Licence Mathématiques 2ème année 1er semestre

COURS D'ALGÈBRE

Nicolas JACON

Université de Franche Comté

Table des matières

Chapitre 1

Rappels d'Algèbre linéaire

Dans ce chapitre, on rappelle les principales notions, définitions et théorèmes liés aux cours d'algèbre linéaire de première année. Dans tout ce qui suit (y compris les prochains chapitres), la lettre $\mathbb{K}$ désignera un corps égale à l'ensemble des nombres rationnelles $\mathbb{Q}$, l'ensemble des nombres réels $\mathbb{R}$ ou l'ensemble des nombres complexes $\mathbb{C}$. On notera cependant que la quasi-totalité du cours s'applique à tout corps commutatif.

1.1 Espaces vectoriels

Nous commençons par la définition d'espace vectoriel E sur un corps $\mathbb{K}$. L'idée ici est de définir une structure algébrique permettant d'effectuer des "combinaisons linéaires" c'est à dire permettant d'une part de multiplier tout élément de $\mathbb{K}$ avec tout élément de E et de pouvoir additionner plusieurs éléments de E. On pourra d'ailleurs penser à $E = \mathbb{R}^2$ sur le corps $\mathbb{K} = \mathbb{R}$

Définition 1.1.1 Un $\mathbb{K}$-*espace vectoriel* est un ensemble E muni d'une loi dite "interne" $+$ et d'une loi dite "externe " $\times$:

$$+ : \begin{array}{ccc} E \times E & \to & E \\ (x,y) & \mapsto & x+y \end{array}, \qquad \times : \begin{array}{ccc} \mathbb{K} \times E & \to & E \\ (\lambda, y) & \mapsto & \lambda y. \end{array}$$

qui vérifient les propriétés suivantes :

1. La loi d'addition est associative, c'est à dire,
$$\forall (x,y,z) \in E^3,\ x + (y+z) = (x+y) + z.$$
2. La loi d'addition est commutative c'est à dire
$$\forall (x,y) \in E^2,\ x + y = y + x.$$
3. Il existe un unique élément neutre (notée 0), c'est à dire
$$\forall x \in E,\ x + 0 = x.$$

4. Tout élément x de E admet un opposé notée $-x$ c'est à dire :
$$\forall x \in E, \exists(-x),\ x + (-x) = 0.$$
5. La loi de multiplication est associative, c'est à dire :
$$\forall(\lambda, \mu) \in \mathbb{K}^2, \forall x \in E,\ (\lambda\mu)x = \lambda(\mu x).$$
6. La multiplication est distributive par rapport à l'addition, c'est à dire
$$\forall(\lambda, \mu) \in \mathbb{K}^2, \forall(x, y) \in E^2,\ (\lambda+\mu)x = (\lambda x)+(\mu x) \text{ et } \lambda(x+y) = (\lambda x)+(\lambda y).$$
7. L'élément unité 1 du corps $\mathbb{K}$ est neutre à gauche pour la loi de multiplication c'est à dire :
$$\forall\, u \in E, \quad 1u = u.$$

Les éléments d'un $\mathbb{K}$-espace vectoriel E sont appelés *des vecteurs* tandis que les éléments de $\mathbb{K}$ sont *les scalaires*.

Quelques remarques quant à la définition ci-dessus.

Remarques.

1. L'utilisation des termes "associative" et "distributive" en 5 et 6 est impropre dans la mesure où ces termes sont usuellement réservés à une loi interne (comme l'addition) et non externe (comme la multiplication par un scalaire). Ces mots sont néanmoins employés car suffisamment parlant.
2. Un ensemble E possédant une loi interne d'addition vérifiant les quatre premiers points est appelés un *groupe commutatif*. Un $\mathbb{K}$-espace vectoriel est donc un groupe commutatif muni d'une loi externe vérifiant 5, 6 et 7.
3. On vérifie facilement :
   (a) $\forall x \in E,\ (-1)x = -x$ c'est à dire que si on multiplie le scalaire -1 avec l'élément x de E, on obtient l'élément opposé de x.
   (b) $\forall x \in E,\ 0x = 0$ c'est à dire que si on multiplie le scalaire 0 avec l'élément x de E, on obtient le neutre de E. Attention ! il ne faut pas confondre le 0 de $\mathbb{K}$ avec le 0 (le neutre) de E. Par exemple, le 0 du $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$ est $(0, 0)$, celui de $\mathbb{R}$ est 0.

Associé à cette notion d'espace vectoriel, on trouve la notion de "sous-espace vectoriel". C'est tout simplement un ensemble F contenu dans un $\mathbb{K}$-espace vectoriel E tel que F a une structure d'espace vectoriel. Pour qu'un sous-ensemble F de E soit un sous-espace vectoriel, on voit facilement qu'il suffit que l'addition de deux éléments de F restent dans F et que la multiplication d'un scalaire avec un élément de F soit dans F, d'où la définition suivante :

Définition 1.1.2 Soit E un $\mathbb{K}$-espace vectoriel. Une partie $F \subset E$ est un *sous-espace vectoriel* si elle contient 0 et est stable par l'addition et multiplication par un scalaire, c'est à dire si, pour tout $(x, y) \in F^2$ et $\lambda \in \mathbb{K}$, les vecteurs $x + y$ et λx sont encore dans F. Il revient au même de vérifier que $0 \in F$ et que pour tout $(x, y) \in F^2$ et $\lambda \in \mathbb{K}$, le vecteur $x + \lambda y$ est dans F.

Bien entendu, comme noté ci-dessus un sous-espace vectoriel d'un espace vectoriel a une structure d'espace vectoriel.

Exemples.

1. $\mathbb{R}$ est un sous-espace vectoriel de $\mathbb{C}$ vu comme $\mathbb{R}$-espace vectoriel. C'est bien sûr faux si on voit $\mathbb{C}$ comme $\mathbb{C}$-espace vectoriel.
2. L'ensemble des polynômes de degré inférieur à 3 est un sous-espace vectoriel de l'ensemble des polynômes.

Etant donnés un $\mathbb{K}$-espace vectoriel E et un ensemble H contenu dans E, la définition ci-dessus nous donne un critère pour vérifier si H est un sous-espace vectoriel. Si H n'en est pas un, il existe forcément un sous-espace vectoriel contenant H (par exemple E). Ce sous-espace vectoriel étant non nécessairement unique, on peut par exemple considérer le plus petit, c'est le but de la prochaine définition.

Définition 1.1.3 Soit E un $\mathbb{K}$-espace vectoriel et soit H une partie quelconque de E. Il existe un plus petit sous-espace vectoriel contenant H. On le note $\mathrm{Vect}(H)$ et on l'appelle *sous-espace vectoriel engendré par H*.

On voudrait maintenant avoir une description explicite de ce sous-espace vectoriel engendré par une partie H. D'après la définition ci-dessus, ce sous-espace contient néssairement l'ensemble des combinaisons linéaires des éléments de H. Une combinaison linéaire d'éléments $x_i \in H$ avec $i = 0, ..., l$ est un vecteur x s'écrivant sous la forme :

$$x = \sum_{i=0}^{l} \lambda_i x_i = \lambda_0 x_0 + ... + \lambda_l x_l.$$

où les éléments λ_i avec $i = 0, ..., l$ sont des éléments de $\mathbb{K}$.

Le théorème suivant montre que l'espace vectoriel $\mathrm{Vect}(H)$ est en fait exactement l'ensemble des combinaisons linéaires des éléments de H.

Théorème 1.1.4 *Soit E un $\mathbb{K}$ espace vectoriel et H une partie de E alors :*

$$\mathrm{Vect}(H) = \left\{ \sum_{i=0}^{l} \lambda_i x_i \mid (\lambda_0, ..., \lambda_l) \in \mathbb{K}^{l+1}, (x_0, ..., x_l) \in H^{l+1},\ l \in \mathbb{N} \right\}$$

Exemple.

1. Si H est une partie finie composée de $l+1$ vecteurs x_0,...,x_l alors :

$$\mathrm{Vect}(H) = \left\{ \sum_{i=0}^{l} \lambda_i x_i \mid (\lambda_0, ..., \lambda_l) \in \mathbb{K}^{l+1},\ (x_0, ..., x_l) \in H^{l+1} \right\}.$$

2. Si H ne contient qu'un vecteur non nul x alors

$$\mathrm{Vect}(H) = \{\lambda x \mid \lambda \in \mathbb{K}\}.$$

   Par exemple, dans $\mathbb{R}^2$, le sous-espace vectoriel engendrée par la partie $\{(1,1)\}$ de $\mathbb{R}^2$ est l'ensemble

$$\{(\lambda, \lambda) \mid \lambda \in \mathbb{K}\}.$$

   Ceci correspond a la droite d'équation $y = x$.

Dans le $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$, le sous-espace vectoriel H_1 engendré par la partie $\{(0,1)\}$ de $\mathbb{R}^2$ est l'ensemble

$$\{(0, \lambda) \mid \lambda \in \mathbb{K}\}.$$

qui correspond a la droite d'équation $x = 0$. Le sous-espace vectoriel H_2 engendrée par la partie $\{(1,0)\}$ de $\mathbb{R}^2$ est l'ensemble

$$\{(\lambda, 0) \mid \lambda \in \mathbb{K}\}.$$

qui correspond a la droite d'équation $y = 0$. le sous-espace vectoriel engendrée par la partie $H_1 \cup H_2 = \{(0,1),(1,0)\}$ de $\mathbb{R}^2$ est l'ensemble

$$\{(\lambda_0, \lambda_1) \mid (\lambda_0, \lambda_1) \in \mathbb{R}^2\}.$$

qui est égale a $\mathbb{R}^2$. En particulier, cet espace n'est pas égale a la réunion des deux droites $x = 0$ et $y = 0$! d'ailleurs, cette réunion ne forme pas un espace vectoriel car non stable par l'addition : $(0,1) + (1,0) = (1,1) \notin H_1 \cup H_2$ (voir la proposition 2.2.1 pour plus de précisions sur la réunion d'espaces vectoriels.)

Dans cet exemple, on voit que l'on arrive a décrire $\mathbb{R}^2$ au moyen de deux éléments $(1,0)$ et $(0,1)$ de $\mathbb{R}^2$ dans le sens ou un élément de $\mathbb{R}^2$ s'écrit comme combinaison linéaire de ces deux vecteurs. On peut maintenant se demander si, étant donné un espace vectoriel E, E peut être décrit (dans un sens a préciser) a l'aide d'un nombre fini d'éléments de E. Ce sont ici que les notions de famille génératrice, de famille libre et la notion de base vont apparaître.

1.2 Base d'un espace vectoriel

Définition 1.2.1 Soit E un $\mathbb{K}$-espace vectoriel et soient e_1, e_2, ..., e_n des vecteurs de E.

1. On dit que la famille $(e_1, ..., e_n)$ est une *famille libre* ou encore est *linéairement indépendente* si pour tous les scalaires λ_1,...,λ_n, on a

$$\sum_{i=1}^{n} \lambda_i e_i = 0 \Rightarrow \lambda_1 = ... = \lambda_n = 0.$$

   Si la famille n'est pas libre, on dit qu'elle est *liée*.

2. On dit que la famille $(e_1, ..., e_n)$ est une *famille génératrice* si

$$\text{Vect}(e_1, ..., e_n) = E,$$

   autrement dit pour tout $x \in E$ il existe des vecteurs x_1,...,x_n et des scalaires λ_1,...,λ_n tels que :

$$x = \lambda_1 x_1 + ... + \lambda_n x_n.$$

3. On dit que famille $(e_1, ..., e_n)$ est *une base* si c'est une famille libre et génératrice de E.

Le théorème suivant permet de definir, en toute généralité des coordonnées pour un vecteur en fonction d'une base de l'espace vectoriel.

Théorème 1.2.2 *Soit E un $\mathbb{K}$-espace vectoriel et soit $(e_1, ..., e_n)$ une famille de vecteurs de E. Les propriétés suivantes sont équivalentes :*

1. *$(e_1, ..., e_n)$ est une base.*
2. *Pour tout $x \in E$, il existe une unique famille $(\lambda_1, ..., \lambda_n)$ de scalaires tels que :*

$$x = \sum_{i=1}^{n} \lambda_i e_i.$$

   Les scalaires λ_i s'appellent les coordonnées du vecteur x dans la base $(e_1, ..., e_n)$.

Ce théorème permet de définir la dimension d'un espace vectoriel :

Définition 1.2.3 Soit E un $\mathbb{K}$-espace vectoriel. Alors on dit que E est de *dimension finie* s'il existe une famille génératrice $(e_1, ..., e_n)$ finie. La *dimension* de E est le plus petit entier n tel qu'il existe une famille génératrice de cardinal n.

En pratique, on se servira plutôt du théorème suivant.

Théorème 1.2.4 *Soit E un $\mathbb{K}$-espace vectoriel de dimensions finie. Alors E possède des bases. De plus, toutes ses bases ont même cardinal égale à la dimension de E.*

La preuve de ce théorème se fait en plusieurs étapes (voir le cours de l'année dernière). Certaines propriétés que nous citons ci-dessous nous seront particulièrement utiles.

Proposition 1.2.5 *Soit E un $\mathbb{K}$-espace vectoriel de dimensions finie et $(e_1, ..., e_n)$ une famille de vecteurs de E. Les propositions suivantes sont équivalentes*

1. *$(e_1, ..., e_n)$ est une base*
2. *$(e_1, ..., e_n)$ est une famille génératrice minimale c'est à dire qu'elle est génératrice et que pour tout $j \in \{1, ..., n\}$, la nouvelle famille obtenue $(e_1, ..., e_{j-1}, e_{j+1}, ..., e_n)$ n'est pas génératrice.*
3. *$(e_1, ..., e_n)$ est une famille libre maximale c'est à dire qu'elle est libre et que pour tout $x \in E$, la famille $(e_1, ..., e_n, x)$ est liée.*

En particulier, si on dispose d'une famille libre de n vecteurs dans un espace vectoriel E de dimension n, cette famille est en fait une base de E.

Le théorème suivant permet (la possibilité) de construire des bases à partir de la donnée d'une famille libre.

Théorème 1.2.6 (Théorème de la base incomplète) *Soit E un espace vectoriel de dimension n et soit $(e_1, ..., e_p)$ une famille libre de E (avec donc $p \leq n$). Alors il existe des vecteurs $e_{p+1}, ..., e_n$ tels que la famille $(e_1, ..., e_n)$ soit une base. En d'autres termes, toute famille libre peut être complétée pour obtenir une base.*

Nous allons maintenant aborder quelques exemples classiques.

1.3 Exemples fondamentaux

Les espaces $\mathbb{K}^n$: Pour tout entier n, on peut considérer l'espace $\mathbb{K}^n := \{(x_1, ..., x_n) \mid \forall i \in \{1, ..., n\},\ x_i \in \mathbb{K}\}$ Cet espace est naturellement muni d'une structure de $\mathbb{K}$-espace vectoriel lorsque l'addition et la multiplication par un scalaire sont définies de la façon suivante. Soit $x = (x_1, ..., x_n) \in \mathbb{K}^n$, $y = (y_1, ..., y_n) \in \mathbb{K}^n$ et $\lambda \in \mathbb{K}$, alors on pose :

$$x + y = (x_1 + y_1, ..., x_n + y_n) \qquad \lambda x = (\lambda x_1, ..., \lambda x_n)$$

On appelle base canonique la base composée des n vecteurs $e_1 = (1, 0, ..., 0)$, $e_2 = (0, 1, 0, ..., 0)$, ..., $e_n = (0, ..., 0, 1)$. Ainsi les coordonnées d'un vecteur

$x = (x_1, ..., x_n) \in \mathbb{K}^n$ dans la base canonique sont les x_i :

$$x = \sum_{i=1}^{n} x_i e_i.$$

Attention !, l'espace $\mathbb{C}^n$ est un $\mathbb{C}$-espace vectoriel comme on vient de le voir mais c'est aussi un $\mathbb{R}$-espace vectoriel. Soit $x = (x_1, ..., x_n) \in \mathbb{K}^n$, $y = (y_1, ..., y_n) \in \mathbb{C}^n$ et $\lambda \in \mathbb{R}$, alors on pose :

$$x + y = (x_1 + y_1, ..., x_n + y_n) \qquad \lambda x = (\lambda x_1, ..., \lambda x_n)$$

(la différence vient du fait que $\lambda \in \mathbb{R}$ alors que la définition de $\mathbb{C}$-espace vectoriel impose que le scalaire soit un nombre complexe).

Ici une base est donnée par les vecteurs $e_1 = (1, 0, .., 0)$, $e'_1 = (i, 0, ..., 0)$, $e_2 = (0, 1, .., 0)$, $e'_2 = (0, i, ..., 0)$,..., $e_n = (0, 0, .., 1)$, $e'_n = (0, 0, ..., 1)$ et vu comme $\mathbb{R}$-espace vectoriel, $\mathbb{C}$ est de dimension $2n$ alors que vu comme $\mathbb{C}$-espace vectoriel, $\mathbb{C}$ est de dimension n.

L'espace $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$: Il s'agit d'un espace vectoriel de dimension infini (l'addition et la multiplication par un scalaire sont définis de façon triviale.) La base canonique est donnée par la base $(1, X, X^2, ..., X^k, ...)$ donc les coordonnées d'un polynôme dans cette base correspondent aux coefficients de celui-ci. On peut également considérer l'espace $\mathbb{K}_n[X]$ des polynômes à coefficients dans $\mathbb{K}$ de degré infèrieur à n. C'est un sous-espace vectoriel de $\mathbb{K}[X]$ et la base canonique est donnée par $(1, X, X^2, ..., X^n)$ et l'espace est donc de dimension $n + 1$.

L'espace des suites $\mathbb{K}^{\mathbb{N}}$ à valeurs dans $\mathbb{K}$: L'addition et la multiplication par un scalaire sont données de la façon suivante. Pour $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ dans $\mathbb{K}^{\mathbb{N}}$, on pose

$$(u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}} = (u_n + v_n)_{n\in\mathbb{N}} \qquad \lambda(u_n)_{n\in\mathbb{N}} = (\lambda u_n)_{n\in\mathbb{N}}$$

Cet espace est de dimension infini.

L'espace des fonctions continues de $I \subset \mathbb{R}$ à valeurs dans $\mathbb{K}$. De même, L'addition et la multiplication par un scalaire étant donnés trivialement (on a ici besoin du fait que la somme de deux fonctions continues est continue et que la multiplication d'un scalaire par une fonction continue donne une fonction continue.)

L'espace des matrices $\mathcal{M}_{n,p}(\mathbb{K})$ à n lignes et p colonnes à coefficients dans $\mathbb{K}$. Un élément quelconque de $\mathcal{M}_{n,p}(\mathbb{K})$ s'écrit

$$\begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}$$

La somme et la multiplication par un scalaire sont définis trivialement :

$$\begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1p}+b_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \cdots & a_{np}+b_{np} \end{pmatrix}$$

$$\lambda \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1p} \\ \vdots & \ddots & \vdots \\ \lambda a_{n1} & \cdots & \lambda a_{np} \end{pmatrix}$$

La base canonique est consituée des np vecteurs E^{kl} pour $1 \leq k \leq n$ et $1 \leq l \leq p$ où E^{kl} est la matrice n'ayant que des 0 sauf en position (k, l) où l'on trouve 1. Ce $\mathbb{K}$-espace vectoriel est donc de dimension np.

Chapitre 2

Applications linéaires

Ayant défini les espaces vectoriel, nous allons maintenant nous intéresser aux applications naturelles préservant la structure de tels espaces.

2.1 Premières définitions

Définition 2.1.1 Soient E et F deux $\mathbb{K}$-espaces vectoriels. Une application $f : E \to F$ est dite *linéaire* si elle vérifie pour tout $(x, y) \in E^2$ et pour tout $\lambda \in \mathbb{K}$

$$f(x + y) = f(x) + f(y) \qquad \text{et} \qquad f(\lambda x) = \lambda f(x).$$

L'ensemble des applications de E dans F est noté $L(E, F)$.

Remarques. Si $f : E \to F$ est une application linéaire alors notons que $f(0) = f(0+0) = f(0) + f(0)$. Par conséquent, $f(0) = 0$ et l'image de 0 par une application linéaire est toujours 0. Par contre, la réciproque est fausse, c'est à dire qu'il existe des applications linéaires tels que $f(x) = 0$ et $x \neq 0$ (à faire en exercice)

Notons que les deux propriétés ci-dessus peuvent être reformulées de sorte que $f : E \to F$ est linéaire si et seulement si elle vérifie pour tout $(x, y) \in E^2$ et pour tout $\lambda \in \mathbb{K}$

$$f(x + \lambda y) = f(x) + \lambda f(y).$$

Définition 2.1.2 Soient E et F deux $\mathbb{K}$-espaces vectoriels.

1. Une application linéaire de E dans E est appelée un *endomorphisme*. On note $L(E) := L(E, E)$ l'ensemble des endomorphismes de E.
2. Une application linéaire bijective est appelée un *isomorphisme*.
3. Une application linéaire bijective d'une espace vectoriel E dans lui-même est appelée un *automorphisme*. L'ensemble des automorphismes de E est noté $GL(E)$.

Soient E et F deux $\mathbb{K}$-espaces vectoriels, considérons l'ensemble $L(E,F)$ Alors on peut vérifier que cet espace est muni d'une structure d'espace vectoriel sur $\mathbb{K}$. Pour ceci, on définit les deux opérations suivantes, soient f et g dans $L(E,F)$ et soit $\lambda \in \mathbb{K}$:

$$\begin{array}{rccc} f+g: & E & \to & F \\ & x & \mapsto & f(x)+g(x) \end{array} \qquad \begin{array}{rccc} \lambda f: & E & \to & F \\ & x & \mapsto & \lambda f(x) \end{array}$$

On doit ici établir les propriétés de la définition 1.1.1 (à faire en exercice).

Exemples.

1. L'application f de $\mathbb{R}$ dans $\mathbb{R}$ définie par $f(t) = t^2$ pour tout $t \in \mathbb{R}$ n'est pas linéaire. En effet $f(2) = 4 \neq 2f(1)$.
2. Si $a \in \mathbb{K}$ est un nombre fixé, alors l'application $f : \mathbb{K} \to \mathbb{K}$ définie par $f(t) = at$ pour tout $t \in \mathbb{K}$ est linéaire. En effet, si $t_1 \in \mathbb{K}$, si $t_2 \in \mathbb{K}$ et si $\lambda \in \mathbb{K}$ alors $f(t_1 + \lambda t_2) = a(t_1 + \lambda t_2) = f(t_1) + \lambda f(t_2)$.
3. Réciproquement, si $f : \mathbb{K} \to \mathbb{K}$ est une application linéaire, alors en posant $a = f(1)$ on obtient, pour tout $t \in \mathbb{K}$, $f(t) = tf(1) = at$. Donc toute application linéaire de $\mathbb{K}$ dans $\mathbb{K}$ est de la forme $f = a\mathrm{Id}_E$ où $a \in \mathbb{K}$.

Exercices.

1. Soit p, q, r trois entiers strictement positifs. Soit $A \in \mathcal{M}_{qr}(\mathbb{K})$ une matrice fixée. Montrer que l'application $f : \mathcal{M}_{pq}(\mathbb{K}) \to \mathcal{M}_{pr}(\mathbb{K})$ définie par $\forall M \in \mathcal{M}_{qr}(\mathbb{K})$, $f(M) = MA$ est une application linéaire. On prendra soin de vérifier que f est bien définie de $\mathcal{M}_{pq}(\mathbb{K})$ dans $\mathcal{M}_{pr}(\mathbb{K})$.
2. Montrer que l'application $f : \mathbb{R}^3 \to \mathcal{M}_{2,2}(\mathbb{R})$ définie par $\forall (x,y,z) \in \mathbb{R}^3$, $f(x,y,z) = \begin{pmatrix} x & x-y \\ x+z & 2y \end{pmatrix}$ est une application linéaire.

Définition 2.1.3 Soient E et F deux espaces vectoriels et soit $f : E \to F$ une application linéaire.

1. On note $\mathrm{Ker}(f)$ et on appelle noyau de f l'ensemble suivant :
$$\mathrm{Ker}(f) := \{x \in E \mid f(x) = 0\} \subset E.$$
2. On note $\mathrm{Im}(f)$ et on appelle image de f l'ensemble suivant :
$$\mathrm{Im}(f) := \{y \in F \mid \exists x \in E,\ f(x) = y\} \subset F.$$

L'étude des deux ensembles ci-dessus va nous permettre de déduire des propriétés importantes de l'application associée, propriétés qui justifient la définition d'application linéaire.

Proposition 2.1.4 *Soient E et F deux $\mathbb{K}$-espaces vectoriels et soit $f : E \to F$ une application linéaire.*

1. *$\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$ sont des sous-espaces vectoriels de respectivement E et F.*
2. *f est injective si et seulement si $\mathrm{Ker}(f) = \{0\}$.*
3. *f est surjective si et seulement si $\mathrm{Im}(f) = F$.*

Preuve.

1. Montrons que $\mathrm{Ker}(f)$ est un sous-espace vectoriel. Comme $f(0) = 0$, on a déjà $0 \in \mathrm{Ker}(f)$. Nous devons maintenant prouver que, étant donnés x et y quelconques dans $\mathrm{Ker}(f)$ et $\lambda \in \mathbb{K}$ quelconque, nous avons $x + \lambda y \in \mathrm{Ker}(f)$. Pour cela, nous calculons :

   $$f(x + \lambda y) = f(x) + \lambda f(y) = 0 + \lambda \cdot 0 = 0.$$

   Ceci prouve que $x+\lambda y \in \mathrm{Ker}(f)$ et donc que $\mathrm{Ker}(f)$ est un sous-espace vectoriel de E.

   Montrons que $\mathrm{Im}(f)$ est un sous-espace vectoriel. On a $f(0) = 0$ donc $0 \in \mathrm{Im}(f)$. Nous devons maintenant prouver que, étant donnés x' et y' quelconques dans $\mathrm{Im}(f)$ et $\lambda \in \mathbb{K}$ quelconque, nous avons $x' + \lambda y' \in \mathrm{Im}(f)$. Puisque x' et y' sont dans $\mathrm{Im}(f)$, ils possèdent des antécédents par f. Il existe donc deux vecteurs x et y dans E tels que $f(x) = x'$ et $f(y) = y'$. Dès lors, nous pouvons remarquer que :

   $$f(x + \lambda y) = f(x) + \lambda f(y) = x' + \lambda y'.$$

   Ceci prouve que $x' + \lambda y'$ possède un antécédent par f (à savoir le vecteur $x + \lambda y$) et donc $x' + \lambda y' \in \mathrm{Im}(f)$. Ceci démontre que $\mathrm{Im}(f)$ est un sous-espace vectoriel de F.

2. Supposons d'abord que f est injective. Comme $f(0) = 0$ et que tout élément de l'espace d'arrivée a au plus un antécédent dans l'espace de départ (définition de l'injectivité), on voit que 0 est le seul antécédent de 0 par f. Par conséquent, $\mathrm{Ker}(f) = \{0\}$.

   Supposons maintenant que $\mathrm{Ker}(f) = \{0\}$. Pour prouver que f est injective, il nous faut montrer que, si deux éléments de l'espace de départ ont la même image, alors ils sont égaux. Par conséquent, il nous faut montrer que, si $x, y \in E$ sont tels que $f(x) = f(y)$, alors $x = y$. Calculons $f(x - y)$. Comme l'application f est linéaire, cela vaut $f(x - y) = f(x) - f(y) = 0$. Comme $\mathrm{Ker}(f) = \{0\}$, ceci va imposer que $x - y = 0$ et donc que $x = y$. Nous avons bien montré que f était injective.

3. Par définition de la surjectivité, f est surjective si et seulement si chaque élément de l'ensemble d'arrivée possède au moins un antécédent

par f. Or $\mathrm{Im}(f)$ est justement l'ensemble des éléments de l'ensemble d'arrivée possédant au moins un antécédent par f. On voit donc que f est surjective si et seulement si $\mathrm{Im}(f) = F$.

2.2 Somme directe de sous-espaces

2.2.A Somme directe de deux sous-espaces

Comme nous l'avons vu dans le chapitre précédent, la réunion de deux sous-espaces vectoriels n'est pas en général un sous-espace vectoriel. En fait la proposition suivante montre que cette réunion n'en est presque jamais un :

Proposition 2.2.1 *Soient F et G deux sous-espaces vectoriels d'un espace vectoriel E. Alors $F \cup G$ est un sous-espace vectoriel de E si et seulement si $F \subset G$ (et alors $F \cup G = G$) ou $G \subset F$ (et alors $G \cup F = F$)*

Preuve. Soient F et G deux sous-espaces vectoriels de E. Raisonnons par contraposée en supposant que F n'est pas contenue dans G et que G n'est pas contenue dans F. Montrons alors que $F \cup G$ n'est pas un sous-espace vectoriel de E. Il existe un élément x_F qui est dans F et qui n'est pas dans G. De même, il existe un élément x_G qui est dans G mais qui n'est pas dans F. Montrons que $y = x_F + x_G \notin F \cup G$. On raisonne par l'absurde : si $y = x_F + x_G \in F \cup G$ alors soit $y \in F$ mais alors $x_G = y - x_F \in F$ ce qui est absurde soit $y \in G$ mais alors $x_F = y - x_G \in G$ ce qui est absurde également. Bref, on a $y = x_F + x_G \notin F \cup G$ ce qui implique que $F \cup G$ n'est pas un sous-espace vectoriel puisque on a trouvé deux éléments de $F \cup G$ dont la somme n'est pas dans $F \cup G$.

La réciproque est évidente : si $F \subset G$, on a $F \cup G = G$ qui est un sous-espace vectoriel de E et si $G \subset F$ alors $G \cup F = F$ qui est un sous-espace vectoriel de E.

□

Puisque la réunion de deux sous-espaces vectoriels n'est "pratiquement jamais" un sous-espace vectoriel, on peut se demander si on peut décrire le plus petit sous-espace vectoriel contenant cette réunion.

Proposition 2.2.2 *Soient F et G deux sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E. L'ensemble*

$$F + G := \{x_F + x_G \mid x_F \in F,\ x_G \in G\}$$

est un sous-espace vectoriel de E. Il s'agit du plus petit sous-espace vectoriel de E contenant $F \cup G$ c'est à dire $\mathrm{Vect}(F \cup G)$.

Preuve. On montre tout d'abord que $F+G$ est bien un sous-espace vectoriel de E. On a déjà $0 \in F+G$ car $0 \in F$ et $0 \in G$. Soient $x \in F+G$, $y \in F+G$

et $\lambda \in \mathbb{K}$. On veut montrer que $x + \lambda y \in F + G$. Comme $x \in F + G$, il existe $x_F \in F$ et $x_G \in G$ tels que $x = x_F + x_G$. De même, comme $y \in F + G$, il existe $y_F \in F$ et $y_G \in G$ tels que $y = y_F + y_G$. On a :

$$x + \lambda y = x_F + x_G + \lambda(y_F + y_G) = x_F + \lambda y_F + y_G + \lambda y_G$$

Or, F est un sous-espace vectoriel de E donc, comme $x_F \in F$ et $y_F \in F$, on a $x_F + \lambda y_F \in F$. De même G est un sous-espace vectoriel de E donc, comme $x_G \in G$ et $y_G \in G$, on a $x_G + \lambda y_G \in F$. Il suit que $x + \lambda y$ s'écrit comme somme d'un élément de E et un de F donc $x + \lambda y \in F + G$.

Il faut maintenant montrer que $F + G$ est le plus petit sous-espace vectoriel de E contenant $F \cup G$. Il est clair que $F + G$ contient F et G donc $F \cup G$. On considère maintenant un sous-espace vectoriel H de E contenant $F \cup G$ et on va montrer que $F + G$ est "plus petit" que H au sens de l'inclusion c'est à dire que $F + G \subset H$. Soit donc $x \in F + G$. Alors, il existe $x_F \in F$ et $x_G \in G$ tels que $x = x_F + x_G$. Mais $x_F \in F \subset F \cup G$ donc $x_F \in F \cup G$ et par conséquant $x_F \in H$. De même, on a $x_G \in H$. Mais comme H a une structure de sous-espace vectoriel, on a $x = x_F + x_G \in H$. On obtient donc $F + G \subset H$.

□

Considérons deux sous-espaces vectoriels F et G d'un espace vectoriel E. Tout élément de $F + G$ s'écrit, par définition comme somme d'un élément de F et d'un élément de G. Cette écriture est-elle unique ? en général la réponse est non, il suffit de considérer les sous-espaces vectoriels $F = \text{Vect}((1,0,0),(0,1,0))$ et $G = \text{Vect}((1,1,0),(0,0,1))$ de $\mathbb{R}^3$. On a par exemple

$$(1,1,1) = ((1,0,0) + (0,1,0)) + (0,0,1) = (0,0,0) + ((1,1,0) + (0,0,1))$$

qui donne deux manières distinctes de décomposer $(1,1,1)$ en somme d'éléments de F et de G. Ce qui semble poser problème ici est que le vecteur $(1,1,0)$ appartient à la fois à F et à G.

Nous allons maintenant déterminer quand nous pourrons parlons d'unicité pour cette décomposition d'un élément de E en somme d'un élément de F et un de G. □

Définition 2.2.3 Soient F et G deux sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E. On dit que F et G sont en *somme directe* lorsque $F \cap G = \{0\}$. La somme de F et de G se note alors $F \oplus G$ et on parle de somme directe de F et de G.

Proposition 2.2.4 *Soient F et G deux sous-espaces vectoriels d'un $\mathbb{K}$ espace vectoriel E. Les assertions suivantes sont équivalentes :*

1. *F et G sont en somme directe*

2. *tout élément de $F+G$ s'écrit de manière unique sous la forme x_F+x_G où $x_F \in F$ et $x_G \in G$.*

Preuve. Supposons tout d'abord que F et G sont en somme directe. Soit $x \in F+G$ alors il existe $x_F \in F$ et $x_G \in G$ tels que $x = x_F + x_G$. Il faut montrer que cette décomposition est unique. Soit donc $x = x'_F + x'_G$ une autre décomposition avec $x'_F \in F$ et $x'_G \in G$. Il suit donc $x_F + x_G = x'_F + x'_G$ soit encore $x_F - x'_F = x_G - x'_G$. Ce dernier élément est donc dans $F \cap G$. Il suit $x_F = x'_F$ et $x_G = x'_G$. D'où le résultat.

Supposons maintenant que tout élément de $F+G$ de décompose de manière unique. Pour montrer que F et G sont en somme directe, il faut montrer que $F \cap G = \{0\}$. Soit donc $x \in F \cap G$. Le vecteur 0 appartient à $F+G$ et il s'écrit comme somme d'élément de F et de G sous la forme $0 = 0 + 0$, il s'écrit aussi $0 = x + (-x)$ avec $x \in F$ et $(-x) \in G$. La décomposition étant unique, on a $x = 0$.

□

Définition 2.2.5 On dit que deux sous-espaces vectoriels F et G d'un $\mathbb{K}$-espace vectoriel E sont *supplémentaires* lorsque $F \oplus G = E$ c'est à dire lorsque $F + G = E$ et $F \cap G = \{0\}$. Il revient au même de dire que tout vecteur de E se décompose de manière unique comme somme d'un élément de F et d'un élément de G.

Exemple.

1. Dans le plan réel, deux droites distinctes sont supplémentaires. En effet, deux droites vectorielles distinctes D_1 et D_2 se coupent uniquement en 0 (attention, on parle ici de sous-espaces vectoriels !) donc elles sont en somme directe. De plus, si x_1 est un vecteur directeur de D_1 et x_2 un vecteur directeur de D_2, alors x_1 et x_2 ne sont pas colinéaires. Ils engendrent donc un espace de dimension 2, c'est-à-dire le plan tout entier.
2. Dans l'espace $E = C(\mathbb{R})$ des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$, on considère le sous-espace vectoriel F des fonctions paires et le sous-espace vectoriel G des fonctions impaires. Ces deux sous-espaces vectoriels sont supplémentaires.

   En effet, si une fonction h est à la fois paire et impaire, elle vérifie, pour tout $x \in \mathbb{R}$: $h(-x) = h(x)$ et $h(-x) = -h(x)$ donc $h(x) = 0$. C'est donc la fonction nulle. Ceci prouve que $F \cap G = \{0\}$. De plus, si nous prenons une fonction h quelconque dans $C(\mathbb{R})$, nous pouvons écrire :

   $$h = f + g,$$

   avec, pour tout $x \in \mathbb{R}$,

   $$f(x) = \frac{h(x) + h(-x)}{2} \text{ et } g(x) = \frac{h(x) - h(-x)}{2}.$$

La fonction f est paire tandis que la fonction g est impaire. Ceci prouve que $F \oplus G = E$. Par exemple, on peut remarquer que la fonction exponentielle se décompose sous la forme

$$\exp = \mathrm{ch} + \mathrm{sh}.$$

où ch est une fonction paire et sh une fonction impaire.

Définition 2.2.6 Soient F et G deux sous-espaces supplémentaires d'un $\mathbb{K}$-espace vectoriel E. Pour tout x dans E, il existe un unique élément $x_F \in F$ et un unique élément $x_G \in G$ tels que $x = x_F + x_G$. On appelle x_F le projeté de x sur F parallèlement à G. On appelle x_G le projeté de x sur G parallèlement à F. Les applications

$$\begin{array}{rccc} p: & E & \to & F \\ & x & \mapsto & x_F \end{array} \qquad \begin{array}{rccc} q: & E & \to & G \\ & x & \mapsto & x_G \end{array}.$$

s'appellent les projecteurs sur F (resp. sur G) parallèlement à G (resp. F)

Proposition 2.2.7 *Sous les notations de la définition ci-dessus, p est une application linéaire avec*

$$\mathrm{Ker}(p) = G \qquad \textit{et} \qquad \mathrm{Im}(p) = F.$$

Preuve. Montrons tout d'abord que p est une application linéaire. Soit x et y des éléments de E et soit $\lambda \in \mathbb{K}$. Il existe des uniques éléments x_F, y_F dans F et x_G, y_G dans G tels que

$$x = x_F + x_G \qquad \text{et} \qquad y = y_F + y_G.$$

On veut déterminer $p(x + \lambda y)$ pour ceci, il faut trouver la décomposition (unique) de $x + \lambda y$ en somme d'un élément de F et d'un élément de G. On a

$$x + \lambda y = x_F + \lambda y_F + x_G + \lambda y_G.$$

La décomposition est trouvée car $x_F + \lambda y_F \in F$ et $x_G + \lambda y_G \in G$, E et F étant des sous-espaces vectoriels. Il suit

$$p(x + \lambda y) = x_F + \lambda y_F = p(x) + \lambda p(y),$$

ce qu'il fallait montrer.

Montrons que $\mathrm{Ker}(p) = G$. Soit $x \in G$, alors x s'écrit $0 + x$ avec $0 \in F$ et $x \in G$. On a donc $p(x) = 0$ et donc $x \in \mathrm{Ker}(p)$. On en déduit $G \subset \mathrm{Ker}(p)$. Soit maintenant $x \in \mathrm{Ker}(p)$. On a donc $p(x) = 0$. D'autre part, il existe $x_F \in F$ et $x_G \in G$ tels que $x = x_F + x_G$. Par définition $p(x) = x_F$. Il suit $x = x_G \in G$. Donc $\mathrm{Ker}(p) \subset G$.

Montrons enfin que $\mathrm{Im}(p) = F$. Soit $x \in F$ alors $p(x) = x$ et donc $x \in \mathrm{Im}(p)$. Réciproquement, si $x \in \mathrm{Im}(p)$ alors il existe $y \in E$ tel que $p(y) = x$. Comme x est le projeté de y sur F, on a $x \in F$. On conclut $\mathrm{Im}(p) = F$. □

2.2.B Somme directe de plusieurs sous-espaces vectoriels

On veut maintenant généraliser cette notion de somme directe à plusieurs sous-espaces vectoriels. Intuitivement, nous voulons par exemple que trois droites D, D' et D'' soient en somme directe si et seulement si elles engendrent un sous-espace de dimension 3. La première idée serait de demander que ces sous-espaces vectoriels soient en somme directes deux à deux. Mais ceci ne convient pas : si on considère un plan avec deux axes formés par D et D' perpendiculaire et si on prend pour D'' la première bissectrice, nous voyons que ces trois droites vérifient $D \cap D'' = \{0\}$, $D' \cap D'' = \{0\}$, $D \cap D' = \{0\}$. Pourtant elles engendrent un sous-espace de dimension 2 : le plan.

On voit donc que la bonne généralisation va venir la proposition 2.2.4 c'est à dire de la propriété de décomposition unique. Notons tout d'abord que si F_1,...,F_p sont p sous-espaces vectoriels d'un sous-espace vectoriel E alors le plus petit sous-espace vectoriel $\text{Vect}(F_1 \cup ... \cup F_p)$ contenant $F_1 \cup ... \cup F_p$ est $F_1 + ... + F_p$ par une récurrence immédiate utilisant la prop 2.2.2.

Définition 2.2.8 Soit E un $\mathbb{K}$-espace vectoriel et F_1,...,F_p des sous-espaces vectoriels de E. On note $\sum_{i=1}^{p} F_i = F_1 + ... + F_p$ l'ensemble des vecteurs $x \in E$ qui s'écrivent sous la forme $x = \sum_{i=1}^{p} x_i$ avec $x_i \in F_i$ pour $i \in \{1, ..., p\}$. On dit que cette somme est *directe* si tout élément de $\sum_{i=1}^{p} F_i$ s'écrit de façon unique sous la forme

$$x = \sum_{i=1}^{p} x_i.$$

où $x_i \in F_i$ pour $i \in \{1, ..., p\}$. On écrit alors :

$$\bigoplus_{i=1}^{p} F_i := F_1 \oplus ... \oplus F_p$$

à la place de $F_1 + ... + F_p$.

Le problème d'une telle définition est qu'elle n'est pas aisée à vérifier ... un analogue à la définition 2.2.5, utilisant les intersections, va aussi être disponible dans ce cadre :

Proposition 2.2.9 *Soit E un $\mathbb{K}$-espace vectoriel et soient F_1, ... F_p des sous-espaces vectoriels de E. Les F_i sont en somme directe si et seulement*

si pour chaque $1 \leq i \leq p$*, on a*

$$F_i \cap \left(\sum_{i \neq j} F_j \right) = \{0\}.$$

Preuve. Supposons d'abord que les F_i sont en somme directe. On sait alors que tout vecteur $x \in \sum_{i=1}^{p} F_i$ s'écrit de façon unique sous la forme

$$x = x_1 + \cdots + x_p.$$

où pour tout $j \in \{1, ..., p\}$, on a $x_j \in F_j$.

Soit x un vecteur appartenant à la fois à F_i et à la somme des F_j pour $j \neq i$. Il s'écrit donc

$$x = x_1 + \cdots + x_{i-1} + x_{i+1} + \cdots + x_p.$$

où pour tout $j \in \{1, ..., p\}$ avec $j \neq i$, on a $x_j \in F_j$.

