MAT 201: Introduction à l'algèbre linéaire

Jean Fasel   Thierry Gallay

Contents

1 Structures algébriques usuelles

On suppose connues la définition et les propriétés principales des ensembles de nombres $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ et $\mathbb{C}$. Si E et F sont deux ensembles, le produit cartésien de E et F est l'ensemble

$$E \times F := \{(e, f) |\ e \in E, f \in F\}.$$

1.1 Lois de composition internes

Définition 1.1.1. Soit E un ensemble. On appelle *loi de composition interne* dans E (parfois notée l.c.i) une application

$$\star : E \times E \to E$$

De manière plus précise, une loi de composition interne associe à toute paire (a, b) d'éléments de E un nouvel élément de E noté $a \star b$. On appelle parfois *opération* sur E une loi de composition interne sur E.

Exemple 1.1.2. L'addition $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ définie par $(n, m) \mapsto n + m$ est une loi de composition interne sur $\mathbb{N}$. La même observation est vraie dans $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$. On peut également vérifier que la multiplication $\cdot : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ est une loi de composition interne (idem en remplaçant $\mathbb{N}$ par un ensemble de nombres comme ci-dessus).

Exemple 1.1.3. On peut également considérer des exemples plus exotiques sur $E = \mathbb{R}$.

1. $\star : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ définie par $a \star b = a$ pour tout $a, b \in \mathbb{R}$.
2. $\star : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ définie par $a \star b = \max\{a, b\}$ pour tout $a, b \in \mathbb{R}$.

Définition 1.1.4. Si $\star$ est une loi de composition interne sur un ensemble E, on dit que:

1. la loi $\star$ est *commutative* si $a \star b = b \star a$ pour tout $a, b \in E$.
2. la loi est *associative* si $a \star (b \star c) = (a \star b) \star c$ pour tout $a, b, c \in E$.

On dit par ailleurs que $e \in E$ est un *élément neutre* si $a \star e = e \star a = a$ pour tout $a \in E$.

Exemple 1.1.5. L'addition considérée dans l'exemple 1.1.2 pour $E = \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ est commutative et associative. Elle admet $0 \in \mathbb{N}$ comme élément neutre. De manière similaire, la multiplication introduite dans le même exemple est commutative et associative, avec élément neutre $1 \in \mathbb{N}$. Par contre la première loi de composition $\star : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ de l'exemple 1.1.3 n'est pas commutative: si $a, b \in \mathbb{R}$ sont tels que $a \neq b$, on a $a \star b = a \neq b = b \star a$. On vérifie par contre qu'elle est associative puisque $a \star (b \star c) = a \star b = a$ et $(a \star b) \star c = a \star c = a$ pour tout $a, b, c \in \mathbb{R}$. La seconde loi de composition de l'exemple 1.1.3 est commutative, associative, mais ne possède pas d'élément neutre.

Remarque 1.1.6. Si la loi $\star$ admet un élément neutre, alors celui-ci est unique. Supposons en effet que e, e' soient des éléments neutres pour $\star$. On a alors

$$e = e' \star e = e',$$

les deux égalités venant du fait que ces éléments sont neutres.

Définition 1.1.7. Soit E un ensemble muni d'une loi de composition interne $\star$. On dit qu'un sous-ensemble $F \subset E$ est *stable par* $\star$ si $a \star b \in F$ pour tout $a, b \in F$. On observe que dans ce cas-là, la loi de composition $\star$ induit une loi de composition $\star_F$ sur F définie par $a \star_F b = a \star b$ pour tout $a, b \in F$. On note souvent $\star$ pour $\star_F$ par un léger abus de notation, et on dit que $\star_F$ est la *loi de composition interne induite par* $\star$.

Exemple 1.1.8. Le sous ensemble $\mathbb{N} \subset \mathbb{Z}$ est stable pour les lois $+$ et $\cdot$. De même le sous ensemble $\mathbb{R}_+ = \{x \in \mathbb{R} | x \geq 0\} \subset \mathbb{R}$ est stable pour l'addition est la multiplication. Enfin, on observe que $\mathbb{R}_- = \{x \in \mathbb{R} | x \leq 0\} \subset \mathbb{R}$ est stable par addition, mais pas par multiplication.

1.2 Groupes

Définition 1.2.1. On appelle *groupe* un ensemble G muni d'une loi de composition interne $\star$ possédant les propriétés suivantes:

1. la loi $\star$ est associative, i.e. $a \star (b \star c) = (a \star b) \star c \ \ \forall a, b, c \in G$.
2. Il existe un élément neutre $e \in G$: $a \star e = e \star a = a \ \ \forall a \in G$.
3. Tout élément $a \in G$ admet un élément symétrique $a' \in G$, i.e. satisfaisant la propriété que $a \star a' = a' \star a = e$.

On dit que le groupe est *commutatif*, ou *abélien*, si de plus

4. la loi $\star$ est commutative: $a \star b = b \star a \ \ \forall a, b \in G$.

Remarque 1.2.2. Par abus de language, on dit parfois que G est un groupe, en omettant la loi $\star$. Cela peut prêter à confusion dans le sens que le même ensemble G peut être muni de deux lois de composition différentes $\star$ et $\star'$ qui en font un groupe (on verra un exemple dans les TDs). Par la suite, on ne fera mention de $\star$ que lorsque une confusion est possible.

Remarque 1.2.3. On remarque que l'élément symétrique du point 3. est unique. En effet, supposons que a' et a'' soient deux symétriques de a. Alors,

$$a' = a' \star e = a' \star (a \star a'') = (a' \star a) \star a'' = e \star a'' = a''$$

où l'on a utilisé respectivement le fait que e est un élément neutre, que a'' est un symétrique de a, l'associativité et le fait que a' est un symétrique de a. Puisqu'on parle d'élément symétrique, on notera que parfois cet élément est appelé *inverse*. On utilisera souvent cette terminologie, même si elle peut a priori porter à confusion.

Exemple 1.2.4 (groupes des nombres additifs). L'ensemble $\mathbb{Z}$ muni de l'addition usuelle est un groupe abélien. Son élément neutre est 0 est le symétrique d'un élément $n \in \mathbb{Z}$ est son opposé $-n \in \mathbb{Z}$. Si $n, m \in \mathbb{Z}$, on pose $m - n = m + (-n)$. De même, le ensembles $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ munis de l'addition sont des groupes commutatifs, avec $0 \in \mathbb{Z}$ comme élément neutre. On observe par contre que $\mathbb{N}$ muni de l'addition n'est *pas* un groupe. En effet, si $n \in \mathbb{N}$ est non nul alors on a $n + m > 0$ pour tout $m \in \mathbb{N}$.

Exemple 1.2.5. On considère l'ensemble $\mathbb{Q}^* := \{x \in \mathbb{Q} | x \neq 0\}$, qu'on munit de la multiplication usuelle comme loi de composition interne. C'est un groupe abélien, avec $1 \in \mathbb{Q}^*$ comme élément neutre et symétrique d'un élément $x = \frac{p}{q}$ l'élément $x^{-1} = \frac{q}{p}$ (avec $p, q \in \mathbb{Z}$ non nuls). De manière consistante avec l'exemple précédent, on pose $\frac{x}{y} := xy^{-1}$ pour tout $x, y \in \mathbb{Q}^*$. De même, les ensembles $\mathbb{R}^* := \mathbb{R} \setminus \{0\}, \mathbb{C}^* = \mathbb{C} \setminus \{0\}$ sont des groupes abéliens pour la multiplication.

Exemple 1.2.6. Si E est un ensemble (non vide), on considère l'ensemble $S(E)$ des bijections de E dans lui même. Si $f : E \to E$ et $g : E \to E$ sont des éléments de $S(E)$, alors la composition $f \circ g$ est encore une bijection de E dans lui-même, et on obtient ainsi une loi de composition interne $\star$ sur $S(E)$ définie par $f \star g = f \circ g$. On vérifie immédiatement que $\star$ est associative, puis que l'identité $\mathrm{Id} : E \to E$ est un élément neutre pour $\star$. Le symétrique de $f \in S(E)$ est la bijection inverse f^{-1} puisque $f \star f^{-1} = f \circ f^{-1} = \mathrm{Id} = f^{-1} \circ f = f^{-1} \star f$. On appelle $S(E)$ (muni de la composition) *groupe symétrique* de E et tout élément $f \in S(E)$ *une permutation de* E. Ce groupe n'est pas abélien si E possède au moins 3 éléments, comme on le verra dans l'exemple suivant.

Exemple 1.2.7. Supposons que $E = \{1, \ldots, n\}$ pour $n \in \mathbb{N}$. Dans ce cas, on note S_n le groupe $S(E)$. Si $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ est une permutation on peut la noter

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

On observe que l'ensemble S_n est de cardinal $n!$. Dans le cas particulier $n = 3$, on peut en particulier considérer les éléments

$$\tau_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}; C_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

On a alors

$$\tau_1 \circ C_1 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = C_1 \circ \tau_1.$$

Si $n \geq 4$, on peut voir τ_1 et C_1 comme des permutations de $\{1, \ldots, n\}$ (fixant $4, \ldots, n$) et le même calcul montre que S_n n'est pas abélien.

Remarque 1.2.8. Les remarques suivantes sont vérifiées:

- Un groupe $(G, \star)$ est toujours non vide puisque $e \in G$.
- Pour tout $a \in G$, on a une équivalence pour tout $x, y \in G$

  $$x = y \iff a \star x = a \star y.$$

  En effet, il suffit de composer à gauche avec a pour passer de gauche à droite, et avec a' pour passer de droite à gauche. De même, on a

  $$x = y \iff x \star a = y \star a$$

- Quels que soient $a, b \in G$, l'équation $a \star x = b$ d'inconnue x a une unique solution, qui est $x = a' \star b$. Par ailleurs, l'unique solution de $x \star a = b$ est $x = b \star a'$.
- Pour tout $a, b \in G$, on a $(a \star b)' = b' \star a'$ et $(a')' = a$.

1.3 Sous-groupes et morphismes de groupes

Définition 1.3.1. Soit $(G, \star)$ un groupe. Un sous-ensemble $H \subset G$ est un sous-groupe de G si les assertions suivantes sont satisfaites:

i) $e \in H$.

ii) F est stable par $\star$, i.e. $a \star b \in F$ pour tout $a, b \in F$.

iii) pour tout $a \in H$, l'élément symétrique a' appartient à H.

Remarque 1.3.2. Les conditions signifient exactement que H, muni de la loi $\star$ induite, est un groupe.

Proposition 1.3.3. *Un sous-ensemble non vide $H \subset G$ est un sous-groupe si et seulement si*

$$a \star b' \in H \ \ \forall a, b \in H. \tag{1.3.1}$$

Proof. Supposons tout d'abord que H soit un sous groupe. Si $a, b \in H$, alors iii) implique que $b' \in H$. D'autre part, ii) donne que $ab' \in H$ et on en conclut que (1.3.1) est vérifiée. Supposons réciproquement que cette équation soit vérifiée. Du fait que H est non vide, il existe $a \in H$. Donc $e = a \star a' \in H$ et i) est vérifiée. Utilisant maintenant (1.3.1) avec e et a, on en déduit que $a' = e \star a' \in H$ et donc que iii) est valide. Finalement, si $a, b \in H$, alors $b' \in H$ comme on l'a vu ci-dessus et ainsi $a \star b = a \star (b')' \in H$. □

Remarque 1.3.4. Un groupe G possède toujours au moins deux sous-groupes: $\{e\}$ et G. Les sous-groupes H de G satisfont $\{e\} \subset H \subset G$ et on dit que H est *non trivial* si les inclusions sont strictes. Si H_1 et H_2 sont deux sous-groupes de G, alors $H_1 \cap H_2$ est un sous-groupe, et il en est de même pour une intersection quelconque de sous-groupes de G.

Exemple 1.3.5. Soit $(G, \star) = (\mathbb{Z}, +)$ le groupe des entiers relatifs muni de l'addition usuelle. Pour $n \in \mathbb{N}$ un entier, on peut considérer $n\mathbb{Z} := \{nx | x \in \mathbb{Z}\} \subset \mathbb{Z}$. On vérifie que $n\mathbb{Z}$ est un sous-groupe de $\mathbb{Z}$ en utilisant la proposition ci-dessus. Tout d'abord, $0 = n \cdot 0 \in n\mathbb{Z}$ qui est donc non vide. Ensuite, si $nx, ny \in n\mathbb{Z}$, alors $(ny)' = -ny$ et il s'ensuit que $(nx) \star (ny)' = nx + (-ny) = nx + n(-y) = n(x - y) \in n\mathbb{Z}$. On en déduit que $n\mathbb{Z}$ est bien un sous-groupe de $\mathbb{Z}$.

Considérons maintenant le groupe $(\mathbb{R}, +)$ et le sous-ensemble $\mathbb{Z} \subset \mathbb{R}$. Evidemment, il est non vide puis on remarque que si $m, n \in \mathbb{Z}$ alors $n' = -n$ et

$$m \star n' = m + (-n) = m - n \in \mathbb{Z}.$$

Ainsi, $\mathbb{Z}$ est un sous groupe de $(\mathbb{R}, +)$. La loi induite sur $\mathbb{Z}$ est la loi $+$ d'addition usuelle, et on déduit de ces deux exemples que $n\mathbb{Z} \subset \mathbb{R}$ est un sous-groupe de $(\mathbb{R}, +)$. Pour information (mais ce ne sera pas utilisé dans le cours), on peut démontrer que tout sous-groupe discret de $(\mathbb{R}, +)$ est de la forme $a\mathbb{Z}$ *avec* $a \in \mathbb{R}$.

Exemple 1.3.6. On considère le groupe $(G, \star) = (\mathbb{C}^*, \cdot)$ et le sous-ensemble $U := \{z \in \mathbb{C} | |z| = 1\} \subset \mathbb{C}^*$. Du fait que $1 \in U$, on voit que U est non vide. Ensuite, si $z_1, z_2 \in \mathbb{C}^*$ alors $z_1 \star z_2' = z_1 z_2^{-1}$ et on a

$$|z_1 z_2^{-1}| = |z_1| \cdot |z_2^{-1}| = \frac{|z_1|}{|z_2|} = 1.$$

Il s'ensuit que U est bien un sous-groupe de $\mathbb{C}^*$.

Dans le même veine, étant donné $n \in \mathbb{N}^*$ on peut définir $U_n := \{z \in \mathbb{C} | z^n = 1\} \subset \mathbb{C}^*$ (l'ensemble des racines n-ièmes de l'unité). On sait d'après le cours de MAT 101 qu'en fait

$$U_n = \{e^{\frac{ik\pi}{n}} | k = 0, \dots, n-1\}$$

et on vérifie que U_n est bien un sous-groupe de $\mathbb{C}^*$ en utilisant les mêmes arguments que ci-dessus.

Définition 1.3.7. Soient $(G, \star)$ et $(\tilde{G}, \tilde{\star})$ deux groupes. On dit qu'une application $f : G_1 \to G_2$ est un *morphisme de groupes* si

$$f(a \star b) = f(a) \tilde{\star} f(b)$$

pour tout $a, b \in G$.

Remarque 1.3.8. De manière informelle, on voit que l'application f est un morphisme de groupe si elle transforme la loi de G en la loi de $\tilde{G}$.

Exemple 1.3.9. Soient $(G, \star) = (\mathbb{R}, +)$ et $(\tilde{G}, \tilde{\star}) = (\mathbb{R}^*, \cdot)$ et $f : \mathbb{R} \to \mathbb{R}^*$ donnée par $f(x) = e^x$. Du fait que $e^{a+b} = e^a \cdot e^b$ pour tout $a, b \in \mathbb{R}$, on voit que f est bien un morphisme de groupes. En effet,

$$f(a \star b) = f(a+b) = e^{a+b} = e^a \cdot e^b = f(a) \cdot f(b) = f(a) \tilde{\star} f(b) \;\; \forall a, b \in \mathbb{R}.$$

Les mêmes arguments montrent que l'exponentielle complexe $f : \mathbb{C} \to \mathbb{C}^*$ définie par $f(z) = \exp(z)$ est bien un morphisme de groupes.

Proposition 1.3.10. *Si $f : G \to \tilde{G}$ est un morphisme de groupes, alors:*

i) $f(e) = \tilde{e}$, où e (respectivement $\tilde{e}$) est l'élément neutre pour la loi $\star$ (respectivement $\tilde{\star}$).

ii) Pour tout $a \in G$, on a $f(a)' = f(a')$.

Proof. Pour la première assertion, on a

$$f(e) = f(e \star e) = f(e) \tilde{\star} f(e).$$

On en déduit que

$$\tilde{e} = f(e) \tilde{\star} f(e)' = (f(e) \tilde{\star} f(e)) \tilde{\star} f(e)' = f(e) \tilde{\star} (f(e) \tilde{\star} f(e)') = f(e) \tilde{\star} \tilde{e} = f(e).$$

En ce qui concerne le second point, on trouve

$$\tilde{e} = f(e) = f(a \star a') = f(a) \tilde{\star} f(a')$$

et de la même manière $\tilde{e} = f(a') \tilde{\star} f(a)$. Ces deux expressions montrent que $f(a')$ est l'élément symétrique de $f(a)$ pour la loi $\tilde{\star}$, et on en déduit que $f(a') = f(a)'$. $\square$

Définition 1.3.11. Soient $(G, \star), (\tilde{G}, \tilde{\star})$ deux groupes et $f : G \to \tilde{G}$ un homomorphisme de groupes. On dit que:

1. f est un endomorphisme si $(G, \star) = (\tilde{G}, \tilde{\star})$.

2. f est un isomorphisme si f est bijectif.

3. f est un automorphisme si $(G,\star) = (\tilde{G},\tilde{\star})$ et f est bijectif.

Définition 1.3.12. Soient $(G,\star)$, $(\tilde{G},\tilde{\star})$ deux groupes et $f : G \to \tilde{G}$ un homomorphisme de groupes.

a) On appelle *noyau de* f et on note $\ker(f)$ le sous-ensemble de G défini par

$$\ker(f) := \{a \in G | f(a) = \tilde{e}\} = f^{-1}(\tilde{e}).$$

b) On appelle *image de* f et on note $\mathrm{Im}(f)$ le sous-ensemble défini par

$$\mathrm{Im}(f) := \{f(a) \in \tilde{G} | a \in G\} = f(G).$$

Remarque 1.3.13. Le sous-ensemble $\ker(f)$ est un sous-groupe de G, et le sous-ensemble $\mathrm{Im}(f)$ est un sous-groupe de $\tilde{G}$. Pour la première assertion, on observe tout d'abord que $e \in \ker(f)$ puisque $f(e) = \tilde{e}$ par la proposition 1.3.10. Si $a, b \in \ker(f)$ alors on a

$$f(a \star b') = f(a)\tilde{\star}f(b') = f(a)\tilde{\star}f(b)'$$

par la proposition 1.3.10. On en tire que $f(a \star b') = f(a)\tilde{\star}f(b)' = \tilde{e}\tilde{\star}(\tilde{e})' = \tilde{e}$, montrant que $a \star b' \in \ker(f)$ qui est donc un sous-groupe. On prouve maintenant que $\mathrm{Im}(f)$ est un sous-groupe de $\tilde{G}$. Pour ce faire, on observe en premier lieu que $\tilde{e} = f(e)$ et donc que $\tilde{e} \in \mathrm{Im}(f)$. Si maintenant $a, b \in G$ alors on doit démontrer que $f(a)\tilde{\star}f(b)' \in \mathrm{Im}(f)$. Cela suit de la proposition 1.3.10 qui donne le calcul

$$f(a)\tilde{\star}f(b)' = f(a)\tilde{\star}f(b') = f(a \star b') \in \mathrm{Im}(f).$$

Remarque 1.3.14. Le noyau et l'image permettent de comprendre certaines propriétés de f. Ainsi, f est injective si et seulement si $\ker(f) = \{e\}$ et f est surjective si et seulement si $\mathrm{Im}(f) = \tilde{G}$. Montrons le premier point, en commençant par supposer que f est injective. Tout d'abord on tire de la remarque précédente que $f(e) = \tilde{e}$. Si $a \in G$ est tel que $f(a) = \tilde{e}$, alors on obtient $f(a) = f(e)$ d'où $a = e$ par injectivité. Bien évidemment, $e \in \ker(f)$ et on en tire que $\ker(f) = \{e\}$. Réciproquement, supposons que $\ker(f) = \{e\}$. Si $a, b \in G$ sont tels que $f(a) = f(b)$ alors on obtient

$$\tilde{e} = f(a)\tilde{\star}f(a)' = f(b)\tilde{\star}f(a)' = f(b \star a')$$

par la proposition 1.3.10. Donc $b \star a' \in \ker(f)$, et ainsi $b \star a' = e$. Ceci permet de conclure, au vu de

$$a = e \star a = (b \star a') \star a = b \star e = b.$$

Le fait que f est surjective si et seulement si $\mathrm{Im}(f) = \tilde{G}$ est laissé en exercice.

Exemple 1.3.15. L'homomorphisme $f : \mathbb{R} \to \mathbb{R}^*$ de l'exemple 1.3.9 est un isomorphisme, donc son noyau est trivial $\ker(f) = \{0\}$ et $\mathrm{Im}(f) = \mathbb{R}^*$. Considérons maintenant l'exponentielle complexe $\exp : \mathbb{C} \to \mathbb{C}^*$, qui est définie comme suit: si $z = a + bi$ avec $a, b \in \mathbb{R}$, on a $\exp(z) = e^a e^{ib}$, ce qui montre que $\mathrm{Im}(\exp) = \mathbb{C}^*$. D'un autre côté, on trouve

$$\begin{aligned} \exp(z) = 1 \quad &\iff \quad e^a = 1 \text{ et } e^{ib} = \cos(b) + i\sin(b) = 1 \\ &\iff \quad a = 0 \text{ et } b = 2\pi n, n \in \mathbb{Z} \\ &\iff \quad z \in 2\pi i\mathbb{Z}. \end{aligned}$$

Ceci donne donc $\ker(\exp) = 2\pi i\mathbb{Z} \subset \mathbb{C}$.

1.4 Anneaux et corps

Définition 1.4.1. On appelle *anneau* un ensemble A muni de deux lois de compositions internes, notées $+$ et $\cdot$, telles que:

i) $(A,+)$ est un groupe abélien, dont l'élément neutre est noté 0.

ii) la loi $\cdot$ est associative et possède un élément neutre noté 1.

iii) La loi $\cdot$ est distributive par rapport à la loi $+$: pour tout $a,b,c \in A$ on a

$$a\cdot(b+c) = a\cdot b + a\cdot b, \text{ et } (b+c)\cdot a = b\cdot a + c\cdot a.$$

On dit que l'anneau est commutatif si de plus la loi $\cdot$ est commutative.

Remarque 1.4.2. Si A est un anneau (avec lois $+$ et $\cdot$), alors on a $a\cdot 0 = 0\cdot a = 0$ pour tout $a\in A$. En effet,

$$a\cdot 0 = a\cdot(0+0) = a\cdot 0 + a\cdot 0$$

et on peut ajouter $-(a\cdot 0)$ des deux côtés de cette équation pour en déduire que $a\cdot 0$. Le même argument permet de montrer que $0\cdot a = 0$.

