ELI 3 MATHEMATIQUES
remise à niveau

Fabrice Bethuel
Université Pierre et Marie Curie

Année Universitaire 2012-2013

1 ESPACES VECTORIELS

Dans tout ce chapitre, l'ensemble $\mathbb{K}$ désigne soit l'ensemble des nombres réels $\mathbb{R}$ soit l'ensemble des nombres complexes $\mathbb{C}$, et n désigne un entier supérieur ou égal à 1.

1.1 L'ensemble $\mathbb{K}^n$

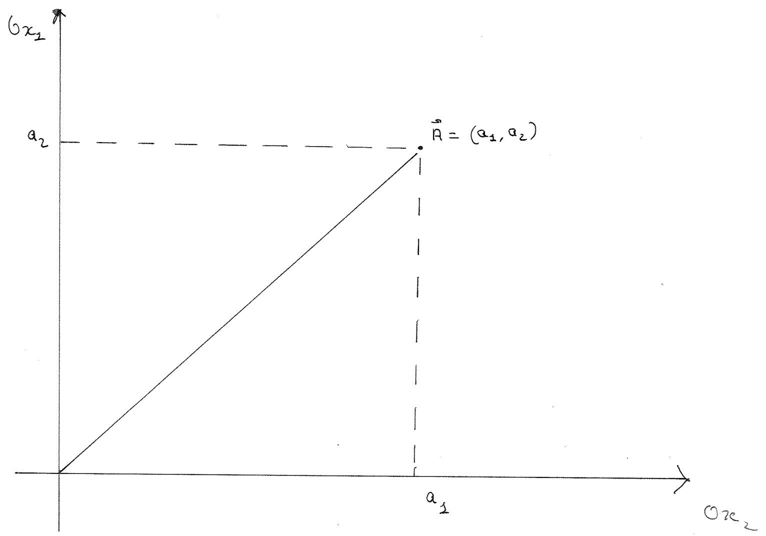
L'ensemble $\mathbb{K}^n$ que nous allons décrire ci-dessous offre l'exemple le plus simple, et peut être le plus parlant de la notion d'espace vectoriel. Rappelons tout d'abord qu'un n-uplet d'éléments de $\mathbb{K}$ est une liste *ordonnée* de n éléments de $\mathbb{K}$ que l'en écrit sous la forme

$$(x_1, x_2, x_3, \ldots, x_n)$$

et que nous noterons souvent $\vec{x} = (x_1, x_2, x_3, \ldots, x_n)$. L'ensemble $\mathbb{K}^n$ désigne alors l'ensemble de tous les n-uplets de $\mathbb{K}$, c'est à dire

$$\mathbb{K}^n = \{\vec{x} = (x_1, x_2, x_3, \ldots, x_n), x_i \in \mathbb{K} \text{ pour } i = 1, \ldots, n\}.$$

Rappelons que dans le cas $n = 2$, et $\mathbb{K} = \mathbb{R}$ les éléments de $\mathbb{R}^2$ peuvent s'identifier aux points du plan, si on a défini au préalable un systèmes d'axe de coordonnées.



1.2 Opérations dans $\mathbb{K}^n$

On définit deux opérations sur $\mathbb{K}^n$: l'addition et la multiplication par un scalaire.

1.2.1 L'addition

Soient deux éléments $\vec{x} = (x_1, x_2, x_3, \ldots, x_n)$ et $\vec{y} = (y_1, y_2, y_3, \ldots, y_n)$ deux éléments de $\mathbb{K}^n$. On définit la somme de ces deux éléments par la formule

$$\vec{x} + \vec{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3, \ldots, x_n + y_n) \in \mathbb{K}^n.$$

Cette opération a alors les propriétés suivantes :

- **Commutativité**, c'est à dire pour tous éléments $\vec{x}$ and $\vec{y}$ dans $\mathbb{K}^n$, on a

$$\vec{x} + \vec{y} = \vec{y} + \vec{x}. \tag{1}$$

- **Associativité**. Pour tous éléments $\vec{x}$, $\vec{y}$ et $\vec{z}$ dans $\mathbb{K}^n$ on a

$$\vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}. \tag{2}$$

- **Existence d'un élément neutre**. En effet l'élément $\vec{0} = (0, \ldots, 0)$ est un élément neutre, c'est à dire que l'on a pour tout $\vec{x} \in \mathbb{K}^n$

$$\vec{x} + \vec{0} = \vec{x}. \tag{3}$$

- **Existence d'un inverse.** Soit $\vec{x} = (x_1, x_2, x_3, \ldots, x_n)$ un élément de $\mathbb{K}^n$. Posons

$$-\vec{x} = (-x_1, -x_2, -x_3, \ldots, -x_n)$$

  Alors

$$-\vec{x} + (-\vec{x}) = \vec{0}. \tag{4}$$

1.2.2 La multiplication par un scalaire

Les éléments de $\mathbb{K}$ seront dorénavant appelés *scalaires*, par opposition aux éléments de $\mathbb{K}^n$ qui seront appelés *vecteurs*. Soit $\lambda \in \mathbb{K}$ un scalaire et soit $\vec{x} = (x_1, x_2, x_3, \ldots, x_n)$ un élément de $\mathbb{K}^n$. On définit la multiplication du vecteur $\vec{x}$ par le scalaire λ et on note $\lambda.\vec{x}$ l'opération suivante :

$$\lambda \cdot \vec{x} = (\lambda x_1, \lambda x_2, \lambda x_3, \lambda \ldots, x_n) \in \mathbb{K}^n,$$

c'est à dire, on multiplie chaque *coordonnée* x_i de $\vec{x}$ par le *scalaire* λ. Cette multiplication a les propriétés suivantes :

- **Element neutre multiplicatif.** On a

$$1 \cdot \vec{x} = \vec{x} \tag{5}$$

- **Associativité multiplicative.** On a pour tous scalaires λ, μ dans $\mathbb{K}$ et tout vecteur $\vec{x}$ dans $\mathbb{K}^n$

$$\lambda \cdot (\mu \vec{x}) = (\lambda\mu) \cdot \vec{x}. \tag{6}$$

- **Distributivités par rapport à l'addition.** Pour tout scalaire $\lambda \in \mathbb{K}$ et pour tous vecteurs $\vec{x}, \vec{y}$ dans $\mathbb{K}^n$ on a

$$\lambda \cdot (\vec{x} + \vec{y}) = \lambda \cdot \vec{x} + \lambda \cdot \vec{y}, \tag{7}$$

  et pour tout vecteur $\vec{x}$ et tous scalaire λ, μ dans $\mathbb{K}$

$$(\lambda + \mu).\vec{x} = \lambda.\vec{x} + \mu.\vec{x}, \tag{8}$$

1.3 Généralisations

Les propriétés élémentaires de l'addition et de la multiplication que nous venons de voir entraînent à leur tour de nombreuses autres propriétés, parfois bien plus subtiles, et permettent des construction qui s'avéreront fort utiles. Ces propriétés et constructions se retrouvent en fait pour bien d'autres objets mathématiques : ces objets correspondent à la notion d'espace vectoriel, développée au court du dixneufième siècle, et que nous présentons ci-dessous.

Soit V un ensemble *non vide*. On suppose que V est muni d'une *addition*, c'est à dire d'une application, notée ici $+$, qui a tout couple d'éléments $\vec{x}$ et $\vec{y}$ de V associe un troisième élément de V, noté $\vec{x}+\vec{y}$. On parle alors d'un *loi interne*. On suppose par ailleurs que V est muni d'une multiplication par les scalaires, c'est à dire d'une application qui à tout scalaire $\lambda \in \mathbb{K}$ et tout vecteur $\vec{x}$ in V associe un élément noté $\lambda.\vec{x}$ de V. On appelle une telle opération une *loi externe*.

Définition 1. *On dit que V muni des opérations précédentes est un espace vectoriel sur K si et seulement si les propriétés* (1),(2), (3), (4), (5), (6), (7) , (8) *sont vérifiées.*

La propriété (3) affirme ici l'existence d'un élément $\vec{0}$ vérifiant (3), et de même, la propriété (4) affirme l'existence d'un élément noté $-\vec{x}$ vérifiant (4), appelé *l'inverse* de $\vec{x}$.

Comme premiers exemples d'espace vectoriel, nous avons bien entendu les espaces $\mathbb{K}^n$, qui sont des espaces vectoriels sur $\mathbb{K}$, en particulier $\mathbb{K}$ est lui même un espace vectoriel sur $\mathbb{K}$. Ainsi, d'après ce qui précéde $\mathbb{C}^n$ est un espacce vectoriel sur $\mathbb{C}$, mais c'est aussi un espace vectoriel sur $\mathbb{R}$ (exercice).

Donnons tout de suite un exemple un peu moins élémentaire d'espace vectoriel. Considérons l'ensemble $\mathcal{P}(x)$ des fonctions polynômiales sur $\mathbb{C}$, c'est à dire de la forme

$$p(z) = a_0 + a_z + a_2 z^2 + a_m z^m,$$

pour un entier $m \in \mathbb{N}$ donné, et des nombres complexes $a_0, \ldots, a_m$ également donnés. On définit l'addition sur $\mathcal{P}(X)$ comme suit : si p et q sont deux polynomes, alors le polynôme $p+q$ est défini par

$$(p+q)(z) = p(z) + q(z), \text{ pour tout } z \in \mathbb{C}.$$

Si λ est un scalaire de $\mathbb{C}$ on définit le polynôme $\lambda \cdot p$ par

$$(\lambda \cdot p)(z) = \lambda p(z) \text{ pour tout } z \in \mathbb{C}.$$

On vérifie que ces deux opérations ont bien les propriétés requises, et munissent ainsi $\mathcal{P}(X)$ d'une structure d'espace vectoriel sur $\mathbb{C}$.

Comme premières propriétés élémentaires d'un espace vectoriel V sur $\mathbb{K}$, on a :

$0 \cdot \vec{x} = \vec{0}$, pour tout vecteur $\vec{x}$ de V
$\lambda \cdot \vec{0} = \vec{0}$, pour tout scalaire $\lambda \in \mathbb{K}$
$(-1) \cdot \vec{x} = -\vec{x}$ pour tout vecteur $\vec{x}$ de V.

1.4 Sous-espace vectoriels

Soit V un espace vectoriel, et $F \subset V$ une partie non vide de V. On dit que F est un sous-espace vectoriel de V si et seulement F, muni de l'addition et de multiplication scalaire de V, est un espace vectoriel. La partie F est donc un sous-espace vectoriel de V si et seulement elle contient $\vec{0}$, si la somme de deux vecteurs de F appartient à F, et si la multiplication de tout vecteur de F par tout scalaire de $\mathbb{K}$ appartient à F. On a donc :

Proposition 1. *Soit F une partie d'un espace vectoriel V. F est un sous espace-vectoriel de V si et seulement :*
- *F contient $\vec{0}$.*
- *$\forall \lambda \in K, \forall \mu \in \mathbb{K}, \forall \vec{x} \in F, \forall \vec{y} \in F$, on a $\lambda\vec{x} + \mu\vec{y} \in F$.*

Comme premier exemple élémentaire de sous-espace vectoriel, notons que le singleton $\{\vec{0}\}$ de V est un sous-espace vectoriel. De même V est évident un sous-espace vectoriel de lui-même.

Exemples. Soit $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ un élément non nul de $\mathbb{K}^n$. Posons

$$F = \{(z_1, z_2, \ldots, z_n) \in \mathbb{K}^n, \text{ tels que } \alpha_1 z_1 + \alpha_2 z_2 + \ldots + \alpha_n z_n = 0\}$$

alors F est un sous-espace vectoriel de $\mathbb{K}^n$. En revanche, si $\beta \in \mathbb{K}$ et si on considère l'ensemble

$$F_\beta = \{(z_1, z_2, \ldots, z_n) \in \mathbb{K}^n, \text{ tels que } \alpha_1 z_1 + \alpha_2 z_2 + \ldots + \alpha_n z_n = \beta\}$$

alors F_β est un sous-espace vectoriel si et seulement si $\beta = 0$ (exercice).

Donnons maintenant quelques propriétés :

Proposition 2. *Soit V un espace vectoriel et F et G deux sous-espaces vectoriels de V. Alors l'intersection $F \cap G$ est un sous-espace vectoriel de V.*

Pour deux sous-espaces vectoriels F et G ont défini par ailleurs leur somme $F + G$ par

$$F + G = \{\vec{u} \in V \text{ tels qu'il existe } \vec{x} \in F \text{ et } \vec{y} \in G \text{ tels que } \vec{u} = \vec{x} + \vec{y}\}.$$

On a alors

Proposition 3. *Soit V un espace vectoriel et F et G deux sous-espaces vectoriels de V. Alors la somme $F + G$ est un sous-espace vectoriel de V.*

Notons que $F + G$ contient chacun des sous-espaces F et G, et donc leur réunion $F \cup G$. C'est même le plus petit sous-espace vectoriel contenant cette réunion $F \cup G$, qui en général n'est pas un sous-espace vectoriel de V(donner des exemples).

Définition 2. *On dit que deux espaces F et G sont en somme directe si et seulement si $F \cap G = \{\vec{0}\}$. On note alors*

$$F \oplus G = F + G.$$

On dit que deux espaces F et G sont supplémentaires si et seulement si ils sont en somme directe et $V = F \oplus G$.

On vérifie (exercice) que si F et G sont en somme directe, alors tout $u \in F \oplus G$ s'écrit de manière *unique* sous la forme

$$u = x + y,\ x \in F, y \in G. \tag{9}$$

En particulier, si F et G sont supplémentaires, alors tout vecteur $u \in V$ se décompose de manière unique sous la forme précédente (9). La réciproque est également vraie.

1.4.1 Combinaisons linéaires

Soit V un espace vectoriel sur $\mathbb{K}$. Soit k un entier, et considérons une famille $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ de k vecteurs de V. On définit alors :

Définition 3. *On dit que $\vec{y}$ est une combinaison linéaire de la famille $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ si et seulement si il existe un k-uplet $(\lambda_1, \ldots, \lambda_k)$ d'éléments de $\mathbb{K}$ tel que*

$$y = \lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \ldots \lambda_k\vec{x}_k.$$

On note $\mathrm{Vect}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k)$ l'ensemble des combinaisons linéaires de la famille $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$. On a donc

$$\mathrm{Vect}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k) = \{y = \lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \ldots \lambda_k\vec{x}_k, \lambda_1, \ldots \lambda_k \in \mathbb{K}\}.$$

On vérifie aisément le résultat suivant :

Proposition 4. *Soit $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ une famille de k vecteurs de V. Alors $\mathrm{Vect}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k)$ est un sous-espace vectoriel de V. De plus, il s'agit du plus petit sous-espace vectoriel de V qui contient la famille $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$.*

Donnons immédiatement quelques propriétes élémentaires :
- $\mathrm{Vect}(\vec{0}) = \{\vec{0}\}$
- $\mathrm{Vect}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k) + \mathrm{Vect}(\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_p) = \mathrm{Vect}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k, \vec{y}_1, \vec{y}_2, \ldots, \vec{y}_p)$
- si $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subset \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_\ell\}$ alors

$$\mathrm{Vect}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k) \subset \mathrm{Vect}(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_\ell).$$

La proposition suivante nous sera fort utile par la suite.

Proposition 5. *Le sous-espace vectoriel engendré par une famille de vecteurs donnés est invariant par les opérations suivante :*
- *$T(1)$ Permutation de deux vecteurs dans la liste*
- *$T(2)$ Multiplication d'un vecteur par un scalaire non nul*
- *$T(3)$ Ajout à l'un des vecteurs d'une combinaison linéaire des autres vecteurs.*

Terminons ce paragraphe en introduisons une nouvelle notion, celle de famille génératrice.