On a alors

$$0 = x - x = x_1 + \cdots + x_{i-1} + x_{i+1} + \cdots + x_p - x.$$

Dans cette expression, on a $x_j \in F_j$ pour tout $j \in \{1, ..., p\}$ avec $j \neq i$ et on a aussi $(-x) \in F_i$. On a donc écrit 0 comme somme d'éléments des F_j. Or 0 s'écrit de façon triviale $0 + ... + 0$ où $0 \in F_j$ pour tout $j \in \{1, ..., p\}$. Cette décomposition étant unique, il suit que $x = 0$ ce qui prouve que $F_i \cap \left(\sum_{j \neq i} F_j \right) = \{0\}$.

Supposons maintenant que la seconde assertion est vérifiée. Soit un vecteur x dans la somme des F_i qui se décompose de deux manières à priori différentes. On peut alors écrire :

$$x = x_1 + \cdots + x_p \text{ et } x = y_1 + \cdots + y_p$$

où pour tout $j \in \{1, ..., p\}$ on a $x_j \in F_j$ et $y_j \in F_j$.

Nous voulons prouver que pour chaque i, nous avons $x_i = y_i$. Fixons-nous donc un $1 \leq i \leq p$. On a alors :

$$0 = x - x = \sum_{j=1}^{p} (x_j - y_j) \text{ donc } y_i - x_i = \sum_{j \neq i} (x_j - y_i).$$

Comme $y_i - x_i \in F_i$ et que, d'autre part, $\sum_{j \neq i} (x_j - y_j) \in \sum_{j \neq i} F_j$. Ceci prouve que $y_i - x_i \in F_i \cap \left(\sum_{j \neq i} F_j \right)$ qui est réduit à $\{0\}$ d'après notre hypothèse. Ainsi $y_i - x_i = 0$ et donc $y_i = x_i$. C'est ce que nous voulions prouver.

□

Exercice. Montrer que des sous-espaces vectoriels $F_1, \ldots, F_p$ sont en somme directe si et seulement si la propriété suivante est vérifiée :

Dès qu'on choisit des vecteurs non nuls $x_1 \in F_1, \ldots, x_p \in F_p$, la famille $(x_1, \ldots, x_p)$ est libre.

Théorème 2.2.10 *Soient $F_1, \ldots, F_p$ des sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel de dimension finie. On suppose que $F_1, \ldots, F_p$ sont en somme directe. Posons $E = F_1 \oplus \cdots \oplus F_p$. Si chaque F_i est muni d'une base $\mathcal{B}_i$, alors le recollement $\mathcal{B}$ des $\mathcal{B}_i$ est une base de E.*

Preuve. Posons $\mathcal{B}_1 = (e_1, \ldots, e_k)$, $\mathcal{B}_2 = (e_{k+1}, \ldots, e_l)$ ... $\mathcal{B}_p = (e_m, \ldots, e_n)$. La famille $\mathcal{B}$ n'est autre que $(e_1, \ldots, e_n)$. Nous devons montrer que cette famille est libre et génératrice.

- Montrons que $\mathcal{B}$ est libre. Soit donc $(\lambda_i)_{i=1,\ldots,n}$ une suite d'éléments de $\mathbb{K}$ telle que :
$$\sum_{i=1}^{n} \lambda_i e_i = 0.$$
On a
$$0 = \lambda_1 e_1 + \cdots + \lambda_k e_k + \cdots + \lambda_m e_m + \cdots + \lambda_n e_n.$$
qui est la décomposition de 0 en somme d'éléments de F_1 (puisque l'on a $\lambda_1 e_1 + \cdots + \lambda_k e_k \in F_1$) ,..., de F_p (car $\lambda_m e_m + \cdots + \lambda_n e_n \in F_p$). Comme la somme des F_i est directe, ceci implique
$$\lambda_1 e_1 + \cdots + \lambda_k e_k = 0, \quad \ldots \quad , \lambda_m e_m + \cdots + \lambda_n e_n = 0$$
La première de ces relations est une combinaison linéaire nulle de la famille $\mathcal{B}_1$. Comme $\mathcal{B}_1$ est libre, on a donc $\lambda_1 = \cdots = \lambda_k = 0$. On procède de même pour chacune des autres combinaisons et on voit donc que tous les λ_i sont nuls.
- Montrons maintenant que $\mathcal{B}$ est génératrice. Comme $\mathcal{B}$ contient toutes les familles $\mathcal{B}_i$, le sous-espace vectoriel $\mathrm{Vect}(\mathcal{B})$ contient les sous-espaces vectoriels $F_1, \ldots, F_p$. Comme nous avions prouvé que $E = F_1 + \cdots + F_p$ est le plus petit sous-espace vectoriel contenant les F_i, ceci prouve que $E \subset \mathrm{Vect}(\mathcal{B})$. Par conséquent, $\mathcal{B}$ est une famille génératrice de E.

□

Corollaire 2.2.11 *Dans un $\mathbb{K}$-espace vectoriel de dimension finie, on considère des sous-espaces vectoriels $F_1, \ldots, F_p$ en somme directe. On a alors :*
$$\dim(F_1 \oplus \cdots \oplus F_p) = \dim(F_1) + \cdots + \dim(F_p).$$

Preuve. C'est clair grâce au théorème 2.2.10 puisque la dimension d'un espace vectoriel est le nombre de vecteurs dans une base.

□

Remarque. Grâce à ce corollaire, on pourra aisément vérifier que certains sous-espaces vectoriels ne sont pas en somme directe en calculant les dimensions. Par exemple, dans l'espace (c'est à dire en dimension 3), deux plans (c'est à dire des sous-espaces vectoriels de dimension 2) ne peuvent être en somme direct.

On va maintenant prouver la réciproque du théorème ci-dessus grâce auxquel on pourra facilement montrer que des sous-espaces vectoriels sont en somme directe

Théorème 2.2.12 *Soient $F_1, \ldots, F_p$ des sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel de dimension finie. Posons $E = F_1 + \cdots + F_p$. Pour tout i, on suppose que chaque F_i est muni d'une base $\mathcal{B}_i$ tel que le recollement $\mathcal{B}$ des $\mathcal{B}_i$ est une base de E. Alors les F_i sont en somme directe, autrement dit $E = F_1 \oplus \cdots \oplus F_p$*

Preuve. Soit $1 \leq i \leq p$, on veut montrer

$$F_i \cap \left(\sum_{i \neq j} F_j \right) = \{0\}$$

Raisonnons par l'absurde en supposant qu'il existe un élément non nul x dans cet ensemble. Posons $\mathcal{B}_i = (e_1, \ldots, e_k)$, $\mathcal{B}_{i+1} = (e_{k+1}, \ldots, e_l)$ ... $\mathcal{B}_{i-1} = (e_m, \ldots, e_n)$. La famille $\mathcal{B}$ n'est autre que $(e_1, \ldots, e_n)$.

Alors x s'écrit comme combinaison linéaire des éléments de la base de F_i :

$$x = \lambda_1 e_1 + ... + \lambda_k e_k$$

Mais x appartient aussi à $\sum_{i \neq j} F_j$ donc il s'écrit comme combinaison linéaire des éléments des $\mathcal{B}_j$ avec $j \neq i$:

$$x = \lambda_{k+1} e_{k+1} + ... + \lambda_n e_n.$$

En écrivant l'égalité obtenu, comme le recollement $\mathcal{B}$ des $\mathcal{B}_i$ est une base de E, il suit que x est nul.

□

Exemple. Ainsi étant donné un espace vectoriel, si on considère une base de cette espace $(e_1, ..., e_n)$, les sous-espaces vectoriels $\text{Vect}(e_i)$ pour $i \in \{1, ..., n\}$ sont en somme directe.

Proposition 2.2.13 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. Soient F_1, ..., F_p tels que*

$$E = F_1 \oplus ... \oplus F_p$$

Soit H un autre $\mathbb{K}$-espace vectoriel et $f_i : F_i \to H$ des applications linéaires. Alors il existe une unique appplication linéaire $u : E \to H$ prolongeant les f_i (dans le sens où $u|_{F_i} = f_i$) pour tout i.)

Preuve. Supposons que $u : E \to H$ soit une application linéaire prolongeant chaque f_i. Pour tout $x \in E$ qui se décompose de manière unique sous la forme $x = x_1 + \cdots + x_p$ où $x_i \in F_i$ pour tout $i \in \{1, ..., p\}$. On a alors :

$$\begin{aligned} u(x) &= u(x_1 + \cdots + x_p) \\ &= u(x_1) + \cdots + u(x_p) && \text{puisque } u \text{ est linéaire} \\ &= f_1(x_1) + \cdots + f_p(x_p) && \text{puisque } u \text{ prolonge chaque } f_i. \end{aligned}$$

Par conséquent, pour que u possède les propriétés ci-dessus, nous devons choisir cette application-ci. Ceci prouve l'unicité. Pour l'existence, il suffit de vérifier que l'application u définie par $u(x) = f_1(x_1)+\cdots+f_p(x_p)$ convient c'est à dire que c'est une application linéeire prolongeant les f_i.

L'application est linéaire : soient $x \in E$, $y \in E$ et $\lambda \in \mathbb{K}$. Alors on a des décompositions $x = x_1 + ... + x_p$ et $y = y_1 + ... + y_p$ avec $x_i \in F_i$ et $y_i \in F_i$ pour tout $i \in \{1, ..., p\}$. On a alors

$$x + \lambda y = x_1 + \lambda y_1 + ... + x_p + \lambda y_p.$$

avec, comme chaque F_i est un sous-espace vectoriel $x_i + \lambda y_i \in F_i$ pour tout $i \in \{1, ..., p\}$. On a ainsi

$$u(x + \lambda y) = f_1(x_1 + \lambda y_1) + \cdots + f_p(x_p + \lambda y_p),$$

et comme les f_i sont des applications linéaires, il suit :

$$\begin{aligned} u(x + \lambda y) &= f(x_1) + ... + f(x_p) + \lambda(f(y_1) + ... + f(y_p)) \\ &= u(x) + \lambda u(y) \end{aligned}$$

Finalement, u prolonge chaque f_i. En effet, il suffit de montrer que pour tout $i \in \{1, ..., p\}$ et $x \in F_i$, on a $u(x) = f(x_i)$ ce qui est clair en décomposant x sous la forme $0 + ... + x + ... + 0$ avec $x \in F_i$ et $0 \in F_j$ si $j \neq i$.

□

2.3 Applications linéaires et matrices

On va maintenant définir et étudier la correspondance entre applications linéaires et matrices.

2.3.A Matrice d'une application linéaire

Théorème 2.3.1 (Définition d'une application linéaire sur une base) *Soit E un $\mathbb{K}$-espace vectoriel de dimension p et soit $(e_1, ..., e_p)$ une base de E. Soit F un autre $\mathbb{K}$-espace vectoriel et $(y_1, ..., y_p)$ une famille quelconque d'éléments de F. Alors il existe une unique application linéaire $u : E \to F$ telle que $u(e_i) = y_i$ pour tout $1 \leq i \leq p$.*

Preuve. Supposons qu'on dispose d'une telle application linéaire. Soit $x \in E$ un vecteur quelconque. Posons $x = \sum_{i=1}^{p} \lambda_i e_i$ avec des scalaires $\lambda_i \in \mathbb{K}$. Alors on a :

$$u(x) = \sum_{i=1}^{p} \lambda_i u(e_i)$$

car u est linéaire. Il suit donc :

$$u(x) = \sum_{i=1}^{p} \lambda_i y_i$$

Ainsi, si u existe, u doit nécessairement vérifier $u(x) = \sum_{i=1}^{p} \lambda_i y_i$. Ceci prouve que si u existe, il est unique.

Réciproquement, posons donc pour $x = \sum_{i=1}^{p} \lambda_i e_i$

$$u(x) = \sum_{i=1}^{p} \lambda_i y_i.$$

Alors on vérifie facilement que u défini une application linéaire.

□

Proposition 2.3.2 *Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension finies. Soit $(e_1, ..., e_p)$ une base de E et soit $(y_1, ..., y_p)$ une famille quelconque de vecteurs de F. Soit $u : E \to F$ l'unique application linéaire envoyant e_i sur y_i pour tout $i \in \{1, ..., p\}$.*

1. *u est injective si et seulement si $(y_1, ..., y_p)$ est une famille libre,*
2. *u est surjective si et seulement si $(y_1, ..., y_p)$ est une famille génératrice de F,*
3. *u est bijective si et seulement si $(y_1, ..., y_p)$ est une base de F,*
4. *E et F sont isomorphes (c'est-à-dire qu'il existe un isomorphisme de E sur F) si et seulement si $dim(E) = dim(F)$.*

Preuve.

1. Comme u est linéaire, u est injective si et seulement si $\mathrm{Ker}(u) = \{0\}$. Soit $x = \sum_{i=1}^{p} \lambda_i e_i$ un vecteur quelconque de l'espace de départ. On a :

$$u(x) = \sum_{i=1}^{p} \lambda_i y_i.$$

Par conséquent, on a :

$$u(x) = 0 \iff \sum_{i=1}^{p} \lambda_i y_i = 0.$$

Ainsi u est injective si et seulement si $\mathrm{Ker}(u) = \{0\}$, ce qui équivaut à dire que $\sum_{i=1}^{p} \lambda_i y_i = 0 \Longrightarrow \sum_{i=1}^{p} \lambda_i e_i = 0$, ce qui équivaut encore à dire que tous les λ_i sont nuls puisque la famille $(e_1, \ldots, e_p)$ est libre. On a donc montré :

$$u \text{ est injective } \iff \begin{array}{c} \text{Pour tous scalaires } \lambda_i \in \mathbb{K}, \\ [\sum_{i=1}^{p} \lambda_i y_i = 0 \Longrightarrow \lambda_1 = \cdots = \lambda_p = 0] \end{array},$$

ce qui signifie bien que u est injective si et seulement si la famille $(y_1, \ldots, y_p)$ est libre.

2. Comme chaque $x \in E$ s'écrit sous la forme $\sum_{i=1}^{p} \lambda_i e_i$, nous voyons que l'ensemble les vecteurs de F dans $\mathrm{Im}(u)$ est exactement l'ensemble

$$\left\{ \sum_{i=1}^{p} \lambda_i y_i \mid (\lambda_1, \ldots, \lambda_p) \in \mathbb{K} \right\} = \mathrm{Vect}(y_1, \ldots, y_p).$$

Par conséquent nous avons

$$\begin{array}{rl} u \text{ est surjective} & \iff \mathrm{Im}(u) = F \\ & \iff \mathrm{Vect}(y_1, \ldots, y_p) = F \\ & \iff (y_1, \ldots, y_p) \text{ est génératrice.} \end{array}$$

3. Il suffit d'utiliser les deux points précédents.
4. Si $\dim(E) = \dim(F) = p$, on peut introduire une base $(f_1, \ldots, f_p)$ de F et l'unique application linéaire $u : E \to F$ telle que $u(e_i) = f_i$ pour tout $1 \leq i \leq p$. Comme $(f_1, \ldots, f_p)$ est une base de F, le point précédent montre que u est bijective, c'est-à-dire que u est un isomorphisme et donc E et F sont isomorphes.

   Réciproquement, si E et F sont isomorphes, il existe un isomorphisme $u : E \to F$. On peut définir une famille $(f_1, \ldots, f_p)$ en posant $f_i = u(e_i)$ pour chaque $1 \leq i \leq p$. D'après le point précédent, la famille $(f_1, \ldots, f_p)$ est une base de F donc $\dim(F) = p = \dim(E)$.

□

Exemple. Dans l'espace vectoriel $E = \mathbb{K}_n[X]$, on peut considérer la base canonique $(1, X, \ldots, X^n)$. L'endomorphisme qui au vecteur de base 1 associe 1 et au vecteur de base X^k $(k \geq 1)$ associe le vecteur $(X+1)^k$ est un isomorphisme. car $\{1, (X+1), ..., (X+1)^n\}$ est une base de E.

En revanche, l'endomorphisme qui au vecteur de base 1 associe 0 et au vecteur de base X^k ($k \geq 1$) associe le vecteur kX^{k-1} n'est ni injectif ni surjectif car $\{0, ..., nX^{n-1}\}$ n'est ni libre ni génératrice de E. Cette application s'appelle la dérivation.

Donnons nous une application linéaire de E dans F, deux $\mathbb{K}$-espaces vectoriels de dimension fini. Soient $\mathcal{B}_1 = (e_1, ..., e_p)$ une base de E et soit $\mathcal{B}_2 = (f_1, ..., f_n)$ une base de F (on suppose donc que E est de dimension p et F de dimension n). On a vu que l'application u est entièrement déterminée pas les valeurs des vecteurs $u(e_j)$ exprimés dans la base $(f_1, ..., f_n)$.

On suppose que pour $i \in \{1, ..., p\}$, il existe des scalaires $a_{ij} \in \mathbb{K}$ avec $j \in \{1, ..., n\}$ tel que :

$$u(e_i) = \sum_{j=1}^{n} a_{ji} f_j$$

On obtient alors la matrice $M \in \mathcal{M}_{np}(\mathbb{K})$ en prenant pour vecteurs colonnes les coordonnées des $u(e_i)$ en fonction des f_j. Cette matrice est appelée la matrice de u dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$. On la note :

$$M = \mathrm{Mat}_{\mathcal{B}_2, \mathcal{B}_1}(u) = \begin{pmatrix} a_{11} & a_{12} & .... & a_{1p} \\ a_{21} & a_{22} & ..... & a_{2p} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{np} \end{pmatrix}$$

Notez bien que sous le symbole Mat, on écrit la base d'arrivée $\mathcal{B}_2$ avant la base de départ $\mathcal{B}_1$. Cette convention est utile pour présenter les formules de changement de base comme on le verra dans la suite de ce chapitre.

Récriproquement, si on se donne une matrice M dans $\mathcal{M}_{np}(\mathbb{K})$ alors pour tout $\mathcal{B}_1 = (e_1, ..., e_p)$ une base de E et tout $\mathcal{B}_2 = (f_1, ..., f_n)$ une base de F, on a une unique application linéaire de E dans F dont la matrice dans ces bases est M.

Si on se donne une application linéaire u, pour chaque choix de base $\mathcal{B}_1$ et $\mathcal{B}_2$, on trouvera une matrice M. Cette matrice sera à priori différente pour chaque choix de bases : à une application, il correspond plusieurs matrices. Réciproquement, pour une matrice donnée, il existe plusieurs applications linéaire (une pour chaque choix de base).

Considérons maintenant trois espaces vectoriels E_1, E_2 et E_3, de dimension n, p et q et munies de bases $\mathcal{B}_1$, $\mathcal{B}_2$ et $\mathcal{B}_3$. Soient également deux applications linéaires :

$$u : E_1 \to E_2$$

$$v : E_2 \to E_3$$

On peut composer ces deux applications pour obtenir l'application

$$v \circ u : E_1 \to E_2 \to E_3$$

Quelle est la matrice $\mathrm{Mat}_{\mathcal{B}_3,\mathcal{B}_1}(v \circ u)$ de cette application composée en fonction de $\mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(u)$ et $\mathrm{Mat}_{\mathcal{B}_3,\mathcal{B}_2}(u)$?

Celle ci s'obtient en multipliant ces deux matrices. Le produit pour les matrices est donné par la formule suivante. Soit $A = (a_{ij}) \in \mathcal{M}_{np}(\mathbb{K})$ et soit $B = (b_{ij}) \in \mathcal{M}_{pq}(\mathbb{K})$ alors A et B sont multipliables et leur produit est la matrice $C = (c_{ij}) \in \mathcal{M}_{nq}(\mathbb{K})$ définie par :

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

On obtient donc :

$$\mathrm{Mat}_{\mathcal{B}_3,\mathcal{B}_1}(v \circ u) = \mathrm{Mat}_{\mathcal{B}_3,\mathcal{B}_2}(v)\mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(u)$$

On constate que dans cette formule, tout se passe comme pour une relation de Chasles d'où l'intérêt de la convention d'écriture.

On vérifiera également que :

$$\mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(u+v) = \mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(u) + \mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(v)$$

et

$$\mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(\lambda u) = \lambda \mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(u).$$

Comme on a une structure naturelle d'espace vectoriel sur $L(E,F)$ et sur $\mathcal{M}_{np}(\mathbb{K})$

$$\begin{array}{rccc} T: & L(E,F) & \to & \mathcal{M}_{np}(\mathbb{K}) \\ & u & \mapsto & \mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1} u \end{array}$$

on a le résultat suivant.

Proposition 2.3.3 *L'application T est un isomorphisme d'espace vectroiels. On a de plus*

$$dim(L(E,F))) = \dim(E)\dim(F).$$

Preuve. Les bases $\mathcal{B}_1$ et $\mathcal{B}_2$ étant fixées, les résultats ci-dessus montrent que T est une application linéaire. Comme à chaque matrice correspond un unique endomorphisme, cette application est bijective. c'est donc un isomorphisme. En particulier, les dimensions des deux espaces sont les mêmes.

□

2.3.B Changement de bases

Dans cette partie, on se donne deux espaces vectoriels E_1 et E_2 munis de bases $\mathcal{B}_1$ et $\mathcal{B}_2$ respectivement. Soit u une application linéaire de E_1 dans E_2 et soit $\mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(u)$ la matrice de cette application dans les bases ci-dessus.

A un vecteur $x \in E_1$, on associe le vecteur $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ où les x_i sont les coordonnées de x dans la base $\mathcal{B}_1$, c'est à dire :

$$x = x_1 e_1 + ... + x_p e_p$$

De même, à un vecteur $y \in E_2$, on associe le vecteur $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ où les y_j sont les coordonnées de y dans la base $\mathcal{B}_2$.

Soit $u : E_1 \to E_2$ une application linéaire et soit $M = \mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(u)$. Nous avons

$$y = u(x) \iff Y = MX$$

Que se passe t-il maintenant si nous changeons de base ? Soit $\mathcal{B}'_1$ une autre base de E_1 et soit $\mathcal{B}'_2$ une autre base de E_2. Soit X' et Y' les vecteurs colonnes des coordonnées de x et y dans les nouvelles bases $\mathcal{B}'_1$ et $\mathcal{B}'_2$

Définition 2.3.4 On appelle *matrice de passage* de la base $\mathcal{B}_1 = (e_1, ..., e_p)$ à la base $\mathcal{B}'_1 = (e'_1, ..., e'_p)$ et on note $P_{\mathcal{B}_1 \to \mathcal{B}'_1}$ la matrice dont les colonnes sont les coordonnées des vecteurs e'_i dans la base $\mathcal{B}_1$ c'est à dire, si

$$e'_j = d_{1j} e_1 + d_{2j} e_2 + ... + d_{pj} e_p$$

pour $j = 1, ..., p$, on a

$$P_{\mathcal{B}_1 \to \mathcal{B}'_1} = \begin{pmatrix} d_{11} & ... & d_{1p} \\ \vdots & \ddots & \vdots \\ d_{p1} & ... & d_{pp} \end{pmatrix}$$

On voit que la matrice de passage de la base $\mathcal{B}_1$ à la base $\mathcal{B}'_1$ n'est autre que la matrice de l'application linéaire identité où la base de départ considérée est $\mathcal{B}'_1$ et la base d'arrivée est $\mathcal{B}_1$. On a donc :

$$P_{\mathcal{B}_1 \to \mathcal{B}'_1} = \mathrm{Mat}_{\mathcal{B}_1,\mathcal{B}'_1}(\mathrm{Id}_E)$$

Le théorème suivant donne les principales formules impliquant ces matrices de passages. Elles permettent de passer d'une base à l'autre aisément. Il est très important de savoir maîtriser ces formules et on pourra facilement les mémoriser en pensant aux relations de Chasles.

Théorème 2.3.5 *Soient $\mathcal{B}_1$, $\mathcal{B}'_1$ et $\mathcal{B}''_1$ trois bases d'un même espace vectoriel E_1 et soit $\mathcal{B}_2$, $\mathcal{B}'_2$ deux bases de E_2. On a :*

$$P_{\mathcal{B}_1 \to \mathcal{B}'_1} = \mathrm{Mat}_{\mathcal{B}_1,\mathcal{B}'_1}(\mathrm{id}_E) \qquad X = P_{\mathcal{B}_1 \to \mathcal{B}'_1} X'.$$

$$P_{\mathcal{B}_1\to\mathcal{B}'_1}P_{\mathcal{B}'_1\to\mathcal{B}''_1}=P_{\mathcal{B}_1\to\mathcal{B}''_1}\qquad P_{\mathcal{B}'_1\to\mathcal{B}_1}=\left(P_{\mathcal{B}_1\to\mathcal{B}'_1}\right)^{-1}.$$
$$\mathrm{Mat}_{\mathcal{B}'_2,\mathcal{B}'_1}(u)=P_{\mathcal{B}'_2\to\mathcal{B}_2}\left[\mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(u)\right]P_{\mathcal{B}_1\to\mathcal{B}'_1}.$$

Preuve. La première formule a été établie précédemment. Pour la seconde, nous considérons l'application linéaire $u=\mathrm{id}_E$ allant de E muni de la base $\mathcal{B}'_1$ dans E_1 muni de la base $\mathcal{B}$. Sa matrice est, d'après la première formule : $P_{\mathcal{B}_1\to\mathcal{B}'_1}$. La formule voulue traduit donc matriciellement la relation $x=u(x)=\mathrm{id}_E(x)$.

Si on remplace maintenant $P_{\mathcal{B}_1\to\mathcal{B}'_1}$ par $\mathrm{Mat}_{\mathcal{B}'_1,\mathcal{B}_1}(\mathrm{id}_E)$ (et de même avec $P_{\mathcal{B}'_1\to\mathcal{B}_1}$ etc), chacune des formules restantes devient évidente. Faites-le.

□

La quatrième formule du théorème ci-dessus montre qu'une matrice de passage est toujours inversible (ce qui n'est pas étonnant étant donnée l'application linéaire qu'elle représente dans les bases appropriées !). Si on note :

$$M=\mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(u),\ M'=\mathrm{Mat}_{\mathcal{B}'_2,\mathcal{B}'_1}(u),\ P:=P_{\mathcal{B}_1\to\mathcal{B}'_1},\ Q:=P_{\mathcal{B}'_2\to\mathcal{B}_2}$$

Alors on voit que :

$$M'=QMP$$

où P et Q sont des matrices inversibles. Réciproquement, si une telle formule existe, alors on peut voir M et M' comme des matrices d'une même application linéaire (dans des bases différentes), les matrices P et Q représentant les matrices de passages dans ces bases.

Définition 2.3.6 Soient M et M' deux matrices de $\mathcal{M}_{np}(\mathbb{K})$. Les conditions suivantes sont équivalentes :

1. M et M' représentent la même application linéaire (dans des bases à priori différentes)
2. il existe des matrices inversibles $P\in\mathrm{GL}_p(\mathbb{K})$ et $Q\in\mathrm{GL}_n(\mathbb{K})$ tels que $M'=QMP$.

On dit alors que M et M' sont *équivalentes.*

Exemple. Dans $\mathbb{R}^2$, on considère deux bases : d'une part la base canonique $(\varepsilon_1,\varepsilon_2)$ et d'autre part la base (e_1,e_2) donnée par $e_1=(1,1)$ et $e_2=(1,2)$. L'application linéaire $u:\mathbb{R}^2\to\mathbb{R}^3$ définie par $u(x_1,x_2)=(2x_1,3x_2,x_1+x_2)$ a pour matrice dans les bases $(\varepsilon_1,\varepsilon_2)$ au départ et la base canonique $(\varepsilon_1,\varepsilon_2,\varepsilon_3)$ à l'arrivée : $M=\begin{pmatrix}2&0\\0&3\\1&1\end{pmatrix}$. Comme $u(e_1)=(2,3,2)$ et $u(e_2)=(2,6,3)$, sa matrice dans les bases (e_1,e_2) au départ et $(\varepsilon_1,\varepsilon_2,\varepsilon_3)$ à l'arrivée est $M'=\begin{pmatrix}2&2\\3&6\\2&3\end{pmatrix}$. Ces deux matrices représentent u dans des bases différentes donc elles sont équivalentes.

2.3.C Cas particulier des endomorphismes

Les résultats ci-dessus se simplifient lorsque l'on parle d'endomorphismes, c'est à dire d'applications linéaires d'un espace vectoriel E dans lui-même. Dans ce cas, il est possible de choisir comme base de départ et d'arrivée pour un endomorphisme u la même base $\mathcal{B}_1$. La matrice obtenue est plus simplement notée $\mathrm{Mat}_{\mathcal{B}_1}(u)$ au lieu de $\mathrm{Mat}_{\mathcal{B}_1,\mathcal{B}_1}(u)$

Si $\mathcal{B}'_1$ est une autre base, on a alors :

$$\mathrm{Mat}_{\mathcal{B}'_1}(u) = P_{\mathcal{B}'_1\to\mathcal{B}_1}\mathrm{Mat}_{\mathcal{B}_1}(u)P_{\mathcal{B}_1\to\mathcal{B}'_1} = P_{\mathcal{B}'_1\to\mathcal{B}_1}\mathrm{Mat}_{\mathcal{B}_1}(u)P^{-1}_{\mathcal{B}'_1\to\mathcal{B}_1}$$

Définition 2.3.7 Soient M et M' deux matrices de $\mathcal{M}_n(\mathbb{K})$. Les assertions suivantes sont équivalentes :

1. M et M' représentent le même endomorphisme dans des bases différentes (en imposant que les bases de départ et d'arrivée soient les mêmes)
2. Il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $M' = P^{-1}MP$.

Dans ce cas, on dit que M et M' sont *semblables*. La relation ainsi définie s'appelle *relation de similitude*.

On voit ainsi que la relation de similitude (qui n'a de sens que pour les matrices carrées) revient à imposer que deux matrices sont équivalentes et que de plus $Q = P^{-1}$. C'est donc une relation plus forte et à priori beaucoup plus délicate que l'équivalence.

2.4 Rang d'une application linéaire

2.4.A Définition

Définition 2.4.1

1. Soit E un $\mathbb{K}$ espace vectoriel et $(x_1, ..., x_p)$ une famille de vecteurs. On appelle *rang* de la famille $(x_1, ..., x_p)$ la dimension du sous-espace vectoriel qu'elle engendre. On la note $\mathrm{rg}(x_1, ..., x_p)$.

$$\mathrm{rg}(x_1, ..., x_p) := \dim(\mathrm{Vect}(x_1, ..., x_p))$$

2. Soient E et F deux $\mathbb{K}$ espaces vectoriels et soit $u : E \to F$ une application linéaire. On appelle rang la dimension de son image. On la note $\mathrm{rg}(u)$:

$$\mathrm{rg}(u) = \dim(\mathrm{Im}(u))$$

3. Soit $M \in \mathcal{M}_{np}(\mathbb{K})$. On note $C_1 \in \mathcal{M}_{n1}(\mathbb{K})$, ..., $C_p \in \mathcal{M}_{n1}(\mathbb{K})$ les vecteurs colonnes de M. On appelle rang de M le rang de la famille $(C_1, ..., C_p)$.

Le lien entre le rang d'une application linéaire et le rang de ses matrices est donné par la proposition suivante.

Proposition 2.4.2 *Soient E et F deux $\mathbb{K}$-espaces vectoriels munis de bases $\mathcal{B}_1$ et $\mathcal{B}_2$. Soit $u \in L(E,F)$ une application linéaire et $M = \mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(u)$ sa matrice dans les bases $\mathcal{B}_1$ au départ et $\mathcal{B}_2$ à l'arrivée. Alors $\mathrm{rg}(M) = \mathrm{rg}(u)$. En particulier, si M et M' sont deux matrices équivalentes alors $\mathrm{rg}(M) = \mathrm{rg}(M')$.*

Preuve. Notons $n = \dim(F)$ et $p = \dim(E)$. Pour $y \in F$, on note $y_1, \ldots, y_n$ ses coordonnées sur la base $\mathcal{B}_2$ et $\phi(y) = Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ le vecteur colonne des coordonnées. L'application $\phi : F \to \mathcal{M}_{n1}(\mathbb{K})$ est un isomorphisme. Par conséquent, il conserve la dimension. Comme les vecteurs colonnes C_j sont les coordonnées des $u(e_j)$ sur la base $\mathcal{B}_2$, on a $C_j = \phi(u(e_j))$. Par conséquent, nous avons :

$$\mathrm{rg}(u) = \dim(u(e_1), \ldots, u(e_p)) = \dim(C_1, ..., C_p) = \mathrm{rg}(C_1, \ldots, C_p) = \mathrm{rg}(M).$$

□

2.4.B Propriétés

Nous allons maintenant établir une version précisée du théorème du rang vu en première année. Pour cela, nous avons besoin de la notion d'application induite : soit $u \in L(E,F)$ une application linéaire et G un sous-espace vectoriel de E. On peut alors restreindre l'application u de façon à obtenir une application $u|_G : G \to F$ de G dans F qui reste bien sûr linéaire.

Si on considère maintenant un sous-espace vectoriel H de F, est-il possible d'obtenir une application de E dans H ? il faut pour ceci que $u(E) \subset H$. On note alors $u|^H : E \to H$ cettte application. Notons que ce n'est pas la même application que u puisque l'espace d'arrivée n'est pas le même.

Définition 2.4.3 Soient E et F deux $\mathbb{K}$ espaces vectoriels et $u \in L(E,F)$ une application linéaire. Soit G un sous-espace vectoriel de E et H un sous-espace vectoriel de F. Lorsque $u(G) \subset H$, on peut définir l'application $u|_G^H : G \to H$ obtenue par restriction de u au départ et à l'arrivée. On dit que u *induit* l'application $u|_G^H$ entre G et H.

Théorème 2.4.4 (Théoreme du rang précisé) *Soient E et F deux $\mathbb{K}$ espaces vectoriels. Soit $u : E \to F$ une application linéaire. Soit G un supplémentaire quelconque de $\mathrm{Ker}(u)$ dans E. Alors u induit un isomorphisme entre G et $\mathrm{Im}(u)$*

Preuve. On considére l'application $u|_G^{\mathrm{Im}(u)} : G \to \mathrm{Im}(u)$ elle est bien définie car $u(G) \subset u(E) = \mathrm{Im}(u)$. Montrons son injectivité en calculant son noyau.

Soit $x \in G$ tel que $u|_G^{\mathrm{Im}(u)}(x) = 0$. On a alors $u(x) = 0$ et donc $x \in \mathrm{Ker}(u)$. On a donc $x \in G \cap \mathrm{Ker}(u)$ et donc $x = 0$ puisque $G \cap \mathrm{Ker}(u) = \{0\}$.

Montrons maintenant la surjectivité. Soit $y \in \mathrm{Im}(u)$. Il existe $x' \in E$ tel que $u(x') = y$. On sait que $\mathrm{Ker}(u) + G = E$ donc il existe $x_K \in \mathrm{Ker}(u)$ et $x' \in G$ tels que $x = x_K + x'$. On obtient $u(x) = u(x_K) + u(x') = 0 + y$. On a donc bien trouvé un antécédent x' de y dans G d'ou la surjectivité.

□

On en déduit le théorème du rang usuel. Rappelons ici que si E est un $\mathbb{K}$-espace vectoriel de dimension n et F un sous-espace vectoriel de E, on note $\mathrm{codim}(F)$ l'entier positif ou nul $n - \dim(E)$

Corollaire 2.4.5 *Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies et soit $u : E \to F$ une application linéaire. Alors*

$$\dim(E) = \dim(\mathrm{Ker}(u)) + \mathrm{rg}(u),$$

soit encore :

$$\mathrm{rg}(u) = \mathrm{codim}(\mathrm{Ker}(u)).$$

Preuve. Soit G un supplémentaire de $\mathrm{Ker}(u)$. On sait que l'on a $\dim(G) = \mathrm{codim}(\mathrm{Ker}(u))$. Le théorème du rang précisé assure que G et $\mathrm{Im}(u)$ sont isomorphes. Ils ont donc la même dimension.

□

Ainsi en dimension finie, l'injectivité, la surjectivité et la bijectivité sont équivalentes comme le montre la proposition suivante.

Corollaire 2.4.6 *Soit E et F deux $\mathbb{K}$-espaces vectoriels de même dimension (finie) n. Soit $u : E \to F$ une application linéaire. Alors pour montrer que u est un isomorphisme, il suffit de montrer que u est injective ou que u est surjective. En particulier, pour un endomorphisme $u \in L(E)$, les assertions suivantes sont équivalentes.*

(i) u est un automorphisme,
(ii) $\mathrm{Ker}(u) = \{0\}$,
(iii) $\mathrm{rg}(u) = n$

Preuve. Si u est injective alors $\mathrm{Ker}(u) = \{0\}$ donc $\mathrm{rg}(u) = \mathrm{codim}\mathrm{Ker}(u) = n = \dim(F)$ donc $\mathrm{Im}(u) = F$ et u est surjective.

Si u est surjective alors $\mathrm{rg}(u) = \dim(F) = n$ donc $\dim(\mathrm{Ker}(u)) = n - \mathrm{rg}(u) = 0$ donc $\mathrm{Ker}(u) = \{0\}$ et u est injective.

Le cas particulier des endomorphismes est alors immédiat.

□

2.4.C Matrices simples d'une application linéaire

La version précisée du théorème du rang va nous permettre de trouver une matrice particulièrement simple pour n'importe quelle application linéaire.

Soient donc E et F deux $\mathbb{K}$-espaces vectoriels de dimensions respectives p et n. Soit $u : E \to F$ une application linéaire. Nous notons $r = \mathrm{rg}(u)$. Soit G un supplémentaire de $\mathrm{Ker}(u)$. D'apres le théorème du rang, nous savons que $\dim(G) = r$. Soit donc $(e_1, ..., e_r)$ une base de G. Soit $(e_{r+1}, ..., e_p)$ une base de $\mathrm{Ker}(u)$. Alors $\mathcal{B}_1 = (e_1, ..., e_p)$ est une base de E. D'après la version précisée du théorème du rang, on sait que u induit un isomorphisme de G sur $\mathrm{Im}(u)$. La famille $(u(e_1), ..., u(e_r))$ est donc une base de $\mathrm{Im}(u)$. On pose alors $f_1 = u(e_1)$, $f_2 = u(e_2)$,...,$f_r = u(e_r)$. Enfin, on complète la famille libre $(f_1, ..., f_r)$ en une base $\mathcal{B}_2 = (f_1, ..., f_r, ..., f_n)$ de F. Ce choix de bases va nous donner une matrice de u particulièrement simple. En effet, pour $1 \leq i \leq r$, nous avons $u(e_i) = f_i$. Pour $r + 1 \leq i \leq n$, nous avons $u(e_i) = 0$. On trouve ainsi :

$$\mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(u) = \begin{pmatrix} 1 & 0 & ... & 0 & 0 & ... & 0 \\ 0 & 1 & ... & 0 & 0 & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & ... & 1 & 0 & ... & 0 \\ 0 & 0 & ... & 0 & 0 & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & ... & 0 & 0 & ... & 0 \end{pmatrix}$$

On note cette matrice J_r. Elle se réécrit par blocs $J_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. On obtient donc le théorème suivant :

Théorème 2.4.7 (Théorème du rang en forme matricielle) *1. Toute application linéaire de rang r a pour matrice J_r dans des bases appropriées.*

2. Toute matrice de rang r est équivalente à J_r. Ainsi, deux matrices sont équivalentes si et seulement si elles ont même rang.

Exemple. Les matrices E^{kl} de la base canonique de $\mathcal{M}_{np}(\mathbb{K})$ sont toutes de rang 1. Elle sont donc toutes équivalentes.

Corollaire 2.4.8 (Rang d'une transposée) *Soit A une matrice. On a $\mathrm{rg}(A) = \mathrm{rg}({}^tA)$. Par conséquent, le rang des vecteurs colonnes de A est égal au rang des vecteurs lignes.*

Preuve. Si A est de rang r alors A est équivalente à J_r. Il existe donc des matrices inversibles P et Q telles que $A = QJ_rP$. On obtient alors

$${}^tA = {}^tP\,{}^tJ_r\,{}^tQ = {}^tPJ_r\,{}^tQ$$

Comme ${}^tP$ et ${}^tQ$ sont inversibles, ceci prouve que ${}^tA$ est équivalente à J_r. Elle est donc de rang r.

□

2.5 Dualité

2.5.A Les formes linéaires

Définition 2.5.1 Soit E un $\mathbb{K}$-espace vectoriel. On appelle *forme linéaire* toute application linéaire de E dans $\mathbb{K}$. Il s'agit d'un cas particulier d'application linéaire lorsque l'espace d'arrivée est le corps de base $\mathbb{K}$.

Exemples.

1. Dans $\mathbb{K}[X]$, l'application qui à un polynôme $\sum_{i=0}^{n} a_i X^i$ associe le scalaire a_2 est une forme linéaire.
2. Dans $\mathbb{K}^n$, Pour tout $i \in \{1, ..., n\}$, l'application qui a un vecteur $(x_1, ..., x_n) \in \mathbb{K}^n$ associe x_i est une forme linéaire.
3. Soit I un intervalle contenue dans $\mathbb{R}$ et soit $a \in I$. Dans $E = C(I)$ l'ensemble des applications continues de I dans $\mathbb{R}$, l'application qui à $f \in C(I)$ associe $f(a) \in \mathbb{R}$ est une forme linéaire, appelée évaluation en a.

Définition 2.5.2 Soit E un $\mathbb{K}$ espace vectoriel de dimension n. On appelle *hyperplan* de E tout sous-espace vectoriel de dimension $n-1$ (c'est à dire de codimension 1).

Ces hyperplans sont, après E lui-même, les plus "gros" sous-espace vectoriel possible. Remarquons également que si F est un hyperplan de E et si x est un vecteur de E qui n'est pas dans F alors

$$F \oplus \mathrm{Vect}(x) = E.$$

En effet, comme $x \notin F$, on a $F \cap \mathrm{Vect}(x) = \{0\}$. Il suit que F et $\mathrm{Vect}(x)$ sont en sommes directes. On remarque également que $\dim(F \oplus \mathrm{Vect}(x)) = \dim(F) + \dim(\mathrm{Vect}(x)) = n = \dim(E)$ d'où le résultat.

Théorème 2.5.3 (Noyau d'une forme linéaire) *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et soit ϕ une forme linéaire non nulle. Alors son noyau est un hyperplan de E*

Preuve. Soit $n = \dim(E)$. Comme $\dim(\mathbb{K}) = 1$, on a $\mathrm{rg}(\phi) = 0$ ou 1. Or $\mathrm{rg}(\phi) = 0$ équivaut à dire que $\mathrm{Im}(\phi) = \{0\}$ c'est à dire que ϕ est nul ce qui est exclue. On a donc $\mathrm{rg}(\phi) = 1$. On applique alors le théorème du rang qui nous donne $\mathrm{Ker}(\phi) = \dim(E) - \mathrm{rg}(\phi) = n - 1$ ce qu'il fallait montrer.

□

2.5.B L'espace dual

Définition 2.5.4 On note E^* l'espace $L(E, \mathbb{K})$ de toutes les formes linéaires sur E. E^* est applelée *l'espace dual* de E.