Par contre, il est possible de trouver $a\cdot b = 0$ avec $a,b\neq 0$ dans A (voir l'exemple 1.4.4 ci-dessous).

Exemple 1.4.3. L'ensemble $\mathbb{Z}$ muni de l'addition et de la multiplication usuelle est un anneau commutatif. Il en est de même pour $\mathbb{Q},\mathbb{R}$ et $\mathbb{C}$.

Exemple 1.4.4. Soit $n\in\mathbb{N}$ tel que $n\geq 2$. On rappelle que tout entier $k\in\mathbb{N}$ s'écrit de manière unique $k = qn + r$ avec $0\leq r < n$. On dit que q est le quotient de k par n, et r est le reste de la division de k par n. On pose par la suite $r = k \pmod n$. Soit $A = \{0,1,\ldots,n-1\}$, qu'on munit de deux lois de la manière suivante:

- L'addition: $A\times A\to A$ définie par $k+l = k+l \pmod n$ pour tout $k,l\in A$.
- La multiplication $A\times A\to A$ définie par $k.l = kl \pmod n$ pour tout $k,l\in A$.

On vérifie que A est un anneau commutatif, qu'on note $\mathbb{Z}/n\mathbb{Z}$. Il est conseillé de dresser les tables de multiplication et d'addition dans le cas $n = 4,5$ pour bien comprendre ces lois. Dans la cas où $n = 4$, on voit que $2\in A$ satisfait $2\cdot 2 = 0$.

Définition 1.4.5. Soit A un anneau, dont les lois sont notées $+$ et $\cdot$ comme d'habitude. On dit qu'un élément $a\in A$ est *inversible* s'il existe un élément $b\in A$ tel que $a\cdot b = b\cdot a = 1$. On dit alors que b est l'inverse de a.

Remarque 1.4.6. En d'autres termes, b est l'élément symétrique de a pour la loi $\cdot$. De ce fait, il est unique s'il existe.

Remarque 1.4.7. Si A est comme ci-dessus, alors on peut considérer l'ensemble $A^\times := \{a\in A |\ a \text{ est inversible}\}$. On vérifie facilement que $A^\times$ est un groupe pour la loi $\cdot$.

Exemple 1.4.8. Si $\mathbb{Z}$ est muni de l'addition et de la multiplication usuelle comme dans l'exemple 1.4.3, alors on trouve $\mathbb{Z}^\times = \{\pm 1\}$. Pour $A = \mathbb{Q},\mathbb{R},\mathbb{C}$, on a $A^\times = A\setminus\{0\}$ dans tous les cas.

Définition 1.4.9. On anneau commutatif $(A, +, \cdot)$ est appelé un corps si $A \neq \{0\}$ et si tout élément de $A \setminus \{0\}$ est inversible, i.e. si $A^\times = A \setminus \{0\}$.

Par exemple, $\mathbb{Q}, \mathbb{R}$ et $\mathbb{C}$ sont des corps au vu de l'exemple ci-dessus.

Exemple 1.4.10. On peut démontrer (mais ce ne sera pas utilisé dans ce cours) que $\mathbb{Z}/n\mathbb{Z}$ est un corps si et seulement si $n \geq 1$ est un nombre premier.

2 Espaces vectoriels

2.1 Définitions et premiers exemples

Les espaces vectoriels sont des ensembles dont les éléments, appelés *vecteurs*, peuvent être additionnés entre eux, ainsi que multipliés par les éléments d'un corps appelés *scalaires*. La définition formelle est la suivante:

Définition 2.1.1. Soit $\mathbb{K}$ un corps. On appelle *espace vectoriel sur* $\mathbb{K}$ un ensemble E muni de:

a) une loi de composition interne $+ : E \times E \to E$, $(v, w) \mapsto v + w$,

b) une loi externe $\cdot : \mathbb{K} \times E \to E$, $(\lambda, v) \mapsto \lambda \cdot v$,

possédant les propriétés suivantes:

1. $(E, +)$ est un groupe abélien,
2. la loi externe est distributive par rapport à la loi interne dans le sens que
$$(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$$
pour tout $\lambda, \mu \in \mathbb{K}$ et tout $v \in E$,
3. la loi externe satisfait pour tout $\lambda, \mu \in \mathbb{K}$ et tout $v \in E$
$$(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v)$$
4. On a $1 \cdot v = v$ pour tout $v \in E$.

Remarque 2.1.2. Comme mentionné dans la brève introduction, les éléments de E sont appelés *vecteur*, et ceux de $\mathbb{K}$ sont appelés *scalaires*. Par ailleurs, la loi interne est notée additivement, et son élément neutre est de fait noté 0_E (ou même parfois 0 si aucune confusion n'est passible). On appelle 0_E le vecteur nul. L'élément symétrique d'un vecteur v est noté $-v$.

Remarque importante 2.1.3. Le corps $\mathbb{K}$ possède lui deux lois de composition internes, malheureusement notées avec les mêmes symboles que la loi interne de E et la loi externe $\mathbb{K} \times E \to E$. L'élément neutre de $\mathbb{K}$ pour l'addition, noté 0, ne doit pas être confondu avec le vecteur nul noté 0_E.

Remarque 2.1.4. Dans ce qui suit, on prendra presque dans tous les cas $\mathbb{K} = \mathbb{R}$ avec les lois usuelles. Il faut néanmoins remarquer que les propriétés développées dans ce cours sont vraies pour un corps $\mathbb{K}$ quelconque.

Proposition 2.1.5. *Pour tout $\lambda \in \mathbb{K}$ et tout $v \in E$, on a:*

i) $0 \cdot v = 0_E$, $(-1) \cdot v = -v$.

ii) $\lambda \cdot 0_E = 0_E$.

Proof. Pour i), on a
$$0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v.$$
On en déduit, additionnant le symétrique $-0 \cdot v$ à gauche et à droite, que $0 \cdot v = 0_E$. En conséquence, on trouve
$$0_E = 0 \cdot v = (1 + (-1)) \cdot v = 1 \cdot v + (-1) \cdot v = v + (-1) \cdot v$$
et on en déduit que $(-1) \cdot v = -v$. Pour montrer ii), on utilise un argument similaire:
$$\lambda \cdot 0_E = \lambda \cdot (0_E + 0_E) = \lambda \cdot 0_E + \lambda \cdot 0_E$$
d'où on déduit que $\lambda \cdot 0_E = 0_E$. □

Notation 2.1.6. On omet le plus souvent le signe $\cdot$, que ce soit pour le produit dans $\mathbb{K}$ ou dans la loi externe. Par exemple, le point 3. dans la définition 2.1.1 s'écrit
$$(\lambda\mu)v = \lambda(\mu v).$$

Exemple 2.1.7. Soit $n \in \mathbb{N}$ tel que $n \geq 1$. On pose $E = \mathbb{R}^n$ et on note en colonne les éléments de E de la manière suivante:
$$v \in E \iff v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \text{ où } v_1, \ldots, v_n \in \mathbb{R}.$$
Les scalaires $v_1, \ldots, v_n$ sont les composantes du vecteur v. L'addition et la multiplication externes sont définies comme suit: si $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n, w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$ et $\lambda \in \mathbb{R}$ alors
$$v + w = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} \text{ et } \lambda v = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \vdots \\ \lambda v_n \end{pmatrix}$$
On vérifie que ces opérations font de $\mathbb{R}^n$ un $\mathbb{R}$-espace vectoriel (on dit aussi que $\mathbb{R}^n$ est un espace vectoriel réel). On notera par ailleurs que lorsque $n = 1$ alors $E = \mathbb{R}$.

Exemple 2.1.8. Soit X un ensemble et $E := \{f \colon X \to \mathbb{R}\}$ l'ensemble des applications de X dans $\mathbb{R}$. On définit les opérations suivantes:

- (addition) Pour tout $f, g \in E$, on définit l'application $f+g$ par $(f+g)(x) = f(x) + g(x)$ pour tout $x \in X$.
- (loi externe) Pour tout $f \in E$ et tout $\lambda \in \mathbb{R}$ on définit l'application λf par $(\lambda f)(x) = \lambda f(x)$ pour tout $x \in X$.

On vérifie que E est un espace vectoriel réel, noté parfois $\mathbb{R}^X$. Si $X = \{1, \ldots, n\}$ pour un certain $n \in \mathbb{N}$, on note que $\mathbb{R}^{\{1,\ldots,n\}}$ coïncide avec l'espace vectoriel $\mathbb{R}^n$ de l'exemple précédent.

Exemple 2.1.9. Soit E l'ensemble des suites réelles $(a_n)_{n \in \mathbb{N}}$ à support fini, i.e. il existe $d \in \mathbb{N}$ tel que $a_n = 0$ si $n > d$. Si $a \in E$, alors on a

$$a = (a_0, a_1, \ldots, a_d, 0, 0, \ldots) \in \mathbb{R}^{\mathbb{N}}.$$

L'entier $d \in \mathbb{N}$ tel que a_d soit le dernier coefficient non nul dépend de a. Si maintenant $b \in E$ on aura

$$b = (b_0, b_1, \ldots, b_m, 0, 0, \ldots) \in \mathbb{R}^{\mathbb{N}}$$

avec en général $d \neq m$. On définit trois opérations de la manière suivante:

- (addition) Si a et b sont comme ci-dessus, alors on définit $a+b$ par $(a+b)_n = a_n + b_n$ pour tout $n \in \mathbb{N}$. Si $N > \max d, m$, alors $(a+b)_n = 0$ et on a donc bien $a + b \in E$.
- (loi externe) Si a est comme ci-dessus et $\lambda \in \mathbb{R}$, on définit λa par $(\lambda a)_n = \lambda a_n$ pour tout $n \in \mathbb{N}$. Comme ci-dessus, on vérifie facilement que $\lambda a \in E$.
- (multiplication) Si a et b sont comme ci-dessus, on définit $a \cdot b$ par

  $$a \cdot b = (a_0 b_0, a_0 b_1 + a_1 b_0, a_0 b_2 + a_1 b_1 + a_2 b_0, \ldots),$$

  c'est à dire $(a \cdot b)_n = \sum_{i=0}^{n} a_i b_{n-i}$.

On vérifie ensuite que E muni de l'addition et de la loi externe ci-dessus est un espace vectoriel sur $\mathbb{R}$. De la même manière, on vérifie que E muni de l'addition et la multiplication ci-dessus est un anneau commutatif. Les éléments neutres de l'addition et de la multiplication sont respectivement

$$0 = (0, 0, \ldots) \text{ et } 1 = (1, 0, 0, \ldots).$$

On note $X = (0, 1, 0, 0, \ldots)$ et on vérifie que $X^2 = (0, 0, 1, 0, 0, \ldots)$, $X^3 = (0, 0, 0, 1, 0, 0, \ldots)$, etc. Ainsi, tout élément $a \in E$ s'écrit de manière unique

$$a = (a_0, a_1, a_2, \ldots, a_d, 0, \ldots) = a_0 1 + a_1 X + a_2 X^2 + \ldots + a_d X^d = \sum_{n=0}^{d} a_n X^n.$$

On dit que a est un *polynôme a une indéterminée à coefficients dans* $\mathbb{R}$ et on note $E = \mathbb{R}[X]$.

Remarque 2.1.10. Si $P = \sum_{n=0}^{d} a_n X^n$ vérifie $a_d \neq 0$, on dit que P est un *polynôme de degré* d, et on note $\deg(P) = d$. Par convention, on pose $\deg(0) = -\infty$. On a alors:

- $\deg(P + Q) \leq \max\{\deg(p), \deg(Q)\}$.
- $\deg(\lambda P) = \deg(P)$ si $\lambda \in \mathbb{R}$ est non nul.
- $\deg(PQ) = \deg(P) + \deg(Q)$.

2.2 Sous-espaces vectoriels

Définition 2.2.1. Soit E un espace vectoriel (réel). On dit qu'un sous-ensemble $F \subset E$ est un *sous-espaces vectoriel* de E si:

i) F n'est pas vide.

ii) F est stable pour l'addition: pour tous $v, w \in F$ on a $v + w \in F$.

iii) F est stable pour la multiplication par un scalaire: pour tout $\lambda \in \mathbb{R}$ et pour tout $v \in F$ on a $\lambda v \in F$.

Remarque 2.2.2. Si $F \subset E$ est un sous-espace vectoriel alors F, muni des lois induites par celles de E, est un espace vectoriel (réel). En effet, on obtient une loi de composition interne

$$+ \colon F \times F \to F$$

donnée par $(v, w) \mapsto v + w$. Si $v, w \in F$, on a que $(-1)w \in F$ d'après la propriété iii) ci-dessus. Du fait que $(-1)w = -w$ par la proposition 2.1.5, on en déduit que $-w \in F$ et on tire de ii) que $v - w \in F$. Comme F est non-vide, on voit que F est donc un sous-groupe de E (et donc en particulier un groupe abélien). Le fait que la loi externe satisfait toutes les propriétés suit immédiatement de iii) et du fait que E est un espace vectoriel.

Exemple 2.2.3. Soit E un espace vectoriel réel. Alors $\{0_E\} \subset E$ est un sous-espace vectoriel. De même, $E \subset E$ en est un. Ce sont respectivement le plus petit sous-espace vectoriel de E et le plus grand.

Exemple 2.2.4. Soit $E = \mathbb{R}^3$, et

$$F := \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} ; x_3 = 0 \right\}.$$

On vérifie que F est bien un sous-espace vectoriel de E. Dans la même veine, soit E l'ensemble des applications de $\mathbb{R}$ dans $\mathbb{R}$ muni de la structure d'espace vectoriel réel défini dans l'exemple 2.1.8. Alors le sous-ensemble F des fonctions *continues* de $\mathbb{R}$ dans $\mathbb{R}$ est un sous-espace vectoriel.

Proposition 2.2.5. *Soit E un espace vectoriel, et $(F_i)_{i \in I}$ une famille quelconque de sous-espaces vectoriels de E. Alors*

$$F := \cap_{i \in I} F_i$$

est un sous-espace vectoriel de E.

Proof. On sait d'après la remarque 2.2.2 que $F_i \subset E$ est un sous-groupe, et en particulier $0_E \in F_i$ pour tout $i \in I$. On en déduit que $0_E \in F$ qui est non vide. On montre maintenant que F est stable par addition. Si $v, w \in F$, alors $v, w \in F_i$ pour tout $i \in I$. Du fait que ce sont des sous-espaces vectoriels, on obtient que $v + w \in F_i$ pour tout $i \in I$. Ainsi, $v + w \in F$ et la seconde propriété d'un sous-espace vectoriel est satisfaite. Si maintenant $\lambda \in \mathbb{R}$ et $v \in F$, alors $v \in F_i$ pour tout $i \in I$ et donc $\lambda v \in F_i$ montrant finalement que $\lambda v \in F$. $\square$

Si l'intersection d'une famille quelconque de sous-espaces vectoriels est encore un sous-espace vectoriel, il n'en est pas de même pour une réunion.

Proposition 2.2.6. *Soient F_1 et F_2 deux sous-espaces vectoriels de E. Alors $F_1 \cup F_2$ est un sous-espace vectoriel de E si et seulement si $F_1 \subset F_2$ ou $F_2 \subset F_1$.*

Proof. Si $F_1 \subset F_2$ alors $F_1 \cup F_2 = F_2$ et la conclusion est évidente. Le même argument fonctionne si $F_2 \subset F_1$. Montrons maintenant la réciproque. Si $F_1 \not\subset F_2$ alors il existe $v \in F_1 \setminus F_2$. Si maintenant $w \in F_2$ alors $v + w \in F_1 \cup F_2$ puisque c'est un sous-espace vectoriel. Si $v + w \in F_2$, alors $v = (v+w)-w \in F_2$ ce qui est impossible. On a donc $v+w \in F_1$ et on en déduit que $w = (v+w)-v \in F_1$, ce qui montre que $F_2 \subset F_1$. Si on suppose que $F_2 \not\subset F_1$, le même argument permet de conclure que $F_1 \subset F_2$. □

Définition 2.2.7. Si F_1, F_2 sont des sous-espaces vectoriels de E, on définit la somme $F_1 + F_2$ par

$$F_1 + F_2 := \{v_1 + v_2 | v_i \in F_i\} \subset E.$$

Proposition 2.2.8. *Le sous-ensemble $F_1 + F_2 \subset E$ est un sous-espace vectoriel.*

Proof. Tout d'abord, on a $0_E \in F_1$ et $0_E \in F_2$, ce qui montre que $0_E = 0_E + 0_E \in F_1 + F_2$. Supposons que $v = v_1 + v_2$ et $w = w_1 + w_2$ soient des éléments de $F_1 + F_2$. Alors,

$$v + w = (v_1 + v_2) + (w_1 + w_2) = (v_1 + w_1) + (v_2 + w_2)$$

en on conclut du fait que $v_1 + w_1 \in F_1$ et $v_2 + w_2 \in F_2$ que $v + w \in F_1 + F_2$. Finalement, supposons que $\lambda \in \mathbb{R}$ et que $v = v_1 + v_2 \in F_1 + F_2$. Alors

$$\lambda v = \lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2 \in F_1 + F_2$$

puisque $\lambda v_1 \in F_1$ et $\lambda v_2 \in F_2$. □

2.3 Combinaisons linéaires, familles libres ou liées

Définition 2.3.1. Soit E un espace vectoriel réel et $v_1, \ldots, v_n \in E$ des vecteurs (où $n \geq 1$). On appelle combinaison linéaire des vecteurs $v_1, \ldots, v_n$ tout vecteur $w \in E$ de la forme

$$w = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = \sum_{i=1}^{n} \lambda_i v_i$$

où $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Remarque 2.3.2. On observe que l'expression $\sum_{i=1}^{n} \lambda_i v_i$ est bien définie grâce à l'associativité de l'addition dans E: les différentes opérations peuvent être effectuées dans un ordre quelconque. Comme l'addition est également commutative, une permutation des vecteurs $v_1, \ldots, v_n$ ne change pas l'ensemble des combinaisons linéaires.

Définition 2.3.3. Si $v_1, \ldots, v_n \in E$, on note $\mathrm{Vect}(v_1, \ldots, v_n)$ l'ensemble des combinaisons linéaires des vecteurs $v_1, \ldots, v_n$.

Proposition 2.3.4. *L'ensemble* $\mathrm{Vect}(v_1, \ldots, v_n)$ *est un sous-espace vectoriel de E, appelé* sous-espace vectoriel engendré par les vecteurs $v_1, \ldots, v_n$. *Il s'agit du plus petit espace vectoriel de* E *contenant les vecteurs* $v_1, \ldots, v_n$.

Proof. Tout d'abord, on sait que $0v_1 = 0v_2 = \ldots = 0v_n = 0_E$. Il s'ensuit que

$$0_E = 0v_1 + 0v_2 + \ldots + 0v_n \in \mathrm{Vect}(v_1, \ldots, v_n)$$

qui est donc non vide. Si maintenant, $v, w \in \mathrm{Vect}(v_1, \ldots, v_n)$, il existe $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n \in \mathbb{R}$ tels que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n, \ w = \mu_1 v_1 + \mu_2 v_2 + \ldots + \mu_n v_n.$$

On en déduit que

$$\begin{aligned} v + w &= (\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) + (\mu_1 v_1 + \mu_2 v_2 + \ldots + \mu_n v_n) \\ &= \lambda_1 v_1 + \mu_1 v_1 + \lambda_2 v_2 + \mu_2 v_2 + \ldots + \lambda_n v_n + \mu_n v_n \\ &= (\lambda_1 + \mu_+)v_1 + (\lambda_2 + \mu_2)v_2 + \ldots + (\lambda_n + \mu_n)v_n, \end{aligned}$$

ce qui montre bien que $v + w \in \mathrm{Vect}(v_1, \ldots, v_n)$. Si maintenant $\mu \in \mathbb{R}$, on a

$$\begin{aligned} \mu v &= \mu(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) \\ &= (\mu\lambda_1)v_1 + (\mu\lambda_2)v_2 + \ldots + (\mu\lambda_n)v_n \end{aligned}$$

qui appartient par conséquent à $\mathrm{Vect}(v_1, \ldots, v_n)$. On en déduit donc que $\mathrm{Vect}(v_1, \ldots, v_n)$ est bien un sous-espace vectoriel de E. On observe maintenant que

$$v_1 = 1v_1 + 0v_2 + \ldots + 0v_n \in \mathrm{Vect}(v_1, \ldots, v_n)$$

et qu'il en est de même pour $v_2, \ldots, v_n$. Supposons finalement que F soit un sous-espace vectoriel de E contenant $v_1, v_2, \ldots, v_n$. Du fait que F est un sous-espace vectoriel, on tire que pour tout $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ les vecteurs $\lambda_1 v_1, \ldots, \lambda_n v_n$ sont dans F. De même, on obtient

$$\lambda_1 v_1 + \ldots + \lambda_n v_n \in F$$

et donc $\mathrm{Vect}(v_1, \ldots, v_n) \subset F$. Ainsi, $\mathrm{Vect}(v_1, \ldots, v_n)$ est bien le plus petit sous-espace vectoriel de E contenant les vecteurs $v_1, v_2, \ldots, v_n$. □

Remarque 2.3.5. Soit $X := \{v_1, \ldots, v_n\}$. Une conséquence de la proposition ci-dessus est que $\mathrm{Vect}(v_1, \ldots, v_n)$ est contenu dans l'intersection de tous les sous-espaces vectoriels F de E contenant X. On démontre facilement qu'en fait c'est une égalité.

On passe maintenant à une définition fondamentale.

Définition 2.3.6. Soit $n \geq 1$ un entier et $v_1, \ldots, v_n$ des vecteurs de E.

a) On dit que la famille $\{v_1, \ldots, v_n\}$ est *libre* si elle vérifie la condition suivante: pour tous $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$,

$$\sum_{i=1}^{n} \lambda_i v_i = 0_E \iff \lambda_1 = \ldots = \lambda_n = 0.$$

b) On dit que la famille $\{v_1, \ldots, v_n\}$ est *liée* si elle n'est pas libre.

Remarque 2.3.7. Si la famille $\{v_1, \ldots, v_n\}$ est libre, on dit aussi que les vecteurs $v_1, \ldots, v_n$ sont *linéairement indépendants*. Dans le cas contraire, on dit que les vecteurs $v_1, \ldots, v_n$ sont *linéairement dépendants*.

Exemple 2.3.8. Supposons que $n = 1$. Dans ce cas, la famille $\{v_1\}$ est libre si et seulement si $v_1 \neq 0$. On a alors $\text{Vect}(v_1) = \{\lambda v_1 | \lambda \in \mathbb{R}\}$, la *droite vectorielle engendrée par* v_1.

Supposons ensuite que $n = 2$. Dans ce cas, la famille $\{v_1, v_2\}$ est liée si et seulement si les deux vecteurs sont *colinéaires*: il existe $\lambda \in \mathbb{R}$ tel que $\lambda v_1 = v_2$ ou $\lambda v_2 = v_1$. Si la famille $\{v_1, v_2\}$ est libre, on dit que $\text{Vect}(v_1, v_2)$ est le plan vectoriel engendré par v_1 et v_2.