Définition 4. *Soit F un sous-espace vectoriel de V, et $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ une famille d'éléments de F. On dit que $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ est une famille génératrice de F si et seulement si*

$$F = \mathrm{Vect}\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\},$$

c'est à dire si et seulement si tout élément de F peut s'écrire comme combinaison linéaire d'éléments de la famille $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$.

2 Familles de vecteurs

Dans la chapitre précédent, nous avons déjà introduit les familles de vecteurs, et terminé sur la notion de famille génératrice. Dans ce chapitre, nous allons introduire une nouvelle notion fondamental, celle de famille libre.

2.1 Familles libres et familles liées

Définition 5. *Soit $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ une famille de k vecteurs de V, espace vectoriel sur $\mathbb{K}$.*

- *On dit que $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ est une famille libre (ou linéairement indépendante) si et seulement si, pour tout k-uplet $(\lambda_1, \lambda_2, \ldots, \lambda_k)$ de scalaires de $\mathbb{K}$, on a*

$$\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \ldots + \lambda_k\vec{x}_k = \vec{0}$$

alors nécessairement

$$\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0.$$

- *On dit que la famille $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ est liée (ou linéairement dépéndante) si elle n'est pas libre, c'est à dire s'il existe un k-uplet $(\lambda_1, \lambda_2, \ldots, \lambda_k) \neq (0, \ldots, 0)$ tel que*

$$\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \ldots + \lambda_k\vec{x}_k = \vec{0}.$$

Par exemple, la famille $\{(2, 31), (1, -1, 2), (7, 3, 8)\}$ est une famille liée de $\mathbb{R}^3$ car

$$2(2, 3, 1) + 3(1, -1, 2) + (-1)(7, 3, 8) = (0, 0, 0).$$

Comme autre exemple, notons qu'une famille contenant le vecteur nul $\vec{0}$ est forcément liée (exercice), qu'une famille contenant deux fois le même vecteur est également liée...

Lorsque la famille est libre, les coefficients des combinaisons linéaires sont déterminés de manière unique, plus précisement :

Proposition 6. *Soit $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ une famille libre de k vecteurs de V. Alors pour tout vecteur $\vec{u}$ de $\mathrm{Vect}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k)$ il existe un unique k-uplet $(\mu_1, \mu_2, \ldots, \mu_k)$ de scalaires de $\mathbb{K}$*

$$\vec{u} = \mu_1\vec{x}_1 + \mu_2\vec{x}_2 + \ldots + \mu_k\vec{x}_k.$$

Preuve L'égalité

$$\vec{u} = \mu_1\vec{x}_1 + \mu_2\vec{x}_2 + \ldots + \mu_k\vec{x}_k = \mu_1'\vec{x}_1 + \mu_2'\vec{x}_2 + \ldots + \mu_k'\vec{x}_k.$$

entraîne

$$(\mu_1 - \mu_1')\vec{x}_1 + (\mu_2 - \mu_2')\vec{x}_2 + \ldots + (\mu_k - \mu_k')\vec{x}_k,$$

et donc, puisque la famille est libre $\mu_1 - \mu_1' = \ldots = \mu_k - \mu_k' = 0$.

Proposition 7. *Soit $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ une famille de k vecteurs de V et soit $\vec{y} \in \mathrm{Vect}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k)$. Alors la famille $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k, y\}$ est liée.*

Réciproquement, si la famille $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k, y\}$ est liée, et que la famille $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ est libre, alors $\vec{y} \in \mathrm{Vect}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k)$.

La première assertion est facile. Pour la seconde, on note que si la famille $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k, y\}$ est liée, alors il existe un $(k+1)$ -uplet $((\lambda_1, \lambda_2, \ldots, \lambda_k, \lambda)$ non identiquement nul tel que

$$\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \ldots + \lambda_k\vec{x}_k + \lambda y = \vec{0}.$$

Vérifions tout d'abord que $\lambda \neq 0$: sinon, on aurait,$\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \ldots + \lambda_k\vec{x}_k = \vec{0}$, ce qui entraîne, comme la famille $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ est supposée libre, que $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0$, ce qui est contraire à l'hypothèse, si $\lambda = 0$.

Comme λ est non nul, on a donc

$$\vec{y} = -\frac{\lambda_1}{\lambda}\vec{x}_1 - \frac{\lambda_2}{\lambda}\vec{x}_2 - \ldots - \frac{\lambda_k}{\lambda}\vec{x}_k,$$

ce qui termine la preuve.

Dans le même esprit, on démontre

Théorème 1. *Une famille* $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ *est liée si et seulement si elle contient un vecteur qui est combinaison linéaire des autres. Si on retire ce vecteur de la famille, l'espace engendré reste inchangé. En particulier, il existe une sous-famille de* $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ *qui soit libre et génératrice de* $\mathrm{Vect}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k)$.

Le résultat qui suit est plus difficile à établir

Théorème 2. *Soit* $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ *une famille d'éléments de* V*, et*

$$F = \mathrm{Vect}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k).$$

Alors toute famille libre de F *contient au plus* k *vecteurs.*

La preuve repose, sur la variante suivante de la Proposition 7 :

Proposition 8. *Soit* $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m\}$ *une famille liée de* V *telle que* $w_1 \neq 0$. *Soit* $F = \mathrm{Vect}(\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m)$.

Alors il existe $j \in 2, \ldots, m$ *telle que*

- $w_j \in \mathrm{Vect}(\vec{w}_1, \ldots, \vec{w}_{j-1})$
- $F = \mathrm{Vect}(\vec{w}_1, \ldots, \vec{w}_{j-1}, w_{j+1}, \ldots, \vec{w}_m)$.

Preuve de la Proposition 8. Comme la famille est liée, il existe un m-uplet non nul $(a_1, \ldots, a_m)$ d'éléments de $\mathbb{K}$ tel que

$$a_1\vec{w}_1 + \ldots + a_m\vec{w}_m = \vec{0}.$$

Les nombres $a_2, \ldots, a_m$ ne peuvent pas tous être nuls, car $\vec{w}_1 \neq 0$. Soit donc j le plus grand entier dans $\{2, \ldots, m\}$tel que $a_j \neq 0$. On a donc

$$\vec{w}_j = -\frac{a_1}{a_j}\vec{w}_1 - \ldots - \frac{a_{j-1}}{a_j}\vec{w}_{j-1}, \tag{10}$$

ce qui prouve la première assertion. Pour établir la seconde assertion, considérons un vecteur $\vec{u} \in F$, de sorte qu'il existe des nombres $c_1, \ldots, c_m$ de $\mathbb{K}$ tels que

$$\vec{u} = c_1\vec{w}_1 + \ldots + c_m\vec{w}_m.$$

Dans cette expression, on remplace alors $\vec{w}_j$ par son expression donnée dans (10), ce qui achève la preuve de la seconde assertion.

Preuve du Théorème 2 Soit $\{\vec{u}_1, \ldots, \vec{u}_m\}$ une famille libre de F. Il nous faut prouver que $m \leq k$. A cet effet, on procède en plusieurs étapes.

Etape 1. D'après la Proposition 7, on sait que la famille $\{\vec{u}_1, \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ est liée. D'après la Proposition 8, on peut enlever de cette collection l'un des x_j, disons x_1 pour simplifier de sorte que

$$F = \mathrm{Vect}(\vec{u}_1, \vec{x}_2, \ldots, x_k)$$

Etape 2. On recommence l'opération. D'après la Proposition 7, on sait que la famille $\{\vec{u}_2, \vec{u}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ est liée. D'après la Proposition 8, l'un de ces vecteurs est combinaison linéaire de ceux qui le précédent dans la liste : comme la famille $\{\vec{u}_1, \ldots, \vec{u}_m\}$ est libre, c'est forcément un des vecteurs $\vec{x}_2, \ \ldots \vec{x}_m$; disons, pour fixer les idées $\vec{x}_2$. On a donc

$$F = \mathrm{Vect}(\vec{u}_1, \vec{u}_2, \vec{x}_3, \ldots, x_k).$$

A l'étape m, on obtient, en raisonnant comme précédemment,

$$F = \mathrm{Vect}(\vec{u}_1, \vec{u}_2, \ldots \vec{u}_m, \vec{x}_{m+1}, \ldots, x_k)$$

et la conclusion en découle.

2.2 La notion de base

Définition 6. *Soit V un espace vectoriel sur $\mathbb{K}$ et F un sous-espace vectoriel de V. On dit que la famille de vecteurs $(\vec{u}_1, \ldots \vec{u}_k)$ est une* base *de F si et seulement si elle est* libre et génératrice.

Tout vecteur $\vec{x}$ de F se décompose alors de manière unique sous la forme

$$\vec{x} = x_1\vec{u}_1 + x_2\vec{u}_2 + \ldots x_k\vec{\vec{u}}_k, \tag{11}$$

où le k-uplet $(x_1, x_2, \ldots, x_k)$ est un élément de $\mathbb{K}^n$, appelé coordonnées de $\vec{x}$ dans la base $(\vec{u}_1, \ldots \vec{u}_k)$.

Notons également que si, réciproquement, tout vecteur se décompose comme une combinaison linéaire unique sur la famille $(\vec{u}_1, \ldots, \vec{u}_k)$ alors cette famille est nécessairement une base (exercice).

Donnons quelques exemples :

1. Dans $\mathbb{R}^3$ (ou $\mathbb{C}^3$) la famille constituée des vecteurs $\vec{e}_1 = (1, 0, 0), \vec{e}_2 = (0, 1, 0), \vec{e}_3 = (0, 0, 1)$ est une base. En effet cette famille est libre et tout vecteur $\vec{x} = (x_1, x_2, x_3)$ de $\mathbb{R}^3$ s'écrit

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3.$$

la base $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ est appelée base canonique de $\mathbb{R}^3$.

2. La famille $(1, 2), (3, 5)$ est une base de $\mathbb{R}^2$.

Le Théorème 1 nous amène directement à l'existence de base pour les espaces vectoriels engendrés par une famille finie.

Lemme 1. *Soit $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m\}$ une famille de V et soit $F = \mathrm{Vect}(\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m)$. Alors F possède une base qui est une sous-famille de la famille $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m\}$.*

Le Théorème 2 a pour sa part une conséquence extrêmement importante et utile.

Théorème 3. *Soit F un sous-espace vectoriel de V, possédant une base $(\vec{u}_1, \ldots \vec{u}_k)$. Alors toutes les autres bases de F ont exactement k éléments.*

On appelle ce nombre commun la *dimension* de F et on le note dim F. On dit alors par ailleurs que F est de *dimension finie.* On conviendra par ailleurs que $\{\vec{0}\}$ est de dimension finie, égale à 0 dim $\{0\}$=0, et que l'ensemble vide en constitue une base.

Preuve. Le Théorème 2 montre que toutes les familles libres ont au plus k éléments. Si on suppose alors par l'absurde qu'il existe une base ayant strictement moins de k éléments, toutes les familles libres auraient alors également moins de k éléments strictement, ce qui est contradictoire avec le fait que $(\vec{u}_1, \ldots \vec{u}_k)$ est libre.

Théorème 4 (Théorème de la base incomplète). *Soit V un espace de dimension finie égale à n. Alors tout sous-espace vectoriel F de V est de dimension finie et*

$$\dim F \leq n.$$

Soit alors $(\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m)$ une base de F. Alors il existe $n-m$ vecteurs de V $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{n-m})$ tels la famille $(\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m, \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_{n-m})$ soit une base de V.

Preuve. Afin de montrer que F est de dimension finie, nous allons construire itérativement une base de F. Si $F = \{\vec{0}\}$, il n'y a rien à faire. Sinon, soit $\vec{w}_1$ un vecteur non nul de F. Si $F = \mathrm{Vect}(\vec{w}_1)$ nous avons terminé. Sinon, il existe un vecteur non nul $\vec{w}_2 \in F \backslash \mathrm{Vect}\, \vec{w}_1$. Notons alors que, grâce à la Proposition 7, la famille $(\vec{w}_1, \vec{w}_2)$ est libre. On recommence alors le raisonnement. Si $F = \mathrm{Vect}(\vec{w}_1, \vec{w}_2)$, nous avons terminé, et dim F=2. Sinon, il existe un vecteur non nul $\vec{w}_3 \in F \backslash \mathrm{Vect}(\vec{w}_1, \vec{w}_2)$, et par la Proposition 7, la famille $(\vec{w}_1, \vec{w}_2, \vec{w}_3)$ est libre. Comme à chaque étape, nous construisons une famille libre de V, et que ces dernières, en vertu du Theorème 2 ont au plus n éléments, l'itération s'arrête en n étapes, au plus.

La deuxième assertion se démontre en utilisant un procédé similaire.

Voyons maintenant quelques conséquences importantes de ce résultat.

Proposition 9. *Soit V un espace vectoriel de dimension finie. Alors pour tout espace vectoriel F de V, il existe un sous-espace vectoriel supplémentaire G de F dans V, c'est à dire tel que $V = F \oplus G$.*

Pour la preuve, en reprenant les notations précédentes, il suffit de vérifier que le sous-espace vectoriel G défini par
$$G = \text{Vect}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{n-m})$$
convient.

Proposition 10. *Soit V un espace vectoriel de dimension finie n, et soit une famille $(\vec{u}_1, \dots, \vec{u}_n)$ de n vecteurs de V. Alors les assertions suivantes sont equivalentes :*
i) La famille $(\vec{u}_1, \dots, \vec{u}_n)$ est libre.
ii) La famille $(\vec{u}_1, \dots, \vec{u}_n)$ est géneratrice.
iii) La famille $(\vec{u}_1, \dots, \vec{u}_n)$ est une base de V.

Preuve Montrons i) $\implies$ ii). Supposons par l'absurde que la famille soit libre, mais ne soit pas géneratrice : alors par le Théorème 4, on pourrait compléter la famille pour former une base de V. Mais, alors cette famille aurait plus de n vecteurs, ce qui est contradictoire avec le Théorème 3. Montrons maintenant que ii) $\implies$ i). Supposons par l'absurde que la famille soit génératrice, mais pas libre. Alors d'après le Théorème 1 on peut trouver une sous-famille qui soit libre et génératrice. Mais ceci de nouveau est contradictoire avec le Théorème 3.

En pratique, il est souvent plus facile de prouver qu'une famille est libre que génératrice, et c'est donc souvent en montrant qu'une famille est libre et qu'elle la bonne taille qu'on vérifie qu'il s'agit bien d'une base.

Dans le même esprit que la Proposition 10, on démontre (exercice) :

Proposition 11. *Soit V un espace vectoriels, F et G deux sous-espaces vectoriels de V. Si $G \subset F$ et si F et de dimension finie, alors dim $G \leq$ dim F, et dim $F =$ dim G si et seulement si $F = G$.*

2.3 Rang d'une famille de vecteurs

Définition 7. *Soit V un espace vectoriel, et $(\vec{u}_1, \dots, \vec{u}_m)$ une famille de m vecteurs de V. On appelle* rang *de la famille $(\vec{u}_1, \dots, \vec{u}_m)$ et on note rg $(\vec{u}_1, \dots, \vec{u}_m)$ la dimension du sous-espace vectoriel* $\text{Vect}(\vec{u}_1, \dots, \vec{u}_m)$.

