalité montre que $v \in \mathrm{Vect}\{u_1, \ldots, u_k, v_1, \ldots v_r\}$.
(ii) *La famille $\{u_1, \ldots, u_k, v_1, \ldots v_r\}$ est libre.*
En effet, soient $x_1, \ldots x_k \in \mathbb{R}$ et $y_1, \ldots y_r \in \mathbb{R}$ tels que $x_1 u_1 + \cdots + x_k u_k + y_1 v_1 + \cdots + y_r v_r = 0_V$. En prenant l'image par f, on obtient que $y_1 w_1 + \cdots + y_r w_r = 0_W$, ce qui impose que les y_j soient tous nuls puisque la famille $\{w_1, \ldots, w_r\}$ est libre (c'est une base de $\mathrm{im}(f)$). En reportant, il vient alors $x_1 u_1 + \cdots + x_k u_k = 0_V$, ce qui implique que les x_j soient également tous nuls puisque la famille $\{v_1, \ldots, v_k\}$ est libre (c'est une base de $\ker(f)$). ■

A noter
Une autre façon d'écrire la formule du rang : $\boxed{\mathrm{rg}(f) = \dim_{\mathbb{R}} V - \dim_{\mathbb{R}} \ker(f)}$

Proposition (bijectivité en dimension finie)
Soient V et W deux espaces vectoriels réels de dimensions finies, et $f = V \to W$ une application linéaire. On suppose que $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{R}}(W)$. Alors, les assertions suivantes sont équivalentes :

(i) f est injective

(ii) f est surjective

(iii) f est bijective

Preuve. Il suffit de montrer que *(i)* et *(ii)* sont équivalentes. On suppose que f est injective. Alors, $\dim_{\mathbb{R}} \ker(f) = 0$ et la formule du rang garantit que $\dim_{\mathbb{R}}(V) = \mathrm{rg}(f)$. Comme V et W ont la même dimension, cela entraîne que $\dim_{\mathbb{R}}(W) = \mathrm{rg}(f)$: les espaces $\mathrm{im}(f)$ et W sont inclus l'un dans l'autre et ont même dimension : ils sont égaux. Donc f est surjective. Inversement, on suppose que f est surjective. Alors, $\mathrm{rg}(f) = \dim(W) = \dim(V)$. Grâce au théorème du rang, il s'ensuit que $\dim_{\mathbb{R}} \ker(f) = 0$, et donc que f est injective. ■

A noter
Il est en général beaucoup plus facile de montrer qu'une application est injective que de montrer qu'elle est surjective. En ce sens, le théorème du rang est un moyen efficace de montrer qu'une application linéaire est surjective – ou de calculer son image.

Définition (hyperplan)
Si V est un espace vectoriel réel, un *hyperplan* de V est le noyau d'une forme linéaire non nulle.

Equations d'un hyperplan
Compte tenu de la forme des formes linéaires, on voit d'abord que H est un hyperplan de $\mathbb{R}^d$ si, et seulement s'il existe $(a_1, \ldots, a_d) \in \mathbb{R}^d \setminus \{(0, \ldots, 0)\}$ tel que

$$H = \left\{ (x_1, \ldots, x_d) \in \mathbb{R}^d,\ \sum_{k=1}^{d} a_k x_k = 0 \right\}.$$

On voit ensuite, si V est un espace vectoriel de dimension finie d dont $(v_1, \ldots v_d)$ est une base, que H est un hyperplan de V si, et seulement s'il existe $(a_1, \ldots, a_d) \in \mathbb{R}^d \setminus \{(0, \ldots, 0)\}$ tel que

$$H = \left\{ \sum_{k=1}^{d} x_k v_k \in V,\ \sum_{k=1}^{d} a_k x_k = 0 \right\}.$$

Dans les deux cas, on dit que $\sum_{k=1}^{d} a_k x_k = 0$ est une *équation* de H.

Exercice
Deux équations $\sum_{k=1}^{d} a_k x_k = 0$ et $\sum_{k=1}^{d} b_k x_k = 0$ définissent le même hyperplan si, et seulement si $(a_1, \ldots, a_d)$ et $(b_1, \ldots, b_d)$ sont colinéaires, *i.e.* si, et seulement s'il existe $\lambda \in \mathbb{R} \setminus \{0\}$, $(b_1, \ldots, b_d) = \lambda\,(a_1, \ldots, a_d)$.

[Si $a, b \in \mathbb{R}^d \setminus \{0\}$ sont proportionnels, ils définissent le même hyperplan ; ça, c'est le sens facile. L'autre implication est plus délicate : si $\left\{(x_1, \ldots, x_d) \in \mathbb{R}^d,\ a_1 x_1 + \cdots + a_d x_d = 0\right\} = \left\{(x_1, \ldots, x_d) \in \mathbb{R}^d,\ b_1 x_1 + \cdots + b_d x_d = 0\right\}$, alors les vecteurs a et b sont colinéaires. On peut procéder par récurrence sur d. Ce résultat trouve une preuve plus conceptuelle (et satisfaisante) dans le cadre de la théorie de la dualité et de la notion d'orthogonal d'un sous-espace, en faisant intervenir le théorème du rang.]

Proposition (dimension d'un hyperplan)
Soient V un espace vectoriel réel de dimension finie d et H une partie de V. Alors, H est un hyperplan de V si, et seulement si H est un sous-espace vectoriel de V de dimension $d-1$.

Preuve. Si H est un hyperplan, c'est un sous-espace vectoriel en tant que noyau et sa dimension est $d-1$ grâce au théorème du rang. Inversement, supposons que H soit un sous-espace de dimension $d-1$. Soit alors $(h_1, \ldots, h_{d-1})$ une base de H que l'on complète en une base $(h_1, \ldots, h_{d-1}, v_d)$ de V, où $v_d \in V \setminus H$. On note v_d^* la d^{e} forme coordonnée relative à cette base. Alors, v_d^* est une forme linéaire sur V. Puisque $v_d^*(v_d) = 1$, cette forme n'est pas nulle. Enfin, H est le noyau de v_d^*. ■

3.5 Equations cartésiennes ou paramétriques d'un sous-espace

Il s'agit ici d'un point de vocabulaire important. Dans un espace vectoriel V dans lequel on se donne une base — qui fournit un unique système de coordonnées à tout vecteur — il y a deux façons standard de décrire un sous-espace vectoriel W par des équations.

Equations paramétriques
Une *équation paramétrique* — ou plus exactement un *système d'équations paramétriques* — est la donnée d'une famille génératrice finie $\{w_1, \ldots, w_r\}$ de W, le plus souvent une base. Elle traduit en termes de coordonnées le

fait que tout vecteur de W est combinaison linéaire de $w_1, \dots, w_r$, ce qui s'écrit :

$$v \in W \Longleftrightarrow \exists s_1, \dots, s_r \in \mathbb{R},\ v = s_1 w_1 + \cdots + s_r w_r.$$

Dans cette formulation, les nombres réels sont les *paramètres* des équations paramétriques.

Equations cartésiennes

Une *équation cartésienne* — ou plus exactement un *système d'équations cartésiennes* — est la donnée d'une famille finie $\{\ell_1, \dots \ell_c\}$ de formes linéaires non nulles dont l'annulation simultanée définit W, ce qui s'écrit :

$$v \in W \Longleftrightarrow \ell_1(v) = \cdots = \ell_s(v) = 0.$$

Exemples

(i) Dans $\mathbb{R}^4$, soit W le sous-espace dont une équation paramétrique est

$$\begin{cases} x = 3v \\ y = u + v \\ z = 0 \\ t = -u. \end{cases}$$

Alors, W est engendré par les vecteurs $(0, 1, 0, -1)$ et $(3, 1, 0, 0)$ dont on voit immédiatement qu'ils sont linéairement indépendants : W est un plan (vectoriel) de $\mathbb{R}^4$. On "élimine"♫ u et v de ces équations : en écrivant d'abord (u, v) en fonction de (x, y, z, t), un raisonnement rapide montre que $(x, y, z, t) \in W$ si, et seulement si

$$\begin{cases} x - 3y - 3t = 0 \\ z = 0. \end{cases}$$

Ces deux équations forment une équation cartésienne de W.

(ii) Dans $\mathbb{R}^4$, soit W le sous-espace dont une équation cartésienne est

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

On peut voir W comme l'intersection des trois hyperlans que chacun des équations définissent. On peut s'attendre à ce que W soit donc une droite, à moins que l'une (au moins) de ces équations ne soit redondante. On arrive à exprimer toutes les variables en fonction de l'un d'entre elles, par exemple t. Pour éviter les fractions rationnelles malcommodes, on prend même $u = t/3$ pour paramètre. Ainsi, $(x, y, z, t) \in W$ si, et seulement si

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

On a trouvé une équation paramétrique de W, qui est la droite engendrée par le vecteur de coordonnées $(2, 0, 2, 3)$.

A noter

(i) Donner une équation paramétrique d'un sous-espace W de V sous la forme $v \in W \Longleftrightarrow \exists s_1, \dots, s_r \in \mathbb{R},\ v = s_1 w_1 + \cdots + s_r w_r$ revient à considérer W comme l'image de l'application linéaire

$$\begin{array}{ccc} \mathbb{R}^r & \longrightarrow & V \\ (s_1, \dots, s_r) & \longmapsto & s_1 w_1 + \cdots + s_r w_r. \end{array}$$

(ii) Donner une équation cartésienne d'un sous-espace W de V sous la forme $v \in W \Longleftrightarrow \ell_1(v) = \cdots = \ell_s(v) = 0$ revient à considérer W comme le noyau de l'application linéaire

$$\begin{array}{ccc} V & \longrightarrow & \mathbb{R}^s \\ v & \longmapsto & \left(\ell_1(v), \dots, \ell_s(v)\right), \end{array}$$

ou encore comme l'intersection des hyperplans dont les équations sont les ℓ_k.

♫Ce vocabulaire est consacré et il existe une théorie de l'élimination que l'on n'aborde pas ici de façon systématique.

3.6 Les attendus du chapitre

Le lire en détail, s'assurer de tout comprendre y compris les exemples, sans laisser de zone d'ombre. Faire, pour soi, une rédaction d'une résolution des exercices énoncés.

Section 3.1
Apprendre soigneusement la définition d'une application linéaire, sa caractérisation par conservation les combinaisons linéaires et retenir qu'une application linéaire est toujours nulle en zéro (preuve comprise). Apprendre le vocabulaire des homomorphismes et des endomorphismes.

Bien comprendre tous les exemples qui ont tous leur importance théorique ou générique ; faire les preuves déléguées en exercices.

Retenir que le composée de deux applications linéaire est linéaire, que la réciproque d'une bijection linéaire est automatiquement linéaire. Apprendre le vocabulaire des isomorphismes et des automorphismes.

Bien comprendre tous les exemples qui ont tous leur importance théorique ou générique ; rédiger les preuves pour soi, même si les notes de cours n'en parlent pas.

Retenir l'action d'une application linéaire sur les sous-espaces. Apprendre la définition du noyau d'une application linéaire, retenir son lien avec l'injectivité.

Section 3.2
Bien comprendre le phénomène de la linéarité qui assure qu'une application linéaire est déterminée par l'image d'une base.

La proposition "définir une application linéaire par l'image d'une base" est délicate, d'usage fréquent et très pratique. Il est important de bien comprendre et de retenir la démarche qui consiste à définir une application linéaire en donnant l'image d'une base et en disant qu'on prolonge l'application à tout l'espace de départ par linéarité.

Apprendre la définition des formes coordonnées qui prennent un rôle fondamental dans ce que l'on appelle la *théorie de la dualité*.

Bien comprendre la proposition sur l'image d'une base et les preuves ; ne laisser aucune zone d'ombre en dégageant avec soin les moments du raisonnement où interviennent les différentes hypothèses.

Bien retenir que deux espaces vectoriels réels de même dimension finie d sont isomorphes, puisqu'ils sont tous les deux isomorphes à $\mathbb{R}^d$. Retenir également comment le choix d'une base permet de montrer cela, *via* les formes coordonnées.

Section 3.3
Apprendre les définitions des lois usuelles sur $\mathrm{Hom}(V, W)$. Retenir qu'elle confèrent à $\mathrm{Hom}(V, W)$ une structure d'espace vectoriel. Se rédiger une preuve de cela.

Retenir la formule de la dimension de $\mathrm{Hom}(V, W)$. On la retrouvera en termes matriciels.

Retenir les opérations liant $+$, $\cdot$ et $\circ$ sur les applications linéaires. Bien comprendre, dans la preuve de la proposition, à quels endroits intervient la linéarité des applications en jeu. Ces propriétés seront transmises sans douleur au calcul matriciel, évitant ainsi des calculs fastidieux.

Section 3.4
Apprendre la définition du rang d'une application linéaire.

Le théorème du rang est un des théorèmes centraux de toute l'UE. Bien en retenir l'énoncé, en comprendre la preuve dans les moindres détails. Essayer de bien dégager pourquoi ce théorème, au fond, repose essentiellement sur le théorème de la base incomplète.

Retenir la proposition sur la bijectivité d'une application linéaire lorsque les espaces de départ et d'arrivée ont la même dimension et en quoi ce résultat repose sur le théorème du rang. C'est un résultat très pratique aux conséquences innombrables. Faire le rapprochement avec la proposition "applications entre deux ensembles finis de même cardinal" de la fin du chapitre 1.

Retenir ce qu'est un hyperplan avec ses deux aspects : noyau d'une forme linéaire non nulle d'un côté, sous-espace vectoriel de dimension un de moins de l'autre. Ces deux aspects seront très utiles pour naviguer dans une interprétation géométrique des systèmes linéaires.

Section 3.5
Connaître le vocabulaire des équations paramétriques et des équations cartésiennes.

Savoir interpréter les deux sorte d'équations en termes géométriques : système générateur d'un côté, intersection d'hyperplans de l'autre. Une autre interprétation géométrique : image d'une application linéaire d'un côté, noyau de l'autre.

Bien réaliser à quel point un système d'équations paramétriques n'est jamais unique. Bien réaliser à quel point un système d'équations cartésiennes n'est jamais unique. C'est pour cela qu'on parle toujours d'*un* système d'équations (paramétriques ou cartésiennes), et jamais *du* système d'équations.

Savoir passer d'une équation paramétrique à une équation cartésienne, et vice-versa.

4 Matrices et applications linéaires

4.1 Matrice d'une application linéaire modulo le choix de bases

Définition (matrice d'une application linéaire modulo le choix de bases)
Soient V et W des espaces vectoriels réels de dimensions finies, et $f \in \operatorname{Hom}(V, W)$. Soient aussi $\mathcal{B} = (v_1, \ldots, v_n)$ une base de V et $\mathcal{C} = (w_1, \ldots, w_m)$ une base de W. La *matrice de f dans les bases $\mathcal{B}$ et $\mathcal{C}$* est la matrice

$$\operatorname{Mat}_{\mathcal{B},\mathcal{C}}(f) \in \mathcal{M}_{m,n}(\mathbb{R})$$

dont la k^{e} colonne est le vecteur-colonne des coordonnées de $f(v_k)$ dans la base $(w_1, \ldots, w_m)$, pour tout $k \in \{1, \ldots, n\}$.