Remarque 2.3.9. Si la famille $\{v_1, \ldots, v_n\}$ est libre, alors $v_i \neq 0$ pour tout $i = 1, \ldots, n$. En effet, si par exemple $v_n = 0$, alors

$$0_E = 0v_1 + \ldots + 0v_{n-1} + 1v_n$$

D'un autre côté, la famille $\{v_1, \ldots, v_n\}$ est liée si et seulement si un des vecteurs s'exprime comme une combinaison linéaire des autres. Si par exemple

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

avec $\lambda_1 \neq 0$, on obtient

$$v_1 = -\frac{1}{\lambda_1} \sum_{i=2}^{n} \lambda_i v_i.$$

Réciproquement, si $v_1 = \sum_{i=2}^{n} \lambda_i v_i$, on obtient

$$(-1)v_1 + \sum_{i=2}^{n} \lambda_i v_i = 0_E.$$

Proposition 2.3.10. *Si* $v \in \text{Vect}(v_1, \ldots, v_n)$ *on a* $\text{Vect}(v_1, \ldots, v_n) = \text{Vect}(v, v_1, \ldots, v_n)$.

Proof. De manière générale, on a $\text{Vect}(v_1, \ldots, v_n) \subset \text{Vect}(v, v_1, \ldots, v_n)$ puisque une combinaison linéaire de la forme $\sum_{i=1}^{n} \lambda_i v_i$ satisfait

$$\sum_{i=1}^{n} \lambda_i v_i = 0v + \sum_{i=1}^{n} \lambda_i v_i.$$

Montrons l'inclusion réciproque. Supposons que $v = \sum_{i=1}^{n} \mu_i v_i$. Dans ce cas, une combinaison de la forme $\lambda v + \sum_{i=1}^{n} \lambda_i v_i$ s'écrit alors

$$\lambda v + \sum_{i=1}^{n} \lambda_i v_i = \lambda(\sum_{i=1}^{n} \mu_i v_i) + \sum_{i=1}^{n} \lambda_i v_i = \sum_{i=1}^{n} (\lambda\mu_i + \lambda_i) v_i$$

ce qui montre bien que $\text{Vect}(v, v_1, \ldots, v_n) \subset \text{Vect}(v_1, \ldots, v_n)$. $\square$

Proposition 2.3.11. *Supposons que la famille $\{v_1, \ldots, v_n\}$ soit libre. Si $v \in E$ est tel que $v \notin \mathrm{Vect}(v_1, \ldots, v_n)$ alors la famille $\{v, v_1, \ldots, v_n\}$ est libre.*

Proof. Supposons que

$$\mu v + \sum_{i=1}^{n} v_i = 0_E.$$

Si $\mu \neq 0$, alors on obtient que

$$v = \frac{1}{\mu}\mu v = -\frac{1}{\mu}\sum_{i=1}^{n} v_i$$

ce qui en fait une combinaison linéaire de $v_1, \ldots, v_n$. C'est exclus par hypothèse, et on en déduit donc que $\mu = 0$. L'égalité ci-dessus devient alors $\sum_{i=1}^{n} v_i = 0_E$ et on en tire que $\lambda_1 = \ldots = \lambda_n = 0$. La famille $\{v, v_1, \ldots, v_n\}$ est donc libre. □

Exemple 2.3.12. Soit $E = \mathbb{R}^3$ et $v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix}$. Supposons que $\lambda v_1 + \mu v_2 = 0_E$ pour certains $\lambda, \mu \in \mathbb{R}$. Ceci donne

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} \lambda \\ 2\mu \\ -2\mu \end{pmatrix}.$$

On obtient donc $\lambda = \mu = 0$ et on en déduit que la famille $\{v_1, v_2\}$ de $\mathbb{R}^3$ est libre. Par ailleurs,

$$\mathrm{Vect}(v_1, v_2) = \left\{ \lambda \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} | \lambda, \mu \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} \lambda \\ 2\mu \\ -2\mu \end{pmatrix} | \lambda, \mu \in \mathbb{R} \right\}.$$

Posant $x_1 = \lambda$ et $x_2 = 2\mu$, on a finalement

$$\mathrm{Vect}(v_1, v_2) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ -x_2 \end{pmatrix} | x_1, x_2 \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 | x_2 + x_3 = 0 \in \mathbb{R} \right\}.$$

2.4 Familles génératrices, bases

Définition 2.4.1. Soit $n \geq 1$ un entier, et soient $v_1, \ldots, v_n$ des vecteurs de E. On dit que la famille $\{v_1, \ldots, v_n\}$ est:

- *génératrice* si $\mathrm{Vect}(v_1, \ldots, v_n) = E$,
- *une base de E* si elle est libre et génératrice.

La caractérisation suivante est fondamentale.

Proposition 2.4.2. *Une famille $\{v_1, \ldots, v_n\}$ est une base de l'espace vectoriel E si et seulement si tout vecteur $v \in E$ admet une écriture unique de la forme*

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n = \sum_{i=1}^{n} \lambda_i v_i.$$

Les scalaires $\lambda_1, \ldots, \lambda_n$ sont appelés les coefficients de v dans la base $v_1, \ldots, v_n$.

Proof. On suppose tout d'abord que $\{v_1, \ldots, v_n\}$ est une base de E. Si $v \in E$, on utilise le fait que la famille est génératrice pour obtenir une expression de la forme

$$v = \sum_{i=1}^{n} \lambda_i v_i \tag{2.4.1}$$

pour certains $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Pour montrer que ces scalaires sont uniquement déterminés, on suppose qu'il existe $\mu_1, \ldots, \mu_n \in \mathbb{R}$ tels que $v = \sum_{i=1}^{n} \mu_i v_i$. On obtient alors

$$0_E = v - v = \sum_{i=1}^{n} \lambda_i v_i - \sum_{i=1}^{n} \mu_i v_i = \sum_{i=1}^{n} (\lambda_i - \mu_i) v_i.$$

Comme la famille est libre, on en déduit que $\lambda_1 - \mu_1 = \ldots = \lambda_n - \mu_n = 0$, ou encore que $\lambda_1 = \mu_1, \ldots, \lambda_n = \mu_n$. Il s'ensuit que l'écriture (2.4.1) est bien unique.

Réciproquement, supposons que tout $v \in E$ s'écrit de manière unique comme combinaison linéaire des vecteurs $v_1, \ldots, v_n$. En particulier, cela entraîne que $E = \mathrm{Vect}(v_1, \ldots, v_n)$ et il ne reste qu'à prouver que la famille $\{v_1, \ldots, v_n\}$ est libre pour conclure. Supposons donc que $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ sont tels que

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

On sait que $0_E = 0v_1 + \ldots + 0v_n = \sum_{i=1}^{n} 0v_i$. Par unicité de l'écriture, on obtient que $\lambda_1 = 0, \ldots, \lambda_n = 0$, et ainsi la famille est bien libre. □

Définition 2.4.3. Soit $\{v_1, \ldots, v_n\} \subset E$ une famille de vecteurs. On dit qu'une famille $\{w_1, \ldots, w_m\} \subset F$ est une *extraction* de $\{v_1, \ldots, v_n\}$ s'il existe une application *strictement croissante*

$$\varphi : \{1, \ldots, m\} \to \{1, \ldots, n\}$$

telle que $w_i = v_{\varphi(i)}$ pour tout $i = 1, \ldots, m$.

Remarque 2.4.4. Du fait que φ est strictement croissante, on a forcément $m \leq n$. En termes plus concrets, une extraction de la famille $\{v_1, \ldots, v_n\}$ est une sous-famille obtenue en retirant certains éléments, l'application φ servant à renuméroter ceux qui restent.

Exemple 2.4.5. La famille $\{v_1, v_3, v_4\}$ est une extraction de la famille $\{v_1, \ldots, v_5\} \subset E$. Ici la fonction φ est déterminée par $\varphi(1) = 1, \varphi(2) = 3$ et $\varphi(3) = 4$.

Dans la suite, on suppose toujours que $E \neq \{0\}$, ceci pour éviter de parler de familles à zéro éléments.

Proposition 2.4.6. *Si E admet une famille génératrice $\{v_1, \ldots, v_n\} \subset E$, alors il existe une extraction $\{w_1, \ldots, w_m\}$ de $\{v_1, \ldots, v_n\}$ qui est une base de E.*

Proof. On construit la famille $\{w_1, \ldots, w_m\}$ inductivement de la manière suivante:

- (Initialisation): Soit $i_1 = \min\{i \in \{1, \ldots, n\} | v_i \neq 0\}$. L'élément $i_1 \in \mathbb{N}$ existe bien car $E \neq \{0\}$. On pose donc $w_1 = v_{i_1}$, et on note que $\mathrm{Vect}(w_1) = \mathrm{Vect}(v_1, \ldots, v_{i_1})$.

- (Étape de récurrence): On suppose qu'on a construit une extraction $\{w_1, \ldots, w_k\} \subset E$ et que la famille $\{w_1, \ldots, w_k\}$ est libre avec $\mathrm{Vect}(w_1, \ldots, w_k) = \mathrm{Vect}(v_1, \ldots, v_{i_k})$. Si $\mathrm{Vect}(w_1, \ldots, w_k) = E$, la famille est bien une base et on a terminé. Sinon, on pose

$$\begin{aligned} i_{k+1} &= \min\{i_k < i \leq n | v_i \notin \mathrm{Vect}(w_1, \ldots, w_k) \\ w_{k+1} &= v_{i_k+1}. \end{aligned}$$

Avec ceci, $\{w_1, \ldots, w_{k+1}\}$ est une famille libre et $\mathrm{Vect}(w_1, \ldots, w_{k+1}) = \mathrm{Vect}(v_1, \ldots, v_{i_{k+1}})$. Cette construction se termine en un nombre fini d'étapes, et fournit une extraction $\{v_1, \ldots, v_n\}$ qui est une base de E. $\square$

Définition 2.4.7. On dit qu'un espace vectoriel $E \neq \{0\}$ est de *dimension finie* s'il existe dans E une famille génératrice finie.

Corollaire 2.4.8. *Tout espace vectoriel $E \neq \{0\}$ de dimension finie admet une base.*

Proposition 2.4.9 (théorème de la base incomplète)**.** *Dans un espace vectoriel E de dimension finie, toute famille libre de E peut être complétée en une base de E.*

Proof. Par hypothèse, E admet une famille génératrice $\{v_1, \ldots, v_n\}$. Si maintenant $\{w_1, \ldots, w_k\}$ est une famille libre, on considère la famille $\{w_1, \ldots, w_k, v_1, \ldots, v_n\} \subset E$, qui est bien génératrice. On peut donc en extraire une base $\{z_1, \ldots, z_m\} \subset E$ en appliquant le procédé décrit dans la proposition 2.4.6. La construction garantit que $m \geq k$ et que $\{z_1, \ldots, z_k\} = \{w_1, \ldots, w_k\}$. $\square$

2.5 Théorie de la dimension

Lemme 2.5.1. *Soit E un espace vectoriel admettant une famille génératrice $\{v_1, \ldots, v_n\}$. Si $\{w_1, \ldots, w_m\}$ est une famille libre, alors $m \leq n$.*

Proof. Supposons au contraire que $m > n$. On va construire de manière inductive une nouvelle famille génératrice de la manière suivante:

- (Initialisation): Comme $\{v_1, \ldots, v_n\}$ est génératrice, il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que $w_1 = \sum_{i=1}^n \lambda_i v_i$. Comme $w_1 \neq 0$, les scalaires λ_i ne sont pas tous nuls, et on peut supposer quitte à permuter les v_i que $\lambda_1 \neq 0$. On trouve ainsi

$$v_1 = \frac{1}{\lambda_1}(w_1 - \sum_{i=2}^n \lambda_i v_i).$$

Ainsi, $E = \mathrm{Vect}(v_1, \ldots, v_n) \subset \mathrm{Vect}(w_1, v_2, \ldots, v_n) \subset E$, et on obtient $\mathrm{Vect}(w_1, v_2, \ldots, v_n) = E$.

- (Récurrence): Supposons que pour un certain $k \in \{1, \ldots, n-1\}$ on ait montré que la famille $\{w_1, \ldots, w_k, v_{k+1}, \ldots, v_n\}$ est génératrice. On a donc

$$w_{k+1} = \sum_{i=1}^{k} \lambda_i w_i + \sum_{i=k+1}^{n} \lambda_i v_i$$

pour certains $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Du fait que la famille $\{w_1, \ldots, w_{k+1}\}$ est libre, on voit que $\lambda_{k+1}, \ldots, \lambda_n$ ne sont pas tous nuls. Quitte à permuter, on peut supposer que $\lambda_{k+1} \neq 0$, et on trouve

$$v_{k+1} = \frac{1}{\lambda_{k+1}} \left(w_{k+1} - \sum_{i=1}^{k} \lambda_i w_i - \sum_{i=k+2}^{n} \lambda_i v_i \right).$$

Il s'ensuit comme ci-dessus que

$$\mathrm{Vect}(w_1, \ldots, w_k, v_{k+1}, \ldots, v_n) = \mathrm{Vect}(w_1, \ldots, w_k, w_{k+1}, v_{k+2}, \ldots, v_n) = E.$$

À l'issue de la récurrence, on obtient que la famille $\{w_1, \ldots, w_n\}$ est génératrice. Comme $m > n$, on a $w_m \notin \{w_1, \ldots, w_n\}$, et il existe donc des scalaires $\lambda_1, \ldots, \lambda_n$ tels que $w_m = \sum_{i=1}^{n} \lambda_i w_i$, contredisant l'hypothèse que la famille $\{w_1, \ldots, w_m\}$ est libre. On a donc finalement $m \leq n$. $\square$

Corollaire 2.5.2. *Si $\{v_1, \ldots, v_n\}$ et $\{w_1, \ldots, w_m\}$ sont deux bases d'un espace vectoriel E, alors $n = m$.*

Définition 2.5.3. Soit E un espace vectoriel de dimension finie. On appelle *dimension de E* et on note $\dim(E)$ l'entier $n \in \mathbb{N}$ défini par:

- Si $E = \{0_E\}$, alors $n = 0$.
- Si $E \neq \{0_E\}$, alors n est le nombre d'éléments d'une base de E.

Exemple 2.5.4. L'espace vectoriel $\mathbb{R}^n$, muni des lois habituelles, est de dimension n. La *base canonique* est la famille $\{e_1, \ldots, e_n\}$ définie par:

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

Exemple 2.5.5. Soit $E = \mathbb{R}[X]$, muni des lois de l'exemple 2.1.9. Pour tout $n \in \mathbb{N}$, on considère la famille $\mathcal{F}_n := \{1, X, \ldots, X^n\}$. On vérifie que cette famille est libre. Soient $\lambda_0, \lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que

$$0 = \lambda_0 1 + \lambda_1 X + \ldots + \lambda_n X^n = \sum_{i=0}^{n} \lambda_i X^i.$$

Supposons que les scalaires ne sont pas tous nuls. Alors, on pose $d := \max\{i \in \mathbb{N} | \lambda_i \neq 0\}$, et on note qu'alors $\lambda_0 1 + \lambda_1 X + \ldots + \lambda_n X^n = \lambda_0 1 + \lambda_1 X + \ldots + \lambda_d X^d$ est un polynôme de degré d. Par ailleurs, ce polynôme est nul par hypothèse, ce qui est une contradiction. On en déduit bien que $\mathcal{F}_n$ est libre. Ceci montre que $E = \mathbb{R}[X]$ n'est pas de dimension finie. Dans le cas contraire, on pourrait trouver une base $\{v_1, \ldots, v_n\}$ de E. Cette famille est en particulier génératrice de cardinal n, alors que la famille $\mathcal{F}_n$ est libre de cardinal $n+1$, contredisant le lemme 2.5.1.

Cet exemple est l'occasion de poser la définition suivante.

Définition 2.5.6. On dit qu'un espace vectoriel E est de *dimension infinie* s'il n'est pas de dimension finie.

Remarque 2.5.7. Dans la veine de l'exemple ci-dessus, on peut démontrer que E est de dimension infinie si et seulement s'il existe pour tout $n \in \mathbb{N}$ une famille libre de n éléments de E. Démontrons le sens direct, la réciproque étant prouvée comme dans l'exemple ci-dessus. Supposons donc que E soit de dimension infinie. On va construire par récurrence pour tout $n \in \mathbb{N}$ une famille libre de n éléments de E notée $\mathcal{F}_n$, procédant de la manière suivante:

- (Initialisation): Comme E est de dimension infinie, il est en particulier non nul et il existe $v_1 \in E$ non nul. On pose alors $\mathcal{F}_1 = \{v_1\}$, qui est bien libre.
- (Récurrence): Supposons avoir construit $\mathcal{F}_n = \{v_1, \ldots, v_n\}$ pour un certain $n \geq 1$. Comme E est de dimension infinie, la famille $\mathcal{F}_n$ n'est pas génératrice. Il existe donc $v_{n+1} \in E$ tel que $v_{n+1} \notin \mathrm{Vect}(v_1, \ldots, v_n)$. On en déduit que la famille $\mathcal{F}_{n+1} := \{v_1, \ldots, v_n, v_{n+1}\} \subset E$ est bien libre.

Revenons-en aux espaces vectoriels de dimension finie.

Proposition 2.5.8. *Soit $n \geq 1$ un entier et E un espace vectoriel de dimension n. Une famille $\mathcal{F} := \{v_1, \ldots, v_n\}$ est libre si et seulement si elle est génératrice. Dans ce cas, elle forme une base.*

Proof. Supposons que $\mathcal{F}$ est libre. Si elle n'était pas génératrice, il existerait $v \in E$ tel que $v \notin \mathrm{Vect}(v_1, \ldots, v_n)$, et ainsi la famille $\{v, v_1, \ldots, v_n\}$ serait libre. On peut la compléter en une base de E, qui aurait au moins $n+1$ éléments. Ceci contredit le corollaire 2.5.2, donc $\mathcal{F}$ est bien génératrice. Supposons maintenant que $\mathcal{F}$ est une famille génératrice. Si elle était liée, on aurait une expression du type

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

avec certains scalaires non nuls. Quitte à permuter les éléments de $\mathcal{F}$, on peut supposer que $\lambda_1 \neq 0$ obtenant ainsi

$$v_1 = -\frac{1}{\lambda_1}(\sum_{i=2}^{n} v_i)$$

Donc $E = \mathrm{Vect}(v_1, \ldots, v_n) = \mathrm{Vect}(v_2, \ldots, v_n)$. On peut extraire de $\{v_2, \ldots, v_n\}$ une base, qui a au plus $n-1$ éléments, contredisant à nouveau le corollaire 2.5.2. On en conclut que la famille $\mathcal{F}$ est libre. □

Définition 2.5.9. Soient F_1, F_2 deux sous-espaces vectoriels de E. On dit que E est la somme directe de F_1 et F_2, et on note $E = F_1 \oplus F_2$, si:

- $E = F_1 + F_2$.
- $F_1 \cap F_2 = \{0_E\}$.

Exemple 2.5.10. Soit $E = \mathbb{R}^2$, $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. On pose $F_1 = \mathrm{Vect}(v_1)$, $F_2 = \mathrm{Vect}(v_2)$ et on vérifie qu'on a bien $E = F_1 \oplus F_2$. Si $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in E$, on vérifie que

$$v = (x_1 - x_2)v_1 + x_2 v_2$$

et ainsi que $E = F_1 + F_2$ puisque $(x_1 - x_2)v_1 \in \mathrm{Vect}(v_1) = F_1$ et $x_2 v_2 \in \mathrm{Vect}(v_2) = F_2$. Supposons maintenant que $v \in F_1 \cap F_2$. Il existe alors $\lambda_1 \in \mathbb{R}$ tel que $v = \lambda_1 v_1$ et $v = \lambda_2 v_2$. On obtient

$$0_E = v - v = \lambda_1 v_1 - \lambda_2 v_2.$$

Or, la famille $\{v_1, v_2\} \subset \mathbb{R}^2$ est génératrice comme on l'a montré ci-dessus. Puisque $\mathbb{R}^2$ est de dimension 2, il s'ensuit que $\{v_1, v_2\}$ est libre, et on tire de l'égalité ci-dessus que $\lambda_1 = \lambda_2 = 0$. Donc $v = 0$ et on a bien $F_1 \cap F_2 = \{0_E\}$.

Proposition 2.5.11. *Soit E un espace vectoriel de dimension finie, et $F \subset E$ un sous-espace vectoriel. Alors:*

1. *F est de dimension finie et $\dim(F) \leq \dim(E)$.*
2. *Il existe un sous-espace vectoriel $G \subset E$ tel que $E = F \oplus G$.*

Proof. On suppose $E \neq \{0_E\}$, auquel cas il n'y a rien à faire, et on pose $n = \dim E$. On remarque tout d'abord que toute famille libre $\{v_1, \ldots, v_m\} \subset F$ est aussi une famille libre de E, et on en déduit que la taille maximale d'une famille libre dans F est n. La remarque 2.5.7 montre alors que F est dimension finie, alors que le lemme 2.5.1 montre que $\dim(F) \leq n$. Pour la seconde assertion, on suppose que $F \neq \{0\}$ (sinon on pose $G = E$) et que $F \neq E$ (sinon on prend $G = \{0_E\}$). On choisit une base $\{v_1, \ldots, v_k\}$ de F (avec $1 \leq k \leq n-1$). Comme cette famille est libre dans E, on peut la compléter en une base $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ de E, et on pose $G = \mathrm{Vect}(v_{k+1}, \ldots, v_n)$. Si $v \in E$, on exploite le fait que $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ est génératrice pour avoir une expression de la forme

$$v = \sum_{i=1}^{k} \lambda_i v_i + \sum_{i=k+1}^{n} \lambda_i v_i$$

On observe que $w_1 := \sum_{i=1}^{k} \lambda_i v_i$ est un élément de $\mathrm{Vect}(v_1, \ldots, v_k) = F$ et que $w_2 := \sum_{i=k+1}^{n} \lambda_i v_i$ est un élément de $G = \mathrm{Vect}(v_{k+1}, \ldots, v_n)$. Ainsi $v = w_1 + w_2 \in F + G$ et on a bien démontré que $E = F + G$. Si $v \in F \cap G$, il existe $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ et $\lambda_{k+1}, \ldots, \lambda_n \in \mathbb{R}$ tels que $v = \sum_{i=1}^{k} \lambda_i v_i$ et $v = \sum_{i=k+1}^{n} \lambda_i v_i$. On obtient alors

$$0_E = v - v = \sum_{i=1}^{k} \lambda_i v_i - \sum_{i=k+1}^{n} \lambda_i v_i$$

La famille $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\} \subset E$ étant libre, on obtient que $\lambda_1 = \ldots = \lambda_n = 0$, prouvant que $v = 0$ et donc que $F \cap G = \{0_E\}$. $\square$

Définition 2.5.12. Si $F \subset E$ est un sous-espace vectoriel, un sous-espace vectoriel $G \subset E$ tel que $F \oplus G = E$ est appelé *un sous-espace supplémentaire de F dans E.*

Proposition 2.5.13. *Soit E un espace vectoriel et $F_1, F_2 \subset E$ des sous-espaces vectoriels tels que $E = F_1 \oplus F_2$. Alors* $\dim(F_1) + \dim(F_2) = \dim(E)$.