Les résultats du chapitre précédent permettent alors d'établir (exercice) les propriétés suivantes :
i) rg $(\vec{u}_1, \dots, \vec{u}_m) \leq m$.
ii) rg $(\vec{u}_1, \dots, \vec{u}_m) \leq$ dim V, si V est de dimension finie, avec égalité si et seulement la famille engendre V.
iii) rg $(\vec{u}_1, \dots, \vec{u}_m) = m$ si et seulement si la famille est libre.

Comment déterminer en pratique le rang d'une famille de vecteurs, lorsque ces derniers sont décomposés dans une base ?

Supposons donnée une base $(\vec{e}_1, \dots, \vec{e}_n)$ de l'espace vectoriel V supposé de dimension finie, et soit une famille $(\vec{u}_1, \dots, \vec{u}_m)$ de m vecteurs. Décomposons chacun de ces vecteurs dans la base :
$$\vec{u}_1 = u_{1,1}\vec{e}_1 + \cdots + u_{1,n}\vec{e}_n$$

$$\dots$$

$$\vec{u}_m = u_{m,1}\vec{e}_1 + \dots + u_{m,n}\vec{e}_n,$$

où les m$\times n$ coefficients $u_{i,j}$ sont des éléments de $\mathbb{K}$, que l'on peut avantageusement ranger dans une *matrice* à m lignes, et n colonnes A

$$A = \begin{pmatrix} u_{1,1} & \dots & u_{1,n} \\ \vdots & & \vdots \\ u_{m,1} & \dots & u_{m,n} \end{pmatrix}, \tag{12}$$

c'est à dire que les lignes de A sont constituée des coordonnées des vecteurs de la famille. On posera par la suite

$$\mathrm{rg} A = \mathrm{rg}(\vec{u}_1, \dots, \vec{u}_m).$$

Notons maintenant que si on fait subir à notre famille les transformations $T(1)$ (permutation de vecteurs), $T(2)$ (multiplication par un scalaire non nul), et $T(3)$ (ajout à l'un des vecteurs une combinaison linéaire des autres) décrites dans la Proposition 5, alors le rang reste inchangé. A chacune de ces opération correspond une opération similaire pour la matrice (permutation de lignes, multiplication d'une ligne par un scalaire, addition à une ligne d'une combinaison linéaire des autre lignes). L'idée centrale est d'utiliser ces opérations sur les lignes pour transformer la matrice initiale en une matrice dont le rang et simple à calculer : comme ce dernier est conservé au cours de ces transformations, il s'agit aussi du rang de la matrice initiale. Les matrices simples que nous voulons obtenir sont les matrices dites *échelonnées*, la méthode décrite étant appelée *méthode du pivot de Gauss.*

Définition 8. *Une matrice est dite échelonnée en lignes si le nombre de zéros précédant la première valeur non nulle d'une ligne augmente ligne par ligne jusqu'à ce qu'il ne reste plus que des zéros.*

Exemple de matrice échelonnée en ligne avec $a \neq 0, b \neq 0$

$$M = \begin{pmatrix} a & \star & \star & \star \\ 0 & 0 & b & \star \\ 0 & 0 & 0 & c \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{13}$$

On vérifie alors facilement que :

Lemme 2. *Le rang d'une matrice echelonnée en ligne correspond au nombre de lignes non identiquement nulles.*

Ainsi le rang de M est 3 si c est non nul, 2 sinon. Donnons maintenant un exemple de calcul de rang d'une famille de vecetur par la méthode du pivot de Gauss.

Exemple : Calculer le rang de la famille $(\vec{V}_1, \vec{V}_2, \vec{V}_3, \vec{V}_4)$ composée des vecteurs de $\mathbb{R}^5$ suivants :

$$\vec{V}_1 = (1, 0, 0, 2, -1),\ \vec{V}_2 = (0, 1, -2, 1, 0),\ \vec{V}_3 = (0, -1, 2, 1, -1),\ \vec{V}_4 = (0, 0, 0, 2, -1).$$

On commence par écrire la matrice A des lignes $\vec{V}_1, \vec{V}_2, \vec{V}_3, \vec{V}_4$

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

puis on applique l'algorithme du pivot de Gauss pour obtenir une matrice échelonnée. En ajoutant (L_2) à (L_3), puis en retranchant (L_3) à (L_4) on trouve

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

La dernière matrice a trois lignes non nulles, et est échelonnée. Son rang est donc trois. Le rang de la famille $(\vec{V}_1, \vec{V}_2, \vec{V}_3, \vec{V}_4)$ est donc 3.

3 Applications linéaires

Nous considérons dans cette partie deux espaces vectoriels V et W sur $\mathbb{K}$ et des applications de l'espace V vers l'espace W

Définition 9. *Soit $T : V \to W$ une application de V vers W. On dit de l'application T qu'elle est linéaire si et seulement si elle vérifie les propriétés suivantes :*
i) additivité *: $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v}), \forall u, v \in V$.*
ii) homogénéîté *: $T(\lambda.\vec{u}) = \lambda.T(\vec{u}), \forall \lambda \in \mathbb{K}, \forall \vec{u} \in V$.*

On utilisera indifféremment la notation $T(\vec{u})$ ou tout simplement $T\vec{u}$. L'ensemble des applications linéaires de V vers W est noté $\mathcal{L}(V, W)$. Lorsque $V = W$, on utilise la notation $\mathcal{L}(V)$ pour $\mathcal{L}(V, V)$.

Exemples. Comme exemple élémentaire, nous avons l'application nulle 0 définie par $0(u) = \vec{0}$, pour tout $u \in V$. Pour $V = W$, un autre exemple élémentaire est fourni par l'identité Id_V, définie par

$$\mathrm{Id}_V(\vec{v}) = \vec{v}.$$

Un exemple un peu moins élémentaire est fourni par les *projections*, qui sont des applications linéaires d'un espace dans lui-même. Supposons que nous puissions décomposer V en somme de deux sous-espaces supplémentaires F et G, c'est à dire que

$$V = F \oplus G.$$

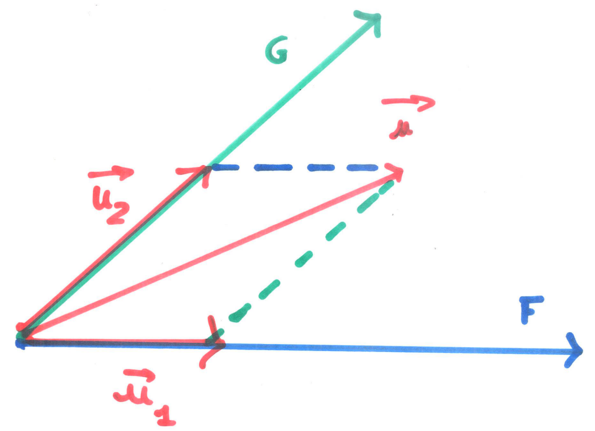
Tout élément v de V se décompose donc de manière unique sous la forme

$$v = \vec{u}_1 + u_2, \text{ où } \vec{u}_1 \in F \text{ et } \vec{u}_2 \in G.$$

L'application qui à v associe $\vec{u}_1$, l'unique élément de F qui réalise la décomposition précédente est appelée *projection sur F parallèlement à G* et est souvent notée π_F. On a donc

$$\pi_F(v) = \vec{u}_1.$$

On pourra vérifier en guise d'exercice que π_F est bien linéaire.



Sur $\mathcal{L}(V,W)$ on peut définir une addition "notée"+ qui a deux applications linéaires en associe une troisième : pour T_1 et T_2 donnés, on définit $T_1 + T_2$ par

$$(T_1 + T_2)(\vec{u}) = T_1(\vec{u}) + T_2(\vec{u}), \forall u \in V.$$

Notons que l'application nulle est l'élément neutre pour cette addition. On définit de même la multiplication par un scalaire comme

$$(\lambda.T)(\vec{u}) = \lambda.T(\vec{u}).$$

Muni de ces deux lois, l'ensemble $\mathcal{L}(V,W)$ devient lui même un espace vectoriel.

Notons enfin que la loi de composition habituelle $\circ$ préserve la linéarité. Rappelons que si T_1 est une application d'un ensemble V vers un ensemble W et que si T_2 est une application de W vers un ensemble Z, on définit l'application $T_2 \circ T_1$ de V vers Z par

$$T_2 \circ T_1(u) = T_2\left(T_1(u)\right), \forall u \in V.$$

On a alors

Proposition 12. *Soient V, W et Z trois espaces vectoriels, T_1 une application linéaire de V vers W, et T_2 une application linéaire de W vers Z. Alors $T_2 \circ T_1$ est une application linéaire de V vers Z.*

Pour $V = W = Z$, on vérifiera que

$$\mathrm{Id}_V \circ T + T \circ \mathrm{Id}\,_V.$$

En guise d'exercice on pourra calculer $\pi_F \circ \pi_F$ et $\pi_f \circ \pi_G$, où π_G est la projection sur F parallèlement à G.

3.1 Noyau et image

Définition 10. *Soit T une application linéaire de V vers W. Le noyau de T noté* $\operatorname{Ker} T$ *est le sous-ensemble de V constitué des éléments de V dont l'image par T est nulle, c'est à dire*

$$\operatorname{Ker} T = \{v \in V, \text{tels que } T\vec{v} = \vec{0}\}.$$

On vérifie immédiatement que

Proposition 13. $\operatorname{Ker} T$ *est un sous-espace vectoriel de V.*

Preuve. En effet par homogénéïté on a $T(\vec{0}) = \vec{0}$, et pour tous $\vec{u}$ et $\vec{v}$ dans $\operatorname{Ker} T$, tous λ, μ dans $\mathbb{K}$, on a

$$T(\lambda.\vec{u} + \mu.\vec{v}) = \lambda.T(\vec{u}) + \mu.T(\vec{v}) = \vec{0} + \vec{0} = \vec{0},$$

de sorte que $\lambda.\vec{u} + \mu.\vec{v} \in \operatorname{Ker} T$.

Proposition 14. *L'application T est injective si et seulement si* $\operatorname{Ker} T = \{0\}$.

Preuve. Rappelons que, de manière générale, qu'une application est injective si et seulement tout élément de W possède au plus un antécédent, c'est à dire que si $T\vec{u}_1 = T\vec{u}_2$, alors nécessairement $\vec{u}_1 = \vec{u}_2$. Or, si $T\vec{u}_1 = T\vec{u}_2$, la linéarité entraîne $T(\vec{u}_1 - \vec{u}_2) = \vec{0}$, c'est à dire $\vec{u}_1 - \vec{u}_2 \ \in \operatorname{Ker} T$. Si $\operatorname{Ker} T = \{0\}$, alors nécessairement $\vec{u}_1 = \vec{u}_2$, et la conclusion en découle.

Soit T une application de V vers W. On définit son image $\operatorname{Im} T$ comme l'ensemble des images des éléments de V par T, c'est à dire

$$\operatorname{Im} T = \{\vec{v} = T\vec{u},\ u \in V\}.$$

Rappelons qu'une application est dite surjective si et seulement si chaque élément de l'ensemble d'arrivée possède au moins un antécédent. Ceci peut aussi s'exprimer par le fait que $\operatorname{Im} T = W$ tout entier.

On a alors

Proposition 15. *Si T est une application linéaire de V vers W alors* $\operatorname{Im} T$ *est un sous-espace vectoriel de W. si de plus V et de dimension finie, et la famille $(\vec{e}_1, \ldots, \vec{e}_n)$ est une base de V alors la famille $(T(\vec{e}_1), T(\vec{e}_2), \ldots, T(\vec{e}_n))$ est une famille génératrice de* $\operatorname{Im} T$.

Preuve. On a vu que $T\vec{0} = \vec{0}$ de sorte que $\vec{0} \in \operatorname{Im} T$ Soit $\vec{v}_1$ et $\vec{v}_2$ deux éléments de V. Par définition de $\operatorname{Im} T$ il existe donc deux éléments $\vec{u}_1$ et $\vec{u}_2$ de V tels que $T\vec{u}_1 = v_1$ et $T\vec{u}_2 = v_2$. Pour λ et μ dans $\mathbb{K}$ on a donc $\lambda\vec{v}_1 + \mu\vec{v}_2 = T(\lambda\vec{u}_1 + \mu\vec{u}_2) \in \operatorname{Im} T$.

Dans le même esprit, il vient :

Proposition 16. *Soit T une application linéaire de V vers W telle que* $\operatorname{Ker} T = \{\vec{0}\}$ *.Alors l'image par T de toute famille libre est libre.*

Preuve Soit $(\vec{u}_1, \ldots, \vec{u}_m)$ une famille libre de V. Considérons son image $(T\vec{u}_1, \ldots, T\vec{u}_m)$ et une combinaison linéaire nulle de cette famille

$$\lambda_1.T(\vec{u}_1) + \ldots \lambda_m.T(\vec{u}_m) = \vec{0}.$$

Par linéarité on a donc

$$T\left(\lambda_1.\vec{u}_1 + \ldots \lambda_m.\vec{u}_m\right) = \vec{0}.$$

Comme on suppose de plus $\operatorname{Ker} T = \{\vec{0}\}$, il en résulte que $\lambda_1.\vec{u}_1 + \ldots \lambda_m.\vec{u}_m = \vec{0}$. La famille $(\vec{u}_1, \ldots, \vec{u}_m)$ étant libre, on en déduit que $\lambda_1 = \lambda_2 = \ldots = 0$, et la conclusion en découle.

En combinant les deux derniers résultats on obtient donc :

Proposition 17. *Si V est de dimension finie et si* $\operatorname{Ker} T = \{0\}$, *alors l'image par T de toute base de V est une base de* $\operatorname{Im} T$. *En particulier* $\dim V = \dim\left(\operatorname{Im} T\right)$.

Le résultat essentiel de cette partie étend ce résultat (Théorème du rang)

Théorème 5. *Soit V un espace de dimension finie, et T une application linéaire de V vers W. Alors* $\operatorname{Im} T$ *est un sous-espace vectoriel de dimension finie de W et*

$$\dim V = \dim\left(\operatorname{Ker} T\right) + \dim\left(\operatorname{Im} T\right).$$

Preuve. Considérons un sous-espace vectoriel F supplémentaire de $\mathrm{Ker}\, T$ de sorte que $V = F \oplus \mathrm{Ker}\, T$. Soit $\tilde{T}$ la restriction de T à F. Comme $F \cap \mathrm{Ker}\, T = \{\vec{0}\}$, on vérifie que $\mathrm{Ker}\, \tilde{T} = \{0\}$, de sorte que par la proposition 17, on a $\dim \left(\mathrm{Im}\, \tilde{T}\right) = \dim F$. Par ailleurs, tout $\vec{u}$ dans V s'écrit

$$u = \vec{u}_1 + \vec{u}_2, \vec{u}_1 \in F, \vec{u}_2 \in \mathrm{Ker}\, T,$$

et ainsi on a $Tu = Tu_1 + Tu_2 = Tu_1 + \vec{0} = Tu_1 = \tilde{T}u_1 \in \mathrm{Im}\, \tilde{T}$, d'ou l'on déduit $\mathrm{Im}\, T = \mathrm{Im}\, \tilde{T}$. Il résulte de ce qui précède

$$\dim\, (\mathrm{Im}\, T) = \dim F.$$

Comme F et $\mathrm{Ker}\, T$ sont supplémentaires

$$\dim V = \dim\, (\mathrm{Ker}\, T) + \dim F,$$

la conclusion en découle en combinant les deux dernières relations.

Tirons immédiatement quelques conséquences de ce résultat.