Autrement dit, si, pour chaque $k \in \{1, \ldots, n\}$, on écrit

$$f(v_k) = a_{1,k}w_1 + a_{2,k}w_2 + \cdots + a_{m,k}w_m = \sum_{j=1}^{m} a_{j,k}w_j$$

alors

$$\operatorname{Mat}_{\mathcal{B},\mathcal{C}}(f) = (a_{\ell,c})_{1 \leq \ell \leq m, 1 \leq c \leq n} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n} \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{R}).$$

A noter
Une manière encore plus condensée d'écrire cette matrice et de définir ses coefficients à l'aide des formes coordonnées dans la base $\mathcal{C}$: pour tout $\ell \in \{1, \ldots m\}$ et pour tout $c \in \{1, \ldots n\}$,

$$\left(\operatorname{Mat}_{\mathcal{B},\mathcal{C}}(f)\right)_{\ell,c} = w_\ell^*\left(f(v_c)\right).$$

Définition (matrice d'un endomorphisme dans *une* (même) base)
Avec les notations de la définition précédente, lorsque $V = W$ et lorsque $\mathcal{B} = \mathcal{C}$, on note seulement

$$\operatorname{Mat}_{\mathcal{B}}(f) = \operatorname{Mat}_{\mathcal{B},\mathcal{B}}(f) \in \mathcal{M}_n(\mathbb{R})$$

et on appelle $\operatorname{Mat}_{\mathcal{B}}(f)$ la *matrice de f dans la base $\mathcal{B}$*.

Exemples
(i) Si V est n'importe quel espace vectoriel réel de dimension d et si $\mathcal{B} = (v_1, \ldots, v_d)$ est n'importe quelle base de V, la matrice de id_V dans la base $\mathcal{B}$ est la matrice dont tous les termes sont nuls à l'exception des termes diagonaux qui valent 1. On la note I_d et on l'appelle *matrice identité* de taille d.

$$I_d = (\delta_{\ell,c})_{1 \leq \ell, c \leq d} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} = \operatorname{diag}(1, 1, \ldots, 1)$$

où $\delta_{\ell,c}$ est le symbole de Kronecker.

(ii) Soit $p : \mathbb{R}^3 \to \mathbb{R}^2$ la projection sur les deux premières coordonnées, définie par $f((x,y,z)) = (x,y)$, pour tout $(x,y,z) \in \mathbb{R}^3$. On note $\mathcal{C}_3 = (v_1, v_2, v_3)$ la base canonique de $\mathbb{R}^3$ et $\mathcal{C}_2\,(w_1, w_2)$ celle de $\mathbb{R}^2$. Alors,

$$\operatorname{Mat}_{\mathcal{C}_3,\mathcal{C}_2}(p) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

(iii) On reprend les exemples (iv) de la section 3.1. Les applications linéaires $s, r : \mathcal{M}_{2,1}(\mathbb{R}) \to \mathcal{M}_{2,1}(\mathbb{R})$ sont définies par

$$\forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathcal{M}_{2,1}(\mathbb{R}), \ s\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix} \quad \text{et} \quad r\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}$$

En notant $\mathcal{C} = (e_1, e_2)$ la base canonique de $\mathcal{M}_{2,1}(\mathbb{R})$ où, on le rappelle ici, $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, on obtient

$$\mathrm{Mat}_{\mathcal{C}}(s) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad \mathrm{Mat}_{\mathcal{C}}(r) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

On note également $b_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ et $b_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Ces deux vecteurs ne sont ni nuls ni proportionnels : ils forment une famille libre. Comme $\mathcal{M}_{2,1}(\mathbb{R})$ est de dimension 2, on en déduit que $\mathcal{B} = (b_1, b_2)$ est une base de $\mathcal{M}_{2,1}(\mathbb{R})$. Alors, $s(b_1) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2e_1 + e_2 = b_1 - b_2$ et $s(b_2) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = e_1 - e_2 = -b_2$, $s(e_1) = e_2 = \frac{1}{3}(b_1 + b_2)$ et enfin $s(e_2) = e_1 = \frac{1}{3}(b_1 - 2b_2)$. On en déduit que

$$\mathrm{Mat}_{\mathcal{B}}(s) = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}, \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(s) = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} \quad \text{et} \quad \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(s) = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}.$$

Proposition (une fois deux bases choisies, une application linéaire *est* une matrice)
Soient V et W des espaces vectoriels réels de dimensions finies respectives d et e, $\mathcal{B}$ une base de V et $\mathcal{C}$ une base de W. Alors, l'application

$$\begin{array}{ccc} \mathrm{Hom}(V, W) & \longrightarrow & \mathcal{M}_{e,d}(\mathbb{R}) \\ f & \longmapsto & \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f) \end{array}$$

est un isomorphisme d'espaces vectoriels.

Preuve. On note $\Phi_{\mathcal{B},\mathcal{C}}$ l'application de l'énoncé. Que $\Phi_{\mathcal{B},\mathcal{C}}$ soit linéaire est une paraphrase des définitions des lois $+$ et $\cdot$ sur $\mathrm{Hom}(V, W)$ d'une part, sur $\mathcal{M}_{e,d}(\mathbb{R})$ d'autre part. Le calcul du noyau de $\Phi_{\mathcal{B},\mathcal{C}}$ est immédiat : seul l'application nulle, neutre pour l'addition dans $\mathrm{Hom}(V, W)$, a pour matrice la matrice nulle. On en déduit que $\Phi_{\mathcal{B},\mathcal{C}}$ est injective. Par ailleurs, on a calculé la dimension des espaces vectoriels réels $\mathrm{Hom}(V, W)$ et $\mathcal{M}_{e,d}(\mathbb{R})$: elles sont toutes les deux égales au produit de. Donc $\Phi_{\mathcal{B},\mathcal{C}}$ est bijective. ■

Exercice
Montrer directement que $\Phi_{\mathcal{B},\mathcal{C}}$ est surjective.

A noter
Le choix de bases de V et de W permet ainsi de trouver un isomorphisme entre l'espace des applications linéaires $V \to W$ et celui des matrices $\mathcal{M}_{\dim W, \dim V}(\mathbb{R})$. Un tel choix ramène alors l'étude d'applications linéaires à celui de matrices. Ce nouveau point de vue sur les applications linéaires est souvent très efficace.

Cela dit, il est important de retenir que ce n'est *que* modulo le choix de bases au départ et à l'arrivée qu'un tel dictionnaire entre applications linéaires et matrices est possible. D'une part, changer de base au départ ou à l'arrivée modifie l'isomorphisme Φ. On obtient ainsi toute une famille d'isomorphismes entre $\mathrm{Hom}(V, W)$ et $\mathcal{M}_{\dim W, \dim V}(\mathbb{R})$. D'autre part, dans le cas général, il n'y pas d'isomorphisme *canonique* (c'est-à-dire ici, sans choix de bases) entre les espaces vectoriels $\mathrm{Hom}(V, W)$ et $\mathcal{M}_{\dim W, \dim V}(\mathbb{R})$.

4.2 Composition des applications linéaires, produit des matrices

Un calcul introductif
Soient U, V et W des espaces vectoriels réels de dimensions finies, $f : U \to V$ et $g : V \to W$ des applications linéaires. Soient également $\mathcal{B} = (u_1, \ldots, u_p)$ une base de U, $\mathcal{C} = (v_1, \ldots, v_q)$ une base de V et $\mathcal{D} = (w_1, \ldots, w_r)$ une base de W.
On sait que l'application $g \circ f : U \to W$ est linéaire. On cherche à calculer $\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(g \circ f)$ en fonction de $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ et de $\mathrm{Mat}_{\mathcal{C},\mathcal{D}}(g)$.

On note $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f) = (a_{i,j})_{1\leq i\leq q, 1\leq j\leq p} \in \mathcal{M}_{q,p}(\mathbb{R})$, $B = \mathrm{Mat}_{\mathcal{C},\mathcal{D}}(g) = (b_{i,j})_{1\leq i\leq r, 1\leq j\leq q} \in \mathcal{M}_{r,q}(\mathbb{R})$ et enfin $C = \mathrm{Mat}_{\mathcal{B},\mathcal{D}}(g\circ f) = (c_{i,j})_{1\leq i\leq r, 1\leq j\leq p} \in \mathcal{M}_{r,p}(\mathbb{R})$.
Alors, pour tout $i \in \{1,\dots,p\}$,

$$\begin{aligned} g\circ f(u_i) = g(f(u_i)) &= g\left(\sum_{j=1}^{q} a_{j,i} v_j\right) \\ &= \sum_{j=1}^{q} a_{j,i} g(v_j) \\ &= \sum_{j=1}^{q} a_{j,i} \left(\sum_{k=1}^{r} b_{k,j} w_k\right) = \sum_{k=1}^{r}\left(\sum_{j=1}^{q} b_{k,j} a_{j,i}\right) w_k \end{aligned}$$

Ainsi, pour tous $i \in \{1,\dots,p\}$ et $k \in \{1,\dots,r\}$, le nombre $\sum_{j=1}^{q} b_{k,j}a_{j,i}$ est la k^{e} coordonnée du vecteur $g\circ f(u_i)$ dans la base $(w_1,\dots,w_r)$: c'est le coefficient correspondant de la matrice de $g\circ f$ dans les bases $\mathcal{B}$ et $\mathcal{D}$. Autrement dit, on a montré que $c_{i,k} = \sum_{j=1}^{q} b_{k,j}a_{j,i}$ pour tous $i \in \{1,\dots,p\}$ et $k \in \{1,\dots,r\}$. C'était l'objectif fixé : on a calculé C en fonction des coefficients de A et B. On pose alors, la définition suivante.

Définition (produit de deux matrices)
Soient $p,q,r \in \mathbb{N}^*$, $M = (m_{i,j})_{1\leq i\leq p, 1\leq j\leq q} \in \mathcal{M}_{p,q}(\mathbb{R})$ et $N = (n_{i,j})_{1\leq i\leq q, 1\leq j\leq r} \in \mathcal{M}_{q,r}(\mathbb{R})$ – noter que le nombre de colonnes de M égale le nombre de lignes de N. Le *produit des matrices M et N*, dans cet ordre, est la matrice $M\times N \in \mathcal{M}_{p,r}(\mathbb{R})$ définie par : $\forall i \in \{1,\dots,p\}$, $\forall k \in \{1,\dots,r\}$,

$$\boxed{(M\times N)_{i,k} = \sum_{j=1}^{q} m_{i,j} n_{j,k}}$$

On note aussi $MN = M\times N$.

A noter
(i) Le produit des matrices définit ainsi des applications

$$\begin{array}{ccc} \mathcal{M}_{p,q}(\mathbb{R}) \times \mathcal{M}_{q,r}(\mathbb{R}) & \longrightarrow & \mathcal{M}_{p,r}(\mathbb{R}) \\ (M,N) & \longmapsto & MN. \end{array}$$

Dans le cas où $p=q=r$, les matrices sont carrées et ce produit définit une loi de composition *interne*

$$\begin{array}{ccc} \mathcal{M}_{p}(\mathbb{R}) \times \mathcal{M}_{p}(\mathbb{R}) & \longrightarrow & \mathcal{M}_{p}(\mathbb{R}) \\ (M,N) & \longmapsto & MN. \end{array}$$

(ii) Lorsque $A \in \mathcal{M}_d(\mathbb{R})$, on note $A^2 = A\times A$ et, par récurrence, $A^n = A^{n-1}\times A$ pour tout $n \in \mathbb{N}^*$.

Exemples

(i) Si $n \in \mathbb{N}^*$, si $(x_1,\dots,x_n) \in \mathcal{M}_{1,n}(\mathbb{R})$ et si $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$, alors

$$\boxed{(x_1,\dots,x_n)\times \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{k=1}^{n} x_k y_k \in \mathbb{R}.}$$

Cette formule, qui calcule le produit d'un vecteur-ligne par un vecteur-colonne dans cet ordre – c'est un nombre – permet de décrire le produit de deux matrices de la façon suivante, conséquence directe de la définition d'un produit de matrices :

le coefficient de la i^e ligne et de la j^e colonne d'un produit MN
est
le produit de la i^e ligne de M par la j^e colonne de N.

(ii) Si $M \in \mathcal{M}_{p,q}(\mathbb{R})$, alors $O_pM = M$, $MO_q = M$, $I_pM = M$ et $MI_q = M$. Cela se voit directement sur la définition du produit et sur celle des matrices nulles O_r et des matrices identité $I_r = (\delta_{ij})_{1\leq i,j\leq r}$ (symbole de Kronecker).

(iii) Pour attraper le mécanisme du produit de deux matrices, on peut le faire sur le calcul de petites matrices. Par exemple,

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \times \begin{pmatrix} a' & c' \\ b' & d' \end{pmatrix} = \begin{pmatrix} aa'+cb' & ac'+cd' \\ ba'+db' & bc'+dd' \end{pmatrix}$$

(iv) Lorsque deux matrices $M, N \in \mathcal{M}_p(\mathbb{R})$ sont carrées, on peut calculer à la fois MN et NM qui sont encore dans $\in \mathcal{M}_p(\mathbb{R})$. Attention :

en général, $MN \neq NM$

Par exemple, si $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$, alors $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ et $BA = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$. Lorsque $MN = NM$, on dit que les matrices M et N *commutent*.

(v) Le produit d'une matrice et de la matrice nulle, à droite ou à gauche, est encore la matrice nulle. Attention, le calcul du produit AB ci-dessus montre qu'

on peut avoir à la fois $A \neq O_p$, $B \neq O_p$ et $AB = O_p$

Cette propriété est une différence marquante entre le produit de nombres réels et le produit matriciel. Du côté des réels (ou complexes), il est vrai que $\forall x, y \in \mathbb{R},\ xy = 0 \Longrightarrow (x = 0 \text{ ou } y = 0)$ alors que cela est faux du côté des matrices.

Proposition (matrice d'une composée d'applications linéaires)
Soient U, V et W des espaces vectoriels réels de dimensions finies, $f : U \to V$ et $g : V \to W$ des applications linéaires. Soient également $\mathcal{B}$ une base de U, $\mathcal{C}$ une base de V et $\mathcal{D}$ une base de W. Alors,

$$\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(g \circ f) = \mathrm{Mat}_{\mathcal{C},\mathcal{D}}(g) \times \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f).$$

Preuve. Cela vient directement de la définition du produit de deux matrices, qui a été justement été choisie pour que cette proposition soit valide. ■

Proposition (propriétés du calcul matriciel)
(i) Si $M \in \mathcal{M}_{p,q}(\mathbb{R})$, $N \in \mathcal{M}_{q,r}(\mathbb{R})$ et $P \in \mathcal{M}_{r,s}(\mathbb{R})$, alors $(MN)P = M(NP)$.
(ii) Si $M \in \mathcal{M}_{p,q}(\mathbb{R})$, alors $I_pM = MI_q = M$.
(iii) Si $M, N \in \mathcal{M}_{p,q}(\mathbb{R})$ et si $P \in \mathcal{M}_{q,r}(\mathbb{R})$, alors $(M+N)P = MP + NP$.
(iv) Si $M \in \mathcal{M}_{p,q}(\mathbb{R})$ et $N, P \in \mathcal{M}_{q,r}(\mathbb{R})$, alors $M(N+P) = MN + MP$.
(v) Si $x \in \mathbb{R}$, $M \in \mathcal{M}_{p,q}(\mathbb{R})$ et $N \in \mathcal{M}_{q,r}(\mathbb{R})$, alors $x \cdot (MN) = (x \cdot M)N = M(x \cdot N)$.

Preuve. Dans chaque cas, modulo le choix de bases, en utilisant la proposition "une fois deux bases choisies, une application linéaire *est* une matrice" de la section 4.1, représenter les matrices de l'énoncé comme des applications linéaires entre des espaces vectoriels de dimensions finies adaptées à la situation – on peut prendre $\mathbb{R}^d$ comme espace de dimension d, c'en est le prototype. Utiliser alors les mêmes formules déjà démontrées dans le cadre des applications linéaires en les traduisant en termes matriciels – voir la section 3.3. ■

A noter

(i) Dans les conditions du *(i)*, on note $MNP = (MN)P = M(NP)$, sans parenthèses. Dans celles du (v), on note indifféremment $xMN = MxN = MNx$, également sans parenthèses.

(ii) Cette proposition a pour conséquence que les règles de calcul – addition, soustraction, multiplication scalaire, produit – sur les matrices sont les mêmes que celles sur les nombres, à l'exception notable de la non commutativité du produit pour les matrices carrées et du fait que deux matrices non nulles peuvent avoir un produit nul, ce qui empêche notamment de pouvoir définir une division par des matrices non nulles.

