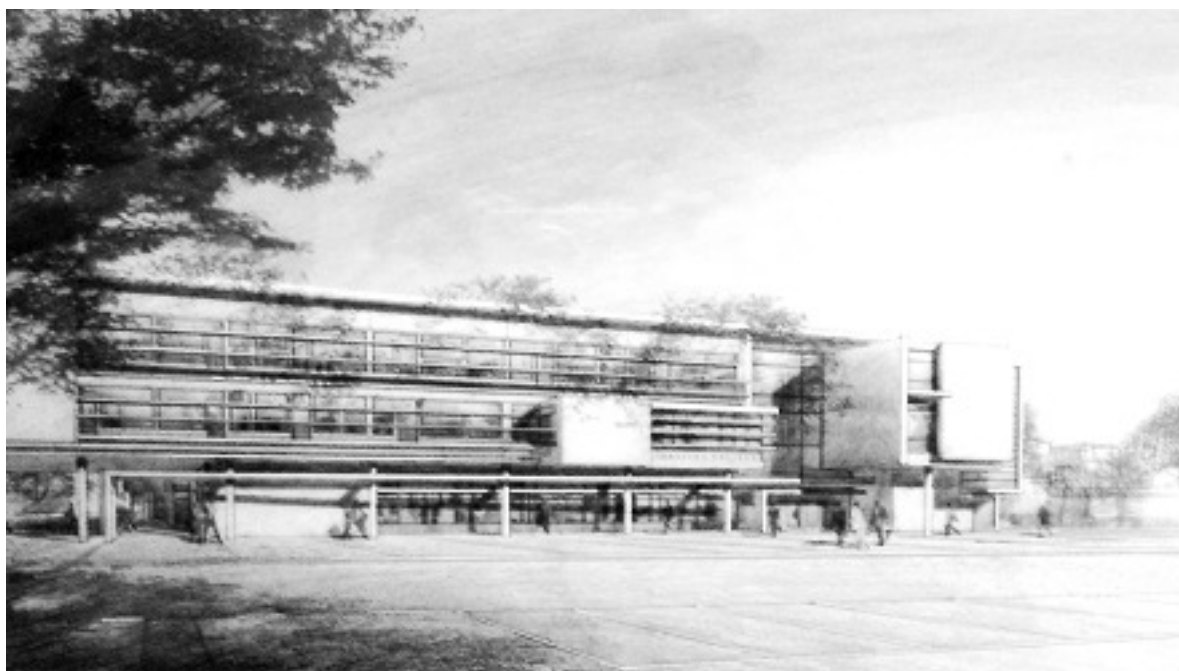


Institut Galilée

Sciences et technologies



Licence 1^{ère} année

Cours d'algèbre linéaire Deuxième semestre

Département de Mathématiques

www.math.univ-paris13.fr/departement/index.php/fr/

I. Les nombres complexes

Pour les deux premiers chapitres, le lecteur pourra consulter le livre de Liret et Martinais¹.

I.1. Les nombres réels ne suffisent pas

I.1.a. L'équation du second degré à coefficients réels

Dans de nombreux problèmes on rencontre une équation du type :

$$(E) \quad ax^2 + bx + c = 0,$$

d'inconnue réelle x , avec a, b, c des réels et $a \neq 0$. Pour résoudre une telle équation du *second degré à une inconnue*, on écrit

$$ax^2 + bx + c = a \left(x^2 + \frac{b}{a}x + \frac{c}{a} \right) = a \left(x^2 + 2\frac{b}{2a}x + \frac{c}{a} \right)$$

et on reconnaît le début du développement de

$$\left(x + \frac{b}{2a} \right)^2 = x^2 + 2\frac{b}{2a}x + \left(\frac{b}{2a} \right)^2$$

d'où

$$(I.1) \quad ax^2 + bx + c = a \left[\left(x + \frac{b}{2a} \right)^2 - \left(\frac{b}{2a} \right)^2 + \frac{c}{a} \right] = a \left[\left(x + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a^2} \right].$$

On note Δ le *discriminant* $b^2 - 4ac$. Puisque a est non nul, l'équation (E) est équivalente à sa *forme canonique* :

$$(I.2) \quad \left(x + \frac{b}{2a} \right)^2 = \frac{\Delta}{4a^2}.$$

et on distingue trois cas :

– **Premier cas** : si $\Delta > 0$, on peut écrire : $\frac{\Delta}{4a^2} = \left(\frac{\sqrt{\Delta}}{2a} \right)^2$ et l'équation (E) devient :

$$(I.3) \quad \left(x + \frac{b}{2a} + \frac{\sqrt{\Delta}}{2a} \right) \left(x + \frac{b}{2a} - \frac{\sqrt{\Delta}}{2a} \right) = 0.$$

1. François Liret et Dominique Martinais. *Algèbre 1re année - Cours et exercices avec solutions*. Dunod, deuxième édition, 2003

I. Les nombres complexes

Il y a alors 2 solutions distinctes

$$x_1 = \frac{-b - \sqrt{\Delta}}{2a} \quad \text{et} \quad x_2 = \frac{-b + \sqrt{\Delta}}{2a}$$

On peut factoriser $ax^2 + bx + c$. D'après (I.1) on a :

$$ax^2 + bx + c = a(x - x_1)(x - x_2).$$

- **Deuxième cas** : si $\Delta = 0$, l'équation (E) devient $a \left(x + \frac{b}{2a}\right)^2 = 0$. Elle admet une seule solution (double) $x_0 = -\frac{b}{2a}$ et on a la factorisation :

$$ax^2 + bx + c = a(x - x_0)^2$$

- **Troisième cas** : si $\Delta < 0$, on a :

$$ax^2 + bx + c = a \underbrace{\left[\left(x + \frac{b}{2a}\right)^2 + \frac{-\Delta}{4a^2} \right]}_{\text{strictement positif car } \Delta < 0}.$$

On en déduit que l'équation (E) n'a pas de racine réelle et que $ax^2 + bx + c$ ne peut pas se factoriser (sur \mathbb{R}). Plus précisément, on ne peut pas écrire $ax^2 + bx + c = a(x - x_1)(x - x_2)$ avec $x_1, x_2 \in \mathbb{R}$.

Exemples I.1.1.

- a. Résoudre $2x^2 - 8x + 6 = 0$. Peut-on factoriser $2x^2 - 8x + 6$?

Le discriminant de cette équation est $\Delta = (-8)^2 - 4 \times 2 \times 6 = 16 = 4^2 > 0$. Les solutions sont donc $x_1 = \frac{-(-8) - 4}{2 \times 2} = 1$ et $x_2 = \frac{-(-8) + 4}{2 \times 2} = 3$. On en déduit :

$$2x^2 - 8x + 6 = 2(x - 1)(x - 3).$$

- b. Résoudre $3x^2 - 12x + 12 = 0$. Peut-on factoriser (sur \mathbb{R}) $3x^2 - 12x + 12$?

Le discriminant de cette équation est $\Delta = (-12)^2 - 4 \times 3 \times 12 = 0$. Elle admet une seule solution $x_0 = -\frac{-12}{2 \times 3} = 2$ et on a :

$$3x^2 - 12x + 12 = 3(x - 2)^2.$$

- c. Résoudre $2x^2 + 8x + 9 = 0$. Peut-on factoriser (sur \mathbb{R}) $2x^2 + 8x + 9$?

Le discriminant de cette équation est $\Delta = 8^2 - 4 \times 2 \times 9 = -8 < 0$.

Il n'y a donc pas de solution réelle et $2x^2 + 8x + 9$ ne peut pas se factoriser (sur \mathbb{R}).

Exercice I.1.2. Résoudre :

$$6x^2 - x - 1 = 0, \quad 3x^2 - 2\sqrt{3}x + 1 = 0, \quad \frac{1}{2}x^2 + 2x + 5 = 0.$$

Peut-on factoriser ?

I.1.b. Un peu d'histoire

Nous venons de voir que toutes les équations de degré 2 n'admettent pas nécessairement de racine réelle. En particulier, l'équation simple $x^2 = -1$ soit :

$$x^2 + 1 = 0$$

n'admet pas de solution dans les nombres réels. Pour y remédier, les mathématiciens ont introduit un nombre dit *imaginaire* noté i tel que $i^2 = -1$ et construit les *nombres complexes*.

L'introduction de ces nouveaux nombres remonte au XVI^{ème} siècle. Les algébristes italiens de l'époque (Del Ferro, Tartaglia, Cardan, ...) ont découvert les formules permettant de résoudre les équations polynomiales du troisième degré, comme par exemple

$$x^3 - 7x + 6 = 0.$$

Ils ont constaté un fait qui leur a paru incompréhensible. Chaque fois qu'une équation de ce type possède trois solutions réelles, comme 1, 2 et -3 pour l'équation précédente, les formules qui leur permettaient de calculer ces solutions faisaient intervenir des racines carrées de nombres négatifs. Ils ont alors considéré ces racines carrées comme *nouveaux nombres* qu'ils ont appelés *nombres impossibles*. Néanmoins l'introduction de ces nouveaux nombres ne s'est pas faite sans mal.

La suite est tirée de *Images, imaginaires, imaginations, une perspective historique pour l'introduction des nombres complexes* IREM, éd. Ellipse. p. 157.

En 1637, Descartes dans sa *Géométrie*, propose d'accepter comme solution d'une équation non seulement les nombres négatifs, mais aussi ceux qui pourraient comporter une racine carrée d'un nombre négatif. Il justifie ceci par un théorème qui ne sera vraiment démontré qu'au XIX^{ème} siècle et qui deviendra le théorème fondamental de l'algèbre :

Une équation de degré n admet n solutions, si on accepte les négatives, celles qui comportent une racine carrée d'un nombre négatif et les multiplicités.

La construction rigoureuse des nombres complexes n'a été achevée qu'à la fin du XVIII^{ème} siècle. La notation définitive est due à Euler. Dans *Eléments d'algèbre* il écrit en 1774 en s'inspirant des règles de calcul pour les racines carrées des nombres positifs :

Maintenant comme $-a$ signifie autant que $+a$ multiplié par -1 , et que la racine carrée d'un produit se trouve en multipliant ensemble les racines des facteurs, il s'ensuit que la racine de a multiplié par -1 , ou $\sqrt{-a}$, est autant que \sqrt{a} multiplié par $\sqrt{-1}$.

Or \sqrt{a} est un nombre possible ou réel, par conséquent ce qu'il y a d'impossible dans une quantité imaginaire, peut toujours se réduire à $\sqrt{-1}$. Par cette raison donc, $\sqrt{-4}$ est autant que $\sqrt{4}$ multiplié par $\sqrt{-1}$ et autant que $2\sqrt{-1}$, à cause de $\sqrt{4}$ égale à 2. Par la même raison $\sqrt{-9}$ se réduit à $\sqrt{9}\sqrt{-1}$, ou $3\sqrt{-1}$ et $\sqrt{-16}$ signifie $4\sqrt{-1}$.

De plus comme \sqrt{a} multipliée par \sqrt{b} fait \sqrt{ab} , on aura $\sqrt{6}$ pour la valeur de $\sqrt{-2}$ multipliée par $\sqrt{-3}$.

Exercice I.1.3.

a. D'après la définition, à quoi est égal $(\sqrt{-1})^2$?

En appliquant les règles du calcul algébrique calculez $\sqrt{-1} \cdot \sqrt{-1}$.

Ces deux résultats sont-ils compatibles ?

b. Euler écrit aussi $\sqrt{-2} \cdot \sqrt{-3} = \sqrt{6}$! Or, suivant la démarche d'Euler, on va écrire

$$\sqrt{-2} \cdot \sqrt{-3} = \sqrt{2} \cdot \sqrt{-1} \cdot \sqrt{3} \cdot \sqrt{-1} = \sqrt{2} \cdot \sqrt{3} \cdot (\sqrt{-1})^2 = \dots = \dots$$

Ces deux égalités sont-elles compatibles ?

I. Les nombres complexes

Il est donc difficile d'utiliser la notation $\sqrt{-a}$ pour un réel $a > 0$, et de continuer à utiliser les règles de calcul connues pour les nombres positifs. Euler va lui-même s'apercevoir de ces contradictions. Aussi décidera-t-il de noter par i (début d'imaginaire ou impossible) la quantité qu'il notait $\sqrt{-1}$.

On peut tout de suite noter la règle suivante :

La notation *racine carrée* $\sqrt{\quad}$ ne s'utilise qu'avec des nombres réels positifs

I.2. Forme cartésienne d'un nombre complexe, addition et multiplication

I.2.a. Rappel : produit cartésien de deux ensembles

Définition I.2.1. Soit A et B deux ensembles. Le produit cartésien $A \times B$ est l'ensemble des couples (a, b) avec $a \in A$ et $b \in B$. On note $A^2 = A \times A$.

En particulier \mathbb{R}^2 est l'ensemble des couples de réels (x, y) . L'ordre est important : par définition, $(x, y) = (x', y')$ si et seulement si $x = x'$ et $y = y'$. Ainsi, $(1, 2) \neq (2, 1)$.

I.2.b. Construction des nombres complexes

Nous avons vu qu'il a été nécessaire d'introduire des nombres ayant un carré négatif. Pour ce faire, on va construire un ensemble muni de deux opérations, l'ensemble des nombres complexes, qui contient les nombres réels et des nombres dits imaginaires, dont le nombre i vérifiant $i^2 = -1$.

Par définition, \mathbb{C} (l'ensemble des *nombres complexes*) est l'ensemble \mathbb{R}^2 muni des deux opérations suivantes.

– *Addition* :

$$(x, y) + (x', y') = (x + x', y + y').$$

– *Multiplication* :

$$(x, y).(x', y') = (xx' - yy', xy' + yx').$$

Si $z = (x, y) \in \mathbb{C}$, le nombre réel x est appelé *partie réelle* de z et noté $\operatorname{Re} z$. Le nombre réel y est appelé *partie imaginaire* de z et noté $\operatorname{Im} z$. Par définition, deux nombres complexes z et z' sont égaux lorsque leurs parties réelle et imaginaire sont égales.

Un nombre complexe de la forme $(x, 0)$ est dit réel et simplement noté x , ce qui permet d'identifier \mathbb{R} à un sous-ensemble de \mathbb{C} . En effet, l'addition et la multiplication complexes restreintes aux nombres réels coïncident avec l'addition et la multiplication réelles :

$$(x, 0) + (x', 0) = (x + x', 0), \quad (x, 0).(x', 0) = (xx', 0).$$

Un nombre complexe de la forme $(0, y)$ est appelé *nombre imaginaire pur* et noté iy . Le seul nombre réel et imaginaire pur est le nombre $(0, 0)$, noté simplement 0 . On note i le nombre complexe $1i$. Par définition de la multiplication sur les nombres complexes, on a :

$$i^2 = -1.$$

I.2. Forme cartésienne d'un nombre complexe, addition et multiplication

Compte tenu de l'identification précédente et de la définition de l'addition complexe, tout nombre complexe s'écrit $(x, y) = x + iy$: cette écriture est appelée *forme cartésienne* d'un nombre complexe. Elle est unique : si x, x', y et y' sont des nombres réels,

$$x + iy = x' + iy' \iff (x = x' \text{ et } y = y').$$

On utilisera désormais *systématiquement* la notation $x + iy$, au lieu de (x, y) .

Pour récapituler, on peut "oublier" la construction précédente, et décrire \mathbb{C} comme suit. **Les nombres complexes sont les nombres $z = x + iy$ où x (la partie réelle de z) et y (la partie imaginaire de z) sont des nombres réels, et $i^2 = -1$. Ils s'additionnent et se multiplient de la manière suivante :**

$$(I.4) \quad (x + iy) + (x' + iy') = x + x' + i(y + y')$$

$$(I.5) \quad (x + iy).(x' + iy') = xx' - yy' + i(xy' + yx').$$

Remarque I.2.2. Le produit $z.z'$ de deux nombres complexes z et z' est aussi noté zz' ou $z \times z'$.

I.2.c. Propriétés de l'addition et de la multiplication

Toutes les propriétés de l'addition et de la multiplication dans \mathbb{R} (commutativité, associativité, distributivité...) restent vraies dans \mathbb{C} . Ainsi, il est facile de vérifier que (à faire au moins une fois) :

- i. $z + 0 = 0 + z = z$, $1z = z1 = z$ et $0z = z0 = 0$.
- ii. $z + z' = z' + z$ et $zz' = z'z$.
- iii. $z + (z' + z'') = (z + z') + z''$ et $z(z'z'') = (zz')z''$.
- iv. $z(z' + z'') = zz' + zz''$.

Remarque I.2.3. On peut écrire la somme (respectivement le produit) de trois nombres complexes z, z' et z'' , sans parenthèse : $z + z' + z''$ (respectivement $z z' z''$). Ces notations ne sont pas ambiguës du fait de l'associativité de l'addition et de la multiplication iii.

Remarque I.2.4. Il est inutile d'apprendre par coeur la formule (I.5) définissant la multiplication de deux nombres complexes. Elle se retrouve immédiatement en utilisant les propriétés précédentes et le fait que $i^2 = -1$:

$$(x + iy)(x' + iy') = xx' + iyx' + xiy' + iyiy' = xx' - yy' + i(xy' + x'y).$$

Remarque I.2.5. La formule du binôme ainsi que les formules des sommes de suites arithmétiques et géométriques découlent des propriétés standard de l'addition et de la multiplication : elles restent vraies pour des nombres complexes. Ces formules sont à savoir. Elles sont rappelées en appendice (cf I.7.b, I.7.c).

I.2.d. Représentation dans le plan

Le plan est muni d'un repère orthonormé $(O, \vec{e}_1, \vec{e}_2)$.

- i. Par construction de \mathbb{C} (I.2.b plus haut), le nombre complexe z de forme cartésienne $x + iy$ est naturellement associée au point M de coordonnées (x, y) .
 M est l'image ponctuelle de z et z est l'affixe de M .

I. Les nombres complexes

- ii. On associe aussi le nombre complexe z de forme cartésienne $x + iy$ avec le vecteur $\vec{v} = \overrightarrow{OM} = x\vec{e}_1 + y\vec{e}_2$. \vec{v} est l'image vectorielle de z et z est l'afixe du vecteur \vec{v} .

Par définition, deux nombres complexes sont égaux s'ils ont même partie réelle et même partie imaginaire.

$$x + iy = x' + iy' \quad \text{si et seulement si} \quad x = x' \text{ et } y = y',$$

c'est à dire si et seulement si leurs images (ponctuelle ou vectorielle) sont confondues.

On peut facilement interpréter géométriquement l'addition sur \mathbb{C} . Étant donnés deux nombres complexes $z = x + iy$ et $z' = x' + iy'$, si on note \vec{v} et \vec{v}' les vecteurs du plan complexe d'affixes z et z' , le nombre $z + z'$ est l'afixe du vecteur $\vec{v} + \vec{v}'$ (la somme s'obtient par la loi du parallélogramme) (c.f. figure I.1 p. 7).

I.3. Autres opérations sur les nombres complexes

On a vu la définition de l'addition et la multiplication de deux nombres complexes. On définit ici d'autres opérations sur ces nombres : opposé, différence, conjugaison et module.

I.3.a. Opposé, différence de nombres complexes

Pour tout $z = x + yi \in \mathbb{C}$, le nombre complexe $-z = -x + (-y)i$ est l'unique nombre complexe z' tel que $z + z' = 0$; ce nombre complexe est l'opposé de z et on le note simplement $-z$. La différence $z - z'$ de deux nombres complexes, est alors définie par $z - z' = z + (-z')$.

I.3.b. Conjugaison et module

Définition I.3.1.

Si $z \in \mathbb{C}$, $x = \operatorname{Re} z$, $y = \operatorname{Im} z$, le nombre complexe $\bar{z} = x - yi$ est appelé le *complexe conjugué* de z .

Le nombre réel positif $|z| = \sqrt{x^2 + y^2}$ est appelé le *module* de z .

La conjugaison est une involution :

$$\overline{\bar{z}} = z.$$

Un calcul simple (à faire) montre :

$$(I.6) \quad |z|^2 = z\bar{z}.$$

Par ailleurs

$$|z|^2 = 0 \iff x = 0 \text{ et } y = 0 \iff z = 0.$$

Interprétation géométrique

Soit z un nombre complexe, d'affixe M . Le conjugué \bar{z} de z a pour affixe M' , l'image de M par la symétrie d'axe Ox . Le module de z est la distance OM (cf figure I.1 p. 7).

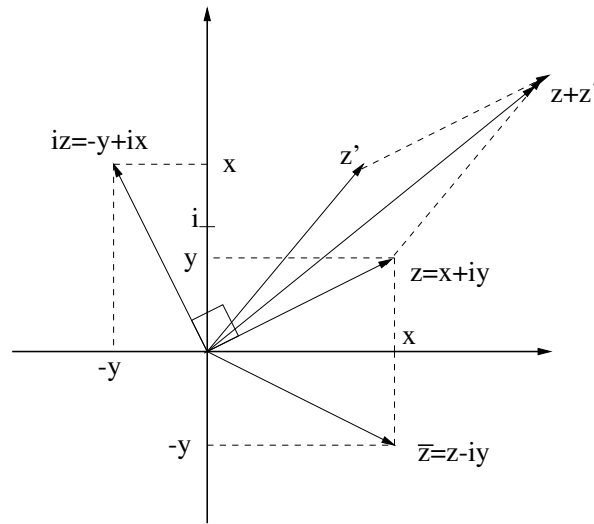


FIGURE I.1.: Le plan complexe : coordonnées cartésiennes

Compatibilité avec l'addition et la multiplication

- Concernant la conjugaison on a :
 - i. $\overline{z + z'} = \bar{z} + \bar{z}'$ et $\overline{z \cdot z'} = \bar{z} \cdot \bar{z}'$.
 - ii. $\operatorname{Re} z = \frac{z + \bar{z}}{2}$ et $\operatorname{Im} z = \frac{z - \bar{z}}{2i}$.
 - iii. $z = \bar{z}$ si et seulement si z est réel.
 - iv. $z = -\bar{z}$ si et seulement si z est imaginaire pur.

Les preuves sont laissées au lecteur. On retiendra que la conjugaison est compatible avec les opérations (propriétés i), et permet avec les trois dernières relations de déterminer si un nombre complexe est réel ou imaginaire pur.

- Concernant le module, on a déjà vu que $|z|^2 = z\bar{z}$ (I.6), on a aussi :
 - v. $|\operatorname{Re} z| \leq |z|$, $|\operatorname{Im} z| \leq |z|$ et $|z| = |\bar{z}|$.
 - vi. $|z \cdot z'| = |z| \cdot |z'|$.
 - vii. $|z + z'| \leq |z| + |z'|$, (inégalité triangulaire).
 - viii. $||z| - |z'|| \leq |z - z'|$.

Le module est donc compatible avec le produit et le quotient (propriété vi). Par contre, on a seulement une inégalité pour la somme, l'inégalité triangulaire vii.

Démonstration. - La propriété v s'obtient par des calculs directs.

- Pour vi on peut par exemple utiliser (I.6) et la propriété i.
- La démonstration de l'inégalité triangulaire n'est pas directe :

On a avec (I.6) : $|z + z'|^2 = (z + z')\overline{(z + z')} = (z + z')(\bar{z} + \bar{z}') = |z|^2 + |z'|^2 + z\bar{z}' + \bar{z}z'$.

Or par i, $\bar{z}z' = \overline{z\bar{z}'}$. Donc par ii, puis v,

$$z\bar{z}' + \bar{z}z' = z\bar{z}' + \overline{z\bar{z}'} = 2 \operatorname{Re}(z\bar{z}') \leq 2|z\bar{z}'| = 2|z||z'|.$$

Ainsi, on obtient avec la première égalité

$$|z + z'|^2 \leq |z|^2 + |z'|^2 + 2|z||z'| = (|z| + |z'|)^2$$

I. Les nombres complexes

ce qui donne l'inégalité triangulaire puisque deux nombres positifs sont dans le même ordre que leurs carrés (car la fonction $x \mapsto x^2$ est croissante sur $[0, +\infty)$).

- La dernière inégalité viii se déduit de la précédente, en écrivant $z = (z - z') + z'$. □

I.3.c. Inverse et quotient

Inverse d'un nombre complexe non nul

Proposition I.3.2. Pour tout nombre complexe $z \in \mathbb{C} \setminus \{0\}$, il existe un unique $z' \in \mathbb{C}$ tel que $zz' = 1$. De plus, $z' = \frac{1}{|z|^2} \bar{z}$.

Remarque I.3.3. L'expression $\frac{1}{|z|^2} \bar{z}$ dans la proposition désigne le produit du nombre réel $\frac{1}{|z|^2}$ (bien défini, car $|z|$ est un réel non nul) par le nombre complexe \bar{z} .

Démonstration. Commençons par montrer l'unicité. Soit $z \in \mathbb{C}$ non nul, et z' tel que $zz' = 1$. En multipliant par \bar{z} et en utilisant la formule (I.6) $z\bar{z} = |z|^2$, on obtient $|z|^2 z' = \bar{z}$. On multiplie ensuite par le nombre réel $\frac{1}{|z|^2}$, ce qui donne

$$(I.7) \quad z' = \frac{1}{|z|^2} \bar{z}.$$

Si z' est un inverse de z , il est donc obligatoirement donné par la formule (I.7) ce qui donne l'unicité.

Réciproquement, en utilisant encore la formule (I.6), on a bien $z \frac{1}{|z|^2} \bar{z} = 1$. □

Définition I.3.4. Pour tout nombre complexe z non nul, l'unique z' tel que $zz' = 1$ est appelé inverse de z et noté $\frac{1}{z}$.

Corollaire I.3.5. Si $zz' = 0$ alors $z = 0$ ou $z' = 0$.

Démonstration. Supposons $zz' = 0$ et $z \neq 0$. En multipliant l'équation $zz' = 0$ par $1/z$ on obtient $z' = 0$. □

Quotient

- Si $z, z' \in \mathbb{C}$ et si $z' \neq 0$, le quotient de z par z' noté $\frac{z}{z'}$ est défini par $\frac{z}{z'} = z \times \frac{1}{z'}$.

- Méthode pour trouver la forme cartésienne d'un quotient :

si $z = x + iy$, $z' = x' + iy' \neq 0$ alors

$$\begin{aligned} \frac{z}{z'} &= \frac{z\bar{z}'}{z'\bar{z}'} = \frac{(x + iy)(x' - iy')}{(x' + iy')(x' - iy')} \\ &= \frac{xx' + yy' + i(x'y - xy')}{(x')^2 + (y')^2} \\ &= \frac{xx' + yy'}{(x')^2 + (y')^2} + i \frac{x'y - xy'}{(x')^2 + (y')^2}. \end{aligned}$$

- Compatibilité avec la conjugaison et le module Si $z, z' \in \mathbb{C}$ avec $z' \neq 0$, alors $\overline{\left(\frac{z}{z'}\right)} = \frac{\bar{z}}{\bar{z}'}$

et $\left|\frac{z}{z'}\right| = \frac{|z|}{|z'|}$. (ceci découle des formules $\overline{ab} = \bar{a}\bar{b}$ et $|ab| = |a||b|$ appliquées à $a = z'$, $b = \frac{z}{z'}$).

Exercices

Exercice I.3.6. Calculer $\frac{z}{z'}$ dans chacun des cas suivants :

- i. $z = 1 - i$, $z' = \sqrt{2} + \sqrt{2}i$.
- ii. $z = 1 - i$, $z' = 1 + i\sqrt{3}$.
- iii. $z = -\sqrt{3} - i$, $z' = i$.
- iv. $z = 3 + 2i$, $z' = 3 - 2i$.

Solution du premier cas :

$$\frac{1 - i}{\sqrt{2} + \sqrt{2}i} = \frac{(1 - i)(\sqrt{2} - \sqrt{2}i)}{(\sqrt{2} + \sqrt{2}i)(\sqrt{2} - \sqrt{2}i)} = \frac{\sqrt{2} - \sqrt{2} + i(-\sqrt{2} - \sqrt{2})}{\sqrt{2}^2 + \sqrt{2}^2} = -i \frac{\sqrt{2}}{2}.$$

Exercice I.3.7.

- a. Soit M un point du plan d'affixe $z \neq 0$. Construire le point M' d'affixe $1/z$.
- b. Comment faut-il choisir z pour que $Z = \frac{5z - 2}{z - 1}$ soit réel ?
- c. Déterminer l'ensemble des nombres complexes z tels que $(z + 1)(\bar{z} - i)$ soit un imaginaire pur.

I.4. Forme polaire d'un nombre complexe**I.4.a. Rappels**

Pour repérer un point M dans le plan, on peut utiliser les coordonnées cartésiennes (x, y) , mais aussi les coordonnées polaires (r, θ) où r est la longueur du segment OM et θ est une mesure de l'angle (\vec{Ox}, \vec{OM}) .

Les relations entre coordonnées cartésiennes et coordonnées polaires sont :

$$x = r \cos \theta \text{ et } y = r \sin \theta.$$

Tout nombre complexe $z = x + iy$ peut donc s'écrire sous la forme $z = r(\cos \theta + i \sin \theta)$, avec r un nombre réel positif.

I.4.b. Définition

- Cette écriture $z = r(\cos \theta + i \sin \theta)$ avec $r \geq 0$ est appelée la *forme polaire* de z .
- $r = \sqrt{x^2 + y^2}$ est exactement le module de z .
- θ , noté $\arg(z)$, est un *argument* de z . Remarquons que θ est défini modulo 2π .

Notation I.4.1. On note $a \equiv b \pmod{2\pi}$ (ou $a = b \pmod{2\pi}$) quand

$$\exists k \in \mathbb{Z}, a = b + 2k\pi.$$

I.4.c. Propriétés

- *Égalité* : deux nombres complexes non nuls, exprimés sous forme polaire, sont égaux si et seulement s'ils ont même module et si leurs arguments diffèrent de $2k\pi$, où k est un nombre entier.
- *Produit* : si $z_1 = r_1(\cos \theta_1 + i \sin \theta_1)$ et $z_2 = r_2(\cos \theta_2 + i \sin \theta_2)$ sont deux nombres complexes non nuls, exprimés sous forme polaire, le calcul de leur produit donne

$$z_1 z_2 = r_1 r_2 \left[\cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 + i (\cos \theta_1 \sin \theta_2 + \cos \theta_2 \sin \theta_1) \right] \\ = r_1 r_2 (\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)),$$

par des formules trigonométriques classiques. On en déduit les relations :

$$(I.8) \quad |z_1 z_2| = |z_1| |z_2| \quad \text{et} \quad \arg(z_1 z_2) \equiv \arg(z_1) + \arg(z_2) \pmod{2\pi}.$$

Si $z \neq 0$ $\left| \frac{1}{z} \right| = \frac{1}{|z|}$ et $\arg\left(\frac{1}{z}\right) = -\arg(z) \pmod{2\pi}$ (d'après l'égalité $z \frac{1}{z} = 1$).

- *Conjugaison* : si $z \neq 0$, $\arg(\bar{z}) = -\arg(z) \pmod{2\pi}$.
Ceci découle immédiatement des formules $\cos(-\theta) = \cos \theta$, $\sin(-\theta) = -\sin \theta$.
- *Méthode pour trouver la forme polaire d'un nombre complexe $z = x + yi \neq 0$* : si θ est un argument de z , on a : $z = |z| \frac{z}{|z|} = |z| \left(\frac{x}{\sqrt{x^2 + y^2}} + i \frac{y}{\sqrt{x^2 + y^2}} \right)$ ainsi :

$$\cos \theta = \frac{x}{\sqrt{x^2 + y^2}} \quad \text{et} \quad \sin \theta = \frac{y}{\sqrt{x^2 + y^2}}.$$

Exercice I.4.2. Écrire sous forme polaire : $1 - i$, $-1 + i$, i , $1 + i\sqrt{3}$, $-\sqrt{3} - i$.

I.4.d. Écriture exponentielle de la forme polaire

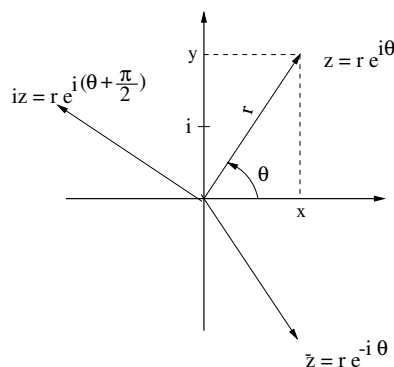


FIGURE I.2.: Le plan complexe : coordonnées polaires

Par convention, on note tout nombre complexe de module 1 sous la forme

$$\cos \theta + i \sin \theta = e^{i\theta}.$$

Cette exponentielle complexe vérifie les mêmes règles de calcul que l'exponentielle réelle. Par (I.8), on a :

$$(I.9) \quad e^{i(\alpha+\beta)} = e^{i\alpha} e^{i\beta}.$$

Pour représenter un nombre complexe sous forme polaire, on utilisera désormais l'écriture exponentielle :

$$z = r e^{i\theta}.$$

(c.f. figure I.2 p. 10). Avec cette écriture, les différentes propriétés que nous avons rencontrées s'écrivent (r, r_1, r_2 sont des nombres réels strictement positifs, $\theta, \theta_1, \theta_2$ des nombres réels) :

- égalité : $r_1 e^{i\theta_1} = r_2 e^{i\theta_2} \Leftrightarrow \begin{cases} r_1 = r_2 \\ \theta_1 = \theta_2 \pmod{2\pi} \end{cases}$
- conjugaison : $\overline{e^{i\theta}} = e^{-i\theta}$
- module : $|r e^{i\theta}| = r$
- produit : $(r_1 e^{i\theta_1}) (r_2 e^{i\theta_2}) = r_1 r_2 e^{i(\theta_1+\theta_2)}$
- quotient : pour $r_2 \neq 0$, $\frac{r_1 e^{i\theta_1}}{r_2 e^{i\theta_2}} = \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)}$

Exercice I.4.3. Écrire sous forme cartésienne : $2e^{i\frac{\pi}{6}}$, $3e^{-2i\frac{\pi}{3}}$, $e^{-i\frac{\pi}{6}}$, $3e^{i\pi}$, $\sqrt{3}e^{i\frac{\pi}{4}}$.

I.4.e. Formule d'Euler

Si $\theta \in \mathbb{R}$ on a la formule d'Euler :

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad \text{et} \quad \sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

Démonstration. On a $\cos \theta = \operatorname{Re} e^{i\theta}$ et $\sin \theta = \operatorname{Im} e^{i\theta}$; on obtient les dites formules en utilisant la relation ii p. 7. □

Linéarisation de $\cos^n \theta$ et $\sin^n \theta$

Linéariser $\cos^n \theta$ ou $\sin^n \theta$, c'est en donner une expression qui ne contient aucun produit de fonctions circulaires. Cette opération est possible, en développant $\cos^n \theta = \left(\frac{e^{i\theta} + e^{-i\theta}}{2} \right)^n$ à l'aide de la formule du binôme (proposition I.7.4 p. 18).

I. Les nombres complexes

Exemple I.4.4.

$$\begin{aligned}\cos^3 \theta &= \left(\frac{e^{i\theta} + e^{-i\theta}}{2} \right)^3 = \frac{(e^{i\theta} + e^{-i\theta})^3}{2^3} \\ &= \frac{1}{2^3} [e^{3i\theta} + 3e^{2i\theta}e^{-i\theta} + 3e^{i\theta}e^{-2i\theta} + e^{-3i\theta}] \\ &= \frac{1}{2^3} [e^{3i\theta} + e^{-3i\theta} + 3(e^{i\theta} + e^{-i\theta})] \\ &= \frac{1}{4} \cos 3\theta + \frac{3}{4} \cos \theta\end{aligned}$$

$$\begin{aligned}\sin^3 \theta &= \left(\frac{e^{i\theta} - e^{-i\theta}}{2i} \right)^3 = \frac{(e^{i\theta} - e^{-i\theta})^3}{2^3 i^3} \\ &= \frac{-1}{2^3 i} [e^{3i\theta} - e^{-3i\theta} - 3(e^{i\theta} - e^{-i\theta})] \\ &= \frac{-1}{4} \sin 3\theta + \frac{3}{4} \sin \theta\end{aligned}$$

Exercice I.4.5. Linéariser $\cos^4 \theta$, $\sin^4 \theta$ et $\cos^3 \theta \sin^2 \theta$.

I.4.f. Formule de Moivre

On établit par récurrence que, pour tout $n \in \mathbb{N}$, $e^{in\theta} = (\cos \theta + i \sin \theta)^n$, soit :

$$\text{Pour tout } n \in \mathbb{N}, \quad \cos n\theta + i \sin n\theta = (\cos \theta + i \sin \theta)^n.$$

C'est ce qu'on appelle la formule de De Moivre.

Application de la formule de Moivre

Calcul de $\cos n\theta$ et $\sin n\theta$ en fonction de $\cos \theta$ et $\sin \theta$.

C'est l'opération inverse de la linéarisation.

Pour la réaliser, on utilise le fait que $\cos n\theta$ est la partie réelle (et $\sin n\theta$ la partie imaginaire) de $(\cos \theta + i \sin \theta)^n$, que l'on développe à l'aide de la formule du binôme (rappelée en §I.7.b, p. 17).

Exemple pour $n = 3$:

$$\begin{aligned}\cos 3\theta + i \sin 3\theta &= (\cos \theta + i \sin \theta)^3 \\ &= \cos^3 \theta + 3i \cos^2 \theta \sin \theta + 3i^2 \cos \theta \sin^2 \theta + i^3 \sin^3 \theta \\ &= \cos^3 \theta - 3 \cos \theta \sin^2 \theta + i(3 \cos^2 \theta \sin \theta - \sin^3 \theta).\end{aligned}$$

En identifiant les parties réelles et imaginaires des deux membres, on obtient

$$\begin{aligned}\cos 3\theta &= \cos^3 \theta - 3 \cos \theta \sin^2 \theta \\ \sin 3\theta &= 3 \cos^2 \theta \sin \theta - \sin^3 \theta.\end{aligned}$$

Si on le souhaite, on peut améliorer ces égalités en utilisant $\cos^2 \theta + \sin^2 \theta = 1$, ce qui donne alors

$$\begin{aligned}\cos 3\theta &= 4 \cos^3 \theta - 3 \cos \theta \\ \sin 3\theta &= 3 \sin \theta - 4 \sin^3 \theta.\end{aligned}$$

Exercice I.4.6. Écrire $\cos 4\theta$, $\sin 4\theta$ et $\cos 3\theta \sin 2\theta$ en fonction de puissances de $\cos \theta$ et $\sin \theta$.

I.5. Racines $n^{\text{ièmes}}$ d'un nombre complexe

I.5.a. Cas général

Définition I.5.1. Soit $z_0 \in \mathbb{C}$ et $n \in \mathbb{N}^*$, on appelle racine $n^{\text{ième}}$ du nombre complexe z_0 tout nombre complexe ω tel que $\omega^n = z_0$.

Par exemple $\frac{\sqrt{2}}{2}(1+i)$ est une racine carrée de i , i est une racine carrée de -1 et donc $\frac{\sqrt{2}}{2}(1+i)$ est une racine quatrième de -1 .

Les racines $n^{\text{ièmes}}$ de z_0 sont les nombres complexes ω solutions de l'équation en z :

$$z^n - z_0 = 0.$$

Le nombre 0 est la seule racine $n^{\text{ième}}$ de 0 car : $z^n = 0 \Leftrightarrow |z^n| = 0 \Leftrightarrow |z|^n = 0 \Leftrightarrow |z| = 0 \Leftrightarrow z = 0$.

Proposition I.5.2. *Tout nombre complexe z_0 non nul possède exactement n racines $n^{\text{ièmes}}$ distinctes. De plus si $\rho_0 e^{i\theta_0}$ est la forme polaire de z_0 , elles sont de la forme*

$$\omega_k = \sqrt[n]{\rho_0} e^{i(\frac{\theta_0}{n} + \frac{k}{n}2\pi)}, \quad k \in \{0, 1, \dots, n-1\}.$$

Démonstration. On cherche tous les nombres complexes $\omega = r e^{it}$ qui vérifient l'équation $\omega^n = \rho_0 e^{i\theta_0}$. D'après la propriété d'égalité de deux nombres complexes, sous forme polaire, on a :

$$\begin{aligned}r^n e^{nit} = \rho_0 e^{i\theta_0} &\iff r^n = \rho_0 \quad \text{et} \quad \exists k \in \mathbb{Z}, \quad nt = \theta_0 + 2k\pi \\ &\iff r = \sqrt[n]{\rho_0} \quad \text{et} \quad \exists k \in \mathbb{Z}, \quad t = \frac{\theta_0}{n} + \frac{2k\pi}{n}.\end{aligned}$$

Ainsi pour tout $k \in \mathbb{Z}$, le nombre complexe $\omega_k = \sqrt[n]{\rho_0} e^{i(\frac{\theta_0}{n} + \frac{k}{n}2\pi)}$ est racine et toutes les racines sont de cette forme.

Montrons qu'il n'y a, en fait, que n racines distinctes : posons $t_k = \frac{\theta_0}{n} + \frac{k}{n}2\pi$ pour $0 \leq k \leq n-1$. Chaque $r e^{it_k}$ est racine, et ils sont deux à deux distincts : en effet, supposons que $r e^{it_k} = r e^{it_\ell}$, $k \neq \ell$, alors $(t_k - t_\ell) = 2p\pi$, soit $k - \ell = pn$ pour un certain $p \in \mathbb{Z}$. Mais comme $0 \leq k, \ell \leq n-1$, on a $|k - \ell| \leq n-1$, donc $|p|n \leq n-1$, ce qui force $p = 0$.

Il reste à montrer que toutes racines $n^{\text{ième}}$ de z_0 est de la forme ω_j , avec $0 \leq j \leq n-1$. Soit ω_k une telle racine, d'argument $t_k = \frac{\theta_0}{n} + \frac{k}{n}2\pi$, $k \in \mathbb{Z}$. Alors avec la division euclidienne de k

I. Les nombres complexes

par n , il existe $0 \leq k' \leq n-1$ et $p \in \mathbb{Z}$ tels que $k = np + k'$, et l'argument t_k s'écrit donc $\frac{\theta_0}{n} + \frac{k'}{n}2\pi + p2\pi$. Ainsi

$$\omega_k = \sqrt[n]{\rho_0} e^{i(\frac{\theta_0}{n} + \frac{k'}{n}2\pi + p2\pi)} = \sqrt[n]{\rho_0} e^{i(\frac{\theta_0}{n} + \frac{k'}{n}2\pi)} = \omega_{k'},$$

ce qui achève la preuve. \square

Exemple I.5.3 (Calcul des racines carrées de $1-i$). $1-i = \sqrt{2}e^{-i\frac{\pi}{4}}$. Si $z = re^{i\theta}$ est solution de $z^2 = 1-i$ alors $z^2 = r^2e^{2i\theta} = \sqrt{2}e^{-i\frac{\pi}{4}}$. D'après ce qui précède, on a deux solutions :

$$z_1 = \sqrt[4]{2}e^{-i\frac{\pi}{8}} \text{ et } z_2 = \sqrt[4]{2}e^{i\frac{7\pi}{8}}$$

I.5.b. Racines $n^{\text{ièmes}}$ de l'unité.

On s'intéresse ici au cas où $z_0 = 1$ c'est à dire aux solutions complexes de l'équation en z

$$z^n - 1 = 0.$$

La précédente proposition nous donne :

Proposition I.5.4. *Le nombre 1 possède exactement n racines $n^{\text{ièmes}}$ distinctes :*

$$\omega_k = e^{i\frac{k}{n}2\pi} \quad k \in \{0, 1, \dots, n-1\}$$

Propriété

Si ω_k est une racine $n^{\text{ième}}$ de l'unité autre que 1 (soit k non multiple de n), alors

$$1 + \omega_k + \omega_k^2 + \dots + \omega_k^{n-1} = 0.$$

Pour la preuve, il suffit d'utiliser la formule de la somme d'une suite géométrique (I.19) rappelée dans l'appendice p. 19. Puisque $\omega_k \neq 1$ et $\omega_k^n = 1$, on a :

$$1 + \omega_k + \omega_k^2 + \dots + \omega_k^{n-1} = \frac{1 - \omega_k^n}{1 - \omega_k} = 0.$$

Lorsque $k = 1$ on obtient la formule :

$$1 + \omega_1 + \omega_2 + \dots + \omega_{n-1} = 0.$$

Cas particulier : racines cubiques de l'unité

D'après I.5.2, les racines cubiques de l'unité sont $1, e^{2i\pi/3}, e^{4i\pi/3}$. Par convention, on pose

$$j = e^{2i\pi/3}.$$

On en déduit : $j^2 = \bar{j} = e^{4i\pi/3}$. On retiendra que les racines cubiques de l'unité sont : $1, j$ et j^2 et qu'elles vérifient la relation

$$1 + j + j^2 = 0.$$

Exercices

Exercice I.5.5. Soit z un nombre complexe non nul et d une racine $n^{\text{ième}}$ de z .

Montrer qu'il suffit de multiplier d par les racines $n^{\text{ièmes}}$ de l'unité pour obtenir les racines $n^{\text{ièmes}}$ de z .

Exercice I.5.6. Déterminer les racines cubiques de -8 .

I.5.c. Racines carrées d'un nombre complexe, sous forme cartésienne

Étant donné un nombre complexe $z_0 = x_0 + iy_0$, on cherche les nombres complexes $z = x + iy$ tels que $z^2 = z_0$.

Comme $z^2 = x^2 - y^2 + 2ixy$, on obtient les équations

$$\begin{cases} x^2 - y^2 = x_0 \\ 2xy = y_0. \end{cases}$$

Remarquons par ailleurs que si $z^2 = z_0$ alors $|z|^2 = |z_0|$ c'est à dire que $x^2 + y^2 = |z_0|$.

On est donc amené à résoudre le système

$$\begin{cases} x^2 - y^2 = x_0 \\ x^2 + y^2 = |z_0| \\ 2xy = y_0. \end{cases}$$

d'où $x^2 = \frac{x_0 + |z_0|}{2} = \frac{|z_0| + \operatorname{Re} z_0}{2}$ et $y^2 = \frac{|z_0| - x_0}{2} = \frac{|z_0| - \operatorname{Re} z_0}{2}$.