La première remarque à faire ici concerne la dimension de l'espace dual : on sait que $\dim(L(E, \mathbb{K})) = \dim(E)\dim(\mathbb{K}) = \dim(E)$. Donc on a toujours : $\dim(E^*) = \dim(E)$. On a donc relié les dimensions de E avec son dual E^*. Peut-on maintenant établir des liens entre les bases de ces espaces ? soit donc $\mathcal{B} = (e_1, ..., e_n)$ une base de E Pour $1 \leq i \leq n$, on définie l'application

$$\begin{array}{rccc} e_i^* : & E & \to & \mathbb{K} \\ & x = \sum_{j=1}^{n} x_j e_j & \mapsto & x_i \end{array}$$

Il est clair qu'il s'agit d'une forme linéaire.

Pour $1 \leq i, j \leq n$, on note δ_i^j le symbole de Kronecker :

$$\delta_i^j : \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Avec cette notation, on constate que e_i^* est définie par $e_i^*(e_j) = \delta_i^j$.

Exemple. Considérons $\mathbb{R}^3$ et sa base canonique $(e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1))$. Par définition, on sait que

$$e_1^*(e_1) = 1,\ e_1^*(e_2) = 0,\ e_1^*(e_3) = 0$$

On obtient alors pour tout $(x, y, z) \in \mathbb{R}^3$:

$$\begin{array}{ll} e_1^*(x,y,z) & = e_1^*(x(1,0,0) + y(0,1,0) + z(0,0,1)) \\ & = xe_1^*(e_1) + ye_1^*(e_2) + ze_1^*(e_3) \\ & = x \end{array}$$

On voit de même que $e_2^*(x, y, z) = y$ et $e_3^*(x, y, z) = z$. Il est clair que (e_1^*, e_2^*, e_3^*) forme une base de E^*. On va voir que ceci reste vrai dans le cadre général.

Théorème 2.5.5 (Base duale) *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie muni d'une base $\mathcal{B} = (e_1, ..., e_n)$. Alors la famille $(e_1^*, ..., e_n^*)$ forme une base de E^* appelée base duale.*

Preuve.

– On montre tout d'abord que la famille $(e_1^*, ..., e_n^*)$ est libre. Soit $(\lambda_1, ..., \lambda_n) \in \mathbb{K}^n$ tels que :

$$\lambda_1 e_1^* + ... + \lambda_n e_n^* = 0$$

Ceci signifie que pour tout $x \in E$, on a :

$$\lambda_1 e_1^*(x) + ... + \lambda_n e_n^* = 0$$

Soit $j \in \{1, ..., n\}$ et choisissons $x = e_j$, on obtient $\lambda_j = 0$ puisque par définition $e_i^*(e_j) = \delta_i^j$. Ceci prouve que cette famile est libre.

– Comme $\dim(E^*) = \dim(E) = n$ et comme on a trouvé une famille libre dans E^* de n éléments. Ceci implique que c'est une base.

□

Gardons les notations précédentes et donnons nous une forme linéaire $\phi \in E^*$. Soit $x = \sum_{i=1}^{n} x_i e_i$ un vecteur de E alors :

$$\phi(x) = \sum_{i=1}^{n} x_i \phi(e_i) = \sum_{i=1}^{n} \phi(e_i) e_i^*(x).$$

On a donc :

$$\phi = \sum_{i=1}^{n} \phi(e_i) e_i^*.$$

C'est tout simplement la décomposition de ϕ comme combinaison linéaire de la base duale !

Exemples.

1. Soit $f : \mathbb{R}^3 \to \mathbb{R}$ définie par $f(x_1, x_2, x_3) = x_1 + 2x_2 - x_3$. f est une forme linéaire qui s'écrit dans la base (e_1^*, e_2^*, e_3^*) duale de la base canonique
$$f = e_1^* + 2e_2^* - e_3^*$$
2. On considère l'espace $\mathbb{R}_3[X]$ muni de la base formée des quatre vecteurs $e_0 = 1$, $e_1 = X$, $e_3 = X^2$ et $e_4 = X^3$. Soit $j \in \{0, 1, 2, 3\}$ et $P(X) = \sum_{k=0}^{3} a_k X^k$, on a
$$e_j^*(P) = a_j$$
On considère la forme linéaire :
$$\begin{array}{rccc} \theta : & \mathbb{R}_3[X] & \to & \mathbb{R} \\ & P & \mapsto & P(2) \end{array}$$
Soit $P(X) = \sum_{k=0}^{3} a_k X^k$. On a alors
$$\theta(P) = P(2) = \sum_{k=0}^{3} a_k 2^k = (e_0^* + 2e_1^* + 4e_2^* + 8e_3^*)(P)$$
Donc on obtient
$$\theta = e_0^* + 2e_1^* + 4e_2^* + 8e_3^*$$

Chapitre 3

Endomorphismes

3.1 La structure d'algèbre

3.1.A Définition d'une algèbre

En plus de pouvoir additionner des endomorphismes et de les multiplier par un scalaire, on peut aussi les composer (ce qui n'est pas possible pour des applications de E dans F avec $E \neq F$). On a donc une troisième loi sur l'ensemble $L(E)$. C'est la structure d'algèbre donnée ci-dessous qui rend compte de ses propriétés.

Définition 3.1.1 Une $\mathbb{K}$-algèbre est un ensemble $\mathcal{H}$ munis de trois lois :

1. Une addition notée $+$,
2. Une multiplication externe par des scalaires de $\mathbb{K}$.
3. Une multiplication interne notée . ou sans symbole.

Ces trois lois doivent vérifier les propriétés suivantes.

1. L'addition et la multiplication externe font de $\mathcal{H}$ un $\mathbb{K}$-espace vectoriel.
2. La multiplication interne possède un élément neutre (encore appelé unité) que l'on note $1_{\mathcal{H}}$ tel que $1_{\mathcal{H}} \neq 0$
3. La multiplication interne est associative :

$$\forall (a,b,c) \in \mathcal{H}^3,\ a(bc) = (ab)c$$

   On note alors ce produit abc.
4. La multiplication interne est distributive par rapport à l'addition c'est à dire :

$$\forall (a,b,c) \in \mathcal{H}^3,\ a(b+c) = ab + bc \text{ et } (a+b)c = ac + bc$$

5. Les multiplications internes et externes commutent :

$$`\forall (a,b) \in \mathcal{H}^2,\ \forall \lambda \in \mathbb{K},\ \lambda(ab) = (\lambda a)b = a(\lambda b)$$

   On note tout simplement λab ce produit.

Les propriétés 2 et 4 traduisent le fait que $\mathcal{H}$ est un *anneau*. Notons que l'on ne suppose pas que le produit est commutatif.

Exemples.

1. Le corps $\mathbb{K}$ est une $\mathbb{K}$-algèbre.
2. L'ensemble $\mathbb{K}[X]$ des polynômes est aussi une $\mathbb{K}$-algèbre (l'unité est le polynôme constant égal a 1.)
3. L'ensemble $\mathcal{M}_n(\mathbb{K})$ des matrices à n lignes et n colonnes à coefficients dans un corps $\mathbb{K}$ est une $\mathbb{K}$-algèbre, l'élément unité est la matrice identité.
4. Si E est un $\mathbb{K}$ espace vectoriel, l'ensemble $L(E)$ des endomorphismes de E est une $\mathbb{K}$-algèbre. La multiplication est donnée par la composition : on notera donc pour u et v dans $L(E)$: $uv = u \circ v$. On notera aussi u^k pour la composition de u par lui-même k fois. L'unité est l'application identité.

Remarquons que dans notre définition d'algèbre, on ne suppose pas que tout élément non nul est inversible (Un élément $a \in \mathcal{H}$ est dit inversible si il existe $b \in \mathcal{H}$ tel que $ab = ba = 1$, on note alors $b = a^{-1}$.) Dans l'algèbre $\mathcal{M}_n(\mathbb{K})$, par exemple, on a des éléments non nul et non inversible.

L'ensemble des éléments inversibles de $\mathcal{M}_n(\mathbb{K})$ est appellée le groupe linéaire d'ordre n et il est noté $GL_n(\mathbb{K})$

L'ensemble des éléments inversibles de $L(E)$ est noté $GL(E)$, ce sont les automorphismes de E (c'est à dire les endomorphismes bijectifs).

3.1.B Polynômes dans une algèbre

Dans une $\mathbb{K}$-algèbre $\mathcal{H}$, étant donné un élément a, on peut calculer l'élément a^k et multiplier cet élément par un scalaire de $\mathbb{K}$: on peut donc calculer des polynômes en a.

Définition 3.1.2 Soit $\mathcal{H}$ une $\mathbb{K}$-algèbre, soit a un élément de $\mathcal{H}$ et soit $P = \sum_{k=0}^{n} \lambda_k X^k$ un polynôme à coefficients dans $\mathbb{K}$. On note $P(a)$ l'élément de $\mathcal{H}$ défini par :

$$P(a) = \lambda_n a^n + \lambda_{n-1} a^{n-1} + ... + \lambda_1 a + \lambda_0 1_{\mathcal{H}}$$

On se méfiera particulièrement dans cette écriture d'une erreur fréquente dans le dernier terme concernant l'unité $1_{\mathcal{H}}$ (qui correspond au terme de degré 0). Dans $L(E)$ cette unité est id_E tandis que dans $\mathcal{M}_n(\mathbb{K})$ l'unité est I_n.

Exemple. Considérons l'endomorphisme $u : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (x,y) & \longmapsto & (x+y, x-y) \end{array} \right.$

Calculons le carré de u. Soit $(x,y) \in \mathbb{R}^2$ un vecteur quelconque. On a

$$\begin{aligned} u^2(x,y) &= u\big(u(x,y)\big) \\ &= u(x+y, x-y) \\ &= \big((x+y)+(x-y), (x+y)-(x-y)\big) \\ &= (2x, 2y). \end{aligned}$$

Ceci prouve que $u^2 = 2\,\mathrm{id}_E$. Considérons maintenant le polynôme suivant

$$P(X) = X^2 - 3X + 4.$$

On a :

$$P(u) = u^2 - 3u + 4\,\mathrm{id}_E = 2\,\mathrm{id}_E - 3u + 4\,\mathrm{id}_E = 6\,\mathrm{id}_E - 3u.$$

Si $(x,y) \in \mathbb{R}^2$ est un vecteur quelconque, on obtient donc :

$$P(u)(x,y) = 6(x,y) - 3u(x,y) = \big(6x - 3(x+y), 6y - 3(x-y)\big) = (3x - 3y, -3x + 9y).$$

Les règles de calcul dans une algèbre (ou plus génèralement dans un anneau) sont les mêmes que dans $\mathbb{R}$ à trois exceptions notables près sur lesquelles il convient d'insister :

1. Tous les éléments non nuls ne sont pas forcément inversibles. Par exemple, dans $L(E)$, "nul" signifie "endomorphisme nul" tandis que "inversible" signifie "endomorphisme bijectif" et nous savons qu'il existe des endomorphismes non identiquement nuls et non bijectifs.
2. Le fait qu'un produit ab soit nul n'implique absolument pas que a ou b soit nul. Prenons un exemple dans $\mathcal{M}_2(\mathbb{K})$. Considérons $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Si vous calculez A^2, vous trouverez bien la matrice nulle et pourtant A est non nul !
3. Tous les éléments ne commutent pas. On ne peut pas remplacer ab par ba (sauf lorsqu'on est justement dans le cas particulier où a et b commutent). Par exemple, calculons $(a+b)^2$. Il vient : $(a+b)^2 = a^2 + ab + ba + b^2$ qui ne se simplifie pas comme dans l'identité remarquable suivante valable dans $\mathbb{R}$: $(x+y)^2 = x^2 + 2xy + y^2$.

3.1.C Eléments nilpotents

Définition 3.1.3 Soit $\mathcal{A}$ une $\mathbb{K}$-algèbre. Un élément $a \in \mathcal{A}$ est dit *nilpotent* lorsqu'il existe un entier p tel que $a^p = 0$. Le plus petit des entiers p tels que $a^p = 0$ s'appelle *l'indice de nilpotence* de a.

Il est clair que 0 est un élément nilpotent. Son indice de nilpotence vaut 1. Dans $\mathcal{M}_2(\mathbb{K})$, nous avons vu précédemment que la matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ était nilpotente, d'indice 2.

3.2 Propriétés de $\mathcal{M}_n(\mathbb{K})$

3.2.A Trace d'une matrice et d'un endomorphisme

Définition 3.2.1 On appelle *trace* d'une matrice $A = (a_{ij})$ le nombre $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$. C'est la somme des termes diagonaux.

Par exemple, $\mathrm{tr}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 2 & 6 \\ 7 & 8 & 4 \end{pmatrix} = 1 + 2 + 4 = 7$. On peut aussi remarquer que $\mathrm{tr}(I_n) = n$.

Les principales propriétés de la trace sont regroupées dans la proposition suivante :

Proposition 3.2.2 *1. La trace est une forme linéaire, c'est-à-dire que, pour toutes matrices $A, B \in \mathcal{M}_n(\mathbb{K})$ et tout scalaire $\lambda \in \mathbb{K}$, on a $\mathrm{tr}(\lambda A) = \lambda\,\mathrm{tr}(A)$ et $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.*

2. La trace d'un produit ne dépend pas de l'ordre : pour toutes matrices $A, B \in \mathcal{M}_n(\mathbb{K})$, $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

3. Deux matrices semblables ont la même trace. On dit que la trace est un invariant de similitude.

Preuve.

1. C'est évident à vérifier.
2. Posons $A = (a_{ij})$, $B = (b_{ij})$, $C = AB = (c_{ij})$ et $D = BA = (d_{ij})$. Nous avons
$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} \text{ et } d_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj}.$$
Par conséquent, la trace vaut :
$$\mathrm{tr}(AB) = \sum_{i=1}^{n} c_{ii} = \sum_{i,k} a_{ik}b_{ki} \text{ et } \mathrm{tr}(BA) = \sum_{i=1}^{n} d_{ii} = \sum_{i,k} b_{ik}a_{ki}.$$
En intervertissant les rôles joués par les indices muets i et k dans ces deux sommes, on voit que $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.
3. Si A et B sont deux matrices semblables, alors il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$. Alors
$$\mathrm{tr}(B) = \mathrm{tr}(P^{-1}AP) = \mathrm{tr}\left(P^{-1}(AP)\right) = \mathrm{tr}\left((AP)P^{-1}\right) = \mathrm{tr}(A).$$

□

Il résulte de la propriété 3 que, si $u \in L(E)$, alors la trace de toutes les matrices de u dans n'importe quelle base est toujours la même. Cette quantité est appelée *la trace de l'endomorphisme u* et notée $\mathrm{tr}(u)$.

3.2.B Produit par blocs

On définit deux matrices par blocs $M = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$ et $M' = \left(\begin{array}{c|c} A' & B' \\ \hline C' & D' \end{array}\right)$ dans $\mathcal{M}_n$. Alors le produit par blocs se calcule comme pour un produit normal, c'est-à-dire :

$$MM' = \left(\begin{array}{c|c} AA' + BC' & AB' + BD' \\ \hline CA' + DC' & CB' + DD' \end{array}\right).$$

On prendra garde tout de même à n'écrire que des matrices multipliables entre elles et on fera attention à ne pas changer l'ordre des matrices puisqu'elles ne commutent pas forcément. En particulier, les matrices diagonales par blocs se multiplient facilement :

$$\left(\begin{array}{c|c} A & 0 \\ \hline 0 & D \end{array}\right)\left(\begin{array}{c|c} A' & 0 \\ \hline 0 & D' \end{array}\right) = \left(\begin{array}{c|c} AA' & 0 \\ \hline 0 & DD' \end{array}\right).$$

De plus, la matrice $\left(\begin{array}{c|c} A & 0 \\ \hline 0 & D \end{array}\right)$ est inversible si et seulement si A et D le sont. Dans ce cas, on a alors

$$\left(\begin{array}{c|c} A & 0 \\ \hline 0 & D \end{array}\right)^{-1} = \left(\begin{array}{c|c} A^{-1} & 0 \\ \hline 0 & D^{-1} \end{array}\right).$$

3.3 Réduction de quelques endomorphismes remarquables

Nous avons vu dans le chapitre précédent que, étant donné un espace vectoriel E muni d'une base $\mathcal{B}$, on pouvait associer à tout endomorphisme $u \in L(E)$ une matrice $M = \mathrm{Mat}_{\mathcal{B}}(u)$. Cette matrice dépend de la base $\mathcal{B}$ choisie. Les autres matrices de u calculées dans d'autres bases sont les matrices semblables à M c'est-à-dire les matrices de la forme $P^{-1}MP$ avec $P \in \mathrm{GL}_n(E)$. Dans cette partie, nous allons chercher, pour certains endomorphismes $u \in L(E)$ particuliers, une base $\mathcal{B}$ dans laquelle la matrice $\mathrm{Mat}_{\mathcal{B}}u$ est le plus simple possible. Dans toute cette partie, E est un $\mathbb{K}$-espace vectoriel fixé de dimension n.

3.3.A Projecteurs

Définition 3.3.1 Un endomorphisme $u \in L(E)$ est dit *idempotent* s'il vérifie $u^2 = u$.

Les endomorphismes idempotents sont ceux qui, quand on réitère leur action, ne font rien de plus que la première fois. Par exemple, 0 et id_E sont deux endomorphismes idempotents. Dans le plan, la projection sur l'axe des

abscisses parallèlement à l'axe des ordonnées correspond à un idempotent. En effet, si on commence par projeter un vecteur z, on obtient un vecteur x situé sur l'axe des abscisses et donc, projeter à son tour x redonne x. Plus généralement, la notion d'endomorphisme idempotent coincide exactement avec celle de projecteur.

Théorème 3.3.2 (Projecteurs et idempotents) *Un endomorphisme $u \in L(E)$ est idempotent si et seulement si c'est un projecteur.*

Preuve. Soit F, G deux sous-espaces supplémentaires dans E. Soit p le projecteur sur F parallèlement à G. Soit $x = x_F + x_G$ un élément quelconque de E. Nous avons $p(x) = x_F$. De plus, le projeté de x_F sur F parallèlement à G est x_F lui-même puisque x_F est déjà dans F. Par conséquent, nous avons $p(x_F) = x_F$ d'où $p^2(x) = p(p(x)) = p(x_F) = x_F = p(x)$. Ceci prouve que $p^2 = p$ et que p est un endomorphisme idempotent.

Soit maintenant u un endomorphisme idempotent. Posons $F = \mathrm{Im}(u)$ et $G = \mathrm{Ker}(u)$. Soit $y \in F$ quelconque. Puisque F est l'image de u, il existe $x \in E$ tel que $u(x) = y$. Alors $u(y) = u(u(x)) = u^2(x) = u(x) = y$. Ceci prouve que u induit l'identité sur F.

Montrons que $E = F \oplus G$. Si $x \in F \cap G$, alors $u(x) = x = 0$ puisque u induit l'identité sur F et s'annule sur G. Ceci prouve que $F \cap G = \{0\}$ et donc la somme est directe. Soit maintenant $x \in E$ quelconque. Nous posons $y = u(x) \in F$ et $z = x - u(x)$. Nous avons $u(z) = u(x) - u^2(x) = u(x) - u(x) = 0$ donc $z \in G$. Comme il est clair que $y + z = x$, nous voyons que tout vecteur x de E admet une décomposition selon la somme $F + G$ et donc finalement que $E = F \oplus G$. Les sous-espaces vectoriels F et G sont supplémentaires dans E.

Nous pouvons alors considérer le projecteur p sur F parallèlement à G. Pour tout $x \in E$, nous avons la décomposition $x = x_F + x_G$. Nous avons

$$u(x) = u(x_F + x_G) = u(x_F) + u(x_G) = u(x_F) = x_F.$$

d'où $p(x) = u(x)$. Cela prouve que p et u sont égaux et donc que u est un projecteur.

□

On peut maintenant se demander quel est la forme des matrices associées à un projecteur fixé et en particulier si il existe une base dans laquel cette matrice est particulièrement simple.

Théorème 3.3.3 (Matrice d'un projecteur) *Soit p un projecteur de rang r sur un $\mathbb{K}$ espace vectoriel E de dimension n. Nous avons $E = \mathrm{Im}(p) \oplus \mathrm{Ker}(p)$. Si on recolle une base de $\mathrm{Im}(u)$ et une base de $\mathrm{Ker}(u)$, on obtient une base $\mathcal{B}$ dans laquelle la matrice de p est égale à J_r où on rappelle que*

$$J_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Preuve. Soit F l'image de p et G son noyau. La démonstration du théorème 3.3.2 prouve que $E = F \oplus G$.[1] De plus $\dim(F) = \mathrm{rg}(p) = r$. Nous pouvons donc prendre une base $(e_1, \ldots, e_r)$ de F et une base $(e_{r+1}, \ldots, e_n)$ de G. Leur recollement $(e_1, \ldots, e_n)$ forme une base de E. Comme $p(e_1) = e_1$, ..., $p(e_r) = e_r$ et $p(e_{r+1}) = \cdots = p(e_n) = 0$, la matrice de p dans la base $\mathcal{B}$ est bien égale à J_r.

□

Un commentaire s'impose. Nous avions vu au théorème 2.4.7 (forme matricielle du théorème du rang) que toute application linéaire avait pour matrice J_r dans des bases convenables. On peut donc se demander l'intérêt de ce nouveau résultat qui a l'air de redire la même chose dans le cas particulier des projecteurs. Il n'en est rien. En effet, dans le théorème 2.4.7, on n'impose pas que la base de départ et la base d'arrivée soient les mêmes. Dans le résultat ci-dessus, on l'impose. La comparaison des deux théorèmes est peut-être plus parlante matriciellement :

- N'importe quelle matrice M de rang r est équivalente à J_r.
- Parmi les matrices M de rang r, seules celles qui vérifient $M^2 = M$ sont semblables à J_r.

Exemple La trace d'un projecteur est égale à son rang. En effet, si p est un projecteur de rang r, nous savons qu'il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(p) = J_r$. Alors $\mathrm{tr}(p) = \mathrm{tr}(J_r) = r = \mathrm{rg}(p)$.

3.3.B Symétries

Dans ce paragraphe, on suppose que $\mathbb{K}$ est égal à $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$.[2]

Définition 3.3.4 On dit qu'un endomorphisme $u \in L(E)$ est une *involution* ou que u est *involutif* lorsque $u^2 = \mathrm{id}_E$. Cela signifie que u est bijectif et égal à son inverse : $u = u^{-1}$.

Les involutions sont les endomorphismes qui "changent les points deux par deux" : en effet, si $u(x) = y$ alors $u(y) = x$. Par exemple, l'application $x \mapsto -x$ est une involution. De même que nous avions caractérisé les endomorphismes idempotents comme des projecteurs, nous allons caractériser les endomorphismes involutifs comme des symétries.

Définition 3.3.5 Soient F et G deux sous-espaces supplémentaires dans E. Etant donné un vecteur x, nous _'ecrivons sa décomposition selon la somme directe $F \oplus G : x = x_F + x_G$ avec $x_F \in F$ et $x_G \in G$. Le *symétrique* de x par

[1] Attention, cette égalité n'est vraie que parce que p est un projecteur ; elle est fausse en général pour un endomorphisme quelconque.

[2] Comme vous le verrez dans la suite, nous aurons besoin à un moment de dire que $x = -x$ entraine $x = 0$, ce qui est vrai dans $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$ mais pas dans un corps quelconque. Par exemple, dans $\mathbb{Z}/2\mathbb{Z}$, nous avons $1 = -1$.

rapport à F parallèlement à G est le point $s(x) = x_F - x_G$. L'application s ainsi définie est nommée *symétrie par rapport à F parallèlement à G.*

Proposition 3.3.6 *Les symétries sont des applications linéaires.*

Preuve. Soit F et G deux sous-espaces vectoriels tels que $E = F \oplus G$. Soit p le projecteur sur F parallèlement à G et q le projecteur sur G parallèlement à F. Nous avons $x_F = p(x)$ et $x_G = q(x)$. Par conséquent, la symétrie par rapport à F parallèlement à G est $s = p - q$. Comme nous savons que p et q sont linéaires, il en est de même de s.

□

Théorème 3.3.7 *Un endomorphisme $u \in L(E)$ est involutif si et seulement si c'est une symétrie.*

Preuve. Supposons que s soit la symétrie par rapport à F parallèlement à G. Nous avons $s(x) = x_F - x_G$ et donc $s^2(x) = s(x_F - x_G) = x_F - (-x_G) = x$. Donc $s^2 = \mathrm{id}_E$ et s est involutif.

Supposons que u soit involutif. Posons $F = \mathrm{Ker}(u - \mathrm{id}_E)$ et $G = \mathrm{Ker}(u + \mathrm{id}_E)$. Ce sont deux sous-espaces vectoriels de E. F est l'ensemble des points fixes de u tandis que G est l'ensemble des points de E que u envoie sur leur opposé. Si $x \in F \cap G$, alors $u(x) = x$ et $u(x) = -x$. Alors $x = -x$ et donc $x = 0$. Par conséquent, $F \cap G = \{0\}$.

Soit $x \in E$. Posons $y = \dfrac{x + u(x)}{2}$ et $z = \dfrac{x - u(x)}{2}$. On a $u(y) = \dfrac{u(x) + u^2(x)}{2} = \dfrac{u(x) + x}{2} = y$ puisque u est involutif. De même, $u(z) = \dfrac{u(x) - u^2(x)}{2} = \dfrac{u(x) - x}{2} = -z$. Par conséquent, $y \in F$ et $z \in G$. Comme $y + z = x$, nous en déduisons que $E = F + G$.

Ces deux informations ensemble prouvent que $E = F \oplus G$. Soit donc maintenant s la symétrie sur F parallèlement à G. Nous avons, pour tout $x \in F$, $s(x) = x = u(x)$ et, pour tout $x \in G$, $s(x) = -x = u(x)$. Par conséquent, les deux applications linéaires coincident sur deux supplémentaires donc partout : $s = u$ et u est une symétrie.

□

Théorème 3.3.8 (Matrice d'une symétrie) *Soit s la symétrie par rapport à un sous-espace F parallèlement à un sous-espace G. Si on recolle une base de F et une base de G, la matrice de s dans la base ainsi obtenue est de la forme (par blocs)*

$$\begin{pmatrix} I_{\dim(F)} & 0 \\ 0 & -I_{\dim(G)} \end{pmatrix}$$

Preuve. Soit $(e_1, \dots, e_p)$ une base de F et $(e_{p+1}, \dots, e_n)$ une base de G. Leur recollement forme une base de E puisque $E = F \oplus G$. On a $s(e_1) = e_1, \dots, s(e_p) = e_p$, et $s(e_{p+1}) = -e_{p+1}, \dots, s(e_n) = -e_n$ donc la matrice de s dans la base $\mathcal{B}$ a bien la forme annoncée.

□

Exemple. Considérons l'application

$$\begin{array}{rccc} s: & \mathbb{R}^3 & \to & \mathbb{R}^3 \\ & (x,y,z) & \to & (x,y,-z) \end{array}$$

Cette application correspond à la symétrie par rapport au sous-espace $F := \{(x,y,0) \mid (x,y) \in \mathbb{R}^2\}$ parrallèlement au sous-espace $G := \{(0,0,z) \mid z \in \mathbb{R}\}$. Une base de F est donné par les vecteurs $(1,0,0)$ et $(0,1,0)$ et une base de G par le vecteur $(0,0,1)$. Dans le recollement $\mathcal{B}$ de ces 2 bases, c'est à dire dans la base canonique ici, la matrice de u obtenue est

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

3.3.C Homothéties

Définition 3.3.9 On appelle *homothétie* tout endomorphisme de E de la forme $\lambda \operatorname{id}_E$ où $\lambda \in \mathbb{K}$.

On peut remarqué que l'homothétie $h = \lambda \operatorname{id}_E$ est inversible si et seulement si $\lambda \neq 0$, auquel cas, $h^{-1} = \dfrac{1}{\lambda} \operatorname{id}_E$. Dans n'importe quelle base la matrice de $h = \lambda \operatorname{id}_E$ est évidemment

$$\lambda I_n = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ 0 & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda \end{pmatrix}.$$

Remarquons que les homothéties commutent avec tous les endomorphismes. En effet, si $h = \lambda \operatorname{id}_E$ et que $u \in L(E)$ est quelconque, alors $uh = hu = \lambda u$.

3.3.D Rotations

Dans le plan $\mathbb{R}^2$, on considère la rotation r d'angle θ (et de centre 0, ce qui est toujours sous-entendu puisque nous parlons ici de rotation vectorielle). Soit (e_1, e_2) la base canonique.

Nous avons $r(e_1) = (\cos\theta, \sin\theta)$ et $r(e_2) = (-\sin\theta, \cos\theta)$. Par conséquent, la matrice de la rotation r dans la base canonique est la matrice suivante

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Une telle matrice est naturellement appelée *matrice de rotation*. Une simple remarque sur les rotations permet bien souvent de simplifier les calculs. La composée de la rotation d'angle θ et de la rotation d'angle α est la rotation d'angle $\theta+\alpha$. Par conséquent, lorsqu'on veut calculer le produit de deux matrices de rotation, il est inutile de faire le calcul, il suffit d'ajouter les angles : $R_\theta R_\alpha = R_{\theta+\alpha} = \begin{pmatrix} \cos(\theta+\alpha) & -\sin(\theta+\alpha) \\ \sin(\theta+\alpha) & \cos(\theta+\alpha) \end{pmatrix}$. De même pour calculer l'inverse d'une matrice de rotation, tout calcul est inutile, il suffit de dire qu'il s'agit de la rotation d'angle opposé : $R_\theta^{-1} = R_{-\theta} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$.

3.3.E Endomorphismes nilpotents

Théorème 3.3.10 (Matrice d'un nilpotent) *1. Un endomorphisme $u \in L(E)$ est dit* nilpotent *si et seulement si il existe une base $\mathcal{B}$ telle que la matrice de u dans la base $\mathcal{B}$ soit triangulaire supèrieure stricte, c'est-à-dire triangulaire supérieure avec des 0 sur la diagonale*[3]*.*

$$\mathrm{Mat}_{\mathcal{B}}\, u = \begin{pmatrix} 0 & * & * \\ \vdots & \ddots & * \\ 0 & \cdots & 0 \end{pmatrix}$$

2. Une matrice M est nilpotente si et seulement si elle est semblable à une matrice triangulaire supérieure stricte.

Preuve. Nous commencons par remarquer que la deuxième assertion ci-dessus n'est qu'une reformulation de la première en termes de matrices.

Supposons tout d'abord qu'il existe une base $\mathcal{B} = (e_1, \dots, e_n)$ dans laquelle la matrice de u soit triangulaire supèrieure stricte. Nous prouvons alors par récurrence sur k que $u^k(e_k) = 0$. Tout d'abord, $e_1 \in \mathrm{Ker}(u)$ donc $u(e_1) = 0$ et la récurrence est initialisée. Ensuite, supposant l'hypothèse vérifiée jusqu'au rang k, nous constatons, d'après la forme de la matrice que $u(e_{k+1}) \in \mathrm{Vect}(e_1, \dots, e_k)$. Par conséquent, il vient $u^{k+1}(e_{k+1}) \in \mathrm{Vect}(u^k(e_1), \dots, u^k(e_k)) = \mathrm{Vect}(0, \dots, 0) = \{0\}$ donc $u^{k+1}(e_{k+1}) = 0$. L'hypothèe de récurrence est ainsi établie au rang $k+1$.

Finalement, nous avons donc $u^n(e_k) = 0$ pour tout $k \leq n$ donc u^n s'annule sur une base d'où $u^n = 0$. Ceci prouve que u est nilpotent.

Supposons maintenant que u est nilpotent. Nous montrons le théorème par récurrence sur n.

- Si $n = 1$, les matrices sont en fait des nombres et le seul élément nilpotent est 0.
- Nous supposons le résultat démontré au rang $n-1$, nous montrons qu'il reste vrai au rang n. Tout d'abord, remarquons qu'il existe un

[3] Les étoiles signifient qu'il y a des coefficients quelconques.

vecteur non nul e_1 de E tel que $u(e_1) = 0$. En effet, soit p l'indice de nilpotence de u. Comme $u^{p-1} \neq 0$, il existe un vecteur $x \in E$ tel que $u^{p-1}(x) \neq 0$. Nous posons $e_1 = u^{p-1}(x)$. Alors $u(e_1) = u^p(x) = 0$ puisque u^p est l'endomorphisme nul.
Nous complétons e_1 en une base $(e_1, \ldots, e_n)$ de E. La matrice de u dans la base obtenue est alors égale à $M = \left(\begin{array}{c|ccc} 0 & * & \ldots & * \\ \hline 0 & & & \\ \vdots & & A & \\ 0 & & & \end{array}\right)$ où les $*$ désignent des termes quelconques que nous ne précisons pas. La matrice A est de taille $n-1$. Comme $u^p = 0$, nous avons $M^p = 0$ et par conséquent, $A^p = 0$. La matrice A est donc nilpotente. Comme elle est de taille $n-1$, on peut appliquer l'hypothèse de récurrence : il existe $Q \in \mathrm{GL}_{n-1}(\mathbb{K})$ et $T \in \mathcal{M}_{n-1}(\mathbb{K})$ triangulaire supèrieure stricte telles que $T = Q^{-1}AQ$. La matrice $P = \left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right)$ est inversible d'inverse $P^{-1} = \left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{array}\right)$. La matrice M est donc semblable à la matrice $N = P^{-1}MP = \left(\begin{array}{c|ccc} 0 & * & \ldots & * \\ \hline 0 & & & \\ \vdots & & Q^{-1}AQ & \\ 0 & & & \end{array}\right) = \left(\begin{array}{c|ccc} 0 & * & \ldots & * \\ \hline 0 & & & \\ \vdots & & T & \\ 0 & & & \end{array}\right)$ qui est triangulaire supèrieure stricte. Comme N est semblable à M, c'est aussi une matrice de u dans une certaine base.

□

Chapitre 4

Le groupe symétrique

Le but de ce chapitre est d'introduire un objet nécessaire à la définition du déterminenat : le groupe symétrique. Les preuves des résultats de ce chapitre sont assez délicates. Elles sont ici pour vous aider à comprendre comment fonctionnent les objets introduits. Il n'est pas souhaitable d'y consacrer trop de temps au détriment des autres parties du cours. Une étude plus approfondie de cet ensemble sera mené en troisième année dans le cours de Théorie des groupes.

4.1 Permutations, transpositions

Définition 4.1.1 Soit n un entier naturel supérieur ou égal à 2.

1. On appelle *permutation* de $\{1, \ldots, n\}$ toute bijection de $\{1, \ldots, n\}$ dans lui-même.
2. On note $\mathfrak{S}_n$ et on appelle *groupe symétrique* [1] l'ensemble des permutations de $\{1, \ldots, n\}$. On rappelle qu'il existe $n!$ permutations différentes de $\{1, \ldots, n\}$ donc $\text{Card}(\mathfrak{S}_n) = n!$.
3. Soit $i \neq j$ dans $\{1, \ldots, n\}$. On note $(i\ j)$ et on appelle *transposition de i et j* la permutation de $\{1, \ldots, n\}$ qui échange i et j et laisse fixes tous les autres nombres. En d'autres termes, $(i\ j)$ est la permutation τ définie par :

$$\text{Pour tout } 1 \leq p \leq n, \quad \tau(p) = \left| \begin{array}{ll} p & \text{si } p \neq i \text{ et } p \neq j \\ j & \text{si } p = i \\ i & \text{si } p = j \end{array} \right.$$

Exemple. Pour $n = 4$, la transposition $\tau = (1\ 3)$ est définie par $\tau(1) = 3$, $\tau(2) = 2$, $\tau(3) = 1$ et $\tau(4) = 4$

Comme une permutation de $\{1, \ldots, n\}$ est entièrement déterminé par les valeurs $\sigma(1)$, $\sigma(2)$, ..., $\sigma(n)$, on notera parfois cette permutation :

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

On peut écrire la composition $\sigma' \circ \sigma$ de deux permutations σ et σ' de $\{1, \ldots, n\}$. C'est encore une permutation de $\{1, \ldots, n\}$. Enfin, on peut définir la réciproque σ^{-1} d'une permutation de $\{1, \ldots, n\}$. La loi de composition fait de $\mathfrak{S}_n$ un groupe dont l'élément neutre est l'identité id. On parlera donc indifféremment de "produit de permutations" ou de "composition de permutations" et on écrira $\sigma'\sigma$ à la place de $\sigma' \circ \sigma$. Remarquons que l'inverse d'une transposition est elle-même : $(i\ j)^{-1} = (i\ j)$.

Théorème 4.1.2 (Décomposition en produit de transpositions) *Toute permutation de $\{1, \ldots, n\}$ peut s'écrire comme un produit de transpositions. En d'autres termes, pour tout $\sigma \in \mathfrak{S}_n$, il existe des transpositions $\tau_1, \ldots, \tau_k$ telles que $\sigma = \tau_1 \tau_2 \ldots \tau_k$.*

Preuve. Nous démontrons ce résultat par récurrence sur n.

- Montrons que le résultat est vrai pour $n = 2$. Les deux permutations possibles de $\{1, 2\}$ sont l'identité et la transposition (1 2). L'identité est égale au produit vide, ou si vous préférez, elle est aussi égale à (1 2)(1 2). C'est donc un produit de transpositions. La transposition (1 2) est bien sûr un produit de transpositions.
- Supposons le résultat vrai pour $n - 1$ et montrons-le au rang n. Soit donc $\sigma \in \mathfrak{S}_n$ une permutation de $\{1, \ldots, n\}$. Nous allons distinguer deux cas dans notre preuve.

  Premier cas : Supposons que $\sigma(n) = n$. Dans ce cas, la permutation envoie $\{1, \ldots, n-1\}$ sur $\{1, \ldots, n-1\}$. Elle induit donc une permutation de $\{1, \ldots, n-1\}$. Par hypothèse de récurrence, elle peut s'écrire comme un produit de transpositions.

  Second cas : Supposons que $\sigma(n) = p < n$. Nous considérons alors la transposition $(p\ n)$. L'image de n par σ est p et l'image de p par $(p\ n)$ est n. Par conséquent, l'image de n par la composée $\sigma' = (p\ n)\sigma$ est égale à n. Comme $\sigma'(n) = n$, le premier cas s'applique et nous voyons que σ' est un produit de transpositions $\sigma' = \tau_1 \ldots \tau_k$ avec $\tau_1, \ldots, \tau_k$ des transpositions. Mais alors, puisque $(p\ n) = (p\ n)^{-1}$, nous trouvons

  $$\sigma = (p\ n)^{-1}\sigma' = (p\ n)\tau_1 \ldots \tau_k.$$

  On a bien réussi à écrire σ comme un produit de transpositions ce qui prouve le résultat.

□

Montrons sur un exemple comment d{eterminer la décomposition d'une permutation en produit de transpositions.

Exemple. Dans $\mathfrak{S}_6$, considérons la permutation σ définie par

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 1 & 6 & 5 \end{pmatrix}$$

On a donc $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 4$, $\sigma(4) = 1$, $\sigma(5) = 6$ et $\sigma(6) = 5$.

On a 1 qui est envoyé sur 2, qui est envoyé sur 3, qui est envoyé sur 4 qui lui même est envoyé sur 1. On retombe donc ici sur le terme de départ : 1. On considère alors le produit de transposition (1 4)(1 3)(1 2).

Ensuite 5 qui est envoyé sur 6, qui est envoyé sur 5 : on retombe donc ici sur le terme de départ : 5. On considère alors le produit de transposition (5 6). Le produit obtenu est alors :

$$(1\ 4)(1\ 3)(1\ 2)(5\ 6)$$

On peut vérifier le résultat en montrant que l'image de chaque $i \in \{1, ..., 6\}$ par le produit ci dessus est bien $\sigma(i)$ ce qui est le cas. Par exemple : l'image de 1 par (5 6) est 1, l'image de 1 par (1 2) est 2, l'image de 2 par (1 3) est 2, l'image de 2 par (1 4) est 2. Li'mage de 1 par (1 4)(1 3)(1 2)(5 6) est donc $2 = \sigma(1)$.

En fait, il y a plusieurs façons d'écrire une permutation comme produit de transpositions. Par exemple, dans l'exemple ci-dessus, σ peut aussi s'écrire

$$(2\ 1)(2\ 4)(1\ 4)(2\ 3)(1\ 4)(5\ 6)$$

On peut utiliser des transpositions différentes et également trouver des écritures qui font intervenir un nombre différent de transpositions. Toutefois, nous constatons que notre première décomposition utilise 3 transpositions et que la deuxième en utilise 5. Si vous essayez, vous verrez que vous n'arrivez pas à écrire σ comme le produit de 2 ou de 4 transpositions. En fait, vous ne réusirez pas à l'écrire comme le produit d'un nombre pair de transpositions. C'est un phénomène général. Il y a plusieurs façons possibles d'écrire une permutation comme produit de transpositions mais la parité du nombre de transpositions nécessaires est toujours la même. Nous allons maintenant expliquer ce phénomène en introduisant la signature d'une permutation.

4.2 Signature d'une permutation

Définition 4.2.1 On appelle *signature* de la permutation $\sigma \in \mathfrak{S}_n$ le nombre

$$\varepsilon(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

En fait, ce nombre vaut ± 1. En effet, si nous calculons sa valeur absolue, il vient :

$$|\varepsilon(\sigma)| = \frac{\displaystyle\prod_{1\leq i<j\leq n} |\sigma(j)-\sigma(i)|}{\displaystyle\prod_{1\leq i<j\leq n} |j-i|} = \frac{\displaystyle\prod_{1\leq I<J\leq n} |J-I|}{\displaystyle\prod_{1\leq i<j\leq n} |j-i|}$$

en appellant I le plus petit des deux nombres $\sigma(i), \sigma(j)$ et J le plus grand. On voit alors que $|\varepsilon(\sigma)| = 1$.

Exemple.

1. Pour $\sigma = \mathrm{id}$, nous trouvons :

$$\varepsilon(\mathrm{id}) = \prod_{1\leq i<j\leq n} \frac{j-i}{j-i} = 1.$$

2. Pour $\tau = (1\ 2)$, nous trouvons :

$$\varepsilon(\tau) = \prod_{2<i<j\leq n} \frac{j-i}{j-i} \times \prod_{2<j\leq n} \frac{j-2}{j-1} \times \prod_{2<j\leq n} \frac{j-1}{j-2} \times \frac{1-2}{2-1}.$$

   Le premier terme correspond au cas où $2 < i < j$, le second au cas $i = 1$ et $j > 2$, le troisième au cas $i = 2$ et $j > 2$, le dernier au cas $i = 1$ et $j = 2$. On voit alors que $\varepsilon(\tau) = -1$.

Nous énonçons maintenant le théorème fondamental sur la signature :

Théorème 4.2.2 (Signature d'un produit de permutations) *La signature d'un produit est le produit des signatures : soit $\sigma, \sigma' \in \mathfrak{S}_n$, alors $\varepsilon(\sigma'\sigma) = \varepsilon(\sigma')\varepsilon(\sigma)$.*

Preuve. Nous avons :

$$\varepsilon(\sigma'\sigma) = \prod_{1\leq i<j\leq n} \frac{\sigma'\sigma(j)-\sigma'\sigma(i)}{j-i} = \prod_{1\leq i<j\leq n} \frac{\sigma'\sigma(j)-\sigma'\sigma(i)}{\sigma(j)-\sigma(i)} \times \frac{\sigma(j)-\sigma(i)}{j-i}.$$

Nous posons $I = \min(\sigma(i), \sigma(j))$ et $J = \max(\sigma(j), \sigma(i))$. Nous obtenons alors :

$$\varepsilon(\sigma'\sigma) = \prod_{1\leq I<J\leq n} \frac{\sigma'(J)-\sigma'(I)}{J-I} \times \prod_{1\leq i<j\leq n} \frac{\sigma(j)-\sigma(i)}{j-i} = \varepsilon(\sigma')\varepsilon(\sigma).$$

□

Ce théorème a pour conséquence la propriété que nous nous proposions de prouver au début de cette partie.