Proposition (action sur les vecteurs-ligne et les vecteurs-colonne)
Soient $p, q \in \mathbb{N}^*$ et $A \in \mathcal{M}_{p,q}(\mathbb{R})$.
(i) L'application "multiplication à droite par un vecteur-colonne"

$$\begin{array}{rccc} d_A : & \mathcal{M}_{q,1}(\mathbb{R}) & \longrightarrow & \mathcal{M}_{p,1}(\mathbb{R}) \\ & X & \longmapsto & AX \end{array}$$

est linéaire.
(ii) L'application "multiplication à gauche par un vecteur-ligne"

$$\begin{array}{rccc} g_A : & \mathbb{R}^p & \longrightarrow & \mathbb{R}^q \\ & X & \longmapsto & XA \end{array}$$

est linéaire.

Preuve. C'est immédiat avec les propriétés du calcul matriciel. ■

Proposition (le calcul matriciel traduit toutes les applications linéaires)
(i) Soient $p, q \in \mathbb{N}^*$ et $f : \mathcal{M}_{p,1}(\mathbb{R}) \to \mathcal{M}_{q,1}(\mathbb{R})$ une application linéaire. On note respectivement $\mathcal{C}_p$ et $\mathcal{C}_q$ les bases canoniques de $\mathcal{M}_{p,1}(\mathbb{R})$ et $\mathcal{M}_{q,1}(\mathbb{R})$. Alors, pour tout $X \in \mathcal{M}_{p,1}(\mathbb{R})$, $f(X) = \mathrm{Mat}_{\mathcal{C}_p,\mathcal{C}_q}(f) \times X$.

(ii) Soient V et W des espaces vectoriels réels de dimensions finies, $\mathcal{B} = (v_1, \dots, v_p)$ une base de V et $\mathcal{C} = (w_1, \dots, w_q)$ une base de W. Soit aussi $f \in \mathrm{Hom}(V, W)$. Alors, pour tout $v \in V$, si on note $X \in \mathcal{M}_{p,1}(\mathbb{R})$ le vecteur-colonne des coordonnées de v dans $\mathcal{B}$ et $Y \in \mathcal{M}_{q,1}(\mathbb{R})$ le vecteur-colonne des coordonnées de $f(v)$ dans $\mathcal{C}$, on a

$$Y = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f) \times X$$

Preuve. *(i)* Par définition de $\mathrm{Mat}_{\mathcal{C}_p,\mathcal{C}_q}(f)$, les applications linéaires f et $d_{\mathrm{Mat}_{\mathcal{C}_p,\mathcal{C}_q}(f)}$ coïncident sur la base $\mathcal{C}_p$ de $\mathcal{M}_{p,1}(\mathbb{R})$. Donc elles sont égales.
(ii) Une fois le produit matriciel défini, l'assertion découle immédiatement de la définition de $\mathrm{Mat}_{\mathcal{C}_p,\mathcal{C}_q}(f)$. ■

4.3 Matrices carrées inversibles

Définition (matrice inversible)
Soit $d \in \mathbb{N}^*$. Une matrice carrée $A \in \mathcal{M}_d(\mathbb{R})$ est dite *inversible* lorsqu'il existe $B \in \mathcal{M}_d(\mathbb{R})$ telle que $AB = BA = I_d$. On note $\mathrm{GL}_d(\mathbb{R})$ l'ensemble des matrices inversibles de $\mathcal{M}_d(\mathbb{R})$ (GL comme "groupe linéaire").

A noter
Si $A \in \mathcal{M}_d(\mathbb{R})$, il existe au plus une matrice B telle $AB = BA = I_d$. En effet, si $AC = CA = I_d$ aussi, alors $C = C(AB) = (CA)B = B$.

Définition (inverse d'une matrice inversible)
Lorsque $A \in \mathcal{M}_d(\mathbb{R})$ est inversible, l'unique matrice B telle que $AB = BA = I_d$ est appelée l'*inverse de A*. On la note A^{-1}. Cela fournit les formules $AA^{-1} = A^{-1}A = I_d$.

Proposition (inversibilité et bijectivité)
Soient V et W deux espaces vectoriels réels de même dimension finie d et $f \in \mathrm{Hom}(V, W)$. Soient $\mathcal{B}$ et $\mathcal{C}$ des bases respectives de V et W. Alors, f est bijective si, et seulement si $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ *est inversible. Dans ces conditions,*

$$\mathrm{Mat}_{\mathcal{C},\mathcal{B}}\left(f^{-1}\right) = \left(\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)\right)^{-1}.$$

PREUVE. Si f est bijective et si on note $g = f^{-1}$ la réciproque de f, alors $f \circ g = \mathrm{id}_W$ et $g \circ f = \mathrm{id}_V$, ce qui entraîne que $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)\,\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(g) = I_d$ et que $\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(g)\,\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f) = I_d$; d'où le résultat. Inversement, si $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ est inversible, on note B son inverse. Soit $g \in \mathrm{Hom}(W, V)$ tel que $B = \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(g)$. Alors, $\mathrm{Mat}_{\mathcal{B}}(g \circ f) = AB = I_d$, ce qui implique que $g \circ f = \mathrm{id}_V$. De la même manière, $\mathrm{Mat}_{\mathcal{C}}(f \circ g) = BA = I_d$ et donc $f \circ g = \mathrm{id}_W$. On a montré que f est bijective et que sa réciproque est g. ■

Proposition (un inverse d'un côté est un inverse tout court)
Soient $d \in \mathbb{N}^*$ et $A, B \in \mathcal{M}_d(\mathbb{R})$. Les assertions suivantes sont équivalentes.
(i) $BA = I_d$;
(ii) $AB = I_d$;
(iii) $AB = BA = I_d$.

PREUVE. On fixe une base $\mathcal{B}$ de $\mathbb{R}^d$. Soient alors $a, b \in \mathrm{End}\left(\mathbb{R}^d\right)$ tels que $A = \mathrm{Mat}_{\mathcal{B}}(a)$ et $B = \mathrm{Mat}_{\mathcal{B}}(b)$.
On suppose que $BA = I_d$. Alors, $b \circ a = \mathrm{id}_{\mathbb{R}^d}$, ce qui implique que a est injective. Comme a est un endomorphisme d'un espace vectoriel de dimension finie, cela entraîne que a est bijective. Alors, $a \circ b = a \circ (b \circ a) \circ a^{-1} = a \circ a^{-1} = \mathrm{id}_{\mathbb{R}^d}$ et donc $AB = I_d$. On a montré que $(i) \Rightarrow (ii)$. En intervertissant les rôles symétriques de A et de B, on a également montré que $(ii) \Rightarrow (i)$ et donc l'équivalence des trois assertions. ■

A noter
Faire une preuve directe de cette proposition en restant au sein du calcul vectoriel est loin d'être facile. Ici, dans l'argumentaire de cette preuve, on mesure la puissance du théorème du rang que l'on a utilisé en interprétant les matrices comme des applications linéaires modulo le choix d'une base.

Exemples
(i) Pour tout d, la matrice I_d est inversible, égale à son inverse, puisque $I_d^2 = I_d$.

(ii) Si $A \in \mathcal{M}_d(\mathbb{R})$ est inversible, alors pour tout $n \in \mathbb{N}^*$, A^n est inversible et $(A^n)^{-1} = \left(A^{-1}\right)^n$ (preuve par récurrence sur n). On note A^{-n} l'inverse de A^n, qui est donc aussi la puissance n^{e} de A^{-1}.

(iii) Soit $A = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. On montre par le calcul que $A^3 - A^2 - A - I_3 = O_3$, ce qui s'écrit encore $A(A^2 - A - I_3) = I_3$. Cela montre que A est inversible et que son inverse est $A^2 - A - I_3$. On peut calculer cette matrice. On trouve $A^{-1} = A^2 - A - I_3 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix}$.

(iv) Soit $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$. Un calcul direct montre que si $ad - bc \neq 0$, alors A est inversible et son inverse est $A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$. Il est utile de retenir cette formule du calcul de l'inverse d'une matrice de dimension 2.

Proposition (inverse d'un produit)
Soient $d \in \mathbb{N}^*$ et $A, B \in \mathcal{M}_d(\mathbb{R})$. Si A et B sont inversibles, alors AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.

PREUVE. Le calcul $(AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = AI_dA^{-1} = AA^{-1} = I_d$ prouve les deux assertions. On peut aussi les montrer en utilisant le fait que la composée de deux bijections f et g est une bijection, et que $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$. ■

4.4 Matrices et changements de bases

Définition (matrice de passage)
Soit V un espace vectoriel réel de dimension finie, $\mathcal{B}$ et $\mathcal{C}$ des bases de V. La *matrice de passage de $\mathcal{B}$ à $\mathcal{C}$* est la matrice dont la j^{e} colonne est formée des coordonnées dans $\mathcal{B}$ du j^{e} vecteur de $\mathcal{C}$. On la notera $\mathrm{Pass}_{\mathcal{B},\mathcal{C}}$.

Autrement dit,
$$\mathrm{Pass}_{\mathcal{B},\mathcal{C}} = \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(\mathrm{id}_V).$$

En particulier, toute matrice de passage est inversible et $\left(\mathrm{Pass}_{\mathcal{B},\mathcal{C}}\right)^{-1} = \mathrm{Pass}_{\mathcal{C},\mathcal{B}}$.

Les matrices de changement de bases permettent de relier les matrices d'une même application linéaire exprimées dans des bases différentes.

Proposition (formule générale de changement de bases)
Soient V et W des espaces vectoriels réels de dimensions finies, $\mathcal{B}$ et $\mathcal{B}'$ des bases de V, $\mathcal{C}$ et $\mathcal{C}'$ des bases de W. Soit aussi $f \in \mathrm{Hom}(V, W)$. Alors,

$$\mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(f) = \mathrm{Pass}_{\mathcal{C}',\mathcal{C}} \times \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f) \times \mathrm{Pass}_{\mathcal{B},\mathcal{B}'}.$$

Preuve. On compose les trois applications linéaires entre espaces vectoriels munis des bases indiquées sur le schéma

$$(V, \mathcal{B}') \xrightarrow{\mathrm{id}_V} (V, \mathcal{B}) \xrightarrow{f} (W, \mathcal{C}) \xrightarrow{\mathrm{id}_W} (W, \mathcal{C}')$$

et on applique deux fois la formule qui calcule la matrice d'une composée. On obtient : $\mathrm{Mat}_{\mathcal{B}',\mathcal{C}}(f \circ \mathrm{id}_V) = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)\,\mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_V)$, puis $\mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(\mathrm{id}_W \circ (f \circ \mathrm{id}_V)) = \mathrm{Mat}_{\mathcal{C},\mathcal{C}'}(\mathrm{id}_W)\,\mathrm{Mat}_{\mathcal{B}',\mathcal{C}}(f \circ \mathrm{id}_V)$ ce qui donne, en combinant les deux formules, $\mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(\mathrm{id}_W \circ f \circ \mathrm{id}_V) = \mathrm{Mat}_{\mathcal{C},\mathcal{C}'}(\mathrm{id}_W)\,\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)\,\mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_V)$, ce qu'il fallait démontrer puisque $\mathrm{id}_W \circ f \circ \mathrm{id}_V = f$. ■

Proposition (formule de changement de base pour un endomorphisme)
Soient V un espace vectoriel de dimension finie et $f \in \mathrm{End}(V)$. Soient $\mathcal{B}$ et $\mathcal{C}$ deux bases de V. Alors,

$$\mathrm{Mat}_{\mathcal{B}}(f) = \mathrm{Pass}_{\mathcal{B},\mathcal{C}} \times \mathrm{Mat}_{\mathcal{C}}(f) \times (\mathrm{Pass}_{\mathcal{B},\mathcal{C}})^{-1}.$$

Preuve. C'est un cas particulier de la proposition précédente. ■

A noter
En particulier, dans la situation de la proposition précédente, si on note P la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, alors P est inversible comme toute matrice de passage, et

$$\mathrm{Mat}_{\mathcal{B}}(f) = P \times \mathrm{Mat}_{\mathcal{C}}(f) \times P^{-1}$$

Il est important de retenir que l'effet d'un changement de base sur une matrice d'endomorphisme est la multiplication à droite et à gauche par une matrice et son inverse.
D'une manière générale, on dit que deux matrices carrées $A, B \in \mathcal{M}_d(\mathbb{R})$ sont *semblables* lorsqu'il existe une matrice $P \in \mathrm{GL}_d(\mathbb{R})$ telle que $B = PAP^{-1}$. La dernière proposition dit en substance que deux matrices sont semblables si, et seulement si elles représentent le même endomorphisme (dans deux bases en général différentes).

4.5 Les attendus du chapitre

Le lire en détail, s'assurer de tout comprendre y compris les exemples, sans laisser de zone d'ombre. Faire, pour soi, une rédaction d'une résolution des exercices énoncés.

Section 4.1
Associer une matrice à une application linéaire modulo le choix de bases et, inversement, associer à une matrice une application linéaire dont les espaces de départ et d'arrivée ont des dimensions adaptées et dans lesquels on choisit des bases est un mécanisme incontournable qu'il faut comprendre en profondeur.

Etant donnés V, W, $f \in \mathrm{Hom}(V, W)$, $\mathcal{B}$ et $\mathcal{C}$, savoir écrire $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ sur des exemples, et aussi de façon générale avec des pointillés et sous forme condensée (signes $\sum$).

Bien saisir ce que signifie $\mathrm{Mat}_{\mathcal{B}}(f)$ lorsqu'il n'est fait mention que d'une seule base en indice du symbole Mat.

Retenir ce qu'est la matrice identité de taille d et qu'elle est la matrice de l'application identique id à condition de prendre la même base au départ et à l'arrivée.

Retenir la proposition *une fois deux bases choisies, etc* qui donne un sens précis au fait qu'une fois des bases de V et de W choisies, la donnée d'une application linéaire $V \to W$ et celle d'une matrice de $\mathcal{M}_{\dim W, \dim V}(\mathbb{R})$ sont équivalentes. Ainsi, selon le contexte, on peut indifféremment raisonner du côté des matrices ou du côté des applications linéaires. Cette dualité de points de vue est très riche ; elle est la source de nombreux arguments fins.

Section 4.2
Apprendre à multiplier des matrices entre elles en prenant garde que les tailles soient adaptées. Savoir faire ces produits sur des exemples ne suffit pas, il faut aussi savoir écrire les formules générales, avec et sans pointillés, en apprenant avec soin comment les indices se manipulent. S'entraîner beaucoup pour développer les automatismes nécessaires.

Retenir que le produit matriciel traduit la composition des applications linéaires – et que c'est l'écriture de cette composition qui préside à la définition du produit matriciel. En particulier, ce produit n'est pas commutatif ; savoir produire des exemples de matrices qui commutent et de matrices qui ne commutent pas.

Assimiler les règles de calcul sur les matrices. En particulier, retenir que le produit de deux matrices non nulles peut être nul. Savoir produire de tels exemples.

Bien garder à l'esprit la multiplication des matrices par des vecteurs-colonne à droite, et par des vecteurs-ligne à gauche. Bien comprendre en quoi cette situation est générique dans la description de toutes les applications linéaires en dimensions finies.

Section 4.3
Retenir la définition d'une matrice inversible et de son inverse. Bien comprendre et retenir que l'inversibilité correspond à la bijectivité dans le dictionnaire qui lie les matrices et les applications linéaires modulo le choix de bases.

Retenir qu'il suffit qu'une matrice ait un inverse à gauche pour que ce dernier soit aussi un inverse à droite (et réciproquement), et que cela est dû, au fond, au théorème du rang.

Savoir lire une équation polynomiale (à coefficient constant non nul) sur une matrice comme l'explicitation de son inverse (exemple (iii)).