Comme on a toujours $|\operatorname{Re} z_0| \leq |z_0|$, $\frac{|z_0| + \operatorname{Re} z_0}{2} \geq 0$ et $\frac{|z_0| - \operatorname{Re} z_0}{2} \geq 0$, on peut donc calculer leur racine carrée dans \mathbb{R} . On trouve

$$\begin{cases} x = \pm \sqrt{\frac{|z_0| + \operatorname{Re} z_0}{2}} \\ y = \pm \sqrt{\frac{|z_0| - \operatorname{Re} z_0}{2}}. \end{cases}$$

La condition $2xy = y_0 = \operatorname{Im} z_0$ permet de déterminer les signes \pm . Ainsi, si $\operatorname{Im} z_0 \geq 0$ alors $xy \geq 0$ donc x et y sont de même signe. Les solutions sont

$$z = x + iy = \sqrt{\frac{|z_0| + \operatorname{Re} z_0}{2}} + i\sqrt{\frac{|z_0| - \operatorname{Re} z_0}{2}} \quad \text{et} \quad z = -\sqrt{\frac{|z_0| + \operatorname{Re} z_0}{2}} - i\sqrt{\frac{|z_0| - \operatorname{Re} z_0}{2}}.$$

Par contre, si $\operatorname{Im} z_0 \leq 0$ alors $xy \leq 0$ donc x et y sont de signe opposé. Les solutions sont

$$z = x + iy = \sqrt{\frac{|z_0| + \operatorname{Re} z_0}{2}} - i\sqrt{\frac{|z_0| - \operatorname{Re} z_0}{2}} \quad \text{et} \quad z = -\sqrt{\frac{|z_0| + \operatorname{Re} z_0}{2}} + i\sqrt{\frac{|z_0| - \operatorname{Re} z_0}{2}}.$$

Ces formules ne sont évidemment pas à savoir, par contre la méthode est à connaître.

Exemple I.5.7 (retour au calcul des racines carrées de $1 - i$). Si $z = x + iy$ est solution de $z^2 = z_0$ alors $z^2 = (x^2 - y^2) + 2ixy = 1 - i$.

On en déduit : $x^2 - y^2 = 1$ et $xy = -\frac{1}{2}$.

Par ailleurs, $|z|^2 = |1 - i|$ donne $x^2 + y^2 = \sqrt{2}$.

I. Les nombres complexes

On trouve $x^2 = \frac{1+\sqrt{2}}{2}$ soit $x = \pm\sqrt{\frac{1+\sqrt{2}}{2}}$ et $y^2 = \frac{-1+\sqrt{2}}{2}$ soit $y = \pm\sqrt{\frac{-1+\sqrt{2}}{2}}$.

Enfin, comme $2xy = -1$, x et y sont de signe opposé.

Ainsi, les solutions sont

$$z_1 = \sqrt{\frac{1+\sqrt{2}}{2}} - \frac{i}{\sqrt{2(1+\sqrt{2})}} \quad \text{et} \quad z_2 = -\sqrt{\frac{1+\sqrt{2}}{2}} + \frac{i}{\sqrt{2(1+\sqrt{2})}}.$$

(comparer avec le résultat de l'exemple I.5.3).

I.6. Équation du second degré à coefficients complexes

Proposition I.6.1. Soient $a, b, c \in \mathbb{C}$, avec $a \neq 0$. Alors l'équation $az^2 + bz + c = 0$ admet deux solutions complexes (qui peuvent être identiques).

On procède comme dans le cas réel :

$$(I.10) \quad az^2 + bz + c = a\left(z + \frac{b}{2a}\right)^2 - \frac{b^2}{4a} + c = a\left[\left(z + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right],$$

donc $az^2 + bz + c = 0$ est équivalent à

$$\left(z + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2} = 0,$$

où Δ est le discriminant $\Delta = b^2 - 4ac$. On distingue deux cas :

- Si $\Delta \neq 0$, alors $\Delta = b^2 - 4ac$ admet deux racines carrées complexes distinctes, notées δ et $-\delta$. Alors les solutions sont

$$(I.11) \quad z_1 = -\frac{b}{2a} + \frac{\delta}{2a} \quad \text{et} \quad z_2 = -\frac{b}{2a} - \frac{\delta}{2a}$$

et elles sont distinctes.

- Si $\Delta = 0$, alors

$$(I.12) \quad -\frac{b}{2a}$$

est racine double (l'expression *racine double* est définie rigoureusement dans le prochain chapitre sur les polynômes).

Les formules (I.11), (I.12) sont à retenir.

Remarque I.6.2. Dans le cas particulier des coefficients réels, $\Delta \in \mathbb{R}$. Si $\Delta > 0$, il y a deux racines réelles distinctes, si $\Delta < 0$, $\delta = i\sqrt{-\Delta}$ et il y a deux racines complexes conjuguées.

Exemple I.6.3. Résoudre l'équation $z^2 - (2+i)z - 1 + 7i = 0$.

Calcul du discriminant : $\Delta = (2+i)^2 - 4(-1+7i) = 7 - 24i$. On cherche un complexe $\delta = x + iy$ tel que $\delta^2 = \Delta$. On utilise la méthode vue en I.5.c.

$$(x + iy)^2 = 7 - 24i \iff (x^2 - y^2 = 7 \quad \text{et} \quad 2xy = -24)$$

$$\text{En rajoutant l'égalité des modules : } x^2 + y^2 = \sqrt{49 + 576} = \sqrt{625} = 25,$$

on obtient le système

$$\begin{cases} x^2 + y^2 = 25 \\ x^2 - y^2 = 7 \\ xy = -12 \end{cases} \Leftrightarrow \begin{cases} x^2 = 16 \\ y^2 = 9 \\ xy = -12 \end{cases} \Leftrightarrow \begin{cases} x = \pm 4 \\ y = \pm 3 \\ xy = -12. \end{cases}$$

Par conséquent, Δ admet deux racines carrées : $\delta = 4 - 3i$ et $-\delta = -4 + 3i$.

Les solutions de l'équation sont :

$$z_1 = \frac{-b + \delta}{2a} = \frac{2 + i + 4 - 3i}{2} = 3 - i \text{ et } z_2 = \frac{-b - \delta}{2a} = \frac{2 + i - 4 + 3i}{2} = -1 + 2i.$$

I.7. Appendice : quelques rappels

I.7.a. Principe du raisonnement par récurrence

Supposons que l'on veuille prouver qu'une propriété P_n , dépendant d'un entier n , est vraie pour tout $n \geq n_0$.

On procède en 2 étapes :

- *Initialisation.* On montre que la propriété P_{n_0} est vraie.
- *Hérédité.* On montre : si P_n est vraie pour un entier n , alors P_{n+1} est vraie.

Cette méthode est la méthode de récurrence simple, mais parfois, elle est insuffisante et il faut utiliser une récurrence généralisée, dans laquelle la seconde étape est remplacée par :

- *Hérédité.* Pour tout $n \geq n_0$, si la propriété P_k est vraie pour tous les entiers k compris entre n_0 et n , alors P_{n+1} est vraie.

Exercice I.7.1. Démontrer par récurrence (simple) sur n les propriétés suivantes :

$$\forall n \geq 0, \quad 2^n \geq n; \quad \forall n \geq 4, \quad n! \geq 2^n$$

où $n!$ (lire *factorielle n*) est défini par

$$0! = 1, \quad \forall n \geq 0, \quad (n+1)! = (n+1)n!$$

i.e. $n! = 1.2 \dots n$.

I.7.b. Formule du binôme

On rappelle que les coefficients du binôme $\binom{n}{k}$ sont définis par $\binom{n}{0} = \binom{n}{n} = 1$, $\binom{n}{k} = 0$ si $k < 0$ et si $k > n$, et la relation :

$$(I.13) \quad \binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}, \quad 0 \leq k \leq n.$$

Remarque I.7.2. On note parfois C_n^k au lieu de $\binom{n}{k}$.

On calcule les coefficients du binôme à l'aide du triangle de Pascal : on met des 1 sur le bord du triangle, et chaque élément est la somme des deux éléments situés au-dessus de lui. Ainsi

I. Les nombres complexes

					1				
					1	1			
				1	2	1			
			1	3	3	1			
		1	4	6	4	1			
	1	5	10	10	5	1			
1	6	15	20	15	6	1			

le coefficient $\binom{n}{k}$ est le k -ième élément de la n -ième ligne (les lignes sont comptées à partir de 0, de même que les éléments de chaque ligne, par exemple $\binom{6}{2} = 15$).

Si n et k sont des entiers tels que $0 \leq k \leq n$, les coefficients du binôme peuvent aussi être définis par la formule suivante

$$(I.14) \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Exercice I.7.3. Démontrer (I.14) par récurrence généralisée sur $n+k$. On utilisera en particulier la formule (I.13).

On a alors la formule du binôme de Newton sur \mathbb{C} :

Proposition I.7.4. Soient $z, z' \in \mathbb{C}$, $n \in \mathbb{N}$,

$$(z + z')^n = \sum_{k=0}^n \binom{n}{k} z^k (z')^{n-k}.$$

La preuve se fait par récurrence sur n , comme dans le cas où z et z' sont réels. Elle est laissée au lecteur.

I.7.c. Suites arithmétiques et géométriques

Une suite *arithmétique* $(u_n)_{n \in \mathbb{N}}$ est une suite définie par une relation de récurrence de la forme :

$$(I.15) \quad u_{n+1} = u_n + a, \quad n \in \mathbb{N}$$

où a est un nombre (réel ou plus généralement complexe) fixé, appelé *raison* de la suite (u_n) . Par une récurrence élémentaire (à faire!), on peut montrer

$$u_n = u_0 + n a, \quad n \in \mathbb{N}.$$

La somme des termes d'une telle suite est donnée par la formule :

$$(I.16) \quad \sum_{n=p}^q u_n = (q-p+1) \frac{u_p + u_q}{2}, \quad 0 \leq p \leq q.$$

On peut retenir que la somme des termes d'une suite arithmétique est égale au nombre de termes $q-p+1$ multiplié par la moyenne $\frac{u_p + u_q}{2}$ du premier terme et du dernier terme.

Exercice I.7.5. Démontrer par récurrence (sur $q \geq p$, p étant fixé) la formule (I.16).

Exemple I.7.6.

$$3 + 5 + 7 + \dots + 17 = \sum_{k=1}^8 (2k + 1) = 8 \times \frac{3 + 17}{2} = 80.$$

Exercice I.7.7. Calculer $\sum_{j=20}^{30} (3j + 5)$.

Une suite *géométrique* $(v_n)_{n \in \mathbb{N}}$ est une suite définie par une relation de récurrence de la forme :

$$(I.17) \quad v_{n+1} = bv_n,$$

où le nombre complexe non nul b est encore appelé *raison* de la suite (v_n) . On démontre, à nouveau par récurrence :

$$v_n = b^n v_0, \quad n \in \mathbb{N}.$$

Si $b = 1$, la suite est constante. Lorsque $b \neq 1$, la somme de ses termes est donnée par la formule

$$(I.18) \quad \sum_{n=p}^q v_n = \frac{1 - b^{q-p+1}}{1 - b} \times v_0 b^p, \quad 0 \leq p \leq q,$$

qui se déduit immédiatement de la formule

$$(I.19) \quad \sum_{n=0}^N b^n = \frac{1 - b^{N+1}}{1 - b}.$$

Pour démontrer (I.19) remarquer que cette formule est équivalente à

$$(1 - b) \sum_{n=0}^N b^n = 1 - b^{N+1}.$$

Or

$$(1 - b) \sum_{n=0}^N b^n = \sum_{n=0}^N b^n - \sum_{n=1}^{N+1} b^n = 1 - b^{N+1}.$$

Exercice I.7.8. On considère un échiquier de 64 cases. On pose 1 grain de blé sur la première case, 2 grains sur la deuxième, 4 sur la troisième et ainsi de suite, en doublant le nombre de grains de blé à chaque nouvelle case. Combien y aura-t-il de grains de blé sur l'échiquier une fois les 64 cases remplies ?

Les formules (I.16) et (I.19) sont à retenir.

II. Les polynômes

Dans ce chapitre, la lettre \mathbb{K} désigne \mathbb{R} ou \mathbb{C} . Les éléments de \mathbb{K} sont appelés “nombres” ou “scalaires”.

II.1. Définitions

II.1.a. Polynômes comme suites finies

Un polynôme P sur \mathbb{K} (ou à coefficients dans \mathbb{K}) est la donnée d'une suite $(a_n)_{n \geq 0}$ d'éléments de \mathbb{K} telle qu'il existe un entier $p \geq 0$ avec

$$\forall n \geq p, \quad a_n = 0.$$

Les nombres a_n sont appelés les *coefficients* de P . Le plus grand nombre d tel que $a_d \neq 0$ est appelé *degré* de P et noté $\deg P$ ou $\deg(P)$. Le coefficient a_d correspondant à ce degré est appelé *coefficient dominant* de P . Le polynôme $(a_n)_{n \geq 0}$ tel que $a_n = 0$ pour tout n est appelé le *polynôme nul* et noté 0 . Par convention $\deg 0 = -\infty$.

Exemple II.1.1. La suite $(1, 2, 0, 0, 0, \dots)$ est un polynôme de degré 1 ($a_0 = 1$, $a_1 = 2$, les ... signifient que $a_n = 0$ si $n \geq 2$).

La suite $(2^n)_{n \geq 0}$ n'est pas un polynôme.

II.1.b. Addition

Soit $P = (a_n)_{n \geq 0}$ et $Q = (b_n)_{n \geq 0}$ deux polynômes. La *somme* $P + Q$ de P et Q est par définition la suite $(a_n + b_n)_{n \geq 0}$. C'est aussi un polynôme. Ceci définit une addition sur l'ensemble des polynômes, qui est commutative et associative :

$$P + Q = Q + P, \quad (P + Q) + R = P + (Q + R).$$

Du fait de l'associativité, on peut noter sans ambiguïté $P + Q + R$ la somme de trois polynômes.

Proposition II.1.2. *Soit P et Q deux polynômes à coefficients dans \mathbb{K} .*

$$\deg(P + Q) \leq \max(\deg P, \deg Q).$$

Si $\deg P \neq \deg Q$,

$$\deg(P + Q) = \max(\deg P, \deg Q).$$

Démonstration. Soit $n = \deg P$, $p = \deg Q$. On a donc :

$$P = \sum_{k=0}^n a_k X^k, \quad Q = \sum_{k=0}^p b_k X^k \text{ avec } a_n \neq 0, \quad b_p \neq 0.$$

II. Les polynômes

Supposons d'abord $p = n$. Alors

$$P + Q = \sum_{k=0}^n (a_k + b_k) X^k,$$

ce qui montre que le degré de $P + Q$ est exactement $n = \deg P = \deg Q$ (si $a_n + b_n \neq 0$) ou strictement plus petit que n (si $a_n + b_n = 0$). On a donc bien, dans ce cas, $\deg(P + Q) \leq \max(\deg P, \deg Q)$.

Supposons maintenant $n \neq p$. Pour fixer les idées, on suppose $n > p$. On note $b_k = 0$ si $k > p$. Alors, par définition de l'addition des polynômes,

$$P + Q = \sum_{k=0}^n (a_k + b_k) X^k.$$

Le coefficient de X^n est $a_n + b_n = a_n$. Le polynôme $P + Q$ est donc exactement de degré $n = \max(\deg P, \deg Q)$. Le cas $n < p$ se traite de la même manière. \square

II.1.c. Indéterminée

On s'empresse d'adopter une notation plus commode que la notation $(a_n)_{n \geq 0}$ pour désigner les polynômes. On fixe une lettre, généralement X , appelée *indéterminée*. Soit $P = (a_n)_{n \geq 0}$ un polynôme de degré d . On note ce polynôme :

$$(II.1) \quad P = a_d X^d + a_{d-1} X^{d-1} + \dots + a_1 X + a_0 = \sum_{j=0}^d a_j X^j.$$

Exemple II.1.3. Le polynôme $(1, 2, 0, 0, 0, \dots)$ est noté $2X + 1$. Le polynôme

$$(0, 0, 0, 9, 4, 0, 3, 0, 0, 0, \dots)$$

est noté $3X^6 + 0X^5 + 4X^4 + 9X^3 + 0X^2 + 0X + 0$ ou plus simplement $3X^6 + 4X^4 + 9X^3$.

Dans (II.1), les puissances sont rangées dans l'ordre décroissant. On peut aussi les ranger dans l'ordre croissant. De fait, par la définition et la commutativité de l'addition des polynômes, on a :

$$a_d X^d + a_{d-1} X^{d-1} + \dots + a_1 X + a_0 = a_0 + a_1 X + \dots + a_{d-1} X^{d-1} + a_d X^d.$$

Notation II.1.4. L'ensemble des polynômes à coefficients dans \mathbb{K} et d'indéterminée X est noté $\mathbb{K}[X]$. Un polynôme P est parfois noté $P(X)$ lorsqu'on veut insister sur le fait que la lettre X désigne l'indéterminée.

II.1.d. Multiplication

Soit $P = \sum_{j=0}^d a_j X^j$ et $Q = \sum_{j=0}^{d'} b_j X^j$ deux polynômes, $d = \deg P$, $d' = \deg Q$. Par définition, le produit PQ de P et Q est le polynôme

$$(II.2) \quad PQ = \sum_{j=0}^{d+d'} \left(\sum_{k+\ell=j} a_k b_\ell \right) X^j.$$

Il s'obtient en développant l'expression $\sum_{k=0}^d a_k X^k \sum_{\ell=0}^{d'} b_\ell X^\ell$ et en utilisant les règles de calcul usuelles sur l'addition et la multiplication, et la règle : $X^k X^\ell = X^{k+\ell}$.

Remarque II.1.5. Dans l'expression (II.2), $\sum_{k+\ell=j}$ signifie que la somme porte sur tous les indices (entiers naturels) k et ℓ tels que $k + \ell = j$. Par exemple :

$$\sum_{k+\ell=2} a_k b_\ell = a_0 b_2 + a_1 b_1 + a_2 b_0.$$

Il est très important de maîtriser ce genre de notation.

La formule

$$(II.3) \quad \deg(PQ) = \deg P + \deg Q$$

découle immédiatement de la définition de la multiplication des polynômes. Le coefficient dominant de PQ est $a_d b_d$. La multiplication des polynômes vérifient aussi les règles de calcul usuelles :

Proposition II.1.6. *La multiplication des polynômes est commutative, associative, et distributive par rapport à l'addition : si $P, Q, R \in \mathbb{K}[X]$,*

$$PQ = QP, \quad (PQ)R = P(QR) \text{ et } P(Q + R) = PQ + PR.$$

Démonstration. On ne démontre que l'associativité, la preuve des autres propriétés est laissée au lecteur. Soit $P = \sum_{k=0}^{d_1} a_k X^k$, $Q = \sum_{\ell=0}^{d_2} b_\ell X^\ell$ et $R = \sum_{m=0}^{d_3} c_m X^m$ trois polynômes. On doit vérifier

$$(II.4) \quad (PQ)R = P(QR).$$

En utilisant la définition de la multiplication, on obtient $PQ = \sum_{j=0}^{d_1+d_2} \left(\sum_{k+\ell=j} a_k b_\ell \right) X^j$, puis

$$(PQ)R = \sum_{r=0}^{d_1+d_2+d_3} \sum_{j+m=r} \left(\sum_{k+\ell=j} a_k b_\ell \right) c_m X^r \stackrel{\star}{=} \sum_{r=0}^{d_1+d_2+d_3} \sum_{k+\ell+m=r} a_k b_\ell c_m X^r.$$

De même, $QR = \sum_{s=0}^{d_2+d_3} \sum_{\ell+m=s} b_\ell c_m X^s$ et donc

$$P(QR) = \sum_{r=0}^{d_1+d_2+d_3} \sum_{k+s=r} a_k \left(\sum_{\ell+m=s} b_\ell c_m \right) X^r \stackrel{\star}{=} \sum_{r=0}^{d_1+d_2+d_3} \sum_{k+\ell+m=r} a_k b_\ell c_m X^r.$$

D'où $(PQ)R = P(QR)$. □

Exercice II.1.7. Se convaincre des deux égalités \star ci-dessus. On pourra commencer par écrire explicitement ces égalités $d_1 = d_2 = d_3 = 1$.

En pratique, pour calculer le produit de deux polynômes, on n'utilise pas directement la formule (II.2), mais simplement les règles de calcul usuelles. Par exemple :

$$\begin{aligned} (X^3 + 2X^2 - 1)(X^2 + 7) &= X^3 X^2 + 7X^3 + 2X^2 X^2 + 14X^2 - X^2 - 7 \\ &= X^5 + 7X^3 + 2X^4 + 13X^2 - 7. \end{aligned}$$

II.2. Premières propriétés

II.2.a. Division euclidienne

Théorème II.2.1. Soient A et B deux polynômes de $\mathbb{K}[X]$, avec B non nul. Il existe un unique polynôme Q et un unique polynôme R dans $\mathbb{K}[X]$ tels que :

$$A = QB + R \quad \deg(R) < \deg(B).$$

Définition II.2.2. Lorsque $R = 0$ dans le théorème ci-dessus, c'est à dire lorsqu'il existe Q dans $\mathbb{K}[X]$ tel que $A = QB$, on dit que B *divise* A , que A est *divisible* par B , ou que B est un *diviseur* de A .

Démonstration. Existence.

Lorsque $A = 0$, on peut choisir $Q = 0$ et $R = 0$. Supposons A non nul, et notons $A = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$, $B = b_p X^p + b_{p-1} X^{p-1} + \dots + b_1 X + b_0$, avec $n = \deg A$, $p = \deg B$. On commence par montrer :

Lemme II.2.3. Soit $G \in \mathbb{K}[X]$ tel que $\deg G \geq \deg B$. Il existe $S \in \mathbb{K}[X]$ tel que

$$(II.5) \quad \deg(G - BS) < \deg G.$$

Preuve du lemme. Soit d le degré de G . On note $G = \sum_{k=0}^d g_k X^k$. On pose $S = \frac{g_d}{b_p} X^{d-p}$. Le polynôme SB est de degré $d - p + p = d$, et son coefficient dominant (le coefficient de X^d) est $\frac{g_d}{b_p} \times b_p = g_d$. On en déduit comme annoncé que $G - BS$ est au plus de degré $d - 1$. \square

Pour montrer l'existence de Q et R , on construit deux suites finies $(A_j)_{j=0 \dots J}$ et $(Q_j)_{j=1 \dots J}$ en posant $A_0 = A$ et en définissant les $A_j, Q_j, j \geq 1$ par récurrence de la manière suivante. Soit $j \geq 0$ tel que A_j est connu.

- ou bien $\deg(A_j) < \deg(B)$, on pose $J = j$ et on arrête la construction ;
- ou bien $\deg(A_j) \geq \deg(B)$. On pose alors $Q_{j+1} = S, A_{j+1} = A_j - BQ_{j+1}$, où S est donné par le lemme avec $G = A_j$.

Par (II.5), la suite $\deg(A_j)$ est strictement décroissante. Puisque c'est une suite d'éléments de $\mathbb{N} \cup \{-\infty\}$, la construction précédente doit s'arrêter pour un certain J . On a $\deg(A_J) < \deg(B)$ et $A_j = A_{j+1} + BQ_{j+1}$ pour tout $j \leq J - 1$. On en déduit :

$$A = A_0 = A_1 + BQ_1 = A_2 + BQ_2 + BQ_1 = \dots = A_J + B(Q_1 + \dots + Q_J).$$

On a bien obtenu une division euclidienne de A par B , de quotient $Q = Q_1 + \dots + Q_J$ et de reste A_J .

Unicité.

Supposons que $A = QB + R = SB + T$, où les degrés de R et T sont strictement inférieurs à celui de B . On en tire :

$$(II.6) \quad R - T = (S - Q)B \quad \text{et} \quad \deg(R - T) < \deg(B).$$

On montre $Q = S$ par l'absurde. Si $Q \neq S$, alors $\deg(S - Q) \geq 0$. Donc

$$\deg((S - Q)B) = \deg(S - Q) + \deg(B) \geq \deg(B),$$

ce qui contredit (II.6).

Donc $Q = S$, d'où $R = T$ par (II.6). \square

En pratique, pour calculer le quotient et le reste de la division euclidienne d'un polynôme par un autre, on calcule les suites A_j, Q_j de la démonstration précédente en posant la division euclidienne

Exemple II.2.4. Division de $X^3 + 2X^2 + X + 1$ par $X^2 + 1$

$$\begin{array}{r|l}
 X^3 + 2X^2 + X + 1 & X^2 + 1 \\
 - X^3 & X + 2 \\
 \hline
 2X^2 + X + 1 & \\
 - 2X^2 & \\
 \hline
 X + 1 & \\
 - X & \\
 \hline
 1 &
 \end{array}$$

Le résultat s'écrit : $X^3 + 2X^2 + X + 1 = (X^2 + 1)(X + 2) - 1$.

Quotient : $X + 2$, reste : -1 .

Dans cet exemple, on a $Q_1 = X, A_1 = 2X^2 + 1, Q_2 = 2$ et $A_2 = -1$.

Exercice II.2.5. Effectuer la division euclidienne de A par B dans les cas suivants :

$$A = X^4 - 2X^2 - X + 1, B = X^2 + X;$$

$$A = X^6 + 4X^4 - X^2 + 1, B = X^2 + 1.$$

II.2.b. Fonctions polynomiales

Définition II.2.6. On appelle *fonction polynomiale* sur \mathbb{K} toute fonction de la forme

$$\tilde{P} : \mathbb{K} \longrightarrow \mathbb{K} : x \mapsto a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0,$$

où les coefficients $(a_j)_{j=0\dots n}$ sont dans \mathbb{K} . Le polynôme $P = a_n X^n + a_{n-1} X^{n-1} + \dots + a_2 X^2 + a_1 X + a_0$ est appelé *polynôme associé* à la fonction \tilde{P} .

Remarque II.2.7. Il faut toujours avoir en mémoire la différence entre la fonction polynomiale \tilde{P} (et sa variable x) et le polynôme P (et son indéterminée X); même si, lorsqu'il n'y a pas de confusion possible, on omet d'écrire le symbole $\tilde{}$.

Ainsi, pour tout $a \in \mathbb{K}$, on notera désormais $P(a)$ la valeur prise par la fonction \tilde{P} au point a .

Proposition II.2.8. *Le reste de la division euclidienne d'un polynôme P par $X - \alpha$ est le polynôme constant égal à $P(\alpha)$.*

Démonstration. La division euclidienne de P par $X - \alpha$ s'écrit :

$$P = (X - \alpha)Q + R \text{ avec } \deg(R) < 1.$$

Puisque $\deg(R) < 1$, le polynôme R est une constante c et $P(\alpha) = c$. □

II.2.c. Polynôme dérivé

Définition II.2.9. Soit $P = a_n X^n + \dots + a_1 X + a_0$ un polynôme sur \mathbb{K} . On appelle *polynôme dérivé* de P le polynôme :

$$P' = n a_n X^{n-1} + (n-1) a_{n-1} X^{n-2} + \dots + 2 a_2 X + a_1 = \sum_{i=0}^{n-1} (i+1) a_{i+1} X^i.$$

On note $P'', P''', P^{(4)}, \dots, P^{(k)}$ la suite des polynômes dérivés successifs. On pose enfin $P^{(0)} = P$.

II. Les polynômes

Exemple II.2.10. Soit $P = 5 + iX^2 + 4X^3$. Alors $P' = 2iX + 12X^2$, $P'' = 2i + 24X$ etc...

Proposition II.2.11. Soient $P, Q \in \mathbb{K}[X]$. Alors

$$(II.7) \quad \deg P \geq 1 \implies \deg(P') = \deg(P) - 1$$

$$(II.8) \quad P' = 0 \iff P \text{ est constant}$$

$$(II.9) \quad (\lambda P + \mu Q)' = \lambda P' + \mu Q', \quad P, Q \in \mathbb{K}[X], \lambda, \mu \in \mathbb{K}$$

$$(II.10) \quad (PQ)' = P'Q + PQ'$$

Démonstration. La propriété (II.7) se déduit immédiatement des définitions du degré et du polynôme dérivé.

Preuve de (II.8). Soit $P = \sum_{k=0}^n a_k X^k$ un polynôme de degré n . Alors

$$P' = 0 \iff \sum_{k=1}^n k a_k X^{k-1} = 0 \iff \forall k \in \{1, \dots, n\}, a_k = 0.$$

En particulier, si $n \geq 1$, $a_n = 0$, une contradiction. Donc $n = 0$ ou $n = -\infty$, ce qui signifie exactement que P est constant.

La preuve de (II.9) est directe et laissée au lecteur.

Preuve de (II.10). On commence par prouver (II.10) lorsque $P = X^n$, $n \geq 0$. On note

$$Q = \sum_{k=0}^p b_k X^k, \quad p = \deg(Q).$$

Alors

$$P' = nX^{n-1}, \quad Q' = \sum_{k=0}^p k b_k X^{k-1}, \quad PQ = \sum_{k=0}^p b_k X^{k+n}.$$

et donc

$$P'Q + PQ' = \sum_{k=0}^p n b_k X^{k+n-1} + \sum_{k=0}^p k b_k X^{k+n-1} = \sum_{k=0}^p (k+n) b_k X^{k+n-1} = (PQ)'$$

On démontre maintenant le cas général. On commence par remarquer que si $N \geq 1$, P_1, \dots, P_N sont des polynômes et $\lambda_1, \dots, \lambda_N$ des scalaires, alors (II.9) et une récurrence élémentaire nous donnent :

$$(II.11) \quad \left(\sum_{k=1}^N \lambda_k P_k \right)' = \sum_{k=1}^N \lambda_k P_k'$$

Supposons P non nul (sinon le résultat est évident). On note $P = \sum_{k=0}^n a_k X^k$ avec $n = \deg(P)$. Par (II.11),

$$\begin{aligned} (PQ)' &= \sum_{k=0}^n a_k (X^k Q)' \\ &= \sum_{k=0}^n a_k \left[(X^k)' Q + X^k Q' \right] = \sum_{k=0}^n a_k k X^{k-1} Q + \sum_{k=0}^n a_k X^k Q' = P'Q + PQ', \end{aligned}$$

ce qui termine la preuve. □

II.3. Racines

II.3.a. Cas général

Définition II.3.1. Soit $P \in \mathbb{K}[X]$. On dit que $\alpha \in \mathbb{K}$ est une *racine* (ou un *zéro*) de P lorsque $P(\alpha) = 0$.

Théorème II.3.2. Soit $P \in \mathbb{K}[X]$. Un élément α de \mathbb{K} est racine de P si, et seulement si, P est divisible par $X - \alpha$.

Démonstration. D'après la proposition II.2.8 $P = Q(X - \alpha) + P(\alpha)$. Par conséquent P est divisible par $X - \alpha$ si, et seulement si, $P(\alpha) = 0$. \square

Exercice II.3.3. Soit P le polynôme sur \mathbb{R} défini par $P = X^3 - X^2 - 3X + 3$.

- Déterminer une racine évidente de P .
- En déduire une expression de P sous la forme d'un produit d'un polynôme de degré 1 par un polynôme de degré 2.
- En déduire l'ensemble des racines de P .

Définition II.3.4. Soit $P \in \mathbb{K}[X]$, $\alpha \in \mathbb{K}$ et $r \in \mathbb{N}^*$. On dit que α est une racine *d'ordre* r , ou *de multiplicité* r , de P si $P = (X - \alpha)^r Q$ avec $Q(\alpha) \neq 0$.

D'après le théorème II.3.2, une racine est toujours de multiplicité au moins 1.

Lorsque $r = 1$, on dit que la racine est simple.

Lorsque $r = 2$, on dit que la racine est double.

Exemple II.3.5. Le polynôme $P = 3(X - 1)^2(X - i)(X + i)^3$ a pour racines $1, i$ et $-i$. 1 est une racine double, i est une racine simple, et $-i$ est une racine d'ordre 3.

Un trinôme complexe du second degré de discriminant non nul a deux racines simples. Un trinôme du second degré de discriminant nul a une racine double.

Définition II.3.6. On dit que le polynôme P a exactement n racines *comptées avec leur ordre de multiplicité* (ou *avec multiplicité*) lorsque la somme des multiplicités de ses racines est exactement n .

Exemple II.3.7. Le polynôme P de l'exemple II.3.5 a 3 racines distinctes, mais 6 racines comptées avec leur ordre de multiplicité.

Avertissement II.3.8. Il y a donc deux manières de compter le nombre de racines d'un polynôme. L'expression "le polynôme P a n racines" est ambiguë et ne doit jamais être utilisée sans précision supplémentaire.

Remarque II.3.9. Si P a r racines et Q a s racines (comptées avec leur ordre de multiplicité), alors PQ a $r + s$ racines (comptées avec leur ordre de multiplicité). Cette propriété n'est plus valable lorsque l'on compte les racines distinctes des polynômes.

Exemple II.3.10. Si l'on compte les racines avec multiplicité, le polynôme $P = (X - 1)^2$ a deux racines, le polynôme $Q = (X - 1)$ a 1 racine, et le polynôme PQ a bien $2 + 1$ racines.

Mais P , Q et PQ ont chacun une seule racine distincte.

Théorème II.3.11. Soit $P \in \mathbb{K}[X]$. La racine $\alpha \in \mathbb{K}$ de P est de multiplicité r si et seulement si, pour tout k entre 0 et $r - 1$, $P^{(k)}(\alpha) = 0$ et $P^{(r)}(\alpha) \neq 0$.

II. Les polynômes

Démonstration. Si α est racine de multiplicité r , alors $P = (X - \alpha)^r Q$ avec $Q(\alpha) \neq 0$.

En dérivant, on obtient :

$$P' = r(X - \alpha)^{r-1}Q + (X - \alpha)^r Q' = (X - \alpha)^{r-1} (rQ + (X - \alpha)Q'),$$

de la forme $(X - \alpha)^{r-1}Q_1$ avec $Q_1(\alpha) = rQ(\alpha) \neq 0$. Donc, lorsque $r > 1$, $P'(\alpha) = 0$.

En itérant la dérivation, on obtient :

$$\text{pour tout } k < r, P^{(k)} \text{ est de la forme } (X - \alpha)^{r-k} Q_k \text{ avec } Q_k(\alpha) \neq 0.$$

Donc lorsque $k < r$, $P^{(k)}(\alpha) = 0$. De plus $P^{(r)}$ est de la forme Q_r avec $Q_r(\alpha) \neq 0$ et donc $P^{(r)}(\alpha) \neq 0$.

Réciproquement, supposons $P(\alpha) = P'(\alpha) = \dots = P^{(r-1)}(\alpha) = 0$ et $P^{(r)}(\alpha) \neq 0$. Soit s la multiplicité de α . D'après ce qui précède

$$(II.12) \quad P(\alpha) = \dots = P^{(s-1)}(\alpha) = 0 \text{ et } P^{(s)}(\alpha) \neq 0$$

On montre $s = r$ par l'absurde :

- si $s > r$ alors, on aurait $P^{(r)}(\alpha) = 0$ par (II.12), ce qui est contraire aux hypothèses.
- si $s < r$ alors, on aurait $P^{(s)}(\alpha) = 0$, contredisant (II.12).

Donc $s = r$ et α est de multiplicité r . □

II.3.b. Polynômes à coefficients complexes

Théorème II.3.12 (Théorème de D'Alembert). (*admis*) Dans $\mathbb{C}[X]$, tout polynôme non constant admet au moins une racine.

Corollaire II.3.13. Tout polynôme P , de degré $n \geq 1$, de $\mathbb{C}[X]$ admet exactement n racines complexes (comptées avec leur ordre de multiplicité).

Démonstration. Par récurrence sur n . Dans toute la démonstration, les racines sont comptées avec leur ordre de multiplicité.

- Initialisation : si $n = 1$, le résultat est immédiat.
- Hérité : supposons que tout polynôme de degré $n - 1$ de $\mathbb{C}[X]$ admette exactement $n - 1$ racines complexes.

Si P est un polynôme de degré n , d'après le théorème de D'Alembert, il admet au moins une racine α .

Il existe donc Q , de degré $n-1$, tel que $P = (X - \alpha)Q$. D'après l'hypothèse de récurrence, Q admet $n-1$ racines $\alpha_1, \dots, \alpha_{n-1}$. Par conséquent, P admet les n racines $\alpha, \alpha_1, \dots, \alpha_{n-1}$. □

II.3.c. Polynômes à coefficients réels

Puisque $\mathbb{R} \subset \mathbb{C}$, un polynôme à coefficients réels peut être considéré comme un polynôme à coefficients complexes. Pour un tel polynôme P , si $\alpha \in \mathbb{C}$, on a $\overline{P(\alpha)} = P(\bar{\alpha})$.

La proposition suivante donne les propriétés des racines complexes d'un polynôme réel :

Proposition II.3.14. Soit $P \in \mathbb{R}[X]$. Si $\alpha \in \mathbb{C}$ est racine de P , alors $\bar{\alpha}$ l'est aussi. De plus, α et $\bar{\alpha}$ ont même ordre de multiplicité.

II.4. Complément : polynômes irréductibles

Démonstration. Soit r l'ordre de multiplicité de α . On a donc $P^{(k)}(\alpha) = 0$ pour tout k entre 0 et $r - 1$, et $P^{(r)}(\alpha) \neq 0$. Donc, $P^{(k)}(\bar{\alpha}) = \overline{P^{(k)}(\alpha)} = \bar{0} = 0$ pour tout $k \leq r - 1$ et $P^{(r)}(\bar{\alpha}) = \overline{P^{(r)}(\alpha)} \neq 0$. \square

Corollaire II.3.15. *Tout polynôme P , de degré $n \geq 1$, de $\mathbb{R}[X]$ admet au plus n racines réelles (comptées avec leur ordre de multiplicité).*

Corollaire II.3.16. *Tout polynôme P , de degré impair de $\mathbb{R}[X]$ admet au moins une racine réelle.*

Démonstration. En effet, un nombre complexe α est réel si et seulement si $\alpha = \bar{\alpha}$. Il découle donc de la proposition II.3.14 que les racines complexes non réelles de P peuvent se ranger par paires de même multiplicité. Il y a donc un nombre pair de racines complexes non réelles (comptées avec leur ordre de multiplicité). Puisque, par le corollaire II.3.13, P a exactement $\deg(P)$ racines, le nombre de racines réelles de P est impair, et donc non nul. \square

Exercice II.3.17. Donner une autre démonstration du corollaire II.3.16, en étudiant la fonction polynôme associée à P et en utilisant le théorème des valeurs intermédiaires.

Proposition II.3.18. *Soient f et g deux fonctions polynomiales sur \mathbb{K} , définies par $f(x) = a_n x^n + \dots + a_1 x + a_0$ et $g(x) = b_p x^p + \dots + b_1 x + b_0$, avec $a_n \neq 0$ et $b_p \neq 0$. Si*

$$\forall x \in \mathbb{K}, \quad f(x) = g(x),$$

alors $p = n$ et $a_i = b_i$ pour tout i .

Démonstration. S'il existait i tel que $a_i \neq b_i$, la fonction polynomiale $f - g$ serait de degré $k \geq i$. Elle aurait au plus k racines et ne serait donc pas nulle. Ce qui est contraire à l'hypothèse. \square

Exercice II.3.19.

- Montrer que i est racine double du polynôme $P = X^6 + X^5 + 3X^4 + 2X^3 + 3X^2 + X + 1$.
- Déterminer les réels a et b tels que le polynôme $P = X^5 + aX^4 + bX^3 - bX^2 - aX - 1$ admette 1 comme racine de plus grande multiplicité possible.

II.4. Complément : polynômes irréductibles

II.4.a. Cas général

Définition II.4.1. Un polynôme P non constant qui vérifie la condition :

si P est produit de deux polynômes de $\mathbb{K}[X]$, l'un des deux est constant

est dit *irréductible* dans $\mathbb{K}[X]$.

Par convention, les polynômes constants ne sont pas irréductibles.

Exemple II.4.2. Un polynôme de degré 1 est irréductible.

Théorème II.4.3. *Dans $\mathbb{K}[X]$, tout polynôme P non constant se décompose en produit de polynômes irréductibles*

II. Les polynômes

Démonstration. Par récurrence sur le degré n de P :

Initialisation. Si $n = 1$ alors le polynôme est irréductible.

Hérédité. Supposons que tout polynôme de degré $< n$ soit produit de polynômes irréductibles. Soit P un polynôme de degré n .

- Si P est irréductible, le résultat est obtenu.
- Si P n'est pas irréductible, $P = P_1 P_2$ avec $\deg(P_1) \geq 1$ et $\deg(P_2) \geq 1$.
 $\deg(P_1) + \deg(P_2) = \deg(P) = n$ donc $\deg(P_1) < n$ et $\deg(P_2) < n$. Par hypothèse de récurrence P_1 et P_2 sont tous les deux produits de polynômes irréductibles donc P est aussi produit de polynômes irréductibles.

□

II.4.b. Polynômes irréductibles de $\mathbb{C}[X]$

Théorème II.4.4. *Les polynômes de degré 1 sont les seuls polynômes irréductibles de $\mathbb{C}[X]$.*

Démonstration. D'après le théorème de d'Alembert, tout polynôme P , de degré ≥ 2 , admet au moins une racine $\alpha \in \mathbb{C}$. P est donc divisible par $(X - \alpha)$, et il n'est pas irréductible. □

Corollaire II.4.5. *(Décomposition dans $\mathbb{C}[X]$). Soit P un polynôme de degré $n \geq 1$ de $\mathbb{C}[X]$. Sa décomposition en produit de facteurs irréductibles est de la forme :*

$$P = \lambda (X - \alpha_1)^{r_1} \cdots (X - \alpha_p)^{r_p} \quad \text{avec } r_1 + \dots + r_p = n,$$

où $\alpha_1, \dots, \alpha_p$ sont les racines distinctes de P , r_1, \dots, r_p leurs multiplicités et λ le coefficient dominant de P .

II.4.c. Polynômes irréductibles de $\mathbb{R}[X]$

Théorème II.4.6. *Les seuls polynômes irréductibles de $\mathbb{R}[X]$ sont*

- les polynômes de degré 1
- les polynômes de degré 2, dont le discriminant est strictement négatif

Démonstration. - Si P est de degré 1, il est irréductible.

- Si P est un polynôme de degré 2 avec $\Delta < 0$, il est irréductible. Sinon, il se décomposerait en produit de deux polynômes, chacun de degré 1 : $P = (aX + b)(cX + d)$. Il aurait deux racines (distinctes ou confondues), et son discriminant serait ≥ 0 . Contradiction.
- Si P est un polynôme de degré 2 avec $\Delta \geq 0$, il admet deux racines réelles (distinctes ou confondues) et s'écrit $P = a(X - \alpha_1)(X - \alpha_2)$. Il n'est pas irréductible.
- Si P est un polynôme de degré $n > 2$, d'après le théorème de d'Alembert, il admet au moins une racine $\alpha \in \mathbb{C}$
 - Ou bien $\alpha \in \mathbb{R}$, P est divisible par $(X - \alpha)$, et il n'est pas irréductible.
 - Ou bien $\alpha \notin \mathbb{R}$ alors $\bar{\alpha}$ est aussi racine de P .

$(X - \alpha)(X - \bar{\alpha}) = (X^2 - 2\operatorname{Re}(\alpha)X + \alpha\bar{\alpha})$ est un polynôme à coefficients réels.

On fait la division euclidienne dans $\mathbb{R}[X]$: $P = (X^2 - 2\operatorname{Re}(\alpha)X + \alpha\bar{\alpha})Q + R$ avec $\deg R \leq 1$. Or $R(\alpha) = R(\bar{\alpha}) = 0$ et $\alpha \neq \bar{\alpha}$. Donc $R = 0$.

Par conséquent, P n'est pas irréductible dans $\mathbb{R}[X]$ car divisible par un polynôme de degré 2.

□

Corollaire II.4.7. *Décomposition dans $\mathbb{R}[X]$*

Soit P un polynôme de degré $n \geq 1$ de $\mathbb{R}[X]$. Sa décomposition en produit de facteurs irréductibles est de la forme :

$$P = \lambda (X - \alpha_1)^{r_1} \cdots (X - \alpha_p)^{r_p} (X^2 + \beta_1 X + \gamma_1)^{s_1} \cdots (X^2 + \beta_k X + \gamma_k)^{s_k}$$

avec $r_1 + \cdots + r_p + 2(s_1 + \cdots + s_k) = n$ et $\beta_i^2 - 4\gamma_i < 0$ pour $i = 1, \dots, k$

Exemple II.4.8. $X^3 + X = X(X^2 + 1) = X(X + i)(X - i)$

La décomposition dans $\mathbb{C}[X]$ est $X(X + i)(X - i)$

La décomposition dans $\mathbb{R}[X]$ est $X(X^2 + 1)$.

Exercice II.4.9. Décomposer en produit de facteurs irréductibles dans $\mathbb{R}[X]$ le polynôme $P = X^4 - 2X^3 + 14X^2 - 18X + 45$ sachant qu'il admet $1 + 2i$ comme racine.

Correction. P est divisible par $(X - 1 - 2i)(X - 1 + 2i) = X^2 - 2X + 5$. En effectuant la division euclidienne on obtient $P = (X^2 - 2X + 5)(X^2 + 9)$ qui est bien le produit de deux polynômes irréductibles de $\mathbb{R}[X]$. La décomposition dans $\mathbb{C}[X]$ s'en déduit immédiatement :

$$P = (X - 1 - 2i)(X - 1 + 2i)(X - 3i)(X + 3i).$$

Exercice II.4.10. Décomposer en produit de facteurs irréductibles le polynôme $X^3 + 1$ dans $\mathbb{R}[X]$ puis dans $\mathbb{C}[X]$.

III. Systèmes linéaires

La référence principale pour ce chapitre est le livre de David C. Lay¹.

On appellera *nombre* ou *scalaire* un nombre réel ou complexe. On posera $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$ l'ensemble de ces nombres. Le choix des nombres réels ou complexes est indifférent dans ce chapitre, sauf pour les interprétations géométriques où l'on privilégiera les nombres réels.