Corollaire 4.2.3 *1. Si τ est une transposition quelconque dans $\mathfrak{S}_n$, alors $\varepsilon(\tau) = -1$.*

2. *Si σ est une permutation quelconque dans $\mathfrak{S}_n$ qui se décompose comme produit de transpositions $\sigma = \tau_1 \dots \tau_k$, alors $\varepsilon(\sigma) = (-1)^k$.*

Preuve.

1. Soit $\tau = (i\ j)$ une transposition quelconque dans $\mathfrak{S}_n$. Nous constatons l'égalité suivante :

$$(2\ j)(1\ i)(i\ j)(1\ i)(2\ j) = (1\ 2).$$

En effet, nous suivons l'action de la permutation à gauche de l'égalité sur les nombres $1, 2, i, j$ (lire les compositions de droite à gauche) :

$$\begin{aligned} &1 \to i \to j \to 2 \\ &2 \to j \to i \to 1 \\ &i \to 1 \to i \\ &j \to 2 \to j \end{aligned}$$

Dès lors, il vient :

$$\varepsilon\big((2\ j)\big)^2 \varepsilon\big((1\ i)\big)^2 \varepsilon\big((i\ j)\big) = \varepsilon\big((1\ 2)\big).$$

Or $\varepsilon\big((2\ j)\big)^2 = \varepsilon\big((1\ i)\big)^2 = 1$ donc $\varepsilon\big((i\ j)\big) = \varepsilon\big((1\ 2)\big) = -1$ d'après l'exemple traité auparavant.

2. Nous avons :

$$\varepsilon(\sigma) = \varepsilon(\tau_1) \dots \varepsilon(\tau_k) = \underbrace{(-1) \dots (-1)}_{k \text{ fois}} = (-1)^k.$$

□

Ainsi, pour calculer la signature d'une permutation σ quelconque, on écrit cette permutation comme produit $\tau_1 \dots \tau_k$ de permutations. La signature vaut alors $(-1)^k$. Dans l'exemple étudié dans ce chapitre, la permutation

$$\sigma = (1\ 4)(1\ 3)(1\ 2)(5\ 6)$$

a donc pour signature $(-1)^4$ c'est à dire 1.

Chapitre 5

Le déterminant

Nous avions remarqué que certaines matrices sont inversibles et que d'autres ne le sont pas. Pourtant, ce n'est pas toujours simple de savoir si une matrice donnée est inversible. Certaines sont clairement inversibles comme $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ et d'autres ne le sont clairement pas comme $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Mais, pour d'autres matrices, c'est beaucoup moins clair. Par exemple, pouvez-vous dire rapidement si la matrice $\begin{pmatrix} 1 & 1 & 0 & 3 \\ 1 & 2 & 1 & 2 \\ 2 & 3 & 2 & 2 \\ 1 & 1 & 1 & 0 \end{pmatrix}$ est inversible [1] ? L'idéal serait de disposer d'une formule permettant, à l'aide des coefficients de la matrice, de calculer un nombre qui nous donne tout de suite l'information "la matrice est inversible ou non". L'objet de ce chapitre est précisément d'introduire ce nombre puis d'apprendre à le calculer. Nous verrons ensuite que ce nombre, appelé *déterminant de la matrice* a de nombreuses applications, au-delà même du problème de départ que nous venons de nous poser.

L'idée est ici d'étudier tout d'abord le cas de la dimension 2 avant de généraliser au cas de la dimension finie quelconque.

5.1 Le cas de la dimension 2

5.1.A Déterminant d'une matrice de taille 2×2

Avant d'entamer le cas général, nous étudions le cas particulier d'une matrice 2×2. Nous considérons une matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$ avec $M \neq 0$ et nous cherchons à étudier son éventuelle inversibilité. Pour ce faire, nous introduisons la matrice complémentaire qui sera très utile par la suite.

[1] La réponse est non car la dernière colonne vérifie $C_4 = C_1 + 2C_2 - 3C_3$ mais il n'est pas évident de le voir !

Définition 5.1.1 On appelle *matrice complémentaire* de M la matrice $\widetilde{M} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$.

Remarque 5.1.2 On a échangé les deux termes diagonaux et changé le signe des autres.

Calculons maintenant $M\widetilde{M}$ et $\widetilde{M}M$:

$$M\widetilde{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & -ab+ab \\ cd-cd & bc+ad \end{pmatrix} = (ad-bc)I_2.$$
$$\widetilde{M}M = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ad-bd & bd-bd \\ -ca+ca & -bc+ad \end{pmatrix} = (ad-bc)I_2.$$

Conclusion : $M\widetilde{M} = \widetilde{M}M = (ad-bc)I_2$. Le nombre $ad-bc$ semble jouer un rôle particulier. Nous le nommons *déterminant* de la matrice M.

Définition 5.1.3 Le *déterminant* de la matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$ est le nombre $ad-bc$. Nous le notons $\det(M)$ ou encore $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ (comme la matrice mais avec des barres verticales).

Etudions maintenant les conséquences de notre calcul.

– Si $\det(M) = 0$, alors M n'est pas inversible. En effet, supposons par l'absurde qu'elle le soit, nous aurions :

$$\widetilde{M} = M^{-1}M\widetilde{M} = M^{-1}.0 = 0.$$

Or, si la matrice $\widetilde{M}$ est nulle, il en est clairement de même de la matrice M elle-même. En supposant que M est inversible, nous concluons que $M = 0$. Cette contradiction prouve que notre hypothèse de départ était fausse et donc que M est non inversible.

– Si $\det(M) \neq 0.$, On peut alors écrire :

$$M\left(\frac{\widetilde{M}}{\det(M)}\right) = \left(\frac{\widetilde{M}}{\det(M)}\right)M = I_2.$$

Ceci prouve que M est inversible et que $M^{-1} = \dfrac{\widetilde{M}}{\det(M)}$.

Nous voyons donc que le déterminant permet très simplement de savoir si une matrice 2×2 est inversible et, le cas échéant, permet de calculer son inverse. Remarquons au passage que, dès qu'une matrice $M \neq 0$ est non inversible, il existe une matrice $\widetilde{M} \neq 0$ telle que $M\widetilde{M} = \widetilde{M}M = 0$. On dit que la matrice M est un *diviseur de zéro*. Remarquons donc encore une fois que "simplifier" par M est faux dès que M n'est pas inversible.

Dans ce paragraphe, on a donc prouvé le théorème suivant :

Théorème 5.1.4 *Soit M une matrice de taille 2×2.*

1. *On a $M\widetilde{M} = \widetilde{M}M = \det(M)I_2$.*
2. *La matrice M est inversible si et seulement si $\det(M) \neq 0$.*
3. *Lorsque M est inversible, son inverse vaut $\dfrac{\widetilde{M}}{\det(M)}$.*
   Lorsque M n'est pas inversible, c'est un diviseur de zéro.

5.1.B Cas de deux vecteurs

Soit E un $\mathbb{K}$-espace vectoriel de dimension 2 muni d'une base $\mathcal{B}$. Soit x_1 et x_2 deux vecteurs de E dont les coordonnées dans la base $\mathcal{B}$ sont données par

$$x_1 = a_{11}e_1 + a_{21}e_2 \text{ et } x_2 = a_{12}e_1 + a_{22}e_2.$$

En écrivant ces coordonnées en colonnes, on obtient la matrice $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$.

Nous savons que la famille (x_1, x_2) est libre si et seulement si la matrice A est inversible, auquel cas la famille (x_1, x_2) est une base de E.

Théorème 5.1.5 (Déterminant dans le plan)

1. *Avec les notations précédentes, on appelle* déterminant *des deux vecteurs x_1, x_2 dans la base $\mathcal{B}$ et on note $\det_{\mathcal{B}}(x_1, x_2)$ le déterminant de la matrice A. Il s'agit du nombre $a_{11}a_{22} - a_{12}a_{21}$.*
2. *La famille (x_1, x_2) est libre si et seulement si $\det_{\mathcal{B}}(x_1, x_2) \neq 0$.*

Remarque 5.1.6 Il est important de remarquer que le déterminant de deux vecteurs se calcule relativement à une base $\mathcal{B}$ donnée au départ (pour avoir des coordonnées) tandis que le déterminant d'une matrice ne fait référence à aucune base. D'ailleurs, il est bon de prendre l'habitude d'éviter de parler de base, de vecteur, de noyau ou d'image d'une matrice.

5.1.C Premières propriétés du déterminant

Nous regroupons dans la proposition suivante des propriétés du déterminant qu'il est très aisé de vérifier grâce à sa définition. Les vérifications sont laissées en exercice.

Proposition 5.1.7 *Soit $M \in \mathcal{M}_2(\mathbb{K})$ et soit x_1, x_2 deux vecteurs d'un plan E rapporté à une base $\mathcal{B}$.*

1. *Pour tout scalaire $\lambda \in \mathbb{K}$, on a $\det(\lambda M) = \lambda^2 \det(M)$.*
2. $\det({}^tM) = \det(M)$.
3. $\det_{\mathcal{B}}(x_2, x_1) = -\det_{\mathcal{B}}(x_1, x_2)$.

4. *Pour tout scalaire $\lambda \in \mathbb{K}$ et tout vecteur $x_3 \in E$, on a :* $\det_{\mathcal{B}}(x_1, x_2 + \lambda x_3) = \det_{\mathcal{B}}(x_1, x_2) + \lambda \det_{\mathcal{B}}(x_1, x_3)$. *On a de même avec la variable de gauche.*
5. $\det_{\mathcal{B}}(x_1, x_1) = 0$. *On dit que le déterminant est* alterné *: si on répète deux fois le même vecteur, le résultat est nul.*

5.1.D Les formes bilinéaires alternées

La propriété 4 ci-dessus traduit le fait que le déterminant dans $\mathbb{K}^2$ est linéaire en chacune des deux variables. On dit pour cela que c'est une forme bilinéaire.

Définition 5.1.8 Soit E un $\mathbb{K}$ espace vectoriel. Une application $f : E \times E \to \mathbb{K}$ est une *forme bilinéaire* lorsqu'elle vérifie :

1. Pour tous $x, x', y \in E$ et $\lambda \in \mathbb{K}$, $f(x + \lambda x', y) = f(x, y) + \lambda f(x', y)$. (On parle de linéarité à gauche)
2. Pour tous $x, y, y' \in E$ et $\lambda \in \mathbb{K}$, $f(x, y + \lambda y') = f(x, y) + \lambda f(x, y')$. (On parle de linéarité à droite)

Ces deux propriétés peuvent également s'exprimer en disant que :

1. Pour tout $y \in E$ fixé, l'application $\begin{array}{rccc} f_y : & E & \to & E \\ & x & \mapsto & f(x, y) \end{array}$ est une forme linéaire.
2. Pour tout $x \in E$ fixé, l'application $\begin{array}{rccc} f_x : & E & \to & E \\ & y & \mapsto & f(x, y) \end{array}$ est une forme linéaire.

Une forme bilinéaire $f : E \times E \to \mathbb{K}$ est dite :

1. *symétrique* lorsque, pour tous $x, y \in E$, $f(x, y) = f(y, x)$.
2. *antisymétrique* lorsque, pour tous $x, y \in E$, $f(x, y) = -f(y, x)$.
3. *alternée* lorsque, pour tout $x \in E$, $f(x, x) = 0$.

Exemple. Le produit scalaire sur $\mathbb{R}^n$, $x \cdot y = \sum_{i=1}^{n} x_i y_i$ est une forme bilinéaire symétrique.

Avec ces définitions, le déterminant de deux vecteurs est une forme bilinéaire alternée et antisymétrique sur $E = \mathbb{K}^2$.

Proposition 5.1.9 *Soit E un $\mathbb{K}$-espace vectoriel et $f : E \times E \to \mathbb{K}$ une forme bilinéaire.*

$$f \text{ est antisymétrique} \iff f \text{ est alternée.}$$

Preuve. Supposons tout d'abord que f soit antisymétrique. Alors, pour tout $x \in E$, nous avons $f(x,x) = -f(x,x)$ soit $2f(x,x) = 0$ d'où il ressort que $f(x,x) = 0$ et donc que f est alternée.

Supposons maintenant que f soit alternée. Soit $x, y \in E$. Nous calculons $f(x+y, x+y)$. En utilisant la bilinéarité de f, cela donne[2] :

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

Or, puisque f est alternée, nous savons que $f(x+y, x+y) = f(x,x) = f(y,y) = 0$ d'où $f(x,y) + f(y,x) = 0$ et $f(x,y) = -f(y,x)$: f est antisymétrique.

□

Le résultat fondamental suivant montre que les formes bilinéaires alternées sont toutes proportionnelles au déterminant.

Proposition 5.1.10 *Soit E un $\mathbb{K}$-espace vectoriel de dimension 2 muni d'une base (e_1, e_2). Soit $\lambda \in \mathbb{K}$ un nombre quelconque. Il existe une unique forme bilinéaire alternée f telle que $f(e_1, e_2) = \lambda$. Elle est donnée par :*

$$\forall x_1, x_2 \in E, \quad f(x_1, x_2) = \lambda \det_{\mathcal{B}}(x_1, x_2).$$

Preuve. Montrons tout d'abord l'existence. Tout d'abord, nous avons prouvé ci-dessus que le déterminant était une forme bilinéaire alternée et que $\det_{\mathcal{B}}(e_1, e_2) = 1$. Par conséquent, $\lambda \det_{\mathcal{B}}$ convient effectivement pour notre problème.

Montrons maintenant l'unicité. Soit f une telle forme bilinéaire alternée. Soit $x_1 = a_{11}e_1 + a_{12}e_2$ et $x_2 = a_{12}e_1 + a_{22}e_2$ deux vecteurs et leurs coordonnées dans la base $\mathcal{B}$. On a alors en développant :

$$\begin{aligned} f(x_1, x_2) &= a_{11}a_{21}\underbrace{f(e_1, e_1)}_{=0} + a_{11}a_{22}f(e_1, e_2) + a_{12}a_{21}\underbrace{f(e_2, e_1)}_{=-f(e_1,e_2)} + a_{12}a_{22}\underbrace{f(e_2, e_2)}_{=0} \\ &= f(e_1, e_2)\,(a_{11}a_{22} - a_{12}a_{21}) \\ &= \lambda \det_{\mathcal{B}}(x_1, x_2). \end{aligned}$$

Ceci démontre l'égalité voulue et donc l'unicité de f.

□

En dimension 2, nous avons trouvé une formule pour le déterminant puis nous avons montré que le déterminant donne toutes les formes bilinéaires alternées en dimension 2. En dimension n, nous allons procéder dans l'autre ordre. Nous commençons par étudier les formes n-linéaires alternées puis nous montrons qu'elles sont toutes proportionnelles entre elles, ce qui permet de définir le déterminant.

[2] Une remarque pour simplifier les calculs : une application bilinéaire se "développe" comme un produit.

5.2 Déterminant dans le cas général

On va essayer ici de généraliser les résultats obtenus en dimension 2. Tout d'abord, il s'agit de donner une "bonne" définition du déterminant grâce à une généralisation des applications bilinéaires en dimension n : les applications n-linéaires.

5.2.A Formes n-linéaires alternés en dimension n

Définition 5.2.1 Soit E un $\mathbb{K}$-espace vectoriel. Une *forme n-linéaire* est une application $f : E^n \to \mathbb{K}$ telle que, pour tous $a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n \in E$, l'application $\left\{ \begin{array}{ccc} E & \longrightarrow & \mathbb{K} \\ x & \longmapsto & f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_n) \end{array} \right.$ est linéaire. En d'autres termes, f est n-linéaire si et seulement si elle est linéaire par rapport à chacune de ses variables.

Soit $f : E^n \to \mathbb{K}$ une forme n-linéaire. On dit :

1. f est *symétrique* lorsque, pour toute permutation $\sigma \in \mathfrak{S}_n$ et pour tous vecteurs $x_1, \dots, x_n \in E$, on a $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n)$.
2. f est *antisymétrique* lorsque, pour toute transposition $\tau \in \mathfrak{S}_n$ et pour tous vecteurs $x_1, \dots, x_n \in E$, on a $f(x_{\tau(1)}, \dots, x_{\tau(n)}) = -f(x_1, \dots, x_n)$ (si on échange deux vecteurs, on change le signe).
3. f est *alternée* lorsque, pour toute famille $(x_1, \dots, x_n)$ de vecteurs de E dont deux d'entre eux sont égaux, alors $f(x_1, \dots, x_n) = 0$.

Remarque. On voit que f est antisymétrique [respectivement alternée] si et seulement si pour tous entiers $i < j$ et tous vecteurs $x_1, \dots, x_n$, l'application

$$\left\{ \begin{array}{ccc} E \times E & \longrightarrow & \mathbb{K} \\ (x, y) & \longmapsto & f(x_1, \dots, \underset{\substack{\uparrow \\ \text{position } i}}{x}, \dots, \underset{\substack{\uparrow \\ \text{position } j}}{y}, \dots, x_n) \end{array} \right.$$

est antisymétrique [respectivement alternée]. Il résulte alors de la proposition 5.1.9 que :

$$f \text{ est alternée } \iff f \text{ est antisymétrique.}$$

Proposition 5.2.2 *Soit E un $\mathbb{K}$-espace vectoriel et $f : E^n \to \mathbb{K}$ une forme n-linéaire. Si f est antisymétrique, alors pour toute permutation $\sigma \in \mathfrak{S}_n$ et pour tous vecteurs $x_1, \dots, x_n$, on a $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \varepsilon(\sigma)\, f(x_1, \dots, x_n)$.*

Preuve. On sait que σ peut s'écrire $\sigma = \tau_1 \dots \tau_k$ où les τ_i sont des transpositions. Alors, pour toute famille $(x_1, \dots, x_n)$, on a :

$$\begin{aligned} f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) &= f(x_{\tau_1 \dots \tau_k(1)}, \dots, x_{\tau_1 \dots \tau_k(n)}) \\ &= -f(x_{\tau_2 \dots \tau_k(1)}, \dots, x_{\tau_2 \dots \tau_k(n)}) \\ &= \dots \\ &= (-1)^k f(x_1, \dots, x_n) \\ &= \varepsilon(\sigma) f(x_1, \dots, x_n). \end{aligned}$$

□

Nous avons maintenant établi tous les résultats nécessaires pour introduire le déterminant de n vecteurs en dimension n.

5.2.B Définition du déterminant de n vecteurs en dimension n

Dans cette partie, nous considérons E un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = (e_1, \dots, e_n)$.

Théorème 5.2.3 (fondamental sur le déterminant) *Soit $\lambda \in \mathbb{K}$. Il existe une et une seule forme n-linéaire alternée $f : E^n \to \mathbb{K}$ telle que $f(e_1, \dots, e_n) = \lambda$.*

Preuve. Dans toute la preuve, si $x_1, \dots, x_n$ sont des vecteurs de E, nous notons a_{ij} leurs coordonnées dans la base $\mathcal{B}$, c'est-à-dire :

$$x_j = \sum_{i=1}^{n} a_{ij} e_i.$$

Montrons l'unicité. Soit f une autre application n-linéaire alternée vérifiant $f(e_1, \dots, e_n) = \lambda$.

Nous avons alors :

$$f(x_1, \dots, x_n) = f\left(\sum_{i_1=1}^{n} a_{i_1 1} e_{i_1}, \dots, \sum_{i_n=1}^{n} a_{i_n n} e_{i_n} \right).$$

On prend soin de nommer différemment $i_1, \dots, i_n$ les indices de sommation dans les différentes sommes ce qui va éviter les confusions entre ces différentes sommes dans un instant. Par linéarité par rapport à chacune des variables,

ceci donne :

$$\begin{aligned} f(x_1,\dots,x_n) &= \sum_{i_1=1}^{n} a_{i_1 1}\, f\left(e_{i_1}, \sum_{i_2=1}^{n} a_{i_2 2} e_{i_2}, \dots, \sum_{i_n=1}^{n} a_{i_n n} e_{i_n}\right) \\ &= \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} a_{i_1 1} a_{i_2 2}\, f\left(e_{i_1}, e_{i_2}, \dots, \sum_{i_n=1}^{n} a_{i_n n} e_{i_n}\right) \\ &\vdots \\ &= \sum_{i_1,\dots,i_n} a_{i_1 1}\dots a_{i_n n} f(e_{i_1},\dots,e_{i_n}), \end{aligned}$$

où la dernière somme porte sur tous les indices $i_1,\dots,i_n$ possibles. Pour simplifier l'écriture et mieux comprendre ce qui se passe, nous introduisons l'application $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$ qui envoie 1 sur i_1, 2 sur i_2, ..., n sur i_n. Notre somme devient alors :

$$f(x_1,\dots,x_n) = \sum_{\sigma} a_{\sigma(1)1}\dots a_{\sigma(n)n} f(e_{\sigma(1)},\dots,e_{\sigma(n)}),$$

cette somme portant sur toutes les applications possibles de $\{1,\dots,n\}$ dans lui-même. Supposons maintenant que σ ne soit pas injective. Alors il existe $i \neq j$ tels que $\sigma(i) = \sigma(j) = k$ et donc le vecteur e_k est présent deux fois dans la famille $(e_{\sigma(1)},\dots,e_{\sigma(n)})$ (en positions i et j). Comme f est alternée, il en résulte que $f(e_{\sigma(1)},\dots,e_{\sigma(n)}) = 0$. Ainsi, dans la somme ci-dessus, tous les termes correspondant à des applications non injectives sont nuls. Mais nous savons qu'une application $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$ est bijective si et seulement si elle est injective. Donc finalement, dans la somme précédente, seules les σ qui sont des permutations donnent un terme non nul. Ceci prouve que

$$f(x_1,\dots,x_n) = \sum_{\sigma\in\mathfrak{S}_n} a_{\sigma(1)1}\dots a_{\sigma(n)n} f(e_{\sigma(1)},\dots,e_{\sigma(n)}).$$

Comme f est antisymétrique, nous avons $f(e_{\sigma(1)},\dots,e_{\sigma(n)}) = \varepsilon(\sigma) f(e_1,\dots,e_n)$. Finalement, nous trouvons :

$$\begin{aligned} f(x_1,\dots,x_n) &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\, a_{\sigma(1)1}\dots a_{\sigma(n)n} f(e_1,\dots,e_n) \\ &= \lambda \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\, a_{\sigma(1)1}\dots a_{\sigma(n)n}. \end{aligned}$$

Par conséquent, la formule $\lambda \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1}\dots a_{\sigma(n)n}$ est **la seule formule possible** pour une forme n-linéaire alternée valant λ en $(e_1,\dots,e_n)$. Pour établir l'existence, il nous faut montrer que cette formule marche effectivement, c'est-à-dire qu'elle définit bien une forme n-linéaire alternée valant λ en $(e_1,\dots,e_n)$.

Montrons maintenant l'existence. Nous posons donc, pour $x_1, \dots, x_n \in E$:

$$f(x_1, \dots, x_n) = \lambda \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}.$$

Nous vérifions que f est une forme n-linéaire alternée qui vaut λ en $(e_1, \dots, e_n)$.

1. f est n-linéaire. Nous montrons la linéarité par rapport à la première variable. La preuve est identique pour les autres variables. Soit $x'_1 = \sum_{i=1}^{n} a'_{i1} e_i$ et $\mu \in \mathbb{K}$. Alors :

$$\begin{aligned} f(x_1 + \mu x'_1, x_2, \dots, x_n) &= \lambda \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \left[a_{\sigma(1)1} + \mu a'_{\sigma(1)1} \right] a_{\sigma(2)2} \dots a_{\sigma(n)n}. \\ &= \lambda \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n} + \mu\lambda \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a'_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n} \\ &= f(x_1, x_2, \dots, x_n) + \mu f(x'_1, x_2, \dots, x_n). \end{aligned}$$

2. f est alternée. Nous montrons en fait plutôt que f est antisymétrique ce qui est équivalent d'après la remarque suivant la définition 5.2.1. Soit $\tau = (i\ j)$ une transposition. Alors :

$$\begin{aligned} f(x_{\tau(1)}, \dots, x_{\tau(n)}) &= \lambda \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)\tau(1)} \dots a_{\sigma(n)\tau(n)} \\ &= \lambda \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(i)j} \dots a_{\sigma(j)i} \dots a_{\sigma(n)n} \\ &= -\lambda \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma\tau) a_{\sigma(\tau(1))1} \dots a_{\sigma(\tau(j))j} \dots a_{\sigma(\tau(i))i} \dots a_{\sigma(\tau(n))n} \\ &= -\lambda \sum_{\sigma' \in \mathfrak{S}_n} \varepsilon(\sigma') a_{\sigma'(1)1} \dots a_{\sigma'(n)n} \qquad [\text{avec } \sigma' = \sigma\tau] \\ &= -f(x_1, \dots, x_n). \end{aligned}$$

   A la troisième ligne de calcul, nous avons utilisé que $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau) = -\varepsilon(\sigma)$ puisque τ est une transposition.

3. $f(e_1, \dots, e_n) = \lambda$. Pour $x_1 = e_1$, ..., $x_n = e_n$, nous avons $a_{ij} = 0$ dès que $i \neq j$ et $a_{ii} = 1$. Par conséquent, dans le produit $a_{\sigma(1)1} \dots a_{\sigma(n)n}$, il y a un terme nul dès qu'il y a un indice i avec $\sigma(i) \neq i$, c'est-à-dire dès que σ est différente de l'identité. Par conséquent, le seul terme non nul de toute la somme qui définit $f(e_1, \dots, e_n)$ est celui qui correspond à $\sigma = \mathrm{id}$, d'où $f(e_1, \dots, e_n) = \lambda\, \varepsilon(\mathrm{id})\, a_{11} \dots a_{nn} = \lambda$.

Par conséquent, l'application f vérifie bien toutes les propriétés requises.

□

Définition 5.2.4 Soit E un $\mathbb{K}$-espace vectoriel de dimension n rapporté à une base $(e_1, \ldots, e_n)$. Le *déterminant dans la base* $\mathcal{B}$ est l'unique application n-linéaire alternée valant 1 en $(e_1, \ldots, e_n)$. Nous le notons $\det_{\mathcal{B}}$.

Définition 5.2.5 Soit $x_1, \ldots, x_n$ n vecteurs de E. Nous introduisons leurs coordonnées a_{ij} dans la base $\mathcal{B}$:

$$x_j = \sum_{i=1}^{n} a_{ij} e_i.$$

Nous avons prouvé dans le théorème précédent la formule

$$\det_{\mathcal{B}}(x_1, \ldots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)\, a_{\sigma(1)1} \ldots a_{\sigma(n)n}.$$

Par analogie, nous définissons le déterminant d'une matrice $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ par la formule

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)\, a_{\sigma(1)1} \ldots a_{\sigma(n)n}.$$

On remarque que le déterminant d'une matrice A est égal au déterminant de ses vecteurs colonnes $C_1, \ldots, C_n$ dans la base canonique de $\mathbb{K}^n$.

Remarquons que la formule ci-dessus n'est pas très simple. Elle a un intérêt purement théorique. Nous allons montrer grâce à elle des propriétés intéressantes du déterminant. Dans un deuxième temps, nous allons voir des méthodes de calcul pratique de déterminants qui n'utilisent pas cette formule.

Tout d'abord, vérifions que cette formule coïncide avec celle que nous avions trouvée pour le cas 2×2. Pour $n = 2$, il n'y a que deux permutations : l'identité et la transposition $\tau = (1\ 2)$. La première a une signature égale à 1 et l'autre à -1. Enfin $a_{\mathrm{id}(1)1} a_{\mathrm{id}(2)2} = a_{11} a_{22}$ tandis que $a_{\tau(1)1} a_{\tau(2)2} = a_{21} a_{12}$. Finalement, nous obtenons :

$$\det(M) = a_{11} a_{22} - a_{21} a_{12}.$$

On retrouve bien la formule énoncée précédemment.

Explicitons maintenant la formule pour les matrices 3×3. Soit $M = (a_{ij})$. Les permutations de $\mathfrak{S}_3$ sont de trois sortes :

- L'identité id, de signature égale à 1.
- Les transpositions $(1\ 2)$, $(1\ 3)$ et $(2\ 3)$, de signature égale à -1.
- Les deux permutations données ci-dessous, de signature égale à 1 :

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

L'identité contribue à la somme $\sum_{\sigma\in\mathfrak{S}_3} \varepsilon(\sigma)\, a_{\sigma(1)1}a_{\sigma(2)2}a_{\sigma(3)3}$ par le terme $a_{11}a_{22}a_{33}$. Les transpositions donnent les termes $-a_{21}a_{12}a_{33}$, $-a_{31}a_{22}a_{13}$ et $-a_{11}a_{32}a_{23}$. Enfin, les deux dernières permutations fournissent les termes $a_{21}a_{32}a_{13}$ et $a_{31}a_{12}a_{23}$. Au final, nous trouvons :

$$\det(M) = a_{11}a_{22}a_{33}+a_{21}a_{32}a_{13}+a_{31}a_{12}a_{23}-a_{21}a_{12}a_{33}-a_{31}a_{22}a_{13}-a_{11}a_{32}a_{23}.$$

La formule est compliquée à retenir mais il y a encore une astuce visuel appelé *règle de Sarrus* qui permet de la retrouver.

Pour ceci on écit la matrice A puis en dessous les deux première lignes de la matrice. Le déterminant correspond alors à la somme du produit des termes diagonaux soustrait par la somme du produit des termes antidiagonaux.

Nous relevons les propriétés suivantes du déterminant qui sont immédiates :

Proposition 5.2.6 *1. Soit $M \in \mathcal{M}_n(\mathbb{K})$, on a $\det(M) = \det({}^tM)$.*

2. $\det_{\mathcal{B}}(e_1,\dots,e_n) = 1$.

3. $\det(I_n) = 1$.

Preuve.

1. Nous savons que ${}^tM = (b_{ij})$ avec $b_{ij} = a_{ji}$. Par conséquent, il vient :

$$\det({}^tM) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\, b_{\sigma(1)1}\dots b_{\sigma(n)n} = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\, a_{1\sigma(1)}\dots a_{n\sigma(n)}.$$

   Faisons le changement d'indice $\sigma' = \sigma^{-1}$ dans la somme ci-dessus. Nous obtenons :

$$\det({}^tM) = \sum_{\sigma'\in\mathfrak{S}_n} \varepsilon(\sigma')\, a_{\sigma'(1)1}\dots a_{\sigma'(n)n} = \det(M).$$

2. Nous avons déjà prouvé cette propriété dans le théorème fondamental.
3. La matrice I_n est la matrice des vecteurs e_i dans la base $\mathcal{B}$. Son déterminant vaut donc 1 d'après le point précédent.

□

Enfin, nous donnons le lien qui existe entre le déterminant et les familles liées :

Proposition 5.2.7 *Soit E un $\mathbb{K}$-espace vectoriel de dimension n rapporté à une base $(e_1,\dots,e_n)$*

1. Si la famille de vecteurs $(x_1,\dots,x_n)$ est liée, alors $\det_{\mathcal{B}}(x_1,\dots,x_n) = 0$.

2. On ne modifie pas la valeur de $\det_{\mathcal{B}}(x_1,\dots,x_n)$ en ajoutant à l'un des vecteurs x_k une combinaison linéaire des autres.

Preuve.

1. Soit $(x_1, \dots, x_n)$ une famille liée. L'un des vecteurs de la famille est alors combinaison linéaire des autres, c'est-à-dire : $x_k = \sum_{j \neq k} \lambda_j x_j$. Alors, en utilisant la linéarité du déterminant par rapport à sa k-ème variable, on trouve :

$$\begin{aligned} \det_{\mathcal{B}}(x_1, \dots, x_n) = \quad & \det_{\mathcal{B}}\left(x_1, \dots, \sum_{j \neq k} \lambda_j x_j, \dots, x_n\right) \\ & = \sum_{j \neq k} \lambda_j \det_{\mathcal{B}}(x_1, \dots, \underset{\substack{\uparrow \\ \text{position } k}}{x_j}, \dots, x_n). \end{aligned}$$

   Dans chacun des termes de la somme ci-dessus, il y a un vecteur répété deux fois (le vecteur x_j est à la fois en position j et en position k). Comme le déterminant est alterné, tous les termes de la somme sont nuls donc $\det_{\mathcal{B}}(x_1, \dots, x_n) = 0$.

2. Considérons u un vecteur combinaison linéaire des x_j $(j \neq k)$. Alors

$$\det_{\mathcal{B}}(x_1, \dots, x_k + u, \dots, x_n) = \det_{\mathcal{B}}(x_1, \dots, x_n) + \det_{\mathcal{B}}(x_1, \dots, u, \dots, x_n).$$

   Or la famille $(x_1, \dots, u, \dots, x_n)$ est liée donc $\det_{\mathcal{B}}(x_1, \dots, u, \dots, x_n) = 0$. Par conséquent, on voit bien que $\det_{\mathcal{B}}(x_1, \dots, x_k + u, \dots, x_n) = \det_{\mathcal{B}}(x_1, \dots, x_n)$ comme annoncé.

□

5.2.C Déterminant et produit

Le déterminant se comporte mal vis-à-vis de la somme puisque $\det(A + B) \neq \det(A) + \det(B)$ en général. En revanche, il se comporte de façon très agréable vis-à-vis du produit.

Théorème 5.2.8 (Déterminant d'un produit) *Soit $A, B \in \mathcal{M}_n(\mathbb{K})$. On a* $\det(AB) = \det(A)\det(B)$.

Preuve. On considère E un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = (e_1, \dots, e_n)$. Soit u et v les endomorphismes de E dont les matrices dans la base $\mathcal{B}$ sont A et B. Les vecteurs colonnes de A, de B et de AB sont les coordonnées des vecteurs $u(e_1), \dots, u(e_n)$, des vecteurs $v(e_1), \dots, v(e_n)$ et des vecteurs $uv(e_1), \dots, uv(e_n)$ dans la base $\mathcal{B}$. Par conséquent, nous avons :

$$\det_{\mathcal{B}}(u(e_1), \dots, u(e_n)) = \det(A), \qquad \det_{\mathcal{B}}(v(e_1), \dots, v(e_n)) = \det(B)$$
$$\text{et} \quad \det_{\mathcal{B}}(uv(e_1), \dots, uv(e_n)) = \det(AB).$$

Or l'application $f : \left\{ \begin{array}{ccc} E^n & \longrightarrow & \mathbb{K} \\ (x_1, \ldots, x_n) & \longmapsto & \det\limits_{\mathcal{B}}(u(x_1), \ldots, u(x_n)) \end{array} \right.$ est n-linéaire alternée. D'après le théorème fondamental sur le déterminant, elle vaut donc $f(e_1, \ldots, e_n) \det\limits_{\mathcal{B}}$.

Or $f(e_1, \ldots, e_n) = \det\limits_{\mathcal{B}}(u(e_1), \ldots, u(e_n)) = \det(A)$. Par conséquent, nous trouvons, pour tous vecteurs $x_1, \ldots, x_n$:

$$\det_{\mathcal{B}}(u(x_1), \ldots, u(x_n)) = \det(A) \det_{\mathcal{B}}(x_1, \ldots, x_n).$$

En appliquant cette formule avec les vecteurs $x_1 = v(e_1)$, ..., $x_n = v(e_n)$, on trouve :

$$\det_{\mathcal{B}}(uv(e_1), \ldots, uv(e_n)) = \det(A) \det_{\mathcal{B}}(v(e_1), \ldots, v(e_n))$$

Ainsi

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

□

Corollaire 5.2.9 *1. Une matrice $P \in \mathcal{M}_n(\mathbb{K})$ est inversible si et seulement si $\det(P) \neq 0$. Si elle est inversible, alors $\det(P^{-1}) = \dfrac{1}{\det(P)}$.*

2. Soit E un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathcal{B}$ et soit $(x_1, \ldots, x_n)$ une famille de n vecteurs. La famille $(x_1, \ldots, x_n)$ est une base si et seulement si $\det\limits_{\mathcal{B}}(x_1, \ldots, x_n) \neq 0$.

3. Deux matrices semblables ont le même déterminant. On dit que le déterminant est un invariant de similitude.

Preuve.

1. Supposons d'abord que P n'est pas inversible. Dans ce cas, les colonnes C_j de P forment une famille liée. Nous avons d'après la proposition 5.2.7, $\det(P) = \det\limits_{(\varepsilon)}(C_1, \ldots, C_n) = 0$.

   Supposons maintenant que P est inversible. On a alors $\det(P^{-1}) \det(P) = \det(P^{-1}P) = \det(I_n) = 1$ donc $\det(P) \neq 0$ et $\det(P^{-1}) = \dfrac{1}{\det(P)}$.

2. Soit P la matrice des coordonnées des x_j dans la base $\mathcal{B}$ écrites en colonnes. On a $\det(P) = \det\limits_{\mathcal{B}}(x_1, \ldots, x_n)$. La famille $(x_1, \ldots, x_n)$ est une base si et seulement si P est inversible d'où le résultat.

3. Soit A et B deux matrices semblables. Il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$. Alors

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

□

Définition 5.2.10 Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ un endomorphisme. Comme det est un invariant de similitude, les matrices de u dans une base quelconque ont toutes le même déterminant. On l'appelle *déterminant de l'endomorphisme* u et on le note $\det(u)$.

J'attire votre attention sur les points suivants :

1. Le déterminant d'une matrice est défini sans faire référence à une base. D'ailleurs, il faut prendre l'habitude de ne pas parler de base pour une matrice.
2. Le déterminant d'un système de n vecteurs en dimension n est défini dans une certaine base.
3. Le déterminant d'un endomorphisme est *a priori* défini grâce à sa matrice dans une certaine base. Il a donc l'air de dépendre de la base choisie. Mais justement non, nous venons de le prouver : c'est le même dans toutes les bases et c'est pour cela que nous le notons juste $\det(u)$ sans faire référence à une base.

5.3 Calculer un déterminant

Dans cette partie nous allons développer des techniques pour nous permettre de calculer un déterminant. Avant toute autre chose, rappelons que nous disposons d'une formule très simple pour le déterminant d'une matrice 2×2 et de la règle de Sarrus pour une matrice 3×3. Pour un n quelconque, nous disposons de la formule

$$\det(M) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)\, a_{\sigma(1)1} \dots a_{\sigma(n)n}.$$

C'est théoriquement satisfaisant mais personne ne souhaite (je l'espère !) calculer un déterminant 4×4 ou 5×5 grâce à cette formule : rappelons que la somme ci-dessus contient $n!$ termes et que $5! = 120$.

5.3.A Techniques de base

Commençons par récapituler les "trucs" de calcul que nous avons déjà prouvés :

Proposition 5.3.1 *1. On ne change pas un déterminant en prenant la transposée d'une matrice.*

2. Si $M \in \mathcal{M}_n(\mathbb{K})$, alors $\det(\lambda M) = \lambda^n \det(M)$

3. Si une matrice contient deux colonnes identiques, alors son déterminant est nul.

4. *Une matrice ayant une colonne nul est de déterminant nul.*
5. *Si on multiplie une colonne par λ, alors le déterminant est multiplié par λ. Si on multiplie deux colonnes par λ, alors le déterminant est multiplié par λ^2, etc.*
6. *Si on échange deux colonnes, on change le déterminant en son opposé.*
7. *On ne change pas le déterminant en ajoutant à une colonne une combinaison linéaire des autres colonnes.*
8. *Les règles de calculs 3 à 7 sont énoncées sur des colonnes. Grâce à la règle 1, on voit que des règles identiques sont valables sur les lignes.*

Exemple. Calculons le déterminant de la matrice $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 4 & -1 \\ 0 & -3 & 1 \end{pmatrix}$ grâce à ces règles de calcul.

$$\begin{vmatrix} 1 & 0 & 0 \\ 2 & 4 & -1 \\ 0 & -3 & 1 \end{vmatrix} \xrightarrow[-2L_1 \text{ à } L_2]{\text{ajout de}} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 4 & -1 \\ 0 & -3 & 1 \end{vmatrix} \xrightarrow[3C_3 \text{ à } C_2]{\text{ajout de}} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{vmatrix}$$

$$\xrightarrow[L_3 \text{ à } L_2]{\text{ajout de}} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

Donc ce déterminant vaut 1.

Vous pouvez constater que, dans l'exemple précédent, le but des manipulations successives sur les lignes et les colonnes de la matrice était de faire apparaître des coefficients nuls. Cette remarque est à la base de la technique développée dans le paragraphe suivant.

5.3.B Utilisation des termes nuls de la matrice

Nous commençons par un lemme :

Lemme 5.3.2 *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice ayant la forme suivante :*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & \vdots & & \vdots \\ | & \vdots & & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & \multicolumn{3}{c}{\multirow{3}{*}{B}} \\ | & & & \\ 0 & & & \end{pmatrix}.$$

La matrice B, extraite de A, est de taille $(n-1) \times (n-1)$. Alors $\det(A) = a_{11} \det(B)$.

De même, si $A = \begin{pmatrix} & & & a_{1n} \\ & \multicolumn{2}{c}{\multirow{2}{*}{B}} & \vdots \\ & & & a_{n-1,n} \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}$, *alors* $\det(A) = a_{nn} \det(B)$.

Preuve. Les deux assertions se prouvent de la même façon. Nous rédigeons la preuve de la seconde.

Nous avons $\det(A) = \det({}^tA) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)\, a_{1\sigma(1)} \dots a_{n\sigma(n)}$. Dès que $\sigma(n) \neq n$, on peut constater que $a_{n\sigma(n)} = 0$ puisque le seul terme non nul sur la dernière ligne de A est a_{nn}. Alors on peut réécrire le déterminant de A :

$$\begin{aligned}\det(A) &= a_{nn} \sum_{\substack{\sigma\in\mathfrak{S}_n \\ \sigma(n)=n}} \varepsilon(\sigma)\, a_{1\sigma(1)} \dots a_{n-1,\sigma(n-1)} \\ &= a_{nn} \sum_{\sigma\in\mathfrak{S}_{n-1}} \varepsilon(\sigma)\, a_{1\sigma(1)} \dots a_{n-1,\sigma(n-1)} \\ &= a_{nn} \det(B)\end{aligned}$$

□

De ce lemme, nous allons déduire la valeur du déterminant d'une matrice triangulaire.

Théorème 5.3.3 (Déterminant d'une matrice triangulaire) *Le déterminant d'une matrice triangulaire est égal aux produits des termes diagonaux. En particulier, une matrice triangulaire est inversible si et seulement si tous ses termes diagonaux sont non nuls.*

$$\det \begin{pmatrix} a_{11} & * & \cdots & * \\ 0 & a_{22} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} = a_{11} \dots a_{nn}.$$

La même formule est valable pour les matrices triangulaires inférieures.