Retenir par cœur la formule de l'inverse d'une matrice inversible de taille 2 – on permute les termes diagonaux, on change les signes des termes antidiagonaux, on divise par $ad - bc$ qui n'est pas nul, exemple (iv).

Bien retenir que l'inverse du produit est le produit des inverses *dans l'autre sens.*

Section 4.4
Etant données deux bases d'un même espace vectoriel, savoir définir et écrire la matrice de changement de bases de l'une à l'autre. Ne pas oublier l'interprétation d'une matrice de passage en termes de la matrice de l'identité dans les deux bases respectives (dans le bon ordre).

Retenir l'existence de la formule générale de changement de base, savoir la retrouver dans le détail lorsque le besoin s'en fait sentir.

Retenir la formule de changement de base pour un endomorphisme. Au delà du détail, en retenir aussi le mécanisme qui assure que changer de base aboutit à une matrice semblable.

5 Systèmes linéaires, pivot de Gauss

5.1 Système linéaire, notation matricielle

Si $n \in \mathbb{N}^*$, une *équation linéaire* entre les *inconnues* $x_1, \dots x_n$ est une équation du type

$$a_1 x_1 + \cdots + a_n x_n = b \tag{4}$$

où $a_1, \dots, a_n$ et b sont des nombres réels. Lorsque $b = 0$, on dit que l'équation est *homogène*. Une *solution* de cette équation est un n-uplet $(x_1, \dots x_n) \in \mathbb{R}^n$ pour lequel l'égalité (4) est vraie. *Résoudre* une telle équation signifie en donner l'ensemble des solutions – de *toutes* les solutions.

Si $n, m \in \mathbb{N}^*$, un *système linéaire de m équations à n inconnues* $x_1, \dots x_n$ est un système constitué d'équations linéaires, du type

$$\left\{\begin{array}{l} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2 \\ \qquad\qquad\qquad \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n = b_m \end{array}\right. \tag{5}$$

où les $a_{\ell,c}$ et les b_k sont des nombres réels. Dans cette écriture en colonne des équations linéaires, *on sous-entend des* "et" qui sont presque toujours implicites. Une solution de ce système linéaire est un n-uplet $(x_1, \dots x_n) \in \mathbb{R}^n$ pour lequel les égalités (5) sont *toutes* vraies. *Résoudre* un tel système (d'équations) signifie en donner l'ensemble des solutions – de *toutes* les solutions. Lorsque $b_1 = \cdots = b_m = 0$, on dit que le système est *homogène*.

Notation matricielle Si on note $A \in \mathcal{M}_{m,n}(\mathbb{R})$ la matrice des coefficients système, $X \in \mathcal{M}_{n,1}(\mathbb{R})$ le vecteur-colonne des inconnues et $B \in \mathcal{M}_{m,1}(\mathbb{R})$ le vecteur-colonne du second-membre :

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}, \ X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \ B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

alors le système (5) s'écrit en notation compacte sous la forme

$$\boxed{AX = B}$$

où la notation AX désigne le produit matriciel. Le *système linéaire homogène associé* au système $AX = B$ est le système linéaire homogène $AX = 0_m$, où 0_m désigne le vecteur nul de l'espace $\mathcal{M}_{m,1}(\mathbb{R})$.

5.2 Structure de l'ensemble des solutions d'un système linéaire

Proposition (structure des solutions d'un système linéaire homogène)
(i) Soient $n \in \mathbb{N}^*$, $(a_1, \dots a_n) \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}$. L'ensemble des solutions de l'équation linéaire

$$a_1 x_1 + \cdots + a_n x_n = 0$$

est un hyperplan de $\mathbb{R}^n$.

(ii) L'ensemble des solutions d'un système linéaire homogène *à n inconnues est un sous-espace vectoriel de $\mathbb{R}^n$.*

Preuve. *(i)* Immédiat, voir chapitre 3.4. *(ii)* C'est une intersection d'hyperplans. ■

A noter Le système est homogène, c'est cela qui conduit à un sous-espace vectoriel. Résoudre une équation linéaire ou un système linéaire homogène revient presque toujours, quelle que soit la description que l'on adopte, à expliciter une base de l'ensemble de ses solutions.

Proposition (structure des solutions d'un système linéaire avec second membre)
Soient $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B \in \mathcal{M}_m(\mathbb{R})$. On suppose que $X_0 \in \mathcal{M}_n(\mathbb{R})$ est une solution du système linéaire $AX = B$.

(i) Si X est une solution du système homogène associé $AX = 0_m$, alors X_0+X est solution du système $AX = B$.
(ii) Si X est solution du système $AX = B$, alors $X - X_0$ est solution du système homogène associé $AX = 0_m$.

Preuve. *(i)* Si $AX = 0_m$, alors $A(X_0 + X) = AX_0 + AX = B + 0_m = B$. *(ii)* Si $AX = B$, alors $A(X - X_0) = AX - AX_0 = B - B = 0_m$. ■

Autrement dit, si on note $\mathcal{S}$ l'ensemble des solutions du système $AX = B$ et $\mathcal{S}_0$ le sous-espace vectoriel de $\mathbb{R}^n$ formé des solutions du système homogène associé $AX = 0_m$, si X_0 est un élément de $\mathcal{S}$ (que l'on suppose donc non vide), alors

$$\mathcal{S} = X_0 + \mathcal{S}_0$$

où la notation $X_0 + \mathcal{S}_0$ désigne $\{X_0 + V,\ V \in \mathcal{S}_0\}$.

Le slogan : *la solution générale d'un système linéaire avec second membre est la somme d'une solution particulière et de la solution générale du système homogène associé.*

A noter Un système linéaire peut n'avoir aucune solution. Prendre par exemple $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ et $B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

5.3 Résolutions des sytèmes linéaires, algorithme du pivot de Gauss

5.3.1 Les méthodes élémentaires de résolution

A noter *Les exemples choisis pour illustrer les méthodes élémentaires de résolution sont pris sur des systèmes linéaires qui ont tous 3 équations, 3 inconnues et une unique solution. Attention, ces exemples purement techniques n'illustrent pas du tout un paysage complet des systèmes linéaires, ni en ce qui concerne le nombre d'équation ou d'inconnues, ni en ce qui concerne la forme de l'ensemble des solutions.*

(i) Combinaisons linéaires adroites de lignes pour supprimer des inconnues
Un exemple : résoudre

$$(s) \begin{cases} 6x - 2y + z = 1 \\ 3x - y - 2z = 0 \\ x + z = 1. \end{cases}$$

On soustrait le double de la deuxième ligne à la première. Cette opération supprime à la fois les inconnues x et y. On obtient $5z = 1$, ce qui impose la valeur de z dans tout triplet de solutions : $z = \frac{1}{5}$. On reporte cette valeur dans la troisième équation, ce qui fournit $x = \frac{4}{5}$. Enfin, on reporte dans la seconde équation et on trouve $y = 2$. On a montré que si (x, y, z) est solution du système, alors nécessairement, $(x, y, z) = \left(\frac{4}{5}, 2, \frac{1}{5}\right)$. Inversement, on vérifie aisément que ce triplet est solution de (s).
Conclusion : l'ensemble des solutions de (s) est le singleton $\left\{\left(\frac{4}{5}, 2, \frac{1}{5}\right)\right\}$.

A noter On a procédé par condition nécessaires dans un premier temps : on a supposé que (x, y, z) était solution, puis on en a déduit que, nécessairement, $(x, y, z) = \left(\frac{4}{5}, 2, \frac{1}{5}\right)$. Il s'agit ensuite de vérifier que cette valeur de (x, y, z) fournit bien une solution, ce qu'on a fait dans un second temps.
D'une manière générale, on peut éviter de procéder en deux temps (analyse puis synthèse, pourrait-on dire) en écrivant une suite de systèmes équivalents au système initial. Encore faut-il, à chaque étape, s'assurer de cette équivalence. Cette démarche sera systématisée plus bas.

(ii) Méthode de résolution par substitution
Il s'agit d'écrire, ligne après ligne, une inconnue en fonction des autres, en éliminant ainsi les inconnues les unes après les autres.

Un exemple : résoudre

$$(s) \begin{cases} 6x - 2y + z = 1 \\ 3x + 2y - 4z = 0 \\ x - z = 1. \end{cases}$$

De la première ligne, on extrait x en fonction de y et de z ; comme on n'a modifié que la première ligne en la remplaçant par une équation équivalente, on obtient un système équivalent :

$$(s) \iff \begin{cases} x = \frac{1}{6}(1 + 2y - z) \\ 3x + 2y - 4z = 0 \\ x - z = 1. \end{cases}$$

On remplace ensuite x par sa valeur en fonction de y et z dans les deux dernières équations. Là encore, on s'assure aisément que le système obtenu est équivalent au précédent :

$$(s) \iff \begin{cases} x = \frac{1}{6}(1+2y-z) \\ 3y - \frac{9}{2}z = -\frac{1}{2} \\ \frac{1}{3}y - \frac{7}{6}z = \frac{5}{6} \end{cases} \iff \begin{cases} x = \frac{1}{6}(1+2y-z) \\ 6y - 9z = -1 \\ 2y - 7z = 5 \end{cases}$$

où la dernière transformation aura consisté à écrire les deux dernières lignes avec des coefficients entiers, en mettant les fractions au même dénominateur. On résout le système de 2 équations à 2 inconnues formé par les deux dernières lignes de façon analogue. Sans paroles :

$$(s) \iff \begin{cases} x = \frac{1}{6}(1-2y-z) \\ y = \frac{1}{6}(-1+9z) \\ -4z = \frac{16}{3} \end{cases} \iff \begin{cases} x = -\frac{1}{3} \\ y = -\frac{13}{6} \\ z = -\frac{4}{3} \end{cases}$$

Conclusion : l'ensemble des solutions de (s) est le singleton $\left\{\left(-\frac{1}{3}, -\frac{13}{6}, -\frac{4}{3}\right)\right\}$.

L'algorithme du pivot de Gauss consiste à systématiser la résolution d'une système linéaire. Par ailleurs, cet algorithme s'implante très bien dans tous les langages de programmation contemporains. L'algorithme appliqué au système lináire $AX = B$ consiste en une suite de transformations de matrices. Une extension – sur le même principe – de l'algorithme de Gauss fournit aussi un algorithme efficace de calcul de l'inverse d'une matrice inversible.

5.3.2 Un exemple introductif

On considère le système linéaire de 3 équations à 3 inconnues x, y, z suivant :

$$(s) \begin{cases} x + 2z = 3 \\ x + 2y + 4z = 5 \\ y + 2z = 2 \end{cases}$$

On transforme ce système en une suite de systèmes équivalents par des opérations successives. La première opération consiste à soustraire un multiple de la première ligne à toutes les autres afin de faire disparaître l'inconnue x.

$$(s) \iff \begin{cases} x + 2z = 3 \\ 2y + 2z = 2 \\ y + 2z = 2 \end{cases} \iff \begin{cases} x + 2z = 3 \\ y + z = 1 \\ y + 2z = 2. \end{cases}$$

La deuxième consiste à ne plus toucher la première ligne, puis à soustraire un multiple de la deuxième lignes à toutes les autres (il n'en reste ici plus qu'une) afin de faire disparaître l'inconnue y.

$$(s) \iff \begin{cases} x + 2z = 3 \\ y + z = 1 \\ z = 1 \end{cases}$$

Enfin, on soustrait un multiple de la troisième ligne aux précédentes afin d'en éliminer les occurences de z.

$$(s) \iff \begin{cases} x = 1 \\ y = 0 \\ z = 1. \end{cases}$$

L'algorithme du pivot de Gauss consiste à généraliser cette méthode à n'importe quel système linéaire. Les manipulations sur les lignes ne font à vrai dire intervenir que les coefficients de la matrice du système. Cela amène à écrire l'algorithme sur des matrices, en faisant abstraction de l'écriture des inconnues.

5.3.3 Opérations élémentaires sur les lignes et les colonnes d'une matrice

On se donne une matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$.
On lui fait subir trois types de transformations sur les lignes : permuter deux lignes, multiplier une ligne par un nombre réel non nul, ajouter à une ligne un multiple d'une autre, ainsi que les mêmes transformations sur les colonnes. Dans cette suite de définitions, on notera A' la matrice obtenue après transformation. Dans chaque cas, toute opération élémentaire sur les *lignes* d'une matrice revient à multiplier *à gauche* la matrice par une matrice inversible bien choisie ; de même, une opération élémentaire sur des *colonnes* d'une matrice revient à multiplier *à droite* la matrice par une matrice inversible bien choisie.

Notation Si $p \in \mathbb{N}^*$ et si $\ell, c \in \{1, \dots, p\}$, on note $E^{(p)}_{\ell,c}$ la matrice de $\mathcal{M}_p(\mathbb{R})$ dont tous les coefficients sont nuls à l'exception de celui de la ℓ^{e} ligne et de la k^{e} colonne qui vaut 1 :

$$E^{(p)}_{\ell,c} = \begin{array}{cl} \begin{pmatrix} & & \vdots & \\ \cdots & \cdots & 1 & \cdots \\ & & \vdots & \\ & & \vdots & \end{pmatrix} & \begin{array}{l} \\ \leftarrow \ell^{\text{e}} \text{ ligne} \\ \\ \\ \end{array} \\ \begin{array}{r} \uparrow \\ k^{\text{e}} \text{ colonne} \end{array} & \end{array}$$

Ce sont, on l'a vu, les matrices de la base canonique de $\mathcal{M}_p(\mathbb{R})$.

5.3.3.1 Permuter deux lignes – ou deux colonnes

Si $i, j \in \{1, \dots, m\}$, il s'agit d'échanger la ligne i et la ligne j de A. On note cette opération élémentaire

$$A \underset{L_i \leftrightarrow L_j}{\rightsquigarrow} A'.$$

De la même façon, pour tout $k, l \in \{1, \dots n\}$, l'échange de la k^{e} et de la ℓ^{e} colonne de A se note

$$A \underset{C_k \leftrightarrow C_\ell}{\rightsquigarrow} A'.$$

Exercice
On note $P^{(p)}_{i,j} \in \mathcal{M}_p(\mathbb{R})$ la matrice $P^{(p)}_{i,j} = I_m - E^{(m)}_{i,i} - E^{(m)}_{j,j} + E^{(m)}_{i,j} + E^{(m)}_{j,i}$ – c'est une *matrice de permutation.*
Alors, $\left(P^{(p)}_{i,j}\right)^2 = I_p$ – ce qui montre que $P^{(p)}_{i,j}$ est inversible – et

(i) $A \underset{L_i \leftrightarrow L_j}{\rightsquigarrow} A'$ si, et seulement si $A' = P^{(m)}_{i,j} \times A$.

(ii) $A \underset{C_k \leftrightarrow C_\ell}{\rightsquigarrow} A'$ si, et seulement si $A' = A \times P^{(n)}_{k,\ell}$.

Exemples

$$\begin{pmatrix} 1 & -1 & 0 \\ 3 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \underset{L_2 \leftrightarrow L_4}{\rightsquigarrow} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & -1 \\ 3 & 1 & 1 \end{pmatrix} = \underset{\substack{\uparrow \\ P^{(4)}_{2,4}}}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}} \begin{pmatrix} 1 & -1 & 0 \\ 3 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

et

$$\begin{pmatrix} 1 & -1 & 0 \\ 3 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \underset{C_2 \leftrightarrow C_3}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & -1 \\ 3 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 3 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \underset{\substack{\uparrow \\ P^{(3)}_{2,3}}}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}}.$$

5.3.3.2 Multiplier une ligne – ou une colonne – par un réel non nul

Si $i \in \{1, \dots m\}$ et si $x \in \mathbb{R} \setminus \{0\}$, il s'agit de remplacer la i^{e} ligne de A par son produit par x ; *idem* pour la j^{e} colonne si $j \in \{1, \dots n\}$. On note ces opérations élémentaires

$$A \underset{L_i \rightsquigarrow xL_i}{\rightsquigarrow} A' \quad \text{et} \quad A \underset{C_j \rightsquigarrow xC_j}{\rightsquigarrow} A'.$$

Exercice

(i) $A \underset{L_i \rightsquigarrow xL_i}{\rightsquigarrow} A'$ si, et seulement si $A' = \left(I_m + (x-1)E_{i,i}^{(m)}\right) \times A = \mathrm{diag}(1, \dots, 1, x, 1, \dots, 1) \times A$, où le x est à la i^{e} place.