Corollaire 1. *Si $\dim W < \dim V$ alors il n'y a pas d'application linéaire injective de V vers W. Si $\dim W > \dim V$ alors il n'y a pas d'application linéaire surjective de V vers W.*

Rappelons enfin qu'une application T entre deux ensemble V et W est dite bijective si elle est injective et surjective, c'est à dire si et seulement si tout élément de W possède un antécédent et un seul par T. On a donc, d'après ce qui précède :

Corollaire 2. *Soit V, W deux espaces vectoriels. On suppose que V et de dimension finie et qu'il existe et T une application linéaire bijective de V vers W. Alors W est de dimension finie et*

$$\dim V = \dim W. \tag{14}$$

Si (14) est vérifiée, alors toute application linéaire S de V vers W est bijective si et seulement si $\mathrm{Ker}\, S = \{\vec{0}\}$.

3.2 Application inverse

Soit T une application bijective entre deux ensembles V et W, alors on note alors T^{-1} l'application qui à un élément w de W associe son antécédent. En d'autres termes pour une application *bijective* entre V et W, on a

$$u = T^{-1}w \text{ si et seulement si } T\, u = w.$$

T^{-1} est alors appelée application *réciproque* ou *inverse* de T. On a :

Proposition 18. *Si T est une application linéaire de V vers W bijective, alors T^{-1} est également linéaire (et bijective).*

Nous laissons la démonstration de ce résultat en exercice. On vérifiera par ailleurs que

$$T^{-1} \circ T = \mathrm{Id}_V \text{ et } T \circ T{-}1 = \mathrm{Id}_W.$$

4 Expression matricielle des applications linéaires

On considère dans tout ce chapitre deux espaces vectoriels V et W de dimension finie,

$$\dim V = n, \ \dim W = m,$$

une base $(\vec{e}_1, \ldots, \vec{e}_n)$ de V, ainsi qu'une base $(\vec{f}_1, \ldots, \vec{f}_m)$ de W. Nous allons voir comment exprimer une application T de V vers W à l'aide d'une matrice $m \times n$, c'est à dire un tableau à m lignes et à n colonnes.

4.1 Construction de la matrice

Décomposons tout d'abord l'image de chaque vecteur de la base $(\vec{e}_1, \ldots, \vec{e}_n)$ dans la base $(\vec{f}_1, \ldots, \vec{f}_m)$:

$$\begin{cases} T(\vec{e}_1) = a_{1,1}\vec{f}_1 + \ldots + a_{m,1}\vec{f}_m \\ T(\vec{e}_2) = a_{1,2}\vec{f}_1 + \ldots + a_{m,2}\vec{f}_m \\ \qquad \vdots \\ T(e_n) = a_{1,n}\vec{f}_1 + \ldots + a_{m,n}\vec{f}_m. \end{cases}$$

Nous pouvons ranger les nombres $a_{j,i}$ dans une matrice à m lignes et n colonnes

$$A = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix}. \tag{15}$$

Pour construire cette matrice, il faut essentiellement noter que les composantes des n colonnes correspondent aux coordonnées des images $T(\vec{e}_1), \ldots, T(\vec{e}_n)$ des n-vecteurs de la base de V, dans la base donnée $(\vec{f}_1, \ldots, \vec{f}_m)$ de W. Si $a_{i,j}$ est un indice de cette matrice, le premier indice i représente le numéro de la ligne, le second indice$j$ celui de la colonne.

Exemple. Considérons l'application T de $\mathbb{R}^2$ vers $\mathbb{R}^3$ définie par :

$$T(x, y) = (x + 3y, 2x + 5y, 7x + 9y).$$

Prenons pour bases de $\mathbb{R}^2$ et $\mathbb{R}^3$ les bases canoniques $\vec{e}_1 = (1, 0)$, $\vec{e}_2 = (0, 1)$ et $\vec{f}_1 = (1, 0, 0)$, $\vec{f}_2 = (0, 1, 0)$, $\vec{f}_3 = (0, 0, 1)$, de sorte que

$$\begin{cases} T\vec{e}_1 = \ \vec{f}_1 + 2\vec{f}_2 + 7\vec{f}_3 \\ T\vec{e}_2 = 3\vec{f}_1 + 5\vec{f}_2 + 7\vec{f}_3. \end{cases}$$

la matrice correspondante s'écrit donc

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix}.$$

Revenons au cas général. A l'aide de la matrice A, on peut calculer aisément les coordonnées dans la base $(\vec{f}_1, \ldots, \vec{f}_m)$. de l'image de tout vecteur $\vec{v}$ de V, pourvu que ce dernier soit décomposé sur la base $(\vec{e}_1, \ldots, \vec{e}_n)$

$$v = v_1\vec{e}_1 + v_2\vec{e}_2 + \ldots v_n\vec{e}_n.$$

On associe alors à cette décomposition une matrice colonne (c'est à dire une $n \times 1$ matrice V définie par

$$V = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$$

L'image $Tv = \begin{pmatrix} w_1 \\ \ldots \\ w_m \end{pmatrix}$ est alors représentée par la matrice colonne $m \times 1$

$$\begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_{1,1}v_1 + \ldots a_{1,n}v_n \\ \vdots \\ a_{m,1}v_1 + \ldots a_{m,n}v_n \end{pmatrix}.$$

En écriture plus synthétique, on a donc, pour $i = 1, \ldots, m$,

$$w_i = \sum_{j=1}^{n} a_{i,j}v_j.$$

4.2 Multiplication des matrices

Considérons dans cette partie un troisième espace vectoriel Z, de dimension finie, muni d'une base $\vec{g}_1, \ldots, \vec{g}_\ell$, une application linéaire T de V vers W et une application linéaire S de W vers Z, dont les matrices correspondantes (respectivement $m \times n$ et $\ell \times m$) sont

$$A = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix} \text{ et } B = \begin{pmatrix} b_{1,1} & \ldots & b_{1,m} \\ \vdots & & \vdots \\ b_{\ell,1} & \ldots & b_{\ell,m} \end{pmatrix}.$$

Considérons maintenant la composée $S \circ T : V \to Z$. Soit C la matrice $\ell \times n$ correspondante. Par définition, nous définirons le produit des matrices B et A comme

$$B\,A \equiv C = \begin{pmatrix} c_{1,1} & \ldots & b_{1,n} \\ \vdots & & \vdots \\ c_{\ell,1} & \ldots & c_{\ell,n} \end{pmatrix}.$$

Comment calculer les éléments $c_{i,j}$ de la matrice C ?

Pour $i = 1, \ldots, \ell$, et $j = 1, \ldots, n$ on vérifie que

$$c_{i,j} = \sum_{k=1}^{m} b_{i,k}a_{k,j},$$

C'est à dire que l'on prend la i-ème ligne de la matrice B, qui est de longueur m, et l'on multiplie terme à terme ses éléments par ceux de la j-ème colonne de la matrice A, qui est également de longueur m.

Exemple

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 6 & 5 & 4 & 3 \\ 2 & 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 10 & 7 & 4 & 1 \\ 26 & 19 & 12 & 5 \\ 42 & 31 & 20 & 9 \end{pmatrix}.$$

4.3 Changement de bases

Comme nous venons de le voir, l'expression d'un application linéaire sous forme matricielle est relative aux choix des bases, à la fois pour l'espace de départ que pour l'espace d'arrivéé. Nous allons voir dans ce chapitre comment cette expression se modifie lorsque que l'on effectue des changements de bases, ainsi que de manière plus générale, les coordonnées d'un vecteur.

Appelons $\mathfrak{e}$ la base initiale $(\vec{e}_1, \ldots, \vec{e}_n)$ de V, et considérons une nouvelle base $\mathfrak{e}'$ formée des vecteurs $(\vec{e}'_1, \ldots, \vec{e}'_n)$. Supposons que nous connaissions les coordonnées des vecteurs de la nouvelle base dans la base initiale, c'est à dire suposons que nous connaissions des décomposition suivantes :

$$\begin{cases} \vec{e}'_1 = p_{1,1}\vec{e}_1 + p_{2,1}\vec{e}_2 + \ldots p_{n,1}\vec{e}_n \\ \vec{e}'_2 = p_{1,2}\vec{e}_1 + p_{2,2}\vec{e}_2 + \ldots p_{n,2}\vec{e}_n \\ \vdots \\ \vec{e}\prime_n = p_{1,n}\vec{e}_1 + p_{2,n}\vec{e}_2 + \ldots p_{n,n}\vec{e}_n \end{cases}$$

Formons la matrice carréé $P(\mathfrak{e} \to \mathfrak{e}')$, appelée *matrice de passage* de la base $\mathfrak{e}$ vers la base $\mathfrak{e}'$

$$P(\mathfrak{e} \to \mathfrak{e}') = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & & & \vdots \\ p_{n,1} & p_{n,2} & \cdots & p_{n,n} \end{pmatrix}$$

Notons que la i-ème colonne de cette matrice est composée des coordonnées du vecteur $\vec{e}'_1$ dans la base initiale $\mathfrak{e}$.

4.3.1 Changements des composantes d'un vecteur

Considérons un vecteur $\vec{v}$ de V, qui se décompose comme

$$\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + \ldots v_n\vec{e}_n$$

dans la base initiale , et comme

$$\vec{v} = v'_1\vec{e}'_1 + v_2\vec{e}'_2 + \ldots v'_n\vec{e}'_n.$$

Un bref calcul montre alors que

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & & & \vdots \\ p_{n,1} & p_{n,2} & \cdots & p_{n,n} \end{pmatrix} \begin{pmatrix} v'_1 \\ v'_2 \\ \vdots \\ v'_n \end{pmatrix},$$

que l'on peut écrire, sous forme plus condensée, en posant

$$V = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \text{ et } V' = \begin{pmatrix} v'_1 \\ v'_2 \\ \vdots \\ v'_n \end{pmatrix}$$

sous la forme

$$V = P(\mathfrak{e} \to \mathfrak{e}')V'. \tag{16}$$

La matrice de passage $P(\mathfrak{e} \to \mathfrak{e}')$ et la relation (16) permettent donc de calculer les coordonnées initiales lorsque l'on connaît les nouvelles coordonnées : en général, c'est plutôt l'inverse que l'on désire faire. Nous déduisons de (16) que

$$V' = P^{-1}(\mathfrak{e} \to \mathfrak{e}')V, \tag{17}$$

et ainsi

$$P(\mathfrak{e}' \to \mathfrak{e}) = P^{-1}(\mathfrak{e} \to \mathfrak{e}').$$

On peut en effet vérifier que les matrices de passage d'une base à l'autre sont toujours inversible.

Ainsi si on veut calculer les nouvelles coordonnées en fonctions des coordonnées initiales, nous sommes amenés à calculer l'inverse de la matrice de passage $P^{-1}(\mathfrak{e} \to \mathfrak{e}')$.

4.3.2 Changements des composantes d'une matrice

Considérons deux espaces vectoriels V et W de dimensions finies, respectivement n et m, ainsi qu'une application linéaire T de V vers W, dont l'expression matricielle dans des bases $\mathfrak{e} = (\vec{e}_1, \ldots, \vec{e}_n)$ et $\mathfrak{f} = (\vec{f}_1, \ldots, \vec{f}_m)$ de V et W est donnée par A, une matrice $m \times n$ représentée par (19). Considérons maintenant deux nouvelles bases de $\mathfrak{e}' = (\vec{e}'_1, \ldots, \vec{e}'_n)$ et $\mathfrak{f}' = (\vec{f}'_1, \ldots, \vec{f}'_m)$ respectivement. On déduit de ce qui précède que l'expression A' de l'application linéaire T dans les nouvelles bases $\mathfrak{e}'$ et $\mathfrak{f}'$ est donnée par la formule :

$$A' = P^{-1}(\mathfrak{f} \to \mathfrak{f}')A\,P(\mathfrak{e} \to \mathfrak{e}'). \tag{18}$$

Un cas particulier important est celui où $V = W$, c'est à dire que l'on considère des applications linéaires d'un espaces dans lui-même. Dans ce cas, on prend, presque toujours $\mathfrak{e} = \mathfrak{f}$, et la matrice A, qui est carrée, fourni l'expression de l'application linéaire dans la base $\mathfrak{e}$. Dans une nouvelle base $\mathfrak{e}'$ l'expression de la matrice est alors donnée par la formule

$$A' = P^{-1}A\,P,$$

où $P = P(\mathfrak{e} \to \mathfrak{e}')$.

4.3.3 Calcul de l'inverse d'une matrice

Le calcul de l'inverse d'une matrice est une opération délicate : la méthode la plus efficace est souvent celle du pivot de Gauss.

5 Le déterminant d'une matrice carrée

Dans cette partie, nous limitons notre attention à l'étude des matrices carrées de taille n, que nous noterons $\mathcal{M}_n(\mathbb{K})$ [Notons au passage qu'il s'agit d'un espace vectoriel sur n^2]. La notion de déterminant, que nous allons bientôt rappeler, est un outil puissant pour déterminer si son rang est maximal.

5.1 Définition du déterminant

Le resultat suivant permet de définir le déterminant

Proposition 19. *Il existe une seule application de* $\mathcal{M}_n(\mathbb{K})$ *vers* $\mathbb{K}$*, que nous noterons* det *et appellerons* déterminant *qui possède les propriétés suivantes :*
1) Linearité par rapport à chaque ligne *:*

$$\det\begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i + \vec{a}'_i \\ \vdots \\ \vec{a}_n \end{pmatrix} = \det\begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} + \det\begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}'_i \\ \vdots \\ \vec{a}_n \end{pmatrix} \det\begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} \leftarrow \quad \text{ligne } i$$

$$\det\begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{\lambda} a_i \\ \vdots \\ \vec{a}_n \end{pmatrix} = \lambda\det\begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} \leftarrow \quad \text{ligne } i$$

2) caractère alterné *: Si la matrice A a deux lignes identiques, alors* $\det A = 0$.
3) Normalisation *:* $\det I_n = 1$.

Remarque 1. *: Dans le résultat précédent, la notation $\vec{a}$ désigne un vecteur ligne.*

Remarque 2. *Toute application de $\mathcal{M}_n(\mathbb{K})$ dans $\mathbb{K}$ qui vérifie les deux propriétés est anti-symétrique, c'est à dire si on permute deux lignes de la matrice, alors le signe change, en particulier*

$$\det\begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ \leftarrow \text{ligne } i \\ \\ \rightarrow \text{ligne } j \\ \\ \\ \end{matrix} = -\det\begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_j \\ \vdots \\ \vec{a}_i \\ \vdots \\ \vec{a}_n \end{pmatrix} \begin{matrix} \\ \\ \leftarrow \text{ligne } i \\ \\ \rightarrow \text{ligne } j \\ \\ \\ \end{matrix}$$

Pour le voir, on remarque que, par le caractère alterné, on a

$$\det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_j + \vec{a}_i \\ \vdots \\ \vec{a}_i + \vec{a}_j \\ \vdots \\ \vec{a}_n \end{pmatrix} = 0.$$

On obtient le résultat en développant grâce à la linéarité par rapport aux lignes.

Idée de la preuve de la proposition 19. Soit $(\vec{e}_1, \dots, \vec{e}_n)$ la base canonique de $\mathbb{K}^n$. Tout vecteur $\vec{a}_i$ de $\mathbb{K}^n$ se décompose comme

$$\vec{a}_i = \sum_{j=1}^{n} a_{i,j} \vec{e}_j$$

de sorte qu'en utilisant la linéarité par rapport aux lignes on en déduit que

$$\det \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_n \end{pmatrix} = \sum_{j_1=1}^{n} \dots \sum_{j_n=1}^{n} a_{1,j_1} \dots a_{1,j_n} \begin{pmatrix} \vec{e}_{j_1} \\ \vdots \\ \vec{e}_{j_n}. \end{pmatrix}.$$

L'application est donc entièrement déterminé par sa valeur sur les vecteurs de la base canonique. Grâce au caractère alterné, on sait que l'image d'une famille contenant deux fois le même vecteur de la base canonique par det est nulle. Par antisymmétrie, on constate que toute les valeurs restantes peuvent se déduire de $\begin{pmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_n \end{pmatrix}$, qui est égal à 1 par normalisation.