Preuve. La preuve du théorème se fait par récurrence sur n.

Supposons $n = 1$. Il y a alors un seul terme dans la matrice : $A = (a_{11})$. On a alors bien $\det(A) = a_{11}$.

Supposons la propriété vraie au rang $n-1$. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice triangulaire supérieure. Elle est de la forme du lemme avec

$$B = \begin{pmatrix} a_{22} & * & \cdots & * \\ 0 & a_{33} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

L'hypothèse de récurrence appliquée à B prouve $\det(B) = a_{22} \dots a_{nn}$. On applique le lemme et on trouve $\det(A) = a_{11} \det(B) = a_{11} \dots a_{nn}$. La propriété est donc établie au rang n.

□

Remarquons que, puisque les matrices diagonales sont triangulaires, la formule précédente est vraie en particulier pour les matrices diagonales :

$$\det\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} = a_{11}\dots a_{nn}.$$

5.3.C Déterminants triangulaires par blocs

Théorème 5.3.4 (Déterminant d'une matrice triangulaire par blocs) *Une matrice triangulaire supérieure par blocs est une matrice $M \in \mathcal{M}_n(\mathbb{K})$ de la forme $M = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)$ où A, B, C sont elles-mêmes des matrices. Alors $\det(M) = \det(A)\det(C)$.*

Preuve. Nous considérons une matrice $M = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)$ avec $A \in \mathcal{M}_p(\mathbb{K})$, $B \in \mathcal{M}_{pq}(\mathbb{K})$ et $C \in \mathcal{M}_q(\mathbb{K})$. Nous distinguons deux cas.

- Supposons d'abord que $\det(A) = 0$. Dans ce cas, la matrice A n'est pas inversible donc ses colonnes sont liées. Dans la matrice M, les colonnes correspondant à A sont liées également (puisqu'elles finissent par des 0 et que leurs débuts sont liés). Par conséquent, la matrice M n'est pas inversible non plus et $\det(M) = 0$. On vérifie bien $\det(M) = \det(A)\det(B)$ dans ce cas particulier.
- Supposons maintenant que $\det(A) \neq 0$. Dans ce cas, la matrice A est inversible. Nous constatons alors qu'un produit par blocs nous donne la formule suivante :

$$M = \underbrace{\left(\begin{array}{c|c} A & 0 \\ \hline 0 & I_q \end{array}\right)}_{=A'} \underbrace{\left(\begin{array}{c|c} I_p & A^{-1}B \\ \hline 0 & C \end{array}\right)}_{=C'}$$

  Dès lors, il vient $\det(M) = \det(A')\det(C')$. En utilisant le lemme 5.3.2 page 69, nous voyons que $\det(A')$ est égal au déterminant de la même matrice dont on a rayé la dernière colonne et la dernière ligne. On peut recommencer jusqu'à obtenir $\det(A') = \det(A)$. De même, dans le calcul de $\det(C')$, on peut utiliser le lemme pour rayer la première colonne et la première ligne. En itérant, il vient $\det(C') = \det(C)$. Finalement, on a bien $\det(M) = \det(A)\det(C)$.

□

Exemple. Calculons le déterminant de la matrice suivante :

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & 0 & 9 & 10 \\ 0 & 0 & 11 & 12 \end{pmatrix}$$

On trouve :

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & 0 & 9 & 10 \\ 0 & 0 & 10 & 11 \end{vmatrix} = \left| \begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ \hline \multicolumn{2}{c|}{0} & 9 & 10 \\ & & 11 & 12 \end{array} \right|$$

$$= \begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix} \times \begin{vmatrix} 9 & 10 \\ 11 & 12 \end{vmatrix}$$

$$= (1 \times 6 - 5 \times 2) \times (9 \times 12 - 11 \times 10)$$

$$= (-4) \times (-2)$$

$$= 8.$$

5.4 Développement par rapport à une ligne ou à une colonne

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$. On définit $A_{ij} \in \mathcal{M}_{n-1}(\mathbb{K})$ la matrice extraite de A obtenue en rayant la ligne i et la colonne j :

$$A_{ij} = \left(\begin{array}{ccc|c|ccc} a_{11} & \dots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{ij} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ \hline a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{array} \right)$$

$$= \left(\begin{array}{ccc|ccc} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ \hline a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{array} \right)$$

Exemple. écrivons quelques unes des matrices A_{ij} relatives à la matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$:

$$A_{11} = \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}, \quad A_{21} = \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} \text{ et } A_{22} = \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix}.$$

Définition 5.4.1
- On appelle *mineur en position* (i, j) de la matrice A le nombre $D_{ij} = \det(A_{ij})$.
- On appelle *cofacteur en position* (i, j) de la matrice A le nombre $\Delta_{ij} = (-1)^{i+j} D_{ij}$.

Calculer Δ_{ij} consiste à calculer le déterminant de la matrice obtenue en rayant la ligne i et la colonne j puis à l'affecter d'un signe correspondant à la place de a_{ij} dans la matrice. Le tableau des signes à mettre s'obtient en partant de $+$ en position $(1,1)$ puis en remplissant la matrice sans jamais mettre le même signe sur deux cases voisines. En guise d'exemples, le tableau des signes est donné ci-dessous dans les cas 2×2, 3×3 et 4×4 :

$$\begin{pmatrix} + & - \\ - & + \end{pmatrix}, \quad \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix} \text{ et } \begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}.$$

Théorème 5.4.2 (Développement par rapport à une ligne ou à une colonne) *Soit $A=(a_{ij})\in\mathcal{M}_n(\mathbb{K})$. Nous avons :*

- *Pour tout $1\leq i\leq n$:* $\det(A)=\sum_{j=1}^{n} a_{ij}\Delta_{ij}$. *(développement par rapport à la ligne i)*
- *Pour tout $1\leq j\leq n$:* $\det(A)=\sum_{i=1}^{n} a_{ij}\Delta_{ij}$. *(développement par rapport à la colonne j)*

Preuve. Nous allons montrer le développement par rapport à une colonne. La preuve est identique pour le développement par rapport à une ligne. Notons C_1, ..., C_n les colonnes de la matrice A. Soit $(\varepsilon)=(\varepsilon_1,\ldots,\varepsilon_n)$ la base canonique de $\mathbb{K}^n$. Nous avons : $C_j=\sum_{i=1}^{n} a_{ij}\varepsilon_i$ et

$$\det(A)=\det_{(\varepsilon)}(C_1,\ldots,C_j,\ldots,C_n)=\sum_{i=1}^{n} a_{ij}\ \det_{(\varepsilon)}(C_1,\ldots,\varepsilon_i,\ldots,C_n).$$

Si nous regardons la formule que nous voulons montrer, nous voyons que le théorème sera établi dès que nous aurons prouvé la formule suivante :

$$\Delta_{ij}=\det_{(\varepsilon)}(C_1,\ldots,\varepsilon_i,\ldots,C_n)$$

Or le déterminant $\det_{(\varepsilon)}(C_1,\ldots,\varepsilon_i,\ldots,C_n)$ est celui que nous obtenons en remplaçant la colonne j de la matrice A par le vecteur $\varepsilon_i=\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}\leftarrow$ ligne i

Nous obtenons donc :

$$\det_{(\varepsilon)}(C_1,\dots,\varepsilon_i,\dots,C_n) = \begin{vmatrix} a_{11} & \dots & a_{1,j-1} & 0 & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ * & & * & 1 & * & & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & 0 & a_{n,j+1} & \dots & a_{nn} \end{vmatrix} \leftarrow \text{ligne } i$$

$$\uparrow$$

$$\text{colonne } j$$

Nous allons ramener le 1 de la colonne j en position $(1,1)$. Pour cela, nous commençons par ramener la colonne j à la place de la colonne 1. Cela nécessite d'intervertir la colonne ε_i et la colonne C_{j-1} puis la colonne ε_i et la colonne C_{j-2} etc, jusqu'à intervertir la colonne ε_i et la colonne C_1. La colonne ε_i a alors été déplacée $j-1$ fois (une fois pour chacune des colonnes C_1, …, C_{j-1}). Comme intervertir deux colonnes multiplie le déterminant par (-1), nous voyons que

$$\det_{(\varepsilon)}(C_1,\dots,\varepsilon_i,\dots,C_n) = (-1)^{j-1}\begin{vmatrix} 0 & a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & * & & * & * & & * \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix}$$

Nous recommençons le même jeu pour ramener le 1 de la première colonne en première place. Cela nécessite d'intervertir la ligne L_i avec L_{i-1}, puis avec L_{i-2} et ainsi de suite jusqu'à L_1. On a alors effectué $i-1$ interversion de lignes donc on a multiplié le déterminant par $(-1)^{i-1}$. Par conséquent :

$$\det_{(\varepsilon)}(C_1,\dots,\varepsilon_i,\dots,C_n) = (-1)^{i-1}(-1)^{j-1}\begin{vmatrix} 1 & * & & * & * & & * \\ 0 & a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix}$$

$$= (-1)^{i+j}\begin{vmatrix} 1 & * & \dots & * \\ 0 & & & \\ \vdots & & A_{ij} & \\ 0 & & & \end{vmatrix}$$

D'après le lemme 5.3.2, cela donne $(-1)^{i+j}\det(A_{ij}) = \Delta_{ij}$. C'est le résultat que nous avions annoncé et cela termine la preuve.

□

En clair, le théorème précédent permet de calculer un déterminant de taille n grâce à des déterminants de taille $n-1$ de la façon suivante :

1. On choisit une colonne (ou une ligne).
2. Pour chaque terme de la colonne, on multiplie le coefficient de la case par le déterminant de la matrice obtenue en rayant la ligne et la colonne correspondant à cette case, le tout affecté du signe donné par le schéma ci-dessus.
3. On additionne les nombres ainsi obtenus pour chaque case de la colonne.
4. Le résultat est le déterminant de la matrice.

Traitons un exemple pour rendre les choses plus claires :

Exemple. Calculons le déterminant de la matrice $A = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 0 & 4 \\ 5 & -1 & 6 \end{pmatrix}$.

1. On choisit une ligne ou une colonne. On a intérêt à choisir la deuxième ligne ou la deuxième colonne en raison du 0 qu'elles contiennent. Choisissons la deuxième ligne.
2. On commence par la première case de la ligne, celle qui contient 2. On raye la ligne et la colonne correspondante. Le mineur est donc

$$\begin{vmatrix} -2 & 3 \\ -1 & 6 \end{vmatrix} = (-2) \times 6 - (-1) \times 3 = -12 + 3 = -9.$$

   On multiplie cela par 2 (coefficient dans la case que nous sommes en train de traiter) et par -1 car le signe correspondant à cette case est $-$. Nous obtenons : $-2 \times (-9) = 18$.

   La deuxième case de la ligne contient un 0 donc sa contribution est nulle.

   La troisième case de la ligne contient 4. Le signe qui lui est affecté est $-$. Le mineur correspondant est $\begin{vmatrix} 1 & -2 \\ 5 & -1 \end{vmatrix} = 1 \times (-1) - 5 \times (-2) = -1 + 10 = 9$. On calcule alors $-4 \times 9 = -36$.
3. On additionne les résultats obtenus ci-dessus : $18 - 36 = -18$.
4. On obtient donc $\det(A) = -18$.

Ce théorème est particulièrement intéressant lorsque la matrice contient de nombreux coefficients nuls. En effet, on développera alors par rapport à une ligne ou une colonne contenant beaucoup de 0 et les calculs seront plus simples. On a tout intérêt à commencer par faire apparatre les 0 par combinaison linéaire de lignes et de colonnes.

Exemple. Calculons le déterminant de la matrice $A = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & -1 \\ 3 & 1 & 3 \end{pmatrix}$.

La deuxième colonne contient un élément nul. Nous avons donc intérêt à

développer par rapport à elle. C'est ce que nous allons faire mais nous commençons par faire apparaître un 0 de plus dans la colonne :

$$\begin{vmatrix} 1 & -1 & 2 \\ 2 & 0 & 1 \\ 3 & 1 & 3 \end{vmatrix} \xrightarrow[L_3 \text{ à } L_1]{\text{ajout de}} \begin{vmatrix} 4 & 0 & 5 \\ 2 & 0 & -1 \\ 3 & 1 & 3 \end{vmatrix} = \underbrace{(-1)}_{\text{signe}} \times \underbrace{1}_{\text{coefficient}} \times \begin{vmatrix} 4 & 5 \\ 2 & -1 \end{vmatrix} = -(-4-10) = 14$$

Comme vous vous en êtes sûrement aperçu, le calcul de déterminants nécessite d'être très soigneux, en particulier sur les signes. Nous allons maintenant prouver le théorème de développement par rapport à une ligne ou une colonne.

5.4.A Application 1 : inverse d'une matrice

Définition 5.4.3 Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$.

- On appelle *comatrice* de A et on note $\operatorname{Com} A$ la matrice des cofacteurs de A :
$$\operatorname{Com} A = (\Delta_{ij}).$$
- On appelle *matrice complémentaire* de A et on note $\widetilde{A}$ la transposée de la comatrice de A :
$$\widetilde{A} = {}^t\operatorname{Com} A.$$

Théorème 5.4.4 *Soit $A \in \mathcal{M}_n(\mathbb{K})$. On a la formule :*

$$A\widetilde{A} = \widetilde{A}A = \det(A)\, I_n.$$

En particulier, si $\det(A) \neq 0$*, alors* $A^{-1} = \dfrac{\widetilde{A}}{\det(A)}$.

Preuve. Soit $\mathcal{B} = (\varepsilon_1, \ldots, \varepsilon_n)$ la base canonique de $\mathbb{K}^n$. Nous nommons $C_1, \ldots, C_n$ les vecteurs colonnes de la matrice A. Soit $u = \sum_{i=1}^{n} b_i \varepsilon_i$ un vecteur colonne quelconque. En utilisant la n-linéarité du déterminant, il vient :

$$\det_{\mathcal{B}}(C_1, \ldots, C_{j-1}, u, C_{j+1}, \ldots, C_n) = \sum_{i=1}^{n} b_i \det_{\mathcal{B}}(C_1, \ldots, \varepsilon_i, \ldots, C_n).$$

Or nous avons vu page 74 que $\det_{\mathcal{B}}(C_1, \ldots, \varepsilon_i, \ldots, C_n) = \Delta_{ij}$. En conclusion, nous avons :

$$\det_{\mathcal{B}}(C_1, \ldots, C_{j-1}, u, C_{j+1}, \ldots, C_n) = \sum_{i=1}^{n} b_i \Delta_{ij}.$$

Nous appliquons tout d'abord ce résultat avec $u = C_j = \sum_{i=1}^{n} a_{ij}\varepsilon_i$. Nous trouvons :

$$\det(A) = \det_{\mathcal{B}}(C_1, \ldots, C_n) = \sum_{i=1}^{n} a_{ij}\Delta_{ij}.$$

Nous retrouvons le développement par rapport à une colonne que nous avions déjà prouvé. Nous appliquons maintenant ce résultat avec $u = C_k = \sum_{i=1}^{n} a_{ik}\varepsilon_i$ pour $k \neq j$. La famille $(C_1, \ldots, C_{j-1}, C_k, C_j, \ldots, C_n)$ contient alors deux fois la colonne C_k (en positions k et j). Par conséquent, il vient :

$$0 = \det_{\mathcal{B}}(C_1, \ldots, C_{j-1}, C_k, C_{j+1}, \ldots, C_n) = \sum_{i=1}^{n} a_{ik}\Delta_{ij}.$$

En résumé, nous avons prouvé la formule suivante :

$$\sum_{i=1}^{n} a_{ik}\Delta_{ij} = \begin{cases} \det(A) & \text{si } k = j \\ 0 & \text{si } k \neq j \end{cases}$$

Soit $\widetilde{A} = (c_{ij})$ et $\widetilde{A}A = (d_{ij})$. Nous avons $c_{ji} = \Delta_{ij}$. Dès lors, il vient :

$$d_{jk} = \sum_{i=1}^{n} c_{ji}a_{ik} = \sum_{i=1}^{n} a_{ik}\Delta_{ij} = \begin{cases} \det(A) & \text{si } k = j \\ 0 & \text{si } k \neq j \end{cases}$$

Ceci prouve que $\widetilde{A}A = \det(A)\, I_n$. Le calcul est identique pour prouver que $A\widetilde{A} = \det(A)\, I_n$.

□

Ce théorème nous donne une formule qui exprime la matrice inverse A^{-1} en fonction des coefficients de A. Cette formule est cependant très théorique et il est fortement déconseillé de l'utiliser pour calculer dans la pratique une matrice inverse. Dans le cas des matrices 2×2, on retrouve la matrice complémentaire précédemment introduite.

Posons $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Nous avons $A_{11} = (d)$, $A_{12} = (c)$, $A_{21} = (b)$ et $A_{22} = (a)$. En tenant compte des signes $\begin{pmatrix} + & - \\ - & + \end{pmatrix}$, il vient $\Delta_{11} = d$, $\Delta_{12} = -c$, $\Delta_{21} = -b$ et $\Delta_{22} = d$. La comatrice vaut donc $\operatorname{Com} A = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$ et enfin la matrice complémentaire vaut $\widetilde{A} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Nous retrouvons bien la formule du paragraphe 5.1.A.

5.4.B Application 2 : le déterminant de Vandermonde

Nous allons, en guise d'application du développement par rapport à une colonne, calculer le déterminant suivant que l'on rencontre assez souvent. Il est appelé déterminant de Vandermonde. On se donne $a_1, \ldots, a_n$ des scalaires et on range leurs puissances successives (de $a_1^0 = 1$ jusqu'à a_1^{n-1}) sur chaque colonne.

Proposition 5.4.5 *Soit $a_1, \ldots, a_n \in \mathbb{K}$. Le déterminant de Vandermonde est le suivant :*

$$V(a_1, \ldots, a_n) = \begin{vmatrix} 1 & \ldots & 1 & 1 \\ a_1 & \ldots & a_{n-1} & a_n \\ a_1^2 & \ldots & a_{n-1}^2 & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & \ldots & a_{n-1}^{n-2} & a_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

On fera attention à l'ordre dans le produit ci-dessus. En particulier, la matrice de Vandermonde est inversible si et seulement si tous les nombres a_i sont distincts deux à deux.

Preuve. Nous prouvons le résultat par récurrence sur n. Pour $n = 2$, le déterminant de Vandermonde s'écrit :

$$V(a_1, a_2) = \begin{vmatrix} 1 & 1 \\ a_1 & a_2 \end{vmatrix} = a_2 - a_1.$$

La propriété est établie au rang $n = 2$.

Supposons maintenant la propriété vraie au rang n et calculons $V(a_1, \ldots, a_{n+1})$. Nous distinguons deux cas :

1. Si deux des nombres a_i sont égaux, les deux lignes correspondantes sont égales dans $V(a_1, \ldots, a_n)$ donc $V(a_1, \ldots, a_n) = 0$, ce qui correspond précisément à la formule annoncée car si deux nombres a_i sont égaux, alors le produit $\displaystyle\prod_{1 \leq i < j \leq n} (a_j - a_i)$ contient 0 donc est nul.
2. Si tous les nombres a_i sont distincts les uns des autres, on introduit la fonction

$$P(x) = V(a_1, \ldots, a_n, x) = \begin{vmatrix} 1 & \ldots & 1 & 1 \\ a_1 & \ddots & a_n & x \\ \vdots & & \vdots & \vdots \\ a_1^{n-1} & \ldots & a_n^{n-1} & x^{n-1} \\ a_1^n & \ldots & a_n^n & x^n \end{vmatrix}$$

Si nous développons par rapport à la dernière colonne, nous voyons que $P(x)$ est un polynôme en x de degré inférieur ou égal à n. Le coefficient dominant de ce polynôme (c'est-à-dire le coefficient devant x^n)

correspond au dernier cofacteur obtenu en rayant la dernière ligne et la dernière colonne, c'est-à-dire $(-1)^{n+1+n+1}\begin{vmatrix} 1 & \dots & 1 & 1 \\ a_1 & \dots & a_{n-1} & a_n \\ a_1^2 & \dots & a_{n-1}^2 & a_n^2 \\ \vdots & & \vdots & \vdots \\ a_1^{n-1} & \dots & a_{n-1}^{n-1} & a_n^{n-1} \end{vmatrix} =$ $V(a_1, \dots, a_n)$.

Par ailleurs, en écrivant $P(a_1)$ on fait apparaître un déterminant contenant 2 lignes égales (la première et la dernière) donc $P(a_1) = 0$. De même, on voit que $P(a_2) = \cdots = P(a_n) = 0$. Comme les a_i sont tous distincts, cela fournit n racines distinctes pour le polynôme P. Comme un polynôme de degré n a au plus n racines, on en déduit que

$$\begin{aligned} P(x) &= V(a_1, \dots, a_n)(x - a_1) \dots (x - a_n) \\ &= \prod_{1 \leq i < j \leq n}(a_j - a_i) \prod_{1 \leq i \leq n}(x - a_i) \end{aligned}$$

Si nous calculons $P(a_{n+1})$, nous trouvons maintenant

$$\begin{aligned} V(a_1, \dots, a_{n+1}) &= \prod_{1 \leq i < j \leq n}(a_j - a_i) \prod_{1 \leq i \leq n}(a_{n+1} - a_i) \\ &= \prod_{1 \leq i < j \leq n+1}(a_j - a_i) \end{aligned}$$

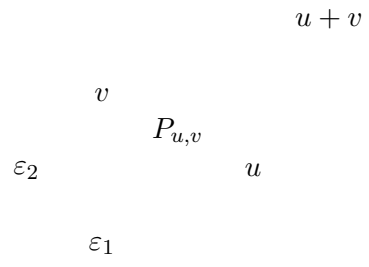
La propriété est établie au rang $n + 1$, ce qui achève la récurrence.

□

5.5 Aires et volumes

5.5.A Aires

Dans le plan usuel $\mathbb{R}^2$ rapporté à sa base canonique $(\varepsilon_1, \varepsilon_2)$, on se donne deux vecteurs u et v et on s'intéresse au paralallélogramme $P_{u,v}$ de côtés u et v. Ses quatre sommets sont 0, u, v et $u + v$:



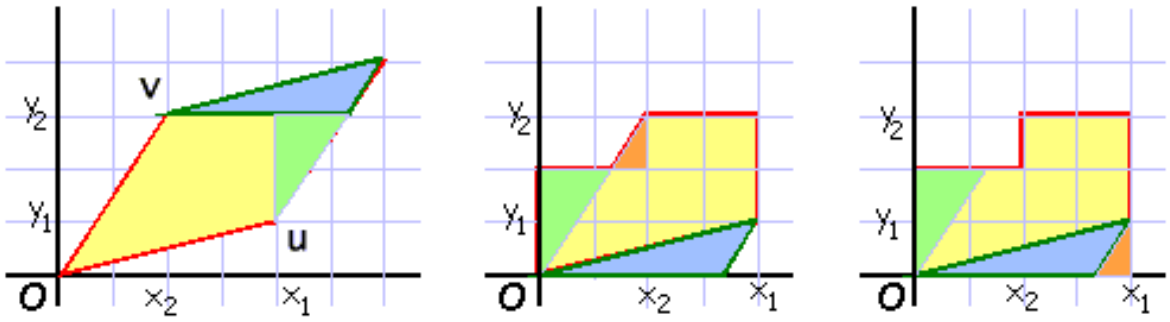
Nous admettons le théorème suivant :

Théorème 5.5.1 (Aire d'un parallélogramême) *Nous avons :*

$$Aire(P_{u,v}) = |\underset{(\varepsilon)}{\det}(u,v)|$$

.

Le dessin suivant n'a pas pour but de fournir une preuve rigoureuse mais juste de se convaincre du résultat. Le parallélogramême $P_{u,v}$ a une aire égale à la somme des aires jaune, verte et bleue (figure 1). En laissant glisser le triangle bleu le long du vecteur v, on le ramène en bas (figure 2). En laissant glisser le triangle vert le long du vecteur u, on le ramène à gauche (figure 2). Enfin, on constate que les deux triangles oranges sont superposables (figures 2 et 3). L'aire de $P_{u,v}$ est donc égale à celle du grand rectangle de côtés x_1 et y_2 moins celle du petit rectangle dont les côtés valent x_2 et y_1. Elle vaut donc $x_1y_2 - x_2y_1 = \underset{(\varepsilon)}{\det}(u,v)$. La configuration spéciale de notre dessin est telle que le déterminant est positif. En toute généralité, il faut prendre la valeur absolue.



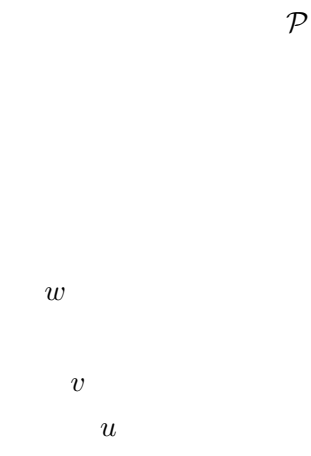
Nous admettons également le théorème suivant :

Théorème 5.5.2 *Soit $f \in L(\mathbb{R}^2)$. Soit $\mathcal{A}$ une partie de $\mathbb{R}^2$ dont l'aire vaut a, alors l'aire de $f(\mathcal{A})$ vaut $|\det f|a$. En d'autres termes, un endomorphisme f multiplie les aires par $|\det f|$.*

5.5.B Volumes

Des résultats identiques sont valables en dimension 3 et nous les admettons également :

Soient u, v, w trois vecteurs de $\mathbb{R}^3$. Nous notons $\mathcal{P}$ le parallélépipède formé sur u, v et w :



Théorème 5.5.3 *1. Le volume de $\mathcal{P}$ vaut $|\det_{(\varepsilon)}(u,v,w)|$.*

2. Un endomorphisme f de $\mathbb{R}^3$ multiplie les volumes par $|\det(f)|$.

5.6 Les systèmes linéaires

5.6.A Généralités

Définition 5.6.1 On appelle *système linéaire* à n équations et p inconnues un système d'équations de la forme

$$\begin{cases} a_{11}x_1 + \cdots + a_{1p}x_p &= b_1 \\ & \vdots \\ a_{n1}x_1 + \cdots + a_{np}x_p &= b_n \end{cases}$$

Les *coefficients* du système sont les nombres $a_{ij} \in \mathbb{K}$. *Le second membre* est donné par les n nombres $b_1, \ldots, b_n \in \mathbb{K}$ et les inconnues sont des nombres x_i dont on cherche à déterminer les valeurs possibles. Un *solution* du système est un p-uplet $(x_1, \ldots, x_p) \in \mathbb{K}^p$ de nombres qui vérifient les équations ci-dessus. Lorsque le second membre est nul, c'est-à-dire lorsque $b_1 = \cdots = b_n = 0$, on dit que le système est *homogène*. Il admet alors évidemment la *solution banale* $(x_1, \ldots, x_p) = (0, \ldots, 0)$.

Il est possible d'interpréter un système de plusieurs façons. Les allers-retours entre les différentes interprétations sont très utiles car certains faits peuvent être évidents dans une interprétation et pas dans une autre.

– **Interprétation matricielle :** On note $A = (a_{ij}) \in \mathcal{M}_{np}((\mathbb{K}))$ la matrice des coefficients, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p1}(\mathbb{K})$ le vecteur colonne

des inconnues et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathcal{M}_{n1}(\mathbb{K})$ le vecteur colonne du second membre. Le système est alors équivalent à l'équation matricielle $AX = b$ d'inconnue $X \in \mathcal{M}_{p1}(\mathbb{K})$.

- **Interprétation vectorielle :** On considère l'application linéaire $u : \mathbb{K}^p \to \mathbb{K}^n$ dont la matrice dans les bases canoniques de $\mathbb{K}^p$ et de $\mathbb{K}^n$ est A. On note $x = (x_1, \ldots, x_p) \in \mathbb{K}^p$ le vecteur des inconnues et $b = (b_1, \ldots, b_n) \in \mathbb{K}^n$ le vecteur du second membre. Le système est alors équivalent à l'équation vectorielle $u(x) = b$ d'inconnue $x \in \mathbb{K}^p$.
- **Interprétation scalaire :**
  On désigne par $C_1, \ldots, C_p \in \mathcal{M}_{n1}(\mathbb{K})$ les vecteurs colonnes de la matrice A. On les assimile à des vecteurs de $\mathbb{K}^n$. Le p-uplet de scalaires $(x_1, \ldots, x_p) \in \mathbb{K}^p$ est solution du système si et seulement si b est combinaison linéaire des vecteurs colonnes C_i avec les x_i comme coefficients, c'est-à-dire si et seulement si $x_1C_1 + \cdots + x_pC_p = b$.

Donnons quelques exemples de l'intérêt des différentes interprétations.

1. Si le système est homogène, dire qu'il existe une solution non banale, c'est dire qu'il existe $(x_1, \ldots, x_p) \neq (0, \ldots, 0)$ solution du système. A l'aide de l'interprétation scalaire, cela revient donc à dire qu'il existe des coefficients non tous nuls x_i tels que $x_1C_1 + \cdots + x_pC_p = 0$. En d'autres termes, cela revient à dire que la famille $(C_1, \ldots, C_p)$ est liée.
2. Dire que, pour un second membre b donné, il existe au moins une solution du système, c'est dire qu'il existe un vecteur x tel que $u(x) = b$. Cela signifie que $b \in \mathrm{Im}(u)$.
3. Dire que, pour n'importe quel b donné, il existe au moins une solution du système, c'est dire que u est surjectif. Dire qu'il existe toujours une unique solution (c'est-à-dire un seul p-uplet de solutions), c'est dire que u est bijectif.
4. Lorsque A est une matrice inversible (ce qui implique en particulier que A est une matrice carrée donc qu'il y a autant d'équations que d'inconnues), l'équation matricielle $AX = B$ peut se réécrire $X = A^{-1}B$. Il existe donc une unique solution au système et elle est donnée par cette formule.

5.6.B Les systèmes de Cramer

Définition 5.6.2 Un système linéaire est *de Cramer* lorsqu'il y a autant d'équations que d'inconnues ($p = n$) et que la matrice A a un déterminant non nul.

Théorème 5.6.3 (Formules de Cramer) *Un système carré (avec autant d'inconnues que d'équations) est de Cramer si et seulement si il admet une*

unique solution. Cette solution unique est alors donnée par les formules :

$$\textit{Pour tout } 1 \leq j \leq n, \quad x_j = \frac{1}{\det(A)} \det_{\mathcal{B}}(C_1, \ldots, C_{j-1}, b, C_{j+1}, \ldots, C_n).$$

Le dernier déterminant désigne le déterminant de la matrice A où on a substitué la colonne j par le second membre b ($\mathcal{B}$ désigne la base canonique).

Preuve.

- Si $\det(A) \neq 0$, alors A est inversible et donc u est bijective. Par conséquent, pour tout $b \in \mathbb{K}^n$, il existe un unique $x \in \mathbb{K}^n$ tel que $u(x) = b$. Le système a donc une unique solution.
- Si $\det(A) = 0$, alors A est non inversible et $\mathrm{rg}(u) = r < n$. Si $b \notin \mathrm{Im}(u)$, alors il n'existe aucun vecteur $x \in \mathbb{K}^n$ tel que $u(x) = b$. Si $b \in \mathrm{Im}(u)$, alors il existe au moins un vecteur $y \in \mathbb{K}^n$ tel que $u(y) = b$. L'équation $u(x) = b$ peut alors se réécrire $u(x) = u(y)$, soit encore $u(x-y) = 0 \iff x - y \in \mathrm{Ker}(u)$. D'après le théorème du rang, nous avons $\dim(\mathrm{Ker}(u)) = n - r > 0$. Par conséquent, il existe des vecteurs non nuls dans $\mathrm{Ker}(u)$ c'est-à-dire qu'il existe des vecteurs $x \neq y$ tels que $x - y \in \mathrm{Ker}(u)$. Il s'ensuit que le système a plusieurs solutions. Plus précisément, l'ensemble des solutions est de la forme $y + \mathrm{Ker}(u)$. C'est ce qui s'appelle un *sous-espace affine* de $\mathbb{K}^n$.

Ceci démontre bien qu'un système carré est de Cramer si et seulement si il a une unique solution. Supposons maintenant que le système est de Cramer et soit $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ l'unique solution. Nous avons :

$$\begin{aligned}\det_{\mathcal{B}}(C_1, \ldots, C_{j-1}, b, C_{j+1}, \ldots, C_n) &= \det_{\mathcal{B}}\left(C_1, \ldots, C_{j-1}, \sum_{i=1}^{n} x_i C_i, C_{j+1}, \ldots, C_n\right) \\ &= \sum_{i=1}^{n} x_i \det_{\mathcal{B}}(C_1, \ldots, C_{j-1}, C_i, C_{j+1}, \ldots, C_n).\end{aligned}$$

Dans cette dernière somme, la colonne C_i est en position j. Par conséquent, lorsque $i \neq j$, la colonne C_i est présente deux fois (en positions i et j) donc le déterminant est nul. Par suite, dans la somme, seul le terme correspondant à $i = j$ est non nul. Celui-là vaut $\det_{(\varepsilon)}(C_1, \ldots, C_{j-1}, C_j, C_{j+1}, \ldots, C_n) = \det(A)$. On en déduit donc :

$$\det_{(\varepsilon)}(C_1, \ldots, C_{j-1}, b, C_{j+1}, \ldots, C_n) = x_j \det(A)$$

d'où les formules de Cramer.

□

Remarques :

1. Cette formule a un intérêt purement théorique car elle supposerait, pour être appliquée, le calcul de $n+1$ déterminants de taille n.

2. Cas d'un système homogène :
   - Un système homogène de Cramer n'admet que la solution banale $(0, \ldots, 0)$ comme unique solution.
   - Pour qu'un système carré homogène admette une solution non banale, il faut et il suffit que $\det(A) = 0$.

Chapitre 6

Le pivot de Gauss

Après avoir développé au chapitre précédent la théorie des déterminants et fourni quelques méthodes de calculs pour calculer un déterminant ou résoudre un système, nous allons nous intéresser à la méthode du pivot de Gauss qui permet la résolution pratique de problèmes donnés avec des matrices explicites. Cette méthode est beaucoup plus efficace que celles que nous avons développées auparavant.

6.1 Transformations élémentaires d'une matrice

6.1.A Les types de transformations élémentaires

Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice dont on note les lignes $L_1, \ldots, L_n$. Les *transformations élémentaires* sur les lignes de A sont de trois types :

Type (I) transvection :

C'est l'ajout de λ fois L_j à la ligne L_i (pour $i \neq j$). On la note schématiquement $L_i \leftarrow L_i + \lambda L_j$. On peut aussi l'exprimer en écrivant $[+\lambda L_j]$ en face de la ligne L_i. Par exemple :

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \qquad \begin{pmatrix} 5 & 7 & 9 \\ 4 & 5 & 6 \\ 3 & 3 & 3 \end{pmatrix} \begin{matrix} \\ \\ [-L_2] \end{matrix}$$

Type (II) échange :

On échange les positions de deux lignes. Schématiquement, cela se note $L_i \leftrightarrow L_j$ ou en écrivant $[L_i]$ devant L_j et $[L_j]$ devant L_i. Par exemple :

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad \begin{pmatrix} 3 & 4 \\ 1 & 2 \\ 5 & 6 \end{pmatrix} \begin{matrix} [L_2] \\ [L_1] \\ \end{matrix}$$

Type (III) dilatation :

On multiplie une ligne par une constante $\lambda \neq 0$. Schématiquement, cela est noté $L_i \leftarrow \lambda L_i$ ou bien en écrivant $[\times\lambda]$ devant L_i. Par exemple :

$$\begin{pmatrix} 2 & 4 & 6 \\ 5 & 6 & 7 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \end{pmatrix} \ [\times 1/2]$$

6.1.B Effets des transformations élémentaires

Nous avons vu aux chapitre précédent les faits suivants :

- Les transformations élémentaires conservent le rang. En particulier, elles conservent l'inversibilité.
- Les transformations élémentaires ne modifient pas les solutions d'un système.
- Les transvections conservent le déterminant. Les échanges de lignes multiplient le déterminant par (-1) et les dilatations par λ multiplient le déterminant par λ. Par exemple :

$$\begin{vmatrix} 2 & 4 & 6 \\ 5 & 6 & 7 \\ 1 & 1 & 0 \end{vmatrix} = 2 \times \begin{vmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 1 & 1 & 0 \end{vmatrix} \ [\times 1/2]$$

6.2 Le principe du pivot de Gauss

6.2.A Premier pas

Soit $A = (a_{ij}) \in \mathcal{M}_{np}(\mathbb{K})$ une matrice non nulle. On choisit un coefficient non nul $P_1 = a_{i_\circ j_\circ}$ que l'on appelle pivot. Après éventuellement un échange de lignes et de colonnes, on peut ramener le pivot P_1 en position $(1,1)$ dans la matrice. La méthode de Gauss consiste alors à effectuer des transvections pour annuler les coefficients de la première colonne situés sous le pivot. On ne modifie pas la ligne du pivot. Par exemple :

$$\begin{pmatrix} 2 & 3 & 4 \\ 5 & 3 & 6 \end{pmatrix} \qquad \begin{pmatrix} 2 & 3 & 4 \\ 0 & -9/2 & -4 \end{pmatrix} \ [-5L_1/2]$$

$$\begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{matrix} [L_2] \\ [L_1] \\ \end{matrix} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix} \begin{matrix} [-4L_1] \\ [-7L_1] \end{matrix} .$$

On constate que, pour la facilité des calculs, si on a le choix du pivot, autant prendre un coefficient valant 1 ou -1.

6.2.B Itération du procédé

Soit $A = (a_{ij})$ une matrice non nulle. On choisit un pivot $P_1 \neq 0$. On fait des échanges de lignes et de colonnes pour le ramener en position $(1,1)$.

On effectue la méthode de Gauss et cela donne une matrice de la forme

$$A' = \begin{pmatrix} P_1 & * & \dots & * \\ 0 & \multicolumn{3}{|c}{\multirow{3}{*}{B}} \\ | & & & \\ 0 & & & \end{pmatrix}$$

Le P_1 désigne le terme que nous avons choisi comme pivot et les $*$ des coefficients que nous ne cherchons pas à préciser. Si la matrice B est nulle, on arrête. Sinon, on peut recommencer le procédé précédent sur la matrice B. Après le choix d'un deuxième pivot P_2 et placement de P_2 en position $(2,2)$, on effectue la méthode de Gauss pour annuler les coefficients de la deuxième colonne situés en-dessous de P_2. En itérant le procédé, on obtient finalement une matrice de la forme :

$$C = \begin{pmatrix} P_1 & * & \dots & * \\ & \ddots & \ddots & \vdots \\ 0 & & P_k & * \end{pmatrix}$$

ou bien la même avec des lignes nulles à la fin si on ne peut plus trouver de pivot non nul à un certain moment. Une telle matrice est dite échelonnée.

6.3 Applications

6.3.A Calcul du rang

On part d'une matrice A et on effectue le pivot de Gauss. A la fin, nous obtenons une matrice échelonnée dont le rang est égale à celui de A puisque les transformations élémentaires et les échanges de lignes conservent le rang. Le rang de la matrice ainsi obtenue est alors clairement le nombre de pivots.

Exemple. Trouvons le rang de la matrice $A = \begin{pmatrix} 1 & 1 & -1 & 3 & 4 \\ 1 & 3 & 0 & 2 & 2 \\ 2 & 2 & -3 & 1 & 3 \\ 0 & 4 & 3 & 3 & 1 \end{pmatrix}$. Nous appliquons la méthode décrite ci-dessus. Les écritures à droite de la matrice donne les transformations effectuées. Les pivots sont entourés avant et après la transformation dans laquelle ils servent.

$$A = \begin{pmatrix} 1 & 1 & -1 & 3 & 4 \\ 1 & 3 & 0 & 2 & 2 \\ 2 & 2 & -3 & 1 & 3 \\ 0 & 4 & 3 & 3 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & -1 & 3 & 4 \\ 0 & \boxed{2} & 1 & -1 & -2 \\ 0 & 0 & -1 & -5 & -5 \\ 0 & 4 & 3 & 3 & 1 \end{pmatrix} \begin{matrix} \\ [-L_1] \\ [-2L_1] \\ \\ \end{matrix}$$

$$\begin{pmatrix} 1 & 1 & -1 & 3 & 4 \\ 0 & \boxed{2} & 1 & -1 & -2 \\ 0 & 0 & -1 & -5 & -5 \\ 0 & 0 & 1 & 5 & 5 \end{pmatrix} \begin{matrix} \\ \\ \\ [-2L_2] \end{matrix} \longrightarrow \begin{pmatrix} 1 & 1 & -1 & 3 & 4 \\ 0 & 2 & 1 & -1 & -2 \\ 0 & 0 & 1 & -5 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \\ \\ \\ [+L_3] \end{matrix}$$

Le nombre de pivots est égal à 3 donc $\mathrm{rg}(A) = 3$.

6.3.B Résolution de systèmes

On applique le pivot de Gauss. Ceci permet d'obtenir un système échelonné qu'il est aisé de résoudre.

Exemple.

$$\left\{\begin{array}{rrrrl} x & - y & + z & + 2t = 1 \\ -3x & + y & - z & + 2t = 4 \\ 2x & - y & + z & + 5t = 3 \\ -3x & + 2y & - 2z & + 3t = 5 \end{array}\right. \iff \left\{\begin{array}{rrrrll} x & - y & + z & + 2t = 1 & \\ & - 2y & + 2z & + 8t = 7 & (L_2 \leftarrow L_2 + 3L_1) \\ & y & - z & + t = 1 & (L_3 \leftarrow L_3 - 2L_1) \\ & - y & + z & + 9t = 8 & (L_4 \leftarrow L_4 + 3L_1) \end{array}\right.$$

$$\iff \left\{\begin{array}{rrrrll} x & - y & + z & + 2t = 1 & \\ & y & - z & + t = 1 & (L_3) \\ & - 2y & + 2z & + 8t = 7 & (L_2) \\ & - y & + z & + 9t = 8 & \end{array}\right.$$

$$\iff \left\{\begin{array}{rrrrll} x & - y & + z & + 2t = 1 & \\ & y & - z & + t = 1 & \\ & & & 10t = 9 & (L_3 \leftarrow L_3 + 2L_2) \\ & & & 10t = 9 & (L_4 \leftarrow L_4 + L_2) \end{array}\right.$$

$$\iff \left\{\begin{array}{l} x = 1 + y - z - 2t \\ y = 1 + z - t \\ t = \dfrac{9}{10} \\ z \text{ est quelconque} \end{array}\right.$$

$$a \iff \left\{\begin{array}{l} x = -\dfrac{7}{10} \\ y = z + \dfrac{1}{10} \\ t = \dfrac{9}{10} \\ z \text{ est quelconque} \end{array}\right.$$

L'ensemble des solutions est donc : $\mathcal{S} = \left\{ \left(-\frac{7}{10}, z + \frac{1}{10}, z, \frac{9}{10} \right) \ ; \ z \in \mathbb{K} \right\}$.