(ii) $A \underset{C_j \rightsquigarrow xC_j}{\rightsquigarrow} A'$ si, et seulement si $A' = A \times \left(I_n + (x-1)E_{j,j}^{(n)}\right) = A \times \mathrm{diag}(1, \dots, 1, x, 1, \dots, 1)$, où le x est à la j^{e} place.

Exemples

$$\begin{pmatrix} 1 & -1 & 0 \\ 3 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \underset{L_2 \rightsquigarrow \frac{1}{3}L_2}{\rightsquigarrow} \begin{pmatrix} 1 & -1 & 0 \\ 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \underset{\substack{\uparrow \\ I_4 + \left(\frac{1}{3} - 1\right) E_{2,2}^{(4)}}}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}} \begin{pmatrix} 1 & -1 & 0 \\ 3 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

et

$$\begin{pmatrix} 1 & -1 & 0 \\ 3 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \underset{C_1 \rightsquigarrow -2C_1}{\rightsquigarrow} \begin{pmatrix} -2 & -1 & 0 \\ -6 & 1 & 1 \\ 0 & 0 & -1 \\ -2 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 3 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \underset{\substack{\uparrow \\ I_3 + (-2-1)\, E_{1,1}^{(3)}}}{\begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}.$$

5.3.3.3 Ajouter à une ligne – ou à une colonne – un multiple d'une autre

Si $i, j \in \{1, \dots m\}$ sont distincts et si $x \in \mathbb{R}$, il s'agit d'ajouter à la i^{e} ligne le produit de x par la j^{e} ligne ; *idem* pour les k^{e} et ℓ^{e} colonnes si $k, \ell \in \{1, \dots n\}$ sont distincts. On note ces opérations élémentaires

$$A \underset{L_i \rightsquigarrow L_i + xL_j}{\rightsquigarrow} A' \quad \text{et} \quad A \underset{C_k \rightsquigarrow C_k + xC_\ell}{\rightsquigarrow} A'.$$

Exercice

(i) $A \underset{L_i \rightsquigarrow L_i + xL_j}{\rightsquigarrow} A'$ si, et seulement si $A' = \left(I_m + E_{i,j}^{(m)}\right) \times A$.

(ii) $A \underset{C_k \rightsquigarrow C_k + xC_\ell}{\rightsquigarrow} A'$ si, et seulement si $A' = A \times \left(I_n + E_{\ell,k}^{(n)}\right)$.

Exemples

$$\begin{pmatrix} 1 & -1 & 0 \\ 3 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \underset{L_3 \rightsquigarrow L_3 + L_2}{\rightsquigarrow} \begin{pmatrix} 1 & -1 & 0 \\ 3 & 1 & 1 \\ 3 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \underset{\substack{\uparrow \\ I_4 + E_{3,2}^{(4)}}}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}} \begin{pmatrix} 1 & -1 & 0 \\ 3 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

et

$$\begin{pmatrix}1&-1&0\\3&1&1\\0&0&-1\\1&1&1\end{pmatrix}\underset{C_2\rightsquigarrow C_2+C_1}{\rightsquigarrow}\begin{pmatrix}1&0&0\\3&4&1\\0&0&-1\\1&2&1\end{pmatrix}=\begin{pmatrix}1&-1&0\\3&1&1\\0&0&-1\\1&1&1\end{pmatrix}\underbrace{\begin{pmatrix}1&1&0\\0&1&0\\0&0&1\end{pmatrix}}_{I_3+(-2-1)\,E^{(3)}_{1,2}}.$$

5.3.3.4 Réversibilité des opérations élémentaires

Proposition (réversibilité des opérations élémentaires)
Les opérations élémentaires sur les lignes et les colonnes d'une matrice sont toutes réversibles. Plus précisément, si $A, A' \in \mathcal{M}_{m,n}(\mathbb{R})$, alors pour tous $i, j \in \{1, \dots, m\}$, pour tous $k, \ell \in \{1, \dots, n\}$ pour tout $x \in \mathbb{R}$,

(i) $A \underset{L_i\leftrightarrow L_j}{\rightsquigarrow} A'$ *si, et seulement si* $A' \underset{L_i\leftrightarrow L_j}{\rightsquigarrow} A$

(ii) $A \underset{C_k\leftrightarrow C_\ell}{\rightsquigarrow} A'$ *si, et seulement si* $A' \underset{C_k\leftrightarrow C_\ell}{\rightsquigarrow} A$

(iii) Si $x \neq 0$, $A \underset{L_i\rightsquigarrow xL_i}{\rightsquigarrow} A'$ *si, et seulement si* $A' \underset{L_i\rightsquigarrow \frac{1}{x}L_i}{\rightsquigarrow} A$

(iv) Si $x \neq 0$, $A \underset{C_k\rightsquigarrow xC_k}{\rightsquigarrow} A'$ *si, et seulement si* $A' \underset{C_k\rightsquigarrow \frac{1}{x}C_k}{\rightsquigarrow} A$

(v) $A \underset{L_i\rightsquigarrow L_i+xL_j}{\rightsquigarrow} A'$ *si, et seulement si* $A' \underset{L_i\rightsquigarrow L_i-xL_j}{\rightsquigarrow} A$

(vi) $A \underset{C_k\rightsquigarrow C_k+xC_\ell}{\rightsquigarrow} A'$ *si, et seulement si* $A' \underset{C_k\rightsquigarrow C_k-xC_\ell}{\rightsquigarrow} A.$

Preuve. Exercice, c'est élémentaire. ■

A noter
Il importe de relier cette réversibilité aux trois exercices précédents, qui montrent que faire subir à A une opération élémentaire sur les lignes (*resp.* sur les colonnes) revient à multiplier A à gauche (*resp.* à droite) par une matrice carrée inversible – exercice : vérifier que les matrices indiquées dans lesdits exercices sont toutes inversibles en calculant explicitement leurs inverses. Cette propriété des opérations élémentaires fournit une autre preuve à la proposition de réversibilité.

5.3.4 Matrices échelonnées et réduites

Partant de n'importe quelle matrice et en faisant une suite d'opérations bien choisies sur les lignes (le fameux algorithme de Gauss, on y vient petit à petit), on aboutit à une matrice de forme très particulière qu'on appelle échelonnée et réduite, particulièrement propice à la résolution des systèmes linéaires.

— Dans la description d'une matrice, le symbole $\star$ désigne un réel quelconque —

Définition (matrice échelonnée, matrice échelonnée et réduite)
Soient $m, n \in \mathbb{N}^*$ et $R \in \mathcal{M}_{m,n}(\mathbb{R})$. On note $L_1, \dots, L_m$ les lignes de R et $C_1, \dots, C_n$ ses colonnes. La matrice R est dite *échelonnée* lorsque :

(i) toute ligne de R est nulle ou de la forme $(0, \dots, 0, 1, \star, \dots, \star)$. L'indice du premier terme non nul d'une ligne est son *pivot* ;

(ii) si $1 \leq i \leq m$ et si la ligne L_i est nulle, toutes les lignes L_j, $j \geq i$ sont nulles également ;

(iii) si $1 \leq i < j \leq m$ et si les lignes L_i et L_j ne sont pas nulles, alors le pivot de L_j est strictement supérieur au pivot de L_i ;

On dit que R est *échelonnée et réduite* lorsqu'elle vérifie en outre la condition :

(iv) si $1 \leq \ell \leq m$, si la ligne L_ℓ est non nulle et si c est son pivot, alors la colonne C_c est nulle à l'exception du 1 de la ℓ^{e} ligne.

Exemples et paraphrases
(i) Le pivot du vecteur-ligne $\left(0, 0, 1, 8, 0, -\frac{7}{2}\right)$ est 3.

(ii) Dans une matrice échelonnée et réduite, toutes les lignes nulles sont regroupées en bas de la matrice.

(iii) Dans une matrice échelonnée et réduite, la configuration $\begin{pmatrix} & & & & \vdots & & & & \\ 0 & \cdots & 0 & 0 & 0 & 1 & \star & \cdots & \star \\ 0 & \cdots & 0 & 1 & 0 & 0 & \star & \cdots & \star \\ & & & & \vdots & & & & \end{pmatrix}$ est interdite.

(iv) Si c est le ℓ^{e} pivot d'une matrice échelonnée et réduite R, alors R est de la forme

$$\begin{array}{cl} \begin{pmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 1 & \star & \cdots & \star \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{pmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow \ell^{\text{e}} \text{ ligne} \\ \\ \\ \\ \end{array} \\ \uparrow & \\ c^{e} \text{ colonne} & \end{array}$$

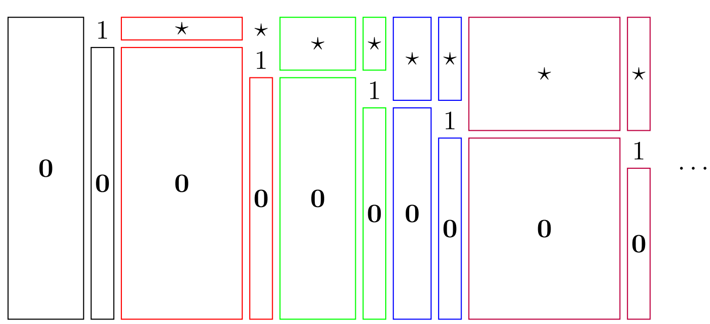
(v) Toute matrice de la forme suivante est échelonnée et réduite ; elle a quatre pivots dont la liste est $(2, 5, 7, 8)$.

$$\begin{pmatrix} 0 & 1 & \star & \star & 0 & \star & 0 & 0 & \star & \star \\ 0 & 0 & 0 & 0 & 1 & \star & 0 & 0 & \star & \star \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \star & \star \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \star & \star \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathcal{M}_{6,10}(\mathbb{R})$$

(vi) Les matrices échelonnées et réduites à moins de deux lignes et moins de deux colonnes sont

$$\begin{pmatrix} 0 \end{pmatrix}, \begin{pmatrix} 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & \star \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & \star \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(vii) Une matrice échelonnée peut (récursivement) se définir comme étant écrite par blocs sous la forme

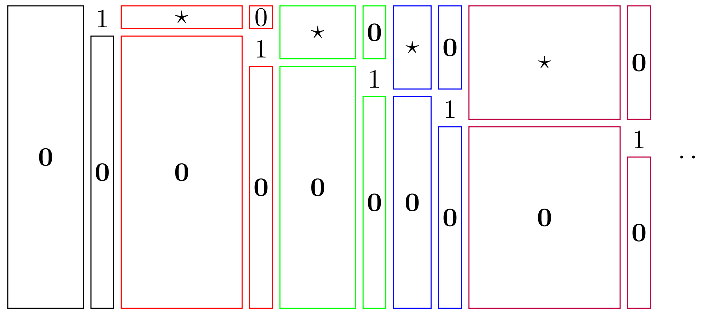


(6)

où les blocs superposés écrits sous la forme $\begin{array}{c} \boxed{\star} \\ \boxed{\mathbf{0}} \end{array}$ peuvent être vides (autrement dit, de largeur nulle).

Dans ces dessins, un rectangle contenant le symbole $\mathbf{0}$ désigne une matrice nulle, un rectangle contenant le symbole $\star$ désigne une matrice dont les coefficients sont des réels arbitraires. Noter aussi que le bloc de 0 le plus à gauche peut aussi être vide, ainsi que la toute dernière colonne.

De même, une matrice échelonnée *et réduite* peut se définir comme étant écrites par blocs sous la forme



(7)

La différence réside en la colonne de 0 au dessus du 1 de chaque pivot.

5.3.5 Echelonner et réduire une matrice, méthode du pivot

Proposition (toute matrice s'échelonne et se réduit, algorithme de Gauss)
Soient $m, n \in \mathbb{N}^*$ et $A \in \mathcal{M}_{m,n}(\mathbb{R})$.

(i) Il existe une suite d'opérations élémentaires sur les lignes de A qui aboutit à une matrice échelonnée et réduite.

(ii) Il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{R})$ telle que PA soit une matrice échelonnée et réduite.

Preuve.
(i) La preuve est constructive. Elle formalise (une version de) l'algorithme de Gauss proprement dit. On procède en deux temps : d'abord, on échelonne la matrice ; ensuite, on la réduit. Ces deux étapes aboutissent à une matrice échelonné et réduite en une succession finie d'opérations élémentaires sur les lignes.

1) Algorithme d'échelonnage
On procède par étapes successives, dont le nombre est inférieur ou égal au nombre n de colonnes de A. Chaque étape est l'*échelonnage à gauche* d'une sous-matrice, qui est une opération sur les lignes de la matrice entière.

L'échelonnage à gauche de A est la suite d'opérations suivantes :
- si A est nulle, on s'arrête, A est échelonnée à gauche (et même échelonnée et réduite !) ;
- si A n'est pas nulle, l'une au moins de ses colonnes n'est pas nulle. On s'occupe de la colonne non nulle la plus à gauche. On la note C_k, $1 \le k \le n$. Le nombre k est appelé premier $\boxed{\textit{pivot}}$ de la matrice A. Puisque C_k n'est pas nulle, au moins un de ses coefficients n'est pas nul. On note $a_{\ell,k}$ son premier coefficient non nul en partant du haut (autrement dit, ℓ est le plus petit entier j tel que $a_{j,k} \neq 0$)[♫]. Si $\ell \neq 1$, la première opération (élémentaire, sur les lignes) consiste à permuter L_1 et L_ℓ.

[Dans l'exemple qui suit cette preuve, cette étape est numérotée ①.]

La matrice obtenue a ses $k-1$ premières colonnes nulles, et le coefficient de sa première ligne et de sa k^{e} colonne n'est pas nul ; on le note a. On divise alors cette première ligne par a, c'est encore une opération sur les lignes. La nouvelle matrice obtenue a ses $k-1$ premières colonnes nulles, et le coefficient de sa première ligne et de sa k^{e} colonne est 1.

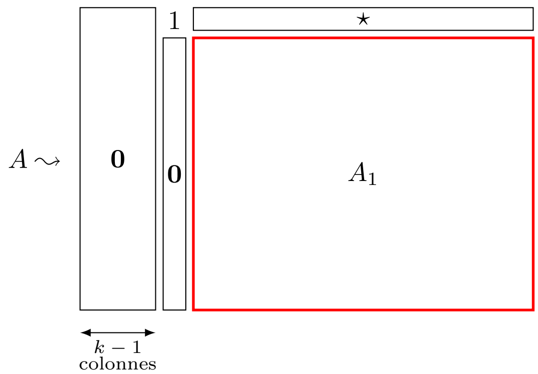
[Dans l'exemple, c'est l'étape ②.]

Il reste à remplacer les coefficients au dessous de ce 1 par des 0. Chaque ligne de cette matrice est de la forme $L = (0, \dots, 0, a, \star, \dots, \star)$, le a étant placé au rang k. On remplace chaque ligne L au dessous du 1 par $L - aL_1$; c'est une opération élémentaire sur les lignes qui remplace la ligne $L = (0, \dots, 0, a, \star, \dots, \star)$ par une ligne de la forme $(0, \dots, 0, 0, \star, \dots, \star)$ (attention, les $\star$ de cette nouvelle ligne ne sont pas les mêmes $\star$ que celles de la ligne L).

[Dans l'exemple, ce sont les étapes regroupées sous le numéro ③. Noter que les coefficients déjà nuls des lignes 2 et 3 de la colonne 2 rendent inutiles les opérations qu'on s'attendait à faire.]