III.1. Définitions et premiers exemples

III.1.a. Définitions

Définition III.1.1. Soit $n \geq 1$. On appelle *équation linéaire* à n inconnues x_1, \dots, x_n une équation de la forme

$$(E) \quad \sum_{j=1}^n a_j x_j = b,$$

où a_1, a_2, \dots, a_n et b sont fixés dans \mathbb{K} . Les scalaires a_1, a_2, \dots, a_n sont appelés *coefficients* de l'équation, b est le *second membre*. Lorsque $b = 0$, on dit que l'équation est *homogène*.

Remarque III.1.2. La notation $\sum_{j=1}^n a_j x_j$ dans (E) signifie $a_1 x_1 + a_2 x_2 + \dots + a_n x_n$. Il est impératif de maîtriser ce type de notation.

Définition III.1.3. L'*ensemble des solutions* de (E) est l'ensemble des (x_1, x_2, \dots, x_n) de \mathbb{K}^n tels que $\sum_{j=1}^n a_j x_j = b$. C'est donc un sous-ensemble de \mathbb{K}^n .

Exemple III.1.4.

$$3x_1 + x_2 - 4x_3 = 2$$

est une équation linéaire non-homogène à 3 inconnues.

$$-ix_1 + (2 + i)x_2 - 4x_3 = 0$$

est une équation linéaire (complexe) homogène à 3 inconnues.

$$2x_1^2 + x_2 x_3 - 4x_3^3 = 1$$

n'est pas une équation linéaire.

1. David C. Lay. *Algèbre linéaire : Théorie, exercices et applications*. Troisième édition, 2004

III. Systèmes linéaires

Définition III.1.5. Si $p \geq 1$, on appelle *système linéaire* à p équations et n inconnues un ensemble de p équations linéaires ayant les mêmes n inconnues :

$$(S) \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pn}x_n = b_p \end{cases}$$

Les scalaires a_{ij} , $1 \leq i \leq p$, $1 \leq j \leq n$ sont encore appelés les *coefficients* du système. Il est d'usage d'utiliser le premier indice pour numéroter les lignes et le deuxième indice pour numéroter les colonnes. Le p -uplet $b = (b_1, \dots, b_p)$ est appelé *second membre* du système. Lorsque tous les b_j sont nuls (on dit encore que b est nul), on dit que le système est *homogène*.

L'ensemble des solutions de (S) est le sous-ensemble de \mathbb{K}^n formé des (x_1, \dots, x_n) qui vérifient toutes les équations de (S). On cherche à résoudre le système (S), c'est à dire décrire précisément cet ensemble.

Remarque III.1.6. On peut réécrire le système (S) sous forme abrégée :

$$\forall i = 1 \dots p, \quad \sum_{j=1}^n a_{ij}x_j = b_i.$$

Remarque III.1.7. Lorsque le nombre d'inconnues n est petit, on note souvent x, y, z, t (au lieu de x_1, x_2, \dots) ces inconnues pour alléger les notations.

Donnons quelques exemples simples.

Exemples III.1.8.

$$\begin{cases} 17(x + 2y) = \sqrt{7}z + 3 \\ \frac{x + z}{2} = 12. \end{cases}$$

est un système linéaire non-homogène, à 2 équations et 3 inconnues, que l'on peut écrire sous la forme (S) (avec $x = x_1, y = x_2, z = x_3$) :

$$\begin{cases} 17x + 34y - \sqrt{7}z = 3 \\ \frac{1}{2}x + 0y + \frac{1}{2}z = 12. \end{cases}$$

Ici $a_{11} = 17, a_{12} = 34, a_{13} = -\sqrt{7}, a_{21} = \frac{1}{2}, a_{22} = 0, a_{23} = \frac{1}{2}, b_1 = 3$ et $b_2 = 12$. Le système :

$$\begin{cases} xyz + 7 = 0 \\ x + 2y = 3. \end{cases}$$

n'est pas linéaire (la première équation ne peut pas être mise sous la forme d'une équation linéaire). L'équation :

$$(III.1) \quad (x + y - 3)^2 + (2x + y + 2)^2 = 0$$

n'est pas linéaire. Toutefois, si l'on cherche à résoudre cette équation sur \mathbb{R} , elle est équivalente au système linéaire :

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

Remarquons que dans le cas $\mathbb{K} = \mathbb{C}$, l'équation (III.1) ne peut pas être ramenée à un système linéaire.

Exercice III.1.9. Mettre les systèmes linéaires suivants sous la forme (S). Déterminer p , n , et les paramètres a_{ij} et b_i .

$$(III.2) \quad \begin{cases} 3x + y = 4z + 3 \\ y = z, \end{cases}$$

$$(III.3) \quad x_1 + x_2 = x_2 + x_3 = x_3 + x_4 = 0.$$

Les systèmes linéaires apparaissent dans tous les domaines d'applications des mathématiques (économie, industrie...) Dans les applications, p et n sont souvent très grands, et on ne peut pas résoudre le système "à la main". Il existe évidemment de nombreux logiciels informatiques qui en sont capables. Les buts de ce chapitre sont :

- savoir résoudre "à la main" un système lorsque p et n sont petits ;
- comprendre une méthode de résolution d'un système général, la méthode du pivot de Gauss ;
- en déduire quelques propriétés de la structure de l'ensemble des solutions. Cette structure sera précisée au Chapitre V, à travers la notion d'espace vectoriel.

On commence par donner des exemples de résolutions de systèmes linéaires à une ou deux équations et une ou deux inconnues, puis une notation commode (la notation matricielle) avant de dégager une méthode générale.

III.1.b. Exemples de petits systèmes linéaires

Une équation à une inconnue

On fixe $(a, b) \in \mathbb{K}^2$. On considère l'équation :

$$(III.4) \quad ax = b.$$

Alors :

- Si $a \neq 0$, $x = \frac{b}{a}$ est la seule solution de (III.4).
- Si $a = 0$ et $b = 0$, l'équation (III.4) s'écrit $0 = 0$ et tous les $x \in \mathbb{K}$ sont solutions.
- Si $a = 0$ et $b \neq 0$, l'équation (III.4) s'écrit $b = 0$, il n'y a donc pas de solution.

Remarque III.1.10. L'ensemble des solutions est ou bien vide, ou bien réduit à un seul élément, ou bien infini. Nous verrons plus tard (proposition III.2.27 p. 47) que cette propriété persiste dans le cas d'un système linéaire général.

Une équation, deux inconnues

On considère maintenant l'équation linéaire :

$$(III.5) \quad ax + by = c.$$

Supposons d'abord $(a, b) \neq (0, 0)$. Si $\mathbb{K} = \mathbb{R}$, on sait que (III.5) est l'équation d'une droite du plan \mathbb{R}^2 et il y a une infinité de solutions. Si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$, on peut résoudre le système algébriquement :

III. Systèmes linéaires

– si $b \neq 0$, on peut réécrire (III.5) $y = \frac{c}{b} - \frac{a}{b}x$, et l'ensemble des solutions est

$$\left\{ \left(x, \frac{c}{b} - \frac{a}{b}x \right), x \in \mathbb{K} \right\}.$$

On dit qu'on a paramétré cet ensemble (x est le paramètre).

– si $b = 0$ et $a \neq 0$, l'équation s'écrit $x = c/a$ et l'ensemble des solutions est donné par $\left\{ \left(\frac{c}{a}, y \right), y \in \mathbb{R} \right\}$.

Lorsque $(a, b) \neq (0, 0)$ il y a donc une infinité de solutions.

Si $(a, b) = (0, 0)$, l'équation (III.5) s'écrit simplement $0 = c$: il y a une infinité de solutions (tous les couples $(x, y) \in \mathbb{K}^2$) si $c = 0$, et aucune solution si $c \neq 0$.

Remarque III.1.11. Dans ce cas, quelles que soient les valeurs de a, b et c , l'équation (III.5) a ou bien une infinité de solutions, ou bien pas de solution du tout (mais jamais un nombre fini, non nul de solutions : comparer avec la remarque III.1.10).

Deux équation, deux inconnues. Opérations sur les lignes

On considère maintenant un système de la forme :

$$(III.6) \quad a_{11}x + a_{12}y = b_1$$

$$(III.7) \quad a_{21}x + a_{22}y = b_2.$$

Lorsque $\mathbb{K} = \mathbb{R}$, $(a_{11}, a_{12}) \neq (0, 0)$ et $(a_{21}, a_{22}) \neq (0, 0)$, (III.6) décrit l'ensemble des points d'intersection de deux droites \mathcal{D}_1 et \mathcal{D}_2 du plan \mathbb{R}^2 . On a donc trois cas :

- Les droites \mathcal{D}_1 et \mathcal{D}_2 ne sont pas parallèles : elles ont un unique point d'intersection, et (III.6) n'a qu'une seule solution.
- Les droites \mathcal{D}_1 et \mathcal{D}_2 sont parallèles. Le système (III.6) n'a pas de solution, sauf si \mathcal{D}_1 et \mathcal{D}_2 sont confondues, auquel cas le système a une infinité de solutions (l'ensemble des coordonnées des points de la droite $\mathcal{D}_1 = \mathcal{D}_2$).

On donne maintenant trois exemples que l'on va résoudre algébriquement, sans utiliser d'argument géométrique.

Exemple III.1.12.

$$(III.8) \quad \begin{cases} 3x + 2y = 8 & (L_1) \\ 2x + 5y = 9 & (L_2). \end{cases}$$

Éliminons l'inconnue " x " de la deuxième ligne à l'aide de la première ligne :

$$(III.9) \quad \begin{cases} 3x + 2y = 8 \\ 0 + \left(5 - \frac{4}{3} \right) y = \left(9 - \frac{16}{3} \right) & (L_2) - \frac{2}{3}(L_1). \end{cases}$$

La notation $(L_2) - \frac{2}{3}(L_1)$ à la deuxième ligne (on écrira parfois $(L_2) \leftarrow (L_2) - \frac{2}{3}(L_1)$) signifie que l'on a remplacé la deuxième ligne par la différence de la deuxième ligne et du produit de $\frac{2}{3}$ par la première ligne.

$$(III.10) \quad \begin{cases} 3x + 2y = 8 \\ y = 1 & \frac{3}{11}(L_2). \end{cases}$$

Une fois connue la valeur de y il suffit de la substituer dans (L_1) pour obtenir la valeur de x : on obtient $3x = 8 - 2 = 6$, i.e. $x = 2$. Le système (III.8) a donc pour unique solution $(2, 1)$ (en d'autres termes, l'ensemble des solutions est le singleton $\{(2, 1)\}$).

On peut aussi conclure, à partir de (III.10), en utilisant des opérations sur les lignes :

$$(III.11) \quad \begin{cases} 3x = 6 \\ y = 1 \end{cases} \quad (L_1) - 2(L_2) \quad \text{puis} \quad \begin{cases} x = 2 \\ y = 1 \end{cases} \quad \frac{1}{3}(L_1)$$

Exemple III.1.13.

$$(III.12) \quad \begin{cases} 3x + 2y = 8 \\ 6x + 4y = 20. \end{cases}$$

On élimine x comme dans l'exemple III.1.12

$$(III.13) \quad \begin{cases} 3x + 2y = 8 \\ 0 = 4 \end{cases} \quad (L_2) - 2(L_1).$$

Il n'y a pas de solution.

Exemple III.1.14.

$$(III.14) \quad \begin{cases} 3x + 2y = 8 \\ 6x + 4y = 16. \end{cases}$$

On utilise la même stratégie :

$$(III.15) \quad \begin{cases} 3x + 2y = 8 \\ 0 = 0 \end{cases} \quad (L_2) - 2(L_1).$$

La deuxième équation est toujours vraie. Le système (III.14) est donc équivalent à l'équation, $3x + 2y = 8$. Il y a, comme pour l'équation (III.5) lorsque $(a, b) \neq (0, 0)$, toute une droite de solutions, l'ensemble

$$\left\{ \left(x, 4 - \frac{3}{2}x \right), x \in \mathbb{K} \right\}.$$

III.1.c. Notation matricielle

Une *matrice* $p \times n$ est un tableau de nombre à p lignes et n colonnes. Nous étudierons les matrices de manière plus systématique dans le chapitre IV de ce cours.

Lors des manipulations sur les lignes des exemples précédents, on peut gagner du temps en décidant de ne pas noter les variables :

Définition III.1.15. On appelle *matrice des coefficients* du système (S) la matrice $p \times n$:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pn} \end{bmatrix}$$

III. Systèmes linéaires

et *matrice augmentée* la matrice $p \times (n + 1)$:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ & \vdots & & \vdots & \\ a_{p1} & a_{p2} & \dots & a_{pn} & b_p \end{array} \right]$$

(le trait vertical, facultatif, permet de séparer les coefficients du système du second membre).

On peut reprendre les opérations précédentes en notation matricielle. Par exemple, les manipulations sur les lignes de l'exemple III.1.12 s'écrivent :

$$\begin{array}{ccc} \left[\begin{array}{cc|c} 3 & 2 & 8 \\ 2 & 5 & 9 \end{array} \right] & \left[\begin{array}{cc|c} 3 & 2 & 8 \\ 0 & 5 - \frac{4}{3} & 9 - \frac{16}{3} \end{array} \right] & (L_2) - \frac{2}{3}(L_1) \\ \left[\begin{array}{cc|c} 3 & 2 & 8 \\ 0 & 1 & 1 \end{array} \right] & \left[\begin{array}{cc|c} 3 & 0 & 6 \\ 0 & 1 & 1 \end{array} \right] & (L_1) - 2(L_2) \quad \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \end{array} \right] & \frac{1}{3}(L_1) \\ & \frac{3}{11}(L_2) & & \end{array}$$

ce qui signifie bien $x = 2$, $y = 1$ (cf (III.11)).

Exercice III.1.16. Résoudre le système

$$(III.16) \quad \begin{cases} 2x + y = 5 \\ x - y = 1 \end{cases}$$

en utilisant la notation matricielle.

III.2. Méthode du pivot

On formalise et généralise maintenant la méthode précédente pour résoudre des systèmes linéaires quelconques. Plus précisément :

- on définit en III.2.a des opérations sur les lignes des systèmes linéaires qui ne changent pas l'ensemble des solutions : les *opérations élémentaires*. Ce sont exactement les opérations apparaissant dans les exemples précédents ;
- on introduit en III.2.b une famille de systèmes pour lesquels il est très facile de décrire l'ensemble des solutions, les systèmes *sous forme échelonnée réduite* ;
- on montre en III.2.c comment ramener, par une série d'opérations élémentaires, tout système linéaire, à un système sous forme échelonnée réduite : c'est la méthode du pivot proprement dite.

Cette partie se termine par l'étude d'une classe de système importante, les systèmes de Cramer (III.2.d).

III.2.a. Systèmes équivalents. Opérations élémentaires

Définition III.2.1. Deux systèmes linéaires ayant le même nombre d'inconnues sont *équivalents* lorsqu'ils ont le même ensemble de solutions.

Exemple III.2.2. Les systèmes suivants sont équivalents :

$$(III.17) \quad \begin{cases} x + y = 0 \\ y = -1 \end{cases} \quad \text{et} \quad \begin{cases} x + y = 0 \\ x = 1 \end{cases}$$

En effet, l'ensemble des solutions de ces deux systèmes est $\{(1, -1)\}$. Les deux systèmes suivants ne sont pas équivalents :

$$(III.18) \quad \begin{cases} x + y = 0 \\ y = -1 \end{cases} \quad \text{et} \quad \begin{cases} x + y = 0 \\ x = -1 \end{cases}$$

En effet, l'ensemble des solutions du premier système est $\{(1, -1)\}$, l'ensemble des solutions du deuxième système est $\{(-1, 1)\}$.

Définition III.2.3. On appelle *opération élémentaire* sur un système (ou sur les lignes d'une matrice) l'une des trois opérations suivantes :

- i. L'échange de deux lignes (L_i) et (L_j) , parfois noté $(L_i) \leftrightarrow (L_j)$.
- ii. La multiplication d'une ligne par un scalaire $k \in \mathbb{K}$ non nul, appelé *cadrage*, parfois noté $(L_i) \leftarrow k(L_i)$.
- iii. L'ajout à une ligne du multiple d'une autre ligne par un scalaire k , appelé *remplacement*, parfois noté $(L_i) \leftarrow (L_i) + k(L_j)$.

Le lecteur est invité à vérifier que toutes les opérations réalisées dans les exemples de la partie III.1 sont des opérations élémentaires au sens de la définition III.2.3. Par exemple, l'opération conduisant à (III.9) est un remplacement, celle conduisant à (III.10) un cadrage.

Remarque III.2.4. Il revient bien sûr au même de faire des opérations élémentaires sur un système d'équations linéaires, ou sur les lignes de la matrice augmentée de ce système.

Les opérations élémentaires ne changent pas l'ensemble des solutions :

Proposition III.2.5. Soient (S) et (S') deux systèmes linéaires ayant le même nombre d'inconnues et d'équations. On suppose que (S') peut être obtenu à partir de (S) par une série d'opérations élémentaires. Alors (S) et (S') sont équivalents.

Démonstration. Par une récurrence simple, il suffit de montrer qu'une seule opération élémentaire ne change pas l'ensemble des solutions.

C'est évident si cette opération élémentaire est l'échange de deux lignes.

Si k est fixé, non nul, alors $ky = 0 \iff y = 0$. Le cadrage (multiplication d'une ligne par un scalaire non nul) ne change donc pas l'ensemble des solutions.

Il reste à traiter le cas du remplacement, i.e. l'opération $(L_j) \leftarrow (L_j) + k(L_i)$, où $i \neq j$, $k \in \mathbb{K}$. En ignorant les lignes autres que la i -ième et la j -ième, qui ne changent pas, on est amené à montrer que les deux systèmes

$$\begin{cases} a_{i1}x_1 + \dots + a_{in}x_n = b_i & (L_i) \\ a_{j1}x_1 + \dots + a_{jn}x_n = b_j & (L_j) \end{cases}$$

et

$$\begin{cases} a_{i1}x_1 + \dots + a_{in}x_n = b_i & (L_i) \\ a_{j1}x_1 + \dots + a_{jn}x_n + k(a_{i1}x_1 + \dots + a_{in}x_n) = b_j + kb_i & (L'_j) \end{cases}$$

III. Systèmes linéaires

sont équivalents, c'est à dire que $x = (x_1, \dots, x_n)$ vérifie (L_i) et (L_j) si et seulement si il vérifie (L_i) et (L'_j) .

On suppose d'abord que x vérifie (L_i) et (L_j) . Alors

$$\underbrace{a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n}_{=b_j \text{ par } (L_j)} + k \underbrace{(a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n)}_{=b_i \text{ par } (L_i)} = b_j + kb_i,$$

ce qui montre (L'_j) .

Supposons réciproquement que x vérifie (L_i) et (L'_j) . Alors

$$\begin{aligned} a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n &= \underbrace{a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n + k(a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n)}_{=b_j + kb_i \text{ par } (L'_j)} \\ &\quad - k \underbrace{(a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n)}_{=b_i \text{ par } (L_i)} = b_j, \end{aligned}$$

d'où (L_j) . □

Avertissement III.2.6. Pour passer d'un système linéaire à un système équivalent, il est très fortement conseillé de s'en tenir à des opérations élémentaires au sens de la définition (III.2.3). Des opérations sur les lignes mal choisies peuvent changer l'ensemble des solutions. Une erreur typique est de réaliser *simultanément* deux remplacements de la forme $(L_i) \leftarrow (L_i) + k(L_j)$ et $(L_j) \leftarrow (L_j) + k'(L_i)$. Par exemple :

$$(III.19) \quad \begin{cases} x + 2y = 2 \\ x - y = 0 \end{cases} \text{ donne : } \begin{cases} 3x = 2 & (L_1) + 2(L_2) \\ \frac{3}{2}x = 1 & (L_2) + \frac{1}{2}(L_1). \end{cases}$$

Les deux systèmes ci-dessus ne sont pas équivalents : le premier a pour unique solution $(\frac{2}{3}, \frac{2}{3})$, le deuxième a une infinité de solutions : $\{(\frac{2}{3}, y), y \in \mathbb{K}\}$. Remarquons que les deux opérations simultanées $(L_2) \leftarrow (L_2) + \frac{1}{2}(L_1)$ et $(L_1) \leftarrow (L_1) + 2(L_2)$ ne peuvent pas s'écrire comme deux remplacements successifs. Le lecteur est invité à méditer cet exemple : pratiquement toutes les erreurs dans les résolutions de système, en dehors des fautes de calcul, sont de cette forme là.

III.2.b. Forme échelonnée

On définit maintenant un ensemble de matrice correspondant à des systèmes dont on peut décrire l'ensemble des solutions sans aucun calcul supplémentaire. Le but de la méthode du pivot sera de se ramener à ce type de matrice.

Définitions

Définition III.2.7. Soit A une matrice $p \times n$. Une *ligne nulle* de A est une ligne de A formée uniquement de zéros. On appelle *élément de tête* d'une ligne non nulle de A l'élément non nul le plus à gauche de cette ligne. On dit que A est *sous forme échelonnée* (ou simplement *échelonnée*) lorsque les deux propriétés suivantes sont vérifiées :

- i. Toutes les lignes non nulles sont situées au-dessus des lignes nulles.

- ii. L'élément de tête de chaque ligne non nulle se trouve dans une colonne (strictement) à droite de l'élément de tête de la ligne précédente.

On dit que A est *sous forme échelonnée réduite* (ou simple *échelonnée réduite*) quand de plus les deux propriétés suivantes sont vérifiées :

- iii. L'élément de tête de chaque ligne non nulle vaut 1.
 iv. L'élément de tête de chaque ligne non nulle est le seul coefficient non nul de sa colonne.

Remarque III.2.8. Le point ii implique que tous les éléments de la matrice situés sous un élément de tête sont nuls.

Définition III.2.9. On dit qu'un système linéaire est *sous forme échelonnée* (respectivement *échelonnée réduite*) quand sa matrice augmentée est sous forme échelonnée (respectivement échelonnée réduite).

Exemples III.2.10. La matrice $A = \begin{bmatrix} 0 & 0 \\ 2 & 3 \end{bmatrix}$ n'est pas échelonnée (i n'est pas vérifiée).

La matrice $B = \begin{bmatrix} 3 & 4 & 5 \\ 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix}$ n'est pas échelonnée (cette fois, ii est faux).

La matrice $C = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}$ est échelonnée, mais pas échelonnée réduite (iii et iv sont faux).

La matrice $D = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}$ est échelonnée, mais pas échelonnée réduite (iii est faux).

La matrice $E = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ est échelonnée réduite.

Le système $\begin{cases} x + 2z = 1 \\ y + 3z = 2 \end{cases}$ est sous forme échelonnée réduite.

Exemples III.2.11. Le premier système de (III.11) est sous forme échelonnée non réduite, le dernier système de (III.11) est sous forme échelonnée réduite.

Le système (III.13) est sous forme échelonnée non réduite.

Le système (III.15) est sous forme échelonnée non réduite. Après multiplication de la première ligne par $\frac{1}{3}$, on obtient la forme échelonnée réduite équivalente :

$$\begin{cases} x + \frac{2}{3}y = \frac{8}{3} \\ 0 = 0. \end{cases}$$

Dans ce système la deuxième ligne $0 = 0$ est bien sûr superflue.

Exercice III.2.12. Déterminer si les systèmes suivants sont sous forme échelonnée (respectivement sous forme échelonnée réduite) :

- i. Un système à une seule équation.

ii.

$$\begin{cases} x_2 + x_3 = 2 \\ x_1 + x_3 = 3 \end{cases}, \text{ puis } \begin{cases} x_1 + x_3 = 2 \\ x_2 + x_3 = 3 \end{cases}$$

iii.

$$\left[\begin{array}{cccc|c} 2 & 1+2i & 0 & 4i & 1 \\ 0 & 0 & -1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right], \quad \left[\begin{array}{cccc|c} 0 & 1 & 0 & 1 & 2 \\ 1 & 0 & 0 & 4 & 3 \\ 0 & 0 & 1 & 2 & 3 \end{array} \right], \quad \left[\begin{array}{cccccc|c} 0 & 0 & 1 & 0 & 2 & 3 & 1 \\ 0 & 0 & 0 & 1 & 0 & -3 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

III. Systèmes linéaires

iv.

$$(S_1) \begin{cases} x = 1 \\ y = 2 \\ z = 3 \\ 0 = 0 \end{cases}, \quad (S_2) \begin{cases} 2x + y + z = 1 \\ e^{i\pi/3}y + 2z = -1 \\ z = -4 \\ 0 = 0 \end{cases}, \quad (S_3) \begin{cases} 2x + y + z = 1 \\ x + y + 2z = -1 \\ x - z = -4 \end{cases}$$

(cf correction p. 51).

Application aux systèmes linéaires

Définition III.2.13. Le système (S) est dit *compatible* si il a au moins une solution et *incompatible* si il n'a pas de solution.

On peut décrire facilement l'ensemble des solutions d'un système linéaire (S) dont la matrice augmentée A est sous forme échelonnée réduite. On rappelle que si (S) est un système à p équations et n inconnue, la matrice A est une matrice $p \times (n + 1)$ (i.e. un tableau de nombres avec p lignes et $n + 1$ colonnes).

On distingue deux cas.

- *Premier cas.* Si la colonne la plus à droite (la $n + 1$ -ème colonne) de A contient un élément de tête (forcément égal à 1) cela se traduit par une équation $0 = 1$, tautologiquement fausse, ce qui montre que le système n'a pas de solution. Le système (S) est incompatible.
- *Deuxième cas.* Supposons que la colonne de droite de A ne contient aucun élément de tête. Dans ce cas, le système est compatible. Donnons une méthode pour décrire l'ensemble des solutions. L'élément de tête de chaque ligne non nulle, situé sur une des n premières colonnes, correspond donc à une des inconnues x_1, \dots, x_n . On appelle *variable de base* du système toute inconnue x_j telle que la j -ième colonne contient un élément de tête non nul. On appelle *variables libres* ou *paramètres* les autres inconnues. Chaque ligne non nulle donne une expression d'une des variables de base en fonction des paramètres. En faisant varier les paramètres dans \mathbb{K} , on obtient exactement l'ensemble des solutions du système. On dit que l'on a obtenu une *description paramétrique* de l'ensemble des solutions. Dans ce cas, le nombre p' de lignes non nulles est exactement le nombre de variable de bases. Le nombre de paramètres est donc $n - p'$.

On peut maintenant donner une définition plus précise de la résolution d'un système linéaire : résoudre le système (S), c'est donner une description paramétrique de l'ensemble des solutions.

Remarque III.2.14. Un cas particulier de système compatible est donné par un système dont la matrice sous forme échelonnée réduite a autant de lignes non nulles qu'il y a d'inconnues. Dans ce cas, toutes les variables sont des variables de base, et il n'y a qu'une seule solution, dont les coordonnées x_1, x_2, \dots, x_n sont données par les valeurs de la colonne de droite.

Remarque III.2.15. Il n'est pas nécessaire de mettre un système sous forme échelonnée réduite pour savoir si il est compatible ou non : une forme échelonnée non réduite convient tout aussi bien. Par le même raisonnement que précédemment, un système sous forme échelonnée est compatible si et seulement si la colonne de droite de sa matrice augmentée ne contient aucun élément de tête.

Lorsque le système est compatible, la forme échelonnée réduite est commode pour décrire l'ensemble des solutions.

Exemple III.2.16. On suppose que la matrice augmentée du système (S) est

$$\left[\begin{array}{cccc|c} \mathbf{1} & 0 & 3 & 0 & 4 \\ 0 & \mathbf{1} & 2 & 0 & 3 \\ 0 & 0 & 0 & \mathbf{1} & 5 \end{array} \right]$$

C'est une matrice de forme échelonnée réduite, dont les éléments de tête sont les "1" en gras. La colonne de droite ne contient aucun élément de tête : le système est compatible. Les variables de base (correspondant aux numéros des colonnes des éléments de tête) sont x_1 , x_2 et x_4 . Le seul paramètre est x_3 . On obtient la description paramétrique suivante de l'ensemble des solutions :

$$x_1 = 4 - 3x_3, \quad x_2 = 3 - 2x_3, \quad x_4 = 5, \quad x_3 \in \mathbb{K}.$$

En d'autres termes, l'ensemble des solutions de (S) est :

$$\{(4 - 3x_3, 3 - 2x_3, x_3, 5), x_3 \in \mathbb{K}\}.$$

Exemple III.2.17. Supposons que le système a pour matrice augmentée

$$\left[\begin{array}{ccc|c} 1 & 3 & 4 & 5 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & \mathbf{2} \end{array} \right]$$

Cette matrice est sous forme échelonnée (non réduite). Le 2 en bas à droite est un élément de tête situé sur la dernière colonne : le système n'a pas de solution, car la dernière ligne se lit $0 = 2$.

Exemple III.2.18. Supposons que le système a pour matrice augmentée :

$$\left[\begin{array}{cccc|c} \mathbf{1} & 2 & 0 & -1 & 4 \\ 0 & 0 & \mathbf{1} & 2 & -5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

Cette matrice est sous forme échelonnée réduite, et le système est compatible. Les variables de base sont x_1 et x_3 , et les variables libres x_2 et x_4 . L'ensemble des solutions est

$$\left\{ (4 - 2x_2 + x_4, x_2, -5 - 2x_4, x_4), (x_2, x_4) \in \mathbb{K}^2 \right\}.$$

Exercice III.2.19. Dire si les systèmes de l'exercice III.2.12, lorsqu'ils sont sous forme échelonnée réduite, sont compatibles. Donner alors une description paramétrique de l'ensemble des solutions.

(cf correction p. 52).

La proposition suivante permet de compter le nombre de paramètres d'un système sous forme échelonnée réduite. Elle découle des arguments précédents.

Proposition III.2.20. *Un système compatible sous forme échelonnée réduite avec n inconnues et p' équations non nulles se décrit avec $n - p'$ paramètres. En d'autres termes, si le système est sous forme échelonnée réduite, on a :*

$$\text{nombre d'inconnues} - \text{nombre d'équations} = \text{nombre de degrés de liberté}.$$

Cette propriété persiste lorsque le système est sous forme échelonnée non réduite (cf Remarque III.2.28).

III. Systèmes linéaires

Exercice III.2.21. Dire si chacune des matrices suivantes est sous forme échelonnée (respectivement sous forme échelonnée réduite). Le système dont c'est la matrice augmentée est-il compatible? Si c'est le cas, donner une description paramétrique de l'ensemble des solutions.

$$\text{a) } \left[\begin{array}{cccc|c} 1 & 2 & 3 & -1 & 2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right], \quad \text{b) } \left[\begin{array}{cccc|c} 1 & 2 & 3 & -4 & 4 \\ 0 & 1 & 1 & 2 & -5 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{array} \right], \quad \text{c) } \left[\begin{array}{cccc|c} 1 & 0 & 0 & -3 & 1 \\ 0 & 1 & 0 & 3 & 2 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

(cf correction p. 52).

III.2.c. Méthode du pivot de Gauss

Nous venons de voir qu'un système linéaire sous forme échelonnée réduite peut être résolu très facilement. Nous allons maintenant montrer :

Théorème III.2.22. *Soit (S) un système linéaire. Alors (S) peut être transformé, par des opérations élémentaires sur les lignes, en un système sous forme échelonnée réduite.*

Il est en fait plus commode de travailler en notation matricielle, et de prouver le théorème équivalent :

Théorème III.2.23. *Soit A une matrice. Alors A peut être transformée, par des opérations élémentaires sur les lignes, en une matrice échelonnée réduite.*

Pour montrer le théorème III.2.23, on utilise une méthode appelée *méthode du pivot* ou *du pivot de Gauss*, qui suit exactement la stratégie ébauchée en III.1 pour la résolution des systèmes à deux équations, deux inconnues. Le nom de cette méthode est un hommage au mathématicien allemand Carl Friedrich Gauss (1777-1855), mais elle était déjà connue des mathématiciens chinois du 1er siècle de notre ère. D'après Wikipedia "Elle est référencée dans l'important livre chinois *Jiuzhang suanshu* ou *Les Neuf Chapitres sur l'art mathématique*, dont elle constitue le huitième chapitre, sous le titre *Fang cheng* (la disposition rectangulaire)".

La méthode du pivot permet de démontrer les théorèmes III.2.22 et III.2.23, mais elle donne aussi et surtout une méthode générale de résolution des systèmes linéaires. Elle doit donc être parfaitement comprise.

Soit A une matrice (p, N) (p lignes, N colonnes). Pour montrer le théorème III.2.23, on veut transformer A en une matrice échelonnée réduite par une série d'opérations élémentaires sur les lignes. La méthode est divisée en une phase de descente (permettant d'obtenir une matrice échelonnée qui n'est pas forcément réduite) et une phase de remontée. Pour illustrer cette méthode, on l'applique à la matrice

$$A = \left[\begin{array}{ccc|c} 1 & 2 & 4 & 7 \\ 2 & 4 & 8 & 14 \\ 1 & 0 & -2 & 3 \end{array} \right].$$

On décrit la méthode pour une matrice générale. Dans le cas d'un système linéaire à n inconnues et p équations, A est la matrice augmentée du système, de taille $p \times (n + 1)$ (et donc $N = n + 1$).

Phase de descente

Cette phase permet de transformer une matrice quelconque en une matrice échelonnée. Elle est divisée en 4 étapes, que l'on doit éventuellement réaliser plusieurs fois.

Étape 1 : choix du pivot. On appelle *colonne pivot* la première colonne non nulle². On choisit sur cette colonne un élément non nul, appelé pivot. Dans notre exemple, la colonne pivot est la première colonne. On peut choisir comme élément pivot le 1 en haut à gauche (en gras).

$$A = \left[\begin{array}{ccc|c} \mathbf{1} & 2 & 4 & 7 \\ 2 & 4 & 8 & 14 \\ 1 & 0 & -2 & 3 \end{array} \right].$$

Étape 2. On échange la première ligne avec la ligne de l'élément pivot. Le pivot devient ainsi l'élément de tête de la première ligne. Dans notre exemple, l'élément pivot est déjà situé sur la première ligne : cette étape ne modifie pas la matrice.

Étape 3. En ajoutant aux autres lignes un multiple adéquat de la première ligne, on annule tous les coefficients de la colonne pivot autre que le pivot. Dans notre exemple, cela donne

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & 7 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & -6 & -4 \end{array} \right] \begin{array}{l} \\ (L_2) - 2(L_1) \\ (L_3) - (L_1) \end{array}.$$

Étape 4. Si la matrice obtenue est sous forme échelonnée, la phase de descente est terminée. Sinon, on applique les étapes 1 à 4 à la matrice à laquelle on a enlevé la première ligne.

Revenons à notre exemple. La matrice A a 3 lignes. On applique les étapes 1 à 4 à la matrice $A' = \left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & -2 & -6 & -4 \end{array} \right]$ obtenue à partir de A en enlevant la première ligne. En pratique, on continue à écrire cette première ligne, que l'on ignore.

Étape 1' : On considère donc la matrice $\left[\begin{array}{ccc|c} 1 & 2 & 4 & 7 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & -6 & -4 \end{array} \right]$. La colonne pivot (première colonne non nulle lorsqu'on ignore la première ligne) est la deuxième colonne. On doit choisir comme pivot le -2 , situé à la troisième ligne de cette colonne.

Étape 2' : on échange la 2ème et la 3ème ligne, la matrice obtenue est

$$(III.20) \quad \left[\begin{array}{ccc|c} 1 & 2 & 4 & 7 \\ 0 & -2 & -6 & -4 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

Étape 3' : les coefficients de la colonne pivot (la deuxième colonne) autre que le pivot sont déjà nuls, il n'y a donc rien à faire. Rappelons que l'on ignore la première ligne. Il n'y a donc que le coefficient situé à la troisième ligne de cette deuxième colonne à considérer. Ce coefficient est bien égal à zéro.

Étape 4' : la matrice obtenue est échelonnée : on arrête ici la phase de descente.

2. c'est à dire la colonne la plus à gauche qui n'est pas composée uniquement de zéros

III. Systèmes linéaires

Si A est une matrice échelonnée $p \times N$, et (a_0, \dots, a_N) sont $N + 1$ scalaires tels que $a_0 \neq 0$, la matrice

$$B = \begin{bmatrix} a_0 & a_1 & \dots & a_N \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{bmatrix},$$

(ou tout autre matrice obtenue à partir de B en rajoutant des colonnes de zéros à sa gauche) est une matrice échelonnée. De cette remarque, et du fait que toute matrice ayant une seule ligne est échelonnée, on déduit que la phase de descente de la méthode du pivot aboutit bien, après un certain nombre³ de passages par les étapes 1 à 4, à une matrice échelonnée.

Pour obtenir une matrice échelonnée réduite, on a besoin de deux étapes supplémentaires, qui constituent la phase de remontée.

Phase de remontée

Etape 5 : cadrage. On multiplie chaque ligne non nulle par l'inverse de son élément de tête, de telle manière que l'élément de tête de la nouvelle ligne vaut 1. Dans notre exemple :

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & 7 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right] \quad -\frac{1}{2}(L_2)$$

Etape 6 : on ajoute à chaque ligne un multiple de la dernière ligne non nulle, pour que la colonne au-dessus de l'élément de tête de la dernière ligne ne soit composée que de zéros. On répète cette opération avec l'avant-dernière ligne, etc, jusqu'à la deuxième ligne. Sur notre exemple :

$$\left[\begin{array}{ccc|c} 1 & 0 & -2 & 3 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right] \quad (L_1) - 2(L_2)$$

Par construction, la matrice obtenue à l'issue de la phase de remontée est bien une matrice échelonnée réduite : on a transformé en 1 tous les éléments de tête, et en 0 tous les coefficients situés au dessus d'un élément de tête.

Ceci termine la démonstration du théorème III.2.23 (et son analogue sur les systèmes, le théorème III.2.22) et la description de la méthode du pivot.

Mentionnons que l'on peut en fait démontrer l'unicité de la forme échelonnée réduite :

Théorème III.2.24. *Soit A et B deux matrices échelonnées réduites. Supposons que B puisse être obtenue à partir de A par des opérations élémentaires sur les lignes. Alors $A = B$.*

On omet la démonstration, cf par exemple David C. Lay⁴, annexe A.

Exercice III.2.25. Programmer dans votre langage informatique préféré la méthode du pivot de Gauss. Tester votre programme sur l'exemple précédent.

3. au plus $p - 1$

4. David C. Lay. *Algèbre linéaire : Théorie, exercices et applications*. Troisième édition, 2004

Remarque III.2.26. On déduit de l'exemple donné une description paramétrique de l'ensemble des solutions du système :

$$\begin{cases} x + 2y + 4z = 7 \\ 2x + 4y + 8z = 14 \\ x - 2z = 3 \end{cases}$$

dont la matrice augmentée est la matrice A . Le système est compatible : il n'y a pas d'élément de tête sur la dernière colonne de la matrice échelonnée réduite obtenue. Il y a deux variables de base, x et y , et une variable libre z . L'ensemble des solutions est donné par

$$\{(3 + 2z, 2 - 3z, z), z \in \mathbb{K}\}.$$

Remarquons que la compatibilité du système peut se vérifier à la fin de la phase de descente, sur la forme échelonnée non réduite (III.20).

En combinant la méthode du pivot de Gauss avec la description de l'ensemble des solutions d'un système sous forme échelonnée réduite (cf la proposition III.2.20), on obtient la propriété suivante, qui généralise la remarque III.1.10 à tous les systèmes linéaires :

Proposition III.2.27. *L'ensemble des solutions d'un système linéaire est vide, ou réduit à un seul élément, ou infini.*

Remarque III.2.28. La phase de remontée du pivot de Gauss montre que tout système sous forme échelonnée est équivalent à un système sous forme échelonnée réduite avec le même nombre de lignes non nulles. Si (S) est un système compatible sous forme échelonnée à n inconnues et p' équations non nulles, l'ensemble des solutions se décrit donc avec exactement $n - p'$ paramètres.

Exercice III.2.29. Combien de paramètres faut-il pour décrire chacune des systèmes suivants ?

$$(S_1) \begin{cases} x + y + z = 1 \\ y + 2z = 2 \end{cases} \quad (S_2) \begin{cases} 2x + 3y = 4 \\ y + z = 1 \\ z = 2 \end{cases} \quad (S_3) \begin{cases} x + (1 + i)y + z = 2 \\ x + y = 1 \\ iy + z = 1 \end{cases}$$

(cf correction p. 52).

Remarque III.2.30. Il y a plusieurs variantes (parfaitement équivalentes) de la méthode du pivot que nous venons de décrire. On peut par exemple échanger les étapes 5 et 6. On peut aussi réaliser l'étape de cadrage 5 pendant la phase de descente. Dans tous les cas il est important de n'utiliser que des opérations élémentaires sur les lignes (cf l'avertissement III.2.6). Ceci est particulièrement important lorsque l'on cherche à résoudre des systèmes avec paramètres (cf §III.3 p. 50).

Récapitulons la méthode générale de résolution d'un système linéaire décrite dans ce chapitre :

- Appliquer la phase de descente de la méthode du pivot de Gauss à la matrice augmentée du système. On obtient une matrice échelonnée.
- Déterminer si le système est compatible : si la colonne de droite contient un élément de tête, le système n'est pas compatible (i.e. il n'y a pas de solution). Sinon il est compatible.
- Si le système est compatible, appliquer la phase de remontée du pivot de Gauss. On obtient une matrice échelonnée réduite. On peut alors donner une description paramétrique de l'ensemble des solutions à l'aide de cette matrice échelonnée réduite.

III. Systèmes linéaires

Donnons un autre exemple, cette fois avec des coefficients complexes. On considère le système :

$$(S) \begin{cases} 4z_1 - (3+i)z_2 - (9+3i)z_3 = 5-3i \\ \quad \quad \quad 2z_1 - 2z_2 - 6z_3 = 2-2i \\ 4z_1 - (2+2i)z_2 - (6+6i)z_3 = 6-2i \end{cases}$$

Appliquons la méthode du pivot à sa matrice augmentée :

$$\left[\begin{array}{ccc|c} 4 & -3-i & -9-3i & 5-3i \\ \mathbf{2} & & -6 & 2-2i \\ 4 & -2-2i & -6-6i & 6-2i \end{array} \right]$$

Phase de descente

Etape 1 : choix du pivot. La colonne pivot est la première colonne. Le pivot est le 2 en gras.

Etape 2 : on place la ligne du pivot en première ligne (opération $(L_1) \leftrightarrow (L_2)$)

$$\left[\begin{array}{ccc|c} \mathbf{2} & -2 & -6 & 2-2i \\ 4 & -3-i & -9-3i & 5-3i \\ 4 & -2-2i & -6-6i & 6-2i \end{array} \right]$$

Etape 3 :

$$\left[\begin{array}{ccc|c} 2 & -2 & -6 & 2-2i \\ 0 & 1-i & 3-3i & 1+i \\ 0 & 2-2i & 6-6i & 2+2i \end{array} \right] \begin{array}{l} \\ (L_2)-2(L_1) \\ (L_3)-2(L_1) \end{array}$$

Etape 4 : on repasse à l'étape 1 en ignorant la première ligne.

Etape 1' : la colonne pivot est la deuxième colonne. On choisit le $1-i$ sur la deuxième ligne de cette colonne comme élément pivot.

Etape 2' : la ligne pivot est la deuxième ligne donc la "première" si on ignore la première ligne. Il n'y a rien à faire.

Etape 3' :

$$\left[\begin{array}{ccc|c} 2 & -2 & -6 & 2-2i \\ 0 & 1-i & 3-3i & 1+i \\ 0 & 0 & 0 & 0 \end{array} \right] \begin{array}{l} \\ \\ (L_3)-2(L_2) \end{array}$$

Etape 4' : la matrice obtenue est échelonnée. La phase de descente est terminée. On remarque que la colonne de droite ne contient aucun élément de tête : le système est compatible. On passe à la phase de remontée pour obtenir une matrice échelonnée réduite.

Phase de remontée

Etape 5 :

$$\left[\begin{array}{ccc|c} 1 & -1 & -3 & 1-i \\ 0 & 1 & 3 & i \\ 0 & 0 & 0 & 0 \end{array} \right] \begin{array}{l} \frac{1}{2}(L_1) \\ \frac{1}{1-i}(L_2) \\ \end{array}$$

Etape 6 : on utilise la ligne (L_2) pour annuler l'élément de la deuxième colonne au-dessus de l'élément de tête :

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 3 & i \\ 0 & 0 & 0 & 0 \end{array} \right] \begin{array}{l} \\ (L_1)+(L_2) \\ \end{array}$$

La matrice obtenue est sous forme échelonnée réduite. Il y a une variable libre, z_3 , et deux variables de base, z_1 et z_2 . Une description paramétrique de l'ensemble des solutions est donnée par :

$$\{(1, i - 3z, z), z \in \mathbb{C}\}.$$

Exercice III.2.31. Résoudre le système

$$\begin{cases} 2x - 2y - 4z = -12 \\ -x + y + 3z = 9 \\ x - y + z = 3. \end{cases}$$

III.2.d. Système de Cramer

Les coefficients (a_{ij}) d'un système linéaire (S) étant fixés, la compatibilité de (S) dépend en général de son second membre (b_i) . On présente ici un cas particulier important, appelé *système de Cramer* pour lequel le système est compatible quelques soient les b_i .

Proposition III.2.32. *Soit (S) un système linéaire de n équations à n inconnues. Supposons que (S) a une et une seule solution. Alors tout système obtenu à partir de (S) en changeant seulement le second membre a une seule solution.*

Définition III.2.33. Un système (S) vérifiant les hypothèses de la proposition III.2.32 est appelé *système de Cramer*.

Remarque III.2.34. Le fait d'être un système de Cramer ne dépend pas du second membre du système, mais seulement de ses coefficients.

Preuve de la proposition III.2.32. Par des opérations élémentaires sur les lignes, on peut, d'après la méthode du pivot de Gauss, se ramener à un système sous forme échelonnée réduite (S') ayant le même ensemble de solutions que (S). Soit p' le nombre de lignes non nulles de ce système. Le système étant compatible, le nombre de paramètres permettant de décrire l'ensemble des solutions est, d'après la proposition III.2.20, $n - p'$. Mais on ne peut pas avoir $n - p' \geq 1$, sinon l'ensemble des solutions serait infini. On a donc $n = p'$: la forme échelonnée réduite du système a n lignes non nulles, aucune ligne nulle, et s'écrit donc :

$$(III.21) \quad \begin{cases} x_1 = b'_1 \\ x_2 = b'_2 \\ \vdots \\ x_n = b'_n. \end{cases}$$

Lorsque l'on change le membre de droite du système (S) sans toucher au membre de gauche, on ne change que le membre de droite du système (III.21), puisque ce dernier est obtenu par des opérations élémentaires sur les lignes qui ne mélangent jamais le membre de gauche et le membre de droite des équations. Ceci montre que tout système obtenu à partir de (S) en ne changeant que le membre de droite n'a qu'une seule solution. \square

Remarque III.2.35. De manière équivalente, on pourrait définir un système de Cramer comme un système à n équations et n inconnues qui a une forme échelonnée réduite du type (III.21).