Notation. Si a est donnée par (19), alors on note le déterminant comme suit

$$\det A = \begin{vmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{vmatrix}. \tag{19}$$

Remarque 3. *Il résulte des propriétés de linéarité et de normalisation que si $n = 1$ et si $A = (a)$ est une matrice 1×1, alors $\det A = 1$. Si $n = 2$, on vérifie en développant les lignes comme ci-dessus que*

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$$

formule bien connue.

5.2 Quelques propriétés

Il résulte immédiatement de la linéarité que

Proposition 20. *i) Pour tout $\lambda \in \mathbb{K}$, et tout matrice A de $\mathcal{M}_n(\mathbb{K})$, on a*

$$\det(\lambda A) = \lambda^n \det A.$$

ii)Si A a une ligne nulle, alors $\det A = 0$.
iii) Si l'on ajoute à une ligne une combinaison linéaire des autres lignes, alors le déterminant ne change pas.

Nous laissons la preuve en exercice. Notons cependant dès à présent que que ces propriétés vont nous permettre de calculer des déterminants à l'aide de la méthode du Pivot de Gauss. Une autre conséquence importante est la suivante :

Corollaire 3. *Soit A une matrice carrée. Les vecteurs ligne forment une famille liée, si et seulement si* $\det A = 0$.

En d'autres termes, si $(\vec{a}_1, \ldots, \vec{a}_n)$ est une famille de n vecteurs de $\mathbb{K}^n$, alors la famille est liée si et seulement si

$$\det \begin{pmatrix} \vec{a}_1 \\ \ldots \\ \vec{a}_n \end{pmatrix} = 0.$$

Nous laissons de nouveau la preuve en exercice.

Proposition 21. *Si A est une matrice diagonale, alors son déterminant est égale au produit des termes diagonaux. de même, si A est triangulaire supérieure, alors* $\det A$ *est égal au produite des termes diagonaux.*

Idée de la preuve Considérons d'abord le cas ou A est diagonale. Alors on a

$$\det A = \begin{vmatrix} a_{1,1} & 0 & \vdots & 0 \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \vdots & \vdots & 0 \\ 0 & \ldots & 0 & a_{n,n} \end{vmatrix} = a_{1,1} \begin{vmatrix} 1 & 0 & \vdots & 0 \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \vdots & \vdots & 0 \\ 0 & \ldots & 0 & a_{n,n} \end{vmatrix}$$

$$= a_{1,1}a_{2,2} \begin{vmatrix} 1 & 0 & \vdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & 0 \\ 0 & \ldots & 0 & a_{n,n} \end{vmatrix} = \ldots$$

$$= \prod_{i=1}^{n} a_{i,i} \begin{vmatrix} 1 & 0 & \vdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & 0 \\ 0 & \ldots & 0 & 1 \end{vmatrix} = \prod_{i=1}^{n} a_{i,i}.$$

Ceci prouve l'assertion dans le cas diagonal. Dans le cas trianngulaire, lorsque aucun des termes sur la diagonale est nulle, on procède de même, pour obtenir

$$\det A = \prod_{i=1}^{n} a_{i,i} \begin{vmatrix} 1 & \star & \ldots & \star \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \ldots & 0 & 1 \end{vmatrix}.$$

Ensuite, on élimine les termes supérieurs en s'inspirant de la méthode du Pivot de Gauss. On adapte cette méthode dans le cas ou un des termes diagonaux est nul.

Rappelons que la matrice transposée d'une matrice A est la matrice déduite de A dont les lignes sont les colonnes de A et réciproquement. On note ${}^tA$ la matrice transposée de A. Ainsi, si

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix} \text{ alors } {}^tA = \begin{pmatrix} a_{1,1} & a_{2,1} & \dots & a_{n,1} \\ a_{1,2} & a_{2,2} & \dots & a_{n,2} \\ \vdots & & & \vdots \\ a_{1,n} & a_{2,n} & \dots & a_{n,n} \end{pmatrix}.$$

On alors le résultat suivant que nous admettrons :

Proposition 22. *On a pour tout matrice* $A \in \mathcal{M}_n(\mathbb{K})$

$$\det {}^tA = \det A.$$

Ce résultat à des conséquences très importantes : il nous permet d'étendre les résultas obtenus pour les vecteurs lignes aux vecteurs colonne. en particulier

Corollaire 4. *Soit A une matrice carrée. Les vecteurs colonne forment une famille liée, si et seulement si* $\det A = 0$.

Corollaire 5. *Soit T une application linéaire d'un espace vectoriel V dans lui-même, et soit A son expression matricielle dans une base donnée. Alors* $\operatorname{Ker} T \neq \{0\}$ *si et seulement si* $\det A = 0$.

Au vu des résultat du Chapitre précédent, il en résulte également que T est un isomorphisme de V dans lui-mêmes si et seulement si $\det A \neq 0$.

Donnons une dernière propriété interésssante, que nous admettrons sans démonstration.

Proposition 23. *Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$. On a*

$$\det AB = \det A . \det B.$$

Il en résulte en particulier que, si A est inversible, alors

$$1 = \det (AA^{-1}) = \det A \det (A^{-1}),$$

et donc

$$\det A^{-1} = (\det A)^{-1}.$$

Voyons une autre conséquence :

Corollaire 6. *Soit T une application linéaire d'un espace vectoriel V dans lui-même.Soit le déterminant de son expression matricielle dans une base est indépendant de la base considérée.*

Nous pouvons alors parler du déterminant d'une application linéaire, en posant

$$\det T = \det A,$$

où A représente l'expression matricielle de T dans une base quelconque.

Preuve du Corollaire 6. Soit $\mathfrak{e}$ et $\mathfrak{e}'$ deux bases de base de V, A et A' les expressions matricielles de T correspondantes. On a $A' = P^{-1}A\,P$, où P désigne la matrice de passage de$\mathfrak{e}$ vers $\mathfrak{e}'$. En prenant le déterminant, et en invoquant ce qui précède, on obtient

$$\det A' = \det(P^{-1}\,A\,P) = \det(P^{-1})\det A\det P = (\det P)^{-1}\det A\det P = \det A.$$

Les dernières partie de ce chapitre sont consacrées aux méthodes de calcul de déterminants.

5.3 Calcul du déterminant par la méthode du pivot de Gauss

Le principe est d'utiliser les transformations élémentaires sur les lignes (permutations, ajout de combinaisons linéaires) pour se ramener au calcul sur les matrices triangulaires, que l'on sait calculer explicitement.

Considérons à cet effet une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$. L'algorithme du pivot de Gauss décrit au chapitre précédent permet d'obtenir une matrice A' échelonnée après une succession de transformations élémentaires. Comme la matrice de départ est carrée, la matrice A' est triangulaire supérieure. Les transformations élémentaires effectuées durant le pivot sont de type (T1) (permutation de deux lignes) ce qui revient à changer le déterminant en son opposé, ou de type (T 3) (ajout d'une combinaison linéaire de lignes), ce qui ne change pas le déterminant. On a donc finalement

$$\det A = (-1)^{\ell}\det A' = (-1)\ell\Pi a'_{i,i}$$

où ℓ désigne le nombre de permutations de lignes effectuées au cours du pivot.

Exemple

$$\begin{vmatrix} 1 & 0 & 4 \\ 2 & 3 & 1 \\ -1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 4 \\ 2 & 3 & 1 \\ 0 & 0 & 5 \end{vmatrix} \quad \text{ajout de } L_1 \text{ à } (L_3))$$

$$= \begin{vmatrix} 1 & 0 & 4 \\ 0 & 3 & -7 \\ 0 & 0 & 5 \end{vmatrix} \quad \text{(on retranche } 2L_1 \text{ à } (L_2))$$

Cette dernière matrice est triangulaire supérieure et le produit de ses termes diagonaux vaut 15 . On a donc $\det A = 15$.

Remarque 4. *On prendra garde au fait que multiplier une ligne (ou une colonne) par un scalaire change la valeur déterminant, quyi est multiplier par le nombre correspondant.*

Remarque 5. *Comme nous l'avons déjà mentionné, on peut faire les opérations indifféremment sur les lignes ou sur les colonnes. Prenons garde néanmoins que ceci n'est vrai* que *pour le calcul du déterminant. Faire un pivot de Gauss sur les colonnes pour résoudre un sytsème linéaire donne en général un résultat erroné !*

5.4 Développement par lignes ou colonnes

Cette méthode consiste à décomposer des déterminants d'ordre n en somme de déterminants d'ordre $n-1$.

Définition 11. *Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$ et soit $1 \leq i, j \leq n$. On appelle mineur de A d'indice (i,j) la matrice $A'_{i,j} \in \mathcal{M}_{n-1}(\mathbb{K})$ obtenue en rayant la i-ième ligne et la j-ième colonne de A.*

On appelle cofacteur d'indice (i, j) le scalaire $\alpha_{i,j} = (-1)^{i+j} \det A'_{i,j}$.

Exemple. soit $A = \begin{pmatrix} 2 & 3 & 1 \\ 2 & -1 & 0 \\ 1 & 2 & 0 \end{pmatrix}$ Alors $A'_{1,2} = \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix}$ et $A'_{2,2} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$.

On a alors le résultat suivant, que nous admettrons

Théorème 6. *Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$ et soit $1 \leq i, j \leq n$. Alors on a le développement suivant du déterminant de A, par rapport à la i-ème ligne*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det A'_{i,j} = \sum_{j=1}^{n} \alpha_{i,j}.$$

Remarque 6. *Comme le déterminant est conservé par transposition, nous pouvons alors également effctuer un développement suivant les colonnes. Ainsi, on peut développer par rapport à la j-ème colonne comme suit*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det A'_{i,j} = \sum_{i=1}^{n} \alpha_{i,j}.$$

En guise d'exemple, voyons comment ces idées conduisent à une formule pour le déterminat d'ordre 3. En développant par rapport à la première ligne, on obtient

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = a_{1,1} \begin{vmatrix} a_{2.2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} - a_{1,2} \begin{vmatrix} a_{2.1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} + a_{1,3} \begin{vmatrix} a_{2.1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix}.$$

On peut bien entendu ensuite développer les déterminants d'ordre 2 pour terminer le calcul.

6 Vecteurs propres et valeurs propres

Dans ce chapitre, nous allons étudier des applications linéaires d'un espace vectoriel V sur $\mathbb{K}$ de dimension finie n dans lui-même. Les vecteurs propres d'un application linéaire de V dans lui-même correspondent à des éléments dont la valeur est particulièrement simple, puisqu'elle est colinéaire au vecteur lui-même. Connaître l'ensemble des vecteurs propres est alors un étape particulière cruciale pour décrire le comportement de T.

6.1 Définitions

Définition 12. *Soit T une application linéaire de V dans lui-même. Soit $\lambda \in K$.On dit que λ est valeur propre de T si et seulement si il existe un vecteur* non nul *$\vec{u}$ de V tel que*

$$T\vec{u} = \lambda\vec{u}.$$

On dit alors que le vecteur $\vec{u}$ est vecteur propre de T associé à la valeur propre λ.

Notons immédiatement que l'ensemble F_λ des vecteurs propres associés à une valeur propre λ donnée, auxquels on adjoint le vecteur nul forme un sous-espace vectoriel de V. En effet, on a

$$F_\lambda = \{\vec{u} \in V, \text{tels que } \; T\vec{u} = \lambda\vec{u}\},$$

La relation $Tu = \lambda u$ peut aussi s'écrire

$$(T - \lambda \mathrm{Id}\,)\vec{u} = 0,$$

de sorte que

$$F_\lambda = \mathrm{Ker}\,(T - \lambda Id) = 0.$$

On a donc

Lemme 3. *Le nombre $\lambda \in \mathbb{K}$ est une valeur propre de T si et seulement si*

$$\mathrm{Ker}\,(T - \lambda Id) \neq \{0\}.$$

Remarque 7. *Si λ est une valeur propre de T, le sous-espace vectoriel F_λ s'appelle alors sous-espace propre de T associé à la valeur propre λ.*

Soit une base $(\vec{e}_1, \ldots, \vec{e}_n)$ de V et soit $A \in \mathcal{M}_n(\mathbb{K})$ l'expression matricielle de T. Il résulte alors du Lemme 3 ainsi que des résultats du chapitre précédent que l'on a

Corollaire 7. *Le nombre $\lambda \in \mathbb{K}$ est une valeur propre de T si et seulement si*

$$\det\,(T - \mathrm{Id}_V) = \det\,(A - \lambda \mathrm{I}_n) = 0,$$

où I_n *désigne la matrice unité d'ordre n.*

Posons pour $\lambda \in \mathbb{K}$

$$P_A(\lambda) = \det\,(A - \lambda \mathrm{I}_n).$$

En développant le déterminant, on s'aperçoit que le nombre de droite est un polynôme de degré n exactement : on l'appelle *polynôme caractéristique de* A. Il en résulte donc :

Le nombre $\lambda \in \mathbb{K}$ est une valeur propre de T si et seulement si λ est une racine du polynôme caractéristique P_A.

Remarque 8. *Au vu de l'invariance par changement de base, le polynôme caractéristique est indépendant de la base choisie, et donc ne dépend que de T. C'est pourquoi on pourra aussi le qualifier de polynôme caractéristique de T, et le noter, indifféremment selon le contexte P_A ou P_T.*

La recherche de valeurs propres conduit donc à la recherche des racines du polynôme caractéristique. Donnons un exemple :

Exemples. Considérons la matrice

$$A = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix}, \tag{20}$$

de sorte que

$$P_A(\lambda) = \det(A - \lambda \mathrm{Id}) = \begin{vmatrix} 1-\lambda & -3 \\ -2 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda) - 6 = \lambda^2 - 3\lambda - 4.$$

Un rapide calcul montre que les racines du polynôme caractéristique sont $\lambda_1 = -1$ et $\lambda_2 = 4$. Les vecteurs propres $\vec{u} = (u_1, u_2)$ associés à λ_1 vérifient

$$\begin{pmatrix} 2 & -3 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

Soit $2u_1 - 3u_2 = 0$. L'espace F_1 est donc la droite engendrée par $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$. De même, les vecteurs propres $\vec{u} = (u_1, u_2)$ associés à λ_2 vérifient

$$\begin{pmatrix} -3 & -3 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

soit $-3u_1 - 3u_2 = 0$. L'espace F_2 est donc la droite engendrée par $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

6.2 Propriétés des vecteurs propres

Théorème 7. *Soit T une application linéaire de V dans lui-même. Soient $\lambda_1, \ldots, \lambda_m$ m- valeurs propres* distinctes *de T, et soient $\vec{u}_1, \ldots, \vec{u}_m$ des vecteurs propres non nuls associés. Alors la famille $(\vec{u}_1, \ldots, \vec{u}_m)$ est libre.*

Preuve. On raisonne par l'absurde et on suppose que la famille $(\vec{u}_1, \ldots, \vec{u}_m)$ est liée. Soit alors k le plus petit entier tel que

$$\vec{u}_k \in \mathrm{Vect}(\vec{u}_1, \ldots, \vec{u}_{k-1}).$$

Il existe donc des coefficients $a_1, \ldots a_{k-1}$ tels que

$$u_k = a_1 \vec{u}_1 + \ldots a_{k-1} \vec{u}_{k-1}. \tag{21}$$

En prenant l'image par T, on obtient

$$\lambda_k \vec{u}_k = a_1 \lambda_1 \vec{u}_1 + \ldots \lambda_{k-1} a_{k-1} \vec{u}_{k-1}. \tag{22}$$

On peut éliminer $\vec{u}_k$ en multipliant la relation (21) par λ_k et en soustrayant (22)

$$\vec{0} = a_1(\lambda_1 - \lambda_k)\vec{u}_1 + \ldots a_{k-1}(\lambda_{k-1} - \lambda_1)\vec{u}_{k-1}.$$

Par notre choix de k, la famille $(\vec{u}_1, \ldots, \vec{u}_{k-1})$ est libre, de sorte que l'on aboutit à une contradiction.