Proof. On choisit une base $\{v_1, \ldots, v_n\}$ de F_1 et une base $\{w_1, \ldots, w_m\}$ de F_2, puis on démontre que $\{v_1, \ldots, v_n, w_1, \ldots, w_m\}$ est une base de E. Si $v \in E$ alors il existe $w_1 \in F_1$ et $w_2 \in F_2$ tels que $v = w_1 + w_2$. Par ailleurs, on peut écrire $w_1 = \sum_{i=1}^n \lambda_i v_i$ et $w_2 = \sum_{i=1}^m \mu_i w_i$ pour obtenir

$$v = w_1 + w_2 = \sum_{i=1}^n \lambda_i v_i + \sum_{i=1}^m \mu_i w_i,$$

ce qui démontre que la famille $\{v_1, \ldots, v_n, w_1, \ldots, w_m\}$ est génératrice. Supposons maintenant que $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_m \in \mathbb{R}$ sont tels que

$$0_E = \sum_{i=1}^n \lambda_i v_i + \sum_{i=1}^m \mu_i w_i$$

Posant $w_1 = \sum_{i=1}^n \lambda_i v_i \in \mathrm{Vect}(v_1, \ldots, v_n) = F_1$ et $w_2 = \sum_{i=1}^m \mu_i w_i \in \mathrm{Vect}(w_1, \ldots, w_m) = F_2$, on obtient $0_E = w_1 + w_2$, ou encore $w_1 = -w_2$. Donc $w_1 \in F_1 \cap F_2 = \{0_E\}$, et de même $w_2 \in F_1 \cap F_2 = \{0_E\}$. On a donc bien que $w_1 = \sum_{i=1}^n \lambda_i v_i = 0_E$ et que $w_2 = \sum_{i=1}^m \mu_i w_i = 0_E$. Du fait que $\{v_1, \ldots, v_n\}$ et $\{w_1, \ldots, w_m\}$ sont respectivement des bases de F_1 et F_2, on en déduit finalement que $\lambda_1 = \ldots = \lambda_n = \mu_1 = \ldots = \mu_m = 0$. Donc la famille $\{v_1, \ldots, v_n, w_1, \ldots, w_m\}$ est bien une base de E, d'où on tire que $\dim(F_1) + \dim(F_2) = \dim(E)$. □

Remarque 2.5.14. Si E est de dimension finie et F est un sous-espace de E, on en déduit que $F = E$ si et seulement si $\dim(E) = \dim(F)$.

3 Applications linéaires

Comme dans le chapitre précédent, on suppose que les espaces vectoriels sont sur $\mathbb{R}$.

Définition 3.0.1. Soient E et F deux espaces vectoriels réels. On dit qu'une application $\varphi \colon E \to F$ est *linéaire* si elle possède les propriétés suivantes:

i) $\varphi(u + v) = \varphi(u) + \varphi(v)$ pour tous $u, v \in E$.

ii) $\varphi(\lambda u) = \lambda \varphi(u)$ pour tout $u \in E$ et tout $\lambda \in \mathbb{R}$.

Remarque 3.0.2. On peut rassembler les conditions ci-dessus en une unique condition: une application $\varphi \colon E \to F$ est linéaire si et seulement si

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

pour tous $u, v \in E$ et tout $\lambda \in \mathbb{R}$.

Remarque 3.0.3. En particulier, les espaces vectoriels E et F sont des groupes abéliens pour les additions respectives. La première condition de la définition ci-dessus dit que φ est un morphisme de groupe, et la seconde condition donne une contrainte supplémentaire. Comme φ est un morphisme de groupe, on a en particulier:

- $\varphi(0_E) = 0_F$.
- $\varphi(-u) = -\varphi(u)$ pour tout $u \in E$.

Par ailleurs, on rappelle que $\ker(\varphi) = \{u \in E | \varphi(u) = 0_F\}$ et $\mathrm{Im}(\varphi) = \{\varphi(u) | u \in E\}$ sont respectivement des sous-groupes de E et F. On verra plus tard que ce sont en fait des sous-espaces vectoriels. En attendant, on peut appliquer la remarque 1.3.14 pour obtenir que:

- φ est injective si et seulement si $\ker(\varphi) = \{0_E\}$.
- φ est surjective si et seulement si $\mathrm{Im}(\varphi) = F$.

Remarque 3.0.4. Si $\varphi\colon E \to F$ est une application linéaire, alors pour tout entier $n \geq 1$, tous $v_1, \ldots, v_n \in E$ et tous $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ on a

$$\varphi(\sum_{i=1}^{n} \lambda_i v_i) = \sum_{i=1}^{n} \lambda_i \varphi(v_i).$$

On peut le démontrer par induction. Ainsi, l'image d'une combinaison linéaire des vecteurs $v_1, \ldots, v_n$ est une combinaison linéaire des vecteurs $\varphi(v_1), \ldots, \varphi(v_n)$.

Notation 3.0.5. Une application linéaire $\varphi\colon E \to F$ est aussi appelée un *morphisme d'espaces vectoriels*. Comme dans le cas des groupes, on dit que:

- φ est un endomorphisme si $E = F$.
- φ est un isomorphisme si elle est bijective.
- φ est un automorphisme si $E = F$ et elle est bijective.

On dit également que φ est une *forme linéaire* si $F = \mathbb{R}$, c'est à dire si $\varphi\colon E \to \mathbb{R}$.

Exemple 3.0.6. Une application $\varphi\colon \mathbb{R} \to \mathbb{R}$ est linéaire s'il existe $a \in \mathbb{R}$ tel que $\varphi(x) = ax$ pour tout $x \in \mathbb{R}$. En effet, on vérifie facilement qu'une telle application est linéaire. Réciproquement, supposons que φ est linéaire, et posons $a = \varphi(1)$. On a alors pour tout $x \in \mathbb{R}$:

$$\varphi(x) = \varphi(x1) = x\varphi(1) = xa = ax.$$

Exemple 3.0.7. Soit X un ensemble non vide, et $E = \mathbb{R}^X$, l'espace vectoriel des fonctions de X dans $\mathbb{R}$. Si $x_0 \in X$, on définit une application $\varphi : E \to \mathbb{R}$ par $\varphi(f) = f(x_0)$ pour tout $f \in E$, et on vérifie que c'est une forme linéaire.

3.1 Opérations sur les applications linéaires

Proposition 3.1.1. *Soient E, F, G des espaces vectoriels et $\varphi\colon E \to F$, $\psi\colon F \to G$ des applications linéaires. Alors, l'application*

$$\psi \circ \varphi\colon E \to G$$

est linéaire.

Proof. Si $u, v \in E$ et $\lambda \in \mathbb{R}$, on obtient (utilisant respectivement que φ puis que ψ sont linéaires)

$$\begin{aligned}(\psi \circ \varphi)(u + \lambda v) &= \psi(\varphi(u + \lambda v)) \\ &= \psi(\varphi(u) + \lambda\varphi(v)) \\ &= \psi(\varphi(u)) + \lambda\psi(\varphi(v)) \\ &= (\psi \circ \varphi)(u) + \lambda(\psi \circ \varphi)(v).\end{aligned}$$

□

Proposition 3.1.2. *Soit $\varphi\colon E \to F$ une application linéaire bijective. Alors, l'application réciproque*

$$\varphi^{-1}\colon F \to E$$

est également linéaire.

Proof. Si $u, v \in F$ et $\lambda \in \mathbb{R}$, il suffit de vérifier l'égalité $\varphi^{-1}(u + \lambda v) = \varphi^{-1}(u) + \lambda\varphi^{-1}(v)$ dans E. Comme φ est bijective, cette égalité est vérifiée si et seulement si

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

Or, $\varphi(\varphi^{-1}(u) + \lambda\varphi^{-1}(v)) = \varphi(\varphi^{-1}(u)) + \lambda\varphi(\varphi^{-1}(v)) = u + \lambda v$ puisque φ est linéaire. □

Proposition 3.1.3. *Soient E et F des espaces vectoriels réels, $\varphi\colon E \to F$ une application linéaire et $\{v_1, \ldots, v_n\} \subset E$ une famille des vecteurs. Alors:*

1. *Si φ est injective et la famille $\{v_1, \ldots, v_n\} \subset E$ est libre, alors la famille $\{\varphi(v_1), \ldots, \varphi(v_n)\} \subset F$ est libre.*

2. *Si la famille $\{v_1, \ldots, v_n\} \subset E$ est génératrice, alors φ est surjective si et seulement si la famille $\{\varphi(v_1), \ldots, \varphi(v_n)\} \subset F$ est génératrice.*

3. *Si la famille $\{v_1, \ldots, v_n\} \subset E$ est une base, alors φ est bijective si et seulement si la famille $\{\varphi(v_1), \ldots, \varphi(v_n)\} \subset F$ est une base.*

Proof. Pour le premier point, supposons que $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ soient tels que $\sum_{i=1}^n \lambda_i\varphi(v_i) = 0_F$. On obtient alors

$$0_F = \sum_{i=1}^n \lambda_i\varphi(v_i) = \varphi(\sum_{i=1}^n \lambda_i v_i).$$

ce qui montre que $\sum_{i=1}^n \lambda_i v_i \in \ker(\varphi)$. Comme l'application est injective, on a $\ker(\varphi) = \{0_E\}$ et ainsi $\sum_{i=1}^n \lambda_i v_i = 0_E$. La famille étant libre, on en déduit que $\lambda_1 = \ldots = \lambda_n = 0$, et donc que la famille $\{\varphi(v_1), \ldots, \varphi(v_n)\} \subset F$ est libre.

Pour le second point, supposons tout d'abord que φ est surjective. Si $w \in F$, il existe alors $v \in E$ tel que $\varphi(v) = w$. La famille $\{v_1, \ldots, v_n\} \subset E$ étant génératrice, il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que $\sum_{i=1}^n \lambda_i v_i = v$. On en tire que

$$w = \varphi(v) = \varphi(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i \varphi(v_i)$$

ce qui montre bien que la famille $\{\varphi(v_1), \ldots, \varphi(v_n)\} \subset F$ est génératrice. Réciproquement, supposons que cette famille soit génératrice. Si $w \in F$, il existe alors $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que $w = \sum_{i=1}^n \lambda_i \varphi(v_i)$. On obtient alors

$$w = \sum_{i=1}^n \lambda_i \varphi(v_i) = \varphi(\sum_{i=1}^n \lambda_i v_i)$$

et donc $w \in \mathrm{Im}(\varphi)$.

On passe finalement à la dernière assertion. Si $\{v_1, \ldots, v_n\} \subset E$ est une base, elle est en particulier libre et génératrice. On démontre dans ce cas que φ est injective si la famille $\{\varphi(v_1), \ldots, \varphi(v_n)\} \subset F$ est libre. Supposons que $v \in E$ soit tel que $\varphi(v) = 0_F$. Il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que $v = \sum_{i=1}^n \lambda_i v_i$, et on obtient

$$0_F = \varphi(v) = \varphi(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i \varphi(v_i).$$

Comme la famille $\{\varphi(v_1), \ldots, \varphi(v_n)\} \subset F$ est libre, on en tire que $\lambda_1 = \ldots = \lambda_n = 0$ et ainsi que $v = 0_E$. On a bien trouvé que $\ker(\varphi) = \{0_E\}$, et donc que φ est injective. Mettant finalement les première et seconde assertion ensemble, on obtient bien que φ est bijective si et seulement si $\{\varphi(v_1), \ldots, \varphi(v_n)\} \subset F$ est une base. $\square$

Corollaire 3.1.4. *Tout espace vectoriel réel de dimension $n \in \mathbb{N} \setminus \{0\}$ est isomorphe à $\mathbb{R}^n$.*

Proof. Choisissons une base $\{v_1, \ldots, v_n\}$ de E. Si $\{e_1, \ldots, e_n\}$ est la base canonique de $\mathbb{R}^n$, on vérifie que l'application $\varphi \colon \mathbb{R}^n \to E$ donnée par

$$\varphi(\sum_{i=1}^n \lambda_i e_i) = \sum_{i=1}^n \lambda_i v_i$$

est bien définie et linéaire. Pour tout $i = 1, \ldots, n$, on a $\varphi(e_i) = v_i$ et on en déduit de la proposition ci-dessus que φ est bijective. $\square$

Remarque 3.1.5. Plus concrètement, on a

$$\varphi(\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}) = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Ceci donne une formule pour φ^{-1}: si $v = \sum_{i=1}^n \lambda_i v_i$, on a

$$\varphi^{-1}(v) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Autrement dit, φ^{-1} associe à un vecteur v ses coordonnées dans la base $\{v_1, \ldots, v_n\}$ de E.

Remarque 3.1.6. L'application linéaire φ ci-dessus dépend du choix de la base $\{v_1, \ldots, v_n\}$ de E. Toute base donne un isomorphisme, et réciproquement tout isomorphisme $\varphi\colon \mathbb{R}^n \to E$ donne une base de E. Ainsi, se donner un isomorphisme entre $\mathbb{R}^n$ et E est exactement la même chose que choisir une base de E.

Corollaire 3.1.7. *Deux espaces vectoriels E et F de dimension finie sont isomorphes si et seulement si* $\dim(E) = \dim(F)$.

Proof. Si $\varphi\colon E \to F$ est un isomorphisme, alors l'image d'une base de E par φ est une base de F au vu de la proposition 3.1.3. On en déduit que $\dim(E) = \dim(F)$. Réciproquement, supposons $\dim(E) = \dim(F) = n$. Si $n = 0$, alors $E = \{0_E\}$ et $F = \{0_F\}$, auquel cas il n'y a rien à faire. Si $n \geq 1$, on trouve deux isomorphismes $\psi_1\colon \mathbb{R}^n \to E$ et $\psi_2\colon \mathbb{R}^n \to F$. On en tire que

$$\psi_2 \circ \psi_1^{-1}\colon E \to F$$

est un isomorphisme. □

Supposons que X soit un ensemble non vide et que F soit un espace vectoriel. Alors, on peut définir sur l'ensemble $F^X = \{f\colon X \to F\}$ les opérations suivantes:

- (addition) Pour tout $f, g \in F^X$, on définit l'application $f + g$ par $(f + g)(x) = f(x) + g(x)$ pour tout $x \in X$.
- (loi externe) Pour tout $f \in F^X$ et tout $\lambda \in \mathbb{R}$ on définit l'application λf par $(\lambda f)(x) = \lambda f(x)$ pour tout $x \in X$.

On vérifie comme dans l'exemple 2.1.8 que F^X est bien un espace vectoriel.

Proposition 3.1.8. *Si E et F sont deux espaces vectoriels, alors l'ensemble $\mathcal{L}(E, F)$ des applications linéaires de E dans F est un sous-espace vectoriel de F^E.*

Proof. Il suffit de vérifier que pour $\varphi, \psi \in \mathcal{L}(E, F)$ et $\lambda \in \mathbb{R}$, les applications $\varphi + \psi$ et $\lambda\varphi$ sont encore linéaires. Il s'agit d'un calcul direct. □

On tire de cette proposition que $\mathcal{L}(E, F)$ (muni des lois ci-dessus) est un espace vectoriel réel.

3.2 Sous-espaces, noyau et image

Proposition 3.2.1. *Soient E, F deux espaces vectoriels réels et $\varphi\colon E \to F$ une application linéaire. Alors:*

a) Si $E' \subset E$ est un sous-espace vectoriel, alors

$$\varphi(E') = \{\varphi(v) | v \in E'\} \subset F$$

est un sous-espace vectoriel.

b) Si $F' \subset F$ est un sous-espace vectoriel, alors

$$\varphi^{-1}(F') = \{v \in E | \varphi(v) \in F'\} \subset E$$

est un sous-espace vectoriel.

Proof. Montrons tout d'abord a). Si $\varphi(v), \varphi(w) \in \varphi(E')$ et $\lambda \in \mathbb{R}$, on a

$$\varphi(v) + \lambda\varphi(w) = \varphi(v + \lambda w) \in \varphi(E')$$

car $v+\lambda w \in E'$ puisque c'est un sous-espace vectoriel de E. Pour b), soient $v, w \in \varphi^{-1}(F')$ et $\lambda \in \mathbb{R}$. Alors on a

$$\varphi(v + \lambda w) = \varphi(v) + \lambda\varphi(w) \in F'$$

puisque $\varphi(v)$ et $\varphi(w)$ sont dans F'. On en déduit que $v + \lambda w \in \varphi^{-1}(F')$. $\square$

Remarque 3.2.2. Notons que $\varphi^{-1}(F')$ est bien défini, même dans le cas où φ n'est pas un isomorphisme!

On s'intéressera en particulier aux cas $E' = E$ et $F' = \{0_F\}$, pour obtenir la définition suivante qui étend la définition 1.3.12:

Définition 3.2.3. Soient E, F deux espaces vectoriels et $\varphi\colon E \to F$ une application linéaire. On appelle:

- *noyau de* φ le sous espace $\ker(\varphi) := \varphi^{-1}(\{0_F\})$.
- *image de* φ le sous-espace $\mathrm{Im}(\varphi) := \varphi(E)$.

Remarque 3.2.4. Plus concrètement, on a

$$\ker(\varphi) = \{x \in E | \varphi(x) = 0_F\}$$

et

$$\mathrm{Im}(\varphi) = \{\varphi(v) | v \in E\}.$$

On a déjà vu la proposition suivante dans le cadre des groupes (voir la remarque 1.3.14). La preuve est ainsi omise.

Proposition 3.2.5. *Une application linéaire $\varphi\colon E \to F$ est:*

a) injective si et seulement si $\ker(\varphi) = \{0_E\}$,

b) surjective si et seulement si $\mathrm{Im}(\varphi) = F$.

Exemple 3.2.6. Soit E un espace vectoriel et $v_1, \dots, v_n \in E$. On considère l'application linéaire $\varphi\colon \mathbb{R}^n \to E$ donnée par

$$\varphi(\begin{pmatrix} \lambda_1 \\ \vdots \\ v_n \end{pmatrix}) = \sum_{i=1}^{n} \lambda_i v_i$$

On a alors $\mathrm{Im}(\varphi) = \mathrm{Vect}(v_1, \dots, v_n)$ et $\ker(\varphi) = \{\lambda_1, \dots, \lambda_n \in \mathbb{R} | \sum_{i=1}^n \lambda_i v_i = 0\}$.

3.3 Rang d'une application linéaire

Définition 3.3.1. Soient E, F deux espaces vectoriels et $\varphi\colon E \to F$ une application linéaire. On dit que φ est de *rang fini* si $\mathrm{Im}(\varphi)$ est de dimension finie. Dans ce cas, on appelle *rang de* φ l'entier naturel

$$\mathrm{rang}(\varphi) := \dim(\mathrm{Im}(\varphi)).$$

Remarque 3.3.2. Bien évidemment, si F est de dimension finie, alors toute application linéaire est de rang fini. De plus, on a que φ est surjective si et seulement si $\mathrm{Im}(\varphi) = F$ si et seulement si $\mathrm{rang}(\varphi) = \dim(F)$. Il est moins évident que toute application linéaire est de rang fini si E est de dimension finie. Cela suit du théorème suivant.

Théorème 3.3.3 (Théorème du rang)**.** *Soient E, F deux espaces vectoriels réels tels que E est de dimension finie. Si $\varphi\colon E \to F$ est une application linéaire, alors elle est de rang finie et on a*

$$\mathrm{rang}(\varphi) + \dim(\ker(\varphi)) = \dim(E). \tag{3.3.1}$$

Proof. Posons $n := \dim(E)$. On remarque tout d'abord que si $\ker(\varphi) = E$, alors $\varphi = 0$ et l'égalité ci-dessus est évidente. Si $\ker(\varphi) = \{0_E\}$, alors $\varphi\colon E \to \mathrm{Im}(\varphi)$ est un isomorphisme et donc $\mathrm{rang}(\varphi) = \dim(E)$. On suppose donc que $\{0_E\} \subsetneq \ker(\varphi) \subsetneq E$. Choisissons une base $\{e_1, \ldots, e_k\}$ de $\ker(\varphi)$, et complétons-la en une base $\{e_1, \ldots, e_k, e_{k+1}, \ldots, e_n\}$ de E. On va démontrer que la famille $\{\varphi(e_{k+1}), \ldots, \varphi(e_n)\}$ est une base de $\mathrm{Im}(\varphi)$. Montrons tout d'abord qu'elle est libre. Si $\lambda_{k+1}, \ldots, \lambda_n \in \mathbb{R}$ sont tels que $\sum_{i=k+1}^n \lambda_i \varphi(e_i) = 0_F$ on a alors on a

$$0_F = \sum_{i=k+1}^{n} \lambda_i \varphi(e_i) = \varphi(\sum_{i=k+1}^{n} \lambda_i e_i)$$

et ainsi $\sum_{i=k+1}^n \lambda_i e_i \in \ker(\varphi)$. Il existe donc $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ tels que $\sum_{i=k+1}^n \lambda_i e_i = \sum_{i=1}^k \lambda_i e_i$. Ainsi

$$\sum_{i=1}^{k} \lambda_i e_i + \sum_{i=k+1}^{n} (-\lambda_i) e_i = 0_E$$

d'où on en déduit que $\lambda_1 = \ldots = \lambda_n = 0$ puisque $\{e_1, \ldots, e_k, e_{k+1}, \ldots, e_n\}$ est libre. Donc $\{\varphi(e_{k+1}), \ldots, \varphi(e_n)\}$ est bien une famille libre de F.

On prouve maintenant que la famille $\{\varphi(e_{k+1}), \ldots, \varphi(e_n)\}$ engendre $\mathrm{Im}(\varphi)$. Soit donc $\varphi(v) \in \mathrm{Im}(\varphi)$. Puisque $v \in E$, il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tels que $v = \sum_{i=1}^n \lambda_i e_i$. Par linéarité de φ, on obtient

$$\varphi(v) = \varphi(\sum_{i=1}^{n} \lambda_i e_i) = \sum_{i=1}^{n} \lambda_i \varphi(e_i)$$

Or $e_1, \ldots, e_k$ sont des éléments de $\ker(\varphi)$, et ainsi $\varphi(e_1) = \ldots = \varphi(e_k) = 0$. L'expression ci-dessus devient donc

$$\varphi(v) = \sum_{i=k+1}^{n} \lambda_i \varphi(e_i),$$

ce qui montre bien que la famille est génératrice et donc une base. Pour conclure, on voit que $\mathrm{Im}(\varphi)$ est engendré par une famille finie, et donc de dimension finie. Ceci montre bien que φ est de rang fini. Par ailleurs, on obtient

$$\dim(E) = n = n - k + k = \mathrm{rang}(\varphi) + \dim(\ker(\varphi)),$$

ce qu'il fallait montrer. $\square$

Exemple 3.3.4. Soit $\varphi\colon \mathbb{R}^3 \to \mathbb{R}^2$ définie par

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + 2x_2 \\ x_2 + 3x_3 \end{pmatrix}.$$

Comme $\varphi(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\varphi(\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ on voit que φ est surjective. On en déduit qu'elle est de rang 2, et donc que $\ker(\varphi)$ est de dimension 1 puisque

$$3 = \dim(\mathbb{R}^3) = \mathrm{rang}(\varphi) + \dim(\ker(\varphi)) = 2 + \dim(\ker(\varphi)).$$

En fait un calcul direct montre que $\ker(\varphi) = \mathrm{Vect}(\begin{pmatrix} 6 \\ -3 \\ 1 \end{pmatrix})$.