III. Systèmes linéaires

Exercice III.2.36. Dire avec le moins de calculs possibles si les systèmes suivants ont une unique solution, pas de solution ou une infinité de solutions. On identifiera en particulier les systèmes homogènes et les systèmes de Cramer. Les systèmes (S_1) , (S_2) et (S_3) ont pour inconnues x, y et z . Les systèmes (S_4) et (S_5) x, y, z et t .

$$(S_1) \begin{cases} x + 3y + z = 2 \\ x + y + 2z = -5 \\ x + 3y + 2z = 2 \end{cases}, \quad (S_2) \begin{cases} x + 3y + z = 4 \\ x + y + 2z = 1 \\ x + 3y + 2z = 3 \end{cases}, \quad (S_3) \begin{cases} x + 3y + z = -17 \\ x + y + 2z = 2 \\ x + 3y + 2z = 5 \end{cases}$$

$$(S_4) \begin{cases} x + 2y + z + t = 0 \\ x + y - z + 5t = 0 \\ 2x + y + 2z + 4t = 0 \end{cases}, \quad (S_5) \quad x = 2y + 2t = 3z + 4(x + y) = 6x + y + z + t.$$

(cf correction p. 52).

III.3. Système avec paramètres

On considère parfois une famille de systèmes (S_λ) dépendant d'un paramètre λ . Le but est de résoudre le système selon les valeurs de λ . Il faut traiter ce type de problème avec précaution. Une erreur classique est de diviser une équation par une quantité, dépendant de λ , qui s'annule pour une certaine valeur de λ . On donne ici deux exemples de résolutions détaillées.

Exercice III.3.1. Résoudre, selon la valeur de $\lambda \in \mathbb{R}$, le système suivant :

$$(III.22) \quad \begin{cases} -x + (2\lambda - 6)y - 2z = -7 \\ -x + (4\lambda - 12)y - 2z = -11 \\ x + (3 - \lambda)y + 2z = 5 \end{cases}$$

Correction. On applique la méthode du pivot de Gauss.

$$\begin{cases} x + (3 - \lambda)y + 2z = 5 & (L_3) \\ -x + (4\lambda - 12)y - 2z = -11 \\ -x + (2\lambda - 6)y - 2z = -7 & (L_1) \end{cases}$$

$$\begin{cases} x + (3 - \lambda)y + 2z = 5 \\ (-9 + 3\lambda)y = -6 & (L_2) + (L_1) \\ (-3 + \lambda)y = -2 & (L_3) + (L_1) \end{cases}$$

On remarque que les lignes (L_2) et (L_3) du dernier système sont équivalentes. Précisément, (L_2) vaut exactement $3(L_3)$. Le système (III.22) est donc équivalent à

$$(III.23) \quad \begin{cases} x + (3 - \lambda)y + 2z = 5 \\ (-3 + \lambda)y = -2 \end{cases}$$

Ce système est sous forme échelonnée. En regardant le coefficient de y dans la deuxième ligne, on voit qu'il faut distinguer deux cas.

1er cas : $\lambda = 3$. La deuxième ligne de (III.23) s'écrit $0 = -2$. Le système n'a pas de solution.

2ème cas : $\lambda \neq 3$ On obtient :

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

En prenant x et y comme variable de base et z comme paramètre, on obtient que l'ensemble des solutions est $\{(3 - 2z, \frac{2}{3-\lambda}, z), z \in \mathbb{R}\}$.

Exercice III.3.2. Résoudre, selon la valeur de $\lambda \in \mathbb{R}$, le système de matrice augmentée :

$$(III.24) \quad \left[\begin{array}{cc|c} 1 & -(1 + \lambda) & \lambda - 4 \\ -2 & 2 + 2\lambda & 12 - 4\lambda \\ 2 & -2\lambda & -4 + 2\lambda \end{array} \right]$$

Correction.

$$\left[\begin{array}{cc|c} 1 & -(1 + \lambda) & \lambda - 4 \\ 0 & 0 & 4 - 2\lambda \\ 0 & 2 & 4 \end{array} \right] \begin{array}{l} \\ (L_2)+2(L_1) \\ (L_3)-2(L_1) \end{array}$$

Le système obtenu est presque sous forme échelonnée (il suffirait d'échanger les lignes 2 et 3). La ligne 2 se lit $4 - 2\lambda = 0$. On distingue deux cas :

1er cas : $\lambda \neq 2$. La ligne 2 est contradictoire. Le système n'est pas compatible.

2ème cas : $\lambda = 2$. La ligne 2 se lit $0 = 0$. En notant x et y les variables, le système est équivalent à :

$$x - 3y = -2 \text{ et } 2y = 4,$$

soit $(x, y) = (2, 4)$.

Remarquons que l'on peut parfois ramener un système non-linéaire à un système linéaire avec paramètre :

Exercice III.3.3. En utilisant les exercices précédents, résoudre les systèmes

$$(S_1) \begin{cases} -x_1 + (2x_2 - 6)x_3 - 2x_4 = -7 \\ -x_1 + (4x_2 - 12)x_3 - 2x_4 = -11 \\ x_1 + (3 - x_2)x_3 + 2x_4 = 5 \end{cases} \quad (S_2) \begin{cases} x - (1 + z)y = z - 4 \\ -2x + (2 + 2z)y = 12 - 4z \\ 2x - 2zy = -4 + 2z \end{cases}$$

(cf correction p. 53).

III.4. Réponse à certains exercices

Exercice III.2.12

- i. Une matrice à une ligne est toujours sous forme échelonnée. Elle est sous forme échelonnée réduite si et seulement si son premier coefficient non nul vaut 1. L'équation (ou le "système" à une équation) correspondant est sous forme échelonnée réduite si et seulement si le coefficient de la première variable qui apparaît dans le système vaut 1.
- ii. La matrice du premier système est $\left[\begin{array}{ccc|c} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 3 \end{array} \right]$, celle du deuxième système est $\left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 3 \end{array} \right]$. Le premier n'est pas sous forme échelonnée. Le deuxième est sous forme échelonnée (non réduite). On peut donc transformer, en échangeant l'ordre des variables, un système non-échelonné en un système échelonné.

III. Systèmes linéaires

- iii. échelonnée non réduite / non échelonnée / échelonnée réduite.
- iv. échelonné réduit / échelonné non réduit / non échelonné.

Exercice III.2.19

- ii. L'ensemble des solutions du deuxième système est donné par : $\{(2 - x_3, 3 - x_3, x_3), x_3 \in \mathbb{K}\}$.
- iii. L'ensemble des solutions du système dont la matrice augmentée est la troisième matrice est donné par $\{(x_1, x_2, 1 - 2x_5 - 3x_6, -2 + 3x_6, x_5, x_6), (x_1, x_2, x_5, x_6) \in \mathbb{K}^4\}$.
- iv. Le système (S_1) a évidemment pour unique solution $(1, 2, 3)$.

Exercice III.2.21

La matrice a) est sous forme échelonnée. La dernière colonne ne contient pas d'élément de tête : le système est compatible. Pour décrire l'ensemble des solutions, on peut facilement mettre la matrice sous forme échelonnée réduite (par le remplacement $(L_1) \leftarrow (L_1) + 2(L_2)$). On obtient la matrice (en omettant la dernière ligne, inutile, qui signifie $0 = 0$) :

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 2 \end{array} \right].$$

Une description paramétrique de l'ensemble des solutions est donnée par :

$$\{(4 - 2x_2 - 3x_3, x_2, x_3, 2), (x_2, x_3) \in \mathbb{K}^2\}.$$

La matrice b) n'est pas sous forme échelonnée. La dernière ligne de cette matrice signifie $3 = 0$: le système n'est pas compatible.

La matrice c) est échelonnée réduite. L'ensemble des solutions est :

$$\{(1 + 3x_4, 2 - 3x_4, -4x_4, x_4), x_4 \in \mathbb{K}\}.$$

Exercice III.2.29

Les systèmes (S_1) et (S_2) sont sous forme échelonnée et compatibles (il est facile de trouver une solution du premier système en fixant z , le deuxième se résout en calculant successivement $z = 2$, $y = -1$ et $x = 7/2$). Le système (S_1) a 3 inconnues et 2 lignes non nulles, on décrit l'ensemble des solutions avec $3 - 2 = 1$ paramètre. Le système (S_2) a 3 inconnues et 3 équations : on décrit l'ensemble des solutions avec $3 - 3 = 0$ paramètre : il a une unique solution (ce que l'on savait déjà).

Le système (S_3) n'est pas sous forme échelonnée. De fait, la ligne (L_1) est la somme des lignes (L_2) et (L_3) . Il est équivalent au système (S'_3) obtenu en retirant la ligne (L_3) au système (S_3) . Ce système (S'_3) est échelonné, on peut décrire l'ensemble des solutions de (S_3) avec $3 - 2 = 1$ paramètre.

Exercice III.2.31 : il y a une infinité de solutions. L'ensemble des solutions peut s'écrire

$$\{(x, x, 3), x \in \mathbb{K}\}.$$

Exercice III.2.36

Par les remplacements $(L_2) \leftarrow (L_2) - (L_1)$ et $(L_3) \leftarrow (L_3) - (L_1)$, on voit que le système (S_1) est équivalent au système

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

Ce dernier système est un système échelonné compatible à trois équations non nulles pour trois inconnues : c'est donc un système de Cramer, qui a une unique solution. De plus, par la Proposition III.2.32 sur les systèmes de Cramer, tout système obtenu à partir de (S_1) en ne changeant que le second membre sont aussi des systèmes de Cramer : on en déduit que (S_2) et (S_3) sont des systèmes de Cramer (et ont donc chacun une et une seule solution).

Le système (S_4) est un système homogène avec 3 équations et 4 inconnues. Il a donc une infinité de solutions. Le système (S_5) est équivalent à :

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

C'est donc aussi un système homogène à 3 équations et 4 inconnues : il a une infinité de solutions.

Exercice III.3.3

Les deux systèmes proposés *ne sont pas* des systèmes linéaires, mais se ramènent à des systèmes linéaires en fixant une des variables.

Si l'on fixe x_2 , le système (S_1) est un système linéaire d'inconnues (x_1, x_3, x_4) . Plus précisément, c'est exactement le système (III.22) avec $x = x_1$, $y = x_3$, $z = x_4$ et le paramètre $\lambda = x_2$. L'ensemble des solutions est donné par la résolution du système (III.22) :

$$\left\{ \left(3 - 2x_4, x_2, \frac{2}{3 - x_2}, x_4 \right), x_2 \in \mathbb{R} \setminus \{3\}, x_4 \in \mathbb{R} \right\}.$$

Si l'on fixe z , le système (S_2) est un système linéaire. Plus précisément, c'est le système d'inconnue (x, y) , dont la matrice augmentée est donnée par (III.24) avec $z = \lambda$. On déduit de la résolution de ce système que l'unique solution du système (S_2) est $(x, y, z) = (4, 2, 2)$.

IV. Introduction aux matrices

Référence : Liret-Martinais¹, chapitre 4.

Nous avons déjà rencontré des tableaux de nombres, ou matrices. Nous allons étudier ici ces matrices de manière plus systématique, en définissant notamment des opérations (additions, multiplications. . .) sur des ensembles de matrices. Les motivations de ce chapitre sont d'une part de mieux comprendre les systèmes linéaires, qui peuvent être vus comme des équations matricielles, d'autre part d'introduire des notions et des méthodes utiles dans l'étude de l'algèbre linéaire proprement dite qui sera l'objet de la suite du cours.

Comme dans le chapitre précédent, on notera $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

IV.1. Définitions. Opérations sur les matrices

IV.1.a. Définitions

Définition IV.1.1. Soient p et n deux entiers ≥ 1 . Une matrice $p \times n$ à coefficients dans \mathbb{K} est un tableau de nombres (c.à.d. d'éléments de \mathbb{K}) à p lignes et n colonnes, que l'on note :

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pn} \end{bmatrix},$$

ou $A = [a_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$, ou encore, quand il n'y a pas d'ambiguïté sur p et n , $A = [a_{ij}]$. Les $a_{ij} \in \mathbb{K}$ sont appelés *coefficients* (ou *éléments*) de A . Le coefficient a_{ij} est situé à la i -ème ligne et j -ième colonne (le premier indice indique toujours la ligne, le deuxième la colonne).

L'ensemble des matrices $p \times n$ est noté $\mathcal{M}_{p,n}(\mathbb{K})$, ou plus simplement $\mathcal{M}_{p,n}$. Lorsque $p = n$, on dit que la matrice est *carrée*. On note simplement $\mathcal{M}_n(\mathbb{K})$ (ou \mathcal{M}_n) au lieu de $\mathcal{M}_{n,n}(\mathbb{K})$ l'ensemble des matrices carrées de taille n . On dit que deux matrices sont *de même taille*, ou *de même dimension* lorsqu'elles ont le même nombre de lignes et le même nombre de colonnes.

Exemples IV.1.2. La *matrice nulle* de $\mathcal{M}_{p,n}(\mathbb{K})$ est la matrice dont tous les coefficients sont nuls. Elle est notée 0 .

$\begin{bmatrix} 1 & i & -5 \\ 3 & 4 & 7+i \end{bmatrix}$ est une matrice complexe 2×3 (*attention ici i est le nombre complexe tel que $i^2 = -1$, ce n'est pas l'indice des lignes!*)

$\begin{bmatrix} -1 & 0 \\ 7 & 3,1 \\ 2 & \sqrt{3} \end{bmatrix}$ est une matrice réelle 3×2 .

1. François Liret et Dominique Martinais. *Algèbre 1re année - Cours et exercices avec solutions*. Dunod, deuxième édition, 2003

IV. Introduction aux matrices

Définition IV.1.3. On dit que deux matrices *de même taille* $A = [a_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ et $B = [b_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ sont égales lorsque tous leur coefficients sont égaux, i.e. $\forall i \in \{1, \dots, p\}, \forall j \in \{1, \dots, n\}, a_{i,j} = b_{i,j}$. On note alors $A = B$.

Exemples IV.1.4. Les matrices $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ et $\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$ ne sont pas égales.

Les matrices $[2i + j]_{\substack{1 \leq i \leq 3 \\ 1 \leq j \leq 2}}$ et $\begin{bmatrix} 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix}$ sont égales.

Définition IV.1.5. Une *matrice colonne* (ou un vecteur colonne) à p coefficients est un élément de $\mathcal{M}_{p,1}(\mathbb{K})$.

Une *matrice ligne* (ou un vecteur ligne) à n coefficients est un éléments de $\mathcal{M}_{1,n}(\mathbb{K})$.

La j -ième colonne de la matrice $[a_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ est la matrice colonne $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{pj} \end{bmatrix}$, sa i -ème ligne la matrice ligne $[a_{i1} \ a_{i2} \ \dots \ a_{in}]$.

Notation IV.1.6. On identifie $\mathcal{M}_{p,1}(\mathbb{K})$ à \mathbb{K}^p en identifiant le vecteur colonne $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$ au p -uplet (x_1, x_2, \dots, x_p) .

IV.1.b. Multiplication par un scalaire et additions

On définit maintenant deux opérations qui se font *coefficient par coefficient*.

Définition IV.1.7. Soit $\lambda \in \mathbb{K}$ et $A \in \mathcal{M}_{p,n}(\mathbb{K})$. Le produit de A par λ , noté λA , est la matrice de $\mathcal{M}_{p,n}(\mathbb{K})$ obtenue en multipliant chaque coefficient de A par λ : si A est la matrice $[a_{ij}]$, λA est la matrice $[\lambda a_{ij}]$.

Définition IV.1.8. Soient $A, B \in \mathcal{M}_{p,n}(\mathbb{K})$ deux matrices. La somme de A et B , notée $A + B$ est la matrice de $\mathcal{M}_{p,n}$ obtenue en sommant les coefficients de A et de B deux à deux : si $A = [a_{ij}]$ et $B = [b_{ij}]$, $A + B$ est la matrice $[a_{ij} + b_{ij}]$.

Avertissement IV.1.9. La somme de deux matrices n'est définie que si ces matrices sont *de même taille*.

Exemples IV.1.10.

$$17 \begin{bmatrix} 0 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -34 \\ 0 & 17 \end{bmatrix}, \quad (3+i) \begin{bmatrix} 1 & 3-i \\ i & -1 \\ 0 & -i \end{bmatrix} = \begin{bmatrix} 3+i & 10 \\ 3i-1 & -3-i \\ 0 & 1-3i \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} 4 & 7 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 0 & 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} 4 & 7 \\ 2 & -2 \\ 3 & 0 \end{bmatrix} \text{ n'est pas définie.}$$

$$[i+j]_{\substack{1 \leq i \leq 3 \\ 1 \leq j \leq 2}} + [2i-j]_{\substack{1 \leq i \leq 3 \\ 1 \leq j \leq 2}} = [3i]_{\substack{1 \leq i \leq 3 \\ 1 \leq j \leq 2}}.$$

Les propriétés suivantes découlent immédiatement des propriétés (commutativité, associativité, distributivité) de l'addition et de la multiplication sur \mathbb{K} .

Proposition IV.1.11. Soient $A, B \in \mathcal{M}_{p,n}(\mathbb{K})$, $\lambda, \mu \in \mathbb{K}$.

- i. (commutativité) $A + B = B + A$.
- ii. (associativité) $(A + B) + C = A + (B + C)$ (on note leur valeur commune $A + B + C$).
- iii. $\lambda(A + B) = \lambda A + \lambda B$.
- iv. $\lambda(\mu A) = (\lambda\mu)A$ (on note la valeur commune $\lambda\mu A$).
- v. $0 + A = A$.
- vi. $A + (-1)A = 0$.

Exercice IV.1.12. Démontrer les propriétés de la proposition.

Notation IV.1.13. On note $-A$ la matrice $(-1)A$ et $A - B$ la somme $A + (-B)$, appelée *différence* de A et B .

IV.1.c. Transposition

Définition IV.1.14. La transposée de la matrice $A = [a_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}} \in \mathcal{M}_{p,n}(\mathbb{K})$ est la matrice $[a_{ji}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in \mathcal{M}_{n,p}(\mathbb{K})$. On la note tA . Les coefficients de la i -ème ligne de tA sont ceux de la i -ème colonne de A , et inversement, les coefficients de la j -ème colonne de tA sont ceux de la j -ème ligne de A .

On rencontre aussi, en particulier dans les ouvrages en anglais, la notation A^T au lieu de tA .

Avertissement IV.1.15. Lorsqu'on transpose une matrice, on inverse le nombre de lignes et le nombre de colonnes. Par exemple, la transposée d'une matrice ligne est une matrice colonne et la transposée d'une matrice colonne est une matrice ligne. La transposée d'une matrice 2×4 est une matrice 4×2 etc...

Exemples IV.1.16.

$${}^t \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 0 \\ -2 & 4 & 5 \end{bmatrix}, \quad {}^t \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = [2 \ 3 \ 4] \quad {}^t \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}.$$

Le dernier exemple de matrice A est une matrice carrée qui vérifie $A = {}^tA$: on dit que A est *symétrique*.

On déduit immédiatement de la définition de la transposée :

Proposition IV.1.17. Soient $A, B \in \mathcal{M}_{p,n}(\mathbb{K})$, $\lambda \in \mathbb{K}$.

$${}^t({}^tA) = A, \quad {}^t(A + B) = {}^tA + {}^tB, \quad {}^t(\lambda A) = \lambda {}^tA.$$

IV.1.d. Multiplication des matrices

La définition de la multiplication est plus délicate que celle des opérations précédentes, et nous allons la diviser en deux étapes.

Multiplication d'une matrice ligne par une matrice colonne

Définition IV.1.18. Soit $A = [a_i]_{1 \leq i \leq n} \in \mathcal{M}_{1,n}(\mathbb{K})$ une matrice ligne et $B = [b_j]_{1 \leq j \leq n} \in \mathcal{M}_{n,1}(\mathbb{K})$ une matrice colonne ayant le même nombre de coefficients. Le produit AB de A et B est le scalaire :

$$AB = \sum_{j=1}^n a_j b_j = a_1 b_1 + a_2 b_2 + \dots + a_n b_n.$$

Avertissement IV.1.19. On ne peut pour l'instant multiplier qu'un vecteur ligne et un vecteur colonne ayant le même nombre de coefficients, et dans cet ordre.

Exemple IV.1.20.

$$[3 \quad 4 \quad 5] \begin{bmatrix} -1 \\ i \\ 2 \end{bmatrix} = -3 + 4i + 10 = 7 + 4i.$$

Cas général

Définition IV.1.21. Soit $A = [a_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}} \in \mathcal{M}_{p,n}(\mathbb{K})$ et $B = [b_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq q}} \in \mathcal{M}_{n,q}(\mathbb{K})$ deux matrices telles que le nombre de colonnes de A est égal au nombre de lignes de B . Le produit de A et B , noté AB , $A.B$ ou $A \times B$ est la matrice de $\mathcal{M}_{p,q}(\mathbb{K})$ dont le coefficient (i, j) est le produit (au sens de la définition IV.1.18) de la i -ème ligne de A par la j -ième colonne de B . Le produit $A \times A$ d'une matrice carrée A est noté A^2 et appelé *carré* de A . On définit de même la puissance n -ième $A^n = A \times \dots \times A$ (n fois) d'une matrice carrée A .

Remarque IV.1.22. On déduit immédiatement de la définition la formule du produit matriciel suivante : si $AB = [c_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}$, alors

$$(IV.1) \quad c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}.$$

Avertissement IV.1.23. La matrice AB n'est définie que lorsque le nombre de colonnes de A est égal au nombre de lignes de B . Le nombre de lignes de AB est le nombre de lignes de A . Le nombre de colonnes de AB est le nombre de colonnes de B .

$$\begin{matrix} \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} & \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} & = & \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \\ \text{(3, 4)} & \text{(4, 2)} & & \text{(3, 2)}. \end{matrix}$$

Exemples IV.1.24. Pour toute matrice A , $0.A = A.0 = 0$. Ici 0 désigne n'importe quelle matrice nulle de dimension appropriée. Attention : les trois 0 ne désignent pas forcément les mêmes matrices nulles !

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \\ 7 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 2 & -1 \end{bmatrix} \text{ n'est pas défini.}$$

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \\ 7 & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 & 1 \\ 2 & -3 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -9 & 0 & 2 \\ 6 & -15 & 0 & 4 \\ -9 & 3 & 0 & 7 \end{bmatrix}.$$

$$\begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \begin{bmatrix} 2 & 1 & -7 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 3 & -21 & 3 \\ 8 & 4 & -28 & 4 \\ 10 & 5 & -35 & 5 \end{bmatrix}$$

En pratique, pour calculer le produit $[c_{ij}]$ de deux matrices A et B , on peut les disposer de telle manière que le coefficient c_{ij} à calculer soit aligné sur la i -ème ligne de A et la j -ième colonne de B :

$$\begin{array}{cc} & \begin{bmatrix} -1 & 0 & 0 & 1 \\ 2 & -3 & 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 2 & 3 \\ 4 & 5 \\ 7 & -1 \end{bmatrix} & \begin{bmatrix} 4 & -9 & 0 & 2 \\ 6 & -15 & 0 & 4 \\ -9 & 3 & 0 & 7 \end{bmatrix} \end{array}.$$

Exercice IV.1.25. On considère les matrices

$$A = \begin{bmatrix} 1 & 2 \\ -3 & i \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1+i & -3 \\ 2 & 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 \\ -3 & -6 \\ -2 & -4 \end{bmatrix}.$$

Donner les valeurs des produits et des carrés de ces matrices (16 opérations potentielles) lorsqu'ils sont définis.

(cf réponses p. 74).

Soit n un entier ≥ 1 . On note I_n la matrice de $\mathcal{M}_n(\mathbb{K})$ dont la diagonale principale est composée de 1 et dont tous les autres coefficients sont nuls :

$$(IV.2) \quad I_n = [\delta_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}, \quad \delta_{ii} = 1, \quad i \neq j \Rightarrow \delta_{ij} = 0.$$

Exemple IV.1.26.

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Proposition IV.1.27. Soient $A \in \mathcal{M}_{p,n}(\mathbb{K})$, alors

$$AI_n = I_p A = A.$$

Exercice IV.1.28. Démontrer la proposition précédente à l'aide de la formule (IV.1).

Définition IV.1.29. La matrice I_n est appelée *matrice identité*.

On donne maintenant les propriétés de base de la multiplication matricielle :

Théorème IV.1.30. Soient A, B et C des matrices.

IV. Introduction aux matrices

i. Associativité : si $A \in \mathcal{M}_{p,n}$, $B \in \mathcal{M}_{n,q}$ et $C \in \mathcal{M}_{q,r}$,

$$(AB)C = A(BC).$$

On note simplement ABC le produit des trois matrices.

ii. Distributivité à gauche : si $A \in \mathcal{M}_{p,n}$ et $B, C \in \mathcal{M}_{n,q}$,

$$A(B + C) = AB + AC.$$

iii. Distributivité à droite : si $A, B \in \mathcal{M}_{p,n}$ et $C \in \mathcal{M}_{n,q}$,

$$(A + B)C = AC + BC.$$

iv. Si $A \in \mathcal{M}_{p,n}$, $B \in \mathcal{M}_{n,q}$ et $\lambda \in \mathbb{K}$,

$$\lambda(AB) = (\lambda A)B = A(\lambda B).$$

v. Si $A \in \mathcal{M}_{p,n}$, $B \in \mathcal{M}_{n,q}$,

$${}^t(AB) = {}^tB {}^tA.$$

Démonstration. On démontre (i). Les preuves (plus simples) des autres propriétés sont laissées au lecteur. Remarquons que les matrices $(AB)C$ et $A(BC)$ sont bien de même taille $p \times r$. On note

$$AB = [d_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}}, \quad (AB)C = [e_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq r}}, \quad BC = [f_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq r}} \text{ et } A(BC) = [g_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq r}}.$$

Notre but est de montrer

$$(IV.3) \quad \forall i \in \{1, \dots, p\}, \forall j \in \{1, \dots, r\}, \quad e_{ij} = g_{ij}.$$

Par la formule du produit matriciel (IV.1),

$$\begin{aligned} d_{ij} &= \sum_{\ell=1}^n a_{i\ell} b_{\ell j}, & e_{ij} &= \sum_{k=1}^q d_{ik} c_{kj} = \sum_{k=1}^q \sum_{\ell=1}^n a_{i\ell} b_{\ell k} c_{kj} \\ f_{ij} &= \sum_{k=1}^q b_{ik} c_{kj}, & g_{ij} &= \sum_{\ell=1}^n a_{i\ell} f_{\ell j} = \sum_{\ell=1}^n a_{i\ell} \sum_{k=1}^q b_{\ell k} c_{kj} = \sum_{k=1}^q \sum_{\ell=1}^n a_{i\ell} b_{\ell k} c_{kj}, \end{aligned}$$

d'où (IV.3). □

Exercice IV.1.31. Refaire le calcul précédent lorsque $p = q = r = n = 2$, sans utiliser le symbole \sum .

Exercice IV.1.32. Démontrer les propriétés (ii), (iii), (iv), (v) du théorème.

Remarque IV.1.33. Soit $A = [a]$ et $B = [b]$ des matrices 1×1 , avec $a \in \mathbb{K}$ et $b \in \mathbb{K}$. Alors $A + B = [a + b]$ et $AB = ab$. L'ensemble $\mathcal{M}_1(\mathbb{K})$ muni de l'addition et de la multiplication matricielles s'identifie à \mathbb{K} , muni de l'addition et de la multiplication usuelles.

On termine cette partie par une mise en garde :

IV.1.f. Formule du binôme

La formule du binôme (cf §I.7.b) reste vraie pour des matrices carrées, *lorsque ces matrices commutent*.

Proposition IV.1.36. Soit $A, B \in \mathcal{M}_n(\mathbb{K})$ telles que

$$AB = BA.$$

Alors

$$(IV.6) \quad (A + B)^n = \sum_{k=1}^n \binom{n}{k} A^k B^{n-k}.$$

La démonstration, qui est exactement la même que dans le cas scalaire, est laissée au lecteur. On peut raisonner par récurrence sur n , en utilisant la formule (I.13).

Exercice IV.1.37. Trouver deux matrices 2×2 A et B , qui ne commutent pas, et telles que la formule (IV.6) soit fautive avec $n = 2$.

IV.2. Matrices inversibles : définitions et exemples

Le but de cette section et de la suivante est l'étude des matrices carrées inversibles pour la multiplication matricielle. Cette section est consacrée à des définitions et des exemples simples. En IV.3, nous verrons une méthode de calculs des inverses et une caractérisation des matrices inversibles qui reposent largement sur la méthode du pivot de Gauss vue au chapitre III.

IV.2.a. Définition

Définition IV.2.1. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. La matrice A est dite *inversible* quand il existe des matrices $B, C \in \mathcal{M}_n(\mathbb{K})$ telles que

$$AB = CA = I_n,$$

où la matrice I_n est la matrice identité $n \times n$ définie en (IV.2).

Exemple IV.2.2. La matrice nulle de $\mathcal{M}_n(\mathbb{K})$ n'est pas inversible. La matrice I_n est inversible ($I_n = I_n \times I_n$). La matrice $\begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}$ est inversible :

$$\begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} = I_2.$$

Proposition IV.2.3. Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, il existe un unique $B \in \mathcal{M}_n(\mathbb{K})$ tel que

$$(IV.7) \quad AB = BA = I_n.$$

La matrice B est appelée inverse de A , et notée A^{-1} . L'unicité signifie que si $M \in \mathcal{M}_n(\mathbb{K})$ vérifie $MA = I_n$ ou $AM = I_n$, alors $M = A^{-1}$.

En d'autres termes, si une matrice est inversible, l'inverse à gauche et l'inverse à droite de cette matrice sont uniques et égaux.

Démonstration. Il suffit de montrer que si $AB = I_n$ et $CA = I_n$, alors $B = C$. La matrice B donnée par la définition IV.2.1 vérifiera alors (IV.7), et sera bien unique au sens donné par la proposition.

En multipliant à gauche l'égalité $AB = I_n$ par C , on obtient

$$\underbrace{CA}_{I_n} B = CI_n = C$$

et donc $B = C$ (on a utilisé l'associativité de la multiplication matricielle). □

Voici une propriété importante des matrices inversibles (cf l'avertissement IV.1.34) :

Proposition IV.2.4. *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice inversible. Alors :*

i. Si $M \in \mathcal{M}_{p,n}(\mathbb{K})$,

$$MA = 0 \implies M = 0.$$

ii. Si $M \in \mathcal{M}_{n,p}(\mathbb{K})$,

$$AM = 0 \implies M = 0.$$

Remarque IV.2.5. La proposition implique que si A est inversible, $(0, 0, \dots, 0)$ est l'unique solution du système homogène $AX = 0$, $X \in \mathbb{K}^n$ qui a n solutions et n inconnues. En d'autres termes, ce système est un système de Cramer. De fait, l'unique solution de l'équation matricielle $AX = B$ est $X = A^{-1}B$. Nous verrons plus loin qu'il y a en fait équivalence : un système de n équations à n inconnues est un système de Cramer si et seulement si la matrice de ses coefficients est inversible.

Démonstration. Démontrons (i), la démonstration de (ii) est similaire. On suppose donc $MA = 0$. En multipliant à droite par A^{-1} , on obtient

$$M \underbrace{AA^{-1}}_{I_n} = 0A^{-1} = 0$$

et donc $M = MI_n = 0$. □

On peut en déduire un exemple typique de matrice non-inversible :

Proposition IV.2.6. *Soit $A \in \mathcal{M}_n(\mathbb{K})$. On suppose qu'une des colonnes, ou une des lignes de A est nulle. Alors A n'est pas inversible.*

Démonstration. On suppose que la i -ème ligne de $A = [a_{ij}]$ est nulle. Soit $Y = [y_j]_{1 \leq j \leq n}$ la matrice ligne de $\mathcal{M}_{1,n}(\mathbb{K})$ dont tous les coefficients sont nuls, sauf le i -ème, qui vaut 1. Alors la matrice ligne YA est nulle : en effet, le ℓ -ième coefficient de cette matrice est donné par

$$\sum_{k=1}^n y_k a_{k\ell} = 0,$$

car $y_k = 0$ si $k \neq i$ par définition de Y , et $a_{i\ell} = 0$ car la i -ème ligne de A est nulle. On en déduit par la Proposition IV.2.4 que A n'est pas inversible.

Dans le cas où la j -ième colonne de A est nulle, on fait le même raisonnement en multipliant A par la matrice colonne $X \in \mathcal{M}_n(\mathbb{K})$ dont tous les coefficients sont nuls, sauf le j -ième qui vaut 1. □

Exemple IV.2.7. La matrice $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ -1 & 2i & 4 \end{bmatrix}$ n'est pas inversible. Dans ce cas, la démonstration de la proposition IV.2.6 signifie que $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix} A = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$, contredisant la proposition IV.2.4.

On donne maintenant deux exemples où il est facile de voir si une matrice est inversible et, le cas échéant, de calculer son inverse.

IV.2.b. Matrices diagonales

Définition IV.2.8. On appelle *matrice diagonale* une matrice carrée $[a_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ dont les coefficients en dehors de la diagonale principale $\{i = j\}$ sont nuls. En d'autres termes :

$$i \neq j \implies a_{ij} = 0.$$

Exemples IV.2.9. Les matrices I_n et $0 \in \mathcal{M}_n(\mathbb{K})$ sont diagonales.

Considérons les matrices

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 0 & 0 \end{bmatrix} \quad \text{et} \quad C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

La matrice A est diagonale. Les matrices B et C ne sont pas diagonales (C n'est même pas carrée).

Remarquons que la somme de deux matrices diagonales de même taille est diagonale, et que si A est diagonale, ${}^tA = A$. On note $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ ou $\text{diag}(\lambda_j)_{1 \leq j \leq n}$ la matrice diagonale dont les coefficients diagonaux sont $\lambda_1, \lambda_2, \dots, \lambda_n$. Par exemple

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \text{diag}(1, 2, 3) = \text{diag}(j)_{1 \leq j \leq 3}, \quad I_n = \text{diag}(1)_{1 \leq j \leq n}.$$

On peut calculer très facilement le produit de deux matrices diagonales :

Proposition IV.2.10. Soit $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ et $B = \text{diag}(\mu_1, \mu_2, \dots, \mu_n)$ deux matrices diagonales de même taille. Alors

$$AB = \text{diag}(\lambda_1\mu_1, \lambda_2\mu_2, \dots, \lambda_n\mu_n).$$

La démonstration, facile, est laissée au lecteur (utiliser (IV.1)).

Corollaire IV.2.11. La matrice diagonale $D = \text{diag}(\lambda_j)_{1 \leq j \leq n}$ est inversible si et seulement si $\lambda_j \neq 0$ pour tout $j \in \{1, \dots, n\}$. Dans ce cas,

$$D^{-1} = \text{diag}(1/\lambda_j)_{1 \leq j \leq n}.$$

Démonstration. Si tous les λ_j sont non nuls, il est facile de vérifier, en utilisant la proposition IV.2.10 que

$$\text{diag}(1/\lambda_j)_{1 \leq j \leq n} D = D \text{diag}(1/\lambda_j)_{1 \leq j \leq n} = I_n.$$

Supposons qu'il existe $k \in \{1, \dots, n\}$ tel que $\lambda_k = 0$. Alors la k -ième ligne de D est nulle, et donc par la Proposition IV.2.6, D n'est pas inversible. \square

Exemples IV.2.12. La matrice

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

est inversible, d'inverse

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 1/3 \end{bmatrix}.$$

IV.2.c. Inversibilité des matrices 2×2

Proposition IV.2.13. Soit $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ une matrice 2×2 . Alors la matrice A est inversible si et seulement si $ad - bc \neq 0$. Dans ce cas, on a $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Remarque IV.2.14. La quantité $ad - bc$ est appelée *déterminant* de A . Le déterminant se généralise à des matrices carrées de plus grande dimension mais les formules sont plus compliquées.

Démonstration. Soit $B = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Par la formule du produit matriciel :

$$AB = BA = (ad - bc)I_2.$$

Lorsque $ad - bc = 0$, on obtient $AB = 0$ et la proposition IV.2.4 montre que A n'est pas inversible. Lorsque $ad - bc \neq 0$, on obtient

$$A \frac{1}{ad - bc} B = \frac{1}{ad - bc} B A = I_2,$$

ce qui montre que A est inversible, d'inverse $\frac{1}{ad-bc}B$. □

Exercice IV.2.15. Inverser les matrices suivantes

$$\begin{bmatrix} -1 & -47 \\ 27 & 1268 \end{bmatrix} \quad \begin{bmatrix} -1 & 25 \\ 33 & -824 \end{bmatrix} \quad \begin{bmatrix} -18 & -19 \\ 701 & 740 \end{bmatrix}.$$

IV.2.d. Stabilité par multiplication et transposition

Proposition IV.2.16. *i.* Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors tA est inversible si et seulement si A est inversible. Dans ce cas, $({}^tA)^{-1} = {}^t(A^{-1})$.

ii. Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ deux matrices inversibles. Alors AB est inversible et

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

Démonstration. On montre d'abord (i). Supposons pour commencer que A est inversible. On transpose les égalités :

$$AA^{-1} = A^{-1}A = I_n,$$

ce qui donne $({}^t(AB) = {}^tB{}^tA)$:

$${}^t(A^{-1}){}^tA = {}^tA{}^t(A^{-1}) = {}^tI_n = I_n.$$

Donc tA est inversible, d'inverse ${}^t(A^{-1})$.

Réciproquement, si tA est inversible, $A = {}^t({}^tA)$ est inversible par ce qui précède, ce qui conclut la preuve du point (i).

Pour montrer (ii), on utilise l'associativité de la multiplication :

$$AB B^{-1}A^{-1} = A I_n A^{-1} = A A^{-1} = I_n,$$

et de même

$$B^{-1}A^{-1} AB = B^{-1} I_n B = B^{-1}B = I_n.$$

□

Avertissement IV.2.17. Il ne faut pas se tromper dans l'ordre des facteurs dans la formule du (ii). Rappelons que la multiplication matricielle n'est pas commutative. On n'a donc pas, en général $(AB)^{-1} = A^{-1}B^{-1}$.

IV.3. Opérations sur les lignes et inversion de matrices

Les matrices inversibles ont plusieurs caractérisations équivalentes. Pour le démontrer, nous allons utiliser plusieurs notions déjà vues au chapitre III du cours, consacré aux systèmes linéaires. En IV.3.a, on introduit des *matrices élémentaires* correspondant aux opérations élémentaires du chapitre III. En IV.3.b on reparle de matrices échelonnées réduites, et en IV.3.c de la méthode du pivot de Gauss. Les matrices inversibles sont caractérisées en IV.3.d, par le théorème IV.3.19. Le IV.3.e est consacrée à l'interprétation matricielle des systèmes de Cramer.

Les deux points les plus importants de ce chapitre sont l'inversion de matrice par la méthode du pivot de Gauss et le théorème IV.3.19.

IV.3.a. Matrices élémentaires

On va montrer que les opérations élémentaires sur les lignes d'une matrice, rencontrée dans le chapitre sur les systèmes linéaires, reviennent à des multiplications à gauche par des matrices bien particulières, appelées matrices élémentaires. On commence par définir ces matrices.

On fixe $n \geq 2$.

Définition IV.3.1. On appelle *matrice élémentaire* une matrice carrée $n \times n$ d'un des trois types suivants.

Soient $\lambda \in \mathbb{K} \setminus \{0\}$ et $k \in \{1, \dots, n\}$. La matrice de *dilatation* $D_k(\lambda)$ est la matrice diagonale de $\mathcal{M}_n(\mathbb{K})$ dont le k -ième coefficient diagonal vaut λ et les autres coefficients diagonaux valent 1.

Soient $k, \ell \in \{1, \dots, n\}$ avec $k \neq \ell$, et $\lambda \in \mathbb{K}$. La matrice de *transvection* $T_{k\ell}(\lambda)$ est la matrice dont les coefficients diagonaux sont égaux à 1, le coefficient (k, ℓ) vaut λ , et les autres coefficients sont nuls.

Si $k, \ell \in \{1, \dots, n\}$ avec $k \neq \ell$, on note $R_{k\ell}$ la matrice dont les coefficients diagonaux valent 1, sauf les coefficients (k, k) et (ℓ, ℓ) , qui valent 0, les coefficients (k, ℓ) et (ℓ, k) valent 1, et les autres coefficients sont nuls. La matrice $R_{k\ell}$ est donc obtenue, à partir de I_n , en échangeant la ligne k et la ligne ℓ . Remarquons que $R_{\ell k} = R_{k\ell}$. Les matrices $R_{k\ell}$ sont parfois appelées *matrices de transposition* (à ne pas confondre avec la transposée tA).

Exemples IV.3.2. On suppose $n = 3$. Alors

$$D_1(\lambda) = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad D_2(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad D_3(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{bmatrix}, \quad T_{12}(\lambda) = \begin{bmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$T_{31}(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \lambda & 0 & 1 \end{bmatrix}, \quad R_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad R_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad R_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Exercice IV.3.3. Ecrire $D_2(\lambda)$, $T_{24}(\lambda)$, $T_{42}(\lambda)$, R_{13} quand $n = 4$.

Proposition IV.3.4. Soit $q \geq 1$ et $A \in \mathcal{M}_{n,q}(\mathbb{K})$.

- i. Soient $k \in \{1, \dots, n\}$ et $\lambda \neq 0$. La matrice $D_k(\lambda)A$ est obtenue à partir de A en multipliant la k -ième ligne de A par λ . La multiplication à gauche par $D_k(\lambda)$ correspond donc à l'opération élémentaire, appelée cadrage et notée $(L_k) \leftarrow \lambda(L_k)$ dans le chapitre précédent du cours.

- ii. Soient $k, \ell \in \{1, \dots, n\}$, $k \neq \ell$ et $\lambda \neq 0$. La matrice $T_{k\ell}(\lambda)A$ est obtenue à partir de A en ajoutant à la k -ième ligne de A le produit de λ et la ℓ -ième ligne de A . La multiplication à gauche par $T_{k\ell}(\lambda)$ correspond donc à l'opération élémentaire, appelée remplacement et notée $(L_k) \leftarrow (L_k) + \lambda(L_\ell)$ dans le chapitre précédent du cours.
- iii. Soient $k, \ell \in \{1, \dots, n\}$. La matrice $R_{k\ell}A$ est obtenue à partir de A en échangeant les lignes k et ℓ . La multiplication à gauche par $R_{k\ell}$ correspond donc à l'opération élémentaire notée $(L_k) \leftrightarrow (L_\ell)$ dans le chapitre précédent.

Remarque IV.3.5. On peut montrer² que la multiplication à droite par les matrices élémentaires correspond à des opérations élémentaires sur les colonnes.

Exercice IV.3.6. Calculer en utilisant la formule du produit matriciel :

$$T_{21}(\lambda) \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

et vérifier que le résultat est cohérent avec la proposition IV.3.4.

Preuve de la proposition IV.3.4. On ne démontre que le point (ii). La démonstration des autres points est laissée au lecteur.

On note $A = [a_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq q}}$, $T_{k\ell}(\lambda) = [b_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq q}}$ et $T_{k\ell}(\lambda)A = [c_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq q}}$. On a donc

$$b_{ii} = 1, \quad i = 1 \dots n, \quad b_{k\ell} = \lambda, \quad b_{ij} = 0 \text{ si } i \neq j \text{ et } (i, j) \neq (k, \ell).$$

Soient $i \in \{1, \dots, n\}$ et $j \in \{1, \dots, q\}$. La formule du produit matriciel (IV.1) donne :

$$(IV.8) \quad c_{ij} = \sum_{r=1}^n b_{ir} a_{rj}.$$

Si $i \neq k$, $b_{ir} = 0$ pour $r \neq i$, et $b_{ii} = 1$. La formule précédente donne donc $c_{ij} = a_{ij}$. La i -ième ligne de $T_{k\ell}(\lambda)A$ est donc exactement la i -ième ligne de A .

On considère maintenant le cas $i = k$. On a $b_{kr} = 0$ pour $r \notin \{k, \ell\}$, $b_{kk} = 1$, et $b_{k\ell} = \lambda$. La formule (IV.8) avec $i = k$ s'écrit donc

$$c_{kj} = a_{kj} + \lambda a_{\ell j},$$

et la k -ième ligne de $T_{k\ell}(\lambda)A$:

$$[c_{k1}, \dots, c_{kq}] = [a_{k1}, \dots, a_{kq}] + \lambda[a_{\ell 1}, \dots, a_{\ell q}] = (L_k) + \lambda(L_\ell),$$

en notant (L_k) et (L_ℓ) la k -ième et la ℓ -ième ligne de A . Le point (ii) est démontré. \square

Exercice IV.3.7. En s'inspirant de la démonstration précédente, montrer les points (i) et (iii) de la proposition IV.3.4.

Proposition IV.3.8. Les matrices $D_k(\lambda)$ ($\lambda \neq 0$), $T_{k\ell}(\lambda)$ ($k \neq \ell$) et $R_{k\ell}$ sont inversibles et :

i. $(D_k(\lambda))^{-1} = D_k\left(\frac{1}{\lambda}\right)$;

2. Par exemple en appliquant la proposition IV.3.4 aux matrices transposées.

IV. Introduction aux matrices

- ii. $(T_{k\ell}(\lambda))^{-1} = T_{k\ell}(-\lambda)$;
- iii. $R_{k\ell}^{-1} = R_{k\ell}$.

Démonstration. Les matrices de dilatation étant diagonales, le point (i) découle immédiatement du Corollaire IV.2.11 (on peut aussi utiliser la proposition IV.3.4 comme dans ce qui suit).