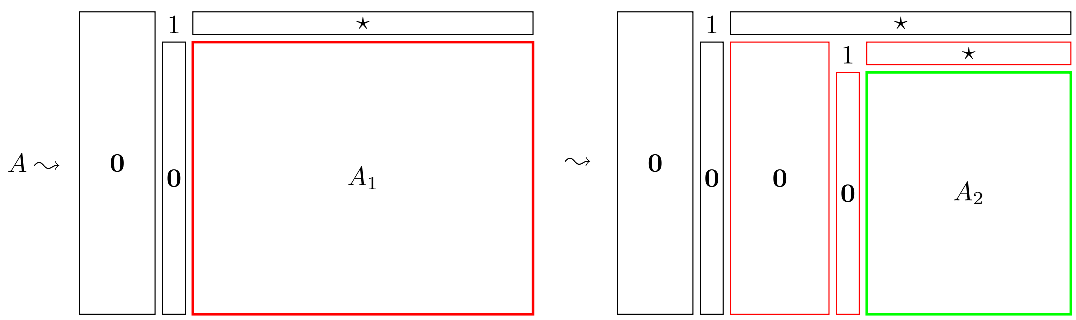
La nouvelle matrice obtenue à la suite de toutes ces opérations est échelonnée à gauche, ce qui signifie qu'elle est de la forme suivante :

[♫] Parfois, on appelle *pivot* ce nombre $a_{\ell,k}$ lui-même et non l'indice k de la première colonne non nulle. Pour ce vocabulaire-là, qui n'est pas universel dans le Landernau mathématique, il s'agit de faire preuve d'un peu de souplesse.



où A_1 est une matrice à coefficients réels qui a une ligne de moins que A et un nombre de colonnes strictement inférieur au nombre de colonnes de A.

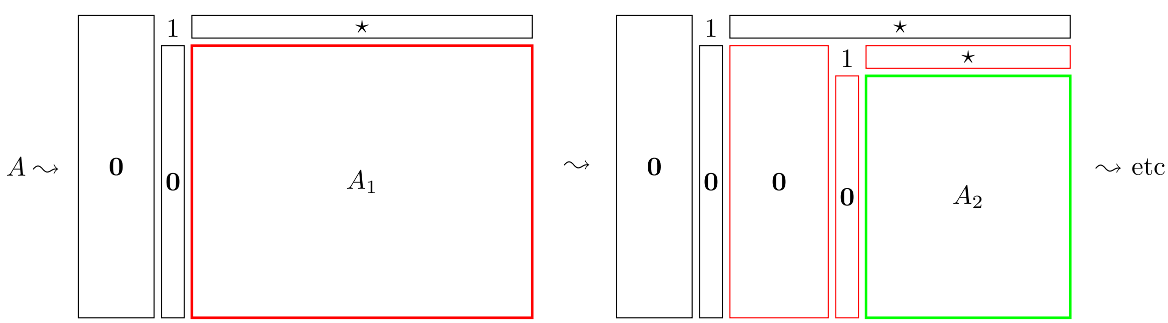
Ensuite, on échelonne à gauche la matrice A_1 selon le procédé précédemment utilisé pour A. Si la matrice A_1 est nulle, l'algorithme s'arrête, la matrice dessinée ci-dessus est échelonnée. Sinon, l'échelonnage de A_1 fait l'objet d'un nouveau pivot ; puisque tous les coefficients de la matrice à gauche de A_1 sont nuls, les opérations sur les lignes de A_1 n'affectent pas ce bloc de 0. Ainsi, lorsque A_1 n'est pas nulle, son échelonnage à gauche aboutit à une matrice de la forme



où A_2 est une matrice à coefficients réels qui a une ligne de moins que A_1 et un nombre de colonnes strictement inférieur au nombre de colonnes de A_1.

[Dans l'exemple, l'échelonnage à gauche de A_1 est la succession des étapes ④ et ⑤.]

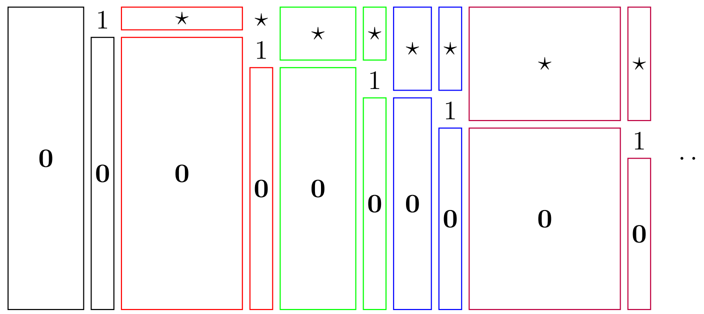
On poursuit récursivement ce procédé, qui s'arrête nécessairement puisque le nombre de ligne et le nombre de colonnes de la matrice obtenue en bas à droite décroissent strictement chaque fois que ladite matrice n'est pas nulle. Au moment où l'algorithme d'échelonnage s'arrête, la matrice obtenue est échelonnée, sous la forme (6).



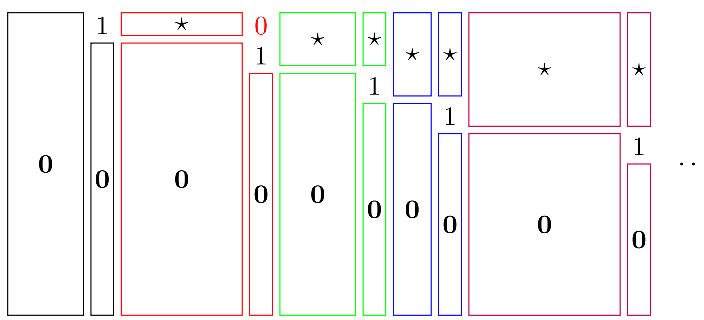
[Dans l'exemple, la matrice obtenue après l'étape ⑩ est échelonnée.]

2) Algorithme de réduction d'une matrice échelonnée

On part d'une matrice échelonnée, qui est ainsi de la forme



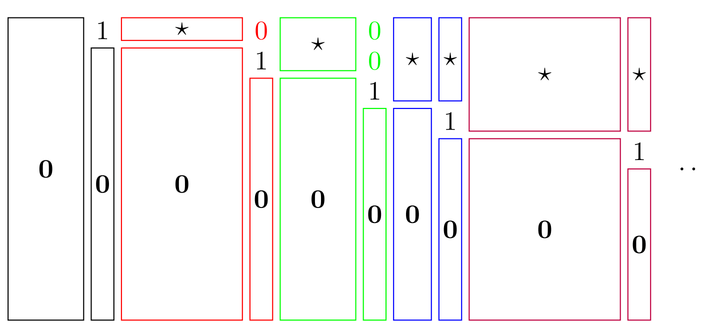
On appelle *colonne d'échelonnage* les colonnes de la forme $\begin{smallmatrix}\boxed{\star}\\ 1\\ \boxed{\mathbf{0}}\end{smallmatrix}$ dans la matrice échelonnée. S'il n'y a pas de colonne d'échelonnage, la matrice est nulle, il n'y a rien à faire (elle est échelonnée et réduite !). Sinon, il s'agit d'annuler $\star$ au dessus des 1 des colonnes d'échelonnage par des opérations élémentaires sur les lignes pour aboutir à une matrice échelonnée et réduite, écrite sous la forme (7). On procéde successivement colonne d'échelonnage par colonne d'échelonnage, de gauche à droite. Sur la première colonne d'échelonnage (en noir sur le dessin), il n'y a rien à faire, le 1 est tout en haut. S'il n'y a pas d'autre colonne d'échelonnage, il n'y a rien à faire, la matrice est est échelonnée et réduite. Sinon, on note a le coefficient au dessus du 1 de la deuxième colonne d'échelonnage (en rouge sur le dessin). On remplace alors la première ligne L_1 de A par $L_1 - aL_2$. La nouvelle matrice est de la forme



(8)

[Dans l'exemple, la première étape de réduction est l'étape ⑪.]

S'il n'y a pas d'autre colonne d'échelonnage, il n'y a rien à faire, la matrice est est échelonnée et réduite. Sinon, si la troisième colonne d'échelonnage s'écrit $\begin{smallmatrix}a\\ b\\ 1\\ \boxed{\mathbf{0}}\end{smallmatrix}$, on procède aux opérations sur les lignes $L_1 \rightsquigarrow L_1 - aL_3$ et $L_2 \rightsquigarrow L_1 - bL_3$. On obtient une matrice de la forme



(9)

[Dans l'exemple, la seconde étape de réduction est la succession d'opérations regroupées sous le numéro ⑫.]

On itère ce procédé jusqu'à épuisement des colonnes d'échelonnage. La matrice obtenue est échelonnée et réduite.

[Dans l'exemple, la dernière étape de réduction est la succession d'opérations regroupées sous le numéro ⑬.]

(ii) Chaque opération élémentaire sur les lignes d'une matrice consiste à multiplier la matrice à gauche par une matrice inversible. Comme le produit de deux matrices inversibles est encore inversible, (ii) est une conséquence directe de (i). ■

Un exemple

$$
\begin{pmatrix}0&0&0&0&0&0\\0&0&0&0&5&2\\0&2&3&-1&1&0\\0&2&3&0&-1&4\\0&-4&-6&1&5&3\end{pmatrix}
\underset{L_1\leftrightarrow L_3}{\overset{\textcircled{1}}{\rightsquigarrow}}
\begin{pmatrix}0&2&3&-1&1&0\\0&0&0&0&5&2\\0&0&0&0&0&0\\0&2&3&0&-1&4\\0&-4&-6&1&5&3\end{pmatrix}
\underset{L_1\rightsquigarrow\frac{1}{2}L_1}{\overset{\textcircled{2}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&-\frac{1}{2}&\frac{1}{2}&0\\0&0&0&0&5&2\\0&0&0&0&0&0\\0&2&3&0&-1&4\\0&-4&-6&1&5&3\end{pmatrix}
$$

$$
\underset{\substack{L_4\rightsquigarrow L_4-2L_1\\ L_5\rightsquigarrow L_5+4L_1}}{\overset{\textcircled{3}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&-\frac{1}{2}&\frac{1}{2}&0\\0&0&0&0&5&2\\0&0&0&0&0&0\\0&0&0&1&-2&4\\0&0&0&-1&7&3\end{pmatrix}
\underset{L_2\leftrightarrow L_4}{\overset{\textcircled{4}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&-\frac{1}{2}&\frac{1}{2}&0\\0&0&0&1&-2&4\\0&0&0&0&0&0\\0&0&0&0&5&2\\0&0&0&-1&7&3\end{pmatrix}
\underset{L_5\rightsquigarrow L_5+L_2}{\overset{\textcircled{5}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&-\frac{1}{2}&\frac{1}{2}&0\\0&0&0&1&-2&4\\0&0&0&0&0&0\\0&0&0&0&5&2\\0&0&0&0&5&7\end{pmatrix}
$$

$$
\underset{L_4\leftrightarrow L_5}{\overset{\textcircled{6}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&-\frac{1}{2}&\frac{1}{2}&0\\0&0&0&1&-2&4\\0&0&0&0&5&2\\0&0&0&0&0&0\\0&0&0&0&5&7\end{pmatrix}
\underset{L_3\leftrightarrow\frac{1}{5}L_3}{\overset{\textcircled{7}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&-\frac{1}{2}&\frac{1}{2}&0\\0&0&0&1&-2&4\\0&0&0&0&1&\frac{2}{5}\\0&0&0&0&0&0\\0&0&0&0&5&7\end{pmatrix}
\underset{L_5\leftrightarrow L_5-5L_3}{\overset{\textcircled{8}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&-\frac{1}{2}&\frac{1}{2}&0\\0&0&0&1&-2&4\\0&0&0&0&1&\frac{2}{5}\\0&0&0&0&0&0\\0&0&0&0&0&5\end{pmatrix}
$$

$$
\underset{L_4\leftrightarrow L_5}{\overset{\textcircled{9}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&-\frac{1}{2}&\frac{1}{2}&0\\0&0&0&1&-2&4\\0&0&0&0&1&\frac{2}{5}\\0&0&0&0&0&5\\0&0&0&0&0&0\end{pmatrix}
\underset{L_4\leftrightarrow\frac{1}{5}L_4}{\overset{\textcircled{10}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&-\frac{1}{2}&\frac{1}{2}&0\\0&0&0&1&-2&4\\0&0&0&0&1&\frac{2}{5}\\0&0&0&0&0&1\\0&0&0&0&0&0\end{pmatrix}
\underset{L_1\leftrightarrow L_1+\frac{1}{2}L_2}{\overset{\textcircled{11}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&0&-\frac{1}{2}&2\\0&0&0&1&-2&4\\0&0&0&0&1&\frac{2}{5}\\0&0&0&0&0&1\\0&0&0&0&0&0\end{pmatrix}
$$

$$
\underset{\substack{L_1\rightsquigarrow L_1+\frac{1}{2}L_3\\ L_2\rightsquigarrow L_2+2L_3}}{\overset{\textcircled{12}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&0&0&11/5\\0&0&0&1&0&24/5\\0&0&0&0&1&2/5\\0&0&0&0&0&1\\0&0&0&0&0&0\end{pmatrix}
\underset{\substack{L_1\rightsquigarrow L_1-\frac{11}{5}L_4\\ L_2\rightsquigarrow L_2-\frac{24}{5}L_4\\ L_2\rightsquigarrow L_3-\frac{2}{5}L_4}}{\overset{\textcircled{13}}{\rightsquigarrow}}
\begin{pmatrix}0&1&\frac{3}{2}&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\\0&0&0&0&0&0\end{pmatrix}.
$$

Remarque (sur la non unicité d'une chaîne d'opérations)
L'algorithme décrit fournit une matrice échelonnée et réduite bien déterminée. On peut faire d'autres choix d'opérations sur les lignes — parfois plus intelligents selon le contexte — qui peuvent aboutir à d'autres matrices échelonnées et réduites.

Corollaire (classes d'équivalence de matrices)
Soient $m, n \in \mathbb{N}^*$ et $A \in \mathcal{M}_{m,n}(\mathbb{R})$.
(i) Il existe une suite d'opérations sur les lignes et sur les colonnes de A qui aboutit à une matrice écrite par blocs sous la forme

$$R_r = \left(\begin{array}{c|c} I_r & O_{r,n-r} \\ \hline O_{m-r,r} & O_{m-r,n-r}\end{array}\right) \tag{10}$$

où $r \in \{0, \ldots, \min\{m, n\}\}$ et où $O_{p,q}$ désigne la matrice nulle à p lignes et q colonnes.
(ii) Il existe $r \in \{0, \ldots, \min\{m, n\}\}$ et deux matrices inversibles $P \in \mathcal{M}_m(\mathbb{R})$ et $Q \in \mathcal{M}_n(\mathbb{R})$ telles que

$$A = PR_rQ, \tag{11}$$

où R_r désigne la matrice (10).

Preuve. Les opérations sur les lignes reviennent à multiplier à gauche par une matrice inversible, celles sur les colonnes reviennent à multiplier à droite par une matrice inversible. Ainsi, il suffit de démontrer (i). On commence par faire l'agorithme de Gauss qui aboutit à une matrice échelonnée et réduite. Il ne reste plus qu'à permuter les colonnes convenablement pour obtenir la forme voulue. ■

Point de vocabulaire (matrices équivalentes)
On dit que deux matrices $A, B \in \mathcal{M}_{m,n}(\mathbb{R})$ sont *équivalentes* lorsqu'il existe deux matrices inversibles $P \in \mathcal{M}_m(\mathbb{R})$ et $Q \in \mathcal{M}_n(\mathbb{R})$ telles que $B = PAQ$. Le corollaire exprime que tout matrice est équivalente à une (unique) matrice de la forme (10).