6.3.C Calculs de déterminants

On applique encore le pivot de Gauss. Comme on part cette fois d'une matrice carrée, on obtient une matrice triangulaire supérieure C dont le déterminant est égal au produit des termes diagonaux (c'est-à-dire au produit des pivots). On fera attention à changer le signe du déterminant pour chaque échange de ligne et à compenser les dilatations.

Exemple. Calculons le déterminant de la matrice $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 2 & -1 & 1 & 3 \\ -1 & 2 & 1 & 2 \end{pmatrix}$.

Nous avons :

$$\begin{aligned}
\det(A) &= \begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 2 & -1 & 1 & 3 \\ -1 & 2 & 1 & 2 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & -5 & -5 & -5 \\ 0 & 4 & 4 & 6 \end{vmatrix} \begin{matrix} \\ [-5L_1] \\ [-2L_1] \\ [+L_1] \end{matrix} \\
&= (-4) \times (-5) \times 2 \times \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & 3 \end{vmatrix} \begin{matrix} \\ [\times(-1/4)] \\ [\times(-1/5)] \\ [\times(1/2)] \end{matrix} \\
&= -40 \times \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 2 & 3 \end{vmatrix} \begin{matrix} \\ [L_3] \\ [L_2] \\ \end{matrix} \\
&= -40 \times \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} \begin{matrix} \\ \\ [-L_2] \\ [-2L_2] \end{matrix} = -40.
\end{aligned}$$

6.3.D Inverse d'une matrice

La première méthode pour inverser une matrice $A = (a_{ij})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{K})$ consiste tout simplement à résoudre un système $AX = Y$ d'inconnue X. On se donne donc deux vecteurs colonnes :

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{et} \qquad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

et on résoud le système $AX = Y$. On obtient :

$$\begin{cases} a_{11}x_1 + ... + a_{1n}x_n = y_1 \\ a_{21}x_1 + ... + a_{2n}x_n = y_2 \\ \cdots \\ a_{n1}x_1 + ... + a_{nn}x_n = y_n \end{cases}$$

On résoud ce système linéaire en utilisant la méthode du pivot de Gauss et on obtient une solution du type :

$$\begin{cases} x_1 = b_{11}y_1 + ... + b_{1n}y_n \\ x_2 = b_{21}y_1 + ... + b_{2n}y_n \\ \cdots \\ x_n = b_{n1}y_1 + ... + b_{nn}y_n \end{cases}$$

La matrice $B = (b_{ij})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{K})$ vérifie donc $X = BY$. C'est donc la matrice A^{-1}.

Une deuxième méthode pratiquement équivalente quoique légèrement plus rapide (grâce aux notations plus efficientes) est donnés comme suit. Commençons par une remarque : les transformations élémentaires s'obtiennent en multipliant A à gauche par certaines matrices. En effet, calculons $E^{ij}A$ pour E^{ij} un élément quelconque de la base canonique de $\mathcal{M}_n(\mathbb{K})$. On constate que $E^{ij}A = \begin{pmatrix} 0 \\ L_j \\ 0 \end{pmatrix} \leftarrow$ position i .

Par conséquent, si on multiple A à gauche par $B = I_n + \lambda E^{ij}$ $(i \neq j)$, il vient :

$$BA = A + \lambda E^{ij}A = \begin{pmatrix} L_1 \\ \vdots \\ L_i + \lambda L_j \\ \vdots \\ L_n \end{pmatrix} \leftarrow \text{ position } i$$

Par conséquent, la multiplication à gauche par B réalise la transvection $L_i \leftarrow L_i + \lambda L_j$.

Posons $T = I_n - E^{ii} - E^{jj} + E^{ij} + E^{ji}$. Alors on obtient TA égale A à laquelle on retranche la ligne L_i en position i, la ligne L_j en position j et à laquelle on ajoute la ligne L_i en position j et la ligne L_j en position i. Finalement, la multiplication à gauche par T réalise l'échange de lignes $L_i \leftrightarrow L_j$.

Enfin, posons, $D = I_n + (\lambda - 1)E^{ii}$. Alors DA égale A à laquelle on ajoute $(\lambda - 1) \times L_i$ dans sa ligne i. Cela revient à multiplier L_i par λ et réalise la dilatation $L_i \leftarrow \lambda L_i$.

Ces préliminaires étant posés, considérons A une matrice inversible. Pour calculer l'inverse de A, on applique le pivot de Gauss en n'agissant que sur les lignes. Toutefois, on améliore légèrement la méthode. Au lieu d'annuler uniquement les termes situés sous le pivot, on va annuler les termes au-dessous et au-dessus du pivot. On obtient ainsi une matrice D diagonale. Par des dilatations, on obtient alors I_n.

Comment est-on passé de A à I_n ? Nous avons effectué des transformations élémentaires sur les lignes de A. Nous venons de voir que cela revient à avoir multiplier A par des matrices à gauche. Nous avons donc $I_n = BA$ où B est le produit des matrices que nous avons utilisées. Par conséquent, il vient : $A^{-1} = B$. Il reste juste à calculer B. Pour cela, nous remarquons que $B = BI_n$ donc la matrice B s'obtient en faisant subir à la matrice I_n les mêmes transformations qu'à la matrice A. Nous disposons d'une présentation efficace pour effectuer cette inversion.

On dispose la matrice A et la matrice I_n l'une à côté de l'autre. Nous effectuons sur A des transformations élémentaires des lignes. Simultanément,

nous effectuons les mêmes transformations sur I_n. A la fin, nous obtenons la matrice I_n et, à côté, la matrice B.

Exemple. Nous cherchons à inverser $A = \begin{pmatrix} 1 & 0 & 1 & 3 \\ -1 & 2 & 2 & -4 \\ 1 & 2 & 1 & 2 \\ 1 & -2 & -1 & 5 \end{pmatrix}$. Nous juxtaposons A et I_n :

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 1 & 3 & 1 & 0 & 0 & 0 \\ -1 & 2 & 2 & -4 & 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 2 & 0 & 0 & 1 & 0 \\ 1 & -2 & -1 & 5 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 1 & 3 & 1 & 0 & 0 & 0 \\ 0 & 2 & 3 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & -1 & -1 & 0 & 1 & 0 \\ 0 & -2 & -2 & 2 & -1 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \\ [+L_1] \\ [-L_1] \\ [-L_1] \end{array}$$

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 1 & 3 & 1 & 0 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & -3 & 0 & -2 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{array}\right) \begin{array}{l} \\ \\ [-L_2] \\ [+L_2] \end{array}$$

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 1 & 3 & 1 & 0 & 0 & 0 \\ 0 & 2 & 3 & -1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & -3 & 0 & -2 & -1 & 1 & 0 \end{array}\right) \begin{array}{l} \\ \\ [L_4] \\ [L_3] \end{array}$$

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 2 & -1 & 1 & 0 & -1 \\ 0 & 2 & 0 & -4 & 1 & -2 & 0 & -3 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 3 & -2 & 2 & 1 & 3 \end{array}\right) \begin{array}{l} [-L_3] \\ [-3L_3] \\ \\ [+3L_3] \end{array}$$

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 7/3 & -7/3 & -2/3 & -3 \\ 0 & 2 & 0 & 0 & -5/3 & 2/3 & 4/3 & 1 \\ 0 & 0 & 1 & 0 & 2/3 & 1/3 & -1/3 & 0 \\ 0 & 0 & 0 & 3 & -2 & 2 & 1 & 3 \end{array}\right) \begin{array}{l} [-2L_4/3] \\ [+4L_4/3] \\ [-L_4/3] \\ \\ \end{array}$$

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 7/3 & -7/3 & -2/3 & -3 \\ 0 & 1 & 0 & 0 & -5/6 & 1/3 & 2/3 & 1/2 \\ 0 & 0 & 1 & 0 & 2/3 & 1/3 & -1/3 & 0 \\ 0 & 0 & 0 & 1 & -2/3 & 2/3 & 1/3 & 1 \end{array}\right) \begin{array}{l} \\ [\times 1/2] \\ \\ [\times 1/3] \end{array}$$

$$\text{d'où} \quad A^{-1} = \begin{pmatrix} 7/3 & -7/3 & -2/3 & -3 \\ -5/6 & 1/3 & 2/3 & 1/2 \\ 2/3 & 1/3 & -1/3 & 0 \\ -2/3 & 2/3 & 1/3 & 1 \end{pmatrix}$$

Chapitre 7

Arithmétique des polynômes

Dans ce chapitre, nos laissons de côté, pour un temps, l'algèbre linéaire pour nous intéresser aux polynômes. Ceci sera à la fois utile pour la suite de ce cours mais aussi dans d'autres matières.

7.1 Polynômes et fonctions polynomiales

On commence par définir les polynômes et étudier leurs principales propriétés.

7.1.A Qu'est-ce qu'un polynôme ?

Une application f correspond à la donnée de deux choses :
- Un ensemble de définition E et un ensemble d'arrivée F.
- Une (ou plusieurs) formule(s) de calcul qui permet, pour chaque élément x de l'ensemble de départ E de calculer son image $f(x)$, qui est un élément de F.

Exemples

1. On peut considérer la fonction
$$f : \left\{ \begin{array}{ccc} \mathbb{R}_+^* & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \ln(2x + \sqrt{x}) \end{array} \right.$$
2. On peut considérer la fonction
$$g : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \left| \begin{array}{ll} x & \text{si } x > 0 \\ x/2 & \text{si } x \leq 0 \end{array} \right. \end{array} \right. \quad \text{de graphe}$$

Dans un polynôme, il n'y a qu'une formule de calcul, sans ensemble de définition ni ensemble d'arrivée.
Exemple. $P(X) = 12X^3 + 7X^2 - 2X + 3 = 12X^3 + 7X^2 - 2X^1 + 3X^0$.

Un polynôme est une "boîte noire" où l'inconnue X est susceptible ensuite d'être remplacée par *presque* n'importe quoi, pourvu que l'élément b qu'on choisit de mettre à la place de X vérifie les propriétés suivantes :

- On peut donner un sens à b^0, ce qui signifie qu'on dispose d'une unité.
- On peut élever b au carré, au cube, etc. En clair, on peut multiplier les éléments entre eux.
- On peut multiplier b et ses puissances par des nombres (ici, $12b^3$, $7b^2$, $-2b$, $3b^0$).
- On peut ajouter les éléments entre eux pour former successivement $12b^3 + 7b^2$, $12b^3 + 7b^2 - 2b$, $12b^3 + 7b^2 - 2b + 3b^0$.

On s'aperçoit que ce sont exactement les propriétés d'une $\mathbb{K}$-algèbre qui sont nécessaires. En résumé, on peut remplacer l'inconnue X d'un polynôme à coefficients dans $\mathbb{K}$ par n'importe quel élément b d'une $\mathbb{K}$-algèbre.

Voyons quelques exemples avec le polynôme $P(X) = X^2 - X + 1$:

Fonction polynomiale : On peut choisir de remplacer X par n'importe quel nombre réel t dans l'écriture du polynôme

$$P(t) = t^2 - t^1 + t^0 = t^2 - t + 1.$$

Le rôle de l'unité est bien entendu joué ici par le nombre 1. Ceci permet alors de définir, pour tout nombre t, le nombre $P(t)$. Nous disposons maintenant de :

- Un ensemble de départ : $\mathbb{R}$;
- Un ensemble d'arrivée : $\mathbb{R}$;
- Une formule de calcul : $t \mapsto P(t)$.

Nous disposons donc d'une fonction, que nous nommons bien sûr *fonction polynomiale*.

Autre fonction polynomiale : Même si notre polynôme est à coefficient réels, nous pouvons décider de calculer $P(z)$ pour tout nombre complexe z (avec la même formule). Ceci définit une autre fonction polynomiale (de $\mathbb{C}$ dans $\mathbb{C}$ ce coup-ci) :

$$\left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & z^2 - z + 1 \end{array} \right.$$

Polynôme de matrice : Considérons la matrice $M = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. On a $M^0 = I_2$. On calcule M^2 :

$$M^2 = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}.$$

Donc $P(M) = M^2 - M + M^0 = M^2 - M + I_2$ donne :

$$P(M) = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}$$

Polynôme d'endomorphisme :

Considérons l'endomorphisme $u : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ (x,y,z) & \longmapsto & (y, y-x, z-y) \end{array} \right.$

Calculons le carré de l'endomorphisme u. Soit $(x,y,z) \in \mathbb{R}^3$ un vecteur quelconque.

$$\begin{aligned} u^2(x,y,z) &= u\big(u(x,y,z)\big) \\ &= u(y, y-x, z-y) \\ &= \big(y-x, (y-x)-y, (z-y)-(y-x)\big) \\ &= (y-x, -x, x-2y+z). \end{aligned}$$

Nous avons alors $P(u) = u^2 - u + u^0 = u^2 - u + \mathrm{id}_E$. Si $(x,y,z) \in \mathbb{R}^3$ est un vecteur quelconque, nous avons donc :

$$\begin{aligned} P(u)(x,y,z) &= (y-x, -x, x-2y+z) - (y, y-x, z-y) + (x,y,z) \\ &= (0,0,x-y+z). \end{aligned}$$

7.1.B Degré

Définition 7.1.1 Soit $P = \sum_{k=0}^{n} a_k X^k$ un polynôme.

1. Les coefficients sont les nombres a_k. Le corps $\mathbb{K}$ auquel ils appartiennent est appelé le corps de base. L'ensemble de tous les polynômes à coefficients dans $\mathbb{K}$ est noté $\mathbb{K}[X]$.
2. Le polynôme X est appelé *indéterminée* ou *inconnue*.
3. Lorsqu'on a pris le soin de prendre $a_n \neq 0$, on dit que P est de degré n et on écrit $\deg(P) = n$. Par convention, le degré du polynôme nul est égal à $-\infty$. Le coefficient a_n est alors appelé *coefficient dominant* ou *directeur* de P. L'ensemble de tous les polynômes à coefficients dans $\mathbb{K}$ et de degré inférieur ou égal à n est noté $\mathbb{K}_n[X]$.
4. Lorsque le coefficient dominant de P vaut 1, on dit que P est un polynôme *unitaire*.

Remarque. Si on écrit $P = \sum_{k=0}^{n} a_k X^k$ sans rien préciser sur le coefficient a_n, alors on ne peut pas être certain que P est de degré n. La seule chose que l'on puisse affirmer avec certitude, c'est $\deg(P) \leq n$.

On rappelle la proposition suivante, vue en première année :

Proposition 7.1.2 *Soit P et Q deux polynômes. On a :*

$$\deg(P+Q) \leq \max[\deg(P), \deg(Q)] \text{ et } \deg(PQ) = \deg(P) + \deg(Q).$$

Remarque. Si P et Q sont de même degré, il est possible que $\deg(P+Q)$ soit strictement inférieur. Cette situation se rencontre chaque fois que les coefficients dominants de P et Q se compensent. Par exemple :

$$P(X) = 2X^3+X^2-X+1, \quad Q(X) = -2X^3-X^2-X+2, \quad (P+Q)(X) = -2X+3.$$

L'unique polynôme de degré $-\infty$ est 0. Les polynômes de degré 0 sont les constantes non nulles.

7.1.C Polynômes dérivés

Définition 7.1.3 Soit $P(X) = \sum_{k=0}^{n} a_k X^k$ un polynôme. Le polynôme dérivé de P est le polynôme P' défini par

$$P'(X) = \sum_{k=1}^{n} k a_k X^{k-1} = \sum_{j=0}^{n-1} (j+1) a_{j+1} X^j.$$

En dérivant m fois de suite, on obtient le polynôme $P^{(m)}$.

Remarques

1. Si $m > \deg(P)$, alors $P^{(m)} = 0$.
2. Soit $P \in \mathbb{C}[X]$. La fonction polynomiale (de $\mathbb{R}$ dans $\mathbb{C}$) associée au polynôme P' est la dérivée de la fonction polynomiale associée au polynôme P.

Théorème 7.1.4 (Formule de Mac-Laurin) *Soit* $P = \sum_{k=0}^{n} a_k X^k$. *Pour tout* $0 \leq k \leq n$, *on a* $a_k = \dfrac{P^{(k)}(0)}{k!}$ *et donc*

$$P(X) = \sum_{k=0}^{n} \frac{P^{(k)}(0)}{k!} X^k.$$

Preuve. On établit (le faire en exercice) tout d'abord par récurrence sur k la formule :

$$(X^m)^{(k)} = m(m-1)\dots(m-k+1)X^{m-k}.$$

Il s'ensuit que la valeur en 0 du polynôme $(X^m)^{(k)}$ est nulle dès que $m \neq k$ et vaut $k!$ si $m = k$. Par conséquent, nous avons :

$$P^{(k)}(0) = \sum_{m=0}^{n} a_m (X^m)^{(k)}(0) = k! a_k$$

ce qui démontre la formule de Mac-Laurin.

□

Remarque. La formule de Mac-Laurin ressemble à la formule de Taylor-Young. Mais il n'y a pas de terme $o(X^n)$. Dans le cas des polynômes, la formule est exacte et non pas approchée.

7.1.D Division euclidienne

Soit n, p deux nombres entiers strictement positifs. La division euclidienne de n par p consiste à trouver le plus gros multiple de p inférieur ou égal à n puis, à compléter avec un reste (nécessairement strictement inférieur à p) pour obtenir n. En d'autres termes, la division euclidienne de n par p est l'écriture suivante :

$$n = pq + r, \text{ avec } q, r \in \mathbb{N}, \ \ 0 \leq r < p.$$

Le nombre q est le *quotient* et le nombre r est le *reste*.

On va procéder de la même manière pour des polynômes. Evidemment, il faut commencer par trouver un équivalent de "strictement inférieur à" : c'est le degré qui va remplacer cette notion dans le cas des polynômes.

Théorème 7.1.5 (Division euclidienne des polynômes) *Soit $S \in \mathbb{K}[X]$ un polynôme quelconque et $P \in \mathbb{K}[X]$ un polynôme non nul. Il existe un unique couple de polynômes (Q, R) tels que*

$$S = PQ + R \text{ avec } \deg(R) < \deg(P).$$

Preuve.

– Commençons par l'existence. Posons $p = \deg P$, $s = \deg S$ et raisonnons par récurrence sur s. Si $s < p$, alors $Q = 0$ et $R = S$ conviennent. Soit $s \geq p$. Nous supposons le résultat prouvé jusqu'au rang $s-1$ et nous le prouvons au rang s. Nous détaillons l'écriture des polynômes S et P :

$$S(X) = a_s X^s + a_{s-1} X^{s-1} + \cdots + a_0 \text{ et } P(X) = b_p X^p + b_{p-1} X^{p-1} + \cdots + b_0.$$

Nous posons $T(X) = \dfrac{a_s}{b_p} X^{s-p}$, ce qui est licite car $b_p \neq 0$ (le polynôme P est non nul) et $s \geq p$. Alors T est un polynôme de degré $s-p$. Par conséquent, le polynôme $S - TP$ a un degré inférieur ou égal à s. De plus, son terme de degré s vaut :

$$a_s - \frac{a_s}{b_p} b_p = 0.$$

Il s'ensuit que $\deg(S - TP) < s$. On peut donc appliquer l'hypothèse de récurrence à ce polynôme et il vient :

$$S - TP = PQ_0 + R \text{ avec } \deg(R) < \deg(P) \text{ d'où } S = P(Q_0 + T) + R.$$

En posant $Q = Q_0 + T$, on obtient le résultat au rang s, ce qui achève notre récurrence.

– Prouvons maintenant l'unicité.
Nous supposons que $PQ_1 + R_1 = PQ_2 + R_2$ avec $\deg(R_1) < \deg(P)$ et $\deg(R_2) < \deg(P)$.
Nous avons alors : $(Q_1 - Q_2)P = R_2 - R_1$. Or $\deg(R_2 - R_1) < \deg(P)$. Par ailleurs, si $Q_1 - Q_2 \neq 0$, alors $\deg((Q_1 - Q_2)P) = \deg(Q_1 - Q_2) + \deg(P) \geq \deg(P)$. C'est absurde donc $Q_1 - Q_2 = 0$. Alors on a aussi $R_2 - R_1 = 0$ donc le couple (Q_1, R_1) est égal au couple (Q_2, R_2).

□

7.1.E Racine et factorisation

Définition 7.1.6 Soit $P(X) = \sum_{k=0}^{n} a_k X^k$ un polynôme à coefficients dans $\mathbb{K}$. Soit $\lambda \in \mathbb{K}$ un nombre quelconque. On pose $P(\lambda) = \sum_{k=0}^{n} a_k \lambda^k = a_n \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1 \lambda + a_0$.

On dit que λ est une *racine* de P si $P(\lambda) = 0$.

Remarque. L'ensemble des racines d'un polynôme donné dépend du corps $\mathbb{K}$ qu'on s'est fixé. Par exemple, prenons le polynôme $X^2 + 1$. Il n'a aucune racine dans $\mathbb{R}$. En revanche, il possède deux racines dans $\mathbb{C}$: les nombres i et $-i$.

Lemme 7.1.7 *Le nombre $\lambda \in \mathbb{K}$ est une racine du polynôme $P \in \mathbb{K}[X]$ si et seulement si $(X - \lambda)$ se factorise dans P, c'est-à-dire si et seulement si il existe un polynôme $Q \in \mathbb{K}[X]$ tel que $P(X) = (X - \lambda)Q(X)$.*

Preuve. Supposons que λ est racine de P. Nous écrivons la division euclidienne du polynôme P par le polynôme $(X - \lambda)$. Il existe alors deux polynômes Q, R avec $\deg R < 1$ tels que $P(X) = (X - \lambda)Q(X) + R(X)$. Comme $\deg R < 1$, le polynôme R est constant. En estimant l'égalité précédente en λ, nous trouvons :

$$0 = P(\lambda) = 0 \times Q(\lambda) + R \text{ donc } R = P(\lambda).$$

Par conséquent, $P(X) = (X - \lambda)Q(X)$.

Réciproquement, si $P(X) = (X - \lambda)Q(X)$, alors $P(\lambda) = 0 \times Q(\lambda) = 0$ donc λ est une racine de P.

□

Dans $\mathbb{R}$ certains polynômes non constants (comme $X^2 + 1$) n'ont aucune racine. Ceci ne se produit jamais dans $\mathbb{C}$. C'est une propriété "miraculeuse" que nous admettons sans démonstration.

Théorème 7.1.8 (de d'Alembert-Gauss) *Tout polynôme complexe non constant a une racine dans* $\mathbb{C}$.

7.2 Arithmétique des polynômes

7.2.A Diviseurs et multiples

Définition 7.2.1 On dit qu'un polynôme P *divise* un polynôme Q ou encore que Q est un *multiple* de P lorsqu'on peut factoriser P dans Q, c'est-à-dire lorsqu'il existe un polynôme R tel que $Q = PR$.

Exemple. Le polynôme $Q(X) = -2X^7 + 2X^6 - 4X^5 + 17X^4 - 13X^3 + 15X^2 - 32X + 14$ est un multiple du polynôme $P(X) = 2X^3 + 2X - 7$. En effet, on constate que :

$$Q(X) = P(X)\left(-X^4 + X^3 - X^2 + 4X - 2\right)$$

Remarque. Si $\deg(P) = 0$ (c'est-à-dire si P est une constante non nulle), alors n'importe quel polynôme est un multiple de P.

Dans tous ces problèmes de divisibilité, on peut toujours se ramener à supposer que les polynômes considérés sont unitaires. En effet, on peut toujours mettre en facteur le coefficient dominant (non nul). En particulier, on peut constater le résultat suivant :

Si P et Q sont deux polynômes **unitaires** de même degré et que P est un multiple de Q, alors $P = Q$.

7.2.B Polynômes irréductibles

On a vu que, comme dans le cas des entiers, on peut définir une division euclidienne sur l'ensemble des polynômes. Dans l'ensemble des entiers, un nombre est dit premier si il n'est divisible que par 1 et lui-même. Le définition suivante donne une notion correspondante dans le cadre des polynômes.

Définition 7.2.2

1. On dit qu'un polynôme $P \in \mathbb{K}[X]$ est *composé* s'il existe deux polynômes Q et R **non constants** de $\mathbb{K}[X]$ tels que $P = QR$.
2. On dit qu'un polynôme $P \in \mathbb{K}[X]$ est *irréductible* s'il n'est ni constant ni composé.

Exemples.

1. Le polynôme $P(X) = 3$ est constant. Il n'est donc pas irréductible.
2. Les polynômes composés sont de degré supérieur ou égal à 2. En effet, si nous avons $P = QR$ avec Q et R non constants, nous avons $\deg(Q) \geq 1$ et $\deg(R) \geq 1$. Par conséquent, nous avons $\deg(P) = \deg(Q) + \deg(R) \geq 2$.

3. Tous les polynômes de degré 1 sont irréductibles. En effet, ils ne sont pas composés d'après ce que nous venons de dire et ils ne sont pas constants non plus.
4. Le polynôme $P(X) = X^2 + 1$ est irréductible si nous le considérons comme un élément de $\mathbb{R}[X]$ mais il est composé si nous le considérons comme un élément de $\mathbb{C}[X]$. La notion de polynômes irréductibles et composés dépend donc du corps $\mathbb{K}$. En effet, nous avons :

$$X^2 + 1 = (X - i)(X + i)$$

ce qui offre une factorisation de $X^2 + 1$ comme produit de deux polynômes non constants de $\mathbb{C}[X]$. Par conséquent, $X^2 + 1$ est composé en tant qu'élément de $\mathbb{C}[X]$.
En revanche, supposons que nous avons écrit $X^2 + 1 = QR$ avec Q et R deux polynômes non constants à coefficients réels. Puisque $\deg X^2 + 1 = 2$, on voit tout de suite que Q et R sont tous les deux de degré 1. Il existe donc des nombres $a, b, c, d \in \mathbb{R}$ tels que $Q(X) = aX+b$ et $R(X) = cX + d$. Il vient alors :

$$X^2 + 1 = (aX + b)(cX + d) = (ac)X^2 + (ad + bc)X + bd.$$

En identifiant les coefficients, il vient : $c = \dfrac{1}{a}$ et $d = \dfrac{1}{b}$ et $ad + bc = 0$ soit encore $\dfrac{a}{b} + \dfrac{b}{a} = \dfrac{a^2 + b^2}{ab} = 0$. Or a et b étant non nuls, on a $a^2 + b^2 > 0$ donc cette dernière égalité est impossible. Nous venons donc de prouver qu'il est impossible de trouver deux polynômes non constants Q et R dans $\mathbb{R}[X]$ tels que $X^2 + 1 = QR$ et donc $X^2 + 1$ est irréductible en tant qu'élément de $\mathbb{R}[X]$.

□

Il existe un lien entre les polynômes irréductibles et les racines mais il est moins évident qu'il n'y paraît.

Lemme 7.2.3 *Tout polynôme de degré supérieur ou égal à 2 ayant une racine dans $\mathbb{K}$ est composé.*

Preuve. Supposons que $\deg(P) \geq 2$ et que P a une racine $\lambda \in \mathbb{K}$. D'après le lemme 7.1.7, il existe un polynôme $Q \in \mathbb{K}[X]$ tel que $P(X) = (X - \lambda)Q(X)$. Puisque $\deg(Q) = \deg(P) - 1 \geq 1$, on voit que Q n'est pas constant. Par conséquent, P est composé.

□

L'erreur classique consiste à lire le lemme précédent à l'envers et à affirmer que si P est un polynôme sans racine, alors P est irréductible. C'est faux comme en témoigne l'exemple suivant :

$$P(X) = X^4 + 2X^2 + 1 = (X^2 + 1)^2 \quad \text{dans } \mathbb{R}[X].$$

Puisque $P(t) > 0$ pour tout $t \in \mathbb{R}$, il est clair que P n'a pas de racine. Pourtant P est un polynôme composé puisque c'est le produit de X^2+1 par lui-même.

Théorème 7.2.4 (– Cas de $\mathbb{C}[X]$) *Les polynômes irréductibles de $\mathbb{C}[X]$ sont exactement les polynômes de degré 1.*

Preuve. Nous avons déjà vu que les polynômes de degré 1 sont irréductibles (cela est vrai que le corps soit $\mathbb{C}$ ou non).

Considérons maintenant un polynôme $P \in \mathbb{C}[X]$ qui n'est pas de degré 1. Si $\deg(P) \leq 0$, alors P est constant et donc P n'est pas irréductible. Si $\deg(P) \geq 2$, alors il admet une racine d'après le théorème de d'Alembert-Gauss. D'après le lemme 7.2.3, il n'est pas irréductible.

□

Théorème 7.2.5 (Cas de $\mathbb{R}[X]$) *Il y a deux sortes de polynômes irréductibles dans $\mathbb{R}[X]$:*

1. *Les polynômes de degré 1.*
2. *Les polynômes de degré 2 dont le discriminant est strictement négatif.*

Preuve. Nous savons déjà que les polynômes de degré 1 sont irréductibles (c'est toujours le cas, que $\mathbb{R}$ soit le corps de base ou pas).

Nous montrons que les polynômes de degré 2 de discriminant positif ou nul sont composés. En effet, si P est de degré 2 et a un discriminant positif ou nul, on sait que P s'annule en un certain nombre $\lambda \in \mathbb{R}$. D'après le lemme 7.2.3, P est composé.

Nous montrons que les polynômes de degré 2 composés ont un discriminant positif ou nul. Soit P un polynôme composé de degré 2. Il existe alors deux polynômes non constants Q et R tels que $P = QR$. On voit directement que Q et R sont alors forcément de degré 1 et donc $Q(X) = aX + b$ et $R(X) = cX + d$. On a alors $P(X) = acX^2 + (ad + bc)X + bd$. Dès lors il vient

$$\Delta = (ad + bc)^2 - 4abcd = (ad)^2 + (bc)^2 - 2abcd = (ad - bc)^2 \geq 0.$$

Il nous reste à prouver que les polynômes de degré supérieur ou égal à 3 sont toujours composés. Soit donc P un polynôme de degré n supérieur ou égal à 3. S'il admet une racine réelle, le lemme 7.2.3 assure qu'il est composé et nous avons fini. Sinon, P étant non constant, il admet une racine complexe non réelle λ. Il existe donc un polynôme $Q \in \mathbb{C}[X]$ tel que

$$P(X) = (X - \lambda)Q(X).$$

Nous avons alors $P(\overline{\lambda}) = \overline{P(\lambda)} = 0$ (puisque les coefficients de P sont réels, ils ne sont pas affectés par la conjugaison). Par conséquent :

$$0 = (\overline{\lambda} - \lambda)Q(\overline{\lambda}).$$

Comme λ n'est pas réel, $\overline{\lambda} \neq \lambda$ et donc $Q(\overline{\lambda}) = 0$. Ceci prouve que $\overline{\lambda}$ est racine de Q et donc qu'il existe un polynôme $R \in \mathbb{C}[X]$ tel que

$$P(X) = (X - \lambda)(X - \overline{\lambda})R(X).$$

Posant $\lambda = a+ib$ avec a, b réels, il vient $(X-\lambda)(X-\overline{\lambda}) = X^2-2aX+a^2+b^2$. Nous avons donc

$$P(X) = \underbrace{(X^2 - 2aX + a^2 + b^2)}_{S(X)} R(X)$$

et le polynôme S est dans $\mathbb{R}[X]$. Nous avons presque démontré que P est composé (nous venons de l'écrire comme produit de deux polynômes). Il nous reste à vérifier deux éléments :

- R est non constant.
- R est à coefficients réels (et non pas complexes).

Comme $\deg P \geq 3$, nous avons $\deg(R) \geq 1$ donc R est non constant. Montrons que R a tous ses coefficients réels. Pour $x \in \mathbb{R}$, on a :

$$R(x) = \frac{P(x)}{S(x)} \in \mathbb{R}.$$

Ceci prouve que la fonction polynomiale associée à R envoie tous les réels dans $\mathbb{R}$ et donc, toutes les dérivées de cette fonction sont aussi des nombres réels. Ainsi, pour tout $k \in \mathbb{N}$, on a $R^{(k)}(0) \in \mathbb{R}$. D'après la formule de Mac-Laurin, on voit que R a tous ses coefficients réels.

□

7.2.C Reconnaître un diviseur commun à plusieurs polynômes

Soit $P \in \mathbb{K}[X]$ un polynôme quelconque. On remarque que le polynôme nul 0 est un multiple de P, que la somme de deux multiples de P est un multiple de P et qu'un multiple d'un multiple de P est encore un multiple de P. Cela nous pousse à la définition suivante :

Définition 7.2.6 Une partie $\mathcal{I}$ de $\mathbb{K}[X]$ est un *idéal* si elle vérifie les trois propriétés suivantes :

1. Le polynôme nul 0 appartient à $\mathcal{I}$.
2. La somme de deux éléments de $\mathcal{I}$ est encore dans $\mathcal{I}$.
3. Tout multiple d'un élément de $\mathcal{I}$ est encore dans $\mathcal{I}$.

L'ensemble des multiples d'un polynôme P donné est un idéal. La notion d'idéal est très utile car elle nous permet de reconnaître un ensemble de multiples sans connaî tre *a priori* leur diviseur commun.

Théorème 7.2.7 (Idéaux de $\mathbb{K}[X]$) *Soit $\mathcal{I}$ un idéal de $\mathbb{K}[X]$ contenant un polynôme non nul. Il existe un unique polynôme unitaire P_0 tel que $\mathcal{I}$ soit exactement l'ensemble des multiples de P_0. On dit que P_0 est* le générateur *de l'idéal $\mathcal{I}$.*

Preuve. Nous désignons par $\mathcal{D}$ l'ensemble : $\mathcal{D} = \{\deg(S)\ ;\ S \in \mathcal{I},\ S \neq 0\}$. Il s'agit d'une partie non vide de $\mathbb{N}$ donc il admet un minimum n. Soit P_1 un polynôme de degré n dans $\mathcal{I}$. Soit $c \neq 0$ le coefficient dominant de P_1. Comme le polynôme $P_0 = \dfrac{P_1}{c}$ est un multiple de P_1, il appartient encore à $\mathcal{I}$ (propriété 3 des idéaux). De plus, P_0 est unitaire. Comme $\mathcal{I}$ est un idéal, tous les multiples de P_0 sont dans $\mathcal{I}$ (à nouveau la propriété 3). Réciproquement, nous voulons montrer que tous les éléments de $\mathcal{I}$ sont des multiples de P_0. Soit donc $P \in \mathcal{I}$. Nous écrivons la division euclidienne de P par P_0 : $P = P_0Q + R$ avec $Q, R \in \mathbb{K}[X]$ et $\deg(R) < n$. Grâce à la propriété 3 des idéaux, on voit que $-P_0Q \in \mathcal{I}$ et donc, en utilisant la propriété 2 : $R = P - P_0Q \in \mathcal{I}$. Or $\deg(R) < n$ donc, d'après la définition de n, on a $R = 0$, c'est-à-dire $P = P_0Q$ et P est bien un multiple de P_0. Ceci prouve l'existence de P_0.

Si P_2 est un autre polynôme unitaire engendrant $\mathcal{I}$, on a alors P_0 multiple de P_2 et P_2 multiple de P_0. Comme P_0 et P_2 sont unitaires, cela implique bien $P_0 = P_2$ et prouve l'unicité.

□

Remarque. L'idéal $\{0\}$ réduit au polynôme nul est engendré par le polynôme nul. Le résultat est donc encore vrai pour l'idéal $\{0\}$ mais le polynôme P_0 n'est pas unitaire.

Donnons-nous deux polynômes non nuls P et Q. L'ensemble $\mathcal{I} = \{PU + QV\ ;\ U, V \in \mathbb{K}[X]\}$, c'est-à-dire l'ensemble des sommes de multiples de P et de multiples de Q est un idéal (le vérifier en exercice). Il contient les polynômes P et Q qui sont non nuls (puisque $P = P \times 1 + Q \times 0$ et $Q = P \times 0 + Q \times 1$). Par le théorème précédent, il existe donc un unique polynôme unitaire P_0 qui engendre $\mathcal{I}$. Nous l'appelons le *plus grand diviseur commun* de P et Q et nous le notons $P \wedge Q$.

Le polynôme $P \wedge Q$ vérifie les trois propriétés suivantes :
- Le polynôme $P \wedge Q$ est unitaire.
- Le polynôme $P \wedge Q$ est un diviseur commun à P et à Q.
- Si D est un autre diviseur commun à P et à Q alors D est un diviseur de $P \wedge Q$.

Preuve. Nous savons déjà que $P \wedge Q$ est unitaire. De plus l'idéal $\mathcal{I}$ est égal à l'ensemble des multiples de $P \wedge Q$. Or $P \in \mathcal{I}$ et $Q \in \mathcal{I}$ donc $P \wedge Q$ est un diviseur de P et de Q.

Si D est un diviseur de P et de Q, alors tout polynôme de la forme $PU + QV$ est aussi un multiple de D donc $P \wedge Q$ (qui est de la forme $PU + QV$ puisqu'il est dans $\mathcal{I}$) est un multiple de D.

□

7.2.D Polynômes premiers entre eux

Définition 7.2.8 On dit que deux polynômes non nuls P et Q sont premiers entre eux lorsque $P \wedge Q = 1$. Cela revient à dire que seules les constantes non nulles divisent à la fois P et Q.

Exemples.

1. Si $P \neq Q$ sont deux polynômes unitaires irréductibles, alors P et Q sont premiers entre eux.

   En effet, supposons par l'absurde qu'il y a un diviseur commun D non constant à ces deux polynômes. En divisant au besoin par le coefficient dominant de D, on peut supposer D unitaire. Il existe alors deux polynômes R et S dans $\mathbb{K}[X]$ tels que $P = DR$ et $Q = DS$. Comme P et Q sont irréductibles et que D n'est pas constant, les polynômes R et S sont forcément constants. Comme P, Q et D sont unitaires, on voit directement que $R = S = 1$ et donc $P = Q = D$. Ceci contredit l'hypothèse.
2. En particulier, si $\lambda \neq \mu$ sont deux nombres, alors les polynômes $(X-\lambda)$ et $(X-\mu)$ sont premiers entre eux.

Le théorème de Bezout ci-dessous permet de caractériser les polynômes premiers entre eux.

Théorème 7.2.9 (de Bezout) *Deux polynômes non nuls $P, Q \in \mathbb{K}[X]$ sont premiers entre eux si et seulement si il existe deux polynômes $U, V \in \mathbb{K}[X]$ tels que $PU + QV = 1$.*

Preuve. Supposons tout d'abord qu'il existe deux polynômes U et V tels que $PU+QV = 1$. Tout diviseur commun de P et de Q divise alors $PU+QV = 1$. Il est donc de degré 0. Tous les diviseurs communs de P et Q sont des constantes non nulles donc P et Q sont premiers entre eux.

Réciproquement, supposons que $P \wedge Q = 1$. Comme $P \wedge Q$ engendre l'idéal $\mathcal{I}$ introduit ci-dessus, on a $1 = P \wedge Q \in \mathcal{I}$. Il existe donc $U, V \in \mathbb{K}[X]$ tels que $PU+QV = 1$. □

Voyons des conséquences du théorème de Bezout.

Corollaire 7.2.10 *Soit A, B, C des polynômes non nuls.*

Lemme de Gauss : Si A est premier avec B et divise BC alors A divise C.

Lemme d'Euclide : A est premier avec le produit BC si et seulement si il est premier avec chacun des facteurs B et C.

Preuve.

Lemme de Gauss : Nous écrivons la relation de Bezout : $AU + BV = 1$ avec $U, V \in \mathbb{K}[X]$. Alors : $ACU + BCV = C$. Comme A divise à la fois BCV et ACU, alors A divise C.

Lemme d'Euclide : Supposons que A est premier avec le produit BC. Alors il existe deux polynômes U et V tels que $AU + BCV = 1$. En posant $V_1 = CV$ et $V_2 = BV$, on constate que $AU + BV_1 = 1 = AU + CV_2$. Ceci prouve que A est premier avec B et premier avec C. Supposons maintenant que A est premier avec chacun des facteurs B et C. On a alors deux relations de Bezout : $AU_1 + BV_1 = 1$ et $AU_2 + CV_2 = 1$. Il s'ensuit que

$$\begin{aligned} BC\underbrace{(V_1V_2)}_{=V} &= (BV_1)(CV_2) = (1 - AU_1)(1 - AU_2) \\ &= 1 - A\underbrace{(U_1 + U_2 - AU_1U_2)}_{=U} \end{aligned}$$

d'où $AU + BCV = 1$ et A est premier avec BC.

□

Remarque. Une conséquence du lemme d'Euclide est le fait suivant : si $\lambda \neq \mu$ sont deux nombres distincts dans $\mathbb{K}$ et $p, q \in \mathbb{N}$, alors $(X - \lambda)^p$ et $(X - \mu)^q$ sont premiers entre eux.

En effet, comme $(X - \lambda)$ est premier avec chacun des termes du produit $\underbrace{(X - \mu)\dots(X - \mu)}_{q \text{ fois}}$, alors $(X - \lambda) \wedge (X - \mu)^q = 1$. Maintenant, comme $(X-\mu)^q$ est premier avec chacun des termes du produit $\underbrace{(X - \lambda)\dots(X - \lambda)}_{p \text{ fois}}$, on a $(X - \lambda)^p \wedge (X - \mu)^q = 1$.

□

7.2.E Calcul pratique du plus grand diviseur commun

On donne ici un algorithme très simple pour déterminer le plus grand diviseur commun de deux polynômes. Cet algorithme est connu sous le nom d'algorithme d'Euclide et il est similaire à celui permettant de trouver le plus grand diviseur commun de deux entiers.

Soit A et B deux polynômes de $\mathbb{K}[X]$. On suppose que $\deg B \geq \deg A$.

– On fait la division euclidienne de B par A : il existe donc un polynôme $Q \in \mathbb{K}[X]$ et un polynôme $A_1 \in \mathbb{K}[X]$ tel que $\deg A_1 < \deg A$ et

$$B = AQ + A_1$$

– On fait la division euclidienne de A par A_1 : il existe donc un polynôme $Q_1 \in \mathbb{K}[X]$ et un polynôme $A_2 \in \mathbb{K}[X]$ tel que $\deg A_2 < \deg A_1$ et

$$A = A_1Q_1 + A_2$$

– ...
– On fait la division euclidienne de A_{n-1} par A_n : il existe donc un polynôme $Q_n \in \mathbb{K}[X]$ et un polynôme $A_{n+1} \in \mathbb{K}[X]$ tel que $\deg A_{n+1} < \deg A_n$ et
$$A_{n-1} = A_n Q + A_{n+1}$$

Comme les degrés des restes diminuent à chaque pas, on arrive nécessairement à un moment donnée à
– soit un reste nul, le plus grand diviseur commun est alors le polynôme unitaire proportionnlle au dernier reste non nul
– Soit à une constante et alors les deux polynômes sont premiers entre eux (le plus grand diviseur commun vaut donc 1).

Exercice. Prouver que cet algorithme calcule effectivement le plus grand diviseur commun.

Exemple. Considérons les polynômes X^3+X^2-X-1 et X^2-X-2. Pour calucler le PGCD de ces deux polynômes, on fait la division euclidienne de X^3+X^2-X-1 par X^2-X-2, on a :

$$X^3+X^2-X-1 = (X+2)(X^2-X-2)+(3X+3)$$

On fait ensuite la division euclidienne de (X^2-X-2) par $(3X+3)$, on obtient :
$$(X^2-X-2) = (\frac{1}{3}X - \frac{2}{3})(3X+3)$$

Donc le dernier reste non nul est $3X+3$. Le PGCD des deux polynômes est donc $X+1$.