D'après la proposition IV.3.4, la matrice

$$T_{k\ell}(\lambda)T_{k\ell}(-\lambda) = T_{k\ell}(\lambda)T_{k\ell}(-\lambda) I_n$$

est obtenue à partir de la matrice I_n par les opérations :

$$(L_k) \leftarrow (L_k) - \lambda L_\ell,$$

puis

$$(L_k) \leftarrow (L_k) + \lambda L_\ell.$$

Puisque $k \neq \ell$, on a donc bien³ $T_{k\ell}(\lambda)T_{k\ell}(-\lambda) = I_n$ et de même $T_{k\ell}(-\lambda)T_{k\ell}(\lambda) = I_n$, ce qui montre le point (ii).

D'après la proposition IV.3.4, la matrice $R_{k\ell}R_{k\ell} = R_{k\ell}R_{k\ell}I_n$ est obtenue à partir de I_n en échangeant deux fois les lignes ℓ et k . C'est donc bien la matrice I_n , ce qui montre le point (iii). \square

Exercice IV.3.9. Calculer lorsque $n = 4$, en utilisant la formule du produit ligne-colonne, $R_{24}(17)R_{24}(-17)$ et vérifier le point (IV.3.4).

IV.3.b. Matrices échelonnées réduites carrées

On renvoie au chapitre précédent ou à David C. Lay⁴ pour la définition d'une matrice échelonnée et d'une matrice échelonnée réduite. Le but de cette partie est de caractériser les matrices échelonnées réduites carrées inversibles.

Définition IV.3.10. La matrice carrée $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{K})$ est dite *triangulaire supérieure* si tous les coefficients de A en dessous de la diagonale principale sont nulle. En d'autres termes :

$$1 \leq j < i \leq n \implies a_{ij} = 0.$$

Exemples IV.3.11. Les matrices diagonales sont triangulaires supérieures. En particulier, la matrice I_n et la matrice $0 \in \mathcal{M}_n(\mathbb{K})$ sont triangulaires supérieures. Considérons les matrices :

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 3 \\ 0 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 1 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & 2 & 3 \end{bmatrix}.$$

La matrice A est triangulaire supérieure. Les matrices B et C ne le sont pas (B n'est pas carrée. Le coefficient $(2, 1)$ de C est non nul).

3. l'hypothèse $k \neq \ell$ montre que la ligne (L_ℓ) n'a pas changé après la première opération.

4. David C. Lay. *Algèbre linéaire : Théorie, exercices et applications*. Troisième édition, 2004

Proposition IV.3.12. *Une matrice échelonnée carrée est triangulaire supérieure.*

Démonstration. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice échelonnée. On note p le nombre de lignes non nulles de A . On a donc $0 \leq p \leq n$, $p = n$ si toutes les lignes de A sont non nulles, $p = 0$ si $A = 0$. De plus, si $1 \leq p \leq n$, A étant échelonnée, les lignes $1, 2, \dots, p$ de A sont non nulles, et les lignes $p + 1, \dots, n$ sont nulles.

On suppose $A \neq 0$, i.e. $p \geq 1$ (sinon $A = 0$ est triangulaire supérieure et la démonstration est finie).

Pour $1 \leq i \leq p$, on note $J(i)$ la colonne de l'élément de tête (le coefficient non nul le plus à gauche) de la i -ème ligne de A . Par propriété des matrices échelonnées, on a $J(k+1) \geq J(k) + 1$ pour $k = 1 \dots p - 1$ et donc, puisque $J(1) \geq 1$, $J(i) \geq i$ pour tout i entre 1 et p . On en déduit

$$(1 \leq i \leq p \text{ et } j < i) \implies (1 \leq i \leq p \text{ et } j < J(i)) \implies a_{ij} = 0,$$

ce qui montre que la matrice A est triangulaire supérieure. □

Exemple IV.3.13. Pour illustrer la proposition et sa démonstration, considérons la matrice échelonnée 5×5 :

$$\begin{bmatrix} 1 & 1+i & 3 & 2 & 0 \\ 0 & 0 & 2 & 3 & i \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

C'est bien une matrice triangulaire supérieure. Le nombre de lignes non nulles est $p = 4$; $J(1) = 1$, $J(2) = 3$, $J(3) = 4$ et $J(4) = 5$.

Théorème IV.3.14. *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice échelonnée réduite. Les conditions suivantes sont équivalentes :*

- i. A est inversible ;*
- ii. aucune ligne de A n'est nulle ;*
- iii. $A = I_n$.*

La seule matrice échelonnée réduite inversible est donc la matrice identité.

Démonstration. La matrice I_n est inversible, donc (iii) \implies (i). De plus, on sait déjà. (i) \implies (ii) (cf proposition IV.2.6).

Il reste à démontrer (ii) \implies (iii). Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice échelonnée réduite sans ligne nulle. Par la proposition IV.3.12, elle est triangulaire supérieure. Par hypothèse, elle a exactement n lignes non nulles. On reprend la notation $J(i)$ de la démonstration de la proposition IV.3.12. Montrons pour commencer

$$(IV.9) \quad \forall i \in \{1, \dots, n\}, \quad J(i) = i.$$

Puisque la matrice est triangulaire supérieure, on a $i \leq J(i) \leq n$ pour tout $i \in \{1, \dots, n\}$. La matrice A étant échelonnée, on a aussi $J(i) \leq J(i + 1) - 1$ pour tout $i \in \{1, \dots, n - 1\}$. Puisque $J(n) \leq n$, on obtient $J(i) \leq i$ pour tout i , par une récurrence descendante sur i . D'où (IV.9).

Par (IV.9), les éléments de tête sont tous sur la diagonale $i = j$. La matrice A étant échelonnée réduite, ces éléments de tête sont tous égaux à 1, et les coefficients au-dessus de chaque élément de tête sont nuls, ce qui montre $A = I_n$. □

IV.3.c. Inversions de matrices par la méthode du pivot de Gauss

Soit A une matrice carrée. Par la méthode du pivot de Gauss (cf chapitre précédent du cours), on peut ramener A à une matrice échelonnée réduite A' par un certain nombre (disons p) de transformations élémentaires sur les lignes de A . Ces transformations élémentaires correspondent à des multiplications par des matrices élémentaires (cf §IV.3.a et Proposition IV.3.4), on a :

$$E_p \dots E_1 A = A',$$

où les matrices E_j correspondent aux p transformations élémentaires appliquées à A . Puisque les matrices élémentaires sont inversibles, on a :

$$A = E_1^{-1} \dots E_p^{-1} A'.$$

D'où (les matrices E_j^{-1} étant elles aussi des matrices élémentaires) :

Théorème IV.3.15. *Toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ est produit de matrices élémentaires et d'une matrice échelonnée réduite A' . Elle est inversible si et seulement si $A' = I_n$, c'est à dire si et seulement si elle est produit de matrices élémentaires.*

(le dernier point découle du théorème IV.3.14).

Application : calcul de matrice inversible

Donnons maintenant une méthode pratique pour étudier l'inversibilité de A , et, lorsque A est inversible, calculer son inverse. On commence par écrire sur deux colonnes la matrice A et la matrice I_n . On ramène ensuite, par la méthode du pivot de Gauss, la matrice A à une matrice échelonnée réduite, tout en appliquant les mêmes opérations élémentaires sur la matrice I_n .

- Si A est inversible, on obtient sur la colonne de gauche la matrice $I_n = E_p \dots E_1 A$ et sur la colonne de droite la matrice $E_p \dots E_1$. L'égalité $I_n = E_p \dots E_1 A$ montre que la matrice obtenue sur la colonne de droite est exactement A^{-1} .
- Si A n'est pas inversible, on obtient sur la colonne de gauche une matrice échelonnée réduite avec au moins une ligne nulle. Dans ce cas, si le but est seulement d'étudier l'inversibilité de A , la colonne de droite est inutile, et on peut arrêter la méthode du pivot dès que l'on a obtenu une ligne nulle.

Exemple IV.3.16. On veut inverser la matrice $A = \begin{bmatrix} 1 & -1 & 1 \\ -2 & 3 & 0 \\ -6 & 8 & -1 \end{bmatrix}$.

$$\begin{array}{ccc}
 \begin{bmatrix} 1 & -1 & 1 \\ -2 & 3 & 0 \\ -6 & 8 & -1 \end{bmatrix} & & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
 \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 5 \end{bmatrix} & \begin{array}{l} (L_2) \leftarrow (L_2) + 2(L_1) \\ (L_3) \leftarrow (L_3) + 6(L_1) \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 6 & 0 & 1 \end{bmatrix} \\
 \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} & (L_3) \leftarrow (L_3) - 2(L_2) & \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & -2 & 1 \end{bmatrix} \\
 \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{array}{l} (L_1) \leftarrow (L_1) - (L_3) \\ (L_2) \leftarrow (L_2) - 2(L_3) \end{array} & \begin{bmatrix} -1 & 2 & -1 \\ -2 & 5 & -2 \\ 2 & -2 & 1 \end{bmatrix} \\
 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & (L_1) \leftarrow (L_1) + (L_2) & \begin{bmatrix} -3 & 7 & -3 \\ -2 & 5 & -2 \\ 2 & -2 & 1 \end{bmatrix}.
 \end{array}$$

Donc A est inversible, d'inverse

$$A^{-1} = \begin{bmatrix} -3 & 7 & -3 \\ -2 & 5 & -2 \\ 2 & -2 & 1 \end{bmatrix}.$$

Exemple IV.3.17. Considérons la matrice $A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 3 & 4 \\ 1 & 3 & 3 \end{bmatrix}$. Par les opérations $(L_2) \leftarrow (L_2) - 2(L_1)$, $(L_3) \leftarrow (L_3) - (L_1)$ puis $(L_3) \leftarrow (L_3) - (L_2)$, on obtient la matrice $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 0 \end{bmatrix}$. La troisième ligne de cette matrice étant nulle, on en déduit que A n'est pas inversible.

On termine cette partie par une remarque qui découle facilement de la méthode précédente :

Proposition IV.3.18. *Soit A une matrice triangulaire supérieure. Alors A est inversible si et seulement si ses coefficients diagonaux sont tous non-nuls.*

En effet, si aucun des coefficients diagonaux de A n'est nul, A est échelonnée, et la phase de remontée de la méthode du pivot permet d'obtenir I_n à partir de A par des opérations élémentaires sur les lignes.

En revanche, si un des coefficients diagonaux est nul, on peut transformer, par des opérations élémentaires sur les lignes, la matrice A en une matrice A' ayant une ligne nulle, ce qui montrera que A n'est pas inversible (sinon A' serait un produit de matrice inversible, donc inversible). Pour obtenir la ligne nulle, on note (i, i) les coordonnées du dernier coefficient diagonal nul, de tel sorte que $a_{ii} = 0$, et $a_{kk} \neq 0$ pour $k > i$. Exactement de la même façon que dans la phase de remontée de la méthode du pivot, en utilisant que les éléments de tête des lignes (L_k) , $k > i$ sont non nuls, on transforme A , par une série de remplacements, en une matrice dont la i -ième ligne est nulle.

IV.3.d. Caractérisation des matrices inversibles

Le théorème fondamental suivant, qui découle de ce qui précède, donne plusieurs critères pour reconnaître une matrice inversible.

Théorème IV.3.19. *Soit $A \in \mathcal{M}_n(\mathbb{K})$. Les conditions suivantes sont équivalentes :*

- i. A est inversible ;
- ii. $\exists B \in \mathcal{M}_n(\mathbb{K})$ t.q. $BA = I_n$;
- iii. $\exists C \in \mathcal{M}_n(\mathbb{K})$ t.q. $AC = I_n$;
- iv. l'équation $AX = 0$, d'inconnue $X \in \mathcal{M}_{n,1}(\mathbb{K})$, a pour seule solution $X = 0$;
- v. pour tout $E \in \mathcal{M}_{n,1}(\mathbb{K})$, l'équation $AX = E$, a une seule solution $X \in \mathcal{M}_{n,1}(\mathbb{K})$;

Démonstration. On commence par montrer que les points (i), (ii), (iv) sont équivalents. Par définition de l'inversibilité, (i) \Rightarrow (ii). Par ailleurs, si (ii) est vrai et $AX = 0$, alors $Y = BAX = B0 = 0$ et donc (ii) \Rightarrow (iv).

Supposons (iv). On veut montrer que A est inversible. Par la méthode du pivot de Gauss, réinterprétée en terme d'opérations élémentaires sur les lignes (cf §IV.3.c), $E_p \dots E_1 {}^t A = R$ où R est une matrice échelonnée réduite et les E_j des matrices élémentaires. On veut montrer que $R = I_n$, ce qui impliquera que ${}^t A$ est inversible et donc que A est inversible. On raisonne par l'absurde : si $R \neq I_n$, par le théorème IV.3.14, la dernière ligne de R est une ligne de 0. Soit $Y' = [0 \ \dots \ 0 \ 1] \in \mathcal{M}_{1,n}(\mathbb{K})$. Alors

$$Y' E_p \dots E_1 {}^t A = Y' R = 0$$

et donc

$$Y {}^t A = 0$$

avec $Y = Y' E_p \dots E_1 \in \mathcal{M}_{1,n}(\mathbb{K})$. Les matrices E_1, \dots, E_p étant inversibles, Y est non nul. En passant à la transposée, on obtient $AX = 0$ avec $X = {}^t Y$, qui est un vecteur colonne non nul. Ceci contredit (iv). Donc A est inversible, ce qui conclut la preuve de (iv) \Rightarrow (i).

On a évidemment (v) \Rightarrow (iv), (i) \Rightarrow (v), et puisque (iv) \Rightarrow (i), le point (v) est équivalent à tous les points précédents.

Enfin, (i) implique (iii) par définition. Réciproquement, supposons (iii). En transposant l'égalité $AC = I_n$, on obtient ${}^t C {}^t A = I_n$. Donc ${}^t A$ vérifie (ii). Par ce qui précède, ${}^t A$ est inversible, ce qui implique, par la proposition IV.2.16 sur l'inversibilité des matrices transposées, que A est inversible. On a montré (iii) \Rightarrow (i), ce qui conclut la preuve du théorème. \square

Remarque IV.3.20. On peut également montrer que l'inversibilité de A est équivalente au fait que l'équation $YA = 0$, d'inconnue $Y \in \mathcal{M}_{1,n}(\mathbb{K})$, a pour seule solution $Y = 0$, ou que pour tout $F \in \mathcal{M}_{1,n}(\mathbb{K})$, l'équation $YA = F$, a une seule solution $Y \in \mathcal{M}_{1,n}(\mathbb{K})$. La démonstration de ces équivalences est laissée au lecteur.

IV.3.e. Système de Cramer et matrice inversible

Soit (S) un système linéaire à n équations et n inconnues, et A la matrice des coefficients. La matrice A est donc une matrice carrée $n \times n$. Le système (S) s'écrit

$$AX = B,$$

où B est la matrice colonne ($B \in \mathcal{M}_{n,1}(\mathbb{K})$) formé du second membre de l'équation, et $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ est la matrice inconnue.

Le point (i) \iff (v) du théorème IV.3.19 signifie exactement :

$$(S) \text{ est un système de Cramer } \iff A \text{ est inversible.}$$

On distingue deux cas :

- La matrice A est inversible et le système (S) est un système de Cramer : il a une unique solution $X = A^{-1}B$, quel que soit le second membre B . Le calcul de A^{-1} permet de résoudre rapidement le système quel que soit B .
- Si A n'est pas inversible, le système homogène $AX = 0$ a une infinité de solutions : en effet, le point (iv) est faux, il y a donc une solution non nulle X et tous les λX , $\lambda \in \mathbb{K}$ sont aussi solutions. Le système (S) a ou bien aucune solution, ou bien une infinité de solutions. Dans ce cas, A étant fixée, la compatibilité du système $AX = B$ dépend du second membre B .

Exemple IV.3.21. Résoudre les systèmes :

$$\begin{cases} x_1 - 3x_2 - 5x_3 - 3x_4 = 1 \\ -x_1 + 2x_2 + 3x_3 + x_4 = 2 \\ -x_1 + 3x_2 + 4x_3 + 2x_4 = 0 \\ -3x_2 - 3x_3 - 4x_4 = 1 \end{cases}, \quad \begin{cases} x_1 - 3x_2 - 5x_3 - 3x_4 = 2 \\ -x_1 + 2x_2 + 3x_3 + x_4 = -1 \\ -x_1 + 3x_2 + 4x_3 + 2x_4 = 1 \\ -3x_2 - 3x_3 - 4x_4 = 0 \end{cases}$$

$$\begin{cases} x_1 - 3x_2 - 5x_3 - 3x_4 = -2 \\ -x_1 + 2x_2 + 3x_3 + x_4 = 0 \\ -x_1 + 3x_2 + 4x_3 + 2x_4 = 0 \\ -3x_2 - 3x_3 - 4x_4 = 1 \end{cases}$$

La matrice des coefficients de ces trois systèmes est

$$A = \begin{bmatrix} 1 & -3 & -5 & -3 \\ -1 & 2 & 3 & 1 \\ -1 & 3 & 4 & 2 \\ 0 & -3 & -3 & -4 \end{bmatrix}.$$

On montre par la méthode du pivot que A est inversible et

$$A^{-1} = \begin{bmatrix} -1 & -9 & 7 & 2 \\ 1 & -1 & 2 & 0 \\ -1 & -3 & 2 & 1 \\ 0 & 3 & -3 & -1 \end{bmatrix}.$$

Les trois systèmes sont des systèmes de Cramer. Les solutions de ces systèmes sont respectivement :

$$A^{-1} \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -17 \\ -1 \\ -6 \\ 5 \end{bmatrix}, \quad A^{-1} \begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 14 \\ 5 \\ 3 \\ -6 \end{bmatrix}, \quad A^{-1} \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ 3 \\ -1 \end{bmatrix}$$

Exercice IV.3.22. Appliquer la méthode du pivot à l'exemple précédent pour calculer A^{-1} .

Appendice : réponse à quelques exercices

Exercice IV.1.25. $AC, AD, BA, B^2, CA, CB, DC$ et D^2 ne sont pas définis.

$$\begin{aligned}
 A^2 &= \begin{bmatrix} -5 & 2+2i \\ -3-3i & -7 \end{bmatrix}, & AB &= \begin{bmatrix} 6 & 3+i & -3 \\ -6+2i & -3-2i & 9 \end{bmatrix}, & BC &= \begin{bmatrix} 5 & 4 & 6 \\ 2 & 4 & 6 \end{bmatrix} \\
 BD &= \begin{bmatrix} 3-3i & 6-6i \\ -3 & -6 \end{bmatrix}, & C^2 &= \begin{bmatrix} -2 & 2 & 3 \\ 0 & 0 & 0 \\ -1 & -2 & -3 \end{bmatrix}, & CD &= \begin{bmatrix} -12 & -24 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \\
 DA &= \begin{bmatrix} 0 & 0 \\ 15 & -6-6i \\ 10 & -4-4i \end{bmatrix}, & DB &= \begin{bmatrix} 0 & 0 & 0 \\ -18 & -9-3i & 9 \\ -12 & -6-2i & 6 \end{bmatrix}
 \end{aligned}$$

Exercice IV.2.15. Les matrices inverses sont $\begin{bmatrix} 1268 & 47 \\ -27 & -1 \end{bmatrix}$, $\begin{bmatrix} 824 & 25 \\ 33 & 1 \end{bmatrix}$, $\begin{bmatrix} -740 & -19 \\ 701 & 18 \end{bmatrix}$.

V. Espaces vectoriels

Comme dans les chapitres précédents, on fixe $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Les éléments de \mathbb{K} sont appelés *nombre*s, ou *scalaires*.

V.1. Définitions et exemples

V.1.a. L'espace vectoriel \mathbb{K}^n

Soit $n \geq 1$ un entier. On rappelle que \mathbb{K}^n est l'ensemble des n -uplets $(x_1, \dots, x_n) = (x_j)_{j=1\dots n}$, avec $x_j \in \mathbb{K}$ pour tout n . Un élément \vec{x} de \mathbb{K}^n est appelé *vecteur*. Les vecteurs seront toujours notés avec une flèche pour les différencier des éléments de \mathbb{K} (les scalaires) notés sans flèche : ainsi \vec{x} , \vec{y} , \vec{x}_1 , \vec{x}_2 seront des vecteurs, x , y , λ , μ , x_1 , x_2 des scalaires. Les lettres grecques (λ , μ , ν etc...) désigneront systématiquement des scalaires (cf l'alphabet grec p. 96).

On considère sur \mathbb{K}^n les deux opérations suivantes :

Addition : soit $\vec{x} = (x_j)_{j=1\dots n}$ et $\vec{y} = (y_j)_{j=1\dots n}$ des vecteurs. Leur somme $\vec{x} + \vec{y}$ est par définition le vecteur $(x_j + y_j)_{j=1\dots n}$.

Multiplication par un scalaire : si λ est un scalaire et $\vec{x} = (x_j)_{j=1\dots n}$ un vecteur, leur produit est par définition le vecteur $(\lambda x_j)_{j=1\dots n}$.

Exemples V.1.1.

$$i(i, -1, 2) + (1, i, 4i) = (0, 0, 6i), \quad (k)_{k=1\dots 10} + \underbrace{(-j + 1)_{j=1\dots 10}}_{10 \text{ fois}} = (1, 1, 1, \dots, 1).$$

Avertissement V.1.2. On a défini le produit d'un scalaire par un vecteur, et pas le produit de deux vecteurs.

L'addition et multiplication par un scalaire vérifient les règles de calcul suivantes :

- i. *Associativité de l'addition* : $\forall \vec{x}, \vec{y}, \vec{z}, \quad (\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$ (on notera $\vec{x} + \vec{y} + \vec{z}$ leur valeur commune).
- ii. *Commutativité de l'addition* : $\forall \vec{x}, \vec{y}, \quad \vec{x} + \vec{y} = \vec{y} + \vec{x}$.
- iii. *Associativité de la multiplication* : Pour tous scalaires λ et μ , pour tout vecteur \vec{x} , $\lambda(\mu\vec{x}) = (\lambda\mu)\vec{x}$.
- iv. *Distributivité* : Pour tous scalaires λ et μ , pour tous vecteurs \vec{x} et \vec{y} , $\lambda(\vec{x} + \vec{y}) = \lambda\vec{x} + \lambda\vec{y}$ et $(\lambda + \mu)\vec{x} = \lambda\vec{x} + \mu\vec{x}$.
- v. Pour tout vecteur \vec{x} , $1\vec{x} = \vec{x}$.

On note $\vec{0}$ le vecteur de \mathbb{K}^n dont toutes les coordonnées sont nulles :

$$\vec{0} = (0, 0, \dots, 0).$$

V. Espaces vectoriels

On a :

$$(V.1) \quad \forall \lambda, \lambda \vec{0} = \vec{0}, \quad \forall \vec{x}, 0\vec{x} = \vec{0}.$$

Soit $\vec{x} = (x_j)_{j=1\dots n} \in \mathbb{K}^n$. On note $-\vec{x}$ le vecteur $(-x_1, \dots, -x_n)$. C'est l'unique vecteur \vec{x}' tel que $\vec{x} + \vec{x}' = \vec{0}$. On notera $\vec{x} - \vec{y} = \vec{x} + (-\vec{y})$ la *différence* de deux vecteurs.

La multiplication par un scalaire a la propriété de régularité suivante :

Proposition V.1.3. *Soit $\lambda \in \mathbb{K}$, $\vec{x} \in \mathbb{K}^n$. Alors*

$$\lambda \vec{x} = \vec{0} \implies (\lambda = 0 \text{ ou } \vec{x} = \vec{0}).$$

Démonstration. Supposons $\lambda \vec{x} = \vec{0}$ et $\lambda \neq 0$. On doit montrer $\vec{x} = \vec{0}$. On note $\vec{x} = (x_1, \dots, x_n)$. Soit $j \in \{1, \dots, n\}$. Puisque $\lambda x_j = 0$ et $\lambda \neq 0$, on a $x_j = 0$. Donc $\vec{x} = (0, 0, \dots, 0) = \vec{0}$. \square

V.1.b. Espaces vectoriels généraux

On commence par donner, pour mémoire, la définition générale d'un espace vectoriel. Conformément au programme, cette définition ne sera jamais utilisée directement dans ce cours.

Définition V.1.4. On appelle \mathbb{K} -*espace vectoriel* tout ensemble E muni d'une addition $E \times E \rightarrow E$ et d'une multiplication par un scalaire $\mathbb{K} \times E \rightarrow E$ qui vérifient les propriétés (i), (ii), (iii), (iv) et (v) ci-dessus, et tel qu'il existe $\vec{0} \in E$ vérifiant (V.1).

Tout sous-ensemble E de \mathbb{K}^n vérifiant :

$$(V.2) \quad \vec{0} \in E$$

$$(V.3) \quad \vec{x} \in E \text{ et } \vec{y} \in E \implies \vec{x} + \vec{y} \in E$$

$$(V.4) \quad \vec{x} \in E \text{ et } \lambda \in \mathbb{K} \implies \lambda \vec{x} \in E$$

est un espace vectoriel : les propriétés (i), (ii), (iii), (iv), (v) et (V.1), vraies sur \mathbb{K}^n , le sont automatiquement sur E . Dans ce cours, conformément au programme de L1 de l'institut Galilée, nous considérerons seulement ces exemples d'espaces vectoriels. Nous adopterons donc comme définition d'un espace vectoriel :

Définition V.1.5. Dans toute la suite de ce cours, on appellera \mathbb{K} -*espace vectoriel* tout ensemble E inclus dans \mathbb{K}^n pour un certain $n \geq 1$ et vérifiant les propriétés (V.2), (V.3) et (V.4).

La plupart des propriétés présentées dans la suite du cours sont en fait valables pour les espaces vectoriels généraux de la définition V.1.4 (nous le préciserons dans le cas contraire). Il existe bien entendu des exemples d'espaces vectoriels au sens de la définition V.1.4 qui ne rentrent pas dans le cadre de la définition V.1.5 : l'ensemble des polynômes sur \mathbb{K} , et l'ensemble des matrices $p \times n$ sur \mathbb{K} en sont deux exemples simples. Remarquons qu'il est facile d'identifier $\mathcal{M}_{p,n}(\mathbb{K})$ à \mathbb{K}^{pn} : l'addition et la multiplication par un scalaire définies sur les matrices correspondent exactement aux opérations de \mathbb{K}^{pn} définies en V.1.a. L'espace vectoriel $\mathbb{K}[X]$, en revanche, est de nature différente¹ et ne rentre pas dans le cadre de la définition V.1.5.

1. il est de dimension infinie, cf §V.4 pour la définition de la dimension.

V.1.c. Exemples

Exemple V.1.6. Soit $(a_1, \dots, a_n) \in \mathbb{K}^n$. L'ensemble des solutions $\vec{x} = (x_1, \dots, x_n)$ de l'équation

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = 0$$

est un \mathbb{K} -espace vectoriel.

Exemple V.1.7. L'ensemble $\{(\lambda, \mu, \lambda + \mu), \lambda \in \mathbb{R}, \mu \in \mathbb{R}\}$ est un \mathbb{R} -espace vectoriel.

V.2. Sous-espace vectoriels

On fixe un \mathbb{K} -vectoriel E .

V.2.a. Deux définitions équivalentes

Proposition V.2.1. *Soit F un sous-ensemble de E . Les propositions suivantes sont équivalentes :*

- i. F est un \mathbb{K} -espace vectoriel.*
- ii. Les trois propriétés suivantes sont vérifiées :*
 - $\vec{0} \in F$;
 - $(\vec{x}, \vec{y}) \in F^2 \implies \vec{x} + \vec{y} \in F$;
 - $(\vec{x} \in F, \lambda \in \mathbb{K}) \implies \lambda\vec{x} \in F$.

Définition V.2.2. Un ensemble F vérifiant les propriétés de la proposition V.2.1 est appelé *sous-espace vectoriel* de E .

Remarque V.2.3. En pratique, la définition donnée par le point (ii) de la proposition est celle que l'on utilise. L'autre définition justifie l'expression "sous-espace vectoriel".

Preuve de la proposition. Cela découle immédiatement de la définition V.1.5 d'un espace vectoriel adoptée dans ce cours. La proposition reste vraie en utilisant la définition générale d'un espace vectoriel (définition V.1.4) : la démonstration est élémentaire mais un petit peu plus longue. \square

Exemple V.2.4. Les espaces vectoriels étudiés dans ce cours (définition V.1.5) sont exactement les sous-espaces vectoriels des espaces vectoriels \mathbb{K}^n .

Exemple V.2.5. Les ensembles $\{\vec{0}\}$ et E sont des sous-espaces vectoriels de E , appelés *sous-espace vectoriels triviaux* de E .

Exemple V.2.6. Si E est un espace vectoriel, et $\vec{u} \in E \setminus \{\vec{0}\}$, l'ensemble

$$\{\lambda\vec{u}, \lambda \in \mathbb{K}\}$$

est un sous-espace vectoriel de E , appelée *droite* (vectorielle) de E . Dans le cas où E est l'espace vectoriel \mathbb{R}^2 ou \mathbb{R}^3 , ces ensembles sont exactement les droites du plan ou de l'espace passant par l'origine.

Exemple V.2.7. L'ensemble $\{(\lambda, \lambda, 2\lambda), \lambda \in \mathbb{R}\}$ est un sous-espace vectoriel du \mathbb{R} -espace vectoriel $\{(\lambda, \mu, \lambda + \mu), \lambda \in \mathbb{R}, \mu \in \mathbb{R}\}$ introduit dans l'exemple V.1.7.

Définition V.2.8. On appelle *combinaison linéaire* de $\vec{u}_1, \dots, \vec{u}_n$ un vecteur de E de la forme $\lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n$, où $\lambda_1, \dots, \lambda_n$ sont des scalaires.

Proposition V.2.9. Soit F un sous-espace vectoriel de E . Alors toute combinaison linéaire d'éléments de F est un élément de F .

Démonstration. C'est immédiat, par récurrence sur le nombre de termes de la combinaison linéaire. \square

V.2.b. Intersection

Proposition V.2.10. Soit F et G des sous-espaces vectoriels du \mathbb{K} -espace vectoriel E . Alors $F \cap G$ est un sous-espace vectoriel de E .

Démonstration. C'est immédiat.

Puisque F et G sont des sous-espaces vectoriels de E , on a $\vec{0} \in F$ et $\vec{0} \in G$ et donc $\vec{0} \in F \cap G$. De même, si \vec{x} et \vec{y} sont des vecteurs de $F \cap G$, ce sont des vecteurs de F , donc $\vec{x} + \vec{y} \in F$ (car F est un sous-espace vectoriel), et des vecteurs de G , donc $\vec{x} + \vec{y} \in G$ (car G est un sous-espace vectoriel). Par suite $\vec{x} + \vec{y} \in F \cap G$. La preuve de $(\vec{x} \in F \cap G, \lambda \in \mathbb{K}) \implies \lambda \vec{x} \in F \cap G$ est identique. \square

Exemple V.2.11. L'intersection des sous-espaces vectoriels de \mathbb{C}^3

$$F = \{(0, x_2, x_3), (x_2, x_3) \in \mathbb{C}^2\} \text{ et } G = \{(x_1, x_2, 0), (x_1, x_2) \in \mathbb{C}^2\}$$

est le sous-espace vectoriel de \mathbb{C}^3 :

$$F \cap G = \{(0, z, 0), z \in \mathbb{C}\}.$$

Proposition V.2.12. Soit F_1, \dots, F_n des sous-espaces vectoriels de E . Alors $F_1 \cap F_1 \cap \dots \cap F_n$ est un sous-espace vectoriel de E .

Démonstration. Par récurrence sur n , à partir de la proposition V.2.10. \square

Donnons un exemple fondamental :

Exemple V.2.13. Soit (H) un système linéaire homogène :

$$(H) \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = 0 \\ \vdots \\ a_{p1}x_1 + a_{22}x_2 + \dots + a_{pn}x_n = 0. \end{cases}$$

alors l'ensemble F des solutions de (H) est un sous-espace vectoriel de \mathbb{K}^n . En effet, notons, pour $i \in \{1, \dots, p\}$, F_i l'ensemble des solutions de l'équation $a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = 0$. Par l'exemple V.1.6, c'est un sous-espace vectoriel de E . Par la proposition V.2.12, $F = F_1 \cap F_2 \cap \dots \cap F_n$ est un sous-espace vectoriel de E .

On peut aussi montrer directement que F est un sous-espace vectoriel de E , en revenant à la définition d'un sous-espace vectoriel.

Avertissement V.2.14. Soit F et G deux sous-espaces vectoriels de E . Alors la réunion $F \cup G$ n'est pas, en général, un sous-espace vectoriel de E . Par exemple, la réunion des sous-espaces vectoriels $\{(x_1, 0), x_1 \in \mathbb{R}\}$ et $\{(0, x_2), x_2 \in \mathbb{R}\}$ de \mathbb{R}^2 n'est pas un sous-espace vectoriel de \mathbb{R}^2 : elle contient les vecteurs $(1, 0)$ et $(0, 1)$ mais pas leur somme $(1, 1)$. Un résultat plus précis est donné par la proposition suivante.

Proposition V.2.15. *Soit F et G des sous-espaces vectoriels de E . Alors $F \cup G$ est un sous-espace vectoriel de E si et seulement si $F \subset G$ ou $G \subset F$.*

Démonstration. Supposons que F n'est pas inclus dans G et que G n'est pas inclus dans F . Il existe alors un vecteur \vec{x} qui est dans F , mais pas dans G , et un vecteur \vec{y} qui est dans G , mais pas dans F . Alors \vec{x} et \vec{y} sont tous les deux dans $F \cup G$. Mais $\vec{x} + \vec{y} \notin F$ (sinon on aurait $\vec{y} = \vec{y} + \vec{x} - \vec{x} \in F$) et $\vec{x} + \vec{y} \notin G$ (sinon on aurait $\vec{x} = \vec{x} + \vec{y} - \vec{y} \in G$). Donc $\vec{x} + \vec{y} \notin F \cup G$, ce qui montre que $F \cup G$ n'est pas un sous-espace vectoriel de E . \square

V.2.c. Sous-espace vectoriel engendré par une famille de vecteurs

On donne maintenant un exemple important de sous-espace vectoriel. Soit $n \geq 1$ et $\vec{u}_1, \dots, \vec{u}_n$ des vecteurs de E (on dit aussi que $(\vec{u}_1, \dots, \vec{u}_n)$ est une famille de vecteurs de E , cf section V.3 plus bas).

Proposition V.2.16. *L'ensemble des combinaisons linéaires des vecteurs $\vec{u}_1, \dots, \vec{u}_n$ de E :*

$$\left\{ \lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n, (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n \right\}$$

est un sous-espace vectoriel de E , appelé espace vectoriel engendré par $\vec{u}_1, \dots, \vec{u}_n$ et noté

$$\text{vect}(\vec{u}_1, \dots, \vec{u}_n) \text{ ou } \text{vect}\{\vec{u}_1, \dots, \vec{u}_n\}.$$

Démonstration. Notons F cet ensemble. On a $\vec{0} = 0\vec{u}_1 + \dots + 0\vec{u}_n \in F$. Il est également très simple de vérifier que F est stable par addition et multiplication par un scalaire, ce qui montre le résultat. \square

Exemple V.2.17.

$$\text{vect}(\vec{0}) = \{\vec{0}\}.$$

Si \vec{u} est un vecteur non nul de E , $\text{vect}(\vec{u})$ est la droite engendrée par \vec{u} (cf exemple V.2.6).

Exemple V.2.18. Soit $\vec{u} = (1, 0, 1)$ et $\vec{v} = (0, 1, 1)$. Le sous-espace vectoriel de \mathbb{R}^3 $\text{vect}(\vec{u}, \vec{v})$ est

$$\{\lambda \vec{u} + \mu \vec{v}, (\lambda, \mu) \in \mathbb{R}^2\} = \{(\lambda, \mu, \lambda + \mu), (\lambda, \mu) \in \mathbb{R}^2\}.$$

C'est exactement le sous-espace vectoriel de \mathbb{R}^3 vu dans l'exemple V.1.7.

L'espace vectoriel engendré par $(\vec{u}_1, \dots, \vec{u}_n)$ est le plus petit espace vectoriel qui contient $(\vec{u}_1, \dots, \vec{u}_n)$:

Proposition V.2.19. *Soit $(\vec{u}_1, \dots, \vec{u}_n)$ des vecteurs d'un espace vectoriel E , et F un sous-espace vectoriel de E tel que*

$$\forall j \in \{1, \dots, n\}, \vec{u}_j \in F.$$

Alors $\text{vect}(\vec{u}_1, \dots, \vec{u}_n) \subset F$.

Démonstration. Soit F un sous-espace vectoriel de E qui contient $\vec{u}_1, \dots, \vec{u}_n$. Par la proposition V.2.9, toute combinaison linéaire d'éléments de F est dans F . Donc

$$\text{vect}(\vec{u}_1, \dots, \vec{u}_n) \subset F.$$

\square

V.2.d. Somme, somme directe, supplémentaires

Somme de deux sous-espaces vectoriels

La démonstration (facile) de la proposition suivante est laissée au lecteur :

Proposition V.2.20. Soit F et G deux sous-espaces vectoriels d'un \mathbb{K} -espace vectoriel E . L'ensemble $H = \{\vec{x} + \vec{y}, \vec{x} \in F, \vec{y} \in G\}$ est un sous-espace vectoriel de E .

Définition V.2.21. Le sous-espace vectoriel H de la proposition précédente est noté $F + G$ et appelé *somme* de F et G .

Exemple V.2.22. Soit $F = \{\vec{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \text{ t.q. } x_1 = 0\}$ et $G = \{\vec{x} \in \mathbb{R}^3 \text{ t.q. } x_3 = 0\}$. Alors

$$F + G = \mathbb{R}^3.$$

En effet, si $\vec{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$, on a :

$$(x_1, x_2, x_3) = \underbrace{(0, x_2, x_3)}_{\in F} + \underbrace{(x_1, 0, 0)}_{\in G},$$

et donc $\vec{x} \in F + G$.

Exemple V.2.23. Soit $n, p \geq 1$ et $\vec{u}_1, \dots, \vec{u}_n, \vec{v}_1, \dots, \vec{v}_p$ des vecteurs de E . Alors

$$\text{vect}(\vec{u}_1, \dots, \vec{u}_n) + \text{vect}(\vec{v}_1, \dots, \vec{v}_p) = \text{vect}(\vec{u}_1, \dots, \vec{u}_n, \vec{v}_1, \dots, \vec{v}_p).$$

En effet, notons F l'espace vectoriel engendré par les vecteurs \vec{u}_j et G l'espace vectoriel engendré par les vecteurs \vec{v}_j . Par définition, $F + G$ est l'ensemble des $\vec{x} + \vec{y}$ avec $\vec{x} \in F$ et $\vec{y} \in G$. En utilisant la définition d'un espace vectoriel engendré, on obtient :

$$F + G = \left\{ \lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n + \mu_1 \vec{v}_1 + \dots + \mu_p \vec{v}_p, (\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_p) \in \mathbb{K}^{n+p} \right\}$$

ce qui donne le résultat annoncé.

Exemple V.2.24. La somme des droites de \mathbb{R}^3 , $\{(x, 0, 0), x \in \mathbb{R}\}$ et $\{(0, y, 0), y \in \mathbb{R}\}$ est le plan de \mathbb{R}^3 :

$$\{(x, y, 0), x \in \mathbb{R}\}.$$

Ceci découle immédiatement de la définition de la somme de deux sous-espaces vectoriels. C'est aussi un cas particulier de l'exemple précédent (avec $n = p = 1$, $\vec{u}_1 = (1, 0, 0)$ et $\vec{v}_1 = (0, 1, 0)$).

Somme directe

Définition V.2.25. Soit F et G deux sous-espaces vectoriels d'un \mathbb{K} -espace vectoriel E . On dit que la somme de F et G est *directe* quand tout élément de $F + G$ s'écrit de manière unique $\vec{x} + \vec{u}$ avec $\vec{x} \in F$ et $\vec{u} \in G$. En d'autres termes :

$$\left(\vec{x} + \vec{u} = \vec{y} + \vec{v}, (\vec{x}, \vec{y}) \in F^2 \text{ et } (\vec{u}, \vec{v}) \in G^2 \right) \implies \left(\vec{x} = \vec{y} \text{ et } \vec{u} = \vec{v} \right).$$

On note alors $F \oplus G$ la somme de F et G .

Proposition V.2.26. La somme $F + G$ est directe si et seulement si $F \cap G = \{\vec{0}\}$.

Démonstration. Supposons que la somme est directe. Soit $\vec{x} \in F \cap G$. Alors

$$\underbrace{\vec{x}}_{\in F} + \underbrace{\vec{0}}_{\in G} = \underbrace{\vec{0}}_{\in F} + \underbrace{\vec{x}}_{\in G},$$

et par unicité de la décomposition d'un vecteur comme somme d'un élément de F et d'un élément de G , on obtient $\vec{x} = \vec{0}$. D'où $F \cap G = \{\vec{0}\}$.

Supposons maintenant $F \cap G = \{\vec{0}\}$. Soit \vec{x} et \vec{y} des vecteurs de F , \vec{u} et \vec{v} des vecteurs de G . On suppose

$$\vec{x} + \vec{u} = \vec{y} + \vec{v}.$$

Alors

$$\vec{x} - \vec{y} = \vec{v} - \vec{u}.$$

Donc $\vec{x} - \vec{y} = \vec{v} - \vec{u} \in F \cap G$ (car $\vec{x} - \vec{y} \in F$ et $\vec{v} - \vec{u} \in G$). Puisque $F \cap G = \{\vec{0}\}$, on en déduit $\vec{x} = \vec{y}$ et $\vec{u} = \vec{v}$, ce qui termine la preuve. \square

Exemple V.2.27. La somme $F + G$ de l'exemple V.2.22 n'est pas directe. En effet, on voit facilement que

$$F \cap G = \{(x_1, x_2, x_3), \text{ t.q. } x_1 = 0 \text{ et } x_3 = 0\} = \text{vect} \{(0, 1, 0)\} \neq \{\vec{0}\}.$$

Exemple V.2.28. La somme des droites de \mathbb{R}^3 , $\{(x, 0, 0), x \in \mathbb{R}\}$ et $\{(0, y, 0), y \in \mathbb{R}\}$ est directe. Le seul point commun à ces droites est bien l'origine $\{\vec{0}\}$.

Supplémentaires

Définition V.2.29. On dit que les sous-espaces vectoriels F et G sont *supplémentaires dans* E lorsque

$$E = F \oplus G.$$

En d'autres termes, la somme de F et G est directe, et égale à E .

Donc par définition, F et G sont supplémentaires dans E si et seulement si tout élément \vec{z} de E s'écrit de manière unique $\vec{z} = \vec{x} + \vec{u}$ avec $\vec{x} \in F$ et $\vec{u} \in G$.

Exemple V.2.30. Les droites $\text{vect}\{(1, 0)\}$ et $\text{vect}\{(0, 1)\}$ sont supplémentaires dans \mathbb{R}^2 .

Exemple V.2.31. Les espaces F et G des exemples V.2.22 et V.2.27 ne sont pas supplémentaires dans \mathbb{R}^3 : leur somme est bien égale à \mathbb{R}^3 , mais elle n'est pas directe.

Les deux droites de \mathbb{R}^3 apparaissant dans l'exemple V.2.28 ne sont pas supplémentaires dans \mathbb{R}^3 : leur somme est directe, mais elle ne vaut pas \mathbb{R}^3 .

Exemple V.2.32. Soit $F = \{(x_1, 0, x_3), (x_1, x_3) \in \mathbb{K}^2\}$ et $G = \text{vect}\{(0, 1, 1)\}$. Vérifier que F et G sont supplémentaires dans \mathbb{R}^3 .

Somme de plusieurs espaces vectoriels

On généralise maintenant ce qui précède au cas de plusieurs espaces vectoriels.

Soit E_1, \dots, E_n des sous-espaces vectoriels de E ($n \geq 2$).

La somme $E_1 + \dots + E_n$ (notée encore $\sum_{j=1}^n E_j$) est le sous-ensemble de E formé des vecteurs $\vec{v}_1 + \dots + \vec{v}_n$, avec $\vec{v}_j \in E_j$ pour $j = 1 \dots n$. On montre facilement que c'est un sous-espace vectoriel de E .

V. Espaces vectoriels

On dit que cette somme est *directe* (et on la note $E_1 \oplus \dots \oplus E_n$) lorsque l'écriture $\vec{v} = \vec{v}_1 + \dots + \vec{v}_n$ avec $\vec{v}_j \in E_j$ pour tout j est unique, i.e. lorsque

$$\left(\vec{v}_1 + \dots + \vec{v}_n = \vec{u}_1 + \dots + \vec{u}_n \text{ et } \forall j \in \{1, \dots, n\}, \vec{v}_j \in E_j, \vec{u}_j \in E_j \right) \\ \implies \forall j \in \{1, \dots, n\}, \vec{v}_j = \vec{u}_j.$$

Cette condition est équivalente (le vérifier!) à

$$\left(\vec{v}_1 + \dots + \vec{v}_n = \vec{0} \text{ et } \forall j \in \{1, \dots, n\}, \vec{v}_j \in E_j \right) \implies \forall j \in \{1, \dots, n\}, \vec{v}_j = \vec{0}.$$

Si la somme est directe, on a $j \neq k \implies E_j \cap E_k = \{\vec{0}\}$, mais cette condition n'est pas suffisante dès que $n \geq 3$ (cf exemple V.2.33 ci-dessous).

Comme dans le cas $n = 2$, on dit que E_1, \dots, E_n sont *supplémentaires* lorsque leur somme est directe et vaut E . Ainsi, les espaces $(E_j)_{j=1 \dots n}$ sont supplémentaires si et seulement si tout élément \vec{v} de E s'écrit de manière unique $\vec{v} = \sum_{j=1}^n \vec{v}_j$, avec $\vec{v}_j \in E_j$ pour tout j .