Remarque 3.3.5. Les mêmes arguments montrent que si φ est de rang fini et si $\ker(\varphi)$ est de dimension finie, alors E est de dimension finie et l'égalité (3.3.1) est vérifiée. En effet, si $\{e_1, \dots, e_k\}$ est une base de $\ker(\varphi)$ et si $\{v_1, \dots, v_l\}$ est une base de $\mathrm{Im}(\varphi)$, alors on choisit $\{u_1, \dots, u_l\} \subset E$ tels que $\varphi(u_i) = v_i$ pour tout $i = 1, \dots, l$. On vérifie ensuite que $\{e_1, \dots, e_k, u_1, \dots, u_l\}$ est une base de E.

Corollaire 3.3.6. *Soient E, F des espaces vectoriels de dimension finie tels que* $\dim(E) = \dim(F)$. *Si $\varphi\colon E \to F$ est une application linéaire, alors*

$$\varphi \text{ est injective} \iff \varphi \text{ est surjective} \iff \varphi \text{ est bijective.}$$

Proof. Supposons que φ est injective. Alors $\dim(\ker(\varphi)) = 0$ et le théorème du rang donne $\mathrm{rang}(\varphi) = \dim(E) = \dim(F)$. On en déduit que $\mathrm{Im}(\varphi) = F$ et donc que φ est surjective. Si maintenant φ est surjective, alors $\mathrm{rang}(\varphi) = \dim(F) = \dim(E)$, d'où l'on déduit immédiatement que $\ker(\varphi) = \{0_E\}$. Ainsi, φ est injective. $\square$

Corollaire 3.3.7. *Soit E un espace vectoriel, et $F_1, F_2 \subset E$ deux sous-espaces vectoriels de dimension finie. Alors $F_1 + F_2$ est aussi de dimension finie et*

$$\dim(F_1) + \dim(F_2) = \dim(F_1 + F_2) + \dim(F_1 \cap F_2).$$

Proof. On peut sans autre supposer que F_1 et F_2 sont non nuls, sans quoi le résultat est évident. Soit $F_1 \times F_2 = \{(v_1, v_2) | v_1 \in F_1, v_2 \in F_2\}$. Cet ensemble est muni d'une addition définie par

$$(v_1, v_2) + (w_1, w_2) = (v_1 + w_1, v_2 + w_2) \;\; \forall v_1, w_1 \in F_1, v_2, w_2 \in F_2$$

et d'une loi externe définie par

$$\lambda(v_1, v_2) = (\lambda v_1, \lambda v_2) \ \forall v_1 \in F_1, v_2 \in F_2, \lambda \in \mathbb{R}.$$

On vérifie immédiatement que ces deux lois munissent $F_1 \times F_2$ d'une structure d'espace vectoriel réel. Si $\{v_1, \ldots, v_n\}$ est une base de F_1 et $\{w_1, \ldots, w_m\}$ est une base de F_2, alors on montre directement que l'ensemble

$$\{(v_1, 0), \ldots, (v_n, 0), (0, w_1), \ldots, (0, w_m)\}$$

est une base de $F_1 \times F_2$. On en déduit que $\dim(F_1 \times F_2) = \dim(F_1) + \dim(F_2)$. Par ailleurs, on peut définir une application

$$\varphi\colon F_1 \times F_2 \to E$$

par $\varphi(v_1, v_2) = v_1 + v_2$, qu'on vérifie être linéaire. On a

$$\begin{aligned} \mathrm{Im}(\varphi) &= \{\varphi(v_1, v_2) | v_1 \in F_1, v_2 \in F_2\} \\ &= \{v_1 + v_2 | v_1 \in F_1, v_2 \in F_2\} \\ &= F_1 + F_2. \end{aligned}$$

Par ailleurs, $(v_1, v_2) \in \ker(\varphi) \iff \varphi(v_1, v_2) = 0 \iff v_1 + v_2 = 0 \iff v_1 = -v_2$. On en déduit que $v_1 \in F_1 \cap F_2$ et que $v_2 = -v_1 \in F_1 \cap F_2$. Ainsi, l'application linéaire $\psi : F_1 \cap F_2 \to \ker(\varphi)$ définie par $\varphi(v) = (v, -v)$ est surjective. On voit immédiatement qu'elle est aussi injective, et on en déduit que c'est un isomorphisme. Ainsi $\dim(\ker(\varphi)) = \dim(F_1 \cap F_2)$. Le théorème du rang devient alors

$$\begin{aligned} \dim(F_1) + \dim(F_2) &= \dim(F_1 \times F_2) \\ &= \dim(\ker(\varphi)) + \dim(\mathrm{Im}(\varphi)) \\ &= \dim(F_1 \cap F_2) + \dim(F_1 + F_2). \end{aligned}$$

□

3.4 Formes linéaires et hyperplans

Définition 3.4.1. Soit E un espace vectoriel réel.

i) On appelle *forme linéaire* sur E toute application linéaire $\varphi\colon E \to \mathbb{R}$. L'ensemble $\mathcal{L}(E, \mathbb{R})$ de toutes les formes linéaires sur E est un espace vectoriel, appelé *espace dual de E*. On le note E^*.

ii) On dit qu'un sous-espace $F \subset E$ est un *hyperplan vectoriel* si F est le noyau d'une forme linéaire non nulle.

Remarque 3.4.2. Si $\varphi\colon E \to \mathbb{R}$ est non nulle, alors $\mathrm{Im}(\varphi) = \mathbb{R}$ et on a $\mathrm{rang}(\varphi) = 1$.

Exemple 3.4.3. Soit $n \in \mathbb{N}$. Si $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$, on obtient une forme linéaire $\varphi_a\colon \mathbb{R}^n \to \mathbb{R}$ définie par

$$\varphi_a(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = \sum_{i=1}^{n} a_i x_i$$

Si a est non nul (i.e. au moins l'un de ses composantes est non nulle), alors φ_a est aussi non nulle. En effet $\varphi_a(a) = \sum_{i=1}^n a_i^2 > 0$. Réciproquement, si $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ est non nulle, alors on obtient un n-uple $(a_1, \dots, a_n)$ défini par $\varphi(e_i) = a_i$, où $\{e_1, \dots, e_n\}$ est la base canonique de $\mathbb{R}^n$. On vérifie que ces deux constructions sont inverses l'une de l'autre, identifiant ainsi $(\mathbb{R}^n)^*$ avec $\mathbb{R}^n$. Nous reviendrons là-dessus un peu plus tard dans le cours.

Quoiqu'il en soit, pour $\varphi_a\colon \mathbb{R}^n \to \mathbb{R}$ comme ci-dessus, on obtient $\ker(\varphi_a) = \{x \in \mathbb{R}^n | \sum_{i=1}^n a_i x_i = 0\}$ qui peut s'interpréter comme l'hyperplan orthogonal à a pour le produit scalaire usuel.

Proposition 3.4.4. *Un sous-espace vectoriel $F \subset E$ est un hyperplan vectoriel si et seulement s'il existe un sous-espace $G \subset E$ de dimension 1 tel que $E = F \oplus G$.*

Proof. Supposons que $F = \ker(\varphi)$ pour une forme linéaire $\varphi\colon E \to \mathbb{R}$ non nulle. Comme on l'a observé ci-dessus, φ est surjective et il existe donc $u \in E$ tel que $\varphi(u) = 1$. En particulier, u est non nul et $\mathrm{Vect}(u) \subset E$ est un sous-espace de dimension 1. Pour conclure, il suffit de montrer que $E = \ker(\varphi) \oplus \mathrm{Vect}(u)$. Suppposons que $v \in \ker(\varphi) \cap \mathrm{Vect}(u)$. Alors $\varphi(v) = 0$ et il existe $\lambda \in \mathbb{R}$ tel que $v = \lambda(u)$. Or

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

et donc $v = 0$. On en tire que $\ker(\varphi) \cap \mathrm{Vect}(u) = \{0_E\}$. Si maintenant $v \in E$, on considère le vecteur $v - \varphi(v)u \in E$. On a

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

ce qui donne $v - \varphi(v)u \in \ker(\varphi)$. Pour finir, on peut écrire $v = v - \varphi(v)u + \varphi(v)u$ ce qui montre que $E = \ker(\varphi) + \mathrm{Vect}(u)$.

Réciproquement, supposons que $E = F \oplus G$ avec G de dimension 1. Soit $u \in G$ non nul. Alors $G = \mathrm{Vect}(u)$ et tout vecteur v de E s'écrit de manière unique

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

avec $\lambda(v) \in \mathbb{R}$ et $w \in F$. On vérifie directement que l'application $\varphi\colon E \to \mathbb{R}$ définie par $\varphi(v) = \lambda(v)$ est linéaire, et que $\varphi(u) = 1$. Cette forme linéaire est donc non nulle, et on termine en constatant que $\ker(\varphi) = F$. □

Corollaire 3.4.5. *Soit E un espace vectoriel de dimension finie. Alors un sous-espace vectoriel $F \subset E$ est un hyperplan vectoriel si et seulement si $\dim(F) = \dim(E) - 1$.*

Proof. Au vu de la proposition, il faut démontrer que $\dim(F) = \dim(E) - 1$ si et seulement si $E = F \oplus G$ avec $\dim(G) = 1$. Si $E = F \oplus G$ avec $\dim(G) = 1$, alors on a bien $\dim(F) = \dim(E) - 1$ par la proposition 2.5.13. Réciproquement, on choisit un supplémentaire G de F dans E, et on utilise la même proposition pour voir que G est de dimension 1. □

La proposition ci-dessous se démontre essentiellement comme la proposition ci-dessus, et sa preuve sera donc omise.

Proposition 3.4.6. *Soit E un espace vectoriel de dimension n, et soit $k \in \{1, \dots, n\}$. Si $F \subset E$ est un sous-espace vectoriel, les affirmations suivantes sont équivalentes:*

i) $\dim(F) = n - k$.

ii) *F admet un supplémentaire G de dimension k dans E.*

iii) *$F = \ker(\varphi)$, où $\varphi \colon E \to \mathbb{R}^k$ est linéaire et surjective.*

3.5 Applications linéaires en dimension finie

Soit E un espace vectoriel réel de dimension finie, et soit $\{e_1, \ldots, e_n\}$ une base de E. On sait alors que tout vecteur $v \in E$ admet une écriture

$$v = \sum_{i=1}^{n} \lambda_i(v) e_i$$

pour d'uniques scalaires $\lambda_1(v), \ldots, \lambda_n(v) \in \mathbb{R}$. On obtient pour tout $i = 1, \ldots, n$ ainsi une application

$$\lambda_i \colon E \to \mathbb{R}, v \mapsto \lambda_i(v).$$

Lemme 3.5.1. *Pour tout $i \in \{1, \ldots, n\}$, l'application $\lambda_i \colon E \to \mathbb{R}$ est linéaire.*

Proof. Si $v, w \in E$ et $\mu \in \mathbb{R}$, on peut écrire $v = \sum_{i=1}^{n} \lambda_i(v) e_i$ et $w = \sum_{i=1}^{n} \lambda_i(w) e_i$. On obtient alors

$$v + \mu w = \sum_{i=1}^{n} \lambda_i(v) e_i + \mu \sum_{i=1}^{n} \lambda_i(w) e_i) \sum_{i=1}^{n} (\lambda_i(v) + \mu \lambda_i(w)) e_i.$$

Par unicité de la décomposition d'un vecteur dans une base, on en conclut que $\lambda_i(v + \mu w) = \lambda_i(v) + \mu \lambda_i(w)$ pour tout $i = 1, \ldots, n$. □

Remarque 3.5.2. Par la remarque 3.1.5, le choix de la base $\{e_1, \ldots, e_n\}$ induit un isomorphisme $\varphi \colon \mathbb{R}^n \to E$ explicitement donné par

$$\varphi(\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}) = \sum_{i=1}^{n} \lambda_i e_i.$$

L'isomorphisme $\varphi^{-1} \colon E \to \mathbb{R}^n$ admet alors la description suivante:

$$\varphi^{-1}(v) = \begin{pmatrix} \lambda_1(v) \\ \vdots \\ \lambda_n(v) \end{pmatrix}$$

et l'application linéaire λ_i du lemme ci-dessus n'est autre que la composition de φ^{-1} suivie de la projection sur le i-ème facteur.

Remarque 3.5.3. Par construction, on a

$$\lambda_i(e_j) = \delta_{ij} := \begin{cases} 1 & \text{si } i = j. \\ 0 & \text{si } i \neq j. \end{cases} \tag{3.5.1}$$

Proposition 3.5.4. *Les formes linéaires* $\{\lambda_1, \ldots, \lambda_n\}$ *forment une base de* $E^* = \mathcal{L}(E, \mathbb{R})$.

Proof. On démontre tout d'abord que cette famille est libre. Supposons que $\mu_1, \ldots, \mu_n \in \mathbb{R}$ soient tels que $\sum_{i=1}^n \mu_i \lambda_i = 0_{E^*}$. On en déduit que pour tout $v \in E$, on a

$$0 = 0_{E^*}(v) = (\sum_{i=1}^n \mu_i \lambda_i)(v) = \sum_{i=1}^n \mu_i \lambda_i(v).$$

En particulier, prenant $v = e_1$ et utilisant (3.5.1), on obtient

$$0 = \sum_{i=1}^n \mu_i \lambda_i(e_1) = \lambda_1.$$

Répétant l'argument avec $e_2, \ldots, e_n$, on obtient bien $\lambda_1 = \ldots = \lambda_n = 0$.

On montre maintenant que la famille $\{\lambda_1, \ldots, \lambda_n\}$ est génératrice. Supposons donc que $\varphi \colon E \to \mathbb{R}$ soit une forme linéaire. On obtient alors des scalaires $\varphi(e_1), \ldots, \varphi(e_n) \in \mathbb{R}$ et on peut considérer la forme linéaire $\psi := \sum_{i=1}^n \varphi(e_i)\lambda_i$. On a

$$\psi(e_1) = (\sum_{i=1}^n \varphi(e_i)\lambda_i)(e_1) = \sum_{i=1}^n \varphi(e_i)\lambda_i(e_1) = \varphi(e_1)$$

Répétant l'argument avec $e_2, \ldots, e_n$, on voit que $\varphi(e_i) = \psi(e_i)$ pour tout $i = 1, \ldots .n$. On en déduit enfin que $\varphi(v) = \psi(v)$ pour toute combinaison linéaire des vecteurs $\{e_1, \ldots, e_n\}$, et donc pour tout $v \in E$. □

Remarque 3.5.5. Ce résultat montre en particulier que si E est de dimension finie, alors E^* est de même dimension que E. On dit souvent que $\{\lambda_1, \ldots, \lambda_n\}$ est la base duale de $\{e_1, \ldots, e_n\}$ et on note $e_i^* := \lambda_i$ pour tout $i = 1, \ldots, n$.

Terminons cette section avec une proposition qui revêt une grande importance pratique.

Proposition 3.5.6. *Soient* E, F *deux espaces vectoriels réels avec* $\dim(E) < \infty$. *Si* $\{e_1, \ldots, e_n\}$ *est une base de* E *et* $\{f_1, \ldots, f_n\}$ *est une famille quelconque de vecteurs de* F*, il existe une unique application linéaire*

$$\varphi \colon E \to F$$

telle que $\varphi(e_i) = f_i$ *pour tout* $i = 1, \ldots, n$.

Proof. On définit $\varphi \colon E \to F$ par

$$\varphi(v) = \sum_{i=1}^n \lambda_i(v) f_i$$

et on observe immédiatement que φ est linéaire puisque les formes $\lambda_1, \ldots, \lambda_n$ le sont. Du fait que $\lambda_i(e_j) = \delta_{ij}$, on obtient bien que

$$\varphi(e_i) = f_i$$

pour tout $i = 1, \ldots, n$. Si maintenant $\psi \colon E \to F$ est une application linéaire satisfaisant la même propriété, on tire que $\varphi(e_i) = \psi(e_i)$ pour tout $i = 1, \ldots, n$, et on en déduit comme dans la preuve précédente que $\varphi = \psi$. □

3.6 Projections et symétries

Dans toute cette section, E désigne un espace vectoriel réel.

Définition 3.6.1. On dit qu'un endomorphisme $p\colon E \to E$ est une projection si $p \circ p = p$ (i.e. p est *idempotente*).

Cette définition est liée au fait de pouvoir décomposer E en somme directe, comme le montre la proposition suivante.

Proposition 3.6.2. *Les assertions suivantes sont vérifiées:*

a) Si $p\colon E \to E$ est une projection, alors $E = \ker(p) \oplus \operatorname{Im}(p)$.

b) Si $E = F \oplus G$, alors il existe une unique projection $p\colon E \to E$ telle que $\ker(p) = F$ et $\operatorname{Im}(p) = G$. On dit que p est la projection de E sur G le long de F (ou parallèlement à F).

Proof. On démontre tout d'abord que $E = \ker(p) \oplus \operatorname{Im}(p)$ si p est une projection. Supposons que $x \in \ker(p) \cap \operatorname{Im}(p)$. Il existe alors $y \in E$ tel que $x = p(y)$. Du fait que $x \in \ker(p)$, on trouve alors

$$0 = p(x) = p(p(y)) = p(y) = x$$

et on en déduit que $\ker(p) \cap \operatorname{Im}(p) = \{0_E\}$. Pour voir que $\ker(p) + \operatorname{Im}(p) = E$, on observe que $x - p(x) \in \ker(p)$. En effet,

$$p(x - p(x)) = p(x) - p(p(x)) = p(x) - p(x) = 0_E.$$

Par ailleurs, $p(x) \in \operatorname{Im}(p)$, et on a donc $x = (x - p(x)) + p(x) \in \ker(p) + \operatorname{Im}(p)$.

Supposons maintenant que $E = F \oplus G$. Tout $x \in E$ se décompose de manière unique en

$$x = y + z \tag{3.6.1}$$

avec $y \in F$ et $z \in G$. On définit donc $p\colon E \to E$ par $p(x) = z$ et on observe que l'application p est bien linéaire par unicité de la décomposition (3.6.1). En effet, si $x' \in E$ admet une décomposition $x' = y' + z'$ avec $y' \in F$ et $z' \in G$ on a

$$x + \lambda x' = (y + z) + \lambda(y' + z') = (y + \lambda y') + (z + \lambda z'),$$

d'où l'on tire que $p(x+\lambda x') = z+\lambda z' = p(x)+\lambda p(x')$. On vérifie ensuite que $p\circ p = p$. Si $x = y+z$ comme dans la décomposition (3.6.1), on a $p(x) = z$. De plus, on peut écrire $z = 0 + z$ puisque $z \in G$ ce qui montre que $p(z) = z$ et ainsi que $p \circ p = p$. Pour finir, on prouve que $\ker(p) = F$ et que $\operatorname{Im}(p) = G$. Si $z \in G$, on a vu ci-dessus que $p(z) = z$, ce qui montre que $G \subset \operatorname{Im}(p)$. Par ailleurs, $\operatorname{Im}(p) \subset G$ par définition. Si $y \in F$, on a $y = y + 0$ et donc $p(y) = 0$. Finalement, si $x \in \ker(p)$, on écrit la décomposition (3.6.1) $x = y + z$ et on observe que $0 = p(x) = z$, et donc que $x = y \in F$.

Supposons finalement que $p'\colon E \to E$ soit une projection avec $\ker(p') = F$ et $\operatorname{Im}(p') = G$. On écrit

$$x = (x - p'(x)) + p'(x)$$

et on voit comme ci-dessus que $x - p'(x) \in \ker(p') = F$ et $p'(x) \in \operatorname{Im}(p') = G$. Par définition de p, on a alors $p(x) = p'(x)$, ce qui démontre l'unicité. $\square$

Remarque 3.6.3. Si $p\colon E \to E$ est une projection, alors $\mathrm{Id} - p$ est également une projection. En effet,

$$(\mathrm{Id} - p) \circ (\mathrm{Id} - p) = \mathrm{Id} - 2p + p \circ p = \mathrm{Id} - 2p + p = \mathrm{Id} - p.$$

On vérifie facilement que $\ker(\mathrm{Id} - p) = \mathrm{Im}(p)$ et que $\mathrm{Im}(\mathrm{Id} - p) = \ker(p)$.

Exemple 3.6.4. Soit $E = \mathbb{R}^3$, $F = \mathrm{Vect}(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix})$ et $G = \{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 | x_1 + x_2 + x_3 = 0\}$. On voit que G est un sous-espace vectoriel de E, puisqu'il s'agit du noyau de la forme linéaire $f\colon \mathbb{R}^3 \to \mathbb{R}$ définie par $f(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}) = x_1 + x_2 + x_3$. Comme $f(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}) = 3 \neq 0$, on tire de la preuve de la proposition 3.4.4 que $E = F \oplus G$. Si $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$, on a

$$x - \frac{x_1 + x_2 + x_3}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \ker(f) = G$$

et on obtient que la projection $p\colon E \to E$ sur G le long de F est donnée par $p(x) = x - \frac{x_1+x_2+x_3}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Plus précisément, on obtient

$$p(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}) = \frac{1}{3} \begin{pmatrix} 2x_1 - x_2 - x_3 \\ -x_1 + 2x_2 - x_3 \\ -x_1 - x_2 + 2x_3 \end{pmatrix}.$$

Proposition 3.6.5. *Si $p\colon E \to E$ est une projection et $f\colon E \to E$ est un endomorphisme, les propriétés suivantes sont équivalentes:*

i) f commute avec p, i.e. $f \circ p = p \circ f$.

ii) f laisse invariants les sous-espaces $\ker(p)$ et $\mathrm{Im}(p)$: $f(\ker(p)) \subset \ker(p)$ et $f(\mathrm{Im}(p)) \subset \mathrm{Im}(p)$.

Proof. Supposons tout d'abord que $f \circ p = p \circ f$. Si $x \in \ker(p)$, on a $p(f(x)) = f(p(x)) = f(0) = 0$, et donc $f(x) \in \ker(p)$. On obtient ainsi $f(\ker(p)) \subset \ker(p)$. De même, si $y = p(x)$, on a $f(y) = f(p(x)) = p(f(x)) \in \mathrm{Im}(p)$.

Réciproquement, supposons que $f(\ker(p)) \subset \ker(p)$ et $f(\mathrm{Im}(p)) \subset \mathrm{Im}(p)$. Si $x \in E$, on écrit $x = y + z$ avec $y \in \ker(p)$ et $z \in \mathrm{Im}(p)$. On obtient

$$(f \circ p)(x) = f(p(x)) = f(p(z)) = f(z).$$

D'un autre côté, on a $f(x) = f(y) + f(z)$ avec $f(y) \in \ker(p)$ et $f(z) \in \mathrm{Im}(p)$. Ainsi

$$(p \circ f)(x) = p(f(x)) = p(f(z)) = f(z).$$

On en déduit bien que $f \circ p = p \circ f$. □

Sous les hypothèses de la proposition, il suffit (pour étudier l'application f) d'étudier les restrictions de f aux sous-espaces $\ker(p)$ et $\mathrm{Im}(p)$. On pose:

$$E = \ker(p) \oplus \mathrm{Im}(p); f = f_1 \oplus f_2$$

avec $f_1 = f_{|\ker(p)}$ et $f_2 = f_{|\mathrm{Im}(p)}$.