Il en résulte immédiatement

Corollaire 8. *Soit T une application linéaires de V dans lui-même. Alors T a au plus n valeurs propres distinctes.*

Si T possède n valeurs propres distinctes $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ *, alors il existe une base $(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$ de V formée de vecteurs propres de T associés aux valeurs propres $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectivement. Dans cette base, l'expression matricielle A de T est diagonale et $A = \mathrm{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$.*

Rappelons qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est diagonale si et seulement si les seuls éléments non nuls de cette matrice se trouvent la diagonale. Elle a donc la forme

$$A = \begin{pmatrix} a_{1,1} & 0 & \ldots & 0 \\ 0 & a_{2,2} & 0 & \ldots \\ \vdots & \ldots & \ddots & 0 \\ 0 & 0 & \ldots & a_{n,n} \end{pmatrix}$$

et non notera $A = \mathrm{diag}\,(a_{1,1}, a_{2,2}, \ldots, a_{n,n})$.

On remarquera par ailleurs que si A est une matrice triangulaire supérieure

$$A = \begin{pmatrix} a_{1,1} & \star & \ldots & \star \\ 0 & a_{2,2} & \star & \ldots \\ \vdots & \ldots & \ddots & \star \\ 0 & 0 & \ldots & a_{n,n} \end{pmatrix}$$

alors ses valeurs propres sont données, repétées eventuellement avec multiplicité, par les nombres $a_{1,1}, a_{2,2}, \ldots, a_{n,n}$.

6.3 Le cas $\mathbb{K} = \mathbb{C}$

Le cas des nombres complexes présente une spécificité remarquable : tous les polynômes de degré non nul en effet des racines. En effet, un théorème fondamental de l'algèbre affirme :

Théorème 8. *Soit P un polynôme à coefficients complexes, de degré $n \geq 1$.Alors*

- *P possède au moins une racine.*
- *De plus, on peut décomposer P sous la forme du produit d'éléments simples suivant*

$$P(z) = c \prod_{i=1}^{n} (z - a_i) \tag{23}$$

où $c \in \mathbb{C}$ est non nul, et où les nombres $a_1, \ldots, a_n$ sont n nombres complexes racines du polynôme.

Remarque 9. *Le résultat est faux si on se restreint aux polynômes à coefficients réels, et si on cherche des racines réelles. Pour s'en convaincre, il suffit de considérer le polynôme*

$$P(x) = x^2 + 1,$$

dont les valeurs sont toujours strictement supérieures à 1 *sur* $\mathbb{R}$, *et qui n'a donc pas de racines réélles. En revanche, il a deux racines complexes* i *et* $-i$.

Notons que la deuxième assertion, à savoir la décomposition (23) est une conséquence de la première, c'est à dire l'existence d'une racine; Pour le démontrer on utilise l'algorithme de division euclidienne des polynômes, qui affirme que si P et Q sont deux polynômes, avec Q non nul, alors on peut écrire de manière unique

$$P(z) = P_1(z)Q(z) + R(z),$$

où P_1 et R sont des polynômes tels que le degré de R est inférieur strictement à celui de P. Pour a dans $\mathbb{C}$, on peut donc écrire, en divisant P par $Q(z) \equiv z - a$ qui est de degré 1

$$P(z) = (z - a)P_1(z) + c, \ c \in \mathbb{C}.$$

Si a est une racine de P alors $P(a) = 0$. En reportant dans l'identité précédente, on obtient

$$P(z) = (z - a)P_1(z),$$

avec P_1 polynome de degré $n-1$, si n est le degré de P. On recommence ensuite avec P_1... pour obtenir finalement la décomposition (23).

Exemple. Cherchons à décomposer

$$P(z) = z^3 + 3z^2 - z - 3.$$

on s'aperçoit que $a_1 = 1$ est une racine évidente. On peut donc écrire

$$P(z) = (z - 1)P_1(z),$$

où P_1 est un polynôme de degré 2. on trouve

$$P_1(z) = z^2 + 4z + 3.$$

Les racines de P_1 sont -1 et -3, on a donc

$$P(z) = (z - 1)(z + 1)(z + 3).$$

□

Remarquons pour finir que dans la décomposition (23) les n racines du polynôme ne sont pas forcément distincts. Considérons alors les k racines distinctes, avec bien entendu $k \leq n$. Quitte à renuméroter la liste, on peut supposer, sans perte de généralité, qu'il s'agit des k premières, et on peut donc écrire

$$P(z) = \prod_{i=1}^{k} (z - a_i)^{m_i}, \tag{24}$$

où l'exposant m_i, appelé *multiplicité* de la racine a_i, désigne le nombre de fois où la racine a_i apparaît dans la liste initiale. Si $m_i = 1$, alors on dit que la racine est simple. Si toutes les racines sont simples, alors il y a exactement m racines distinctes.

Nous appliquons maintenant le Théorème 8 au polynôme caractéristique P_A. En rapprochant ce résultat du Thérème 7, on conclut immédiatement :

Théorème 9. *Soit T une application de V, espace vectoriel sur $\mathbb{C}$ dans lui-même. Alors T a au moins une valeurs propre.*

De plus, si toutes les racines du polynôme caractéristique P_A sont simples, alors il existe une base de V formée de vecteurs propres de T, associés à n valeurs propres distinctes, et l'expression matricielle de T dans cette base est alors diagonale.

Remarque 10. *La condition de simplicité des valeurs propres n'est bien entendu pas une condition* nécessaire *pour être diagonalisable, elle est juste suffisante : il existe des matrices à valeurs propres multiples qui sont diagonalisables. Pour s'en convaincre, il suffit de considérer la matrice unité associé à l'identité de V dont la polynôme caractéristique est $P_{\mathrm{Id}}(z) = (1-z)^n$.*

Exemples. Considérons la matrice

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

Cette matrice est triangulaire supérieure, et son polynôme caractéristique est donné par $P_A(z) = (1-z)(2-z)$. Elle a donc deux valeurs propres simples et est donc diagonalisable. Une fonction propre associée à la valeur propre 1 est $\vec{e}_1' = \vec{e}_1$. Pour trouver une fonction propre associée à la valeur propre 2, on résout :

$$\begin{cases} x + y = 2x \\ \quad 2y = 2y \end{cases},$$

qui se résume à $y = -x$. Un vecteur propre associé à la valeur propre 2 est donc $\vec{e}_2' = (- - 1,_1)$. Dans la base $(\vec{e}_1', \vec{e}_2')$ la matrice associée prend alors la forme :

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

Considérons maintenant la matrice

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

Elle est triangulaire supérieure également, et son polynôme caractéristique est donné par $P_B(z) = (1-z)^2$. Il y a maintenant une seule valeur propre 1, de multiplicité 2. Le noyau $\mathrm{Ker}\,(B - \mathrm{Id}\,)$ est donné par l'équation

$$\begin{cases} x + y = x \\ \quad\; y = y \end{cases}$$

qui se résume à $y = x$. Le noyau $\mathrm{Ker}\,(B - \mathrm{Id}\,)$ est donc de dimension 1, engendré par le vecteur $\vec{e}_1$. La matrice n'est donc pas diagonalisable.

Si toutes les matrices ne sont pas diagonalisables lorsqu'il y a des valeurs propres de multiplicité multiple, elles sont en revanche toujours trigonalisables, comme le montre le résultat suivant.

Théorème 10. *Soit T une application linéaire de V, espace vectoriel sur $\mathbb{C}$, dans lui-même. alors il existe une base de V dans laquelle l'expression matricielle de T est triangulaire supérieure.*

Preuve. On peut démontrer ce résultat en utilisant une récurrence sur la dimension n de l'espace vectoriel On introduit dons la propriété $\mathcal{P}(n)$ suivante : " Pour tout espace vectoriel sur $\mathbb{C}$ de dimension n, il existe une base dans laquelle son expression matricielle est triangulaire supérieure".

La propriété $\mathcal{P}$ clairement vérifiée pour $n = 1$.

Supposons la propriété $\mathcal{P}(n)$vérifiée à l'ordre n, et démontrons alors qu'elle l'est aussi à l'ordre $n+1$. Soit donc V_{n+1} un espace vectoriel sur $\mathbb{C}$ de dimension $n+1$, et T une application linéaire de V_{n+1} dans lui-même. Par le théorème 9, T a au moins une valeur propre λ_1 et un vecteur propre associé $\vec{e}_1$. D'après le théorème de la base incomplète, on peut compléter ce vecteur par n autre vecteurs pour former une base $(\vec{e}_1, \vec{e}_2', \ldots, \vec{e}_{n+1}')$ de V_{n+1}. Dans cette base la matrice A'_{n+1} associée à T a alors la forme

$$A'_{n+1} = \begin{pmatrix} \lambda_1 & B \\ 0 & A'_n \end{pmatrix}$$

où B est une matrice $1 \times n$, 0 désigne la matrice $n \times 1$ nulle, et A_n est une matrice carrée $n \times n$. On a donc, en développant le déterminant donnant le polynôme caractéristique

$$P_{A_{n+1}}(z) = (z - \lambda_1) P_{A_n}(z).$$

Considérons maintenant le sous-espace vectoriel $F_n = \text{Vect}(\vec{e}_2', \ldots, \vec{e}_n)$, la projection p sur F_n parallèlement à $\text{Vect}(\vec{e}_1)$, et enfin l'application T_n de F_n dans lui-même définie par $T_n = p \circ T_{|F_n}$. Par l'hypothèse de récurrence, il existe une base $(\vec{e}_2, \ldots, \vec{e}_{n+1})$ dans laquelle la matrice associée à T_n est triangulaire supérieure. Dans la base $(\vec{e}_1, \ldots, \vec{e}_n)$ l'expression matricielle de T_{n+1} a alors la propriété désirée.

6.4 Le théorème de Cayley-Hamilton

6.4.1 Polynôme de matrices carrée

Commencons par une remarque élémentaire, mais néanmoins cruciale : dans l'espace $\mathcal{L}(V) \equiv \mathcal{L}V, V)$ des applications linéaires d'un espace vectoriel V dans lui-même, la loi de composition $\circ$ definie une loi interne : si $T, S \in \mathcal{L}(V)$, alors $T \circ S \in \mathcal{L}(V)$. Ainsi, dans $\mathcal{L}(V)$ on peut definir $T^2 = T \circ T$, et de manière générale, pour $m \in \mathbb{N}^*$

$$T^m = \underbrace{T \circ T \ldots\ldots T}_{\text{m fois}}.$$

on pose par ailleurs par convention

$$T^0 = \text{Id}_V.$$

Notons que l'on a alors les relations pour des entiers m et k

$$T^m \circ T^k = T^{m+k} \text{ et } (T^m)^k = T^{mk}.$$

Soit alors p un polynôme de degré m donnée par

$$p(z) = a_m z^m + a_{m-1} z^{n-1} + \ldots + a_1 z + a_0.$$

On peut alors former le polynôme en T donné par

$$p(T) = a_m z^m + a_{m-1} z^{n-1} + \ldots + a_1 z + a_0,$$

qui définit une application de $\mathcal{L}(V)$ dans lui-même. Les mêmes considérations s'appliquent bien entendu aux matrices, c'est à dire, on pourra considérer les polynômes de matrices.

Parmi les propriétés élémentaires , notons d'abord que si P_1 et P_2 sont deux polynômes, $P_1(T)$ et $P_2(T)$ commutent, c'est à dire

$$P_1(T) \circ P_2(T) = P_2(T) \circ P_1(T). \tag{25}$$

Par ailleurs, si un polynôme Q se décompose comme

$$Q = \prod_{i=1}^{k} P_i,$$

alors, pour toute application $T \in \mathcal{L}(V)$, on

$$\operatorname{Ker} P_1(T) + \ldots \operatorname{Ker} P_2(T) + \ldots + \operatorname{Ker} P_k(T) \subset \operatorname{Ker} Q(T). \tag{26}$$

En effet , soit $\vec{u}_i \in \operatorname{Ker} P_i(T)$. On a donc $P_i(T)\vec{u}_i = \vec{0}$, et aussi

$$Q(T)(\vec{u}_i) = \prod_{j \neq i} P_j(T)[P_i(T)(\vec{u}_i)] = \vec{0}.$$

Il en résulte que $\forall i = 1, \ldots, k$, on a

$$\operatorname{Ker} P_i(T) \subset \operatorname{Ker} Q(T),$$

et la relation (26) en résulte. Le prochain résultat, que nous admettrons permet de pousser plus loin cette analyse :

Théorème 11 (dit des noyaux)**.** *Soit Q un polynôme. On suppose que Q peut s'écrire comme*

$$Q = \prod_{i=1}^{k} P_i.$$

Si on suppose de plus que les polynômes $P_1, \ldots, P_n$ sont premiers entre eux deux à deux, alors

$$\operatorname{Ker} Q(T) = \operatorname{Ker} P_1(T) \oplus \operatorname{Ker} P_2(T) \oplus \ldots \oplus \operatorname{Ker} P_k(T). \tag{27}$$

Rappelons que deux polynômes sont premiers entre eux si les seul polynômes qui les divisent tous les deux sont les constantes (non nulles).

Terminons ce paragraphe par une dernière remarque :

Proposition 24. *Soit T une application linéaire de $\mathcal{L}(V)$. Si $\vec{e}_i$ est un vecteur propre de valeur propre λ_i, on a*

$$P(T)(\vec{e}_i) = P(\lambda_i)\vec{e}_i.$$

Si de plus T est diagonalisable, de valeurs propres $\lambda_1, \ldots, \lambda_n$, et si $(\vec{e}_1, \ldots, \vec{e}_n)$ est une base de vecteurs propres associée à $(\lambda_1, \ldots, \lambda_n)$, de sorte que l'expression matricielle de T dans cette base est $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$*, alors celle de $p(T)$ est donnée dans la base $(\vec{e}_1, \ldots, \vec{e}_n)$ par*

$$\operatorname{diag}(P(\lambda_1), \ldots, P(\lambda_n)).$$

Preuve On remarque que

$$T\vec{e}_i = \lambda_i \vec{e}_i, \ T^2 \vec{e}_i = \lambda_i T(\vec{e}_i) = \lambda^2 \vec{e}_i, \ldots, T^m(\vec{e}_i) = \lambda_i^m \vec{e}_i,$$

et donc

$$P(\vec{e}_i) = P_(\lambda_i)(\vec{e}_i),$$

ce qui permet d'énonver le résultat.