7.2.F Décomposition en facteurs irréductibles

Théorème 7.2.11 *Soit P un polynôme unitaire de $\mathbb{K}[X]$ de degré supérieur ou égal à 1. Il existe des polynômes irréductibles unitaires $P_1, \ldots, P_m \in \mathbb{K}[X]$ tels que $P = P_1 P_2 \ldots P_m$.*

De plus, cette décomposition est unique (à l'ordre près des facteurs). Elle s'appelle la décomposition en facteurs irréductibles *de P.*

Preuve. Montrons tout d'abord l'existence de cette décomposition. Nous raisonnons par récurrence sur $d = \deg(P)$. Si $d = 1$, P est un polynôme de degré 1 donc irréductible ; il suffit de prendre $m = 1$ et $P_1 = P$.

Supposons le résultat acquis pour tous les polynômes de degré $< d$ et soit P un polynôme de degré d. Si P est irréductible, nous prenons $m = 1$ et $P_1 = P$. Si P n'est pas irréductible, il est composé donc il existe deux polynômes Q et R non constants tels que $P = QR$. Comme Q et R sont de degrés ≥ 1, on a $1 \leq \deg(Q), \deg(R) < d$. L'hypothèse de récurrence s'applique donc à Q et à R. Il existe alors des polynômes unitaires irréductibles $P_1, \ldots, P_m$

et $P_{m+1}, \ldots, P_r$ tels que $Q = P_1 \ldots P_m$ et $R = P_{m+1} \ldots P_r$. On a alors $P = P_1 \ldots P_r$ et le résultat est acquis.

Montrons maintenant l'unicité. Nous nous donnons deux décompositions possibles $P_1 \ldots P_m = Q_1 \ldots Q_n$ où les P_k et les Q_j sont des polynômes unitaires irréductibles et nous cherchons à prouver que $n = m$ et les Q_j sont les P_k (éventuellement listés dans un autre ordre). Nous raisonnons par récurrence sur m.

Supposons $m = 1$ et donc $P_1 = Q_1 \ldots Q_n$. Si P_1 était différent de chacun des Q_j, il serait premier avec chacun d'entre eux (nous avons déjà prouvé que des polynômes unitaires irréductibles distincts étaient premiers entre eux) et il serait alors premier avec leur produit d'après le lemme d'Euclide. Nous en déduirions que P_1 est premier avec $Q_1 \ldots Q_n = P_1$. C'est absurde ! Par conséquent, P_1 est l'un des Q_j. Quitte à changer l'ordre des Q_j, nous pouvons supposer que $Q_1 = P_1$. Si $n \geq 2$, il vient alors $\deg(P_1) = \deg(P_1) + \deg(Q_2 \ldots Q_n) > \deg(P_1)$. C'est absurde donc $n = 1$ et le résultat est acquis au rang $m = 1$.

Supposons $m > 1$ et le résultat démontré au rang $m - 1$. Si le polynôme P_m ne se trouvait pas dans la liste $Q_1, \ldots, Q_n$, il serait premier avec chacun des Q_j et donc avec leur produit (cf raisonnement précédent). Nous aurions donc P_m premier avec $Q_1 \ldots Q_n = P_1 \ldots P_m$. C'est absurde et donc P_m est l'un des Q_j. Quitte à réécrire les Q_j dans un autre ordre, nous pouvons supposer que $P_m = Q_n$. En simplifiant par $P_m = Q_n$, il vient alors :

$$P_1 \ldots P_{m-1} = Q_1 \ldots Q_{n-1}.$$

Par hypothèse de récurrence, cela prouve que $m - 1 = n - 1$ et que les P_k sont les Q_j (à l'ordre près). Ceci achève la récurrence.

□

Dans l'écriture précédente, il est tout à fait possible qu'un polynôme irréductible unitaire apparaisse plusieurs fois dans la liste $P_1, P_2, \ldots, P_m$.

Remarques.

1. Si P n'est pas unitaire, il s'écrit $P = cP_1 \ldots P_m$ avec c son coefficient dominant et les P_j des polynômes unitaires irréductibles.
2. En regroupant les P_j qui sont répétés dans le produit $P_1 \ldots P_m$, on obtient une forme équivalente de la décomposition en facteurs irréductibles : soit P un polynôme de degré supérieur ou égal à 1, il existe alors une constante c, des polynômes unitaires irréductibles $P_1, \ldots, P_n$ distincts deux à deux et des entiers $\alpha_1, \ldots, \alpha_n \in \mathbb{N}$ tels que

$$P = c\, P_1^{\alpha_1} \ldots P_n^{\alpha_n}.$$

Notons que l'on peut rajouter artificiellement des polynômes à la liste des P_j en prenant des puissances α_j nulles.

3. Enfin, on convient (comme d'habitude) que $P_j^0 = 1$ et donc les polynômes constants s'écrivent aussi sous la forme précédente avec c la constante et les α_j nuls.

Corollaire 7.2.12 *1. Tout polynôme complexe s'écrit de façon unique (à l'ordre près des facteurs) sous la forme :*

$$P(X) = c\prod_{j=1}^{n}(X - a_j).$$

2. Tout polynôme réel s'écrit de façon unique (à l'ordre près des facteurs) sous la forme :

$$P(X) = c\prod_{j=1}^{k}(X - a_j)\prod_{i=1}^{m}(X^2 + b_iX + c_i)$$

où les a_j, b_i, c_i sont des réels vérifiant $b_i^2 - 4c_i < 0$ pour tout i.

Preuve. Ce n'est qu'une réécriture du théorème précédent en tenant compte de la liste des polynômes irréductibles dans $\mathbb{R}[X]$ et $\mathbb{C}[X]$ énumérée aux théorèmes 7.2.4 et 7.2.5.

□

Proposition 7.2.13 *Soit P et Q deux polynômes non nuls. Soit $P = aP_1^{\alpha_1}\dots P_n^{\alpha_n}$ et $Q = bP_1^{\beta_1}\dots P_n^{\beta_n}$ leurs décompositions en facteurs irréductibles. Alors P divise Q si, et seulement si, pour chaque j, on a $\alpha_j \leq \beta_j$.*

Preuve. Tout d'abord il est licite de supposer que les mêmes P_j sont utilisés pour P et pour Q. En effet, on peut éventuellement rajouter dans Q les P_j utilisés dans P (avec puissance nulle) et inversement.

Si, pour chaque j, on a $\alpha_j \leq \beta_j$, alors il est clair que P divise Q. Réciproquement, si P divise Q, alors il existe un polynôme R tel que $Q = PR$. Les polynômes irréductibles qui divisent R divisent également Q donc la décomposition en facteurs irréductibles de R n'utilise que les polynômes $P_1, \dots, P_n$:

$$R = c\,P_1^{\gamma_1}\dots P_n^{\gamma_n}$$

avec c une constante et $\gamma_j \in \mathbb{N}$. On a alors :

$$Q = PR = (ac)P_1^{\alpha_1+\gamma_1}\dots P_n^{\alpha_n+\gamma_n}.$$

Par unicité de la décomposition en facteurs irréductibles, il vient

$$ac = b \text{ et } \forall j, \quad \alpha_j + \gamma_j = \beta_j$$

et donc en particulier $\alpha_j \leq \beta_j$.

□

7.3 Équations polynomiales

7.3.A Multiplicités des racines

Définition 7.3.1 On dit qu'un nombre $\lambda \in \mathbb{K}$ est racine du polynôme $P \in \mathbb{K}[X]$ avec *multiplicité* m si $(X-\lambda)^m$ divise P mais que $(X-\lambda)^{m+1}$ ne divise pas P. En d'autres termes, la multiplicité m d'une racine est le plus grande puissance de $(X-\lambda)$ que l'on puisse mettre en facteur dans P.

Théorème 7.3.2 *Le nombre $\lambda \in \mathbb{K}$ est une racine du polynôme P de multiplicité m si et seulement si*

$$P(\lambda) = 0, \ldots, P^{(m-1)}(\lambda) = 0 \text{ et } P^{(m)}(\lambda) \neq 0.$$

Preuve. Quitte à remplacer l'étude du polynôme $P(X)$ par celle du polynôme $P(X+\lambda)$, on peut supposer que la racine étudiée est 0. Nous supposons donc $\lambda = 0$ et nous posons $P(X) = \sum_{k=0}^{n} a_k X^k$.

Supposons tout d'abord que $P(0) = \cdots = P^{(m-1)}(0) = 0$ et $P^{(m)}(0) \neq 0$ D'après la formule de Mac-Laurin, nous avons alors $P(X) = \sum_{k=0}^{n} \frac{P^{(k)}(0)}{k!} X^k = \sum_{k=m}^{n} \frac{P^{(k)}(0)}{k!} X^k = X^m Q(X)$ avec $Q(0) \neq 0$. On voit donc bien que X^m divise P. Comme $Q(0) \neq 0$, X ne divise par Q et donc X^{m+1} ne divise pas P.

Supposons maintenant que 0 soit racine de P avec multiplicité m. Alors $P(X) = X^m Q(X)$ avec $Q(0) \neq 0$ puisque X^{m+1} ne divise pas P. Alors Q est de degré $n-m$ et s'écrit $Q(X) = b_0 + b_1 X + \cdots + b_{n-m} X^{n-m}$ avec $b_0 \neq 0$. Alors $P(X) = b_0 X^m + b_1 X^{m+1} + \cdots + b_{n-m} X^n$, ce qui avec la formule de Mac-Laurin, montre que $P(0) = \cdots = P^{(m-1)}(0) = 0$ et $P^{(m)}(0) \neq 0$.

□

Remarque. Il existe deux méthodes pour compter les racines d'un polynôme P :

- Racines distinctes : chaque racine compte une fois. Il y a alors autant de racines qu'il existe de nombres λ distincts tels que $P(\lambda) = 0$.
- Avec multiplicité : on recompte chaque racine suivant sa multiplicité (si une racine est double, on la compte deux fois etc).

Par exemple, le polynôme $(X-1)(X-7)^2$ a deux racines distinctes (les nombres 1 et 7) mais il a trois racines comptées avec multiplicités (les nombres 1, 7 et 7 à nouveau).

Théorème 7.3.3 (Nombre de racines) *Un polynôme non nul de degré n a au plus n racines comptées avec multiplicités.*

Preuve. Soit P un polynôme non nul de degré n. On énumère les racines distinctes de P que nous avons trouvées : ce sont les nombres λ_1, …, λ_p avec multiplicités $m_1, \ldots, m_p$. Le nombre de racines de P comptées avec multiplicités est égal à $m = m_1 + \cdots + m_p$. Nous savons que P s'écrit $(X-\lambda_1)^{m_1}Q_2(X)$. Comme $(X-\lambda_2)^{m_2}$ est premier avec $(X-\lambda_1)^{m_1}$ et divise $P = (X-\lambda_1)^{m_1}Q_2(X)$, le lemme de Gauss assure que $(X-\lambda_2)^{m_2}$ divise $Q_2(X)$. On peut donc factoriser également $(X-\lambda_2)^{m_2}$. En recommençant, on finit par écrire $P = (X-\lambda)^{m_1}\ldots(X-\lambda_p)^{m_p}Q(X)$ où Q est un polynôme non nul puisque P est non nul. Dès lors, il est évident que $n = \deg(P) = m + \deg(Q) \geq m$. Le polynôme P de degré n a donc au plus n racines comptées avec multiplicités.

□

Définition 7.3.4 On dit qu'un polynôme P est *scindé sur* $\mathbb{K}$ s'il peut s'écrire sous la forme

$$P(X) = c\prod_{i=1}^{p}(X-\lambda_i)^{m_i}$$

avec $c \in \mathbb{K}$, $\lambda_i \in \mathbb{K}$ et $m_i \in \mathbb{N}^*$.

- Dans l'écriture précédente, c est le coefficient dominant de P, les λ_i sont ses racines et les m_i sont les multiplicités.
- Le polynôme nul est scindé sur $\mathbb{K}$ (prendre $c = 0$). Plus généralement, toutes les constantes sont scindées sur $\mathbb{K}$ (prendre c cette constante et $p = 0$: un produit vide vaut 1).
- Un polynôme non nul est scindé sur $\mathbb{K}$ si et seulement si il a autant de racines comptées avec multiplicité que son degré. Un polynôme est donc scindé sur $\mathbb{K}$ s'il "ne manque aucune racine dans $\mathbb{K}$".
- D'après le corollaire 7.2.12, tous les polynômes sont scindés sur $\mathbb{C}$. Ce n'est pas toujours le cas sur $\mathbb{R}$. Par exemple, le polynôme X^2+1 n'a aucune racine dans $\mathbb{R}$: il n'est pas scindé sur $\mathbb{R}$.

Remarque. Si un polynôme $P \neq 0$ est scindé sur $\mathbb{K}$, alors tout diviseur D de P est également scindé sur $\mathbb{K}$. En effet, nous écrivons $P = DQ$ avec $Q \in \mathbb{K}[X]$. Notons $p = \deg(P)$, $d = \deg(D)$ et $q = \deg(Q)$. Nous avons $p = d+q$. Or P a p racines comptées avec multiplicités, Q en a au plus q (théorème précédent) donc D en a au moins d. D'après le théorème précédent, il ne peut pas en avoir plus de d donc il en a exactement d. Ceci prouve que D est scindé sur $\mathbb{K}$.

Le calcul du plus grand diviseur commun de deux polynômes complexes est aisé. Soit P et Q deux polynômes complexes non constants. Nous nommons $\lambda_1, \ldots, \lambda_k$ les nombres complexes qui sont racines **à la fois** de P et de Q. Soit $\alpha_1, \ldots, \alpha_k$ les multiplicités de ces racines pour P et $\beta_1, \ldots, \beta_k$ les multiplicités de ces racines pour Q. Alors :

$$(P \wedge Q)(X) = (X-\lambda_1)^{\min(\alpha_1,\beta_1)}\ldots(X-\lambda_k)^{\min(\alpha_k,\beta_k)}.$$

En particulier, nous retenons que deux polynômes complexes sont premiers entre eux si, et seulement si, ils n'ont aucune racine en commun. Ce dernier résultat tombe en défaut dans $\mathbb{R}$ puisque X^2+1 et $(X^2+1)^2$ n'ont aucune racine en commun sur $\mathbb{R}$ (ils n'ont aucune racine) mais ils ne sont pas premiers entre eux puisque le second est le carré du premier.

Exemple.

$$P(X) = (X-1)(X-i)^3(X-2i)^4 \text{ et } Q(X) = (X-3)(X-i)^7(X-2i)$$

Les racines communes sont i et $2i$, les multipicités de i sont 3 et 7 tandis que celles de $2i$ sont 4 et 1. On garde la plus petite multiplicité à chaque fois et il vient :

$$P \wedge Q(X) = (X-i)^3(X-2i).$$

Corollaire 7.3.5 *Si deux polynômes P et Q de degrés inférieurs ou égaux à n prennent les mêmes valeurs en $n+1$ points, alors $P = Q$.*

Preuve. Soit $R = P - Q$ et $d = \deg(R)$. Nous avons $d \leq n$. Si R est non nul, alors R possède au plus d racines. Mais nous avons justement supposé qu'il en possédait au moins $n+1 > d$. C'est absurde donc $R = 0$ et $P = Q$. □

7.4 Lemme des noyaux

Nous allons maintenant utiliser les résultats vus dans ce chapitre pour prouver une propriété fondamentale.

Définition 7.4.1 Des polynômes $Q_1, \ldots Q_k$ sont dits *premiers entre eux deux à deux* si, pour tout $i \neq j$, les polynômes Q_i et Q_j sont premiers entre eux. Il revient au même de dire que chaque polynôme Q_i est premier avec le produit de tous les autres $Q_1 \ldots Q_{i-1}Q_{i+1} \ldots Q_k$ (lemme d'Euclide).

Théorème 7.4.2 (Lemme des noyaux) *Soit E un $\mathbb{K}$-espace vectoriel et $u \in L(E)$. Soit $Q_1, \ldots, Q_k \in \mathbb{K}[X]$ des polynômes et $Q = Q_1 \ldots Q_k$ leur produit. Nous posons $q = Q(u)$, $q_1 = Q_1(u)$, …, $q_k = Q_k(u)$. Si les polynômes $Q_1, \ldots, Q_k$ sont premiers entre eux deux à deux, alors*

$$\mathrm{Ker}(q) = \bigoplus_{i=1}^{k} \mathrm{Ker}(q_i).$$

En particulier, si $Q(u) = 0$, alors $E = \mathrm{Ker}(q_1) = \oplus \cdots \oplus \mathrm{Ker}(q_k)$.

Preuve. Nous prouvons le résultat par récurrence sur k. Commençons avec $k = 2$.

Tout d'abord, $\mathrm{Ker}(q_1) \subset \mathrm{Ker}(q)$ et $\mathrm{Ker}(q_2) \subset \mathrm{Ker}(q)$. En effet, si $x \in \mathrm{Ker}(q_1)$, alors $q(x) = q_2 q_1(x) = q_2(0) = 0$ donc $x \in \mathrm{Ker}(q)$.

Comme Q_1 et Q_2 sont premiers entre eux, en utilisant le théorème de Bezout, nous voyons qu'il existe des polynômes $V_1, V_2 \in \mathbb{K}[X]$ tels que $Q_1V_1 + Q_2V_2 = 1$. En substituant X par u dans cette relation et en posant $v_1 = V_1(u)$, $v_2 = V_2(u)$, nous trouvons :

$$q_1v_1 + q_2v_2 = \mathrm{id}_E \text{ et donc} x = q_1v_1(x) + q_2v_2(x) \quad \text{pour tout } x \in E.$$

Montrons que la somme $\mathrm{Ker}(q_1) + \mathrm{Ker}(q_2)$ est directe. Pour cela, nous considérons $x \in \mathrm{Ker}(q_1) \cap \mathrm{Ker}(q_2)$. Alors : $x = v_1 \underbrace{q_1(x)}_{=0} + v_2 \underbrace{q_2(x)}_{=0} = 0 + 0 = 0$. Nous avons prouvé $\mathrm{Ker}(q_1) \cap \mathrm{Ker}(q_2) = \{0\}$ donc la somme est directe.

Montrons que $\mathrm{Ker}(q) = \mathrm{Ker}(q_1) \oplus \mathrm{Ker}(q_2)$. Prenons $x \in \mathrm{Ker}(q)$. Posons $x_1 = q_2v_2(x)$ et $x_2 = q_1v_1(x)$. Nous avons : $q_1(x_1) = (q_1q_2v_2)(x) = v_2q(x) = 0$ et $q_2(x_2) = (q_2q_1v_1)(x) = v_1q(x) = 0$. Par conséquent $x_1 \in \mathrm{Ker}(q_1)$ et $x_2 \in \mathrm{Ker}(q_2)$. Enfin, $x_1 + x_2 = q_2v_2(x) + q_1v_1(x) = x$. Ceci prouve que tout vecteur de $\mathrm{Ker}(q)$ admet une décomposition selon la somme $\mathrm{Ker}(q_1) \oplus \mathrm{Ker}(q_2)$ d'où le résultat annoncé.

Prouvons maintenant l'hérédité de la récurrence. Nous supposons le résultat établi au rang k. Soit $Q_1, \ldots, Q_{k+1}$ des polynômes premiers entre eux deux à deux. Nous savons que Q_{k+1} est premier avec le produit $P = Q_1 \ldots Q_k$ donc, en utilisant le résultat au rang 2, nous avons : $\mathrm{Ker}(q) = \mathrm{Ker}(P(u)) \oplus \mathrm{Ker}(q_{k+1})$. Par ailleurs, d'après notre hypothèse de récurrence au rang k, nous avons $\mathrm{Ker}(P(u)) = \mathrm{Ker}(q_1) \oplus \cdots \oplus \mathrm{Ker}(q_k)$. Finalement, nous avons donc :

$$\mathrm{Ker}(q) = \mathrm{Ker}(q_1) \oplus \cdots \oplus \mathrm{Ker}(q_k) \oplus \mathrm{Ker}(q_{k+1}) = \bigoplus_{i=1}^{k+1} \mathrm{Ker}(q_i).$$

□

Exemples.

1. Considérons, dans un $\mathbb{K}$-espace vectoriel E, un projecteur p d'image F et de noyau G. Nous avons $p^2 = p$. Posons $Q(X) = X^2 - X$. Nous pouvons écrire $Q(X) = X(X-1)$. Nous avions vu précédemment que X et $X-1$ sont premiers entre eux. Par ailleurs, le polynôme X appliqué sur p donne tout simplement p et le polynôme $X-1$ appliqué sur p donne $p - \mathrm{id}_E$. Enfin, $Q(p) = p^2 - p = 0$. Par conséquent, le lemme des noyaux donne :

   $$E = \mathrm{Ker}(Q(p)) = \mathrm{Ker}(p) \oplus \mathrm{Ker}(p - \mathrm{id}_E).$$

   Vous pouvez vérifier en guise d'exercice que $\mathrm{Ker}(p) = G$ et $\mathrm{Ker}(p - \mathrm{id}_E) = F$. On retrouve donc $E = F \oplus G$.

2. Considérons, dans un $\mathbb{K}$-espace vectoriel E, une symétrie s par rapport à F parallèlement à G. Nous avons $s^2 = \mathrm{id}_E$. Posons $Q(X) = X^2 - 1$. Nous pouvons écrire $Q(X) = (X-1)(X+1)$. Les polynômes $X-1$ et $X+1$ sont premiers entre eux. Par ailleurs, le polynôme $X-1$ appliqué à s donne $s - \mathrm{id}_E$ et le polynôme $X+1$ appliqué à s donne $s + \mathrm{id}_E$. Enfin $Q(s) = s^2 - \mathrm{id}_E = 0$ donc le lemme des noyaux donne :

$$E = \mathrm{Ker}(Q(s)) = \mathrm{Ker}(s - \mathrm{id}_E) \oplus \mathrm{Ker}(s + \mathrm{id}_E).$$

Là encore, vous pouvez vérifier en exercice que $\mathrm{Ker}(s - \mathrm{id}_E) = F$ et $\mathrm{Ker}(s + \mathrm{id}_E) = G$ donc on retrouve $E = F \oplus G$.

Chapitre 8

Sous-espaces stables d'un endomorphismes

L'objet de ce chapitre est d'étudier les sous-espaces stables d'un endomorphisme et leur intérêt. Nous introduisons également des objets fondamentaux associés aux applications linéaires : les valeurs propres et sous-espaces propres et nous étudions les propriétés de ces objets.

8.1 Généralités sur les sous-espaces stables

Définition 8.1.1 Soit E un $\mathbb{K}$-espace vectoriel et $u \in L(E)$ un endomorphisme. Un sous-espace vectoriel F est dit *stable* par u lorsque $u(F) \subset F$. On peut alors définir l'endomorphisme $\hat{u}$ induit par u sur F. Il est obtenu par restriction de u à F au départ et à l'arrivée. Il est défini tout simplement par :

$$\forall x \in F, \quad \hat{u}(x) = u(x).$$

Le procédé semble évident mais il faut bien se rendre compte de l'intérêt de la notion. Comme $u(F) \subset F$, les images $u(x)$ pour $x \in F$ sont dans F et pas simplement dans E comme c'est le cas en général. Par conséquent, $\hat{u}$ est un endomorphisme de F (de F dans lui-même).

Exemples

1. Soit $u : \mathbb{K}^2 \to \mathbb{K}^2$ qui au couple (x_1, x_2) associe le couple $(x_1 + x_2, 2x_1)$. Soit F la diagonale $F = \{(t,t) \ ; \ t \in \mathbb{K}\}$. Pour $x = (t,t) \in F$, on a $u(x) = (2t, 2t) = 2x \in F$. On a donc $u(F) \subset F$ et le sous-espace F est stable par u. L'endomorphisme induit par u sur F est $\widehat{u} = 2\,\mathrm{id}_F$.
2. Soit $E = \mathcal{C}^\infty(\mathbb{R})$ le $\mathbb{R}$-espace vectoriel des fonctions de classe C^∞ sur $\mathbb{R}$. Soit d l'endomorphisme de E qui à une fonction f associe sa dérivée f'. Soit enfin F le sous-espace formé des combinaisons linéaires des

fonctions $t \mapsto e^{\lambda t}$ avec $\lambda \in \mathbb{R}$:

$$F = \left\{ t \mapsto \sum_{i=1}^{n} \alpha_i e^{\lambda_i t} \ ; \ n \in \mathbb{N},\ \alpha_i \in \mathbb{R},\ \lambda_i \in \mathbb{R} \right\}$$

Comme $\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=1}^{n} \alpha_i e^{\lambda_i t}\right) = \sum_{i=1}^{n} \lambda_i \alpha_i e^{\lambda_i t}$, on voit que $d(F) \subset F$. Par conséquent, le sous-espace F est stable par d.

Parmi les sous-espaces stables d'un endomorphisme u, on trouve toujours $\{0\}$ et E tout entier. C'est évident (et sans intérêt) et on les nomme *sous-espaces stables triviaux*. Sont également stables $\mathrm{Ker}(u)$ et $\mathrm{Im}(u)$.

Proposition 8.1.2 (Expression matricielle d'un sous-espace stable) *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$. Soit F un sous-espace stable non trivial de u. Soit $(e_1, \ldots, e_p)$ une base de F complétée en une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E. Alors la matrice de u dans la base $\mathcal{B}$ est triangulaire supérieure par blocs : $Mat_{\mathcal{B}}(u) = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)$.*

Preuve. La preuve est évidente puisque les images des vecteurs $e_1, \ldots, e_p$ sont dans F donc ont des coordonnées nulles sur les vecteurs $e_{p+1}, \ldots, e_n$. □

Parmi les sous-espaces stables, les droites stables sont particulièrement intéressantes.

8.2 Valeurs propres et vecteurs propres

8.2.A Cas d'un endomorphisme

Dans toute cette partie, E est un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ est un endomorphisme.

Définition 8.2.1 Un vecteur non nul $x \in E$ est un *vecteur propre* de u si $u(x)$ est colinéaire à x ce qui, compte tenu du fait que $x \neq 0$, revient à dire qu'il existe un nombre $\lambda \in \mathbb{K}$ appelé *valeur propre* tel que $u(x) = \lambda x$.

Quelques remarques importantes :
- Il est capital dans la définition ci-dessus que le vecteur x soit non nul. En effet, $u(0) = 0$ est bien entendu colinéaire à 0 mais 0 n'est pas un vecteur propre.
- A un vecteur propre x donné est associée une seule valeur propre λ. En effet, si λ' est une autre valeur propre pour le même vecteur propre, alors $u(x) = \lambda x = \lambda' x$, ce qui implique $\lambda' = \lambda$ puisque $x \neq 0$.

- A une valeur propre λ donnée sont toujours associés plusieurs vecteurs propres. En effet, si $x \neq 0$ vérifie $u(x) = \lambda x$, alors $u(2x) = 2u(x) = 2\lambda x$ donc $2x$ est aussi vecteur propre et il est distinct de x puisque $x \neq 0$.
- Si 0 est valeur propre de u alors il existe $x \neq 0$ tel que $u(x) = 0$, le noyau de u est donc non réduit à 0 et u est non inversible. Réciproquement, si u est non inversible, il existe $x \neq 0$ tel que $u(x) = 0$ et on conclut que x est vecteur propre associé à la valeur propre 0 de u. On a donc

$$u \text{ inversible} \iff 0 \text{ valeur propre}$$

Définition 8.2.2 On appelle *spectre* de u et on note $\mathrm{Sp}(u)$ l'ensemble des nombres $\lambda \in \mathbb{K}$ qui sont des valeurs propres de u.

Exemples

1. Il est clair que $\mathrm{Sp}(\mathrm{id}_E) = \{1\}$ et que tout vecteur non nul est un vecteur propre de id_E. De même, il est clair que $\mathrm{Sp}(0) = \{0\}$ et, plus généralement, $\mathrm{Sp}(\lambda\,\mathrm{id}_E) = \{\lambda\}$.
2. Dans le plan, soit u la rotation d'angle $\theta \in]0, \pi[$. Pour $x \neq 0$, le vecteur $u(x)$ n'est jamais colinéaire à x donc il n'y a aucun vecteur propre et aucune valeur propre : $\mathrm{Sp}(u) = \emptyset$.

Comme nous l'avons déjà signalé, à une valeur propre $\lambda \in \mathbb{K}$ donnée, on associe plusieurs vecteur propres. En fait, l'ensemble des vecteurs propres de u associés à une valeur propre λ donnée est un sous-espace vectoriel de E auquel on a retiré 0. En effet :

$$u(x) = \lambda x \iff (u - \lambda\,\mathrm{id}_E)(x) = 0 \iff x \in \mathrm{Ker}(u - \lambda\,\mathrm{id}_E).$$

Ainsi, $\mathrm{Ker}(u - \lambda\,\mathrm{id}_E)$ est la réunion du vecteur nul et des vecteurs propres de u associés à λ.

Définition 8.2.3 Soit $\lambda \in \mathrm{Sp}(u)$ une valeur propre de u. On appelle sous-espace propre de u associé à λ et on note $E_\lambda(u)$, ou plus simplement E_λ lorsqu'il n'y a pas de confusion possible, le sous-espace vectoriel $E_\lambda = \mathrm{Ker}(u - \lambda\,\mathrm{id}_E)$.

- Remarquons que les sous-espaces propres de u sont stables par u. En effet, si $x \in E_\lambda$, alors $u(x) = \lambda x$ donc $u(u(x)) = u(\lambda x) = \lambda u(x)$. Ceci prouve que $u(x) \in E_\lambda$.
- Un vecteur non nul x est un vecteur propre de u si et seulement si la droite $\mathrm{Vect}(x)$ est stable par u.

Nous terminons ce paragraphe par un résultat important : les sous-espaces propres d'un endomorphisme sont toujours en somme directe.

Théorème 8.2.4 (Somme directe des sous-espaces propres) *Soit E un $\mathbb{K}$-espace vectoriel de dimension n, soit $u \in L(E)$ un endomorphisme et soit $\lambda_1, \ldots, \lambda_p$ des valeurs propres distinctes de u.*

– *Les sous-espaces propres $E_{\lambda_1}, \ldots, E_{\lambda_p}$ sont en somme directe. On note donc :*
$$\sum_{\lambda \in \mathrm{Sp}(u)} E_\lambda = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda$$

– *Si x_1 est un vecteur propre associé à λ_1, ..., x_p est un vecteur propre associé à λ_p, alors $(x_1, \ldots, x_p)$ est une famille libre.*

Preuve. Soir P le polynôme $P(X) = (X - \lambda_1) \ldots (X - \lambda_p)$. Puisque les valeurs propres λ_i sont distinctes, nous savons que les polynômes $(X - \lambda_1), \ldots, (X - \lambda_p)$ sont premiers entre eux deux à deux. Nous pouvons donc appliquer le lemme des noyaux et il vient :

$$\mathrm{Ker}(P(u)) = \mathrm{Ker}(u - \lambda_1 \mathrm{id}_E) \oplus \cdots \oplus \mathrm{Ker}(u - \lambda_p \mathrm{id}_E) = \bigoplus_{i=1}^{p} E_{\lambda_i}.$$

En particulier les sous-espaces propres sont en somme directe.

Montrons maintenant la seconde assertion. Soit x_1 un vecteur propre associé à la valeur propre λ_1, ..., x_p un vecteur propre associé à la valeur propre λ_p. Pour prouver que la famille $(x_1, \ldots, x_p)$ est libre nous nous donnons une combinaison linéaire nulle $\sum_{i=1}^{p} \mu_i x_i$ et nous montrons qu'elle est triviale (c'est-à-dire que tous les μ_i sont nuls).

Le vecteur 0 est un vecteur de la somme des sous-espaces vectoriels E_{λ_i} et la combinaison linéaire précédente nous en donne une décomposition :

$$0 = \mu_1 x_1 + \cdots + \mu_p x_p$$

avec $\mu_i x_i \in E_{\lambda_i}$ pour tout $i \in \{1, ..., p\}$. Comme nous avons également la décomposition $0 = 0 + \cdots + 0$ et que la décomposition est unique (car la somme des sous-espace vectoriel est directe), nous en déduisons que pour chaque i, on a $\mu_i x_i$. Or x_i est un vecteur propre donc c'est un vecteur non nul. Par conséquent, $\mu_i = 0$. C'est ce qu'on voulait prouver.

□

8.2.B Cas d'une matrice

Soit E un $\mathbb{K}$-espace vectoriel de dimension n rapporté à une base $\mathcal{B}$. Soit $u \in L(E)$ et A la matrice de u dans la base $\mathcal{B}$. Soit $x \in E$ et X le vecteur colonne des coordonnées de x dans la base $\mathcal{B}$. Les coordonnées de $u(x)$ sont données par le vecteur colonne AX. Par conséquent, il est clair que $x \neq 0$ est un vecteur propre pour la valeur propre λ si et seulement si le vecteur colonne $X \neq 0$ vérifie $AX = \lambda X$. La définition suivante est donc légitime.

Définition 8.2.5 Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée.

1. Un nombre $\lambda \in \mathbb{K}$ est une valeur propre de A s'il existe un vecteur colonne non nul X tel que $AX = \lambda X$.
2. Le spectre de A est l'ensemble de ses valeurs propres.

On fera attention à ne pas confondre matrices et endomorphismes (une matrice peut représenter plusieurs endomorphismes dans des bases différentes et un endomorphisme possède plusieurs matrices suivant la base choisie). En particulier, on ne parle pas de vecteur propre d'une matrice ni de sous-espace propre d'une matrice. C'est un abus de langage (que vous ferez sûrement par la suite mais que je vous déconseille formellement de faire tant que vous débutez dans ce domaine) que de parler de noyau, d'image ou de sous-espace propre d'une matrice. Rigoureusement, ces objets seraient définis dans le $\mathbb{K}$-espace vectoriel $\mathcal{M}_{n1}(\mathbb{K})$...

8.3 Polynôme caractéristique

8.3.A Polynôme caractéristique d'une matrice

Définition 8.3.1 Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. On appelle *polynôme caractéristique* de A le polynôme $P_A(X) = \det(A - XI_n)$.

Remarquons tout d'abord qu'il s'agit bien d'un polynôme :

$$P_A(X) = \begin{vmatrix} a_{11} - X & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - X & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - X \end{vmatrix}$$

Les termes a_{ij} de part et d'autre de la diagonale sont mis pour signifier que les termes en dehors de la diagonale sont ceux de la matrice A. Pour voir que P_A est un polynôme, on utilise la formule pour le déterminant : $P_A(X) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)\, a'_{1\sigma(1)} \dots a'_{n\sigma(n)}$ où les a'_{ij} sont les coefficients de la matrice $A - XI_n$, c'est-à-dire $a'_{ij} = a_{ij}$ pour $i \neq j$ et $a'_{ii} = a_{ii} - X$. Les coefficients a'_{ij} sont donc des polynômes en X de degré au plus 1. Par conséquent, comme ils apparaissent n par n dans la somme ci-dessus, on voit que $P_A(X)$ est un polynôme de degré au plus n. Pour obtenir un terme de degré n, il faut que chacun des termes $a'_{i\sigma(i)}$ vaille $a_{ii} - X$. Ceci n'est possible que pour la permutation $\sigma = \mathrm{id}$. De même, pour obtenir un terme de degré $n-1$, il faut que la permutation σ laisse fixe au moins $n-1$ des n nombres $1, \dots, n$. Mais si elle laisse fixes $n-1$ nombres, elle doit laisser fixe également le dernier. Par conséquent, le terme de degré n et le terme de degré $n-1$ s'obtiennent pour $\sigma = \mathrm{id}$. Le terme correspondant à $\sigma = \mathrm{id}$ est

$$a'_{11} \dots a'_{nn} = \prod_{i=1}^{n} (a_{ii} - X) = (-1)^n \prod_{i=1}^{n} (X - a_{ii}) = (-1)^n \left[X^n - \sum_{i=1}^{n} a_{ii} X^{n-1} + \dots \right]$$

Par conséquent, on en déduit que $P_A(X) = (-1)^n X^n + (-1)^{n-1}\operatorname{tr}(A)X^{n-1} + \dots$. Le terme constant de ce polynôme correspond à $P_A(0) = \det(A)$. Nous retenons :

Proposition 8.3.2 *Le polynôme caractéristique de $A \in \mathcal{M}_n(\mathbb{K})$ est un polynôme de degré n de la forme*

$$P_A(X) = (-1)^n X^n + (-1)^{n-1}\operatorname{tr}(A)X^{n-1} + \cdots + \det(A).$$

Exemples.

1. Soit A une matrice triangulaire. La matrice $A - XI_n$ est elle aussi triangulaire donc $P_A(X) = (a_{11} - X)\dots(a_{nn} - X)$.
2. On appelle *matrice de Froebenius* toute matrice de la forme suivante :

$$F = \begin{pmatrix} 0 & & & & a_0 \\ & \ddots & & & \vdots \\ & & 0 & & \\ 1 & & & & a_1 \\ & & & & \\ & & & & \\ & & & 0 & a_{n-2} \\ & 0 & & & \\ & & & 1 & a_{n-1} \end{pmatrix}$$

On a alors :

$$P_F(X) = \begin{vmatrix} -X & & & & a_0 \\ & \ddots & & & \vdots \\ & & 0 & & \\ 1 & & & & a_1 \\ & & & & \\ & & & & \\ & & & -X & a_{n-2} \\ & 0 & & & \\ & & & 1 & a_{n-1} - X \end{vmatrix}$$

Dans F, la diagonale contient des 0 sauf le terme a_{n-1}. La sous-diagonale contient des 1.

Nous montrons par récurrence sur n que $P_F(X) = (-1)^n\big[X^n - a_{n-1}X^{n-1} - \cdots - a_1X - a_0\big]$.

Pour $n = 2$, on a $F = \begin{pmatrix} 0 & a_0 \\ 1 & a_1 \end{pmatrix}$ donc $P_F(X) = X^2 - \operatorname{tr}(A)X + \det(A) = X^2 - a_1X - a_0$. Ceci initialise la récurrence.

Supposons la propriété démontrée au rang $n-1$. En développant le déterminant $\det(F - XI_n)$ par rapport à la première ligne, on obtient la somme de $(-1)^{n+1}a_0$ avec $-X$ fois le déterminant de Frobenius au rang $n-1$ (avec pour coefficients $a_1, \ldots, a_{n-1}$) . Par hypothèse de récurrence, il vaut $(-1)^{n-1}\big[X^{n-1} - a_{n-1}X^{n-2} - \cdots - a_1\big]$ donc

$$\begin{aligned} P_F(X) &= (-1)^n\big[X^n - a_{n-1}X^{n-1} - \cdots - a_1X\big] + (-1)^{n+1}a_0 \\ &= (-1)^n\big[X^n - a_{n-1}X^{n-1} - \cdots - a_1X - a_0\big]. \end{aligned}$$

□

Notons qu'on retrouve bien $\mathrm{tr}(F) = a_{n-1}$ et $\det(F) = a_0$ (développer par rapport à la première ligne).

L'intérêt du polynôme caractéristique repose dans le théorème suivant :

Théorème 8.3.3 (Valeurs propres et polynôme caractéristique) *Soit $A \in \mathcal{M}_n(\mathbb{K})$. Un nombre λ est valeur propre de A si et seulement si $P_A(\lambda) = 0$. En d'autres termes, les valeurs propres de A sont les racines du polynôme caractéristique de A.*

Preuve. Un nombre $\lambda \in \mathbb{K}$ est une valeur propre de A s'il existe un vecteur colonne X non nul tel que $AX = \lambda X$. Cela revient exactement à dire que la matrice $A - \lambda I_n$ n'est pas inversible donc que $\det(A - \lambda I_n) = 0$, c'est-à-dire $P_A(\lambda) = 0$.

□

- Nous avons une technique pour déterminer les valeurs propres d'une matrice A. Il "suffit" de calculer le déterminant $P_A(X)$ puis de résoudre l'équation polynômiale $P_A(X) = 0$.
- On fera attention au problème suivant : résoudre l'équation $P_A(X) = 0$ dépend du corps $\mathbb{K}$. En effet, prenons par exemple $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Nous avons $P_A(X) = \begin{vmatrix} -X & -1 \\ 1 & -X \end{vmatrix} = X^2 + 1$. Si on considère la matrice A comme un élément de $\mathcal{M}_2(\mathbb{R})$, alors on résout l'équation $X^2 + 1 = 0$ dans $\mathbb{R}$. Elle n'a aucune solution et la matrice A n'a donc aucune valeur propre. En revanche, si on considère A comme un élément de $\mathcal{M}_2(\mathbb{C})$, alors on résout l'équation $X^2 + 1 = 0$ dans $\mathbb{C}$. Elle a deux solutions, i et $-i$ donc A possède deux valeurs propres. La notion de valeur propre dépend donc du corps de base sur lequel on s'est placé.
- Si $A = (a_{ij})$ est une matrice triangulaire (ou *a fortiori* une matrice diagonale), alors il est clair que $P_A(X) = (a_{11} - X)\ldots(a_{nn} - X)$. Dans ce cas, les valeurs propres de A sont les coefficients diagonaux de A.
- Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ possède au plus n valeurs propres. En effet, le polynôme P_A étant de degré n, il possède au plus n racines. En particulier, le spectre est toujours un ensemble fini.

– Si $A \in \mathcal{M}_n(\mathbb{R})$ avec n impair, alors A possède au moins une valeur propre. En effet, le polynôme P_A a un degré impair et un coefficient dominant égal à $(-1)^n = -1$ donc $\lim_{X\to-\infty} P_A(X) = +\infty$ et $\lim_{X\to+\infty} P_A(X) = -\infty$. Par le théorème des valeurs intermédiaires, on voit alors que P_A s'annule.
– En revanche, pour n pair, il peut exister $A \in \mathcal{M}_n(\mathbb{R})$ sans valeur propre comme nous l'avons vu juste au-dessus.

Proposition 8.3.4 *Le polynôme caractéristique est un invariant de similitude, c'est-à-dire que deux matrices semblables ont le même polynôme caractéristique.*

Preuve. Soit $A, B \in \mathcal{M}_n(\mathbb{K})$, $P \in \mathrm{GL}_n(\mathbb{K})$ telles que $B = P^{-1}AP$. Alors

$$\begin{aligned} P_B(X) &= \det(P^{-1}AP - XI_n) \\ &= \det(P^{-1}(A - XI_n)P) \\ &= \det(P^{-1})\det(A - XI_n)\det(P) \\ &= P_A(X). \end{aligned}$$

□

8.3.B Polynôme caractéristique d'un endomorphisme

Dans ce paragraphe, E est un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ est un endomorphisme. Grâce à la proposition précédente, on voit que toutes les matrices de u ont le même polynôme caractéristique. On l'appelle polynôme caractéristique de l'endomorphisme u et on le note P_u. On a :

$$P_u(X) = \det(u - X\,\mathrm{id}_E).$$

Exactement comme pour les matrices, nous voyons que $\lambda \in \mathbb{K}$ est une valeur propre de u si et seulement si $P_u(\lambda) = 0$. En d'autres termes, les valeurs propres de u sont les racines du polynôme caractéristique de u. Les remarques précédentes sur les matrices sont valables sans changement pour les endomorphismes.