Définition 3.6.6. On dit qu'une application linéaire $s\colon E \to E$ est une symétrie (ou une involution) si $s \circ s = \mathrm{Id}$.

Proposition 3.6.7. *Les assertions suivantes sont vérifiées:*

i) Si $s\colon E \to E$ est une symétrie, alors $E = F \oplus G$ où

$$F := \{x \in E | s(x) = -x\} = \ker(s + \mathrm{Id}), G := \{x \in E | s(x) = x\} = \ker(s - \mathrm{Id}).$$

ii) Si $E = F \oplus G$ où F est G sont des sous-espaces vectoriels, il existe une unique symétrie $s\colon E \to E$ telle que $F = \ker(s + \mathrm{Id})$ et $G = \ker(s - \mathrm{Id})$. On dit que s est une symétrie par rapport à G parallèlement à F.

Proof. Soit $x \in F \cap G$. Alors $s(x) = x$ et $s(x) = -x$. On obtient alors $x = -x$, i.e. $2x = 0$. Comme 2 est inversible dans $\mathbb{R}$, on obtient $x = 0$. Soit maintenant $x \in E$. On a

$$x = \frac{1}{2}(x - s(x)) + \frac{1}{2}(x + s(x))$$

et on vérifie que $\frac{1}{2}(x - s(x)) \in F$ et $\frac{1}{2}(x + s(x)) \in G$. En effet,

$$s(\frac{1}{2}(x - s(x))) = \frac{1}{2}s(x - s(x)) = \frac{1}{2}(s(x) - s(s(x)) = \frac{1}{2}(s(x) - x) = -\frac{1}{2}(x - s(x))$$

et le même calcul montre que $\frac{1}{2}(x + s(x)) \in G$. En conclusion, on a bien $E = F \oplus G$.

Supposons maintenant que $E = F \oplus G$. Si $x \in E$, on écrit $x = y + z$ avec $y \in F$ et $z \in G$. On pose alors $s(x) = -y + z$, et on vérifie immédiatement que s est linéaire, satisfait $s \circ s = \mathrm{Id}$ et que $F = \ker(Id + s)$, $G = \ker(\mathrm{Id} - s)$. Il reste encore à vérifier que s est unique avec ces propriétés. Supposons que $s'\colon E \to E$ satisfasse les mêmes propriétés. On a alors pour tout $x \in E$

$$x = \frac{1}{2}(x - s'(x)) + \frac{1}{2}(x + s'(x))$$

avec $\frac{1}{2}(x - s'(x)) \in F$ et $\frac{1}{2}(x + s'(x)) \in G$. Ainsi

$$s(x) = -\frac{1}{2}(x - s'(x)) + \frac{1}{2}(x + s'(x)) = s'(x).$$

□

Remarque 3.6.8. Si $s\colon E \to E$ est une symétrie, alors $p := \frac{1}{2}(\mathrm{Id} + s)$ et $\mathrm{Id} - p = \frac{1}{2}(\mathrm{Id} - s)$ sont des projections. Inversement, si p est une projection, alors $2p - \mathrm{Id}$ et $\mathrm{Id} - 2p$ sont des symétries.

Remarque 3.6.9. Si $s\colon E \to E$ est la symétrie par rapport à G le long de F, alors $-s$ est la symétrie par rapport à F le long de G, et $p = \frac{1}{2}(\mathrm{Id} + s)$ est la projection sur G le long de F.

4 Calcul matriciel

4.1 Définitions générales

Définition 4.1.1. Soient $m, n \in \mathbb{N} \setminus \{0\}$. Une *matrice à m lignes et n colonnes* à coefficients réels est une famille $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$ de $m \cdot n$ scalaires notée sous forme d'un tableau rectangulaire

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i,1} & \dots & a_{i,j} & \dots & a_{i,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,j} & \dots & a_{m,n} \end{pmatrix}.$$

L'indice $i \in \{1, \dots, m\}$ est l'indice de la ligne, et l'indice $j \in \{1, \dots, m\}$ est celui de la colonne. Le réel $a_{i,j}$ est le coefficient d'indices i, j de la matrice $A = (a_{i,j})$. L'ensemble des matrices à m lignes et n colonnes est noté $\mathcal{M}_{m,n}(\mathbb{R})$.

Exemple 4.1.2.
$$A = \begin{pmatrix} 1 & 2 & 5 \\ 0 & 7 & -3 \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{R}).$$

Remarque 4.1.3. Lorsqu'il n'y a pas de confusion possible, on note aussi a_{ij} le scalaire $a_{i,j}$. Attention toutefois que $a_{1,23} \neq a_{12,3}$ en général.

L'ensemble $\mathcal{M}_{m,n}(\mathbb{R})$ des matrices à m lignes et n colonnes est muni d'une structure d'espace vectoriel, comme le montre la proposition suivante dont la preuve est laissée en exercice.

Proposition 4.1.4. *L'ensemble $\mathcal{M}_{m,n}(\mathbb{R})$ muni des opérations suivantes*

- *(Addition) Si $A = (a_{i,j}) \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B = (b_{i,j}) \in \mathcal{M}_{m,n}(\mathbb{R})$, on pose*
$$A + B = (a_{i,j} + b_{i,j}).$$
- *(Loi externe) Si $A = (a_{i,j}) \in \mathcal{M}_{m,n}(\mathbb{R})$ et $\lambda \in \mathbb{R}$, alors on pose*
$$\lambda A = (\lambda a_{i,j}).$$

est un espace vectoriel réel.

En d'autres termes, l'addition et la multiplication externe sont définies coefficient par coefficient.

Définition 4.1.5. Si $k \in \{1, \dots, m\}$ et $l \in \{1, \dots, n\}$, on note $E_{k,l} \in \mathcal{M}_{m,n}(\mathbb{R})$ la matrice définie par

$$E_{k,l} = (m_{i,j}), \text{ où } m_{i,j} = \begin{cases} 1 & \text{si } (i,j) = (k,l). \\ 0 & \text{si } (i,j) \neq (k,l) \end{cases}$$

Exemple 4.1.6. Si $m = 2$ et $n = 3$, on a

$$E_{1,1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, E_{1,2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \dots, E_{2,3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Avec cette définition, une matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ s'écrit de la manière suivante:

$$A = \sum_{i=1}^{m} \sum_{j=1}^{m} a_{i,j} E_{i,j},$$

et cette écriture est unique. On a ainsi obtenu le résultat suivant.

Proposition 4.1.7. *La famille de matrices* $\{E_{i,j} | 1 \leq i \leq m, 1 \leq j \leq m\}$ *est une base de* $\mathcal{M}_{m,n}(\mathbb{R})$*, appelée* base canonique.

Corollaire 4.1.8. *L'espace vectoriel* $\mathcal{M}_{m,n}(\mathbb{R})$ *est de dimension* $m \cdot n$*.*

Remarque 4.1.9. L'espace vectoriel $\mathcal{M}_{m,1}$ est isomorphe à $\mathbb{R}^m$. Les notions de bases canoniques coïncident.

4.2 Produit matriciel

Définition 4.2.1. Soit $m, n, p \geq 1$ des entiers, $A = (a_{i,j}) \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B \in \mathcal{M}_{n,p}(\mathbb{R})$. On appelle *produit matriciel de* A *par* B la matrice $C = AB = (c_{i,j}) \in \mathcal{M}_{m,p}(\mathbb{R})$ définie par

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}, \text{ pour } 1 \leq i \leq m \text{ et } 1 \leq j \leq p.$$

Remarque 4.2.2. Notons que le produit matriciel n'est défini que lorsque le nombre de colonnes n de A est égal au nombre de lignes de B.

Exemple 4.2.3. Soit $n \geq 1$ un entier, $A = (a_{i,j}) \in \mathcal{M}_{1,n}(\mathbb{R})$ et $B = (b_{i,j}) \in \mathcal{M}_{n,1}(\mathbb{R})$. Alors

$$AB = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \end{pmatrix} \begin{pmatrix} b_{1,1} \\ b_{2,1} \\ \vdots \\ b_{n,1} \end{pmatrix} = \left(\sum_{k=1}^{n} a_{1,k} b_{k,1} \right)$$

n'est autre que le produit scalaire de A et B. Plus généralement, si $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B \in \mathcal{M}_{n,p}(\mathbb{R})$ et $C = AB$, alors $c_{i,j}$ est le produit scalaire de la i-ème ligne de A avec la j-ième colonne de B.

Exemple 4.2.4. Si $A = \begin{pmatrix} 1 & 2 & 5 \\ 0 & 7 & -3 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ 1 & 2 \\ -1 & 0 \end{pmatrix}$ alors

$$AB = \begin{pmatrix} -3 & 5 \\ 10 & 14 \end{pmatrix}, BA = \begin{pmatrix} 0 & 7 & -3 \\ 1 & 16 & -1 \\ -1 & -2 & -5 \end{pmatrix}.$$

Nous admettrons la proposition suivante, dont la démonstration est une vérification directe.

Proposition 4.2.5. *Soient* $m, n, p, q \geq 1$ *des entiers.*

1. *(Associativité) Si* $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $B \in \mathcal{M}_{n,p}(\mathbb{R})$ *et* $C \in \mathcal{M}_{p,q}(\mathbb{R})$ *alors*
$$A(BC) = (AB)C \in \mathcal{M}_{m,q}(\mathbb{R}).$$
2. *(Linéarité à droite) Si* $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $B, C \in \mathcal{M}_{n,p}(\mathbb{R})$ *et* $\lambda \in \mathbb{R}$, *alors*
$$A(B + \lambda C) = AB + \lambda AC \in \mathcal{M}_{m,p}(\mathbb{R}).$$
3. *(Linéarité à gauche) Si* $A, B \in \mathcal{M}_{m,n}(\mathbb{R})$, $C \in \mathcal{M}_{n,p}(\mathbb{R})$ *et* $\lambda \in \mathbb{R}$, *alors*
$$(A + \lambda B)C = AC + \lambda BC \in \mathcal{M}_{m,p}(\mathbb{R}).$$

Définition 4.2.6. Si $A = (a_{i,j})_{1\leq i\leq m, 1\leq j\leq n} \in \mathcal{M}_{m,n}(\mathbb{R})$, alors on définit la *transposée de* A et on note A^T la matrice
$$A^T = (a_{j,i})_{1\leq j\leq n, 1\leq i\leq m} \in \mathcal{M}_{n,m}(\mathbb{R}).$$
En d'autres termes, les lignes de A^T sont les colonnes de A et les colonnes de A^T sont les lignes de A.

Exemple 4.2.7. Si $A = \begin{pmatrix} 1 & 2 & 5 \\ 0 & 7 & -3 \end{pmatrix}$, alors $A^T = \begin{pmatrix} 1 & 0 \\ 2 & 7 \\ 5 & -3 \end{pmatrix}$.

Remarque 4.2.8. La transposition est une opération involutive: on a $(A^T)^T = A$ pour toute matrice A.

Proposition 4.2.9. *Si* $A \in \mathcal{M}_{m,n}(\mathbb{R})$ *et* $B \in \mathcal{M}_{n,p}(\mathbb{R})$ *alors*
$$(AB)^T = B^T A^T \in \mathcal{M}_{p,m}(\mathbb{R}).$$

Proof. Si $C = AB$, alors $c_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j}$ et ainsi
$$(C^T)_{i,j} = C_{j,i} = \sum_{k=1}^n a_{j,k} b_{k,i}.$$
Par ailleurs,
$$(B^T A^T)_{i,j} = \sum_{k=1}^n (B^T)_{i,k} (A^T)_{k,j} = \sum_{k=1}^n b_{k,i} a_{j,k} = \sum_{k=1}^n a_{j,k} b_{k,i}.$$
□

Exemple 4.2.10. Reprenant l'exemple ci-dessus avec $A = \begin{pmatrix} 1 & 2 & 5 \\ 0 & 7 & -3 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ 1 & 2 \\ -1 & 0 \end{pmatrix}$, on obtient
$$(AB)^T = \begin{pmatrix} -3 & 10 \\ 5 & 14 \end{pmatrix}$$
et
$$B^T A^T = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 7 \\ 5 & -3 \end{pmatrix} = \begin{pmatrix} -3 & 10 \\ 5 & 14 \end{pmatrix}.$$

4.3 L'anneau des matrices carrées

Définition 4.3.1. Soit $n \geq 1$ un entier. On note $\mathcal{M}_n(\mathbb{R})$ l'espace vectoriel $\mathcal{M}_{n,n}(\mathbb{R})$ des matrices carrées à n lignes et n colonnes(et coefficients réels).

On note en particulier que si $A \in \mathcal{M}_n(\mathbb{R})$, alors $A^T \in \mathcal{M}_n(\mathbb{R})$. Ceci conduit à la définition suivante:

Définition 4.3.2. On dit qu'une matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$ est

- *symétrique* si $A = A^T$.
- *antisymétrique* si $A = -A^T$.

Exemple 4.3.3. Si $A \in \mathcal{M}_n(\mathbb{R})$, alors AA^T est symétrique. En effet,

$$(AA^T)^T = (A^T)^T A^T = AA^T.$$

Si $A = \begin{pmatrix} 1 & 2 & 5 \\ 0 & 7 & -3 \end{pmatrix}$, alors $A^T = \begin{pmatrix} 1 & 0 \\ 2 & 7 \\ 5 & -3 \end{pmatrix}$ et

$$AA^T = \begin{pmatrix} 30 & -1 \\ -1 & 58 \end{pmatrix}, \quad A^T A = \begin{pmatrix} 1 & 2 & 5 \\ 2 & 53 & -11 \\ 5 & -11 & 34 \end{pmatrix}.$$

En plus de l'addition et de la loi externe, l'ensemble $\mathcal{M}_n(\mathbb{R})$ est muni d'une multiplication interne:

$$\mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R}), (A, B) \mapsto AB.$$

Ce produit admet pour élément neutre la matrice identité

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

En d'autres termes

$$(I_n)_{i,j} = \begin{cases} 0 & \text{si } i \neq j. \\ 1 & \text{si } i = j. \end{cases}.$$

On vérifie directement que $AI_n = I_n A = A$ pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$.

Proposition 4.3.4. *L'ensemble $\mathcal{M}_n(\mathbb{R})$, muni de l'addition et de la loi de multiplication interne ci-dessus est un anneau non commutatif.*

La démonstration découle de la discussion ci-dessus et de la proposition 4.2.5.

Remarque 4.3.5. Comme énoncé dans la proposition ci-dessus, la multiplication n'est pas commutative. Si $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, alors

$$AB = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \neq BA = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

De plus, l'anneau $\mathcal{M}_n(\mathbb{R})$ n'est pas intègre, dans le sens que le produit de deux éléments non nuls peut être nul. Par exemple,

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

Définition 4.3.6. On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *inversible* s'il existe une matrice $B \in \mathcal{M}_n(\mathbb{R})$ telle que $AB = BA = I_n$. Cette matrice B est alors unique et on la note A^{-1}. En d'autres termes, une matrice est inversible si et seulement si c'est un élément inversible de l'anneau $\mathcal{M}_n(\mathbb{R})$.

On vient de voir que l'anneau $\mathcal{M}_n(\mathbb{R})$ contient des matrices non nulles et non inversibles si $n \geq 1$. Par contre, on a la propriété suivante remarquable.

Proposition 4.3.7. *Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est inversible si et seulement s'il existe une matrice $B \in \mathcal{M}_n(\mathbb{R})$ telle que $AB = I_n$* ou *$BA = I_n$. Dans ce cas, on a $B = A^{-1}$.*

Proof. Si A est inversible, alors on prend $B = A^{-1}$ et la conclusion est évidente. Réciproquement, supposons qu'il existe $B \in \mathcal{M}_n(\mathbb{R})$ telle que $AB = I_n$. Considérons l'application linéaire

$$\varphi \colon \mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R})$$

définie par $\varphi(X) = XA$. On vérifie tout d'abord que φ est injective: si $X \in \mathcal{M}_n(\mathbb{R})$ satisfait $\varphi(X) = 0$, alors

$$0 = XA = (XA)B = X(AB) = XI_n = X.$$

Comme φ est un endomorphisme, on déduit du corollaire 3.3.6 que φ est un isomorphisme. En particulier, elle est surjective. Il existe donc $C \in \mathcal{M}_n(\mathbb{R})$ telle que

$$I_n = \varphi(C) = CA.$$

On obtient alors

$$C = CI_n = C(AB) = (CA)B = I_nB = B$$

et on en déduit bien que A est inversible avec $B = C = A^{-1}$.

Si on suppose maintenant qu'il existe $B \in \mathcal{M}_n(\mathbb{R})$ telle que $BA = I_n$ alors on considère l'application linéaire

$$\varphi \colon \mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R})$$

définie par $\varphi(X) = AX$ et on procède comme ci-dessus. □

Proposition 4.3.8. *Soient $A, B \in \mathcal{M}_n(\mathbb{R})$ deux matrices inversibles. Alors AB est aussi inversible, d'inverse $B^{-1}A^{-1}$.*

Proof. On a
$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$
et on conclut en utilisant la proposition ci-dessus (ou on calcule $(B^{-1}A^{-1})(AB)$ comme ci-dessus). $\square$

Définition 4.3.9. Soit $n \geq 1$ un entier. On note
$$\mathrm{GL}_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) |\ A \text{ inversible}\}.$$
Alors $\mathrm{GL}_n(\mathbb{R})$ est un groupe (dont l'opération est la multiplication de matrices) non abélien si $n \geq 2$.

Remarque 4.3.10. Il n'est en général pas facile de déterminer si une matrice A est inversible, et si c'est le cas de calculer son inverse. Nous verrons au chapitre 5 des méthodes pour faire cela. Si $n = 1$, on voit néanmoins facilement que $\mathrm{GL}_1(\mathbb{R}) = \mathbb{R} \setminus \{0\} = \mathbb{R}^\times$.

Exemple 4.3.11. La matrice $A = \begin{pmatrix} -1 & -2 \\ 3 & 5 \end{pmatrix}$ est inversible puisque
$$\begin{pmatrix} -1 & -2 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 5 & 2 \\ -3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$
De même, la matrice $B = \begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$ est inversible, comme le montre le calcul suivant:
$$\begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 1 & -2 & 1 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

4.4 Coordonnées d'un vecteur, matrice d'une application linéaire

Dans cette section, E est un espace vectoriel de dimension (finie) $m \geq 1$. On se donne une base $\{e_1, \ldots, e_m\}$ de E, et on rappelle que tout vecteur $v \in E$ admet une écriture unique
$$v = \sum_{i=1}^{m} v_i e_i, \text{ où } v_1, \ldots, v_m \in \mathbb{R}.$$
Les scalaires $v_1, \ldots, v_m$ sont les *coordonnées du vecteur v dans la base* $\{e_1, \ldots, e_n\}$.

Notation 4.4.1. On note
$$V = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} \in \mathcal{M}_{m,1}(\mathbb{R})$$

le vecteur colonne des coordonnées de v dans la base $\{e_1, \dots, e_m\}$. Si on souhaite mentionner la base $\mathbf{e} = \{e_1, \dots, e_m\}$ dans laquelle les coordonnées du vecteur sont calculées, on utilisera la notation

$$V = M_{\mathbf{e}}(v) = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}$$

Remarque 4.4.2. Il est important d'observer que la matrice V des coordonnées dépend de l'ordre dans lequel les éléments de la base sont pris. Par convention, on considère que $\{e_1, \dots, e_m\}$ est un ensemble ordonné.

Remarque 4.4.3. Quelle que soit la base $\mathbf{e}$ de E, l'application linéaire de la remarque 3.1.5

$$E \to \mathcal{M}_{m,1}(\mathbb{R}),\ v \mapsto M_{\mathbf{e}}(v) = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}$$

est un isomorphisme d'espaces vectoriels.

La notation ci-dessus s'étend à une famille de vecteurs.

Définition 4.4.4. Si $u, v, \dots, w$ une famille de n vecteurs de E, on note

$$M_{\mathbf{e}}(u, v, \dots, w) = \begin{pmatrix} u_1 & v_1 & \cdots & w_1 \\ u_2 & v_2 & \cdots & w_2 \\ \vdots & \vdots & \ddots & \vdots \\ u_n & v_n & \cdots & w_n \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{R})$$

la matrice dont les colonnes sont les coordonnées des vecteurs $u, v, \dots, w$ dans la base $\mathbf{e}$.

Supposons à présent que E et F sont des espaces vectoriels réels, munis respectivement d'une base $\mathbf{e} = \{e_1, \dots, e_n\}$ et d'une base $\mathbf{f} = \{f_1, \dots f_m\}$.

Définition 4.4.5. On appelle *matrice de l'application linéaire* $\varphi\colon E \to F$ *dans les bases* $\mathbf{e}, \mathbf{f}$ la matrice à m lignes et n colonnes donnée par

$$M_{\mathbf{f}}^{\mathbf{e}}(\varphi) = M_{\mathbf{f}}(\varphi(e_1), \dots, \varphi(e_n)).$$

En d'autres termes, les colonnes de la matrice $A := M_{\mathbf{e}}^{\mathbf{f}}(\varphi)$ contiennent les coordonnées dans la base $\mathbf{f}$ de l'image par φ des vecteurs de la base $\mathbf{e}$. Si $A = (a_{i,j})$, on a donc

$$\varphi(e_j) = \sum_{i=1}^{m} a_{i,j} f_i,\ \text{pour tout } j \in \{1, \dots, n\}. \tag{4.4.1}$$

Cette formule peut être utilisée comme la définition de la matrice $A = M_{\mathbf{f}}^{\mathbf{e}}(\varphi)$.

Exemple 4.4.6. Soit $\varphi\colon \mathbb{R}^3 \to \mathbb{R}^2$ l'application linéaire définie par $\varphi(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}) = \begin{pmatrix} x_1 + 2x_2 \\ x_1 - x_2 + 3x_3 \end{pmatrix}$. Si on munit $\mathbb{R}^3$ de la base canonique $\mathbf{e} = \{e_1, e_2, e_3\}$ et $\mathbb{R}^2$ de la base canonique $\mathbf{f} = \{f_1, f_2\}$, alors

$$M_{\mathbf{f}}^{\mathbf{e}}(\varphi) = \begin{pmatrix} 1 & 2 & 0 \\ 1 & -1 & 3 \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{R}).$$

Exemple 4.4.7. Soit $E = \mathbb{R}_3[X]$ l'espace vectoriel des polynômes de degré inférieur ou égal à 3, et $\varphi\colon E \to E$ l'endomorphisme défini par

$$\varphi(P) = P', \text{ pour tout } P \in E.$$

Si on munit E de la base $\{1, X, X^2, X^3\}$, on obtient une matrice de la forme

$$M_{\mathbf{e}}^{\mathbf{e}}(\varphi) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in \mathcal{M}_{4,4}(\mathbb{R}).$$

Remarque 4.4.8. Lorsque $\varphi\colon E \to E$ est. un endomorphisme, il est usuel de prendre une même base $\mathbf{e}$ pour l'espace de départ et d'arrivée. On parle alors de la matrice $M_{\mathbf{e}}^{\mathbf{e}}(\varphi)$ de l'endomorphisme φ dans la base $\mathbf{e}$.