A noter (preuve géométrique)
On peut faire des preuves alternatives du *(ii)* de cette proposition, sans passer par la méthode du pivot. En voici une, géométrique, instructive. Soit $f_A : \mathcal{M}_{n,1}(\mathbb{R}) \to \mathcal{M}_{m,1}(\mathbb{R})$, $X \mapsto AX$. La matrice A est la matrice de f_A dans les bases canoniques respectives de $\mathcal{M}_{n,1}(\mathbb{R})$ et de $\mathcal{M}_{m,1}(\mathbb{R})$. On note r le rang de f_A. Soit $(w_1 \dots, w_r)$ une base de $\operatorname{im}(f)$. Pour chaque $j \in \{1, \dots, r\}$, soit $v_j \in f_A^{-1}(w_j)$. Le théorème du rang assure que la dimension de $\ker(f_A)$ est $n - r$. Soit ainsi $(v_{r+1}, \dots, v_n)$ une base de $\ker(f_A)$. On montre facilement que la famille $\{v_1, \dots, v_n\}$ est libre. Puisqu'elle a n vecteurs, $\mathcal{B} = (v_1, \dots, v_n)$ est une base de $\mathcal{M}_{n,1}(\mathbb{R})$. On complète enfin $(w_1 \dots, w_r)$ en une base $\mathcal{C} = (w_1 \dots, w_m)$ de $\mathcal{M}_{m,1}(\mathbb{R})$. Alors, on vérifie aisément que la matrice de f_A dans les bases $\mathcal{B}$ et $\mathcal{C}$ est la matrice R de la formule (10). La formule de changement de bases du paragraphe (4.4) montre alors que A et R sont équivalentes.

A noter (unicité du nombre de pivots)
La preuve alternative ci-dessus montre que *le nombre r de la formule* (11) *est le rang de l'application linéaire f_A*. Ce nombre est donc déterminé par A : deux matrices de même taille R_r et R_s sont équivalentes si, et seulement si $r = s$.
On revient à l'algorithme du pivot de Gauss : le *nombre de pivots* de l'algorithme est le nombre de colonnes d'échelonnage définies dans la longue preuve de la proposition “toute matrice s'échelonne et se réduit” du paragraphe 5.3.5 ; on l'a noté r dans les lignes qui précèdent. Ce nombre ne change pas lors des opérations élémentaires sur les colonnes qui amènent ensuite à la formule (10) : c'est toujours le rang de f_A.
Cela montre en particulier que le nombre de pivots sera toujours le même, quels que soient les choix opérés lors d'une variante de l'algorithme de Gauss qui, partant de A, aboutit à une matrice échelonnée et réduite par une suite d'opérations élémentaires sur les lignes. Ce nombre est le *rang de A*, qui sera complètement défini au dernier chapitre du cours.

5.3.6 Résolution des systèmes linéaires par la méthode du pivot de Gauss

Définition (matrice augmentée d'un système linéaire)
Soient $m, n \in \mathbb{N}^*$, $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B \in \mathcal{M}_{m,1}(\mathbb{R})$. La *matrice augmentée* du système linéaire $AX = B$ est la matrice de $A \in \mathcal{M}_{m,n+1}(\mathbb{R})$ dont les n premières colonnes sont celles de A et dont la dernière colonne est le vecteur-colonne B. On la note

$$\left(\begin{array}{c|c} A & B \end{array}\right).$$

Proposition (invariance des solutions par opérations sur les lignes)
Soient $m, n \in \mathbb{N}^*$, $A, A' \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B, B' \in \mathcal{M}_{m,1}(\mathbb{R})$.

(i) S'il existe une opération élémentaire sur les lignes telle que $\left(\begin{array}{c|c} A & B \end{array}\right) \rightsquigarrow \left(\begin{array}{c|c} A' & B' \end{array}\right)$, alors les sytèmes $AX = B$ et $A'X = B'$ ont le même ensemble de solutions.

(ii) On ne change pas l'ensemble des solutions d'un système linéaire en faisant une suite finie d'opérations élémentaires sur les lignes.

(iii) Si $\left(\begin{array}{c|c} R & S \end{array}\right)$ est la matrice échelonnée et réduite obtenue à partir de $\left(\begin{array}{c|c} A & B \end{array}\right)$ après application de l'algorithme de Gauss, les systèmes linéaires $AX = B$ et $RX = S$ ont le même ensemble de solutions.

Preuve.
(i) Soit $L \in \mathcal{M}_m(\mathbb{R})$ la matrice inversible telle que $\left(\begin{array}{c|c} A' & B' \end{array}\right) = L \times \left(\begin{array}{c|c} A & B \end{array}\right)$. Alors, $A' = LA$ et $B' = LB$ (exercice : vérifier cela en détail). Si $X \in \mathcal{M}_{n,1}(\mathbb{R})$ est une solution du système $AX = B$, alors $A'X = LAX = LB = B'$, ce qui montre que X est aussi solution du système $A'X = B'$. Inversement, puisque $\left(\begin{array}{c|c} A & B \end{array}\right) = L^{-1} \times \left(\begin{array}{c|c} A' & B' \end{array}\right)$, le même raisonnement montre que toute solution de $A'X = B'$ est aussi solution de $AX = B$.

(ii) Procéder par récurrence sur le nombre d'opérations élémentaires.

(iii) L'algorithme de Gauss consiste en une suite finie d'opérations élémentaires sur les lignes. ■

A noter
Pour que la proposition soit vraie, il importe de ne faire sur la matrice augmentée que des opérations élémentaires sur les lignes.

Exemples
(i) On re-considère le système linéaire de 3 équations à 3 inconnues x, y, z suivant :

$$(s)\left\{\begin{array}{l} x + 2z = 3 \\ x + 2y + 4z = 5 \\ y + 2z = 1 \end{array}\right.$$

On écrit sa matrice augmentée et on lui applique l'algorithme de Gauss, ce qui s'écrit

$$\left(\begin{array}{ccc|c} 1 & 0 & 2 & 3 \\ 1 & 2 & 4 & 5 \\ 0 & 1 & 2 & 1 \end{array}\right) \underset{\text{alg. Gauss}}{\leadsto} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right).$$

Ainsi, le système (s) est équivalent au nouveau système

$$(s')\left\{\begin{array}{l} x = 3 \\ y = 1 \\ z = 0 \end{array}\right.$$

ce qui montre que le système initial (s) a une unique solution qui est le vecteur-colonne $\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$.

(ii) On considère le système linéaire de 3 équations à 3 inconnues x, y, z suivant :

$$(s)\left\{\begin{array}{l} x + 2z = 3 \\ x + 2y + 4z = 5 \\ y + z = 2 \end{array}\right.$$

On écrit sa matrice augmentée et on lui applique l'algorithme de Gauss, ce qui s'écrit

$$\left(\begin{array}{ccc|c} 1 & 0 & 2 & 3 \\ 1 & 2 & 4 & 5 \\ 0 & 1 & 1 & 2 \end{array}\right) \underset{\text{alg. Gauss}}{\leadsto} \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

Ainsi, le système (s) est équivalent au nouveau système

$$(s')\left\{\begin{array}{l} x + 2z = 0 \\ y + z = 1 \\ 0 = 1 \end{array}\right.$$

ce qui montre que le système initial (s) n'a pas de solution.

(iii) On considère le système linéaire de 3 équations à 3 inconnues x, y, z suivant :

$$(s)\left\{\begin{array}{l} x + 2z = 3 \\ x + 2y + 4z = 5 \\ y + z = 1 \end{array}\right.$$

On écrit sa matrice augmentée et on lui applique l'algorithme de Gauss, ce qui s'écrit

$$\left(\begin{array}{ccc|c} 1 & 0 & 2 & 3 \\ 1 & 2 & 4 & 5 \\ 0 & 1 & 1 & 1 \end{array}\right) \underset{\text{alg. Gauss}}{\leadsto} \left(\begin{array}{ccc|c} 1 & 0 & 2 & 3 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Ainsi, le système (s) est équivalent au nouveau système de 2 équations à trois inconnues x, y, z

$$(s')\left\{\begin{array}{l} x + 2z = 3 \\ y + z = 1 \end{array}\right.$$

ce qui montre que le système initial (s) a pour ensemble de solutions

$$\left\{ \begin{pmatrix} 3-2z \\ 1-z \\ z \end{pmatrix},\ z \in \mathbb{R} \right\} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix}.$$

Noter en passant que le vecteur $\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$ est une solution ("particulière") du système, et que système homogène associé a pour ensemble de solutions la droite vectorielle de $\mathcal{M}_{3,1}(\mathbb{R})$ engendrée par $\begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix}$. Relier cette manière de voir à la proposition "structure des solutions d'un système linéaire avec second membre".

A noter
La proposition précédente montre que l'algorithme de Gauss permet de ramener efficacement la résolution de n'importe que système linéaire à celle d'un système dont la matrice augmentée est échelonnée et réduite. *Encore faut-il savoir se dépêtrer de ces derniers.* Sur le modèle de l'exemple ci-dessus, il est élémentaire de décrire l'ensemble des solutions d'un système linéaire dont la matrice augmentée est échelonnée et réduite. Sa vacuité éventuelle se lit immédiatement. Lorsque cet ensemble n'est pas vide, la lecture d'une solution particulière et d'une base de l'espace vectoriel des solutions du système homogène associé est également immédiate, à condition de bien savoir faire le lien entre le système linéaire initial d'une part, d'autre part avec la matrice augmentée échelonnée et réduite obtenue après application de l'algorithme de Gauss.

5.3.7 Algorithme de calcul de l'inverse d'une matrice inversible

Soient $d \in \mathbb{N}^*$ et $A \in \mathcal{M}_d(\mathbb{R})$. On fabrique la matrice rectangulaire $\left(A \mid I_d \right)$ formée de la concaténation en ligne de A et de la matrice identité de dimension d. On appelle ici cette nouvelle matrice la matrice suraugmentée de A – ce vocabulaire n'est pas universel du tout du tout.

Proposition (inversiblité et algorithme de Gauss)
Soient $d \in \mathbb{N}^*$ et $A \in \mathcal{M}_d(\mathbb{R})$. La matrice A est inversible si, et seulement si la matrice échelonnée et réduite obtenue en appliquant l'algorithme de Gauss à la matrice suraugmentée $\left(A \mid I_d \right)$ est de la forme $\left(I_d \mid B \right)$ où $B \in \mathcal{M}_d(\mathbb{R})$. Dans ces conditions, $B = A^{-1}$.

Preuve. La matrice A est inversible si, et seulement si pour tout $Y \in \mathcal{M}_{d,1}(\mathbb{R})$, le système linéaire de d équations à d inconnues $AX = Y$ admet une solution unique $X \in \mathcal{M}_{d,1}(\mathbb{R})$; cette unique solution s'écrit alors $X = A^{-1}Y$. Par linéarité, cet énoncé reste valide en remplaçant Y par un vecteur arbitraire de la base canonique de $\mathcal{M}_{d,1}(\mathbb{R})$; on note δ_k le k^{e} vecteur de ladite base canonique. Ainsi, la matrice A est inversible si, et seulement si pour tout $k \in \{1, \dots, d\}$, le système linéaire $AX = \delta_k$ admet une unique solution, qui s'écrit alors $X_k = A^{-1}\delta_k$. Autrement dit, la matrice A est inversible si, et seulement si pour tout k, l'application de l'algorithme de Gauss à la matrice augmentée $\left(A \mid \delta_k \right)$ aboutit à une matrice échelonnée et réduite de la forme $\left(I_d \mid X_k \right)$ où $X_k \in \mathcal{M}_{d,1}(\mathbb{R})$. Puisque l'algorithme de Gauss appliqué à n'importe quelle matrice augmentée de la forme $\left(A \mid B \right)$ où $B \in \mathcal{M}_{d,1}(\mathbb{R})$ n'est guidé que par les coefficients de A, on peut regrouper l'écriture des algorithmes de Gauss appliqués aux d matrices augmentées $\left(A \mid \delta_k \right)$ en un seul algorithme appliqué à la matrice suraugmentée $\left(A \mid I_d \right)$. On obtient ainsi l'assertion : la matrice A est inversible si, et seulement si l'application de l'algorithme de Gauss à la matrice suraugmentée $\left(A \mid I_d \right)$ aboutit à une matrice échelonnée et réduite de la forme $\left(I_d \mid B \right)$ où $B \in \mathcal{M}_d(\mathbb{R})$. Dans ces conditions, pour tout $k \in \{1, \dots, d\}$, la k^{e} colonne de B est le vecteur-colonne X_k, unique solution du système linéaire $AX = \delta_k$; autrement dit, $B = A^{-1}$ puisque $X_k = A^{-1}\delta_k = B\delta_k$, pour tout k. ■

5.4 Les attendus du chapitre

Lire tout le chapitre en détail, s'assurer de tout comprendre y compris les exemples, sans laisser de zone d'ombre.

Comprendre ce qu'est un système linéaire.

Savoir écrire un système linéaire sous forme compacte $AX = B$. Inversement, savoir reconnaître un système linéaire dans une équation linéaire matricielle $AX = B$ où A est une matrice, B un vecteur-colonne et X le vecteur-colonne inconnu, tous de tailles adaptées.

Savoir ce qu'est le système linéaire *homogène* associé à un système linéaire quelconque.

Savoir interpréter l'ensemble des solutions d'un système linéaire homogène en termes géométriques, à savoir une intersection d'autant d'hyperplans (vectoriels) qu'il y a d'équations.

Retenir l'énoncé qui précise la structure de l'ensemble des solutions d'un système linéaire. Savoir exprimer ce résultat sous la forme du slogan : *la solution générale d'un système linéaire avec second membre est la somme d'une solution particulière et de la solution générale du système homogène associé.* S'assurer de comprendre avec exactitude tous les recoins du sens de ce slogan, en termes d'ensembles de solutions.

Savoir résoudre des systèmes très simples de façon élémentaire voire naïve – y compris par méthode de substitution. Avant de se lancer dans un algorithme du pivot de Gauss en bonne et due forme, lever la tête un instant et s'assurer que des moyens élémentaires de résolution ne sont pas plus efficaces.

Savoir associer à un système linéaire sa matrice des coefficients, sa matrice du second membre et sa matrice augmentée. Inversement, savoir associer un système linéaire à une matrice augmentée.

Savoir décrire et reconnaître une matrice échelonnée et une matrice échelonnée et réduite.

Savoir décrire l'algorithme du pivot de Gauss sur les matrices par le formalisme des trois opérations élémentaires sur les lignes $L_i \rightsquigarrow L_i + aL_j$, $L_i \rightsquigarrow aL_i$ et $L_i \leftrightarrow L_j$.

Savoir résoudre un système linéaire à l'aide de l'algorithme du pivot de Gauss, en mesurant le fait que c'est une méthode générale, algorithmiquement efficace – tous les ordinateurs sont programmés pour utiliser cette méthode, éventuellement améliorée par quelques variantes. En particulier, après lecture directe de la matrice échelonnée et réduite obtenue, savoir – c'est un point délicat :
- décrire une base de l'espace vectoriel des solutions du système linéaire homogène associé ;
- reconnaître si l'ensemble des solutions est vide et, dans le cas contraire, en donner une solution ("particulière").

Bien garder à l'esprit le lien entre la suite des matrices successives obtenues par opérations sur les lignes et les ré-écritures successives du système initial. Cela intervient notamment dans la lecture de la matrice échelonnée et réduite obtenue *in fine*.

Ne pas se contenter de savoir pratiquer l'algorithme du pivot sur des exemples simples, mais comprendre et s'approprier les résultats et les preuves du cas général.

Bien comprendre en quoi effectuer une opération élémentaire *sur les lignes* d'une matrice revient à multiplier cette matrice *à gauche* par une matrice inversible. Bien comprendre aussi en quoi effectuer une opération élémentaire *sur les colonnes* d'une matrice revient à multiplier cette matrice *à droite* par une matrice inversible.