Exemple V.2.33. On considère

$$P = \{(x, y, z) \in \mathbb{R}^3 \text{ t.q. } x = y\}, \quad D_1 = \text{vect}\{(1, 0, 1)\}, \quad D_2 = \text{vect}\{(0, 1, 0)\}.$$

Alors $P + D_1 + D_2 = \mathbb{R}^3$. En effet, $(x, y, z) \in \mathbb{R}^3$ s'écrit :

$$(x, y, z) = \underbrace{(z - x)(0, 0, 1)}_{\in P} + \underbrace{x(1, 0, 1)}_{\in D_1} + \underbrace{y(0, 1, 0)}_{\in D_2}.$$

On a $P \cap D_1 = \{\vec{0}\}$, $P \cap D_2 = \{\vec{0}\}$ et $D_1 \cap D_2 = \{\vec{0}\}$, mais la somme n'est pas directe :

$$\underbrace{(1, 1, 1)}_{\in P} - \underbrace{(1, 0, 1)}_{\in D_1} - \underbrace{(0, 1, 0)}_{\in D_2} = (0, 0, 0).$$

Exemple V.2.34. Soit $n \geq 1$. Notons \vec{e}_j le vecteur de \mathbb{K}^n dont toutes les coordonnées sont nulles, sauf la j -ième qui vaut 1. Alors les n droites $\text{vect}(\vec{e}_j)$, $j = 1 \dots n$ sont supplémentaires dans \mathbb{K}^n .

V.3. Familles de vecteurs

Dans toute cette partie, E désigne un \mathbb{K} -espace vectoriel. On étudie ici certaines propriétés des familles (finies) de vecteurs de E , qui seront utiles pour traiter, dans la partie V.4, la notion de dimension d'un espace vectoriel.

V.3.a. Familles de vecteurs : définition

Définition V.3.1. Soit n un entier ≥ 1 . Une *famille* (finie) \mathcal{F} de n vecteurs de E est un n -uplet $(\vec{u}_1, \dots, \vec{u}_n)$ de vecteurs de E . On note $\vec{u} \in \mathcal{F}$ quand \vec{u} est un des vecteurs \vec{u}_j . Le nombre n est le *cardinal* de \mathcal{F} , et on note $n = |\mathcal{F}|$. On convient qu'il existe une seule famille de cardinal 0, notée \emptyset .

Soit $\mathcal{F} = (\vec{u}_1, \dots, \vec{u}_n)$ et $\mathcal{G} = (\vec{v}_1, \dots, \vec{v}_p)$ deux familles de vecteurs. On note $\mathcal{F} \cup \mathcal{G} = (\vec{u}_1, \dots, \vec{u}_n, \vec{v}_1, \dots, \vec{v}_p)$.

Si $\mathcal{F} = (\vec{u}_1, \dots, \vec{u}_n)$ est une famille de vecteurs, et \mathcal{G} est une autre famille de la forme $(\vec{u}_{k_1}, \dots, \vec{u}_{k_p})$ avec $1 \leq k_1 < k_2 < \dots < k_p \leq n$, on dit que \mathcal{G} est *extraite* de \mathcal{F} , et on note $\mathcal{G} \subset \mathcal{F}$. On dit aussi que la famille \mathcal{F} *complète* la famille \mathcal{G} .

Exemple V.3.2. Soit

$$\mathcal{F} = \left((1, 0, 0), (0, 1, 0), (0, 1, 0), (1, 0, 1) \right) \text{ et } \mathcal{G} = \left((1, 0, 0), (0, 1, 0), (1, 0, 1) \right).$$

Alors \mathcal{F} et \mathcal{G} sont des familles de vecteurs de \mathbb{R}^3 , $|\mathcal{F}| = 4$, $|\mathcal{G}| = 3$, et \mathcal{G} est extraite de \mathcal{F} .

V.3.b. Familles libres

Définition V.3.3. Une famille $\mathcal{F} = (\vec{u}_1, \dots, \vec{u}_n)$ de E est *libre* si la propriété suivante est vérifiée :

$$\forall (x_1, \dots, x_n) \in \mathbb{K}^n, \quad x_1\vec{u}_1 + x_2\vec{u}_2 + \dots + x_n\vec{u}_n = \vec{0} \implies x_1 = x_2 = \dots = x_n = 0.$$

On dit aussi que les vecteurs $\vec{u}_1, \dots, \vec{u}_n$ sont *linéairement indépendants*. Si la famille \mathcal{F} n'est pas libre, on dit qu'elle est *liée*.

Exemple V.3.4. La famille² $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ est libre dans \mathbb{R}^3 . En effet, supposons

$$x_1 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

i.e $x_2 = 0$, $x_1 + x_2 = 0$ et $x_1 + 2x_2 = 0$. Alors $x_1 = x_2 = 0$.

On voit sur cet exemple que montrer qu'une famille de p vecteurs de \mathbb{K}^n est libre revient à montrer qu'un certain système linéaire homogène à p inconnues et n équations a pour seule solution la solution nulle.

Exemple V.3.5. Soit $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ et $\vec{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. La famille $\mathcal{G} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$ est libre dans \mathbb{C}^3 . Comme dans l'exemple précédent, on est ramené à étudier un système homogène. Soit $(x_1, x_2, x_3) \in \mathbb{C}^3$ tels que $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 = \vec{0}$. Alors

$$x_1 + x_2 + x_3 = 0, \quad x_2 = 0 \text{ et } x_1 = 0,$$

ce qui donne facilement $x_1 = x_2 = x_3 = 0$.

Exemple V.3.6. Soit $\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\vec{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ et $\vec{u}_3 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$. Alors la famille $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ est liée.

En effet :

$$\vec{u}_1 + \vec{u}_2 - \vec{u}_3 = \vec{0}.$$

(Si l'on ne voit pas immédiatement cette relation, on peut la retrouver en résolvant le système $x_1\vec{u}_1 + x_2\vec{u}_2 + x_3\vec{u}_3 = \vec{0}$, d'inconnues x_1, x_2 et x_3).

Exemple V.3.7. Si $\vec{0} \in \mathcal{F}$, alors \mathcal{F} est liée. En effet $1 \cdot \vec{0} = \vec{0}$, et $1 \neq 0$.

2. On rappelle que l'on identifie un élément de \mathbb{K}^n à une matrice colonne. Dans ce chapitre on utilisera donc indifféremment la notation (x_1, \dots, x_n) ou la notation matricielle en colonne pour un élément de \mathbb{K}^n .

V. Espaces vectoriels

Exemple V.3.8. Une famille à un élément \vec{u} est libre si et seulement si $\vec{u} \neq \vec{0}$. En effet, si $\vec{u} = \vec{0}$ la famille n'est pas libre (cf exemple précédent). En revanche, si $\vec{u} \neq \vec{0}$, alors, par la Proposition V.1.3

$$\lambda \vec{u} = \vec{0} \implies \lambda = 0$$

ce qui montre que la famille (\vec{u}) est libre.

La proposition suivante découle immédiatement de la définition d'une famille libre :

Proposition V.3.9. *Si \mathcal{F} est une famille libre, toute famille extraite de \mathcal{F} est libre.*

On termine cette partie sur les familles libres par le *lemme utile sur les familles libres* suivant :

Lemme V.3.10. *Soit \mathcal{F} une famille libre et $\vec{v} \in E$. Alors la famille $\mathcal{F} \cup (\vec{v})$ est libre si et seulement si $\vec{v} \notin \text{vect } \mathcal{F}$.*

Démonstration. On note $\mathcal{F} = (\vec{u}_1, \dots, \vec{u}_n)$. On rappelle la définition de l'espace vectoriel engendré par \mathcal{F} (cf §V.2.c) :

$$\text{vect } \mathcal{F} = \left\{ x_1 \vec{u}_1 + \dots + x_n \vec{u}_n, (x_1, \dots, x_n) \in \mathbb{K}^n \right\}.$$

Supposons d'abord que $\mathcal{F} \cup (\vec{v})$ est libre. On a

$$(V.5) \quad \forall (x_1, \dots, x_n) \in \mathbb{K}^n, \quad \sum_{j=1}^n x_j \vec{u}_j - \vec{v} \neq \vec{0}.$$

En effet, c'est une combinaison linéaire d'éléments de la famille libre $\mathcal{F} \cup (\vec{v})$ avec au moins un des coefficients (celui de \vec{v}) non nul. Par (V.5), $\vec{v} \notin \text{vect } \mathcal{F}$.

Réciproquement, on suppose $\vec{v} \notin \text{vect } \mathcal{F}$. Soit $(x_1, \dots, x_n, y) \in \mathbb{K}^{n+1}$. On suppose

$$\sum_{j=1}^n x_j \vec{u}_j + y \vec{v} = \vec{0}$$

Alors $y = 0$: sinon on aurait $\vec{v} = -\frac{1}{y} \sum_{j=1}^n x_j \vec{u}_j \in \text{vect } \mathcal{F}$. Donc

$$\sum_{j=1}^n x_j \vec{u}_j = \vec{0},$$

ce qui implique, la famille étant libre, $x_1 = x_2 = \dots = x_n = 0$. On a montré que la famille $(\vec{u}_1, \dots, \vec{u}_n, \vec{v})$ était libre, ce qui conclut la preuve du lemme. \square

Exemple V.3.11. Une famille de deux vecteurs (\vec{u}, \vec{v}) est libre si et seulement si $\vec{u} \neq \vec{0}$ et \vec{v} n'appartient pas à la droite $\text{vect } \vec{u}$. Ceci découle immédiatement des deux propositions précédentes.

V.3.c. Familles génératrices

Définition V.3.12. La famille $\mathcal{F} = (\vec{u}_1, \dots, \vec{u}_n)$ est une *famille génératrice de E* (ou simplement *est génératrice* quand il n'y a pas d'ambiguïté) quand pour tout vecteur \vec{v} de E , il existe $(x_1, \dots, x_n) \in \mathbb{K}^n$ tel que

$$\vec{v} = x_1\vec{u}_1 + \dots + x_n\vec{u}_n.$$

On dit aussi que \mathcal{F} engendre E .

Remarque V.3.13. La famille \mathcal{F} est génératrice si et seulement si $\text{vect } \mathcal{F} = E$.

Exemple V.3.14. La famille $\mathcal{F}_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ est une famille génératrice de \mathbb{R}^2 : si $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$, $\begin{bmatrix} x \\ y \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, et donc $\mathbb{R}^2 = \text{vect } \mathcal{F}_1$.

Exemple V.3.15. La famille $\mathcal{F}_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ n'est pas génératrice de \mathbb{R}^3 . En effet, la

dernière coordonnée de tout élément de $\text{vect } \mathcal{F}$ est nulle, et donc par exemple $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ n'appartient pas à $\text{vect } \mathcal{F}_2$.

Exemple V.3.16. La famille \mathcal{G} de l'exemple V.3.5 p. 83 est une famille génératrice de \mathbb{C}^3 . En effet, il s'agit de montrer que pour tout $(b_1, b_2, b_3) \in \mathbb{C}^3$, il existe (x_1, x_2, x_3) tel que

$$x_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

ou encore :

$$x_1 + x_2 + x_3 = b_1, \quad x_2 = b_2, \quad x_1 = b_3.$$

Il est facile de résoudre ce système. On peut aussi remarquer que c'est un système de Cramer par l'exemple V.3.5 : il a 3 équations, 3 inconnues, et une seule solution lorsque $b_1 = b_2 = b_3 = 0$. Il a donc une unique solution quel que soit (b_1, b_2, b_3) .

Plus généralement, montrer qu'une famille de p vecteurs $(\vec{v}_1, \dots, \vec{v}_p)$ de \mathbb{K}^n est génératrice revient à montrer, pour tout $\vec{b} \in \mathbb{K}^n$, la compatibilité du système à n équations et p inconnues $(x_1, \dots, x_p) : x_1\vec{v}_1 + \dots + x_p\vec{v}_p = \vec{b}$.

Le résultat suivant est une conséquence directe de la définition d'une famille génératrice :

Proposition V.3.17. *Soit \mathcal{F} une famille génératrice. Alors toute famille qui complète \mathcal{F} est encore génératrice.*

V.3.d. Bases

Définition V.3.18. Une famille \mathcal{F} de E est une *base* quand elle est à la fois libre et génératrice.

Exemple V.3.19. La famille $\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$ est une base de \mathbb{K}^2 . Plus généralement, si $n \geq 1$ et $j \in \{1, \dots, n\}$, définissons \vec{e}_j comme le vecteur de \mathbb{K}^n dont toutes les coordonnées sont nulles, sauf la j -ième, qui vaut 1. Alors la famille $(\vec{e}_1, \dots, \vec{e}_n)$ est une base de \mathbb{K}^n , appelée *base canonique de \mathbb{K}^n* .

V. Espaces vectoriels

Exemple V.3.20. La famille \mathcal{G} de l'exemple V.3.5 p. 83 est une base de \mathbb{C}^3 : on a montré qu'elle était libre (exemple V.3.5) et génératrice (exemple V.3.16).

Exemple V.3.21. La famille $\left(\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right)$ n'est pas une base de \mathbb{R}^3 : elle est libre mais pas génératrice (justifier).

Exemple V.3.22. La famille $\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right)$ n'est pas une base de \mathbb{R}^3 : elle est génératrice mais pas libre (justifier).

Proposition et définition V.3.23. Soit $\mathcal{B} = (\vec{e}_1, \dots, \vec{e}_n)$ une base de E . Alors tout vecteur \vec{x} de E s'écrit de manière unique

$$\vec{x} = \sum_{j=1}^n x_j \vec{e}_j.$$

Les x_j sont appelés *coordonnées* de \vec{x} dans la base E .

Démonstration. La famille \mathcal{B} étant génératrice, tout vecteur \vec{x} de E s'écrit

$$\vec{x} = \sum_{j=1}^n x_j \vec{e}_j.$$

Montrons l'unicité de cette écriture. Supposons $(x_1, \dots, x_n) \in \mathbb{K}^n$, $(y_1, \dots, y_n) \in \mathbb{K}^n$ et

$$\sum_{j=1}^n y_j \vec{e}_j = \vec{x} = \sum_{j=1}^n x_j \vec{e}_j.$$

Alors

$$\sum_{j=1}^n (x_j - y_j) \vec{e}_j = \vec{0}.$$

La famille \mathcal{B} étant libre, on en déduit $x_j - y_j = 0$ pour tout $j \in \{1, \dots, n\}$, ce qui conclut la preuve. \square

Exemple V.3.24. Les coordonnées du vecteur (x_1, \dots, x_n) de \mathbb{K}^n dans la base canonique de \mathbb{K}^n sont (x_1, \dots, x_n) .

V.4. Espaces vectoriels de dimension finie

V.4.a. Définition

Définition V.4.1. On dit que E est de *dimension finie* quand il admet une famille génératrice finie. On convient que l'espace vectoriel $\{\vec{0}\}$ est de dimension finie, et que la famille vide est une famille génératrice de $\{\vec{0}\}$.

Nous verrons plus loin que tous les espaces vectoriels étudiés dans ce cours (les sous-espaces vectoriels de \mathbb{K}^n) sont de dimension finie. Mentionnons toutefois qu'il existe des espaces vectoriels (au sens de la définition générale V.1.4) qui sont de dimension infinie : c'est le cas par exemple de l'espace vectoriel des polynômes sur \mathbb{K} . Ce type d'espace vectoriel n'est pas au programme de la première année de licence à l'Institut Galilée.

Exemple V.4.2. L'espace vectoriel \mathbb{K}^n est de dimension finie : la base canonique de \mathbb{K}^n est une famille génératrice finie.

V.4.b. Existence de bases

Proposition V.4.3. *Soit E un espace vectoriel de dimension finie. Alors E admet une base. Plus précisément, de toute famille génératrice \mathcal{F} on peut extraire une base.*

Démonstration. Le cas $E = \{\vec{0}\}$ est immédiat. On suppose donc $E \neq \{\vec{0}\}$. Si $n \geq 1$, on note (\mathcal{P}_n) la propriété suivante :

De toute famille génératrice de E de cardinal n on peut extraire une base de E .

(Si il n'existe aucune famille génératrice finie de cardinal n , on convient que la propriété (\mathcal{P}_n) est vraie.) On va montrer que (\mathcal{P}_n) est vraie pour tout n , par récurrence sur n .

Montrons (\mathcal{P}_1) . Si E n'admet pas de famille génératrice à 1 élément, (\mathcal{P}_1) est vrai. Dans le cas contraire, on se donne $\mathcal{G} = (\vec{u})$ une famille génératrice de E à 1 élément. Puisque $E \neq \{\vec{0}\}$, $\vec{u} \neq \vec{0}$. Donc (\vec{u}) est libre (cf exemple V.3.8), ce qui montre que \mathcal{G} est une base.

Soit $n \geq 1$. Supposons (\mathcal{P}_n) et montrons (\mathcal{P}_{n+1}) . Soit $\mathcal{G} = (\vec{u}_1, \dots, \vec{u}_{n+1})$ une famille génératrice de E à $n+1$ éléments (là encore, si une telle famille n'existe pas, (\mathcal{P}_{n+1}) est automatiquement vraie). Si \mathcal{G} est libre, c'est une base (extraite de \mathcal{G}). Sinon, il existe $n+1$ scalaires $\lambda_1, \dots, \lambda_{n+1}$, non tous nuls, tels que

$$\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \dots + \lambda_{n+1} \vec{u}_{n+1} = \vec{0}.$$

Soit $j \in \{1, \dots, n+1\}$ tel que $\lambda_j \neq 0$. On a alors :

$$\vec{u}_j = -\frac{1}{\lambda_j} \sum_{\substack{1 \leq k \leq n+1 \\ k \neq j}} \lambda_k \vec{u}_k.$$

Soit \mathcal{G}' la famille obtenue à partir de \mathcal{G} en retirant \vec{u}_j . L'égalité précédente montre que $\text{vect } \mathcal{G} = \text{vect } \mathcal{G}'$, et donc que \mathcal{G}' est aussi génératrice. Mais $|\mathcal{G}'| = n$. Par (\mathcal{P}_n) , on peut extraire une base \mathcal{B} de \mathcal{G}' . La base \mathcal{B} est aussi extraite de \mathcal{G} , ce qui conclut la preuve. \square

V.4.c. Dimension d'un espace vectoriel.

On note E un \mathbb{K} -espace vectoriel de dimension finie. On rappelle que le *cardinal* d'une famille \mathcal{F} , noté $|\mathcal{F}|$, est le nombre d'éléments de \mathcal{F} .

Le théorème suivant est crucial pour définir la dimension d'un espace vectoriel. On utilise pour le démontrer un résultat du chapitre III du cours sur les systèmes linéaires.

Théorème V.4.4. *Soit \mathcal{F} une famille libre et \mathcal{G} une famille génératrice de E . Alors $|\mathcal{F}| \leq |\mathcal{G}|$.*

V. Espaces vectoriels

Démonstration. On note $\mathcal{F} = (\vec{u}_1, \dots, \vec{u}_n)$ et $\mathcal{G} = (\vec{e}_1, \dots, \vec{e}_p)$. On veut montrer $p \geq n$. On raisonne par l'absurde.

Supposons $n > p$. Puisque \mathcal{G} est une famille génératrice de E , il existe, pour tout indice $j \in \{1, \dots, n\}$, p scalaires a_{1j}, \dots, a_{pj} tels que

$$(V.6) \quad \vec{u}_j = \sum_{i=1}^p a_{ij} \vec{e}_i.$$

Considérons le système homogène

$$\forall i = 1 \dots p, \quad \sum_{j=1}^n a_{ij} x_j = 0$$

qui a p équations et n inconnues x_1, \dots, x_n (donc strictement plus d'inconnues que d'équations). Ce système est équivalent, par la méthode du pivot du chapitre III, à un système homogène (donc compatible) sous forme échelonnée réduite ayant n inconnues et $p' \leq p < n$ lignes non nulles : il a donc une infinité de solutions, que l'on peut décrire par $n - p'$ paramètres. Notons (x_1, \dots, x_n) une solution non nulle de ce système. Alors, d'après (V.6),

$$x_1 \vec{u}_1 + \dots + x_n \vec{u}_n = \sum_{j=1}^n \sum_{i=1}^p x_j a_{ij} \vec{e}_i = \sum_{i=1}^p \left(\sum_{j=1}^n x_j a_{ij} \right) \vec{e}_i = \vec{0},$$

car les p termes entre parenthèse dans la somme précédente sont nulles, par définitions des x_j . Puisque $(\vec{u}_1, \dots, \vec{u}_n)$ est libre, on doit avoir $x_1 = \dots = x_n = 0$, une contradiction. \square

On peut maintenant définir la dimension d'un espace vectoriel de dimension finie :

Théorème et définition V.4.5. Soit E un espace vectoriel de dimension finie. Alors :

- i. Toutes les bases de E ont le même cardinal, appelé *dimension* de E et noté $\dim E$.
- ii. Le cardinal de toute famille libre de E est inférieur ou égal à $\dim E$.
- iii. Le cardinal de toute famille génératrice de E est supérieur ou égal à $\dim E$.

Démonstration. Les trois points sont conséquences du théorème V.4.4

Soit \mathcal{B} et \mathcal{B}' deux bases de E . Puisque \mathcal{B} est libre et \mathcal{B}' génératrice, le théorème V.4.4 implique $|\mathcal{B}| \leq |\mathcal{B}'|$. De plus, \mathcal{B}' est libre et \mathcal{B} est génératrice, donc $|\mathcal{B}'| \leq |\mathcal{B}|$. D'où (i).

Les points (ii) et (iii) découlent immédiatement du théorème V.4.4, en utilisant encore qu'une base est une famille libre et génératrice. \square

Exemples V.4.6. L'espace vectoriel $\{\vec{0}\}$ est de dimension 0 (il a pour base l'ensemble vide \emptyset par convention). L'espace vectoriel \mathbb{K}^n est de dimension n sur \mathbb{K} : la base canonique a n éléments.

Soit E un espace vectoriel et \vec{u} un vecteur non nul de E . Alors l'espace vectoriel $\text{vect}(\vec{u})$ est de dimension 1 : il a pour base \vec{u} .

La famille de \mathbb{C}^3 :

$$\mathcal{F} = \left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ i \end{bmatrix}, \begin{bmatrix} 2 \\ i \\ 2 \end{bmatrix}, \begin{bmatrix} \sqrt{3} \\ i+2 \\ 3 \end{bmatrix} \right)$$

n'est pas une famille libre du \mathbb{C} -espace vectoriel \mathbb{C}^3 d'après le point (ii) : cette famille a 4 éléments, alors que \mathbb{C}^3 est de dimension 3.

La famille de \mathbb{R}^4 :

$$\mathcal{G} = \left(\begin{bmatrix} 1 \\ 2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix} \right)$$

n'est pas une famille génératrice de \mathbb{R}^4 .

Exemple V.4.7. Soit \mathcal{F} une famille libre d'un espace vectoriel E . Alors $\text{vect } \mathcal{F}$ est un sous-espace vectoriel de E , qui est de dimension finie : il découle immédiatement des définitions que \mathcal{F} est une base de $\text{vect } \mathcal{F}$.

V.4.d. Caractérisation des bases

Théorème V.4.8. Soit E un espace vectoriel de dimension finie. Soit \mathcal{F} une famille de vecteurs de E . Les conditions suivantes sont équivalentes :

- i. \mathcal{F} est une base de E .
- ii. \mathcal{F} est une famille libre, et $|\mathcal{F}| = \dim E$.
- iii. \mathcal{F} est une famille génératrice, et $|\mathcal{F}| = \dim E$.

Démonstration. Les implications (i) \implies (ii) et (i) \implies (iii) découlent de la définition d'une base et du théorème/définition V.4.5.

Notons $n = \dim E$.

Montrons (iii) \implies (i). Soit \mathcal{F} une famille génératrice à n éléments. Alors, par la proposition V.4.3, il existe une base \mathcal{B} de E extraite de \mathcal{F} . Mais $|\mathcal{B}| = n = |\mathcal{F}|$ par le théorème/définition V.4.5. Donc $\mathcal{B} = \mathcal{F}$, et \mathcal{F} est une base.

Montrons maintenant (ii) \implies (i). Soit $\mathcal{F} = (\vec{u}_1, \dots, \vec{u}_n)$ une famille libre de E à n éléments. Montrons que \mathcal{F} est génératrice. Soit $\vec{x} \in E$. Par le théorème V.4.5, la famille $\mathcal{F} \cup (\vec{x})$ n'est pas libre (elle a $n + 1$ éléments). Par le *lemme utile sur les familles libres* p.84, $\vec{x} \in \text{vect } \mathcal{F}$. On a bien montré que \mathcal{F} est génératrice, ce qui termine la preuve. \square

Remarque V.4.9. Soit E un espace vectoriel de dimension finie n . On se donne une famille \mathcal{F} de vecteurs de E , de cardinal n . Alors \mathcal{F} est une base de E si et seulement si \mathcal{F} est libre. Dans le cas $E = \mathbb{K}^n$, il suffit donc, pour déterminer si \mathcal{F} est une base, de savoir si un système homogène de n équations à n inconnues est un système de Cramer (ou encore si la matrice des coefficients du système est inversible). On évite ainsi de résoudre un système non-homogène (ce que l'on doit faire pour montrer directement qu'une famille de \mathbb{K}^n engendre \mathbb{K}^n). Le lecteur est par exemple invité à montrer que

$$\left(\begin{bmatrix} -1 \\ -1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -2 \\ 3 \\ -2 \end{bmatrix} \right)$$

est une base de \mathbb{R}^4 .

V.4.e. Théorème de la base incomplète

On termine cette section sur la dimension finie par un résultat important, le théorème de la base incomplète :

Théorème V.4.10. *Soit E un espace vectoriel de dimension finie n . Soit $\mathcal{L} = (\vec{u}_1, \dots, \vec{u}_p)$ une famille libre de E , de cardinal p . Alors $p \leq n$ et on peut compléter \mathcal{L} en une base $(\vec{u}_1, \dots, \vec{u}_p, \vec{u}_{p+1}, \dots, \vec{u}_n)$ de E .*

Démonstration. L'inégalité $p \leq n$ découle du théorème V.4.5. Montrons par récurrence descendante sur $p \in \{1, \dots, n\}$ que toute famille libre de cardinal p peut être complétée en une base.

C'est vrai lorsque $p = n$: par le théorème V.4.8, une famille libre de cardinal n est une base.

Soit $p \in \{1, \dots, n-1\}$. Supposons le résultat vrai pour les familles libres de cardinal $p+1$. Soit $\mathcal{L} = (\vec{u}_1, \dots, \vec{u}_p)$ une famille libre de cardinal p . Puisque $p < n$, \mathcal{L} n'est pas une base. Elle n'est donc pas génératrice, et il existe $\vec{u}_{p+1} \in E$ tel que $\vec{u}_{p+1} \notin \text{vect } \mathcal{L}$. Par le *lemme utile sur les familles libres* p.84, la famille $(\vec{u}_1, \dots, \vec{u}_p, \vec{u}_{p+1})$ est libre. Par hypothèse de récurrence, on peut la compléter en une base de E , ce qui termine la preuve. \square

Exemple V.4.11. La famille libre $((1, 0, 0), (1, 1, 0))$ de \mathbb{R}^3 peut-être complétée en une base $((1, 0, 0), (1, 1, 0), (0, 0, 1))$ de \mathbb{R}^3 . Nous verrons en §V.5.d une méthode systématique pour compléter une famille libre en une base.

V.5. Sous-espaces vectoriels et dimension

V.5.a. Dimension d'un sous-espace vectoriel

Théorème V.5.1. *Soit E un \mathbb{K} -espace vectoriel de dimension finie, et F un sous-espace vectoriel de E . Alors F est de dimension finie et $\dim F \leq \dim E$.*

Démonstration. Le principe de la preuve est de trouver une base de F , en choisissant une famille libre maximale de F . On suppose E non réduit à $\{\vec{0}\}$, sinon le résultat est trivial.

Soit $n = \dim E \geq 1$. On commence par remarquer que toute famille libre de F est de cardinal $\leq n$. En effet, une telle famille est aussi une famille libre de E , qui est de dimension n , et le résultat découle du théorème V.4.5, (ii).

Soit

$$p = \max \left\{ |\mathcal{L}|, \mathcal{L} \text{ famille libre de } F \right\} \leq n.$$

L'entier p est bien défini (c'est le maximum d'une famille non vide d'entiers majorée par n). Soit \mathcal{L} une famille libre de F , de cardinal p . Montrons que \mathcal{L} engendre F . On en déduira que \mathcal{L} est une base de F , et donc que F est de dimension finie $p \leq \dim E$.

On note $\mathcal{L} = \{u_1, \dots, u_p\}$. Soit $\vec{v} \in F$. La famille $(\vec{u}_1, \dots, \vec{u}_p, \vec{v})$ n'est pas libre (car c'est une famille de cardinal $p+1 \geq p$). Par le *lemme utile sur les familles libres* p.84, $\vec{v} \in \text{vect}(\vec{u}_1, \dots, \vec{u}_p)$.

Ceci montre que \mathcal{L} engendre F et donc, comme annoncé, que \mathcal{L} est une base de F . \square

Exemple V.5.2. Tout sous-espace vectoriel de \mathbb{K}^n est de dimension finie, inférieure ou égale à n . En d'autres termes, tous les espaces vectoriels étudiés dans ce cours sont de dimension finie.

Exemple V.5.3. Le sous-espace vectoriel F de \mathbb{R}^3 d'équation $x = y$ est un sous-espace vectoriel de dimension 2 (un plan) de \mathbb{R}^3 . En écrivant l'ensemble des solutions de cette équation sous forme paramétrique, on obtient

$$F = \{(x, x, z), (x, z) \in \mathbb{R}^2\} = \{x(1, 1, 0) + z(0, 0, 1), (x, z) \in \mathbb{R}^2\} = \text{vect}((1, 1, 0), (0, 0, 1)).$$

La famille $((1, 1, 0), (0, 0, 1))$ engendre F . Puisque c'est une famille libre, c'est une base de F , ce qui montre le résultat annoncé.

Le seul sous-espace vectoriel de E de dimension $\dim E$ est E lui-même :

Proposition V.5.4. *Soit E un espace vectoriel de dimension finie et F un sous-espace vectoriel de même dimension $\dim E$. Alors $E = F$.*

Démonstration. Soit \mathcal{B} une base de F . Alors \mathcal{B} est une famille libre de E , de dimension $\dim E$. Par le théorème V.4.8, \mathcal{B} est une base de E . Donc $\dim E = |\mathcal{B}| = \dim F$. \square

Définition V.5.5. On appelle *rang* d'une famille de vecteurs la dimension du sous-espace vectoriel engendré par cette famille.

V.5.b. Dimension de la somme de deux sous-espaces vectoriels

Théorème V.5.6. *Soit E de dimension finie, F et G des sous-espaces vectoriels de E . Alors*

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

Corollaire V.5.7. *Sous les hypothèses du théorème, si la somme $F \oplus G$ est directe, $\dim(F + G) = \dim F + \dim G$. Si de plus F et G sont supplémentaires dans E , $\dim F + \dim G = \dim E$.*

Preuve du théorème V.5.6. On note p la dimension de F , q celle de G et k celle de $F \cap G$. On sait (cf Théorème V.5.1), que $k \leq p$ et $k \leq q$. On se donne une base $(\vec{e}_1, \dots, \vec{e}_k)$ de $F \cap G$. Par le théorème de la base incomplète, on peut compléter cette famille libre en une base $(\vec{e}_1, \dots, \vec{e}_k, \vec{f}_{k+1}, \dots, \vec{f}_p)$ de F et une base $(\vec{e}_1, \dots, \vec{e}_k, \vec{g}_{k+1}, \dots, \vec{g}_q)$ de G . Il suffit de montrer que $\mathcal{A} = (\vec{e}_1, \dots, \vec{e}_k, \vec{f}_{k+1}, \dots, \vec{f}_p, \vec{g}_{k+1}, \dots, \vec{g}_q)$ est une base de $F + G$: on aurait alors que $\dim(F + G)$ est égal au cardinal de \mathcal{A} , soit $k + (p - k) + (q - k) = p + q - k$ comme annoncé.

La famille \mathcal{A} engendre $F + G$: tout vecteur de $F + G$ s'écrit $\vec{v} = \vec{f} + \vec{g}$ avec $\vec{f} \in F = \text{vect}(\vec{e}_1, \dots, \vec{e}_k, \vec{f}_{k+1}, \dots, \vec{f}_p)$ et $\vec{g} \in G = \text{vect}(\vec{e}_1, \dots, \vec{e}_k, \vec{g}_{k+1}, \dots, \vec{g}_q)$.

Il reste à prouver que la famille \mathcal{A} est libre. On se donne des scalaires $(x_i)_{i=1, \dots, k}$, $(y_i)_{i=k+1, \dots, p}$ et $(z_i)_{i=k+1, \dots, q}$ tels que

$$(V.7) \quad \sum_{i=1}^k x_i \vec{e}_i + \sum_{i=k+1}^p y_i \vec{f}_i + \sum_{i=k+1}^q z_i \vec{g}_i = \vec{0}.$$

On en déduit $\sum_{i=k+1}^q z_i \vec{g}_i = -\sum_{i=1}^k x_i \vec{e}_i - \sum_{i=k+1}^p y_i \vec{f}_i \in F \cap G$. Notons (t_1, \dots, t_k) les coordonnées de ce vecteur dans la base $(e_i)_{i=1 \dots k}$ de $F \cap G$. On a donc, par (V.7),

$$\sum_{i=1}^k (x_i + t_i) \vec{e}_i + \sum_{i=k+1}^p y_i \vec{f}_i = \vec{0},$$

ce qui implique, la famille $(\vec{e}_1, \dots, \vec{e}_k, \vec{f}_{k+1}, \dots, \vec{f}_p)$ étant libre, que $y_{k+1} = \dots = y_p = 0$. En revenant à (V.7), on obtient

$$\sum_{i=1}^k x_i \vec{e}_i + \sum_{i=k+1}^q z_i \vec{g}_i = \vec{0},$$

ce qui montre, en utilisant que la famille $(\vec{e}_1, \dots, \vec{e}_k, \vec{g}_{k+1}, \dots, \vec{g}_q)$ est libre, que $x_1 = \dots = x_k = z_{k+1} = \dots = z_q = 0$. Finalement, tous les coefficients de la combinaison linéaire (V.7) sont bien nuls, ce qui montre comme annoncé que la famille \mathcal{A} est libre. \square

Exemple V.5.8. Le théorème V.5.6 donne une information “gratuite” (la dimension) pour calculer $F + G$. Considérons par exemple les deux sous-espaces vectoriels de \mathbb{R}^3 : $F = \{(x, y, z) \in \mathbb{R}^3 \text{ t.q. } x = y\}$ et $G = \text{vect}\{(1, 1, 0), (1, 0, 1)\}$. Montrons que $F + G = \mathbb{R}^3$.

Pour cela, on remarque que $\dim F = 2$ (exemple V.5.3), $\dim G = 2$ (G est engendré par une famille libre de dimension 2). De plus $\dim(F \cap G) = 1$. En effet, si $\vec{x} = (x, y, z) \in G$, alors \vec{x} s'écrit $\lambda(1, 1, 0) + \mu(1, 0, 1)$. De plus, $\vec{x} \in F \iff x = y \iff \mu = 0$. Donc $F \cap G = \text{vect}\{(1, 1, 0)\}$ est bien de dimension 1. Par le théorème V.5.6,

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G) = 2 + 2 - 1 = 3 = \dim \mathbb{R}^3.$$

Par la proposition V.5.4, $F + G = \mathbb{R}^3$.

V.5.c. Description des sous-espaces vectoriels de \mathbb{K}^n

On connaît deux façons de décrire un sous-espace vectoriel F de \mathbb{K}^n : comme l'espace vectoriel engendré par une de ses bases, ou comme l'ensemble des solutions d'un système linéaire homogène (on parle dans ce deuxième cas de *description par des équations cartésiennes*, ou simplement de *description cartésienne* de F). On explique ici comment passer d'une de ces écritures à l'autre.

Passer d'un système d'équations à une base

Soit (S) un système linéaire homogène sur \mathbb{K} à n inconnues. L'ensemble F des solutions de (S) est un sous-espace vectoriel de \mathbb{K}^n . La méthode du pivot de Gauss, vue au chapitre III du cours, permet de déterminer une base de F : on trouve, par cette méthode, un système (S') équivalent à (S) et sous forme échelonnée réduite. Soit p' le nombre de lignes non nulles de (S') . D'après le chapitre III, on peut décrire l'ensemble F avec $n - p'$ paramètres (les variables libres du système). Cette description donne une base de (S') à $n - p'$ éléments.³ L'espace vectoriel F est de dimension $n - p'$.

Considérons par exemple l'espace vectoriel

$$F = \{(x, y, z, t) \in \mathbb{R}^4 \text{ t.q. } x + 2y - t = 0 \text{ et } z + 2t = 0\}.$$

On veut trouver une base et déterminer la dimension de l'espace vectoriel F . Celui-ci est décrit par un système linéaire qui est déjà sous forme échelonnée réduite. Les variables libres sont y et t , les variables de base x et z . L'ensemble F est donné par

$$F = \left\{ \begin{bmatrix} -2y + t \\ y \\ -2t \\ t \end{bmatrix}, (y, t) \in \mathbb{R}^2 \right\} = \left\{ y \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix}, (y, t) \in \mathbb{R}^2 \right\},$$

3. Le fait que cette famille est libre résulte de la forme échelonnée de (S') .

ou encore

$$F = \text{vect} \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} \right\}.$$

La famille $\left(\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} \right)$ est une base de F .

Passer d'une famille génératrice à un système d'équations

Soit maintenant F un sous-espace vectoriel de \mathbb{K}^n dont on connaît une famille génératrice $(\vec{f}_1, \dots, \vec{f}_k)$. On cherche une description cartésienne de F . Soit $\vec{x} \in \mathbb{K}^n$. On écrit

$$\begin{aligned} \vec{x} \in F &\iff \exists(\lambda_1, \dots, \lambda_k) \in \mathbb{K}^k, \lambda_1 \vec{f}_1 + \dots + \lambda_k \vec{f}_k = \vec{x} \\ &\iff \text{Le système } (S) : \lambda_1 \vec{f}_1 + \dots + \lambda_k \vec{f}_k = \vec{x}, \text{ d'inconnues } \lambda_1, \dots, \lambda_k \text{ est compatible.} \end{aligned}$$

On transforme alors, par la méthode du pivot, le système (S) en un système sous forme échelonnée réduite (S') . La compatibilité des systèmes (S) et (S') est équivalente à la nullité des membres de droite des lignes de (S') dont le membre de gauche est nul, ce qui donne un système linéaire sur les coordonnées (x_1, \dots, x_n) de \vec{x} , donc une description cartésienne de F .

Exemple V.5.9. Soit $F = \text{vect} \left\{ \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} \right\}$. Alors

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in F \iff \exists(\lambda, \mu) \in \mathbb{K}^2 \text{ t.q. } (S) \begin{cases} \lambda + 2\mu = x \\ 3\lambda - \mu = y \\ -4\lambda - \mu = z \end{cases}$$

Par les opérations $(L_2) \leftarrow (L_2) - 3(L_1)$, $(L_3) \leftarrow (L_3) + 4(L_1)$, puis $(L_3) \leftarrow (L_3) + (L_2)$, on obtient le système sous forme échelonnée réduite, équivalent au système (S) :

$$(S') \quad \begin{cases} \lambda + 2\mu = x \\ -7\mu = y - 3x \\ 0 = x + y + z. \end{cases}$$

On voit que (S') admet une solution (λ, μ) si et seulement si $x + y + z = 0$, ce qui donne une description cartésienne de F :

$$F = \{(x, y, z) \in \mathbb{R}^3 \text{ t.q. } x + y + z = 0\}.$$

Pour résumer :

Proposition V.5.10. *Tout sous-espace vectoriel F de \mathbb{K}^n admet une description cartésienne. Si $p = \dim F$, F s'écrit comme l'ensemble des solutions d'un système homogène sous forme échelonnée réduite à n inconnues et $n - p$ équations.*

En particulier, une droite de \mathbb{K}^n est l'ensemble des solutions d'un système homogène sous forme échelonnée à $n - 1$ équations. Un sous-espace vectoriel de \mathbb{K}^n de dimension $n - 1$ (un tel sous-espace vectoriel est appelé *hyperplan* de \mathbb{K}^n) s'écrit comme l'ensemble des solutions d'une seule équation linéaire homogène. En particulier, un plan de \mathbb{R}^3 peut toujours s'écrire comme l'ensemble des (x, y, z) tels que $ax + by + cz = 0$ pour un certain triplet de réels non tous nuls (a, b, c) .

V.5.d. Manipulation de familles de vecteurs de \mathbb{K}^n

Calcul du rang d'une famille. Extraction d'une base

On rappelle que le *rang* d'une famille de vecteurs \mathcal{F} de E est la dimension de l'espace vectoriel engendré par cette famille. Pour rechercher le rang de \mathcal{F} , il suffit donc de trouver une base de $\text{vect } \mathcal{F}$. La démonstration facile de la proposition suivante est laissée au lecteur :

Proposition V.5.11. Soit $\mathcal{F} = (\vec{e}_1, \dots, \vec{e}_k)$ une famille de vecteurs de E . Soit \mathcal{F}' une des familles suivantes :

- \mathcal{F}' est la famille obtenue à partir de \mathcal{F} en échangeant les vecteurs \vec{e}_j et \vec{e}_k , où $j \neq k$.
- \mathcal{F}' est la famille obtenue à partir de \mathcal{F} en remplaçant le j -ième vecteur \vec{e}_j par le vecteur $\vec{e}_j + \lambda \vec{e}_k$ où $j \neq k$ et $\lambda \in \mathbb{K}$.
- \mathcal{F}' est la famille obtenue à partir de \mathcal{F} en remplaçant le j -ième vecteur \vec{e}_j par le vecteur $\lambda \vec{e}_j$, où $\lambda \in \mathbb{K} \setminus \{0\}$.

Alors $\text{vect } \mathcal{F} = \text{vect } \mathcal{F}'$.

En d'autres termes, les opérations élémentaires sur les vecteurs (analogues des opérations élémentaires sur les lignes du chapitre III) ne changent pas $\text{vect } \mathcal{F}$. Lorsque $E = \mathbb{K}^n$, on peut alors trouver une base de $\text{vect } \mathcal{F}$ en appliquant la méthode du pivot de Gauss sur les éléments de \mathcal{F} , pour ramener \mathcal{F} à une famille de vecteurs échelonnée, au sens de la définition suivante :

Définition V.5.12. Une famille de vecteurs $(\vec{v}_1, \dots, \vec{v}_p)$ de \mathbb{R}^n est dite *échelonnée* lorsque la matrice $p \times n$ obtenue en mettant à la i -ième ligne les coordonnées de \vec{v}_i est échelonnée.

Il est facile de voir que le rang d'une famille de vecteurs échelonnée est égal au nombre de vecteurs non nuls de cette famille.

Exemple V.5.13. Considérons la famille de \mathbb{R}^4 $\mathcal{F} = \left(\begin{bmatrix} -1 \\ -2 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -2 \\ -2 \end{bmatrix} \right)$. On ap-

plique la méthode du pivot⁴ à \mathcal{F} :

$$\left(\begin{bmatrix} -1 \\ -2 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -4 \\ -5 \end{bmatrix} \right) \quad \left(\begin{bmatrix} -1 \\ -2 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ -3 \end{bmatrix} \right)$$

$$\begin{array}{ll} (C_3) \leftarrow (C_3) + (C_1) & (C_3) \leftarrow (C_3) - (C_2) \\ (C_4) \leftarrow (C_4) - (C_1) & (C_4) \leftarrow (C_4) + 2(C_2). \end{array}$$

4. Le lecteur gêné par les opérations sur les colonnes pourra écrire les vecteurs de \mathcal{F} en lignes plutôt qu'en colonnes, et remplacer les opérations élémentaires sur les colonnes par des opérations élémentaires sur les lignes.

La famille \mathcal{F} est donc de même rang que la famille $\mathcal{F}' = \left(\begin{bmatrix} -1 \\ -2 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix} \right)$, qui est une famille de vecteurs échelonnée. Donc \mathcal{F} est de rang 3, et $\text{vect } \mathcal{F}$ a pour base \mathcal{F}' .

Compléter une famille libre en une base

Soit \mathcal{F} une famille libre de \mathbb{K}^n . Par le théorème de la base incomplète, il existe une base \mathcal{B} de \mathbb{K}^n qui complète \mathcal{F} . Pour trouver une telle base, on peut “échelonner” la famille comme précédemment, à l’aide de la proposition V.5.11, puis compléter par les vecteurs appropriés de la base canonique de \mathbb{K}^n .

Exemple V.5.14. La famille $\left(\begin{bmatrix} -1 \\ -2 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix} \right)$ de \mathbb{R}^4 , obtenue plus haut, est libre et échelonnée. Elle se complète de manière triviale en une base $\left(\begin{bmatrix} -1 \\ -2 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right)$ de \mathbb{R}^4 .

Exemple V.5.15. La famille $\mathcal{F} = \left(\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix} \right)$ de \mathbb{R}^3 est libre. Par l’opération $(C_2) \leftarrow (C_2) - 2(C_1)$, on obtient la famille échelonnée $\mathcal{F}' = \left(\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ -8 \\ -1 \end{bmatrix} \right)$, qui vérifie, par la proposition V.5.11, $\text{vect } \mathcal{F} = \text{vect } \mathcal{F}'$. On complète la famille \mathcal{F}' en une base $\left(\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ -8 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right)$ de \mathbb{R}^3 .

On en déduit que $\left(\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right)$ est une base de \mathbb{R}^3 , qui complète \mathcal{F} .

Appendice : alphabet grec

Minuscule	Majuscule	Nom	Minuscule	Majuscule	Nom
α	A	alpha	ν	N	nu
β	B	bêta	ξ	Ξ	xi
γ	Γ	gamma	o	O	omicron
δ	Δ	delta	π	Π	pi
ε	E	epsilon	ρ	P	rhô
ζ	Z	zêta	σ	Σ	sigma
η	H	êta	τ	T	tau
θ	Θ	thêta	υ	Υ	upsilon
ι	I	iota	ϕ	Φ	phi
κ	K	kappa	χ	X	khi
λ	Λ	lambda	ψ	Ψ	psi
μ	M	mu	ω	Ω	oméga

VI. Applications linéaires

Référence pour ce chapitre : Liret-Martinais¹.