Exemple 4.4.9. Si on considère l'application $\mathrm{Id}_E\colon E \to E$ (où $\dim(E) = n$), alors on vérifie immédiatement que

$$M_{\mathbf{e}}^{\mathbf{e}}(\mathrm{Id}_E) = I_n$$

et ceci quelle que soit la base $\mathbf{e}$ choisie.

Proposition 4.4.10. *Soient $\mathbf{e} = \{e_1, \ldots, e_n\}$ une base de E et $\mathbf{f} = \{f_1, \ldots, f_m\}$ une base de F. Alors, l'application*

$$\Psi\colon \mathcal{L}(E, F) \to \mathcal{M}_{m,n}(\mathbb{R})$$

donnée par $\Psi(\varphi) = M_{\mathbf{f}}^{\mathbf{e}}(\varphi)$ est un isomorphisme.

Proof. Supposons que $\varphi_1, \varphi_2 \in \mathcal{L}(E, F)$ et que $\lambda \in \mathbb{R}$. Si $A = M_{\mathbf{f}}^{\mathbf{e}}(\varphi_1)$ et $B = M_{\mathbf{f}}^{\mathbf{e}}(\varphi_2)$, l'équation (4.4.1) donne pour tout $j = 1, \ldots, n$

$$(\varphi_1 + \lambda\varphi_2)(e_j) = \varphi_1(e_j) + \lambda\varphi_2(e_j) = \sum_{i=1}^{m} a_{i,j} f_i + \lambda \sum_{i=1}^{m} b_{i,j} f_i = \sum_{i=1}^{m} (a_{i,j} + \lambda b_{i,j}) f_i,$$

d'où l'on tire que

$$M_{\mathbf{f}}^{\mathbf{e}}(\varphi_1 + \lambda\varphi_2) = M_{\mathbf{f}}^{\mathbf{e}}(\varphi_1) + \lambda M_{\mathbf{f}}^{\mathbf{e}}(\varphi_2).$$

Ainsi, Ψ est bien linéaire. On vérifie maintenant qu'elle est injective. Supposons que $\varphi \in \mathcal{L}(E, F)$ soit tel que $\Psi(\varphi) = 0$. Nous obtenons alors que $M_{\mathbf{f}}^{\mathbf{e}}(\varphi) = 0$ et donc, par (4.4.1) une fois encore, que

$$\varphi(e_j) = \sum_{i=1}^{m} 0 f_i = 0$$

pour tout $j = 1, \ldots, n$. Ceci montre bien que $\varphi = 0$, et donc que Ψ est injective. Finalement, on montre qu'elle est surjective. Pour ce faire, on prend une matrice $A = (a_{i,j}) \in \mathcal{M}_{m,n}(\mathbb{R})$, et on considère les vecteurs

$$v_j = \sum_{i=1}^{m} a_{i,j} f_i$$

pour tout $j = 1, \ldots, n$. Par la proposition 3.5.6, il existe une (unique) application linéaire $\varphi \colon E \to F$ telle que $\varphi(e_j) = v_j = \sum_{i=1}^{m} a_{i,j} f_i$ pour $j \in \{1, \ldots, n\}$. Il reste à vérifier que $\Psi(\varphi) = A$, ce qui est immédiat. □

Corollaire 4.4.11. *Si E et F sont de dimension finie, on a*

$$\dim(\mathcal{L}(E, F)) = \dim(E) \dim(F).$$

La proposition suivante montre que l'image d'un vecteur par une application linéaire peut être calculée à l'aide de matrices.

Proposition 4.4.12. *Soit $\varphi \colon E \to F$ une application linéaire, $\mathbf{e} = \{e_1, \ldots, e_n\}$ une base de E et $\mathbf{f} = \{f_1, \ldots, f_m\}$ une base de F. Si $x \in E$ et $\varphi(x) = y \in F$, on a*

$$Y = AX$$

où $Y = M_{\mathbf{f}}(y)$, $A = M_{\mathbf{f}}^{\mathbf{e}}(\varphi)$ et $X = M_{\mathbf{e}}(x)$.

Proof. Si $x = \sum_{j=1}^{n} x_j e_j$, on a $M_{\mathbf{e}}(x) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ et

$$\varphi(x) = \varphi(\sum_{j=1}^{n} x_j e_j) = \sum_{j=1}^{n} x_j \varphi(e_j) = \sum_{j=1}^{n} x_j (\sum_{i=1}^{m} a_{i,j} f_i) = \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{i,j} x_j) f_i$$

Ceci donne les coordonnées de $y = \varphi(x)$ dans la base $\mathbf{f}$, qui sont de la forme

$$M_{\mathbf{f}}(y) = \begin{pmatrix} \sum_{j=1}^{n} a_{1,j} x_j \\ \sum_{j=1}^{n} a_{2,j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{m,j} x_j \end{pmatrix} = AX.$$

□

On montre maintenant que la composition de deux applications linéaires est donnée par le produit des deux matrices correspondantes.

Proposition 4.4.13. *Soient E, F, G des espaces vectoriels réels de bases respectives $\mathbf{e} = \{e_1, \ldots, e_n\}$, $\mathbf{f} = \{f_1, \ldots, f_m\}$ et $\mathbf{g} = \{g_1, \ldots, g_p\}$. Si $\varphi \colon E \to F$ et $\psi \colon F \to G$ sont des applications linéaires, alors*

$$M_{\mathbf{g}}^{\mathbf{e}}(\psi \circ \varphi) = M_{\mathbf{g}}^{\mathbf{f}}(\psi) M_{\mathbf{f}}^{\mathbf{e}}(\varphi) \in \mathcal{M}_{p,n}(\mathbb{R})$$

Proof. On pose $A := M^{\mathbf{e}}_{\mathbf{f}}(\varphi)$, $B := M^{\mathbf{f}}_{\mathbf{g}}(\psi)$, $C := M^{\mathbf{e}}_{\mathbf{g}}(\psi \circ \varphi)$ et on calcule:

$$(\psi \circ \varphi)(e_j) = \psi(\sum_{i=1}^{m} a_{i,j} f_i) = \sum_{i=1}^{m} a_{i,j}\psi(f_i) = \sum_{i=1}^{m} a_{i,j}(\sum_{k=1}^{p} b_{k,i} g_k) = \sum_{k=1}^{p}(\sum_{i=1}^{m} b_{k,i}a_{i,j})g_k.$$

Or, $\sum_{i=1}^{m} b_{k,i}a_{i,j} = (BA)_{k,j}$ par définition et on en déduit que $C = BA$. □

Proposition 4.4.14. *Soient E, F des espaces vectoriels réels de bases respectives $\mathbf{e} = \{e_1, \ldots, e_n\}$ et $\mathbf{f} = \{f_1, \ldots, f_m\}$. Supposons que $\varphi\colon E \to F$ est une application linéaire. Alors φ est un isomorphisme si et seulement si $A := M^{\mathbf{e}}_{\mathbf{f}}(\varphi)$ est inversible. Dans ce cas, on a $M^{\mathbf{f}}_{\mathbf{e}}(\varphi^{-1}) = A^{-1}$.*

Proof. Supposons tout d'abord que φ est bijective. Alors $\varphi^{-1}\colon F \to E$ existe et on a $\varphi \circ \varphi^{-1} = \mathrm{Id}_F$, $\varphi^{-1} \circ \varphi = \mathrm{Id}_E$. Par la proposition 4.4.13 et l'exemple 4.4.9, on obtient

$$M^{\mathbf{e}}_{\mathbf{f}}(\varphi)M^{\mathbf{f}}_{\mathbf{e}}(\varphi^{-1}) = M^{\mathbf{f}}_{\mathbf{f}}(\mathrm{Id}_F) = I_n$$

ainsi que

$$M^{\mathbf{f}}_{\mathbf{e}}(\varphi^{-1})M^{\mathbf{e}}_{\mathbf{f}}(\varphi) = M^{\mathbf{e}}_{\mathbf{e}}(\mathrm{Id}_E) = I_n.$$

On en déduit que $A = M^{\mathbf{e}}_{\mathbf{f}}(\varphi)$ est inversible et que $M^{\mathbf{f}}_{\mathbf{e}}(\varphi^{-1}) = A^{-1}$. Réciproquement, supposons que A est inversible. Par la proposition 4.4.10, il existe une unique application linéaire $\psi\colon F \to E$ telle que $M^{\mathbf{f}}_{\mathbf{e}}(\psi) = A^{-1}$. On utilise encore la proposition 4.4.13 pour en conclure que

$$M^{\mathbf{f}}_{\mathbf{f}}(\mathrm{Id}_F) = I_n = AA^{-1} = M^{\mathbf{e}}_{\mathbf{f}}(\varphi)M^{\mathbf{f}}_{\mathbf{e}}(\psi) = M^{\mathbf{f}}_{\mathbf{f}}(\varphi \circ \psi)$$

et ainsi que $\varphi \circ \psi = \mathrm{Id}_F$. Le même argument montre que $\psi \circ \varphi = \mathrm{Id}_E$. On obtient finalement que φ est inversible, que $\psi = \varphi^{-1}$ et que $M^{\mathbf{f}}_{\mathbf{e}}(\varphi^{-1}) = A^{-1}$. □

4.5 Matrice de changement de base

Il arrive fréquemment qu'on considère deux bases différentes sur un même espace vectoriel E. Les matrices des coordonnées des vecteurs de E sont donc différentes dans ces deux bases, et il est indispensable de savoir comment transformer l'une en l'autre (le même problème s'applique aux applications linéaires).

Notation 4.5.1. Soit E un espace vectoriel, muni de deux bases $\mathbf{e} = \{e_1, \ldots, e_n\}$ et $\mathbf{f} = \{f_1, \ldots, f_n\}$. On note

$$P^{\mathbf{f}}_{\mathbf{e}} = M^{\mathbf{f}}_{\mathbf{e}}(\mathrm{Id}_E) = M_{\mathbf{e}}(f_1, \ldots, f_n).$$

En d'autres termes, les colonnes de $P^{\mathbf{f}}_{\mathbf{e}}$ sont les coordonnées des éléments de $\mathbf{f}$ dans la base $\mathbf{e}$. Il est clair que $P^{\mathbf{f}}_{\mathbf{e}}$ est inversible et que $(P^{\mathbf{f}}_{\mathbf{e}})^{-1} = P^{\mathbf{e}}_{\mathbf{f}}$. On appellera $P^{\mathbf{f}}_{\mathbf{e}}$ la *matrice de passage de la base* $\mathbf{f}$ *à la base* $\mathbf{e}$.

Proposition 4.5.2. *Pour tout vecteur $x \in E$, on a*

$$M_{\mathbf{e}}(x) = P^{\mathbf{f}}_{\mathbf{e}} M_{\mathbf{f}}(x).$$

Proof. C'est une application immédiate de la proposition 4.4.12, avec $E = F$ and $\varphi = \mathrm{Id}_E$. □

En ce qui concerne les applications linéaires, on a la proposition suivante.

Proposition 4.5.3. *Soient E et F deux espaces vectoriels réels. Soient $\mathbf{e} = \{e_1, \dots, e_n\}$ et $\mathbf{e}' = \{e'_1, \dots, e'_n\}$ deux bases de E, $\mathbf{f} = \{f_1, \dots, f_m\}$ et $\mathbf{f}' = \{f'_1, \dots, f'_m\}$ deux bases de F. Si $\varphi\colon E \to F$ est une application linéaire, on a*

$$M^{\mathbf{e}'}_{\mathbf{f}'}(\varphi) = P^{\mathbf{f}}_{\mathbf{f}'} M^{\mathbf{e}}_{\mathbf{f}}(\varphi) P^{\mathbf{e}'}_{\mathbf{e}}.$$

Proof. On considère les applications linéaires suivantes: $E \xrightarrow{\mathrm{Id}_E} E \xrightarrow{\varphi} F \xrightarrow{\mathrm{Id}_F} F$, où les espaces vectoriels sont munis respectivement des bases $\mathbf{e}', \mathbf{e}, \mathbf{f}, \mathbf{f}'$. La proposition 4.4.13 donne

$$M^{\mathbf{e}'}_{\mathbf{f}'}(\varphi) = M^{\mathbf{f}}_{\mathbf{f}'}(\mathrm{Id}_F) M^{\mathbf{e}}_{\mathbf{f}}(\varphi) M^{\mathbf{e}'}_{\mathbf{e}}(\mathrm{Id}_E) = P^{\mathbf{f}}_{\mathbf{f}'} M^{\mathbf{e}}_{\mathbf{f}}(\varphi) P^{\mathbf{e}'}_{\mathbf{e}}$$

□

Corollaire 4.5.4. *Soit E un espace vectoriel réel muni de deux bases $\mathbf{e} = \{e_1, \dots, e_n\}$ et $\mathbf{f} = \{f_1, \dots, f_n\}$. Si $\varphi\colon E \to E$ est un endomorphisme, on a*

$$M^{\mathbf{f}}_{\mathbf{f}}(\varphi) = P^{\mathbf{e}}_{\mathbf{f}} M^{\mathbf{e}}_{\mathbf{e}}(\varphi) P^{\mathbf{f}}_{\mathbf{e}} = (P^{\mathbf{f}}_{\mathbf{e}})^{-1} M^{\mathbf{e}}_{\mathbf{e}}(\varphi) P^{\mathbf{f}}_{\mathbf{e}}.$$

Exemple 4.5.5. Soit $E = \mathbb{R}^3$, $\mathbf{e} = \{e_1, e_2, e_3\}$ la base canonique et $\mathbf{f} = \{f_1, f_2, f_3\}$ avec

$$f_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \ f_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \ f_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

On observe que $e_1 = f_3 - f_2$, $e_2 = f_1 + f_2 - f_3$ et $e_3 = f_3 - f_1$, ce qui donne une matrice de passage

$$P^{\mathbf{e}}_{\mathbf{f}} = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

Comme $f_1 = e_1 + e_2$, $f_2 = e_2 + e_3$ et $f_3 = e_1 + e_2 + e_3$, on trouve

$$P^{\mathbf{f}}_{\mathbf{e}} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

On vérifie que $P^{\mathbf{f}}_{\mathbf{e}} = (P^{\mathbf{e}}_{\mathbf{f}})^{-1}$. Si maintenant $x = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \in E$, on a

$$M_{\mathbf{e}}(x) = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \ M_{\mathbf{f}}(x) = P^{\mathbf{e}}_{\mathbf{f}} \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}.$$

Si maintenant $\varphi\colon E \to E$ est l'endomorphisme défini par $\varphi(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}) = \begin{pmatrix} 3x_1 - 2x_2 + 2x_3 \\ x_1 + 2x_3 \\ x_1 - x_2 + 3x_3 \end{pmatrix}$, on obtient

$$M^{\mathbf{e}}_{\mathbf{e}}(\varphi) = \begin{pmatrix} 3 & -2 & 2 \\ 1 & 0 & 2 \\ 1 & -1 & 3 \end{pmatrix},$$

alors que

$$M_{\mathbf{f}}^{\mathbf{f}}(\varphi) = P_{\mathbf{f}}^{\mathbf{e}} M_{\mathbf{e}}^{\mathbf{e}}(\varphi) P_{\mathbf{e}}^{\mathbf{f}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

On peut également vérifier par calcul direct que $\varphi(f_1) = f_1, \varphi(f_2) = 2f_2, \varphi(f_3) = 3f_3$.

Définition 4.5.6. On dit que deux matrices $A, B \in \mathcal{M}_n(\mathbb{R})$ sont *semblables* s'il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbb{R})$ telle que

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

En vertu du corollaire 4.5.4, deux matrices sont semblables si et seulement si elles représentent le même endomorphisme dans deux bases différentes (en effet, toute matrice inversible est une matrice de passage).

Remarque 4.5.7. La relation *être semblables* est une relation d'équivalence. En effet:

- (Réflexivité) Toute matrice est semblables à elle-même.
- (Symétrie) Si A est semblables à B, alors B est semblable à A.
- (Transitivité) Si A est semblable à B, et B est semblable à C, alors A est semblable à C.

Définition 4.5.8. On dit que deux matrices $A, B \in \mathcal{M}_{m,n}(\mathbb{R})$ sont *équivalentes* s'il existe des matrices inversibles $P \in \mathrm{GL}_n(\mathbb{R})$ et $Q \in \mathrm{GL}_m(\mathbb{R})$ telles que

$$B = Q^{-1}AP$$

Il suit de la proposition 4.5.3 que deux matrices sont équivalentes si et seulement si elles représentent la même application linéaire à un changement de base près dans l'espace de départ et dans l'espace d'arrivée.

Remarque 4.5.9. Si $A, B \in \mathcal{M}_n(\mathbb{R})$, alors

$$A,\ B \text{ semblables } \implies A,\ B \text{ équivalentes.}$$

L'implication réciproque est fausse.

4.6 Rang d'une matrice

Définition 4.6.1. Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$. On appelle *rang de la matrice* A la dimension du sous-espace vectoriel de $\mathbb{R}^m$ engendré par les colonnes de A. On note cet entier $\mathrm{rang}(A)$, et on observe que $\mathrm{rang}(A) \leq \min(m, n)$.

Exemple 4.6.2. Si $A = \begin{pmatrix} 1 & 2 & 5 \\ 0 & 7 & -3 \end{pmatrix}$, on a $\mathrm{rang}(A) = 2$.

Proposition 4.6.3. *Supposons que* $A = M_{\mathbf{f}}^{\mathbf{e}}(\varphi)$ *pour une application linéaire* $\varphi\colon E \to F$*,* $\mathbf{e}$ *une base de* E *et* $\mathbf{f}$ *une base de* F*. Alors*

$$\mathrm{rang}(A) = \mathrm{rang}(\varphi).$$

Proof. On a

$$\operatorname{rang}(\varphi) = \dim(\operatorname{Im}(\varphi)) = \dim(\operatorname{Vect}(\varphi(e_1), \dots, \varphi(e_n))) = \dim \operatorname{Vect}(M_{\mathbf{f}}(\varphi(e_1)), \dots, M_{\mathbf{f}}(\varphi(e_n))))$$

et le dernier terme est égal à $\operatorname{rang}(A)$. $\square$

En particulier, on a $\operatorname{rang}(A) = \operatorname{rang}(\varphi_A)$ où $\varphi_A \colon \mathbb{R}^n \to \mathbb{R}^m$ est l'application linéaire canoniquement associée à A (i.e. définie par $X \mapsto AX$).

Proposition 4.6.4. *Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est inversible si et seulement si* $\operatorname{rang}(A) = n$.

Proof. On sait que A est inversible ssi $\varphi_A \colon \mathbb{R}^n \to \mathbb{R}^n$ est bijective ssi $\varphi_A \colon \mathbb{R}^n \to \mathbb{R}^n$ est surjective ssi $\operatorname{rang}(\varphi_A) = n$ ssi $\operatorname{rang}(A) = n$. $\square$

On a vu que deux matrices équivalentes représentent la même application linéaire, à des choix de bases près; elles ont donc le même rang. La réciproque est également vraie:

Proposition 4.6.5. *Deux matrices $A, B \in \mathcal{M}_{m,n}(\mathbb{R})$ sont équivalentes si et seulement si elles ont le même rang.*

Proof. Soit r le rang de la matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$. On rappelle que $r \le \min(m, n)$. Considérons l'application linéaire $\varphi_A \colon \mathbb{R}^n \to \mathbb{R}^m$ canoniquement associée à A. Par le théorème du rang, on obtient

$$n = \dim(\ker(\varphi_A)) + \dim(\operatorname{Im}(\varphi_A)) = \dim(\ker(\varphi_A)) + \operatorname{rang}(A) = \dim(\ker(\varphi_A)) + r.$$

Ainsi, $\dim(\ker(\varphi_A)) = n - r$, et on choisit une base $\{e_{r+1}, \dots, e_n\}$ de $\ker(\varphi_A)$ qu'on complète en une base $\mathbf{e} = \{e_1, \dots, e_r, e_{r+1}, \dots, e_n\}$ de E. On a vu dans la preuve du théorème du rang que si on pose $f_j = \varphi_A(e_j)$ pour tout $j = 1, \dots, r$ alors $\{f_1, \dots, f_r\}$ est une base de $\operatorname{Im}(\varphi_A)$. On complète cette famille en une base $\mathbf{f} = \{f_1, \dots, f_r, f_{r+1}, \dots, f_m\}$ de $\mathbb{R}^m$. Par construction, la matrice de φ_A dans les bases $\mathbf{e}$ et $\mathbf{f}$ est de la forme

$$M_{\mathbf{f}}^{\mathbf{e}}(\varphi_A) := J_r = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

On a ainsi que A est équivalente à J_r par la proposition 4.5.3: il existe $Q \in \mathrm{GL}_m(\mathbb{R})$ et $P \in \mathrm{GL}_n(\mathbb{R})$ telles que $A = Q^{-1} J_r P$. Il en est de même pour B puisque $r = \operatorname{rang}(A) = \operatorname{rang}(B)$, donnant l'existence de $Q' \in \mathrm{GL}_m(\mathbb{R})$ et $P' \in \mathrm{GL}_n(\mathbb{R})$ telles que $B = (Q')^{-1} J_r P'$. Ceci donne finalement

$$A = Q^{-1} Q' B (P')^{-1} P = ((Q')^{-1} Q)^{-1} B (P^{-1} P')^{-1}$$

qui montre que A et B sont équivalentes. $\square$

Corollaire 4.6.6. *Une matrice et sa transposée ont le même rang, i.e. si $A \in \mathcal{M}_{m,n}(\mathbb{R})$ alors* $\operatorname{rang}(A) = \operatorname{rang}(A^T)$.

Proof. Pour $A \in \mathcal{M}_{m,n}(\mathbb{R})$, on a vu que $A = Q^{-1}J_rP$ où $r = \operatorname{rang}(A)$. Ceci donne

$$A^T = (Q^{-1}J_rP)^T = P^TJ_r^T(Q^{-1})^T$$

avec $P^T \in \mathrm{GL}_n(\mathbb{R})$ et $(Q^{-1})^T \in \mathrm{GL}_m(\mathbb{R})$, montrant que A^T est équivalente à J_r^T. On conclut en observant directement que J_r^T est de rang r. $\square$

5 Matrices et systèmes linéaires

Cette section est destinée à expliquer comment calculer efficacement le rang d'une matrice donnée, et de résoudre des systèmes d'équations linéaires de grande dimension.