6.4.2 Enoncé du théorème

Le théorème de Cayley-Hamilton s'énonce comme suit :

Théorème 12. *Soit T una application d'un espace vectoriel sur $\mathbb{C}$ dans lui-même. Alors on a*

$$P_T(T) = 0, \tag{28}$$

c'est à dire, T annule son polynôme caractéristique.

Si on préfère le langage des matrices, cet énoncé est équivalent à

$$P_A(A) = 0, \tag{29}$$

ou A est une matrice $n \times n$.

Idée de la preuve. Considérons d'abord un cas "facile".

Preuve dans le cas diagonalisable. La matrice associée à $P_T(T)$ dans une base de vecteurs propres est donnée, d'après la Proposition 24

$$\mathrm{diag}\,(P_T(\lambda_1), \dots, P_T(\lambda_n)) = \mathrm{diag}\,(0, 0, \dots, 0) = 0,$$

car par définition, le polynôme caractéristique prend des valeurs nulles pour les valeurs propres.

Preuve dans le cas général Pour le cas général, on utilise une base dans laquelle l'expression matricielle A est triangulaire supérieur, de la forme

$$A = \begin{pmatrix} \lambda_1 & \dots & \star \\ & \vdots & \\ 0 & \dots & \lambda_n \end{pmatrix},$$

de sorte que

$$p_A(z) = (z - \lambda_1) \dots (z - \lambda_n),$$

où les valeurs propres sont repétées en fonction de leur multiplicité. L'énoncé du théorème se ramène à montrer que

$$(T - \lambda_1 \mathrm{Id}\,) \circ \dots \circ (T - \lambda_n \mathrm{Id}\,) = 0. \tag{30}$$

A cet effet, posons pour $i = 1, \dots, n$

$$T_i = (T - \lambda_1 \mathrm{Id}\,) \circ \dots \circ (T - \lambda_i \mathrm{Id}\,).$$

Pour établir (30), il suffit de vérifier que pour tout $i = 1, \dots, n$, on a

$$T_i(\vec{e}_1) = T_i(\vec{e}_2) = \dots = T_i(\vec{e}_i) = \vec{0}. \tag{31}$$

Nous allons vérifier (31) par récurrence. Pour $i = 1$, la propriété est facile à vérifier car

$$T_1(\vec{e}_1) = (T - \lambda_1 \mathrm{Id}\,)(\vec{e}_1) = T(\vec{e}_1 - \lambda_1 \vec{e}_1) = \vec{0}.$$

Supposons maintenant que la propriété soit vérifiée à l'ordre $i - 1$, c'est à dire que

$$T_{i-1}(\vec{e}_1) = T_{i-1}(\vec{e}_2) = \dots = T_{i-1}(\vec{e}_{i-1}) = \vec{0} \tag{32}$$

et montrons que $T_i(\vec{e}_1) = T_i(\vec{e}_2) = \ldots = T_i(\vec{e}_i) = \vec{0}$. Comme

$$T_i = (T - \lambda_1 \mathrm{Id}\,) \circ T_{i-1} = T_{i-1} \circ (T - \lambda_1 \mathrm{Id}\,),$$

il résulte immédiatement de l'hypothèse de récurrence (32) que

$$T_i(\vec{e}_1) = \ldots = T_i(\vec{e}_{i-1}) = \vec{0}.$$

Il reste donc à montrer que $T_i(\vec{e}_i) = \vec{0}$. On écrit

$$T_i(\vec{e}_i) = T_{i-1}(T(\vec{e}_i) - \lambda_i \vec{e}_i))$$

Comme la matrice est triangulaire supérieur, on a

$$T(e_i) = a_{1,i}\vec{e}_1 + a_{2,i}\vec{e}_2 + \cdots + \lambda_i \vec{e}_i,$$

et donc

$$T_i(\vec{e}_i) = T_{i-1}(a_{1,i}\vec{e}_1 + a_{2,i}\vec{e}_2 + \ldots + a_{i-1,i}\vec{e}_{i-1}) = \vec{0},$$

ce qui termine la récurrence, et donc la preuve.

7 Décomposition spectrale

Nous poursuivons dans ce chapitre l'étude des applications linéaires T d'un espace vectoriel V sur $\mathbb{C}$ dans lui-même. Lorsque l'application se diagonalise, ce qui est par exemple le cas lorsque toutes les racines du polynôme caractéristique sont simples, alors on peut raisonnablement prétendre que l'étude est en grande partie terminée : ce chapitre se concentre donc sur le cas non diagonalisable, notre but étant de trouver et de construire des bases dans lesquelles l'expression matricielle de l'application est simple et facile à utiliser.

Nous avons déjà franchi au chapitre précédent un premier pas dans cette direction, en montrant que toute application lineaire de V dans lui-même est trigonalisable. Ici nous désirons aller plus loin dans cette démarche.

Le concept que nous présentons maintenant est central dans la suite de notre analyse.

7.1 Applications nilpotentes

Définition 13. *une application T de V dans lui-même est dite nilpotente s'il existe un entier naturel $m \in \mathbb{N}$ tel que*

$$T^m = 0.$$

le plus petit entier m tel que cette relation soit vérifiée s'appelle l'indice de nilpotence de T.

Exemples. Les matrices suivantes sont nilpotentes :

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

d'ordre de nilpotence 2, 3 et 3 respectivement.

On a

Proposition 25. *Les assertions suivantes sont équivalentes :*
i) T est nilpotente
ii) 0 est la seul valeur propre de T
iii) le polynôme caractéristique P_T s'ecrit $P_T(z) = (-1)z^n$.

Preuve. Le fait que ii) et iii) soient équivalents résulte immediatement du Theorème de Cayley-Hamilton. Par ailleurs, si λ est valeur propre de T, alors λ^m est valeur propre de T^m, ce qui montre que i) entraîne ii).

□

Remarque 11. Notons qu'il résulte de iii) que l'indice de nilpotence est toujours inférieur à la dimension n de V.

Proposition 26. *Soit T une application linéaire de V dans lui-même d'indice de nilpotence m, et soit $\vec{e} \in \operatorname{Ker} T^m \backslash \operatorname{Ker} T^{m-1}$ un vecteur non nul. Alors la famille $\{\vec{e}, T(\vec{e})), T^2(\vec{e}), \ldots, T^{m-1}(\vec{e})\}$ est libre. Posons*

$$N = \operatorname{Vect}\{\vec{e}, T(\vec{e})), T^2(\vec{e}), \ldots, T^{m-1}(\vec{e})\}.$$

Alors N est stable par T, c'est à dire

$$T(N) \subset N.$$

Preuve. Pour la première assertion, on considère

$$\lambda_1\vec{e}+\lambda_2 T(\vec{e})+\lambda_3 T^2(\vec{e})+\ldots\lambda_{m-1}T^{m-1}(\vec{e})=0$$

une combinaison linéaire nulle de ces vecteurs. En prenant l'image par T^{m-1} des deux membres de cette identité on trouve

$$\lambda_1 T^{m-1}(\vec{e})=0,$$

car les autres termes sont nuls, en raison de l'hypothèse de nilpotence. Comme $T^{m-1}(e)\neq\vec{0}$, il en résulte que $\lambda_1=0$. On recommence alors on prenant l'image par T^{m-2}, T^{m-3}, etc.. pour montrer que les autres coefficients sont nuls également.

Pour la deuxième assertion on verifie que l'image de N est donnée par

$$\mathrm{Im}(N)=\mathrm{Vect}\{T(\vec{e})),T^2(\vec{e}),\ldots,T^{m-1}(\vec{e})\},$$

Elle est donc incluse dans N strictement.

Remarque 12. Si $m=n$, alors la famille précédente est une base de V. Si on pose $\vec{e}_1=T^{n-1}(\vec{e}),e_2=T^{n-2}(\vec{e}),\ldots e_{n-1}=T(\vec{e}),e_n=e$, alors l'expression matricielle de V dans la base $(e_1,\ldots,e_n)$ est donnée par

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

Si S est une application linéaire quelconque de V dans lui- même et $m\in\mathbb{N}$, alors on a

$$\mathrm{Ker}\,S\subset\mathrm{Ker}\,S^2\subset\ldots\mathrm{Ker}\,S^m, \tag{33}$$

et que

$$\mathrm{Im}\,T^m\subset\mathrm{Im}\,T^{m-1}\ldots\mathrm{Im}\,T \tag{34}$$

Dans le cas des applications nilpotentes, on a le résultat suivant :

Proposition 27. *Soit T une application linéaire de V dans lui-même d'indice de nilmpotence m. Alors pour tout entier $k\leq m$, on a $\mathrm{Ker}\,T^{k-1}$ est strictement inclus dans $\mathrm{Ker}\,T^k$, et $\mathrm{Im}\,T^k$ est strictement inclus dans $\mathrm{Im}\,T^{k-1}$.*

Preuve. Pour la première assertion, raisonnons par l'absurde et supposons que $\mathrm{Ker}\,T^{k-1}=\mathrm{Ker}\,T^k$: cec signifie que si $u\in V$ est tel que $T^k(u)=0$ alors nécessairement $T^{k-1}(u)=0$. Comme $T^k(u)=T(T^{k-1}u)$, on en déduit donc que

$$\mathrm{Im}\,T^{k-1}\cap\mathrm{Ker}\,T=\{\vec{0}\}.$$

Il en résulterait que la restriction de T à $\mathrm{Im}\,T^{k-1}$ est une bijection de $\mathrm{Im}\,T^{k-1}$ sur son image $\mathrm{Im}\,T^k$. Ceci serait alors vrai pour toutes les itérées T^2, T^3, etc..., et on aurait $\mathrm{Im}\,T^j=\mathrm{Im}\,T^{k-1}$ pour tout $j\geq k-1$. Ceci est bien entendu contradictoire avec l'hypothèse de nilpotence.

La deuxième assertion peut se déduire de la première par le théorème noyau-image.

Elle est triangulaire supérieure, avec une diagonale nulle, et des 1 au dessus de la diagonale, tous les autres coefficients étant nuls.

Exercice. Soit T une application linéaire de V dans lui-même, d'indice de nilpotence V. Soit $(e_1, e_2, \dots, e_k)$ une famille libre de vecteur de $\mathrm{Ker}\, T^m \setminus \mathrm{Ker}\, T^{m-1}$. Que peut-on dire de la famille

$$\{e_1, e_2, \dots e_k, T(e_1), T(e_2), \dots T(e_k), T^2(e_1), T^2(e_2), \dots T^2(e_k), \dots, T^{m-1}(e_1), T^{m-1}(e_2), \dots T^{m-1}(e_k)\}?$$

7.2 Les sous-espaces caractéristiques

Soit T une application linéaire de V dans lui-même, dont le polynôme caractéristique à la forme

$$P_T(z) = (z - \lambda_1)^{m_1} \dots (z - \lambda_k)^{m_k}$$

où $\lambda_1, \dots, \lambda_k$ désigne les valeurs propres distinctes de T, les nombres $m_1, \dots, m_k$ leur multiplicité comme racine du polynôme P_T de sorte que

$$\sum_{i=1}^{k} m_i = n = \dim V.$$

Comme $P_(T) = 0$, on a donc

$$(T - \lambda_1)^{m_1} \circ \dots \circ (z - \lambda_k)^{m_k} = 0,$$

et il résulte ainsi du Théorème des noyaux que

$$V = N_1 \oplus N_2 \oplus \dots \oplus N_k,$$

où $N_i = \mathrm{Ker}\,(T - \lambda_i \mathrm{Id}\,)^{m_i}$.

Etudions les propriétés de ces espaces.

Définition 14. *On appelle sous-espace caractéristique associé à la valeur propre λ_i le sous espace vectoriel N_i défini par*

$$N_i = \mathrm{Ker}\,(T - \lambda_i \mathrm{Id}\,)^{m_i}.$$

Rappelons que l'on appelle sous-espace propre de T associé à la valeur propre λ_i le sous-espace vectoriel F_i composé des vecteurs propres de T associés à la valeur propre λ_i, auxquels on ajoute le vecteur nul, c'est à dire

$$F_i = \mathrm{Ker}\,(T - \lambda_i \mathrm{Id}\,).$$

Notons immédiatement qu'il résulte de l'inclusion (33)

$$F_i \subset N_i.$$

On a

Proposition 28. *Pour $i = 1, \dots, k$, le sous-espace vectoriel caractéristique N_i est stable par T, c'est à dire que*

$$T(N_i) \subset N_i.$$

Preuve Soit $\vec{u} \in N_i$. Par définition de N_i, on a $(T - \lambda_i \mathrm{Id}\,)^{m_i}(\vec{u}) = 0$. Il nous faut montrer que $T(\vec{u}) \in N_i$, c'est à dire que $(T - \lambda_i \mathrm{Id}\,)^{m_i}(T(\vec{u})) = 0$. Or

$$(T - \lambda_i \mathrm{Id}\,)^{m_i}(T(\vec{u})) = (T - \lambda_i \mathrm{Id}\,)^{m_i} \circ T(\vec{u}) = T \circ (T - \lambda_i \mathrm{Id}\,)^{m_i}(u) = T\left((T - \lambda_i \mathrm{Id}\,)^{m_i}(\vec{u})\right) = T(\vec{0}) = \vec{0}.$$

ce qui termine la preuve.

On a alors le résultat central suivant, dont certaines parties viennent d'être démontré (nous admettrons les autres) :

Théorème 13. *L'espace V se décompose comme*

$$V = N_1 \oplus N_2 \oplus \ldots \oplus N_k.$$

Chaque espace caractéristique N_i est de dimension m_i exactement et stable par T. De plus la restriction à N_i de $T - \lambda_i \mathrm{Id}$ est nilpotente d'indice $r_i \leq m_i$.

Remarquons que si on considère la restriction de Tà N_i, on peut trouver une base de N_I dans laquelle sa matrice associée A_i est triangulaire supérieure, c'est à dire de la forme

$$A_i = \begin{pmatrix} \lambda_i & & \star \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix}$$

de sorte que la matrice A s'ecrit dans la base réunion des bases précédentes

$$A = \begin{pmatrix} \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ 0 & & \lambda_1 \end{pmatrix} & & & 0 \\ & \begin{pmatrix} \lambda_2 & & \star \\ & \ddots & \\ 0 & & \lambda_2 \end{pmatrix} & & \\ & & \ddots & \\ 0 & & & \begin{pmatrix} \lambda_n & & \star \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \end{pmatrix}.$$

Cette matrice est bien entendu triangulaire supérieure.

7.3 La décomposition diagonale+ nilpotente

Les résultats précédents peuvent se traduire sous la forme suivante, dont nous sera particulièrement utile dans la chapitre suivante.

Théorème 14 (Décomposition de Dunford)**.** *Soit T une application linéaire de V dans lui-même. Alors, il existe un unique couple $(D, \mathcal{N})$ d'applications de $\mathcal{L}(V)$ tels que :*
i) $T = \mathcal{D} + \mathcal{N}$.
ii) $\mathcal{D}$ est diagonalisable et $\mathcal{N}$ est nilpotente.
ii) $\mathcal{D}$ et $\mathcal{N}$ commutent, c'est à dire $\mathcal{D} \circ \mathcal{N} = \mathcal{N} \circ D$.

Idée de la preuve Montrons l'existence de cette décomposition, l'unicité étant admise.