On notera comme d'habitude $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

Une application linéaire est une application d'un espace vectoriel dans un autre qui est compatible avec la structure d'espace vectoriel. Une base de l'espace vectoriel de départ et une base de l'espace vectoriel d'arrivée étant données, ces applications linéaires s'identifient aux matrices, vues au chapitre IV, la composition des applications s'identifiant au produit matriciel. La partie VI.1 de ce chapitre est consacrée à des définitions et des propriétés élémentaires, la partie VI.2 à la correspondance entre les applications linéaires et les matrices. En VI.3, on définit des opérations sur les applications linéaires et on identifie ces opérations aux opérations matricielles. La suite du chapitre est consacrée au lien des applications linéaires avec les objets étudiés au chapitre V : bases, sous-espaces vectoriels, dimension.

VI.1. Définitions et propriétés

VI.1.a. Définitions et exemples

Soit E et F des espaces vectoriels sur \mathbb{K} . On notera $\vec{0}_E$ (respectivement $\vec{0}_F$) le vecteur nul de E (respectivement F).

Définition VI.1.1. Soit f une application de E dans F . On dit que f est une *application linéaire* de E dans F quand les deux conditions suivantes sont respectées :

$$(VI.1) \quad \forall(\vec{x}, \vec{y}) \in E^2, \quad f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$$

$$(VI.2) \quad \forall\vec{x} \in E, \forall\lambda \in \mathbb{K}, \quad f(\lambda\vec{x}) = \lambda f(\vec{x}).$$

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

Une application linéaire de E dans E est appelée *endomorphisme* de E . On note pour simplifier $\mathcal{L}(E)$ (au lieu de $\mathcal{L}(E, E)$) l'ensemble des endomorphismes de E .

Exemples VI.1.2. i. Si $\lambda \in \mathbb{K} \setminus \{0\}$, l'application $h_\lambda : \vec{x} \mapsto \lambda\vec{x}$ est un endomorphisme de E , appelée *homothétie* de rapport λ . L'application h_1 est simplement l'identité de E .

ii. L'application constante nulle, $\vec{x} \mapsto \vec{0}_F$ est une application linéaire de E dans F , notée 0 .

iii. L'application f définie par $f(x, y) = 3x - 4y$ est une application linéaire de \mathbb{R}^2 dans \mathbb{R} . La même formule définit une application linéaire de \mathbb{C}^2 dans \mathbb{C} .

iv. La fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(x) = x^3$ n'est pas une application linéaire de \mathbb{R} sur \mathbb{R} . En effet,

$$f(2 \times 1) = f(2) = 8 \text{ mais } 2f(1) = 2 \neq 8.$$

1. François Liret et Dominique Martinais. *Algèbre 1re année - Cours et exercices avec solutions*. Dunod, deuxième édition, 2003

VI. Applications linéaires

v. Soit $\theta \in \mathbb{R}$. La rotation R_θ de \mathbb{R}^2 de centre $(0, 0)$ et d'angle θ est une application linéaire de \mathbb{R}^2 dans \mathbb{R}^2 . En effet, on a la formule (à vérifier) :

$$R_\theta(x, y) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta),$$

dont on déduit facilement le résultat annoncé. Voir aussi l'exercice VI.1.4 plus loin.

VI.1.b. Quelques propriétés

Proposition VI.1.3. Si $f \in \mathcal{L}(E, F)$, $f(\vec{0}_E) = \vec{0}_F$.

Démonstration.

$$f(\vec{0}_E) = f(0 \times \vec{0}_E) = 0f(\vec{0}_E) = \vec{0}_F.$$

□

Exercice VI.1.4. Soit M un point du plan \mathbb{R}^2 , différent de l'origine $(0, 0)$, et $\theta \in (0, 2\pi)$. La rotation f de \mathbb{R}^2 de centre M et d'angle θ est-elle une application linéaire?

(cf correction p. 121).

Proposition VI.1.5. Soit $f \in \mathcal{L}(E, F)$ et $\vec{u}_1, \dots, \vec{u}_k$ des éléments de E . Alors

$$f(\lambda_1 \vec{u}_1 + \dots + \lambda_k \vec{u}_k) = \lambda_1 f(\vec{u}_1) + \dots + \lambda_k f(\vec{u}_k).$$

Démonstration. Si $k = 2$, on démontre la formule en utilisant successivement les conditions (VI.1) et (VI.2) de la définition VI.1.1 :

$$f(\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2) = f(\lambda_1 \vec{u}_1) + f(\lambda_2 \vec{u}_2) = \lambda_1 f(\vec{u}_1) + \lambda_2 f(\vec{u}_2).$$

Le cas général se démontre par récurrence sur k .

□

Corollaire VI.1.6. On suppose E de dimension n . Soit $\mathcal{B} = (\vec{u}_1, \dots, \vec{u}_n)$ une base de E , et $(\vec{v}_1, \dots, \vec{v}_n)$ des vecteurs de F . Alors il existe une unique application linéaire $f \in \mathcal{L}(E, F)$ telle que

$$(VI.3) \quad \forall j = 1 \dots n, \quad f(\vec{u}_j) = \vec{v}_j.$$

Démonstration. Supposons que $f \in \mathcal{L}(E, F)$ vérifie (VI.3). Soit \vec{x} un élément de E et (x_1, \dots, x_n) ses coordonnées dans la base \mathcal{B} . Alors par la proposition VI.1.5

$$(VI.4) \quad f(x_1 \vec{u}_1 + x_2 \vec{u}_2 + \dots + x_n \vec{u}_n) = x_1 \vec{v}_1 + \dots + x_n \vec{v}_n,$$

ce qui montre l'unicité de f vérifiant (VI.3). Pour montrer l'existence, il suffit de définir f par la formule (VI.4) : on vérifie aisément que l'application obtenue est un élément de $\mathcal{L}(E, F)$. □

Exemple VI.1.7. Soit $\vec{u} = (1, 3)$, $\vec{v} = (1, 4)$. Alors il existe une unique application linéaire $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ telle que $f(\vec{u}) = (-1, 2, 3)$, $f(\vec{v}) = (0, 0, 0)$.

Exercice VI.1.8. Donner une expression de l'application f de l'exemple précédent. (Correction p. 121).

VI.2. Applications linéaires et matrices

VI.2.a. Matrice de représentation d'une application linéaire

On montre ici que sur les espaces vectoriels de dimension finie², une base de l'espace de départ et une base de l'espace d'arrivée étant données, une application linéaire s'identifie à une matrice.

On rappelle que l'on identifie \mathbb{K}^n à l'espace vectoriel $\mathcal{M}_{n,1}(\mathbb{K})$ des matrices colonnes à n lignes : un élément de \mathbb{K}^n sera noté indifféremment (x_1, x_2, \dots, x_n) ou $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ (cf par exemple

la formule (VI.5) plus bas : la première notation est utilisée pour le terme de gauche de la première égalité, la seconde notation est utilisée pour tous les autres termes).

Définition VI.2.1. Soit E, F des espaces vectoriels de dimensions respectivement n et p . Soit $\mathcal{B} = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)$ une base de E et $\mathcal{C} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p)$ une base de F . Soit $f \in \mathcal{L}(E, F)$. La *matrice de représentation de f* dans les bases \mathcal{B} et \mathcal{C} (ou plus simplement *matrice de f* dans les bases \mathcal{B} et \mathcal{C}), notée $\text{Mat}_{\mathcal{B},\mathcal{C}}(f)$ est la matrice $p \times n$ dont le coefficient (i, j) est donné par la i -ème coordonnée de $f(\vec{u}_j)$ dans la base \mathcal{C} . En d'autres termes, la j -ième colonne de $\text{Mat}_{\mathcal{B},\mathcal{C}}(f)$ est donnée par les coordonnées de $f(\vec{u}_j)$ dans la base \mathcal{C} .

Les bases \mathcal{B} et \mathcal{C} étant fixées, on montre facilement que pour toute matrice A de $\mathcal{M}_{p,n}(\mathbb{K})$, il existe une unique application linéaire $f \in \mathcal{L}(E, F)$ telle que $A = \text{Mat}_{\mathcal{B},\mathcal{C}}(f)$ ³. Il est donc équivalent de considérer les applications linéaires de E dans F et les matrices $p \times n$. Le point de vue "application linéaire" est intrinsèque : il ne dépend pas du choix de bases de E et F . Le point de vue matricielle n'a pas cette propriété, mais est très utile pour faire des calculs explicites.

VI.2.b. Exemple : cas des bases canoniques

Soit f une application linéaire de \mathbb{K}^n dans \mathbb{K}^p , $A = [a_{i,j}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ sa matrice dans les bases canoniques de \mathbb{K}^n et \mathbb{K}^p . Alors :

$$(VI.5) \quad f(x_1, x_2, \dots, x_n) = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{p1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{p2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{pn} \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pn}x_n \end{bmatrix}.$$

On peut donc facilement reconnaître une application linéaire de \mathbb{K}^n dans \mathbb{K}^p : chacune de ses coordonnées dans \mathbb{K}^p est donnée, comme dans (VI.5), par une combinaison linéaire de coordonnées dans \mathbb{K}^n . De plus, on lit sur cette formule les coefficients de la matrice de l'application linéaire dans les bases canoniques :

2. On rappelle que tous les exemples d'espaces vectoriels vus dans ce cours sont de dimension finie.
3. En d'autres termes, l'application $f \mapsto \text{Mat}_{\mathcal{B},\mathcal{C}}(f)$ de $\mathcal{L}(E, F)$ dans $\mathcal{M}_{p,n}(\mathbb{K})$ est une bijection, cf VI.5.

VI. Applications linéaires

Exemple VI.2.2. La formule

$$(VI.6) \quad f(x_1, x_2, x_3, x_4, x_5) = (\pi x_1 + 4x_3 + x_5, ix_3 + \sqrt{2}x_4 + x_5, -x_1 + x_2, x_1 + 3x_2 + 4x_3 + x_4)$$

que l'on peut encore écrire :

$$(VI.7) \quad f(x_1, x_2, x_3, x_4, x_5) = \begin{pmatrix} \pi x_1 + 4x_3 + x_5 \\ ix_3 + \sqrt{2}x_4 + x_5 \\ -x_1 + x_2 \\ x_1 + 3x_2 + 4x_3 + x_4 \end{pmatrix}$$

définit une application linéaire de \mathbb{C}^5 dans \mathbb{C}^4 . Sa matrice dans les bases canoniques est :

$$\begin{pmatrix} \pi & 0 & 4 & 0 & 1 \\ 0 & 0 & i & \sqrt{2} & 1 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & 3 & 4 & 1 & 0 \end{pmatrix}.$$

Exemple VI.2.3. La matrice de représentation de la rotation R_θ de centre 0 et d'angle θ (cf Exemple VI.1.2, v) est :

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

VI.2.c. Cas général

Donnons un exemple de calcul de matrice de représentation dans des bases autres que les bases canoniques.

Exemple VI.2.4. Soit $f \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$ défini par

$$(VI.8) \quad f(x_1, x_2) = \frac{1}{2} \begin{bmatrix} x_1 + 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + x_2 \end{bmatrix}$$

Soit $\mathcal{B} = (\vec{u}_1, \vec{u}_2)$, $\mathcal{C} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$, où

$$\vec{u}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

On vérifie facilement que \mathcal{B} et \mathcal{C} sont des bases de \mathbb{R}^2 et \mathbb{R}^3 respectivement. Calculons

$$A = \text{Mat}_{\mathcal{B}, \mathcal{C}}(f).$$

La formule (VI.8) nous donne :

$$f(\vec{u}_1) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \vec{v}_1, \quad f(\vec{u}_2) = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} = 2\vec{v}_2.$$

D'où

$$\text{Mat}_{\mathcal{B},\mathcal{C}}(f) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}.$$

Remarquons que cette matrice est complètement différente de la matrice de f dans les bases canoniques de \mathbb{R}^2 et \mathbb{R}^3 , $\frac{1}{2} \begin{bmatrix} 1 & 3 \\ 3 & 5 \\ -1 & 1 \end{bmatrix}$, qui se lit sur la formule (VI.8).

La matrice de f dans les bases \mathcal{B} et \mathcal{C} est beaucoup plus simple que sa matrice dans les bases canoniques. On voit ici l'intérêt de choisir une autre base que la base canonique pour écrire la matrice d'une application linéaire.

Comme le montre l'exemple précédent, le calcul de la matrice de représentation d'une application linéaire f dans les bases $(\vec{u}_1, \dots, \vec{u}_n)$, $(\vec{v}_1, \dots, \vec{v}_p)$ se décompose en deux étapes :

- le calcul de $f(\vec{u}_1), \dots, f(\vec{u}_n)$;
- le calcul des coordonnées de $f(\vec{u}_1), \dots, f(\vec{u}_n)$ dans la base $(\vec{v}_1, \dots, \vec{v}_p)$: soit, en général, la résolution de n systèmes de Cramer, chacun à p équations et p inconnues. Dans l'exemple précédent, ce calcul était immédiat.

Il est aussi possible de déterminer la matrice de représentation de f par des calculs matriciels, à l'aide de la formule de changement de bases qui n'est pas au programme du cours cette année mais est donnée en complément (cf partie VI.8).

VI.2.d. Un exemple de changement de base

Soit E et F des espaces vectoriels de dimensions finies et $f \in \mathcal{L}(E, F)$. On se donne deux bases de E , \mathcal{B} et $\tilde{\mathcal{B}}$ et deux bases de F , \mathcal{C} et $\tilde{\mathcal{C}}$. Peut-on exprimer $\text{Mat}_{\mathcal{B},\mathcal{C}}(f)$ en fonction de $\text{Mat}_{\tilde{\mathcal{B}},\tilde{\mathcal{C}}}(f)$? La formule de changement de bases (cf partie VI.8), hors programme cette année, permet de répondre à cette question. On peut aussi le faire par des calculs directs. On donne ici un exemple d'un tel calcul.

On suppose $E = \mathbb{R}^2$, $F = \mathbb{R}^3$, $\mathcal{B} = (\vec{u}_1, \vec{u}_2)$, $\tilde{\mathcal{B}} = (\tilde{\vec{u}}_1, \tilde{\vec{u}}_2)$, $\mathcal{C} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$, $\tilde{\mathcal{C}} = (\tilde{\vec{v}}_1, \tilde{\vec{v}}_2, \tilde{\vec{v}}_3)$ avec

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \tilde{\vec{u}}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \tilde{\vec{u}}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

et

$$\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \tilde{\vec{v}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \tilde{\vec{v}}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \tilde{\vec{v}}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

On suppose de plus

$$\text{Mat}_{\mathcal{B},\mathcal{C}}(f) = \begin{bmatrix} -1 & 1 \\ 2 & 0 \\ 3 & 2 \end{bmatrix}.$$

On cherche à calculer $\text{Mat}_{\tilde{\mathcal{B}},\tilde{\mathcal{C}}}(f)$.

Première étape. On commence par calculer $f(\tilde{\vec{u}}_1)$ et $f(\tilde{\vec{u}}_2)$. Puisqu'on connaît $f(\vec{u}_1)$ et $f(\vec{u}_2)$, il suffit de trouver les coordonnées de $\tilde{\vec{u}}_1$ et $\tilde{\vec{u}}_2$ dans la base \mathcal{B} . On a :

$$\tilde{\vec{u}}_1 = \vec{u}_1 \text{ et } \tilde{\vec{u}}_2 = \frac{1}{2}\vec{u}_1 + \frac{1}{2}\vec{u}_2$$

VI. Applications linéaires

(la deuxième égalité s'obtient par résolution du système $x_1 + x_2 = 1$, $x_1 - x_2 = 0$). Donc, en utilisant l'expression de $\text{Mat}_{\mathcal{B},\mathcal{C}}(f)$,

$$\begin{aligned} f(\vec{u}_1) &= f(\vec{u}_1) = -\vec{v}_1 + 2\vec{v}_2 + 3\vec{v}_3, \\ f(\vec{u}_2) &= \frac{1}{2}f(\vec{u}_1) + \frac{1}{2}f(\vec{u}_2) = \frac{1}{2}(-\vec{v}_1 + 2\vec{v}_2 + 3\vec{v}_3) + \frac{1}{2}(\vec{v}_1 + 2\vec{v}_3) = \vec{v}_2 + \frac{5}{2}\vec{v}_3. \end{aligned}$$

Deuxième étape. On a obtenu les coordonnées de $f(\vec{u}_1)$ et $f(\vec{u}_2)$ dans la base \mathcal{C} . On a besoin des coordonnées de ces vecteurs dans la base $\tilde{\mathcal{C}}$. Il suffit pour cela d'écrire les vecteurs \vec{v}_1 , \vec{v}_2 et \vec{v}_3 en fonction de \vec{v}_1 , \vec{v}_2 et \vec{v}_3 . En résolvant pour chaque vecteur un système linéaire (dans ce cas précis, échelonné) de trois équations à trois inconnues, on obtient

$$\vec{v}_1 = 2\vec{v}_1 - \vec{v}_2, \quad \vec{v}_2 = \vec{v}_2, \quad \vec{v}_3 = -\vec{v}_1 + \vec{v}_3.$$

D'où, en utilisant la première étape,

$$f(\vec{u}_1) = -(2\vec{v}_1 - \vec{v}_2) + 2\vec{v}_2 + 3(\vec{v}_3 - \vec{v}_1) = -5\vec{v}_1 + 3\vec{v}_2 + 3\vec{v}_3,$$

et

$$f(\vec{u}_2) = \vec{v}_2 + \frac{5}{2}(-\vec{v}_1 + \vec{v}_3) = -\frac{5}{2}\vec{v}_1 + \vec{v}_2 + \frac{5}{2}\vec{v}_3.$$

On en déduit

$$\text{Mat}_{\tilde{\mathcal{B}},\tilde{\mathcal{C}}}(f) = \begin{bmatrix} -5 & -\frac{5}{2} \\ 3 & 1 \\ 3 & \frac{5}{2} \end{bmatrix}.$$

VI.2.e. Applications linéaires et multiplication par une matrice

La proposition suivante donne une façon plus concrète d'interpréter $\text{Mat}_{\mathcal{B},\mathcal{C}}(f)$.

Proposition VI.2.5. *Sous les conditions de la définition VI.2.1, on se donne un vecteur \vec{x} de E de coordonnées $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ dans la base \mathcal{B} . On note $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}$ les coordonnées de $f(\vec{x})$ dans la base \mathcal{C} . Alors*

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

Démonstration. Notons $\text{Mat}_{\mathcal{B},\mathcal{C}}(f) = A = [a_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$. Alors

$$\begin{aligned} f(\vec{x}) &= f(x_1\vec{u}_1 + x_2\vec{u}_2 + \dots + x_n\vec{u}_n) = x_1f(\vec{u}_1) + x_2f(\vec{u}_2) + \dots + x_nf(\vec{u}_n) \\ &= x_1 \left(\sum_{i=1}^p a_{i1}\vec{v}_i \right) + x_2 \left(\sum_{i=1}^p a_{i2}\vec{v}_i \right) + \dots + x_n \left(\sum_{i=1}^p a_{in}\vec{v}_i \right) \end{aligned}$$

En regroupant les termes de la dernière ligne de l'inégalité précédente, on voit que pour $i = 1 \dots p$, la coordonnée y_i de \vec{v}_i est donnée par

$$y_i = \sum_{j=1}^n a_{ij}x_j,$$

ce qui signifie bien que $Y = AX$ au sens du produit matriciel. □

Exemple VI.2.6. Reprenons l'application linéaire f de l'exemple VI.2.4. Les coordonnées de $f(2\vec{u}_1 + \vec{u}_2)$ dans la base $\mathcal{C} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$ sont données par le produit matriciel :

$$\text{Mat}_{\mathcal{B},\mathcal{C}}(f) \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}.$$

En d'autres termes :

$$f(2\vec{u}_1 + \vec{u}_2) = 2\vec{v}_1 + 2\vec{v}_2.$$

En identifiant \mathbb{K}^n à $\mathcal{M}_{n,1}(\mathbb{K})$, une matrice $A \in \mathcal{M}_{p,n}(\mathbb{K})$ définit une application linéaire

$$X \mapsto AX$$

de \mathbb{K}^n dans \mathbb{K}^p . En fait, toute application linéaire de \mathbb{K}^n dans \mathbb{K}^p est de cette forme :

Proposition VI.2.7. *Soit f une application linéaire de \mathbb{K}^n dans \mathbb{K}^p . Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$ la matrice de f dans les bases canoniques de \mathbb{K}^n et \mathbb{K}^p . Alors f est l'application linéaire*

$$X \mapsto AX.$$

Démonstration. Ceci découle immédiatement de la proposition VI.2.5. Si X est un vecteur de \mathbb{R}^n (respectivement de \mathbb{R}^p), le vecteur de ses coordonnées dans la base canonique de \mathbb{R}^n (respectivement \mathbb{R}^p) est également X . □

Exemple VI.2.8. L'application linéaire f de \mathbb{R}^2 dans \mathbb{R}^3 définie par $f(x_1, x_2) = \begin{bmatrix} 2x_1 + x_2 \\ x_1 - x_2 \\ 3x_2 \end{bmatrix}$ est

l'application $X \mapsto AX$, avec

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

VI.3. Opérations sur les applications linéaires

VI.3.a. Addition et multiplication par un scalaire

Définition VI.3.1. Soit $f, g \in \mathcal{L}(E, F)$, $\mu \in \mathbb{K}$. On définit les applications $f + g$ et μf de E dans F par :

$$(f + g)(\vec{x}) = f(\vec{x}) + g(\vec{x}), \quad (\mu f)(\vec{x}) = \mu f(\vec{x}).$$

Dans les deux égalités de la ligne précédente, l'addition et la multiplication par μ apparaissant dans le membre de droite de chaque égalité sont l'addition et la multiplication par μ de l'espace vectoriel F .

Proposition VI.3.2. *Si $f, g \in \mathcal{L}(E, F)$ et $\mu \in \mathbb{K}$, les applications $f + g$ et μf données par la définition VI.3.1 sont des éléments de $\mathcal{L}(E, F)$.*

VI. Applications linéaires

Démonstration. Soit \vec{x} et \vec{y} deux éléments de E . En utilisant successivement la définition de $f + g$, la linéarité de f et g , la commutativité de l'addition sur F puis à nouveau la définition de $f + g$, on obtient :

$$\begin{aligned}(f + g)(\vec{x} + \vec{y}) &= f(\vec{x} + \vec{y}) + g(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}) + g(\vec{x}) + g(\vec{y}) \\ &= f(\vec{x}) + g(\vec{x}) + f(\vec{y}) + g(\vec{y}) = (f + g)(\vec{x}) + (f + g)(\vec{y}).\end{aligned}$$

De même, si $\lambda \in \mathbb{K}$,

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

Donc $f + g \in \mathcal{L}(E, F)$. La démonstration du fait que μf est un élément de $\mathcal{L}(E, F)$ si $\mu \in \mathbb{K}$ est très proche et laissée au lecteur. \square

Remarque VI.3.3. On montre aisément que l'addition et la multiplication par un scalaire ainsi définies sur $\mathcal{L}(E, F)$ lui confèrent une structure de \mathbb{K} -espace vectoriel, au sens de la définition générale V.1.4. Cette propriété ne sera pas utilisée dans ce cours.

Exemple VI.3.4. Soit $f, g \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$ définis par

$$f(x_1, x_2) = \begin{bmatrix} 2x_1 - x_2 \\ x_1 + x_2 \\ 0 \end{bmatrix}, \quad g(x_1, x_2) = \begin{bmatrix} x_2 - 2x_1 \\ 0 \\ x_2 \end{bmatrix}.$$

Alors :

$$(2f)(x_1, x_2) = \begin{bmatrix} 4x_1 - 2x_2 \\ 2x_1 + 2x_2 \\ 0 \end{bmatrix}$$

et

$$(f + g)(x_1, x_2) = \begin{bmatrix} 0 \\ x_1 + x_2 \\ x_2 \end{bmatrix}.$$

VI.3.b. Composition

On rappelle que si E, F et G sont des ensembles, f est une application de E dans F et g est une application de F dans G , la *composée* de f et g , noté $g \circ f$, est définie par

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

C'est une application de E dans G .

Proposition VI.3.5. Soient E, F, G des espaces vectoriels sur \mathbb{K} . Soit $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$. Alors $g \circ f \in \mathcal{L}(E, G)$.

Démonstration. Si $\vec{x}, \vec{y} \in E$, et $\lambda \in \mathbb{K}$,

$$g \circ f(\vec{x} + \vec{y}) = g(f(\vec{x} + \vec{y})) = g(f(\vec{x}) + f(\vec{y})) = g(f(\vec{x})) + g(f(\vec{y})) = (g \circ f)(\vec{x}) + (g \circ f)(\vec{y}).$$

et

$$g \circ f(\lambda\vec{x}) = g(f(\lambda\vec{x})) = g(\lambda f(\vec{x})) = \lambda(g \circ f(\vec{x})).$$

\square

Exercice VI.3.6. Soit $f \in \mathcal{L}(\mathbb{R}^2, \mathbb{R})$, $g \in \mathcal{L}(\mathbb{R}, \mathbb{R}^3)$ donnés par $f(x_1, x_2) = x_1 + x_2$ et $g(x) = (x, -2x, 3x)$. Calculer $g \circ f$.

Exercice VI.3.7. On note R_θ la rotation de \mathbb{R}^2 d'angle $\theta \in \mathbb{R}$ et de centre $\vec{0}_{\mathbb{R}^2}$. Soit $\theta, \sigma \in \mathbb{R}$. Déterminer $R_\theta \circ R_\sigma$.

Quelle est la composée de la symétrie orthogonale de \mathbb{R}^2 d'axe Ox et de la symétrie orthogonale de \mathbb{R}^2 d'axe Oy ?

VI.3.c. Effet des opérations sur les matrices

Les opérations sur les applications linéaires se traduisent facilement en terme de matrices de représentations :

Proposition VI.3.8. *i.* Soit E, F des espaces vectoriels de dimension finie, \mathcal{B}, \mathcal{C} des bases de E et F respectivement. Soit $f_1, f_2 \in \mathcal{L}(E, F)$, $\lambda, \mu \in \mathbb{K}$. Alors :

$$\text{Mat}_{\mathcal{B}, \mathcal{C}}(\lambda f_1 + \mu f_2) = \lambda \text{Mat}_{\mathcal{B}, \mathcal{C}}(f_1) + \mu \text{Mat}_{\mathcal{B}, \mathcal{C}}(f_2).$$

ii. Soit E, F, G des espaces vectoriels de dimension finie, $\mathcal{B}, \mathcal{C}, \mathcal{D}$ des bases de E, F et G respectivement. Soit $f \in \mathcal{L}(E, F)$, $g \in \mathcal{L}(F, G)$. Alors :

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

Le point (ii) signifie que la composition des applications se traduit par le produit matriciel de leurs matrices de représentations. Remarquons que le produit matriciel de la dernière ligne de la proposition est bien défini : si $n = \dim E$, $p = \dim F$ et $q = \dim G$, le premier facteur du membre de droite est une matrice $q \times p$, le deuxième une matrice $p \times n$. Le produit obtenu est bien une matrice $q \times n$.

Avertissement VI.3.9. Attention à l'ordre des bases \mathcal{B}, \mathcal{C} et \mathcal{D} dans le terme de droite de la formule du point (ii).

Démonstration. Nous ne démontrerons que le point (ii). La preuve (plus facile) du point (i) est laissée au lecteur.

On note n, p et q les dimensions respectives des espaces vectoriels E, F et G . Soit $X \in \mathbb{K}^n = \mathcal{M}_{n,1}(\mathbb{K})$. Soit \vec{x} le vecteur de E de coordonnées X dans la base \mathcal{B} . Soit $Y \in \mathcal{M}_{p,1}(\mathbb{K})$ le vecteur colonne des coordonnées de $f(\vec{x}) \in F$ dans la base \mathcal{C} , et $Z \in \mathcal{M}_{q,1}(\mathbb{K})$ le vecteur colonne des coordonnées de $g \circ f(\vec{x}) = g(f(\vec{x})) \in G$ dans la base \mathcal{D} . Par la proposition VI.2.5, appliquée successivement à f, g et $g \circ f$

$$Y = \text{Mat}_{\mathcal{B}, \mathcal{C}}(f)X, \quad Z = \text{Mat}_{\mathcal{C}, \mathcal{D}}(g)Y = \text{Mat}_{\mathcal{C}, \mathcal{D}}(g) \text{Mat}_{\mathcal{B}, \mathcal{C}}(f)X \text{ et } Z = \text{Mat}_{\mathcal{B}, \mathcal{D}}(g \circ f)X.$$

On en déduit

$$(VI.9) \quad \forall X \in \mathcal{M}_{1,n}(\mathbb{K}), \quad \text{Mat}_{\mathcal{C}, \mathcal{D}}(g) \text{Mat}_{\mathcal{B}, \mathcal{C}}(f)X = \text{Mat}_{\mathcal{B}, \mathcal{D}}(g \circ f)X,$$

ce qui montre l'égalité

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

En effet, pour passer de (VI.9) à cette dernière ligne, il suffit d'appliquer (VI.9) aux vecteurs $X = \vec{e}_1, \dots, X = \vec{e}_n$, où $(\vec{e}_1, \dots, \vec{e}_n)$ est la base canonique de $\mathbb{K}^n = \text{Mat}_{1,n}(\mathbb{K})$. \square

VI. Applications linéaires

Remarque VI.3.10. Le (ii) justifie a posteriori la définition du produit matriciel.

Exercice VI.3.11. Soit $E = \mathbb{R}^2$, $F = \mathbb{R}^3$, $G = \mathbb{R}^2$. On suppose

$$\text{Mat}_{\mathcal{B},\mathcal{C}}(f) = A = \begin{bmatrix} -1 & 1 \\ 2 & 0 \\ 3 & 2 \end{bmatrix}, \quad \text{Mat}_{\mathcal{C},\mathcal{D}}(g) = B = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \end{bmatrix}.$$

a. Calculer $\text{Mat}_{\mathcal{B},\mathcal{D}}(g \circ f)$.

b. La matrice AB est-elle la matrice de représentation d'une certaine application linéaire dans des bases que l'on précisera ?

(cf correction p. 122)

Exercice VI.3.12. Soit, pour $\theta \in \mathbb{R}$,

$$M_\theta = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}.$$

Exprimer, pour $\theta, \sigma \in \mathbb{R}$, la matrice $M_\theta M_\sigma$ en fonction d'une certaine matrice M_ρ (ρ à déterminer). Comparer avec l'exercice VI.3.7.

VI.4. Applications linéaires et sous-espaces vectoriels

VI.4.a. Image et image réciproque d'un ensemble par une application

Commençons par rappeler la définition de l'image et de l'image réciproque d'un ensemble par une application :

Définition VI.4.1. Soit f une application de E dans F , où E et F sont des ensembles.

i. Si $A \subset E$, l'image de A par f , notée $f(A)$, est le sous-ensemble de F :

$$f(A) = \{f(a), \quad a \in A\} = \{b \in F, \text{ t.q. } \exists a \in E, b = f(a)\}.$$

ii. Si $B \subset F$, l'image réciproque⁴ de B par f , notée $f^{-1}(B)$ est le sous-ensemble de E :

$$f^{-1}(B) = \{a \in E \text{ t.q. } f(a) \in B\}.$$

Avertissement VI.4.2. La notation $f^{-1}(B)$ ne signifie pas que f est bijective et admet une application réciproque f^{-1} (notion rappelée p. 109). L'ensemble $f^{-1}(B)$ est défini pour toute application f de E dans F , dès que B est un sous-ensemble de F . Par ailleurs, on n'a pas toujours $f(f^{-1}(B)) = B$ ou $f^{-1}(f(A)) = A$ (cf l'exercice suivant).

Exercice VI.4.3. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ la fonction cosinus $x \mapsto \cos x$. Calculer $f^{-1}(\mathbb{R})$, $f(\mathbb{R})$, $f^{-1}([-1, 1])$, $f^{-1}(\{0\})$, $f([0, \pi])$, $f([0, \pi/2])$, $f(\{0\})$, puis $f(f^{-1}(\mathbb{R}))$ et $f^{-1}(f([0, \pi]))$.

(Réponse p. 122).

4. Le lecteur pressé pourra faire l'impasse en première lecture sur cette notion d'image réciproque, qui sera surtout utilisée dans la suite pour définir le noyau d'une application linéaire (cf la remarque VI.4.5 plus bas qui donne une définition du noyau sans utiliser cette notion).

VI.4.b. Cas des applications linéaires

Définition VI.4.4. Soit $f \in \mathcal{L}(E, F)$. L'image $f(E)$ de E par f est appelée *image de f* et notée $\text{Im}(f)$. L'image réciproque $f^{-1}(\vec{0}_F)$ de $\{\vec{0}_F\}$ par f est appelée *noyau de f* et notée⁵ $\text{Ker}(f)$.

Remarque VI.4.5. En d'autres termes, l'image de f est l'ensemble :

$$\text{Im}(f) = \{\vec{y} \in F, \text{ t.q. } \exists \vec{x} \in E, \vec{y} = f(\vec{x})\},$$

et le noyau de f est l'ensemble :

$$\text{Ker}(f) = \left\{ \vec{x} \in E \text{ t.q. } f(\vec{x}) = \vec{0}_F \right\}.$$

Théorème VI.4.6. Soit $f \in \mathcal{L}(E, F)$.

- i. L'image d'un sous-espace vectoriel de E par f est un sous-espace vectoriel de F .
- ii. L'image réciproque d'un sous-espace vectoriel de F par f est un sous-espace vectoriel de E .

En particulier, le noyau et l'image de f sont des sous-espaces vectoriels (de E et F respectivement).

Démonstration. Soit G un sous-espace vectoriel de E . Alors $\vec{0}_E \in G$, donc $\vec{0}_F = f(\vec{0}_E) \in f(G)$.

Soit maintenant $\vec{u}, \vec{v} \in f(G)$. Par définition de $f(G)$, il existe des vecteurs \vec{x} et \vec{y} de G tels que $\vec{u} = f(\vec{x})$ et $\vec{v} = f(\vec{y})$. On a

$$\vec{u} + \vec{v} = f(\vec{x}) + f(\vec{y}) = f(\vec{x} + \vec{y}).$$

Puisque G est un sous-espace vectoriel de E , $\vec{x} + \vec{y} \in G$ et on déduit de l'égalité précédente $\vec{u} + \vec{v} \in f(G)$.

Si de plus $\lambda \in \mathbb{K}$, on a $\lambda\vec{x} \in G$ et donc

$$\lambda\vec{u} = \lambda f(\vec{x}) = f(\lambda\vec{x}) \in f(G),$$

ce qui achève de montrer que $f(G)$ est bien un sous-espace vectoriel de F .

Montrons maintenant le deuxième point. Soit H un sous-espace vectoriel de F . Alors $f(\vec{0}_E) = \vec{0}_F \in H$, et donc $\vec{0}_E \in f^{-1}(H)$. De plus, si $\vec{u}, \vec{v} \in f^{-1}(H)$. Alors

$$f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v}),$$

et puisque $f(\vec{u})$ et $f(\vec{v})$ sont dans H et H est un sous-espace vectoriel de F , $f(\vec{u}) + f(\vec{v}) \in H$, ce qui montre que $\vec{u} + \vec{v} \in f^{-1}(H)$.

Enfin, si $\lambda \in \mathbb{K}$ et $\vec{u} \in f^{-1}(H)$, alors $f(\lambda\vec{u}) = \lambda f(\vec{u})$ est un élément de H puisque $f(\vec{u})$ est un élément de H . Donc $\lambda\vec{u} \in f^{-1}(H)$. On a bien démontré que $f^{-1}(H)$ est un sous-espace vectoriel de E . \square

5. de l'allemand *Kern* signifiant noyau. La traduction anglaise de noyau (au sens mathématique) est *kernel*.

VI. Applications linéaires

Exemple VI.4.7. Soit A le sous-ensemble de \mathbb{R}^3 :

$$A = \left\{ (x, y, z) \in \mathbb{R}^3, \text{ t.q. } x + y + z = 0 \text{ et } x - 2y = 0 \right\}.$$

Alors A est le noyau de l'application linéaire $f \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$ définie par

$$f(x, y, z) = (x + y + z, x - 2y).$$

C'est donc un sous-espace vectoriel de \mathbb{R}^3 .

Exemple VI.4.8. Soit B le sous-ensemble de \mathbb{R}^3 :

$$B = \left\{ (x_1 + x_2, x_1 - x_2, 2x_1 + x_2), (x_1, x_2) \in \mathbb{R}^2 \right\}.$$

Alors B est l'image de l'application linéaire $g \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$ définie par

$$g(x_1, x_2) = (x_1 + x_2, x_1 - x_2, 2x_1 + x_2).$$

Donc A est un sous-espace vectoriel de \mathbb{R}^3 .

Exercice VI.4.9. Soit A le sous-ensemble de \mathbb{R}^3 des triplets $(x_1 + x_2, x_3 - x_4, x_1 - x_3)$ tels que $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ et $2x_1 - x_2 + x_3 - x_4 = 0$. Montrer que A est un sous-espace vectoriel de \mathbb{R}^3 .

(cf solution p. 122).

VI.5. Injectivité et surjectivité

VI.5.a. Injection, surjection : rappels

Rappel VI.5.1. Soit A et B deux ensembles et f une application de A dans B .

- i. On dit que f est *injective* lorsque tous les éléments dans l'ensemble d'arrivée ont au plus un antécédent dans l'ensemble de départ, i.e :

$$\forall x, y \in A, \quad f(x) = f(y) \implies x = y.$$

Ce qui peut encore se lire :

$$\forall x, y \in A, \quad x \neq y \implies f(x) \neq f(y)$$

(deux éléments distincts de l'ensemble de départ ont des images distinctes).

- ii. On dit que f est *surjective* lorsque tous les éléments dans l'ensemble d'arrivée ont au moins un antécédent dans l'ensemble de départ, i.e.

$$\forall z \in B, \exists x \in A, \text{ t.q. } z = f(x).$$

- iii. On dit que f est *bijective* lorsqu'elle est injective et surjective. C'est à dire :

$$\forall z \in B, \exists! x \in A, \text{ t.q. } z = f(x).$$

Si f est bijective, on appelle application réciproque de f , et on note f^{-1} l'application de B dans A qui à z associe l'unique x de A tel que $f(x) = z$. On a alors

$$f \circ f^{-1} = \text{Id}_B, \quad f^{-1} \circ f = \text{Id}_A.$$

Remarque VI.5.2. Il découle immédiatement de la définition d'une bijection qu'une application $f : A \rightarrow B$ est bijective si et seulement si elle admet une application réciproque, i.e. une application $g : B \rightarrow A$ telle que $f \circ g = \text{Id}_B$ et $g \circ f = \text{Id}_A$.

Exemple VI.5.3. L'application $x \mapsto x^2$ de \mathbb{R} dans \mathbb{R} n'est pas injective : si $(-1)^2 = 1^2$ et $-1 \neq 1$. Elle n'est pas surjective : il n'existe aucun réel x tel que $x^2 = -1$.

L'application $x \mapsto x^3$ est une bijection de \mathbb{R} dans \mathbb{R} , d'application réciproque $x \mapsto \sqrt[3]{x}$.

Avertissement VI.5.4. La notion d'injectivité dépend du choix de l'espace de départ. Les notions de surjectivité et de bijectivité dépendent des espaces de départ et d'arrivée. Par exemple, l'application $f : x \mapsto x^2$ est bijective de $[0, +\infty[$ dans $[0, +\infty[$, surjective (mais pas injective) de \mathbb{R} dans $[0, +\infty[$, injective (mais pas surjective) de $[0, +\infty[$ dans \mathbb{R} .

VI.5.b. Injectivité, surjectivité, image et noyau

Théorème VI.5.5. Soit E, F deux espaces vectoriels et $f \in \mathcal{L}(E, F)$. Alors :

- i. f est surjective si et seulement si $\text{Im}(f) = F$.
- ii. f est injective si et seulement si $\text{Ker}(f) = \{0\}$.

Démonstration. Le point (i) découle immédiatement de la définition de la surjectivité et de celle de l'image d'une application linéaire.

Démontrons (ii).

Supposons f injective. On sait que $\vec{0}_E$ est un élément de $\text{Ker}(f)$. Soit $\vec{x} \in \text{Ker}(f)$. Alors $f(\vec{x}) = \vec{0}_F = f(\vec{0}_E)$, et donc par injectivité $\vec{x} = \vec{0}_E$. D'où $\text{Ker}(f) = \{\vec{0}_E\}$.

Réciproquement, on suppose $\text{Ker}(f) = \{\vec{0}_E\}$. Soit \vec{x} et \vec{y} deux éléments de E tels que $f(\vec{x}) = f(\vec{y})$. Alors

$$f(\vec{x} - \vec{y}) = f(\vec{x}) - f(\vec{y}) = \vec{0}_F,$$

donc $\vec{x} - \vec{y} \in \text{Ker}(f)$, et, puisque $\text{Ker}(f) = \{\vec{0}_E\}$, $\vec{x} - \vec{y} = \vec{0}_E$. On a bien montré que f était injective. \square

Exemple VI.5.6. On suppose $\mathbb{K} = \mathbb{C}$. Soit $f \in \mathcal{L}(\mathbb{C}^3, \mathbb{C}^3)$ définie par

$$f(x_1, x_2, x_3) = (x_1 - x_2 + x_3, x_2 - 2x_3, 4x_3).$$

Alors f est injective. En effet, si $(x_1, x_2, x_3) \in \text{Ker}(f)$, on a $x_1 - x_2 + x_3 = x_2 - x_3 = 4x_3 = 0$, dont on déduit facilement $x_1 = x_2 = x_3 = 0$.

VI.5.c. Injectivité, surjectivité et familles de vecteurs

On peut également caractériser l'injectivité ou la surjectivité d'une application par des propriétés de l'image d'une base de l'espace de départ.

Proposition VI.5.7. Soit E et F des espaces vectoriels. On suppose E de dimension n . Soit $\mathcal{B} = (\vec{e}_1, \dots, \vec{e}_n)$ une base de E , et $f \in \mathcal{L}(E, F)$. Alors :

VI. Applications linéaires

- i. La famille $(f(\vec{e}_1), \dots, f(\vec{e}_n))$ engendre $\text{Im}(f)$. En particulier, l'application f est surjective si et seulement si $(f(\vec{e}_1), \dots, f(\vec{e}_n))$ est une famille génératrice de F .
- ii. L'application f est injective si et seulement si $(f(\vec{e}_1), \dots, f(\vec{e}_n))$ est une famille libre.

Démonstration. Montrons le point (i).

Les vecteurs $f(\vec{e}_1), f(\vec{e}_2), \dots, f(\vec{e}_n)$ sont bien dans $\text{Im}(f)$. De plus, un vecteur $\vec{y} \in \text{Im}(f)$ s'écrit $f(\vec{x})$ avec $\vec{x} \in E$. Soit (x_1, \dots, x_n) les coordonnées de \vec{x} dans la base \mathcal{B} . On a

$$\vec{x} = x_1\vec{e}_1 + \dots + x_n\vec{e}_n, \quad \vec{y} = f(\vec{x}) = x_1f(\vec{e}_1) + \dots + x_nf(\vec{e}_n),$$

ce qui montre bien

$$\text{Im}(f) = \text{vect}(f(\vec{e}_1), f(\vec{e}_2), \dots, f(\vec{e}_n)).$$

On en déduit immédiatement :

$$f \text{ surjective} \iff \text{Im}(f) = F \iff \text{vect}(f(\vec{e}_1), \dots, f(\vec{e}_n)) = F,$$

Montrons maintenant (ii). L'application f est injective si et seulement si $\text{Ker}(f) = \{\vec{0}_E\}$, i.e.

$$\forall \vec{x} \in E, \quad f(\vec{x}) = \vec{0}_F \implies \vec{x} = \vec{0}_E$$

ou, en utilisant que $(\vec{e}_1, \dots, \vec{e}_n)$ est une base de E :

$$f(\lambda_1\vec{e}_1 + \lambda_2\vec{e}_2 + \dots + \lambda_n\vec{e}_n) = \vec{0}_F \implies \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

Par linéarité de f , on peut réécrire la propriété précédente :

$$\lambda_1f(\vec{e}_1) + \lambda_2f(\vec{e}_2) + \dots + \lambda_nf(\vec{e}_n) = \vec{0}_F \implies \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

L'injectivité de f est donc bien équivalente à la liberté de la famille $(f(\vec{e}_1), f(\vec{e}_2), \dots, f(\vec{e}_n))$. \square

VI.6. Rang d'une application linéaire

Soit E un espace vectoriel de dimension n .

VI.6.a. Définition

On rappelle que le *rang* d'une famille \mathcal{F} de vecteurs de E est la dimension du sous-espace vectoriel $\text{vect } \mathcal{F}$ engendré par \mathcal{F} .

D'après le point (i) de la proposition VI.5.7, si $(\vec{e}_1, \dots, \vec{e}_n)$ est une base de E , alors $(f(\vec{e}_1), \dots, f(\vec{e}_n))$ est une famille génératrice de l'image de f , ce qui montre que l'image de f est de dimension inférieure ou égale à la dimension de E .

Définition VI.6.1. Soit $f \in \mathcal{L}(E, F)$ une application linéaire. Le *rang* de f , noté $\text{rg}(f)$, est la dimension de l'image de f . Si $A \in \mathcal{M}_{p,n}$, le rang de A est le rang de l'application linéaire $f \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ de matrice A dans les bases canoniques de \mathbb{R}^n et \mathbb{R}^p .

Remarque VI.6.2. Le rang de $f \in \mathcal{L}(E, F)$ est donc égal au rang de la famille de vecteurs

$$(f(\vec{e}_1), f(\vec{e}_2), \dots, f(\vec{e}_n)).$$

De même, le rang d'une matrice A est égal au rang de la famille de ses vecteurs colonnes.