5.1 Opérations élémentaires sur les matrices

Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$ une matrice. On note $L_1, \dots, L_m \in \mathcal{M}_{1,n}(\mathbb{R})$ les lignes de A et $C_1, \dots, C_n \in \mathcal{M}_{n,1}(\mathbb{R})$ les colonnes de A:

$$A = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{pmatrix} = \begin{pmatrix} C_1 \mid & C_2 \mid & \dots \mid & C_n \end{pmatrix}$$

Considérons les *opérations élémentaires* suivantes:

- Opérations sur les lignes:
  a) $L_i \mapsto \lambda L_i$, multiplication de la ligne L_i par $\lambda \in \mathbb{R}$ *non nul*.
  b) $L_i \mapsto L_j$, $L_j \mapsto L_i$, échange des lignes i et j ($i \neq j$).
  c) $L_i \mapsto L_i + \lambda L_j$, ajout de λL_j à la ligne L_i ($i \neq j$).
- Opérations sur les colonnes:
  a) $C_i \mapsto \lambda C_i$, multiplication de la colonne C_i par $\lambda \in \mathbb{R}$ *non nul*.
  b) $C_i \mapsto C_j$, $C_j \mapsto C_i$, échange des colonnes i et j ($i \neq j$).
  c) $C_i \mapsto C_i + \lambda C_j$, ajout de λC_j à la colonne C_i ($i \neq j$).

Lemme 5.1.1. *Les opérations élémentaires ci-dessus transforment $A \in \mathcal{M}_{m,n}(\mathbb{R})$ en une matrice équivalente.*

Proof. On montre le résultat pour les opérations élémentaires sur les colonnes, la preuve pour les opérations sur les lignes étant quasi identique. Rappelons que pour tout $i = 1, \dots, n$ et $j = 1, \dots, n$ la matrice $E_{i,j} \in \mathcal{M}_n(\mathbb{R})$ est la matrice définie par

$$(E_{i,j})_{k,l} = \begin{cases} 1 & \text{si } (i,j) = (k,l). \\ 0 & \text{si } (i,j) \neq (k,l). \end{cases}$$

Un calcul direct montre que

$$AE_{i,j} = \begin{pmatrix} 0 \mid & 0 \mid & \dots & \mid & C_i \mid & 0 \mid & \dots \mid & 0 \end{pmatrix}$$

où C_i est devenu la j-ième colonne de $AE_{i,j}$. On donne maintenant les formules des opérations élémentaires à l'aide de ces matrices:

a) $C_i \mapsto \lambda C_i$, multiplication de la colonne C_i par $\lambda \in \mathbb{R}$ *non nul*. Cette opération est obtenue via $A \mapsto AP$ avec $P = \lambda E_{1,1} + E_{2,2} + \ldots + E_{n,n}$.

b) $C_1 \mapsto C_2$, $C_2 \mapsto C_1$, échange des colonnes 1 et 2. Cette opération est obtenue via $A \mapsto AP$ avec $P = E_{1,2} + E_{2,1} + E_{3,3} + \ldots + E_{n,n}$.

c) $C_1 \mapsto C_1 + \lambda C_2$, ajout de λC_2 à la colonne C_1. Cette opération est obtenue via $A \mapsto AP$ avec $P = E_{1,1} + \lambda E_{2,1} + E_{2,2} + \ldots + E_{n,n}$.

On vérifie que toutes matrices sont bien inversibles:

a) $P^{-1} = \lambda^{-1} E_{1,1} + E_{2,2} + \ldots + E_{n,n}$.

b) $P^{-1} = P = E_{1,2} + E_{2,1} + E_{3,3} + \ldots + E_{n,n}$.

c) $P^{-1} = E_{1,1} - \lambda E_{2,1} + E_{2,2} + \ldots + E_{n,n}$.

□

Remarque 5.1.2. On peut aussi vérifier directement que toutes les opérations élémentaires sur les colonnes laissent inchangé le sous-espace vectoriel

$$\mathrm{Vect}(C_1, \ldots, C_n) \subset \mathbb{R}^m$$

dont la dimension est précisément le rang de A. Les opérations élémentaires sur les colonnes transforment donc A en une matrice de même rang, i.e. équivalente à A.

Montrons maintenant à l'aide d'exemples comment transformer, par une suite d'opérations élémentaires sur les lignes, une matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ en une matrice plus simple, dite *échelonnée* pour lequel le rang est facile à calculer.

Exemple 5.1.3. Soit

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

On observe que $a_{1,1} = 1 \neq 0$, et on peut donc éliminer les facteurs $a_{i,1}$ pour $i = 2, 3, 4$ en enlevant un multiple de L_1 aux autres lignes. Ici, on effectue les opérations $L_2 \mapsto L_2 - 2L_1$, $L_3 \mapsto L_3 - L_1$ et $L_4 \mapsto L_4 + L_1$ pour obtenir une matrice

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

À ce point, on voit que $a_{2,2} = -1 \neq 0$, et on peut se servir de ce fait pour maintenant éliminer $a_{3,2}$ et $a_{4,2}$ en utilisant les opérations $L_3 \mapsto L_3 + L_2$ et $L_4 = L_4 + L_2$. Ceci donne

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

avec $a_{3,3} = a_{3,4} = 0$. On ne peut pas se ramener à $a_{3,3} = 0$ en permutant les lignes, et on laisse donc la colonne 3 intacte. On utilise enfin le fait que $a_{3,4} \neq 0$ pour éliminer $a_{4,4}$, utilisant cette fois $L_4 \mapsto L_4 - 2L_3$ pour finalement obtenir

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

qui est de rang 3.

Exemple 5.1.4. On considère maintenant

$$A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 2 & 2 & 4 & 1 \\ 3 & 4 & 7 & 0 \end{pmatrix}.$$

Comme dans l'exemple précédent, on a $a_{1,1} \neq 0$ et on effectue les opérations $L_2 \mapsto L_2 - 2L - 1$, $L_3 \mapsto L_3 - 3L_1$ pour avoir

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

Comme $a_{2,2} = 0$, on permute les deux dernières lignes donnant ainsi une matrice

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

qui est de rang 3.

Définition 5.1.5. On dit qu'une matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ est *échelonnée* s'il existe des entiers $r \in \{0, \ldots, \min(m,n)\}$ et $1 \leq j_1 < j_2 < \ldots < j_r \leq n$ tels que

- Si $i \in \{1, \ldots, r\}$: $a_{i,j_i} \neq 0$ et $a_{i,j} = 0$ si $j < j_i$.
- Si $i \in \{r+1, \ldots, m\}$: $a_{i,j} = 0$ pour tout $j \in \{1, \ldots, n\}$.

En d'autres termes, une matrice échelonnée est de la forme

$$A = \begin{pmatrix} a_{1,j_1} & \dots & a_{1,j_2} & \dots & a_{1,j_3} & \dots & a_{1,j_r} & \dots & a_{1,n} \\ 0 & \dots & a_{2,j_2} & \dots & a_{2,j_3} & \dots & a_{2,j_r} & \dots & a_{2,n} \\ 0 & \dots & 0 & \dots & a_{3,j_3} & \dots & a_{3,j_r} & \dots & a_{3,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & a_{r,j_r} & \dots & a_{r,n} \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{pmatrix} \tag{5.1.1}$$

Proposition 5.1.6. *Si $A \in \mathcal{M}_{m,n}(\mathbb{R})$ est échelonnée au sens de la définition ci-dessus, alors on a $\operatorname{rang}(A) = r$.*

Proof. On sait d'après le corollaire 4.6.6 que $\operatorname{rang}(A) = \operatorname{rang}(A^T)$. Transposant la matrice (5.1.1), on obtient une matrice avec r colonnes non nulles, ce qui montre que le rang de A est au plus r. Il s'agit ensuite de démontrer que ces colonnes sont linéairement indépendantes, ce qui est facile. □

Proposition 5.1.7. *Toute matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ peut être mise sous forme échelonnée par une suite d'opérations élémentaires sur les lignes.*

Proof. Si $A = 0$, A est déjà échelonnée (avec $r = 0$). On suppose donc que la matrice est non nulle. Dans ce cas, on pose $j_1 = \min(j \in \{1, \dots, n\} | C_j \neq 0\}$ (C_{j_1} est donc la première colonne non nulle). Quitte à permuter les lignes, on peut supposer que $a_{1,j_1} \neq 0$. Ensuite, on effectue les opérations élémentaires

$$L_i \mapsto L_i - \frac{a_{i,j_1}}{a_{1,j_1}} L_1$$

pour $i = 2, \dots, m$, obtenant ainsi une matrice où $a_{i,j_1} = 0$ pour tout $i = 2, \dots, m$. Elle est donc de la forme

$$\begin{pmatrix} 0 & \mathbf{a} \\ 0 & A' \end{pmatrix}$$

avec $\mathbf{a} = \begin{pmatrix} a_{1,j_1} & \dots & a_{n,j_1} \end{pmatrix}$ et $A' = (a'_{i,j})_{2 \le i \le m, j_1+1 \le j \le n}$. On recommence le procédé, et on voit qu'on obtient une matrice échelonnée en un nombre fini d'étapes. □

5.2 Généralités sur les systèmes linéaires

Définition 5.2.1. Soit $m, n \geq 1$ des entiers. On appelle *système linéaire à m équations et n inconnues* une équation de la forme

$$AX = b \tag{5.2.1}$$

où la matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et le vecteur $b \in \mathbb{R}^m$ sont donnés, et l'inconnue est le vecteur $X \in \mathbb{R}^n$.

Remarque 5.2.2. Notant $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ et $A = \begin{pmatrix} C_1 | & \ldots | C_n \end{pmatrix}$ (où les C_j sont les colonnes de A), le système (5.2.1) s'écrit en fait

$$x_1 C_1 + x_2 C_2 + \ldots + x_n C_n = b. \tag{5.2.2}$$

où les lignes de ce système sont les m équations considérées.

Exemple 5.2.3. On cherche à résoudre le sytème

$$\begin{cases} x + y + z = 6 \\ 3x - 3y + z = 0 \\ x - y + z = 2 \end{cases}$$

correspondant à $AX = b$ avec $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $b = \begin{pmatrix} 6 \\ 0 \\ 2 \end{pmatrix}$ et $A = \begin{pmatrix} 1 & 1 & 1 \\ 3 & -3 & 1 \\ 1 & -1 & 1 \end{pmatrix}$.

Dans ce qui suit, on va toujours identifier une matrice $A \in \mathrm{M}_{m,n}(\mathbb{R})$ avec l'application linéaire $X^n \to \mathbb{R}^m$ donnée par $X \mapsto AX$.

Proposition 5.2.4. *Les assertions suivantes sont vérifiées:*

a) Le système (5.2.1) *possède une solution si et seulement si* $b \in \mathrm{Im}(A) = \mathrm{Vect}(C_1, \ldots, C_n)$. *En particulier,* (5.2.1) *possède une solution pour tout* $b \in \mathbb{R}^m$ *si et seulement si* $\mathrm{rang}(A) = m$.

b) Si elle existe, la solution de (5.2.1) *est unique si et seulement si* $\ker(A) = \{0\}$.

Proof. Pour le premier point: d'après l'équation (5.2.2), on sait que le système d'équation a une solution si set seulement si b est combinaison linéaire des colonnes de A, ce qui est le cas si et seulement si $b \in \mathrm{Im}(A)$. Pour la seconde assertion, supposons que X et X' soient deux solutions du système $AX = b$. On obtient alors $A(X - X') = b - b = 0$, i.e. $X - X' \in \ker(A)$. La conclusion s'ensuit. □

Proposition 5.2.5 (Système équilibré ($m = n$))**.** *Si* $m = n$*, le système* (5.2.1) *possède une unique solution quel que soit* $b \in \mathbb{R}^m$ *si et seulement si* $A \in \mathrm{GL}_n(\mathbb{R})$.

Proof. D'après la proposition ci-dessus, le système possède une unique solution si et seulement si $\mathrm{rang}(A) = n$ et $\ker(A) = \{0\}$. □

Remarque 5.2.6. Utilisant le théorème du rang, on voit que si $m = n$ le système (5.2.1) possède une solution quel que soit $b \in \mathbb{R}^m$ si et seulement si $A \in \mathrm{GL}_n(\mathbb{R})$. De plus les solutions sont uniques.

Remarque 5.2.7. Si $A \in \mathrm{GL}_n(\mathbb{R})$, alors l'unique solution du système $AX = b$ est $X = A^{-1}b$.

5.3 Résolution effective des systèmes linéaires

La méthode générale pour résoudre le système (5.2.1) repose sur le principe suivant:

Proposition 5.3.1. *Les opérations élémentaires sur les lignes, effectuées simultanément sur la matrice A et le second terme b, ne changent pas les solutions du système (5.2.1).*

Proof. On a vu que faire des opérations élémentaires sur les lignes de A correspond à multiplier A à gauche par une matrice inversible $Q \in \mathrm{GL}_m(\mathbb{R})$. Si $A' = QA$ et $b' = Qb$, on a alors

$$AX = b \Leftrightarrow QAX = Qb \Leftrightarrow A'X = b'.$$

□

L'idée est de mettre à présent la matrice A sous forme échelonnée par une suite d'opérations élémentaires sur les lignes.

Exemple 5.3.2. Soit le système

$$\begin{cases} x+y+z=6 \\ 3x-3y+z=0 \\ x-y+z=2 \end{cases}$$

Effectuant des opérations élémentaires sur les lignes, on obtient successivement

$$\begin{cases} x+y+z=6 \\ -6y-2z=-18 \\ -2y=-4, \end{cases} \qquad L_2 \mapsto L_2-3L_1, L_3 \mapsto L_3-L_1$$

$$\begin{cases} x+y+z=6 \\ 3y+z=9 \\ y=2 \end{cases} \qquad L_2 \mapsto -\frac{1}{2}L_2, L_3 \mapsto -\frac{1}{2}L_2$$

$$\begin{cases} x+y+z=6 \\ y=2 \\ 3y+z=9 \end{cases} \qquad L_2 \mapsto L_3, L_3 \mapsto L_2$$

et enfin

$$\begin{cases} x+y+z=6 \\ y=2 \\ z=3 \end{cases} \qquad L_3 \mapsto L_3-3L_2$$

d'où l'on tire que $x=1$. Finalement, nous obtenons la solution unique $X = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

Exemple 5.3.3. On considère maintenant le système

$$\begin{cases} x+y+z=3 \\ 2x-3y+7z=1 \\ x-9y+11z=-7 \end{cases}$$

On trouve dans ce cas:

$$\begin{cases} x+y+z=3 \\ -5y+5z=-5 \\ -10y+10z=-10 \end{cases} \qquad L_2 \mapsto L_2-2L_1, L_3 \mapsto L_3-L_1$$

$$\begin{cases} x+y+z=3 \\ y-z=1 \\ y-z=1 \end{cases} \qquad L_2 \mapsto -\frac{1}{5}L_2, L_3 \mapsto -\frac{1}{10}L_3$$

$$\begin{cases} x+y+z=3 \\ y-z=1 \\ 0=0 \end{cases} \qquad L_3 \mapsto L_3-L_2$$

La matrice est maintenant échelonnée, de rang 2. Passant z au membre de droite, ce système devient

$$\begin{cases} x+y=3-z \\ y=1+z \end{cases}$$

ou encore

$$\begin{cases} x=2-2z \\ y=1+z \end{cases} \qquad L_1 \mapsto L_1-L_2$$

L'ensemble des solutions est donc de la forme

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2-2z \\ 1+z \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}.$$

C'est une droite affine. On peut noter que le vecteur $\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \in \ker(A)$.

Remarque 5.3.4. Le même système avec un second terme différent

$$\begin{cases} x+y+z=3 \\ 2x-3y+7z=1 \\ x-9y+11z=13 \end{cases}$$

ne possède pas de solutions. Il suffit de faire les mêmes opérations que ci-dessus pour s'en apercevoir. Ceci vient du fait que $\begin{pmatrix} 3 \\ 1 \\ 13 \end{pmatrix} \notin \mathrm{Im}(A)$.

On retourne maintenant au cas général, i.e. celui d'un système $AX = b$ avec $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $b \in \mathbb{R}^m$. On effectue une suite d'opérations élémentaires sur les lignes pour en arriver à un système de la forme:

$$\begin{aligned}
a_{1,j_1}x_{j_1} + \ldots + a_{1,j_2}x_{j_2} + \ldots + a_{1,j_r}x_{j_r} + \ldots + a_{1,n}x_n &= b_1 \\
a_{2,j_2}x_{j_2} + \ldots + a_{2,j_r}x_{j_r} + \ldots + a_{2,n}x_n &= b_2 \\
a_{r,j_r}x_{j_r} + \ldots + a_{r,n}x_n &= b_r \\
0 &= b_{r+1} \\
\vdots \quad &\quad \vdots \\
0 &= b_m
\end{aligned}$$

Premier cas: Supposons que $r < m$. Dans ce cas, le système ci-dessus n'a aucune solution si $(b_{r+1}, \ldots, b_m) \neq (0, \ldots, 0)$. Si au contraire on a $(b_{r+1}, \ldots, b_m) = (0, \ldots, 0)$, alors les $m - r$ dernières lignes sont tautologiques et peuvent être omises. Dans la discussion ci-dessous, *on pourra donc supposer que* $m = r$, i.e. que $\mathrm{Im}(A) = \mathbb{R}^m$.

Deuxième cas: On a $n = m(= r)$. Dans ce cas, on a $(j_1, \ldots, j_r) = (1, \ldots, n)$ et le système s'écrit

$$\begin{aligned}
a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,n}x_n &= b_1 \\
a_{2,2}x_2 + \ldots + a_{2,n}x_n &= b_2 \\
\ddots \qquad \vdots \quad &\quad \vdots \\
a_{n,n}x_n &= b_n
\end{aligned}$$

avec $a_{i,i} \neq 0$ pour tout $i = 1, \ldots, n$. Le système possède donc une unique solution, que l'on calcule facilement en trouvant tout d'abord x_n, puis x_{n-1}, ..., et enfin x_1.

Troisième cas: Supposons que $n > m(= r)$. Dans ce cas, on distingue:

- *les inconnues principales* $x_{j_1}, \ldots, x_{j_r}$.
- *les inconnues libres* $x_{k_{r+1}}, \ldots, x_{k_n}$ où $\{k_{r+1}, \ldots, k_n\} = \{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\}$.

Les inconnues libres peuvent prendre des valeurs quelconques, et une fois ces valeurs fixées les inconnues principales sont déterminées par le système. Passant les inconnues libres à droite dans le système, on obtient:

$$\begin{aligned}
a_{1,j_1}x_{j_1} + a_{1,j_2}x_{j_2} + a_{1,j_3}x_{j_3} + \ldots + a_{1,j_r}x_{j_r} &= b_1 - \sum_{i=r+1}^{n} a_{1,k_i}x_{k_i} \\
a_{2,j_2}x_{j_2} + a_{2,j_3}x_{j_3} + \ldots + a_{2,j_r}x_{j_r} &= b_2 - \sum_{i=r+1}^{n} a_{2,k_i}x_{k_i} \\
\ddots \qquad \vdots \quad &\quad \vdots \\
a_{r,j_r}x_{j_r} &= b_r - \sum_{i=r+1}^{n} a_{r,k_i}x_{k_i}
\end{aligned}$$

Ce système permet de déterminer les variables principales en fonction des inconnues libres $x_{k_{r+1}}, \ldots, x_{k_n}$. On obtient ainsi toutes les solutions du système qui forment un sous-espace de $\mathbb{R}^n$ de la forme

$$\mathcal{S} := \{X \in \mathbb{R}^n | \ X - C \in \ker(A)\}$$

où C est la solution du système ci-dessus obtenu en posant $x_{k_{r+1}} = \ldots = x_{k_n} = 0$.

5.4 Instabilités numériques, pivot de Gauss

Rappelons que la première étape de l'échelonnement d'une matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ est d'identifier la première colonne non nulle C_{j_1}, puis de choisir un terme non nul dans cette colonne qui donne a_{1,j_1} quitte à permuter les lignes. La seconde étape est de faire des opérations élémentaires de la forme

$$L_i \mapsto L_i - \frac{a_{i,j_1}}{a_{1,j_1}} L_1$$

pour $i = 2, \ldots, m$, qui sont possibles puisque $a_{1,j_1} \neq 0$. Dans la pratique, on fait des calculs avec une précision finie, et on voit que si par exemple $a_{1,j_1} = 10^{-6}$ les opérations ci-dessus peuvent amplifier les erreurs d'arrondi, et finalement mettre en danger la précision du résultat.

Pour minimiser ce risque, on a intérêt à ce que le pivot a_{1,j_1} soit de valeur absolue aussi grande que possible. Ainsi, quitte à permuter les lignes, on peut demander que

$$|a_{1,j_1}| \geq \max(|a_{2,j_1}|, \ldots, |a_{m,j_1}|) > 0$$

On appelle *pivot de Gauss* un pivot possédant cette propriété. La *méthode du pivot de Gauss* consiste à résoudre des systèmes linéaires par la méthode décrite dans la section 5.3, en choisissant à chaque étape le pivot de Gauss pour minimiser les amplifications des erreurs d'arrondi commises lors des calculs des diverses opérations élémentaires. Cette méthode est largement utilisée en analyse numérique matricielle.

5.5 Calcul de l'inverse d'une matrice

Si $A \in \mathcal{M}_n(\mathbb{R})$ est inversible (i.e. appartient à $\mathrm{GL}_n(\mathbb{R})$), on écrit

$$A^{-1} = \begin{pmatrix} C_1 \mid & C_2 \mid & \ldots \mid & C_n \end{pmatrix}$$

où les C_i sont les colonnes de la matrice A^{-1} que l'on cherche à déterminer. Utilisant l'équation

$$AA^{-1} = \mathrm{I}_n = \begin{pmatrix} e_1 \mid & e_2 \mid & \ldots \mid & e_n \end{pmatrix}$$

où $\{e_1, \ldots, e_n\}$ est la base canonique de $\mathbb{R}^n$. On obtient donc n systèmes d'équations de la forme

$$AC_i = e_i, \ i = 1, \ldots, n.$$

Le calcul de l'inverse de A revient donc à résoudre n systèmes de taille n. Cette opération est coûteuse pour n grand, même si l'échelonnement de la matrice A peut se faire une seule fois.

En pratique, pour n petit, on peut résoudre génériquement le système $AX = b$ pour $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ générique, puis en spécialisant en $b = e_1$, $b = e_2, \ldots, b = e_n$.

Exemple 5.5.1. Supposons que $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$. On obtient donc un système de la forme

$$\begin{cases} x + y = b_1 \\ 2x + y = b_2. \end{cases}$$

Echelonnant, on obtient

$$\begin{cases} x + y = b_1 \\ -y = b_2 - 2b_1 \end{cases} \qquad L_2 \mapsto L_2 - 2L_1$$

$$\begin{cases} x + y = b_1 \\ y = 2b_1 - b_2 \end{cases} \qquad L_2 \mapsto -L_2$$

$$\begin{cases} x = b_2 - b_1 \\ y = 2b_1 - b_2 \end{cases} \qquad L_1 \mapsto L_1 - L_2$$

Ceci donne donc génériquement $x = b_2 - b_1$ et $y = 2b_1 - b_2$. Pour $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, on a l'unique solution $C_1 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, alors que pour $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ on obtient $C_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Ceci donne finalement

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