Retenir l'énoncé sur les classes d'équivalence de matrices. Ce résultat à l'importance théorique avérée trouvera une résonnance naturelle dans le dernier chapitre sur le rang.

Savoir utiliser l'algorithme du pivot de Gauss pour reconnaître une matrice inversible et, dans ce cas, pour calculer son inverse.

6 Rang

6.1 Rang d'une application linéaire

On a déjà défini le rang d'une application linéaire comme étant la dimension de son image. En particulier, le rang d'un isomorphisme entre deux espaces vectoriels de dimension finie d est d.

Proposition (les isomorphismes ne modifient pas le rang)
Soit $f : V \to W$ une application linéaire. Si $p \in \mathrm{Aut}(V)$ et $q \in \mathrm{Aut}(W)$ sont des applications linéaires bijectives, alors

$$\mathrm{rg}\,(q \circ f \circ p) = \mathrm{rg}\,(f)\,.$$

Preuve. D'abord, puisque p est une application linéaire bijective, $\mathrm{im}\,(q \circ f \circ p) = \mathrm{im}\,(q \circ f)$ – les détails de cette double inclusion sont laissés en exercice. Il suffit donc de montrer que $\mathrm{rg}\,(q \circ f) = \mathrm{rg}\,(f)$. On note r le rang de f. Soit $(w_1 \ldots, w_r)$ une base de $\mathrm{im}\,(f)$. Alors, puisque q est une application linéaire bijective, on montre aisément que la famille $\{q\,(w_1)\,, \ldots, q\,(w_r)\}$ est libre et qu'elle engendre $\mathrm{im}\,(q \circ f)$: elle en forme une base. ■

6.2 Rang d'une matrice

Définition (rang d'une matrice)
Soient $m, n \in \mathbb{N}^*$. Si $A \in \mathcal{M}_{m,n}\,(\mathbb{R})$, le *rang* de A est le rang de l'application linéaire qu'elle représente dans les bases canoniques de $\mathcal{M}_{n,1}\,(\mathbb{R})$ et de $\mathcal{M}_{m,1}\,(\mathbb{R})$. Autrement dit, c'est le rang de l'application linéaire

$$f_A : \begin{array}{ccc} M_{n,1}\,(\mathbb{R}) & \longrightarrow & M_{m,1}\,(\mathbb{R}) \\ X & \longmapsto & AX. \end{array} \tag{12}$$

Proposition (le rang d'une matrice est le rang de n'importe quelle application linéaire qu'elle représente)
Soient $m, n \in \mathbb{N}^*$ et $A \in \mathcal{M}_{m,n}\,(\mathbb{R})$. S'il existe deux espaces vectoriels réels V et W de dimensions respectives n et m, une base $\mathcal{B}_V$ de V, une base $\mathcal{B}_W$ de W et une application linéaire $f : V \to W$ telle que

$$A = \mathrm{Mat}_{\mathcal{B}_V, \mathcal{B}_W}\,(f)\,,$$

alors $\mathrm{rg}\,(A) = \mathrm{rg}\,(f)$.

Preuve. On note respectivement $\mathcal{C}_m$ et $\mathcal{C}_n$ les bases canoniques de $\mathcal{M}_{m,1}\,(\mathbb{R})$ et de $\mathcal{M}_{n,1}\,(\mathbb{R})$. Soit $p : \mathcal{M}_{n,1}\,(\mathbb{R}) \to V$ l'application linéaire qui envoie la base $\mathcal{C}_n$ sur la base $\mathcal{B}_V$ – autrement dit, p envoie le vecteur colonne X sur le vecteur de V dont les coordonnées dans la base $\mathcal{B}_V$ sont les composantes de X. De même, soit $q : \mathcal{M}_{m,1}\,(\mathbb{R}) \to W$ l'application linéaire qui envoie la base $\mathcal{C}_m$ sur la base $\mathcal{B}_W$. Puisqu'elles envoient une base sur une base, p et q sont des isomorphismes. Par définition de $\mathrm{Mat}_{\mathcal{B}_V, \mathcal{B}_W}\,(f)$, on a la relation $f_A = q^{-1} \circ f \circ p$. Alors, $\mathrm{rg}\,(A) = \mathrm{rg}\,(f_A) = \mathrm{rg}\,(f)$ selon la proposition "les isomorphismes ne modifient pas le rang" du paragraphe 6.1. ■

Proposition (les matrices équivalentes sont les matrices de même rang)
Soient $m, n \in \mathbb{N}^*$ et $A, B \in \mathcal{M}_{m,n}\,(\mathbb{R})$. Alors, A et B sont équivalentes si, et seulement si elles ont le même rang.

Preuve. Pour tout $r \in \{1, \ldots, \min\{m, n\}\}$, on note

$$R_r = \left(\begin{array}{c|c} I_r & O_{r,n-r} \\ \hline O_{m-r,r} & O_{m-r,n-r} \end{array}\right)\ ; \tag{13}$$

l'écriture de f_{R_r} montre directement que $\mathrm{rg}\,(R_r) = r$. La proposition "classes d'équivalences de matrices" du paragraphe 5.3.5 montre alors que toute matrice de rang r est équivalente à la matrice R_r. Ainsi, si A et B sont équivalentes, elles sont toutes les deux équivalentes à une même R_r et donc ont le même rang r. Inversement, si A et B ont le même rang r, elles sont toutes les deux équivalentes à R_r et sont donc équivalentes entre elles. ■

A noter
Deux matrices sont équivalentes si, et seulement si elles représentent la même application linéaire dans des bases adaptées – cela se voit en interprétant les matrices comme des applications linéaires. La proposition "le rang

d'une matrice est le rang de n'importe application linéaire qu'elle représente" fournit ainsi une autre preuve de la proposition "les matrices équivalentes sont les matrices de même rang".

Définition (transposée d'une matrice)
Soient $m, n \in \mathbb{N}^*$ et $A = (A_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathcal{M}_{m,n}(\mathbb{R})$. La *transposée* de A est la matrice

$$^tA = (A_{j,i})_{1 \leq j \leq n, 1 \leq i \leq m} \in \mathcal{M}_{n,m}(\mathbb{R})$$

En particulier, la transposée d'un vecteur-colonne est un vecteur-ligne et inversement.

Proposition (propriétés opératoires de la transposition)
Soient $m, n, p \in \mathbb{N}^*$, $x \in \mathbb{R}$, $A, B \in \mathcal{M}_{m,n}(\mathbb{R})$, $C \in \mathcal{M}_{n,p}(\mathbb{R})$ et $P \in \mathcal{M}_m(\mathbb{R})$, inversible. Alors,

$$^t(A+B) = {^tA} + {^tB}, \quad ^t(xA) = x\,^tA, \quad ^t(A \times C) = {^tC} \times {^tA}.$$

En outre, ^tP est encore inversible et

$$\left(^tP\right)^{-1} = {^t\left(P^{-1}\right)}.$$

Preuve. Pour les trois premiers énoncés : exercice, c'est élémentaire. Pour le dernier, soit $Q = P^{-1}$. On transpose la relation $PQ = I_m$, ce qui fournit $^tQ\,^tP = I_m$. Cela montre à la fois que ^tP est inversible et que son inverse est $^tP^{-1}$. Noter en passant qu'on a utilisé ici la proposition "un inverse d'un côté est un inverse tout court" du paragraphe 4.3. ■

Corollaire (rang de la transposée)
Une matrice et sa transposée ont le même rang.

Preuve. Soient $m, n \in \mathbb{N}^*$ et $A \in \mathcal{M}_{m,n}(\mathbb{R})$. On note $r = \mathrm{rg}(A)$. On note $R_r^{(m,n)} \in \mathcal{M}_{m,n}(\mathbb{R})$ la matrice de la formule (13) ci-dessus, en en spécifiant le nombre de lignes et le nombre de colonnes en exposant. D'après la proposition "classes d'équivalence de matrices" du paragraphe 5.3.5, soient $P \in \mathcal{M}_m(\mathbb{R})$ et $Q \in \mathcal{M}_n(\mathbb{R})$, inversibles, telles que $A = PR_r^{(m,n)}Q$. Alors, ^tP et ^tQ sont encore inversibles et $^tA = {^tQ} \times {^tR_r^{(m,n)}} \times {^tP} = {^tQ} \times R_r^{(n,m)} \times {^tP}$, ce qui montre que le rang de ^tA est encore r. ■

Proposition (voir le rang d'une matrice sur ses sous-matrices inversibles)
Le rang d'une matrice est la dimension d'une sous-matrice carrée inversible de taille maximale.

Preuve. Soient $A = (A_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathcal{M}_{m,n}(\mathbb{R})$ et r son rang. (i) Puisque les colonnes de A sont dans l'image de f_A (voir (12) pour une définition de f_A), r d'entre elles forment une base de $\mathrm{im}(f_A)$: soit $(C_{i_1}, \ldots, C_{i_r})$ un r-uplet de colonnes qui forment une base de $\mathrm{im}(f_A)$. Alors, la matrice extraite (m lignes, r colonnes) dont les colonnes sont dans cet ordre $C_{i_1}, \ldots, C_{i_r}$ est encore de rang r. Sa transposée est donc également de rang r. Soient donc $(L_{j_1}, \ldots, L_{j_r})$ un r-uplet de lignes de cette matrice de $\mathcal{M}_{m,r}(\mathbb{R})$ dont les transposées forment une base de $\mathrm{im}(f_{^tA})$. Alors, la matrice carrée $(A_{i_k,j_\ell})_{1 \leq k, \ell \leq r} \in \mathcal{M}_r(\mathbb{R})$ a des lignes toutes indépendantes : elle est de rang r, donc inversible. (ii) Inversement, si A a une sous-matrice carrée inversible de dimension d, alors les d colonnes de A qui correspondent aux colonnes de cette sous-matrice forment une famille libre contenue dans $\mathrm{im}(f_A)$. Donc $d \leq r$. ■

6.3 Rang d'une famille de vecteurs

Définition (rang d'une famille de vecteurs)
Soient V un espace vectoriel réel et $v_1, \ldots, v_n \in V$. Le *rang de la famille* $\{v_1, \ldots, v_n\}$ est la dimension du sous-espace qu'elle engendre :

$$\mathrm{rg}(v_1, \ldots, v_n) = \dim_{\mathbb{R}} \mathrm{Vect}\{v_1, \ldots, v_n\}.$$

Proposition
(i) Le rang d'une famille de vecteur est le cardinal d'une sous-famille libre maximale.

(ii) Le rang d'une famille de vecteur est le cardinal d'une sous-famille génératrice minimale du sous-espace qu'elle engendre.

(iii) Le rang d'une matrice est le rang de la famille de ses vecteurs-colonne.

(iv) Le rang d'une matrice est le rang de la famille de ses vecteurs-ligne.

PREUVE. *(i)* et *(ii)* récoulent immédiatement de la définition de la dimension. *(iii)* vient de la définition du rang d'une matrice A comme étant la dimension de l'image de f_A (voir (12)) qui est engendrée par les vecteurs-colonnes de A. *(iv)* est une conséquence de *(iii)* et du fait qu'une matrice a le même rang que sa transposée. ■

6.4 Rang d'un système linéaire (homogène)

Définition (rang d'un système linéaire)
Soient $m, n \in \mathbb{N}^*$, $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B \in \mathcal{M}_{m,1}(\mathbb{R})$. Le *rang du système linéaire* écrit sous forme compacte $AX = B$ est le rang de A.

Proposition (le rang d'un système linéaire est le rang de ses équations)
Soient $m, n \in \mathbb{N}^*$. Soient $\ell_1, \ldots \ell_m$ une famille de formes linéaires sur $\mathbb{R}^n$ et $(b_1, \ldots, b_m) \in \mathbb{R}^m$. Le rang du système linéaire

$$\begin{cases} \ell_1(x_1, \ldots, x_n) = b_1 \\ \ell_2(x_1, \ldots, x_n) = b_2 \\ \qquad \vdots \\ \ell_m(x_1, \ldots, x_n) = b_m \end{cases}$$

est le rang de la famille de vecteurs $\{\ell_1, \ldots \ell_m\}$ dans l'espace dual $(\mathbb{R}^n)^* = \mathrm{Hom}(\mathbb{R}^n, \mathbb{R})$.

PREUVE. C'est une paraphrase de la définition du rang d'un système et du fait que le rang d'une matrice est le rang de la famille de ses vecteurs-ligne. ■

Proposition (le rang est le nombre de pivots)
Le rang d'un système linéaire est le nombre de pivots qui apparaissent lors de l'application de l'algorithme du pivot de Gauss.

PREUVE. Si A est la matrice du système, le nombre de pivots est le rang de l'application linéaire f_A (voir le dernier "A noter" du paragraphe 5.3.5), c'est-à-dire le rang de A. ■

Proposition (dimension de l'espace des solutions d'un système homogène)
La dimension de l'espace vectoriel des solutions d'un système linéaire homogène *est le nombre de ses inconnues auquel on soustrait le rang du système.*

PREUVE. Soient $m, n \in \mathbb{N}^*$ et $A \in \mathcal{M}_{m,n}(\mathbb{R})$. L'énoncé stipule que la dimension de l'espace vectoriel des solutions du système linéaire homogène $AX = O_m$ est $n - \mathrm{rg}(A)$. C'est le théorème du rang, puisque l'espace de ces solutions est le noyau de f_A. ■

Proposition (compatibilité d'un système linéaire avec second membre)
Un système linéaire avec second membre admet au moins une solution si, et seulement si son rang est aussi le rang de sa matrice augmentée.

Autrement dit, si un système linéaire s'écrit $AX = B$ sous forme compacte, son ensemble de solutions est non vide si, et seulement si

$$\mathrm{rg}(A) = \mathrm{rg}\left(\begin{array}{c|c} A & B \end{array}\right).$$

Cette égalité est parfois appelée *condition de Rouché-Fontené*[♫].

PREUVE. Il suffit de montrer la proposition dans le cas où $\left(\begin{array}{c|c} A & B \end{array}\right)$ est une matrice échelonnée et réduite puisque le rang de A et le rang de $\left(\begin{array}{c|c} A & B \end{array}\right)$ ne changent pas lors de l'application de l'algorithme du pivot de Gauss. Dans ces conditions, la matrice augmentée du système s'écrit sous la forme

$$\left(\begin{array}{c|c} A & \star \\ \hline \mathbf{0} & C \end{array}\right)$$

[♫]Eugène Rouché, 1832 – 1910 et Georges Fontené, 1848 – 1923.

où $\star$ est un vecteur colonne et C un vecteur colonne nul ou de la forme $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. Le système admet alors une solution si, et seulement si C est nul, ce qui équivaut à l'égalité des rangs de A et de $\left(\begin{array}{c|c} A & B \end{array} \right)$. ■

6.5 Les attendus du chapitre

Lire tout le chapitre en détail, s'assurer de tout comprendre y compris les exemples, sans laisser de zone d'ombre.

Ne pas considérer ce chapitre comme secondaire. Bien au contraire, le rang est une notion fondamentale, tant d'un point de vue opératoire que sur le plan théorique.

Bien connaître les définitions des diverses notions de rang et savoir les relier entre elles. C'est cette diversité de points de vues qui en font la richesse.

Reconnaître les énoncés opératoires qui permettent le calcul effectif du rang.

Relire tous les matins l'ensemble de tous ces chapitres qui fondent l'algèbre linéaire. S'assurer de ne pas en perdre un miette et de savoir faire tous les exercices des listes travaillées en TD et des énoncés d'examens passés et, surtout, à venir. Se régaler à inventer des exercices et à les résoudre. Prendre goût à ce paysage mathématique – et aussi aux autres. S'assurer de savoir raconter le détail de cet enthousiasme avec rigueur et avec une justesse d'argumentation sans faille, même à trois heures du matin avec trente-neuf degrés de fièvre. Alors, et seulement alors, cueillir le fruit si gratifiant du lourd travail fourni et en jouir sans retenue.

Fin du cours