Exemple VI.6.3. Soit $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ défini par

$$f(x_1, x_2, x_3) = \begin{bmatrix} x_1 + 2x_2 + x_3 \\ 2(x_1 + 2x_2 + x_3) \\ -(x_1 + 2x_2 + x_3) \end{bmatrix}.$$

Alors $f(1, 0, 0) = (1, 2, -1)$, $f(0, 1, 0) = (2, 4, -2)$ et $f(0, 0, 1) = (1, 2, -1)$. Donc

$$\text{rg}(f) = \text{rg} \left(\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \right) = 1.$$

VI.6.b. Théorème du rang

Théorème VI.6.4 (Théorème du rang). Soit $f \in \mathcal{L}(E, F)$, où E et F sont des espaces vectoriels (de dimensions finies). Alors

$$\text{rg}(f) + \dim \text{Ker}(f) = \dim E.$$

Démonstration. Soit k la dimension de $\text{Ker}(f)$, et ℓ la dimension de $\text{Im } f$. On doit montrer

$$\dim E = k + \ell.$$

Soit $(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k)$ une base de $\text{Ker}(f)$, $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_\ell)$ une base de $\text{Im } f$. Il existe donc, pour $j = 1 \dots \ell$, un vecteur $\vec{w}_j \in E$ tel que $f(\vec{w}_j) = \vec{v}_j$. Montrons que $\mathcal{B} = (\vec{u}_1, \dots, \vec{u}_k, \vec{w}_1, \dots, \vec{w}_\ell)$ est une base de E (ce qui impliquera immédiatement le résultat annoncé).

Montrons d'abord que \mathcal{B} engendre E . Soit $\vec{x} \in E$. Puisque $f(\vec{x}) \in \text{Im } f$ et $\vec{v}_1, \dots, \vec{v}_\ell$ est une base $\text{Im } f$, il existe $(y_1, y_2, \dots, y_\ell) \in \mathbb{K}^\ell$ tels que

$$f(\vec{x}) = y_1 \vec{v}_1 + \dots + y_\ell \vec{v}_\ell.$$

On en déduit

$$f(\vec{x}) = y_1 f(\vec{w}_1) + \dots + y_\ell f(\vec{w}_\ell) = f(y_1 \vec{w}_1 + \dots + y_\ell \vec{w}_\ell),$$

et donc

$$\vec{x} - y_1 \vec{w}_1 - \dots - y_\ell \vec{w}_\ell \in \text{Ker}(f).$$

Il existe donc $(x_1, \dots, x_k) \in \mathbb{K}^k$ tels que

$$\vec{x} - y_1 \vec{w}_1 - \dots - y_\ell \vec{w}_\ell = x_1 \vec{u}_1 + \dots + x_k \vec{u}_k,$$

soit

$$\vec{x} = y_1 \vec{w}_1 + \dots + y_\ell \vec{w}_\ell + x_1 \vec{u}_1 + \dots + x_k \vec{u}_k.$$

La famille \mathcal{B} engendre bien E .

VI. Applications linéaires

Montrons maintenant que \mathcal{B} est libre. On suppose que l'on a

$$(VI.10) \quad x_1 \vec{u}_1 + \dots + x_k \vec{u}_k + y_1 \vec{w}_1 + \dots + y_\ell \vec{w}_\ell = \vec{0}_E$$

avec $(x_1, \dots, x_k, y_1, \dots, y_\ell) \in \mathbb{K}^{k+\ell}$. En appliquant f à la ligne précédente et en utilisant que $f(\vec{u}_j) = \vec{0}_F$ pour $j = 1 \dots k$ et $f(\vec{w}_j) = \vec{v}_j$ pour $j = 1 \dots \ell$, on obtient

$$y_1 \vec{v}_1 + y_2 \vec{v}_2 + \dots + y_\ell \vec{v}_\ell = f(\vec{0}_E) = \vec{0}_F.$$

On en déduit, puisque la famille $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_\ell)$ est libre :

$$y_1 = y_2 = \dots = y_\ell = 0.$$

En revenant à (VI.10), et en utilisant que la famille $(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k)$ est libre, on obtient

$$x_1 = x_2 = \dots = x_k = 0,$$

ce qui montre que \mathcal{B} est libre et termine la preuve du théorème du rang. □

Exemple VI.6.5. Soit $f : \mathbb{R}^4 \rightarrow \mathbb{R}^5$ définie par

$$f(x_1, x_2, x_3, x_4) = (x_1 + x_2, x_1 - x_2, 2x_1 + x_2, x_3 + x_4, x_3 - x_4).$$

On veut déterminer le rang de f .

On calcule pour cela la dimension du noyau de f .

$$(x_1, x_2, x_3, x_4) \in \text{Ker}(f) \iff x_1 + x_2 = x_1 - x_2 = 2x_1 + x_2 = x_3 + x_4 = x_3 - x_4 = 0.$$

On résout facilement ce système linéaire homogène de 5 équations :

$$(x_1, x_2, x_3, x_4) \in \text{Ker}(f) \iff x_1 = x_2 = x_3 = x_4 = 0.$$

Donc $\text{Ker}(f) = \{(0, 0, 0, 0)\}$. Le noyau de f est de dimension 0. Par le théorème du rang,

$$\text{rg}(f) = \dim \mathbb{R}^4 - \dim \text{Ker}(f) = 4 - 0 = 4.$$

Remarque VI.6.6. En dimension finie, on peut souvent utiliser le théorème du rang et des arguments de dimension pour étudier l'injectivité, la surjectivité et la bijectivité d'une application. Reprenons l'application injective f de l'exemple VI.5.6 p. 109. Le théorème du rang s'écrit :

$$\text{rg}(f) + \underbrace{\dim \text{Ker}(f)}_0 = \underbrace{\dim \mathbb{C}^3}_3,$$

et donc $\text{rg}(f) = 3$. On en déduit que $\text{Im}(f)$ est un sous-espace vectoriel de dimension 3 de \mathbb{C}^3 , c'est donc exactement \mathbb{C}^3 . Donc f est surjective, et par conséquent bijective. On voit ici, comme dans l'exemple VI.6.5, que la dimension donne une information gratuite et facilement exploitable.

VI.6.c. Rang et systèmes linéaires

L'image et le noyau d'une application linéaire s'interprète en terme de systèmes linéaires. On considère un système linéaire non-homogène à p équations et n inconnues :

$$(S) \quad \forall i \in \{1, \dots, p\}, \quad \sum_{j=1}^n a_{ij}x_j = b_i$$

et le système linéaire homogène correspondant, à p équations et n inconnues :

$$(H) \quad \forall i \in \{1, \dots, p\}, \quad \sum_{j=1}^n a_{ij}x_j = 0.$$

Alors :

Proposition VI.6.7. *Fixons les coefficients a_{ij} . Soit $F \subset \mathbb{K}^n$ l'ensemble des solutions de (H) et G le sous-ensemble de \mathbb{K}^p formé des (b_1, \dots, b_p) tel que (S) a au moins une solution. Alors F et G sont des sous-espaces vectoriels de \mathbb{K}^n et \mathbb{K}^p respectivement et*

$$\dim F + \dim G = n.$$

Si $(b_1, \dots, b_p) \in G$, l'ensemble des solutions \mathcal{S} du système compatible (S) s'écrit :

$$(VI.11) \quad \mathcal{S} = \{\vec{y} + \vec{x}, \quad \vec{x} \in F\},$$

où \vec{y} est une solution, fixée, du système (S).

Démonstration. L'ensemble F solutions de (H) est le noyau $\text{Ker } f$ de l'application linéaire $f : \mathbb{K}^n \rightarrow \mathbb{K}^p$ de matrice $[a_{ij}]_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ dans les bases canoniques de \mathbb{K}^n et \mathbb{K}^p . L'ensemble G est l'image de f . Ceci montre que F et G sont des sous-espaces vectoriels de \mathbb{K}^n et \mathbb{K}^p respectivement. L'égalité sur les dimensions découle du théorème du rang.

Supposons maintenant $(b_1, \dots, b_p) \in G$, et fixons une solution \vec{y} de (S). Soit \mathcal{S} l'ensemble des solutions de (S). On voit facilement que \vec{x} est solution de (S) si et seulement si $\vec{x} - \vec{y}$ est solution de (H), ce qui donne exactement (VI.11). \square

Remarque VI.6.8. Fixons les coefficients a_{ij} du système (S), et notons q la dimension de F . L'affirmation (VI.11) implique que si (S) est compatible, on peut décrire l'ensemble des solutions avec exactement q paramètres. La théorie générale des systèmes linéaires (cf chapitre III) montre que $q = n - p'$, où p' est le nombre de lignes non nulles⁶ que l'on obtient lorsque l'on échelonne le système (S). Le nombre p' est donc indépendant du second membre de (S) (seule la compatibilité du système dépend de ce second membre).

Exemple VI.6.9. Considérons la famille de systèmes, d'inconnues $(x_1, x_2) \in \mathbb{R}^2$:

$$(S) \quad \begin{cases} x_1 - x_2 = b_1 \\ x_1 + x_2 = b_2 \\ 2x_1 + x_2 = b_3. \end{cases}$$

6. C'est à dire autres que $0 = 0$.

Ici $n = 2$, $p = 3$. Il est facile de voir que le système homogène n'a que la solution nulle $(0, 0)$. Donc $F = \{\vec{0}\}$ est de dimension 0. L'espace vectoriel G formé des (b_1, b_2, b_3) tels que (S) est compatible est de dimension $2 - 0 = 2$. Lorsque $(b_1, b_2, b_3) \in G$, le système (S) a une unique solution. Il est facile de déterminer G en échelonnant le système (S). Par les opérations $(L_2) \leftarrow (L_2) - (L_1)$, $(L_3) \leftarrow (L_3) - 2(L_1)$, puis $(L_3) \leftarrow (L_3) - \frac{3}{2}(L_2)$, on obtient le système équivalent

$$(S') \quad \begin{cases} x_1 - x_2 = b_1 \\ 2x_2 = b_2 - b_1 \\ 0 = b_3 - \frac{3}{2}b_2 - \frac{1}{2}b_1. \end{cases}$$

On en déduit que (S) est compatible si et seulement si $2b_3 - 3b_2 - b_1 = 0$, ce qui donne la représentation cartésienne de G :

$$G = \left\{ (b_1, b_2, b_3) \in \mathbb{R}^3 \text{ t.q. } 2b_3 - 3b_2 - b_1 = 0 \right\}.$$

VI.7. Isomorphismes

VI.7.a. Définition

Soit E et F deux \mathbb{K} -espaces vectoriels

Définition VI.7.1. On appelle *isomorphisme* entre E et F une application linéaire bijective de E dans F . Lorsqu'il existe un isomorphisme de E dans F , on dit que E et F sont isomorphes. On appelle *automorphisme* de E un isomorphisme de E dans lui-même. Un automorphisme est donc un endomorphisme bijectif.

Exemple VI.7.2. L'application linéaire f de l'exemple VI.5.6 est, par la remarque VI.6.6 un isomorphisme de \mathbb{C}^3 dans \mathbb{C}^3 , donc un automorphisme de \mathbb{C}^3 .

Exemple VI.7.3. L'identité de E est un automorphisme de E (donc E est isomorphe à lui-même).

Le fait que deux espaces vectoriels sont isomorphes signifie qu'ils ont exactement la même structure d'espace vectoriel : ce sont en quelques sortes deux copies du même espace vectoriel.

VI.7.b. Application réciproque d'un isomorphisme

Proposition VI.7.4. Soit E et F des espaces vectoriels et f un isomorphisme de E dans F . Alors f^{-1} est une application linéaire (donc un isomorphisme) de F dans E .

Remarque VI.7.5. Il découle de la proposition VI.7.4 que la relation " E est isomorphe à F " est symétrique i.e. :

$$E \text{ isomorphe à } F \iff F \text{ isomorphe à } E.$$

Exercice VI.7.6. Vérifier que c'est aussi une relation transitive, i.e. :

$$\left(E \text{ isomorphe à } F \text{ et } F \text{ isomorphe à } G \right) \implies E \text{ isomorphe à } G.$$

Démonstration de la proposition. L'application réciproque d'une application bijective est également bijective. Si f^{-1} est une application linéaire, ce sera donc aussi un isomorphisme. Montrons que f^{-1} est linéaire.

Soit $\vec{x}, \vec{y} \in F$. Alors (par définition de l'application réciproque) :

$$f(f^{-1}(\vec{x} + \vec{y})) = \vec{x} + \vec{y}.$$

De plus

$$f(f^{-1}(\vec{x}) + f^{-1}(\vec{y})) = f(f^{-1}(\vec{x})) + f(f^{-1}(\vec{y})) = \vec{x} + \vec{y}.$$

En combinant ces deux lignes, on obtient :

$$f(f^{-1}(\vec{x} + \vec{y})) = f(f^{-1}(\vec{x}) + f^{-1}(\vec{y})),$$

et donc, par injectivité de f ,

$$f^{-1}(\vec{x} + \vec{y}) = f^{-1}(\vec{x}) + f^{-1}(\vec{y}).$$

On démontre de même que si $\lambda \in \mathbb{K}$, $f^{-1}(\lambda\vec{x}) = \lambda f^{-1}(\vec{x})$. □

Exemple VI.7.7. Soit f l'automorphisme de \mathbb{R}^2 défini par

$$f(x_1, x_2) = (x_1 + x_2, x_1 - x_2).$$

Calculons l'application réciproque de f . Si $(y_1, y_2) \in \mathbb{R}^2$, $(x_1, x_2) = f^{-1}(y_1, y_2)$ est l'unique solution du système

$$y_1 = x_1 + x_2, \quad y_2 = x_1 - x_2.$$

On résout ce système :

$$x_1 = \frac{y_1 + y_2}{2}, \quad x_2 = \frac{y_1 - y_2}{2}.$$

On a donc

$$f^{-1}(y_1, y_2) = \left(\frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2} \right).$$

Plus généralement, calculer la réciproque d'un isomorphisme f dans un système de coordonnées revient à résoudre le système de Cramer inhomogène $f(x_1, \dots, x_n) = (y_1, \dots, y_n)$, d'inconnues (x_1, \dots, x_n) , de second membre (y_1, \dots, y_n) . Cela revient aussi à calculer l'inverse de la matrice de f dans une base donnée (cf plus bas le point de vue matriciel).

On rappelle qu'une application est bijective si et seulement si elle admet une application réciproque. On peut ainsi montrer qu'une certaine application linéaire est un isomorphisme en donnant son application réciproque :

Exemple VI.7.8. Soit $\theta \in \mathbb{R}$. La rotation de \mathbb{R}^2 de centre $(0, 0)$ et d'angle θ :

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

est un automorphisme de \mathbb{R}^2 : elle admet pour application réciproque la rotation $R_{-\theta}$.

VI.7.c. Condition sur les dimensions

Proposition VI.7.9. Soit $f \in \mathcal{L}(E, F)$, où E et F sont de dimension finie.

- i. Si f est injective, $\dim E \leq \dim F$.
- ii. Si f est surjective, $\dim E \geq \dim F$.
- iii. Si f est un isomorphisme, $\dim E = \dim F$.

Démonstration. Le point (iii) découle immédiatement des points (i) et (ii).

Si f est injective, le noyau de f est de dimension 0 et le théorème du rang s'écrit $\dim E = \operatorname{rg}(f) = \dim \operatorname{Im}(f)$. Puisque $\operatorname{Im}(f)$ est un sous-espace vectoriel de F , on a $\dim \operatorname{Im}(f) \leq \dim F$, et le point (i) en découle.

Supposons maintenant f surjective. Alors $\dim \operatorname{Im}(f) = \dim F$, et le théorème du rang s'écrit : $\dim E = \dim F + \dim \operatorname{Ker}(f) \geq \dim F$, ce qui donne le point (ii). \square

Remarque VI.7.10. Si $\dim F > \dim E$, il n'existe donc aucune application linéaire surjective de E dans F . De même, si $\dim F < \dim E$, il n'existe aucune application linéaire injective de E dans F .

Exemple VI.7.11. L'application linéaire f de \mathbb{R}^4 dans \mathbb{R}^5 définie par

$$f(x_1, x_2, x_3, x_4) = (x_1 + x_2 - \sqrt{2}x_3, x_1 - x_2 + x_3 - x_4, x_2 + 3x_4, \pi x_1 + e^\pi x_2 + x_3, x_1 - \sqrt{7}x_4)$$

n'est pas surjective. En effet, il n'existe aucune application linéaire surjective de \mathbb{R}^4 dans \mathbb{R}^5 .

La proposition suivante est une application immédiate du théorème du rang et sa démonstration est laissée au lecteur.

Proposition VI.7.12. Soit E et F de dimensions finies, et $f \in \mathcal{L}(E, F)$. Les conditions suivantes sont équivalentes :

- i. f est un isomorphisme.
- ii. f est injective et $\dim E = \dim F$.
- iii. f est surjective et $\dim E = \dim F$.

Exemple VI.7.13. Soit F le sous-espace vectoriel de \mathbb{R}^4

$$F = \{(y_1, y_2, y_3, y_4) \in \mathbb{R}^4, y_1 - y_2 + y_3 - y_4 = 0\}.$$

Considérons l'application linéaire

$$\begin{aligned} f : \mathbb{R}^3 &\rightarrow F \\ (x_1, x_2, x_3) &\mapsto (x_1 + x_2, x_1 - x_2, x_3 - x_2, x_3 + x_2). \end{aligned}$$

L'application est bien définie : on vérifie facilement que $(x_1 + x_2, x_1 - x_2, x_3 - x_2, x_3 + x_2)$ est toujours un élément de F . Montrons que f est bijective :

- L'espace vectoriel F est de dimension 3 (c'est le noyau d'une application linéaire de rang 1). Donc $\dim F = \dim \mathbb{R}^3$.
- On vérifie facilement que $\ker f = \{\vec{0}\}$, donc que f est injective.
- Par les deux points précédents et la proposition VI.7.12, f est un isomorphisme de \mathbb{R}^3 dans F .

VI.7.d. Matrices inversibles et isomorphismes

Proposition VI.7.14. *Soit E et F deux espaces de dimension finie, \mathcal{B} et \mathcal{C} des bases respectives de E et F , et $f \in \mathcal{L}(E, F)$. Soit $A = \text{Mat}_{\mathcal{B}, \mathcal{C}}(f)$. Les conditions suivantes sont équivalentes :*

- i. l'application f est un isomorphisme de E dans F ;*
- ii. $\dim E = \dim F$ et la matrice A est inversible.*

Sous ces conditions, on a alors :

$$\text{Mat}_{\mathcal{C}, \mathcal{B}}(f^{-1}) = A^{-1}.$$

Démonstration. Supposons que f est un isomorphisme. Par la proposition VI.7.9, $\dim E = \dim F$. Notons n leur dimension commune.

Par la formule donnant la matrice de la composée de deux applications linéaires (point (ii) de la proposition VI.3.8),

$$\text{Mat}_{\mathcal{B}, \mathcal{C}}(f) \text{Mat}_{\mathcal{C}, \mathcal{B}}(f^{-1}) = \text{Mat}_{\mathcal{C}, \mathcal{C}}(f \circ f^{-1}) = \text{Mat}_{\mathcal{C}, \mathcal{C}}(\text{Id}_F) = I_n,$$

ce qui montre que $\text{Mat}_{\mathcal{B}, \mathcal{C}}(f)$ est inversible.

Réciproquement, on suppose que $\dim E = \dim F$ et que $\text{Mat}_{\mathcal{B}, \mathcal{C}}(f)$ est inversible. Soit g l'application linéaire de F dans E telle que $\text{Mat}_{\mathcal{C}, \mathcal{B}}(g) = \text{Mat}_{\mathcal{B}, \mathcal{C}}(f)^{-1}$. Alors

$$\text{Mat}_{\mathcal{C}, \mathcal{C}}(f \circ g) = \text{Mat}_{\mathcal{B}, \mathcal{C}}(f) \text{Mat}_{\mathcal{C}, \mathcal{B}}(g) = I_n, \quad \text{Mat}_{\mathcal{B}, \mathcal{B}}(g \circ f) = \text{Mat}_{\mathcal{C}, \mathcal{B}}(g) \text{Mat}_{\mathcal{B}, \mathcal{C}}(f) = I_n,$$

où $n = \dim E = \dim F$, ce qui montre que $f \circ g$ est l'identité de F et $g \circ f$ est l'identité de E : f est donc bijective, d'application réciproque g . □

VI.7.e. Isomorphisme entre espaces de mêmes dimensions

Le théorème suivant est un résultat simple, mais important sur les isomorphismes entre espaces vectoriels de dimension finie :

Théorème VI.7.15. *Soit E et F deux espaces vectoriels de dimension finie. Alors E et F sont isomorphes si et seulement si $\dim E = \dim F$.*

Démonstration. Si E et F sont isomorphes, $\dim E = \dim F$ par le point (iii) de la proposition VI.7.9.

Réciproquement, supposons $\dim E = \dim F$, et notons n leur dimension commune. Soit $(\vec{u}_1, \dots, \vec{u}_n)$ une base de E , $(\vec{v}_1, \dots, \vec{v}_n)$ une base de F , et f l'unique application linéaire de E dans F telle que

$$f(\vec{u}_j) = \vec{v}_j, \quad j = 1 \dots n$$

(on rappelle que l'on peut définir une application linéaire sur un espace vectoriel de dimension finie en donnant l'image d'une base, cf le corollaire VI.1.6). Vérifions que f est injective. Soit \vec{x} un élément de $\text{Ker}(f)$, x_1, \dots, x_n ses coordonnées dans la base $\vec{u}_1, \dots, \vec{u}_n$. Alors

$$f(\vec{x}) = \vec{0}_F \text{ i.e. } f(x_1\vec{u}_1 + x_2\vec{u}_2 + \dots + x_n\vec{u}_n) = \vec{0}_F.$$

VI. Applications linéaires

En développant le terme de gauche de la dernière égalité par linéarité, et en utilisant que $f(\vec{u}_j) = \vec{v}_j$ pour $j = 1 \dots n$, on obtient

$$x_1 \vec{v}_1 + \dots + x_n \vec{v}_n = \vec{0}_F,$$

et comme la famille $\vec{v}_1, \dots, \vec{v}_n$ est libre,

$$x_1 = x_2 = \dots = x_n = 0.$$

On en déduit $\vec{x} = \vec{0}_E$. On a montré $\text{Ker}(f) = \{\vec{0}_E\}$, c'est à dire que f est injective. Par le théorème du rang, $\dim \text{Im}(f) = \dim E - \dim \text{Ker}(f) = \dim(E)$ donc f est surjective. Finalement f est bijective. \square

Exemple VI.7.16. Les espaces vectoriels \mathbb{R}^3 et F de l'exemple VI.7.13 sont isomorphes car tous les deux de dimension 3. L'application linéaire f de l'exemple VI.7.13 donne un isomorphisme explicite.

Exercice VI.7.17. Montrer de la même manière que $\dim E \leq \dim F$ si et seulement si il existe une injection de E dans F , ou encore si et seulement si il existe une surjection de F dans E .

(cf correction p. 122).

VI.8. Complément : changement de bases

On donne ici une formule générale permettant de changer les bases de la matrice de représentation d'une application linéaire. Pour déduire $\text{Mat}_{\tilde{\mathcal{B}}, \tilde{\mathcal{C}}}(f)$ de $\text{Mat}_{\mathcal{B}, \mathcal{C}}(f)$ dans l'exemple de §VI.2.d, on a eu besoin :

- des coordonnées des vecteurs de la base $\tilde{\mathcal{B}}$ dans la base \mathcal{B} ;
- des coordonnées des vecteurs de la base \mathcal{C} dans la base $\tilde{\mathcal{C}}$.

Ces informations sont données par les *matrices de passage* d'une base à une autre, que nous allons définir maintenant. La formule de changement de base proprement dite, qui utilise ces matrices de passage, sera vue en §VI.8.b

VI.8.a. Matrices de passage

Définition VI.8.1. Soit E un espace vectoriel de dimension finie n , $\mathcal{B} = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)$ et $\tilde{\mathcal{B}} = (\vec{\tilde{u}}_1, \vec{\tilde{u}}_2, \dots, \vec{\tilde{u}}_n)$ deux bases de E . La *matrice de passage* de \mathcal{B} à $\tilde{\mathcal{B}}$, notée $\text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}}$, est la matrice $n \times n$ dont les vecteurs colonnes sont les coordonnées des vecteurs de la base $\tilde{\mathcal{B}}$ dans la base \mathcal{B} . En d'autres termes, si l'on note $\text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} = [p_{ij}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, on a

$$\forall j \in \{1, \dots, n\}, \quad \vec{\tilde{u}}_j = \sum_{i=1}^n p_{ij} \vec{u}_i.$$

Exemple VI.8.2. Soit \mathcal{B} la base canonique de \mathbb{K}^n et $\mathcal{C} = (\vec{u}_1, \dots, \vec{u}_n)$ une autre base de \mathbb{K}^n . La matrice de passage de \mathcal{B} à \mathcal{C} est la matrice $n \times n$ dont le j -ième vecteur colonne est constitué des coordonnées de \vec{u}_j . Supposons $n = 3$ et considérons la famille

$$\mathcal{C} = \left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \right).$$

On vérifie facilement que \mathcal{C} est une base de \mathbb{R}^3 . On obtient sans calcul la matrice de passage de \mathcal{B} à \mathcal{C} :

$$\text{Mat}_{\mathcal{B} \rightarrow \mathcal{C}} = \begin{bmatrix} 1 & -2 & 4 \\ 1 & 3 & 5 \\ 0 & 0 & 6 \end{bmatrix}.$$

Exemple VI.8.3. On reprend les notations de l'exemple de l'exemple de VI.2.d p.101. Alors, d'après la première étape de cet exemple :

$$\text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}.$$

D'après la deuxième étape :

$$\text{Mat}_{\tilde{\mathcal{C}} \rightarrow \mathcal{C}} = \begin{bmatrix} 2 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Proposition VI.8.4. *Sous les hypothèses de la définition précédente, on note*

$$P = \text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}}.$$

Soit $\vec{x} \in E$, $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ ses coordonnées dans \mathcal{B} et $\tilde{X} = \begin{bmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{bmatrix}$ ses coordonnées dans $\tilde{\mathcal{B}}$. Alors

$$X = P\tilde{X}.$$

Démonstration. On a

$$\vec{x} = \sum_{j=1}^n \tilde{x}_j \vec{u}_j = \sum_{j=1}^n \tilde{x}_j \sum_{i=1}^n p_{ij} \vec{u}_i = \sum_{i=1}^n \left(\sum_{j=1}^n p_{ij} \tilde{x}_j \right) \vec{u}_i,$$

et donc

$$x_i = \sum_{j=1}^n p_{ij} \tilde{x}_j,$$

ce qui montre la formule annoncée. □

Avertissement VI.8.5. L'appellation "matrice de passage de \mathcal{B} à $\tilde{\mathcal{B}}$ " peut être source de confusion : la matrice de passage de \mathcal{B} à $\tilde{\mathcal{B}}$ transforme les coordonnées d'un vecteur dans $\tilde{\mathcal{B}}$ en celles de ce même vecteur dans \mathcal{B} (et non l'inverse, comme on pourrait s'y attendre).

Le lecteur pourra vérifier que la matrice de passage de \mathcal{B} à $\tilde{\mathcal{B}}$ est en fait la matrice de représentation, dans la base \mathcal{B} , de l'application linéaire qui envoie les vecteurs de \mathcal{B} dans ceux de $\tilde{\mathcal{B}}$, ce qui explique son nom :

$$\text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} = \text{Mat}_{\mathcal{B}, \mathcal{B}}(f), \quad f \in \mathcal{L}(E, E), \quad \forall i = 1 \dots n, \quad f(\vec{u}_i) = \vec{\tilde{u}}_i.$$

Corollaire VI.8.6. *Toute matrice de passage est inversible. De plus :*

$$\text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} = \left(\text{Mat}_{\tilde{\mathcal{B}} \rightarrow \mathcal{B}} \right)^{-1}.$$

VI. Applications linéaires

Démonstration. Soit P la matrice de passage de \mathcal{B} à $\tilde{\mathcal{B}}$ et Q la matrice de passage de $\tilde{\mathcal{B}}$ à \mathcal{B} . Alors, si $\vec{x} \in E$ a pour coordonnées X dans la base \mathcal{B} et \tilde{X} dans la base $\tilde{\mathcal{B}}$, on a, par la proposition VI.8.4 :

$$QX = \tilde{X} \text{ et } X = P\tilde{X}.$$

D'où

$$\forall X \in \mathbb{K}^n, \quad X = PQX.$$

On en déduit que l'application linéaire $X \rightarrow PQX$ est l'identité de \mathbb{K}^n , c'est à dire :

$$PQ = I_n.$$

□

Exemple VI.8.7. On revient à l'exemple de VI.2.d. Alors,

$$\text{Mat}_{\tilde{\mathcal{B}} \rightarrow \mathcal{B}} = \left(\begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} \right)^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}.$$

En d'autres termes,

$$\vec{u}_1 = \vec{\tilde{u}}_1 \text{ et } \vec{u}_2 = -\vec{\tilde{u}}_1 + 2\vec{\tilde{u}}_2,$$

ce que l'on peut vérifier par un calcul direct à l'aide des expressions de \vec{u}_1 , \vec{u}_2 , $\vec{\tilde{u}}_1$ et $\vec{\tilde{u}}_2$.

De même, en calculant l'inverse de la matrice $\text{Mat}_{\tilde{\mathcal{C}} \rightarrow \mathcal{C}}$, on obtient

$$\text{Mat}_{\mathcal{C} \rightarrow \tilde{\mathcal{C}}} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}.$$

VI.8.b. Formule de changement de bases

Théorème VI.8.8. Soit $f \in \mathcal{L}(E, F)$, \mathcal{B} et $\tilde{\mathcal{B}}$ des bases de E , \mathcal{C} et $\tilde{\mathcal{C}}$ des bases de F . Alors

$$\begin{aligned} \text{Mat}_{\tilde{\mathcal{B}}, \tilde{\mathcal{C}}}(f) &= \text{Mat}_{\tilde{\mathcal{C}} \rightarrow \mathcal{C}} \text{Mat}_{\mathcal{B}, \mathcal{C}}(f) \text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} \\ &= (\text{Mat}_{\mathcal{C} \rightarrow \tilde{\mathcal{C}}})^{-1} \text{Mat}_{\mathcal{B}, \mathcal{C}}(f) \text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} \end{aligned}$$

Démonstration. Ceci découle de la formule de changement de coordonnées de la proposition VI.8.4.

Soit $\vec{x} \in E$, X le vecteur colonne de ses coordonnées dans \mathcal{B} et \tilde{X} le vecteur colonne de ses coordonnées dans $\tilde{\mathcal{B}}$. Alors, par la proposition VI.8.4, les coordonnées de $f(\vec{x})$ dans \mathcal{C} sont :

$$Y = \text{Mat}_{\mathcal{B}, \mathcal{C}}(f)X = \text{Mat}_{\mathcal{B}, \mathcal{C}}(f) \text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} \tilde{X}$$

et les coordonnées de $f(\vec{x})$ dans $\tilde{\mathcal{C}}$ sont donc :

$$\tilde{Y} = \text{Mat}_{\tilde{\mathcal{C}} \rightarrow \mathcal{C}} Y = \text{Mat}_{\tilde{\mathcal{C}} \rightarrow \mathcal{C}} \text{Mat}_{\mathcal{B}, \mathcal{C}}(f) \text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} \tilde{X}.$$

Or par définition de $\text{Mat}_{\tilde{\mathcal{B}}, \tilde{\mathcal{C}}}(f)$,

$$\tilde{Y} = \text{Mat}_{\tilde{\mathcal{B}}, \tilde{\mathcal{C}}}(f) \tilde{X}.$$

On a donc :

$$\forall \tilde{X} \in \mathbb{K}^n, \quad \text{Mat}_{\tilde{\mathcal{C}} \rightarrow \mathcal{C}} \text{Mat}_{\mathcal{B}, \mathcal{C}}(f) \text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} \tilde{X} = \text{Mat}_{\tilde{\mathcal{B}}, \tilde{\mathcal{C}}}(f) \tilde{X},$$

ce qui montre comme annoncé :

$$\text{Mat}_{\tilde{\mathcal{C}} \rightarrow \mathcal{C}} \text{Mat}_{\mathcal{B}, \mathcal{C}}(f) \text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} = \text{Mat}_{\tilde{\mathcal{B}}, \tilde{\mathcal{C}}}(f).$$

□

Remarque VI.8.9. On peut résumer le théorème VI.8.8 par le schéma suivant

$$\begin{array}{ccc} (E, \mathcal{B}) & \xrightarrow{\text{Mat}_{\mathcal{B}, \mathcal{C}}(f)} & (F, \mathcal{C}) \\ \uparrow P & & \downarrow Q^{-1} \\ (E, \tilde{\mathcal{B}}) & \xrightarrow{\text{Mat}_{\tilde{\mathcal{B}}, \tilde{\mathcal{C}}}(f)} & (F, \tilde{\mathcal{C}}) \end{array}$$

où $P = \text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}}$ permet de passer des coordonnées dans $\tilde{\mathcal{B}}$ aux coordonnées dans \mathcal{B} , et $Q = \text{Mat}_{\mathcal{C} \rightarrow \tilde{\mathcal{C}}}$ permet de passer des coordonnées dans $\tilde{\mathcal{C}}$ aux coordonnées dans \mathcal{C} (cf proposition VI.8.4).

Exemple VI.8.10. On revient à l'exemple donné en VI.2.d. On a :

$$\text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}, \quad \text{Mat}_{\tilde{\mathcal{C}} \rightarrow \mathcal{C}} = \begin{bmatrix} 2 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

et

$$\text{Mat}_{\mathcal{B}, \mathcal{C}}(f) = \begin{bmatrix} -1 & 1 \\ 2 & 0 \\ 3 & 2 \end{bmatrix}.$$

On en déduit, par le théorème VI.8.8, puis un calcul simple de produits matriciels :

$$\begin{aligned} \text{Mat}_{\tilde{\mathcal{B}}, \tilde{\mathcal{C}}}(f) &= \text{Mat}_{\tilde{\mathcal{C}} \rightarrow \mathcal{C}} \text{Mat}_{\mathcal{B}, \mathcal{C}}(f) \text{Mat}_{\mathcal{B} \rightarrow \tilde{\mathcal{B}}} \\ &= \begin{bmatrix} 2 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & 0 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -5 & -\frac{5}{2} \\ 3 & 1 \\ 3 & \frac{5}{2} \end{bmatrix}. \end{aligned}$$

On retrouve bien le résultat de l'exemple de VI.2.d, p.101.

VI.9. Correction de quelques exercices

Correction de l'exercice VI.1.4

Soit R la distance entre l'origine et M , et \mathcal{C} le cercle de centre M et de rayon R . Notons $O = (0, 0)$ l'origine, $A = f(0, 0)$. Par définition d'une rotation, A est le point de \mathcal{C} tel que l'angle $(\overrightarrow{MO}, \overrightarrow{MA})$ soit égal à θ . Puisque θ n'est pas congru à 0 modulo 2π , on en déduit que O et A ne sont pas confondus, et donc que $f(0, 0) \neq (0, 0)$. Par la proposition VI.1.3, l'application f n'est pas linéaire .

Correction de l'exercice VI.1.8.

VI. Applications linéaires

On exprime les vecteurs de la base canonique en fonction de $\vec{u} = (1, 3)$ et $\vec{v} = (1, 4)$. En résolvant deux systèmes linéaires à deux équations et deux inconnues, on obtient :

$$(1, 0) = 4\vec{u} - 3\vec{v}, \quad (0, 1) = -\vec{u} + \vec{v}.$$

D'où

$$f(x, y) = f(x(1, 0) + y(0, 1)) = f(x(4\vec{u} - 3\vec{v}) + y(-\vec{u} + \vec{v})) = (4x - y)(-1, 2, 3),$$

où pour la dernière égalité on a utilisé que $f(\vec{u}) = (-1, 2, 3)$ et $f(\vec{v}) = (0, 0, 0)$.

Correction de l'exercice VI.3.11.

a. Par la proposition VI.3.8,

$$\text{Mat}_{\mathcal{B}, \mathcal{D}}(g \circ f) = \text{Mat}_{\mathcal{C}, \mathcal{D}}(g) \text{Mat}_{\mathcal{B}, \mathcal{C}}(f) = \begin{pmatrix} -3 & 1 \\ 7 & 2 \end{pmatrix}.$$

b. On a envie de dire que AB est une matrice de représentation de l'application $f \circ g$ (qui est bien définie, comme application linéaire de \mathbb{R}^3 dans \mathbb{R}^3). Mais

$$AB = \text{Mat}_{\mathcal{B}, \mathcal{C}}(f) \text{Mat}_{\mathcal{C}, \mathcal{D}}(g),$$

la base \mathcal{D} de \mathbb{R}^2 considéré comme l'"espace d'arrivée" de g n'est donc pas la même que la base \mathcal{B} de \mathbb{R}^2 considéré comme "espace de départ" de f . Le produit AB n'a donc pas d'interprétation particulière en terme de composition d'applications linéaires.

Réponses à l'exercice VI.4.3.

On a :

$$\begin{aligned} f^{-1}(\mathbb{R}) &= \mathbb{R}, & f(\mathbb{R}) &= [-1, 1], & f^{-1}([-1, 1]) &= \mathbb{R}, & f^{-1}(\{0\}) &= \left\{ k\pi + \frac{\pi}{2}, k \in \mathbb{Z} \right\} \\ f([0, \pi]) &= [-1, 1], & f([0, \pi/2]) &= [0, 1], & f(\{0\}) &= \{1\} \\ f(f^{-1}(\mathbb{R})) &= [-1, 1], & f^{-1}(f([0, \pi])) &= \mathbb{R}. \end{aligned}$$

Correction de l'exercice VI.4.9.

Soit f l'application linéaire de \mathbb{R}^4 dans \mathbb{R}^3 définie par

$$f(x_1, x_2, x_3, x_4) = (x_1 + x_2, x_3 - x_4, x_1 - x_3).$$

Soit

$$G = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \text{ t.q. } 2x_1 - x_2 + x_3 - x_4 = 0\}.$$

C'est un sous-espace vectoriel de \mathbb{R}^4 (c'est le noyau d'une application linéaire de \mathbb{R}^4 dans \mathbb{R}).

On a

$$A = f(G).$$

Donc A est l'image d'un sous-espace vectoriel de \mathbb{R}^4 par l'application linéaire f , ce qui montre que c'est un sous-espace vectoriel de \mathbb{R}^3 .

Correction de l'exercice VI.7.17.

Soit E et F des espaces vectoriels, de dimensions respectives n et p .

On sait déjà (proposition VI.7.9) que s'il existe une application linéaire injective de E dans F ou une application linéaire surjective de F dans E , $\dim E \leq \dim F$. Montrons la réciproque. On note $\mathcal{B} = (\vec{u}_1, \dots, \vec{u}_n)$ une base de E , et $\mathcal{C} = (\vec{v}_1, \dots, \vec{v}_p)$ une base de F . Supposons $n \leq p$.

- Soit f l'application linéaire de E dans F définie par

$$f(\vec{u}_j) = \vec{v}_j, \quad j = 1 \dots n.$$

La famille $(f(\vec{u}_1), \dots, f(\vec{u}_n)) = (\vec{v}_1, \dots, \vec{v}_n)$, extraite d'une famille libre, est libre. Par la proposition VI.5.7, f est injective.

- Soit g l'application linéaire de F dans E définie par

$$g(\vec{v}_j) = \vec{v}_j, \quad j = 1 \dots n, \quad g(\vec{v}_j) = \vec{0}, \quad j = n + 1 \dots p.$$

La famille $(g(\vec{v}_1), \dots, g(\vec{v}_n)) = (\vec{u}_1, \dots, \vec{u}_n, 0, \dots, 0)$, qui complète la base $(\vec{u}_1, \dots, \vec{u}_n)$ est génératrice. Par la proposition VI.5.7, f est surjective.

Table des matières

I. Les nombres complexes	1
I.1. Les nombres réels ne suffisent pas	1
I.1.a. L'équation du second degré à coefficients réels	1
I.1.b. Un peu d'histoire	3
I.2. Forme cartésienne d'un nombre complexe, addition et multiplication	4
I.2.a. Rappel : produit cartésien de deux ensembles	4
I.2.b. Construction des nombres complexes	4
I.2.c. Propriétés de l'addition et de la multiplication	5
I.2.d. Représentation dans le plan	5
I.3. Autres opérations sur les nombres complexes	6
I.3.a. Opposé, différence de nombres complexes	6
I.3.b. Conjugaison et module	6
I.3.c. Inverse et quotient	8
I.4. Forme polaire d'un nombre complexe	9
I.4.a. Rappels	9
I.4.b. Définition	9
I.4.c. Propriétés	10
I.4.d. Écriture exponentielle de la forme polaire	10
I.4.e. Formule d'Euler	11
I.4.f. Formule de Moivre	12
I.5. Racines $n^{\text{ièmes}}$ d'un nombre complexe	13
I.5.a. Cas général	13
I.5.b. Racines $n^{\text{ièmes}}$ de l'unité.	14
I.5.c. Racines carrées d'un nombre complexe, sous forme cartésienne	15
I.6. Équation du second degré à coefficients complexes	16
I.7. Appendice : quelques rappels	17
I.7.a. Principe du raisonnement par récurrence	17
I.7.b. Formule du binôme	17
I.7.c. Suites arithmétiques et géométriques	18
II. Les polynômes	21
II.1. Définitions	21
II.1.a. Polynômes comme suites finies	21
II.1.b. Addition	21
II.1.c. Indéterminée	22
II.1.d. Multiplication	22
II.2. Premières propriétés	24
II.2.a. Division euclidienne	24
II.2.b. Fonctions polynomiales	25

II.2.c. Polynôme dérivé	25
II.3. Racines	27
II.3.a. Cas général	27
II.3.b. Polynômes à coefficients complexes	28
II.3.c. Polynômes à coefficients réels	28
II.4. Complément : polynômes irréductibles	29
II.4.a. Cas général	29
II.4.b. Polynômes irréductibles de $\mathbb{C}[X]$	30
II.4.c. Polynômes irréductibles de $\mathbb{R}[X]$	30
III. Systèmes linéaires	33
III.1. Définitions et premiers exemples	33
III.1.a. Définitions	33
III.1.b. Exemples de petits systèmes linéaires	35
III.1.c. Notation matricielle	37
III.2. Méthode du pivot	38
III.2.a. Systèmes équivalents. Opérations élémentaires	38
III.2.b. Forme échelonnée	40
III.2.c. Méthode du pivot de Gauss	44
III.2.d. Système de Cramer	49
III.3. Système avec paramètres	50
III.4. Réponse à certains exercices	51
IV. Introduction aux matrices	55
IV.1. Définitions. Opérations sur les matrices	55
IV.1.a. Définitions	55
IV.1.b. Multiplication par un scalaire et additions	56
IV.1.c. Transposition	57
IV.1.d. Multiplication des matrices	57
IV.1.e. Systèmes linéaires et matrices	61
IV.1.f. Formule du binôme	62
IV.2. Matrices inversibles : définitions et exemples	62
IV.2.a. Définition	62
IV.2.b. Matrices diagonales	64
IV.2.c. Inversibilité des matrices 2×2	65
IV.2.d. Stabilité par multiplication et transposition	65
IV.3. Opérations sur les lignes et inversion de matrices	66
IV.3.a. Matrices élémentaires	66
IV.3.b. Matrices échelonnées réduites carrées	68
IV.3.c. Inversions de matrices par la méthode du pivot de Gauss	70
IV.3.d. Caractérisation des matrices inversibles	72
IV.3.e. Système de Cramer et matrice inversible	72
V. Espaces vectoriels	75
V.1. Définitions et exemples	75
V.1.a. L'espace vectoriel \mathbb{K}^n	75
V.1.b. Espaces vectoriels généraux	76

V.1.c. Exemples	77
V.2. Sous-espace vectoriels	77
V.2.a. Deux définitions équivalentes	77
V.2.b. Intersection	78
V.2.c. Sous-espace vectoriel engendré par une famille de vecteurs	79
V.2.d. Somme, somme directe, supplémentaires	80
V.3. Familles de vecteurs	82
V.3.a. Familles de vecteurs : définition	82
V.3.b. Familles libres	83
V.3.c. Familles génératrices	85
V.3.d. Bases	85
V.4. Espaces vectoriels de dimension finie	86
V.4.a. Définition	86
V.4.b. Existence de bases	87
V.4.c. Dimension d'un espace vectoriel	87
V.4.d. Caractérisation des bases	89
V.4.e. Théorème de la base incomplète	90
V.5. Sous-espaces vectoriels et dimension	90
V.5.a. Dimension d'un sous-espace vectoriel	90
V.5.b. Dimension de la somme de deux sous-espaces vectoriels	91
V.5.c. Description des sous-espaces vectoriels de \mathbb{K}^n	92
V.5.d. Manipulation de familles de vecteurs de \mathbb{K}^n	94
VI. Applications linéaires	97
VI.1. Définitions et propriétés	97
VI.1.a. Définitions et exemples	97
VI.1.b. Quelques propriétés	98
VI.2. Applications linéaires et matrices	99
VI.2.a. Matrice de représentation d'une application linéaire	99
VI.2.b. Exemple : cas des bases canoniques	99
VI.2.c. Cas général	100
VI.2.d. Un exemple de changement de base	101
VI.2.e. Applications linéaires et multiplication par une matrice	102
VI.3. Opérations sur les applications linéaires	103
VI.3.a. Addition et multiplication par un scalaire	103
VI.3.b. Composition	104
VI.3.c. Effet des opérations sur les matrices	105
VI.4. Applications linéaires et sous-espaces vectoriels	106
VI.4.a. Image et image réciproque d'un ensemble par une application	106
VI.4.b. Cas des applications linéaires	107
VI.5. Injectivité et surjectivité	108
VI.5.a. Injection, surjection : rappels	108
VI.5.b. Injectivité, surjectivité, image et noyau	109
VI.5.c. Injectivité, surjectivité et familles de vecteurs	109
VI.6. Rang d'une application linéaire	110
VI.6.a. Définition	110
VI.6.b. Théorème du rang	111

Table des matières

VI.6.c. Rang et systèmes linéaires	113
VI.7. Isomorphismes	114
VI.7.a. Définition	114
VI.7.b. Application réciproque d'un isomorphisme	114
VI.7.c. Condition sur les dimensions	116
VI.7.d. Matrices inversibles et isomorphismes	117
VI.7.e. Isomorphisme entre espaces de mêmes dimensions	117
VI.8. Complément : changement de bases	118
VI.8.a. Matrices de passage	118
VI.8.b. Formule de changement de bases	120
VI.9. Correction de quelques exercices	121