On construit l'application $\mathcal{D}$ comme l'application linéaire dont la restruction à N_i est exactement $\lambda_i \mathrm{Id}_{|N_i}$. Il en résulte que $\mathcal{D}$ est diagonalisable, et admet chaque N_I comme sous-espace vectoriel propre pour la valeur propre λ_i. On pose ensuite

$$\mathcal{N} = T - \mathcal{D}.$$

On veut d'abord montrer que N est nilpotente et commute avec $\mathcal{D}$. Comme chaque sous-espace caractéristique N_I est stable par T et par D, il suffit donc de le vérifier pour les restrictions aux N_I.On a $T_i = \mathcal{D}_i + \mathcal{N}_i$, et $\mathcal{D}_i = \lambda_I \mathrm{Id}_{N_i}$. Par définition de N_i $(T_i - \lambda_i \mathrm{Id}_{n_i})^{m_i} = 0$, ce qui entraîne que $\mathcal{N}_i$ est nilpotente, et commute avec l'homothétie $\mathcal{D}_i$. Le couple construit conbvent donc.

Exemple. On considère la matrice

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{pmatrix}. \tag{35}$$

Commencons par calculer son polynôme caractéristique. On a

$$P_A(\lambda) = \begin{vmatrix} 2-\lambda & -1 & 1 \\ 0 & 3-\lambda & -1 \\ 2 & 1 & 3-\lambda \end{vmatrix} = (2-\lambda)\begin{vmatrix} 3-\lambda & -1 \\ 1 & 3-\lambda \end{vmatrix} + 2\begin{vmatrix} -1 & 1 \\ 3-\lambda & -1 \end{vmatrix}$$

d'où

$$\begin{aligned} P_A(\lambda) &= (2-\lambda)[(3-\lambda)^2 + 1] + 2[1 - (3-\lambda)] = (2-\lambda)[(3-\lambda)^2 - 1] \\ &= (2-\lambda)^2(4-\lambda). \end{aligned}$$

Nous avons donc deux valeurs propres distinctes $\lambda_1 = 2$, qui est double, et $\lambda_2 = 4$, qui est simple. Calculons maintenant les sous-espaces caractéristiques associés.

Détermination de N_1. La valeur propre λ_1 est de multiplication 2, calculond donc $(A - \lambda_1 \mathrm{I}_3)^2$. On a d'abord

$$A - \lambda_1 \mathrm{I}_3 = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 2 & 1 & 1 \end{pmatrix} \text{ et } (A - \lambda_1 \mathrm{I}_3)^2 = \begin{pmatrix} -2 & 0 & -2 \\ -2 & 0 & -2 \\ 2 & 0 & 2 \end{pmatrix}$$

Le vecteur (x, y, z) appartient donc à $N_{\lambda_1} = \mathrm{Ker}\ (A - \lambda_1 \mathrm{I}_3)^2$ si et seulement si $x = -z$. Choisissont maintenant un vecteur dans N_1, Par exemple $\vec{e} = (0, 1, 0)$. On a

$$(A - \lambda_1 \mathrm{I}_3)(\vec{e}) = (-1, 1, 1) \neq 0.$$

ll résulte de la Proposition 26 que $(A - \lambda_1 \mathrm{I}_3)(\vec{e}) \in F_1 = \mathrm{Ker}\, (A - \lambda_1 \mathrm{Id}_3)$, que la famille

$$(u_1, u_2) \equiv ((A - \lambda_1 \mathrm{Id}_3)(\vec{e}), \vec{e}) = ((-1, 1, 1), (0, 1, 0))$$

est une base de N_1 l'expression de la restriction de T à N_1 dans la base (u_1, u_2) est donnée par

$$A - \lambda_1 \mathrm{I}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

c'est à dire

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

Détermination de $N_2 = F_{\lambda_2}$.

De même, le vecteur (x, y, z) appartient à $F_{\lambda_2} = \mathrm{Ker}\,(A - 4\mathrm{Id}_3)$ si et seulement si

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

c'est à dire $y = -z$ et $-2x - y + z = 0$, soit $x = -y$. On a donc

$$F_{\lambda_2} = \mathrm{Vect}(\vec{u}_3),$$

où $\vec{u}_3 = (1, -1, 1)$.

On vérifie donc finalement que la base $(\vec{u}_1, \vec{u}_2, \vec{u}_3) = ((-1, 1, 1), (0, 1, 0), (1, -1, 1))$ l'expression matricielle de l'application linéaire est donnée par

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

Dans cette nouvelle base

$$A' = \mathcal{D}' + \mathcal{N}', \tag{36}$$

avec

$$\mathcal{D}' = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad \text{et } \mathcal{N}' = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

où $\mathcal{D}'$ est diagonale, $\mathcal{N}'$ est nilpotente, et où les deux matrices commutent.

8 Exponentielle de matrice

Rappelons que l'exponentielle d'un nombre complexe peut se définir à l'aide d'une suite convergente

$$\exp z = 1 + z = \frac{z^2}{2!} + \ldots \frac{z^i}{i!} + \ldots = \sum_{i=0}^{\infty} \frac{z^i}{i!}.$$

De la même facon que nous avions défini des polynômes d'applications linéaires d'un espace vectoriel dans lui-même, nous pouvons définir leur exponentielle par la formule suivante :

Définition 15. *Soit $T \in \mathcal{L}(V)$, où V est un espace vectoriel de dimension finie. On pose*

$$\exp T = \mathrm{Id}\ + T + \frac{T^2}{2!} + \ldots + \frac{T^i}{i!} + \ldots = \sum_{i=0}^{\infty} \frac{T^i}{i!}. \tag{37}$$

Bien entendu, cette définition n'a de sens que si la suite précédente converge. Plus précisement, comme T peut être représentée par une matrice en coordonnées, cela signifie que si A est l'expression matricielle de T, alors la série de matrices

$$\exp A = \mathrm{Id}\ + A + \frac{A^2}{2!} + \ldots + \frac{A^i}{i!} + \ldots = \sum_{i=0}^{\infty} \frac{A^i}{i!}$$

converge, au sens ou chacun des éléments de la matrice possède une limite : nous admettrons ce résultat. Notonqs également que $\exp A$ est une limite de polynôme de A.

Donnons maintenant quelques propriétés.

Proposition 29. *Soient T et S deux éléments de $\mathcal{L}(V)$.*
i) Si S et T commutent, alors $\exp T \circ \exp T = \exp(S + T)$.
ii) $\exp T$ *est inversible d'inverse* $\exp(-T)$.
iii) Si S est inversible alors $\exp(S^{-1}TS) = S^{-1}(\exp T)S$.

Preuve. Pour la première assertion, on multiplie les séries donnant les exponentielles, et on regroupe les termes de manière adéquate, en utilisant le fait que les deux applications commutent. Pour la seconde, on remarque que

$$\exp(T + (-T)) = \exp 0 = \mathrm{Id}\,.$$

et on utilise la première assertion pour conclure. pour la troisième, on utilise de nouveau la formule-définition.

8.1 Calculs d'exponentielles

Commençons par quelques cas particuliers simples.

Le cas des applications diagonalisables. Soit $T \in \mathcal{L}(V)$ une application diagonalisables, et soient $\lambda_1, \lambda_2, \ldots \lambda_n$), la suite des valeurs propres, répétées avec leur multiplicité. Dans une base de vecteurs propres $(\vec{u}_1, \ldots, \vec{u}_n)$, La matrice associéé dans cette base est alors

$$D = \mathrm{diag}\,[\lambda_1, \ldots, \lambda_n).$$

Rappelons que si P est un polynôme, alors

$$P(D) = \mathrm{diag}\,[P(\lambda_1), \dots, P(\lambda_n)).$$

Comme l'exponentielle est une limite de polynômes, on en déduit

$$\exp D = \mathrm{diag}\,(\exp \lambda_1, \dots, \exp \lambda_n).$$

c'est à dire

$$\exp D = \begin{pmatrix} \exp \lambda_1 & & 0 \\ 0 & \ddots & \\ 0 & \dots & \exp \lambda_n \end{pmatrix}.$$

Si $(\vec{e}_1, \dots, \vec{e}_n)$ est une autre base de V et que l'expression matricielle de T est donnée dans cette base par une matrice A, alors on a

$$D = P^{-1}AP,$$

où P est la matrice de passage de la base $(e_1, \dots, \vec{e}_n)$ vers la base $(\vec{u}_1, \dots, \vec{u}_n)$

$$\exp A = P\,\mathrm{diag}\,(\exp \lambda_1, \dots, \exp \lambda_n)\,P^{-1}.$$

Le cas des applications nilpotente. Si N est nilpotente, alors $N^m = 0$, où m est l'indice de nilpotence, d'où il résulte que

$$N^k = 0, \quad \text{pour } k \geq m.$$

si on revient à la formule (37), on s'aperçoit alors que cette dernière ne contient pour une matrice qu'un nombre fini de termes. on a donc

$$\exp N = \mathrm{Id} + N + \frac{^2}{2!} + \dots + \frac{N^{k-1}}{(k-1)!} = \sum_{i=0}^{k-1} \frac{T^i}{i!}.$$

Le cas général. Dans le cas général, on peut utiliser la décomposition diagonale+ nilpotente, c'est à dire écrire

$$T = D + N,$$

où D est diagonale, N nilpotente, et où D et N commutent.

Exemple. Reprenons l'exemple de la matrice (35) donnée au chapitre précédent . On a dans la nouvelle base $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$

$$\exp \mathcal{D}' = \begin{pmatrix} \exp 2 & 0 & 0 \\ 0 & \exp 2 & 0 \\ 0 & 0 & \exp 4 \end{pmatrix} \tag{38}$$

et comme $\mathcal{N}'^2 = 0$

$$\exp \mathcal{N}' = \mathrm{I}_3 + \mathcal{N}'.$$

On a donc

$$\begin{aligned} \exp A' &= \exp \mathcal{D}'.(\mathrm{I}_3 + \mathcal{N}') = \exp \mathcal{D}' + \exp \mathcal{D}'.\mathcal{N}' \\ &= \begin{pmatrix} \exp 2 & 0 & 0 \\ 0 & 2\exp 2 & 0 \\ 0 & 0 & \exp 4 \end{pmatrix}. \end{aligned} \tag{39}$$

Bien entendu, si on veut revenir aux coordonnées initiale il faut introduire la matrice de passage de l'ancienne base vers la nouvelle base $\mathcal{P}$, qui s'écrit ici

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

On trouve par ailleurs

$$\mathcal{P}^{-1} = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

de sorte que

$$\exp A = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \exp 2 & 0 & 0 \\ 0 & 2\exp 2 & 0 \\ 0 & 0 & \exp 4 \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

9 Applications aux équations différentielles

Les systèmes d'équations différentielles linéaires du premier ordre les plus simples peuvent se mettre la forme matricielle suivante : chercher une fonction inconnue V de $\mathbb{R}$ dans $\mathbb{C}^n$ qui vérifie pour tout $t \in \mathbb{R}$, la condition différentielle

$$\frac{d}{dt} V(t) = AV(t), \tag{40}$$

Ici A désigne une matrice $n \times n$ qui est donnée, n étant un entier.

Exemple. On recherche trois fonctions inconnues x, y z, à valeurs réelles telles que

$$\begin{cases} x'(t) = 2x(t) - y(t) + z(t) \\ y'(t) = 3y(t) - z(t) \\ z'(t) = 2x(t) + y(t) + 3z(t). \end{cases} \tag{41}$$

Si on introduit la matrice A définie par

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

et on pose, pour $t \in \mathbb{R}$ $V(t) = (x(t), y(t), z(t)$, alors le système (41) se ramène à (40), pour $n = 3$ et A donné ci-dessus (cette matrice a été étudiée et decomposée dans (35)).

9.1 Le cas réél $n = 1$

C'est le cas évidemment le plus facile, car il correspondant au cas scalaire. dans ce cas $A = a$ où a est un nombre que l'on suppose de plus réél., et $V(t) = x(t)$ une fonction supposée également réelle. L'équation s'écrit dans ce cas

$$\frac{d}{dt} x(t) = ax(t), \tag{42}$$

Pour résoudre l'équation différentielle (42) on peut utiliser, au moins formellement, la méthode de separation des variables, qui s'écrit :

$$\frac{dx(t)}{x(t)} = adt$$

qui par intégration donne

$$\ln |x(t)| = at + C,$$

où $C \in \mathbb{C}$ est une constante d'intégration quelconque réelle. Ensuite on prenant l'exponentielle, on trouve donc

$$x(t) = K \exp at, \tag{43}$$

où $K = \exp C$. On vérifie ensuite que la forme (43) avec $k \in \mathbb{R}$ donne toutes les solutions.

9.2 Le cas général

La formule (43) suggère de considérer l'application $\Phi : \mathbb{R} \to \mathcal{M}_n(\mathbb{C})$ définie par $\Phi(t) = \exp tA$. En utilisant le développement (37) on a pour $t \in \mathbb{R}$

$$\exp tA = \mathrm{I}_n + tA + t^2\frac{A^2}{2!} + \ldots + t^i\frac{A^i}{i!} + \ldots = \sum_{i=0}^{\infty} t^i\frac{A^i}{i!}.$$

En dérivant, au moins formellement, terme à terme on obtient

$$\begin{aligned}\frac{d}{dt}\exp tA &= A + tA^2 + \ldots + t^{i-1}\frac{A^i}{(i-1)!} + \ldots = \sum_{i=1}^{\infty} t^{i-1}\frac{A^i}{(i-1)!}\\ &= A\sum_{i=0}^{\infty} t^i\frac{A^i}{i!}\\ &= A\exp tA.\end{aligned}$$

Si $\vec{X}$ est un vecteur quelconque de $\mathbb{C}^n$, alors pour tout $t \in \mathbb{R}$

$$V(t) \equiv \exp tA\vec{X}$$

est un vecteur de $\mathbb{R}^n$ et

$$\frac{d}{dt}V(t) = \frac{d}{dt}[\exp tA]X = A\exp tAX = AV(t). \tag{44}$$

On a donc construit une solution de (40). En fait, on démontre qu'on les obtient toutes de cette manière. On a

Théorème 15. *Toutes les solutions de l'équation* (40) *sont de la forme*

$$V(t) = \exp tA\,\vec{X},$$

où $\vec{X} \in \mathbb{R}^n$. L'ensemble des solutions forme donc un espace vectoriel de dimension n. Pour X_0 donné, la solution $V_0(t) = \exp TA\,\vec{X}_0$ est la seule solution de (40) *telle que*

$$V_0(0) = \vec{X}_0.$$

Exemple. Reprenons l'exemple (41), qui implique la matrice A étudiée en (35) : dans la nouvelle base $(\vec{u}_1, \vec{e}_2, \vec{u}_2)$, et en raisonnant comme dans (45) on voit que

$$\exp t\mathcal{D}' = \begin{pmatrix} \exp 2t & 0 & 0 \\ 0 & \exp 2t & 0 \\ 0 & 0 & \exp 4t \end{pmatrix} \tag{45}$$

et comme $\mathcal{N}'^2 = 0$

$$\exp t\mathcal{N}' = \mathrm{I}_3 + t\mathcal{N}'.$$

On a donc

$$\begin{aligned}\exp A' &= \exp t\mathcal{D}'.(\mathrm{I}_3 + t\mathcal{N}') = \exp t\mathcal{D}' + t\exp t\mathcal{D}'.\mathcal{N}'\\ &= \begin{pmatrix} \exp 2t & 0 & 0 \\ 0 & (1+t)\exp 2t & 0 \\ 0 & 0 & \exp 4t \end{pmatrix}.\end{aligned} \tag{46}$$

Dans les coordonnées initiales, il vient donc

$$\exp tA = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \exp 2t & 0 & 0 \\ 0 & (1+t)\exp 2t & 0 \\ 0 & 0 & \exp 4t \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix},$$

qu'il ne reste plus qu'à développer (exercice), pour trouver la solution du systeme (41).