L'intérêt du polynôme caractéristique est clair. *A priori*, pour savoir si un nombre $\lambda \in \mathbb{K}$ est une valeur propre de u, il faut chercher un vecteur x non nul tel que $u(x) = \lambda x$. Grâce au polynôme caractéristique, c'est beaucoup plus simple : on calcule $P_u(X)$ et on résout dans $\mathbb{K}$ l'équation $P_u(X) = 0$.
Exemples.

1. Soit p un projecteur de rang r. Il existe une base $\mathcal{B}$ de E telle que $\mathcal{M}_{\mathcal{B}}(u) = J_r$. Il est alors clair (écrivez le déterminant $P_p(X)$) que $P_p(X) = (1 - X)^r(-X)^{n-r}$. Les valeurs propres de p sont donc 1 (répétée r fois) et 0 (répétée $n - r$ fois).

2. Soit $u \in L(E)$ un endomorphisme nilpotent. Nous avons vu qu'il existait une base de E dans laquelle la matrice de u est triangulaire supérieure stricte. Le déterminant $P_u(X)$ est donc triangulaire supérieur avec des $-X$ sur la diagonale. Par conséquent, $P_u(X) = (-X)^n$. L'unique valeur propre de u est 0 et elle est répétée n fois.
3. Si $s \in L(E)$ est une symétrie, alors il existe une base où la matrice de u est de la forme
   Ainsi $P_s(X) = (-1)^n(X-1)^p(1+X)^{n-p}$ et les valeurs propres de s sont 1 et -1.

Dans les exemples ci-dessus, on voit que certaines valeurs propres sont plusieurs fois racines de P_u. Nous introduisons donc la définition suivante :

Définition 8.3.5 Soit λ une valeur propre de u.
- La *multiplicité algébrique* de λ, notée α_λ, est le nombre de fois que λ est racine de P_u.
- La *multiplicité géométrique* de λ, notée β_λ, est la dimension du sous-espace propre E_λ.

Pourquoi introduire ces deux entiers ? Nous avons deux points de vue pour les valeurs propres. Primo, ce sont des nombres auxquels correspondent des sous-espaces propres. Il semble donc naturel de compter chaque valeur propre en fonction de la dimension du sous-espace propre correspondant : c'est la multiplicité géométrique β_λ. Secundo, les valeurs propres sont les racines du polynôme P_u. Il semble donc naturel de compter plusieurs fois une valeur propre lorsqu'elle est racine multiple de P_u : c'est la multiplicité algébrique α_λ.

Le théorème 8.3.3, très simple à prouver, établit néanmoins un lien "inattendu" entre les deux points de vue (algébrique via P_u et géométrique via les vecteurs propres). Tout le problème, c'est que ce lien n'est pas "parfait". Nous allons voir que, certes, les valeurs propres sont les racines de P_u mais les multiplicités algébrique et géométrique ne sont pas toujours les mêmes. Lorsque ce sont les mêmes, il en résulte une propriété très intéressante pour l'endomorphisme u. Ce sera l'objet du chapitre suivant. Mais n'anticipons pas et étudions les premières propriétés du polynôme caractéristique.

Proposition 8.3.6 *1. Soit $A \in \mathcal{M}_n(\mathbb{K})$. On a $P_A = P\,{}^tA$.*

2. Soit $u \in L(E)$ et F un sous-espace stable par u. On note $\hat{u}$ l'endomorphisme induit par u sur F. Alors le polynôme P_u est un multiple de $P_{\hat{u}}$.

3. Soit λ une valeur propre de u. Alors $1 \leq \beta_\lambda \leq \alpha_\lambda$.

Preuve.

1. On a

$$P_{{}^tA}(X) = \det({}^tA - XI_n) = \det\left({}^t(A - XI_n)\right) = \det(A - XI_n) = P_A(X).$$

2. On introduit une base $(e_1, \dots, e_p)$ de F que l'on complète en une base $\mathcal{B} = (e_1, \dots, e_n)$ de E. Dans cette base, la matrice de u est triangulaire supérieure par blocs de la forme $\left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)$, où A est la matrice de $\hat{u}$ dans la base $(e_1, \dots, e_p)$. Par conséquent, il vient :

$$\begin{aligned} P_u(X) &= \left|\begin{array}{c|c} A - XI_p & B \\ \hline 0 & C - XI_{n-p} \end{array}\right| \\ &= \det(A - XI_p)\det(C - XI_{n-p}) \\ &= P_{\hat{u}}(X)Q(X) \end{aligned}$$

   en posant $Q(X) = P_C(X)$ donc P_u est un multiple de $P_{\hat{u}}$.

3. E_λ est stable par u. Pour $x \in E_\lambda$, on a $u(x) = \lambda x$ donc l'endomorphisme induit $\hat{u}$ vaut $\lambda \operatorname{id}_{E_\lambda}$. Par conséquent, on a $P_{\hat{u}}(X) = (\lambda - X)^{\beta_\lambda}$. D'après le point précédent, nous savons que $P_u(X) = (\lambda - X)^{\beta_\lambda}Q(X)$. Par conséquent, la racine λ est présente au moins β_λ fois dans P_u, ce qui prouve que $\alpha_\lambda \geq \beta_\lambda$.

□

Chapitre 9

Trigonalisation et Diagonalisation

L'objet de ce chapitre est d'utiliser les sous-espaces stables d'un endomorphisme et en particulier les sous-espaces propres pour obtenir une matrice particulièrement simple de cet endomorphisme. Le but va donc être de trouver les bases adéquates dans laquelle écrire la matrice associée à l'endomorphisme.

9.1 Trigonalisation

Définition 9.1.1 1. Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ un endomorphisme. On dit que u est *trigonalisable* lorsqu'il existe une base $\mathcal{B}$ dans laquelle la matrice de u est triangulaire supérieure.

2. Une matrice carrée $M \in \mathcal{M}_n(\mathbb{K})$ est dite *trigonalisable* lorsqu'elle est semblable à une matrice triangulaire supérieure. Cela revient à dire qu'elle est la matrice d'un endomorphisme trigonalisable dans une certaine base ou encore qu'il existe une matrice inversible P telle que $P^{-1}MP$ soit triangulaire supérieure.

Exemples.

1. $\lambda \operatorname{id}_E$ est trigonalisable puisque sa matrice dans n'importe quelle base est diagonale. On remarque que $P_{\lambda \operatorname{id}_E}(X) = (\lambda - X)^n$ est un polynôme scindé sur $\mathbb{K}$.
2. Un projecteur p est toujours trigonalisable puisqu'il existe une base dans laquelle sa matrice est J_r qui est diagonale. On remarque que $P_p(X) = (-1)^n(X-1)^r X^{n-r}$ est scindé sur $\mathbb{K}$.
3. Une symétrie s est toujours trigonalisable puisqu'il existe une base dans laquelle sa matrice est de la forme $\left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & -I_q \end{array}\right)$ qui est diagonale.

On remarque que $P_s(X) = (-1)^n(X-1)^p(X+1)^q$ est scindé sur $\mathbb{K}$.

4. Un endomorphisme nilpotent u est toujours trigonalisable puisqu'il existe une base dans laquelle sa matrice est triangulaire supérieure stricte. On remarque que $P_u(X) = (-1)^n X^n$ est scindé sur $\mathbb{K}$.

Nous avons remarqué que, dans chacun des exemples d'endomorphismes trigonalisables précédents, le polynôme caractéristique est scindé sur $\mathbb{K}$. C'est un fait général qui caractérise la trigonalisation :

Théorème 9.1.2 (Caractérisation des endomorphismes trigonalisables) *Soit E un $\mathbb{K}$-espace vectoriel de dimension n. Un endomorphisme $u \in L(E)$ est trigonalisable si et seulement si son polynôme caractéristique P_u est scindé sur $\mathbb{K}$.*

Remarque. Si nous formulons ce théorème avec des matrices, nous voyons qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est trigonalisable si et seulement si P_A est un polynôme scindé sur $\mathbb{K}$.

Preuve. Soit $u \in L(E)$ un endomorphisme trigonalisable. Il existe donc une base $\mathcal{B}$ de E telle que

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

soit une matrice triangulaire supérieure. Il est alors clair que $P_u(X) = (\lambda_1 - X)\dots(\lambda_n - X)$ est un polynôme scindé sur $\mathbb{K}$.

Intéressons-nous maintenant à la réciproque. Nous supposons maintenant que $u \in L(E)$ est un endomorphisme tel que P_u soit un polynôme scindé sur $\mathbb{K}$ et nous montrons que u est trigonalisable. Nous montrons ce résultat par récurrence sur $n = \dim(E)$.

Si $\dim(E) = 1$, alors les matrices sont de taille 1×1. Elles sont toutes triangulaires supérieures donc tous les endomorphismes sont trigonalisables.

Nous supposons le résultat connu pour $\dim(E) = n-1$ et nous le prouvons pour $\dim(E) = n$. Comme le polynôme P_u est scindé, il admet au moins une racine λ, c'est-à-dire que u possède une valeur propre λ et un vecteur propre e_1. Nous complétons e_1 en une base $\mathcal{B} = (e_1, \dots, e_n)$ de E. Dans cette base, la matrice de u est $M = \begin{pmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$

Nous avons alors $P_u(X) = P_M(X) = (\lambda - X)P_B(X)$. Comme P_u est scindé, il s'ensuit que P_B est scindé sur $\mathbb{K}$ également. La matrice B est de taille $(n-1) \times (n-1)$ et a un polynôme caractéristique scindé sur

$\mathbb{K}$. D'après notre hypothèse de récurrence, elle est donc trigonalisable. Il existe alors $Q \in \mathrm{GL}_{n-1}(\mathbb{K})$ et une matrice triangulaire supérieure $T \in \mathcal{M}_{n-1}(\mathbb{K})$ telles que $Q^{-1}BQ = T$. Nous décidons maintenant de poser $P = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{pmatrix}$. C'est une matrice diagonale par blocs. Si nous calculons le produit par blocs $\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{pmatrix}$, nous trouvons I_n donc P est inversible et $P^{-1} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{pmatrix}$. Nous calculons maintenant le produit $P^{-1}MP$ par blocs :

$$
\begin{aligned}
P^{-1}MP &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{pmatrix} \begin{pmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{pmatrix} \\
&= \begin{pmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & Q^{-1}BQ & \\ 0 & & & \end{pmatrix} \\
&= \begin{pmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & T & \\ 0 & & & \end{pmatrix}
\end{aligned}
$$

Cette matrice $P^{-1}MP$ est triangulaire supérieure (puisque T l'est). Elle est de plus semblable à M donc c'est aussi une matrice de u (dans une autre base). Par conséquent, u est trigonalisable.

□

L'hypothèse "P_u est scindé sur $\mathbb{K}$" est particulièrement simple si le corps de base est $\mathbb{C}$ car alors tout polynôme est scindé. Nous en déduisons le corollaire suivant, qui constitue le principal cas d'application du théorème précédent :

Corollaire 9.1.3 *1. Si E est un $\mathbb{C}$-espace vectoriel de dimension n, alors tout endomorphisme sur E est trigonalisable.*

2. *Toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ est trigonalisable.*

Voici un exemple d'application :

Proposition 9.1.4 (Trace, déterminant et valeurs propres) *Soit $u \in L(E)$ un endomorphisme trigonalisable et soit $\lambda_1, \ldots, \lambda_n$ les valeurs propres de u (répétées selon leurs multiplicités algébriques). On a :*

$$\mathrm{tr}(u) = \sum_{i=1}^{n} \lambda_i \text{ et } \det(u) = \prod_{i=1}^{n} \lambda_i.$$

On a un énoncé identique sur les matrices trigonalisables.

Preuve. Si $u \in L(E)$ est trigonalisable, alors il existe une base $\mathcal{B}$ dans laquelle $T = \mathrm{Mat}_{\mathcal{B}}(u)$ est triangulaire supérieure. De plus, les termes diagonaux de T sont les valeurs propres de u. Dès lors, il est clair que $\mathrm{tr}(u) = \mathrm{tr}(T) = \lambda_1 + \cdots + \lambda_n$ et $\det(u) = \det(T) = \lambda_1 \ldots \lambda_n$.

□

9.2 Diagonalisation

Définition 9.2.1

1. Soit E un $\mathbb{K}$-espace vectoriel de dimension n. Un endomorphisme $u \in L(E)$ est dit *diagonalisable* lorsqu'il existe une base $\mathcal{B}$ de E dans laquelle la matrice de u est diagonale.
2. Une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ est dite *diagonalisable* lorsqu'elle est semblable à une matrice diagonale. Cela revient à dire que A est la matrice d'un endomorphisme diagonalisable dans une certaine base ou encore qu'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}AP$ soit une matrice diagonale.

Remarque. J'attire votre attention sur le fait qu'il ne faut pas confondre une matrice diagonalisable et une matrice diagonale. De plus, le terme "endomorphisme diagonal" n'a aucun sens. Considérons la matrice $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$. Elle n'est pas diagonale. Toutefois, on peut vérifier (faites-le en exercice) que $P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ est inversible d'inverse $P^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ et que $P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ qui est diagonale. Par conséquent, la matrice A est diagonalisable.

Nous arrivons au théorème qui permet d'exprimer la diagonalisation de plusieurs façons différentes :

Théorème 9.2.2 (Différentes expressions de la diagonalisation) *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ un endomorphisme. Les propriétés suivantes sont équivalentes :*

1. *u est diagonalisable.*
2. *Il existe une base de E formée de vecteurs propres de u.*
3. *$E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda$.*
4. *P_u est scindé sur $\mathbb{K}$ et, pour toute valeur propre λ, on a $\alpha_\lambda = \beta_\lambda$.*

Preuve. Nous montrons successivement les équivalences suivantes : $1 \iff 2$, $2 \iff 3$, et $3 \iff 4$.

$\boxed{1 \iff 2}$ Dire que la matrice de u dans la base $\mathcal{B}$ est diagonale (de coefficients diagonaux λ_i), c'est exactement dire que, pour chaque i, $u(e_i) = \lambda_i e_i$. Cela revient donc exactement à dire que la base $\mathcal{B}$ est constituée de vecteurs propres.

$\boxed{2 \iff 3}$ Supposons tout d'abord que $E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda$. Soit $\lambda_1, \ldots, \lambda_p$ les valeurs propres distinctes de u. On commence par choisir une base $\mathcal{B}_1$ de E_{λ_1}. On choisit ensuite une base $\mathcal{B}_2$ de E_{λ_2}, on continue ainsi de suite jusqu'à choisir une base $\mathcal{B}_p$ de E_{λ_p}. Comme $E = \bigoplus_{i=1}^{p} E_{\lambda_i}$, on sait que le recollement de toutes ces bases forme une base de E. Comme les vecteurs de base sont des vecteurs non nuls des E_{λ_i}, ce sont tous des vecteurs propres. On a donc trouvé une base de E formée de vecteurs propres de u.

Supposons maintenant que E admet une base $\mathcal{B} = (e_1, \ldots, e_n)$ formée de vecteurs propres de u. Comme tous les e_i sont situés dans un sous-espace propre, on a $E = \mathrm{Vect}(e_1, \ldots, e_n) \subset \sum_{i=1}^{p} E_{\lambda_i}$. Donc $E = \sum_{i=1}^{p} E_{\lambda_i}$. Par ailleurs, on sait que la somme des sous-espaces propres est directe, d'où le résultat.

$\boxed{3 \iff 4}$ Soit $F = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda$. Rappelons une proposition que nous avions vue sur les sommes directes de plusieurs sous-espaces. Si p sous-espaces sont en somme directe, alors la dimension de leur somme est égale à la somme des dimensions. Par conséquent, nous avons $\dim(F) = \sum_{\lambda \in \mathrm{Sp}(u)} \dim(E_\lambda) = \sum_{\lambda \in \mathrm{Sp}(u)} \beta_\lambda$.

Par ailleurs, comme P_u est scindé sur $\mathbb{K}$, le polynôme P_u a n racines (comptées avec multiplicités), ce qui signifie exactement que $\sum_{\lambda \in \mathrm{Sp}(u)} \alpha_\lambda = n$.

Par conséquent, si on suppose que $F = E$, alors $\sum_{\lambda \in \mathrm{Sp}(u)} \beta_\lambda = n = \sum_{\lambda \in \mathrm{Sp}(u)} \alpha_\lambda$. En sommant des quantités a priori plus petites (les β_λ sont toujours inférieurs ou égaux aux α_λ voir la prop. 8.3.6), nous trouvons la même somme. Ceci n'est possible que si, en fait, chaque β_λ vaut α_λ. Ceci prouve $3 \Rightarrow 4$.

Supposons maintenant que chaque β_λ vaut α_λ. Alors $\sum_{\lambda\in\mathrm{Sp}(u)} \beta_\lambda = \sum_{\lambda\in\mathrm{Sp}(u)} \alpha_\lambda = n$. D'après notre rappel, il vient alors $\dim(F) = n$, c'est-à-dire $F = E$. Ceci prouve $4 \Rightarrow 3$.

□

Remarques.

1. Dans le point 2, notons que c'est précisément dans une base de vecteurs propres de u que la matrice de u est diagonale.
2. Dans le point 3, notons que nous avions déjà prouvé au théorème 8.2.4 page 115 que la somme des sous-espaces propres était directe. La nouvelle information ici, c'est que cette somme directe recouvre l'espace tout entier.
3. Dans le point 4, nous voyons que le fait d'être diagonalisable signifie que l'identification des points de vue algébrique et géométrique esquissée avec le théorème 8.3.3 page 119 est complète. Voir aussi la discusi et seulement sion qui suit la définition 8.3.5 page 121.

Proposition 9.2.3 *Dans un $\mathbb{K}$-espace vectoriel E de dimension n, tout endomorphisme ayant n valeurs propres distinctes est automatiquement diagonalisable.*

Preuve. Si u est un endomorphisme ayant n valeurs propres distinctes (en dimension n), alors ces valeurs propres sont simples ($\alpha_\lambda = 1$). Par ailleurs, comme P_u possède n racines, il est scindé sur $\mathbb{K}$. Enfin, comme $1 \leq \beta_\lambda \leq \alpha_\lambda = 1$, il vient $\beta_\lambda = \alpha_\lambda$. Par conséquent, u est diagonalisable.

□

Nous allons maintenant passer à la diagonalisation concrète d'un endomorphisme donné.

9.3 Diagonalisation concrète

Nous allons diagonaliser un endomorphisme. Pour diagonaliser une matrice M, on considère l'endomorphisme de $\mathbb{K}^n$ dont la matrice dans la base canonique (ε) est M. La méthode est la suivante :

1. On choisit une base (ε) de E (c'est souvent la base canonique lorsque $E = \mathbb{K}^n$ mais ce peut être une autre si on préfère). On écrit ensuite la matrice M de u dans cette base (ε). On est bien entendu dispensé de cette étape lorsqu'on cherche à diagonaliser une matrice.
2. On calcule le polynôme caractéristique $P_u(X)$ de u. Pour cela, on écrit la matrice $M - XI_n$ et on calcule son déterminant.
3. On résout l'équation $P_u(X) = 0$ en prenant soin de la résoudre dans le corps $\mathbb{K}$. Pour cela, on factorise dans $P_u(X)$ tous les termes de degré 1

possibles. On trouve ainsi les valeurs propres λ de u et leur multiplicités algébriques α_λ. Si P_u admet moins de n racines (avec multiplicités), alors u n'est pas diagonalisable et on s'arrête. Sinon, on continue.

4. Pour chaque valeur propre λ, on résout le système linéaire $MX = \lambda X$ avec $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Ce système admet **toujours** plusieurs solutions (un sous-espace propre de solutions de dimension $\beta_\lambda \leq \alpha_\lambda$).
5. Si $\beta_\lambda < \alpha_\lambda$ pour une valeur propore λ, la matrice n'est pas diagonalisable. On peut s'arrêter. Sinon, elle est diagonalisable et on continue.
6. On trouve une base de chaque sous-espace propre (autant de vecteurs que la dimension du sous-espace propre).
7. On recolle ces bases et on obtient une base $\mathcal{B}$ dans laquelle la matrice de u est diagonale. Ecrire cette matrice diagonale D ainsi que la base $\mathcal{B}$.
8. Dans le cas d'une matrice, on s'intéresse en plus à la matrice de passage P telle que $M = P^{-1}DP$. Pour cela, on remarque que P est la matrice de passage de $\mathcal{B}$ à (ε). Par conséquent, P^{-1} est la matrice de passage de (ε) à $\mathcal{B}$. Elle est donc constituée des coordonnées des e_i dans la base (ε), écrites en colonnes [1]. Par conséquent :
   - A l'étape 6, on a trouvé les coordonnées des vecteurs e_i. Ecrire ces coordonnées en colonne. Cela fournit la matrice P^{-1}.
   - Inverser la matrice P^{-1}. Cela fournit la matrice P.

Traitons un exemple. Diagonalisons $M = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \\ -1 & 0 & 2 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$.
Pour répondre à la question, il nous faut trouver une matrice diagonale D, une matrice inversible P et son inverse P^{-1} telles que $M = P^{-1}DP$.

1. On appelle $u : \mathbb{K}^n \to \mathbb{K}^n$ l'endomorphisme dont la matrice dans la base canonique est M.
2. Calculons $P_M(X) = P_u(X)$. On a :

$$P_M(X) = \begin{vmatrix} 1-X & 0 & 0 \\ 1 & 1-X & -1 \\ -1 & 0 & 2-X \end{vmatrix} = (1-X)\begin{vmatrix} 1-X & 0 \\ -1 & 2-X \end{vmatrix} = (1-X)^2(2-X).$$

   (on a développé par rapport à la deuxième colonne)
3. Grâce à l'écriture précédente, les racines de P_M dans $\mathbb{R}$ sont toutes trouvées : 2 est racine simple et 1 est racine double.

[1] Si tout cela ne vous semble pas clair, vous devriez vous reporter aux formules de changement de bases §2.3.B

4. Nous résolvons les deux systèmes $MX = X$ et $MX = 2X$ avec $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. La résolution de $MX = X$ donne :

$$(M - I_3)X = 0 \iff \begin{cases} 0 = 0 \\ x - z = 0 \\ -x + z = 0 \end{cases} \iff x = z.$$

   E_1 est un sous-espace vectoriel déterminé par une équation. C'est donc un plan. Il est engendré par deux vecteurs linéairement indépendants vérifiant $x = z$, par exemple $e_1 = (1, 0, 1)$ et $e_2 = (1, 1, 1)$.
   La résolution de $MX = 2X$ donne :

$$(M - 2I_3)X = 0 \iff \begin{cases} -x = 0 \\ x - y - z = 0 \\ -x = 0 \end{cases} \iff \begin{cases} x = 0 \\ y = -z \end{cases}$$

   E_2 est un sous-espace vectroiel déterminé par deux équations indépendantes. C'est donc une droite. Il est engendré par exemple par le vecteur $e_3 = (0, -1, 1)$.
   NB : On peut trouver d'autres vecteurs, du moment qu'ils engendrent bien E_1 et E_2.

5. Nous constatons que $\alpha_1 = \beta_1 = 2$ et que $\alpha_2 = \beta_2 = 1$. Par conséquent, u (et donc M) est diagonalisable.

6. Nous l'avons déjà fait à l'étape 4

7. La base $\mathcal{B}$ obtenue est (e_1, e_2, e_3). Nous pourrions vérifier que c'est bien une base (vous pouvez le faire pour vous exercer) mais cela est inutile du point de vue logique car nous savons par ailleurs que des vecteurs propres associés à des valeurs propres différentes forment une famille libre. Dans cette base, la matrice de u est particulièrement simple puisque $u(e_1) = e_1$, $u(e_2) = e_2$ et $u(e_3) = 2e_3$. Donc $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

8. Comme la matrice D représente le même endomorphisme u dans une nouvelle base $\mathcal{B}$, nous savons que M et D sont semblables. Plus précisément, nous avons $M = P^{-1}DP$ avec $P = P_{\mathcal{B}\to(\varepsilon)}$. Ainsi P^{-1} est constituée des coordonnées des vecteurs e_i dans la base (ε) donc $P^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$.
   Pour obtenir P, nous inversons P^{-1}, ce qui se fait par exemple grâce

au pivot de Gauss.

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \qquad \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right) [-L_1]$$

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

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

Nous avons donc $M = P^{-1}DP$.

9.4 Application aux systèmes différentiels

9.4.A Résolution de l'équation différentielle $x' = \lambda x$

Soit $\lambda \in \mathbb{R}$ un nombre fixé. L'équation différentielle $x' = \lambda x$ est une équation dont l'inconnue est une fonction. Une solution de cette équation est une fonction dérivable $x : t \mapsto x(t)$ définie sur $\mathbb{R}$ et telle que, pour tout $t \in \mathbb{R}$, on ait $x'(t) = \lambda x(t)$. *Il ne faut pas se laisser abuser par l'écriture (guidée par l'usage) : ici x est une fonction et non pas un nombre.*

Si $x : \mathbb{R} \to \mathbb{R}$ est une fonction dérivable quelconque, alors la dérivée de la fonction $t \mapsto e^{-\lambda t}x(t)$ est $t \mapsto (x'(t) - \lambda x(t))e^{-\lambda t}$. Par conséquent, la fonction x est solution de l'équation différentielle $x' = \lambda x$ si et seulement si la fonction $t \mapsto e^{-\lambda t}x(t)$ a une dérivée nulle donc est constante. Si nous notons c cette constante, il vient : $\forall t \in \mathbb{R}, \quad x(t) = ce^{\lambda t}$. Nous voyons alors que $x(0) = c$. Par conséquent, nous retenons :

Théorème 9.4.1 (Solution de l'équation différentielle $x' = \lambda x$) *Soit λ un nombre réel fixé. Quel que soit le nombre réel c, il existe une unique fonction dérivable $x : \mathbb{R} \to \mathbb{R}$ solution de l'équation différentielle $x' = \lambda x$ et vérifiant la condition initiale $x(0) = c$. Cette fonction est donnée par la formule*

$$\forall t \in \mathbb{R}, \quad x(t) = ce^{\lambda t}.$$

9.4.B Résolution de $X' = DX$

On se donne maintenant n nombres réels $\lambda_1, \ldots, \lambda_n$ et on veut résoudre simultanément les n équations différentielles $x_1' = \lambda_1 x_1, \ldots, x_n' = \lambda_n x_n$. Afin

de simplifier les écritures, on écrit pour chaque $t \in \mathbb{R}$:

$$X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \quad X'(t) = \begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{pmatrix}$$

Dès lors, chaque fonction x_j est solution de l'équation différentielle $x_j' = \lambda_j x_j$ si et seulement si X est solution de $X' = DX$. A l'aide du théorème précédent, on voit que, pour chaque vecteur unicolonne $C = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$, il existe une unique solution $t \mapsto X(t)$ de $X' = DX$ qui vérifie la condition initiale $X(0) = C$. Cette solution est donnée, pour tout $t \in \mathbb{R}$ par : $x_1(t) = c_1 e^{\lambda_1 t}, \ldots, x_n = c_n e^{\lambda_n t}$ soit :

$$X(t) = \begin{pmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix},$$

9.4.C Résolution d'un système différentiel diagonalisable

Un système différentiel linéaire homogène à coefficients constants (on écrire plus rapidement "système différentiel") est un système d'équations différentielles à n fonctions inconnues $x_1, \ldots, x_n$ de la forme

$$\begin{cases} x_1' = a_{11}x_1 + \cdots + a_{1n}x_n \\ \qquad \vdots \\ x_n' = a_{n1}x_1 + \cdots + a_{nn}x_n \end{cases} \tag{9.1}$$

où les a_{ij} sont des nombres fixés. Une solution est une famille de n fonctions dérivables $x_1, \ldots, x_n$ telles que

$$\begin{cases} x_1'(t) = a_{11}x_1(t) + \cdots + a_{1n}x_n(t) \\ \qquad \vdots \\ x_n'(t) = a_{n1}x_1(t) + \cdots + a_{nn}x_n(t) \end{cases}$$

pour chaque réel t. On note $X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$, $X'(t) = \begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix}$ et $A = (a_{ij})$.

La famille de fonctions $(x_1, \ldots, x_n)$ est solution du système différentiel (9.1) si et seulement si X est solution de $X' = AX$.

Nous supposons dans toute la suite que A est une matrice diagonalisable. Soit $P \in \mathrm{GL}_n(\mathbb{R})$ et

$$D = \begin{pmatrix} \lambda_1 & 0 & \ddots & 0 \\ 0 & \lambda_2 & 0 & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

telles que $A = P^{-1}DP$. On pose $Y = PX$ et $Y' = PX'$. On a alors :

$$X' = AX \iff X' = P^{-1}DPX \iff PX' = DPX \iff Y' = DY.$$

Le système différentiel $Y' = DY$ est diagonal comme dans la partie précédente. On peut donc le résoudre comme ci-dessus et en déduire les solutions de $X' = AX$.

Exemple. Trouver les fonctions x_1, x_2, x_3 solutions du système différentiel suivant et vérifiant les conditions initiales prescrites :

$$\begin{cases} x_1' = -x_2 - x_3 \\ x_2' = x_1 + x_3 \\ x_3' = -x_1 + x_2 \end{cases} \qquad \begin{cases} x_1(0) = 0 \\ x_2(0) = 1 \\ x_3(0) = 2 \end{cases} \tag{9.2}$$

En suivant le schéma proposé ci-dessus, nous écrivons le système sous forme matricielle. Nous posons donc pour tout $t \in \mathbb{R}$:

$$X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}, \quad X'(t) = \begin{pmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \end{pmatrix} \text{ et } A = \begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}$$

Le système différentiel (9.2) avec condition initiale équivaut alors à $X' = AX$ et $X(0) = C$ avec $C = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$.

Voyons maintenant si la matrice A est diagonalisable et diagonalisons la si possible. Soit u l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique (ε) est A. Nous calculons le polynôme caractéristique :

$$P_u(X) = P_A(X) = \begin{vmatrix} -X & -1 & -1 \\ 1 & -X & 1 \\ -1 & 1 & -X \end{vmatrix}$$

Si nous utilisons (par exemple) la règle de Sarrus, il vient :

$$P_u(X) = (-X)^3 + (-1) + (-1)^2 - (-1)^2(-X) - (-X) - (-1)(-X) = -X^3 + X$$

On factorise

$$P_u(X) = X(1 - X^2) = X(1 - X)(1 + X)$$

Les trois valeurs propres de u sont donc 0, -1 et 1, toutes simples. On cherche maintenant les vecteurs propres. Pour cela, on résout les systèmes $AX = 0$, $AX = X$ et $AX = -X$. Cela donne :

$$AX = 0 \iff \begin{cases} -x_2 - x_3 = 0 \\ x_1 + x_3 = 0 \\ -x_1 + x_2 = 0 \end{cases} \iff \begin{cases} x_2 = -x_3 \\ x_1 = x_2 \end{cases}$$

Un vecteur propre de u associé à la valeur propre 0 est donc $e_1 = (1, 1, -1)$. En résolvant de même, on trouve comme vecteur propre de u associé à la valeur propre 1 le vecteur $e_2 = (1, 0, -1)$, comme vecteur propre de u associé à la valeur propre -1, le vecteur $e_3 = (0, 1, -1)$. Nous avons donc $A = P^{-1}DP$ avec $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ et, en écrivant les coordonnées des e_i en colonnes : $P^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & -1 & -1 \end{pmatrix}$.

Le calcul de P se fait par la méthode du pivot de Gauss :

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ -1 & -1 & -1 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 1 \end{array}\right) \begin{matrix} [-L_1] \\ [+L_1] \end{matrix} \longrightarrow$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 1 \end{array}\right) [+L_2] \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & -1 & 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 & 0 & 1 \end{array}\right) \begin{matrix} [+L_3] \\ [+L_3] \end{matrix} \longrightarrow$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & -1 & 0 & -1 \end{array}\right) \begin{matrix} \\ [\times(-1)] \\ [\times(-1)] \end{matrix} \quad \text{donc} P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -1 \\ -1 & 0 & -1 \end{pmatrix}$$

On pose alors $Y = PX = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, $Y' = PX'$ et le système différentiel (9.2) équivaut à $Y' = DY$ et $Y(0) = PC = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -1 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \\ -2 \end{pmatrix}$.

Les solutions sont alors, pour tout $t \in \mathbb{R}$:

$$y_1(t) = 3e^0 = 3, \quad y_2(t) = -3e^t, \quad y_3(t) = -2e^{-t}.$$

En revenant à X, il vient :

$$X = P^{-1}Y \text{donc} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ -3e^t \\ -2e^{-t} \end{pmatrix}$$

soit finalement, pour tout $t \in \mathbb{R}$:

$$x_1(t) = 3 - 3e^t, \quad x_2(t) = 3 - 2e^{-t}, \quad x_3(t) = -3 + 3e^t + 2e^{-t}.$$

9.5 Application aux suites récurrentes

Nous traiterons un exemple. Trouver toutes les suites $(u_k)_{k\geq 0}$ et $(v_k)_{k\geq 0}$ de nombres complexes tels que

$$\begin{cases} u_{k+1} = u_k - v_k \\ v_{k+1} = u_k + v_k \end{cases}$$

Nous cherchons à diagonaliser la matrice $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Soit $u : \mathbb{C}^2 \to \mathbb{C}^2$ l'endomorphisme dont la matrice dans la base canonique $(\varepsilon_1, \varepsilon_2)$ de $\mathbb{C}^2$ est A. Son polynôme caractéristique vaut $P_A(X) = X^2 - \text{tr}(A)X + \det(A) = X^2 - 2X + 2$.

Résolvons l'équation $P_A(X) = 0$ dans $\mathbb{C}$. Le discriminant vaut $\Delta = 4 - 8 = -4 = (2i)^2$. Les deux solutions sont donc $\dfrac{2 \pm 2i}{2} = 1 \pm i$.

En posant $X = \begin{pmatrix} x \\ y \end{pmatrix}$ avec $x, y \in \mathbb{C}$, nous résolvons les systèmes $AX = (1+i)X$ et $AX = (1-i)X$. Cela donne :

$$AX = (1+i)X \iff \begin{cases} x - y = (1+i)x \\ x + y = (1+i)y \end{cases} \iff x = iy$$

$$AX = (1-i)X \iff \begin{cases} x - y = (1-i)x \\ x + y = (1-i)y \end{cases} \iff y = ix.$$

Ceci nous permet d'écrire $E_{1+i} = \text{Vect}\big((i,1)\big)$ et $E_{1-i} = \text{Vect}\big((1,i)\big)$.

En écrivant ces coordonnées en colonnes, on trouve la matrice $P^{-1} = \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}$ telle que $A = P^{-1}DP$ avec $D = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix} = \begin{pmatrix} z & 0 \\ 0 & \overline{z} \end{pmatrix}$, $z = 1 + i = \sqrt{2}e^{i\pi/4}$. Le calcul de l'inverse de la matrice P^{-1} est direct :

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

Nous avons alors $A = P^{-1}DP$ et donc, pour tout $k \in \mathbb{N}$, $A^k = P^{-1}D^kP$. Cela donne :

$$\begin{aligned} A^k &= \frac{1}{2}\begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}\begin{pmatrix} z^k & 0 \\ 0 & \overline{z}^k \end{pmatrix}\begin{pmatrix} -i & 1 \\ 1 & -i \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} z^k + \overline{z}^k & i[z^k - \overline{z}^k] \\ -i[z^k - \overline{z}^k] & z^k + \overline{z}^k \end{pmatrix} \\ &= \begin{pmatrix} \Re(z^k) & -\operatorname{Im}(z^k) \\ \operatorname{Im}(z^k) & \Re(z^k) \end{pmatrix} \\ &= \sqrt{2^k}\begin{pmatrix} \cos(k\pi/4) & -\sin(k\pi/4) \\ \sin(k\pi/4) & \cos(k\pi/4) \end{pmatrix} \end{aligned}$$

où $\Re(x)$ et $\operatorname{Im}(x)$ désigne ici la partie réel et imaginaire du nombre complexe x.

Résolvons maintenant le problème de départ. Nous posons $X_k = \begin{pmatrix} u_k \\ v_k \end{pmatrix} \in \mathcal{M}_{21}(\mathbb{C})$. Nous avons $X_{k+1} = AX_k$ d'où on déduit facilement par récurrence que $X_k = A^kX_0$. Grâce à la formule démontrée ci-dessus, il vient :

$$u_k = \sqrt{2^k}\left[\cos\left(\frac{k\pi}{4}\right)u_0 - \sin\left(\frac{k\pi}{4}\right)v_0\right] \text{ et } v_k = \sqrt{2^k}\left[\cos\left(\frac{k\pi}{4}\right)v_0 + \sin\left(\frac{k\pi}{4}\right)u_0\right].$$

Chapitre 10

Polynômes d'endomorphismes

10.1 Polynômes annulateurs d'un endomorphisme

10.1.A Généralités

Nous avions vu au chapitre 7 comment définir un polynôme d'endomorphisme. Rappelons-le :

Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $P(X) = \sum_{i=1}^{d} a_i X^i$ un polynôme à coefficients dans $\mathbb{K}$. On définit $P(u) = \sum_{i=1}^{d} a_i u^i = a_d u^d + a_{d-1} u^{d-1} + \cdots + a_1 u + a_0 \mathrm{id}_E$. J'avais déjà attiré votre attention sur le fait de ne pas oublier dans cette écriture que $u^0 = \mathrm{id}_E$.

Définition 10.1.1 Un polynôme P est dit *annulateur de* u si $P(u) = 0$ (c'est-à-dire est l'endomorphisme nul).

La proposition suivante donne les premières propriétés des polynômes d'endomorphismes.

Proposition 10.1.2 *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ un endomorphisme.*

1. *Si $P \in \mathbb{K}[X]$ est un polynôme quelconque, alors les endomorphismes u et $P(u)$ commutent. Plus généralement, si P et Q sont deux polynômes, alors $P(u)$ et $Q(u)$ commutent.*
2. *Il existe des polynômes non nuls P tels que $P(u) = 0$.*

Preuve.

1. Soit $R(X) = XP(X) = P(X)X$. Il est clair que $uP(u) = R(u) = P(u)u$. Plus généralement, on a $P(u)Q(u) = (PQ)(u) = (QP)(u) = Q(u)P(u)$.
2. La famille $(u^0, u^1, \ldots, u^{n^2})$ contient n^2+1 éléments dans l'espace vectoriel $L(E)$ de dimension n^2. Par conséquent, elle est liée. Il existe alors des nombres $a_0, \ldots, a_{n^2}$, non tous nuls, tels que $a_0u^0 + \cdots + a_{n^2}u^{n^2} = 0$. Posons $P(X) = \sum_{k=0}^{n^2} a_k X^k$. P n'est pas le polynôme nul puisque les a_k ne sont pas tous nuls et $P(u) = 0$.

□

Exemples

1. Soit p un projecteur. Nous rappelons que c'est un endomorphisme idempotent, c'est-à-dire que $p^2 = p$. Considérons le polynôme non nul $Q(X) = X^2 - X$. Alors $Q(p) = p^2 - p = 0$. Donc Q est un polynôme annulateur de p.
2. Soit s une symétrie. Nous posons $Q(X) = X^2 - 1$. Comme $s^2 = \mathrm{id}_E$, nous avons $Q(s) = 0$ donc Q est un polynôme annulateur de s.
3. Soit u un endomorphisme nilpotent d'indice p. Posons $Q(X) = X^p$. On a alors $Q(u) = u^p = 0$ donc Q est un polynôme annulateur de u.

Proposition 10.1.3 *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ un endomorphisme. Soit x_0 un vecteur propre de u associé à la valeur propre λ. Soit $P \in \mathbb{K}[X]$ un polynôme. Alors :*

$$P(u)(x_0) = P(\lambda)x_0.$$

En particulier,
- *Si λ est une valeur propre de u, alors $P(\lambda)$ est une valeur propre de $P(u)$.*
- *Si λ est une valeur propre de u, alors c'est une racine de n'importe lequel des polynômes annulateurs de u.*

Preuve. Nous constatons que $u^0(x_0) = \mathrm{id}_E(x_0) = x_0 = \lambda^0 x_0$. De plus $u(x_0) = \lambda x_0$,

$$u^2(x_0) = u(u(x_0)) = u(\lambda x_0) = \lambda u(x_0) = \lambda^2 x_0.$$

Par une récurrence facile, on constate que, pour tout $k \in \mathbb{N}$, $u^k(x_0) = \lambda^k x_0$. Soit $P(X) = \sum_{i=0}^{d} a_i X^i$. Alors :

$$P(u)(x_0) = \sum_{i=1}^{d} a_i u^i(x_0) = \sum_{i=1}^{d} a_i \lambda^i x_0 = P(\lambda)x_0.$$

□

Remarque A gauche de cette égalité, il s'agit de l'endomorphisme $P(u)$ appliqué au vecteur x_0. A droite, c'est le vecteur x_0 multiplié par le scalaire $P(\lambda)$.

10.1.B Un exemple important : le théorème de Cayley-Hamilton

Le théorème suivant fournit de façon systématique un polynôme annulateur de u.

Théorème 10.1.4 (de Cayley-Hamilton) *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ un endomorphisme. Alors $P_u(u) = 0$. En d'autres termes, le polynôme caractéristique est un polynôme annulateur.*

Preuve. Nous posons $v = P_u(u) \in L(E)$. Soit $x_0 \in E \setminus \{0\}$. On veut prouver que $v(x_0) = 0$. Désignons par J l'ensemble des entiers naturels k tels que la famille $(x_0, u(x_0), \ldots, u^k(x_0))$ soit liée.

- Tout d'abord, l'ensemble J est non vide. Prenons par exemple $k = n$. La famille $(x_0, u(x_0), \ldots, u^n(x_0))$ contient $n+1$ vecteurs en dimension n. Elle est donc liée. Par conséquent, $n \in J$ et $J \neq \emptyset$.
- Nous posons alors $p = \min J$. L'entier p vérifie à la fois que la famille $(x_0, u(x_0), \ldots, u^p(x_0))$ est liée et que la famille $(x_0, u(x_0), \ldots, u^{p-1}(x_0))$ est libre.

Par conséquent, nous avons $u^p(x_0) \in \mathrm{Vect}(x_0, \ldots, u^{p-1}(x_0))$. Il existe donc des scalaires $a_0, \ldots, a_{p-1}$ tels que $u^p(x_0) = a_0 x_0 + a_1 u(x_0) + \cdots + a_{p-1} u^{p-1}(x_0)$.

Nous introduisons $F = \mathrm{Vect}(x_0, u(x_0), \ldots, u^{p-1}(x_0))$. Les images des vecteurs $x_0, u(x_0), \ldots, u^{p-2}(x_0)$ valent respectivement $u(x_0), u^2(x_0), \ldots, u^{p-1}(x_0)$. Elles sont donc toutes dans F. De plus, l'image de $u^{p-1}(x_0)$ vaut $u^p(x_0)$ qui est dans F (nous venons de le montrer). Ceci prouve que F est stable par u. Soit donc $\hat{u}$ l'endomorphisme de F induit par u. La famille $(x_0, u(x_0), \ldots, u^{p-1}(x_0))$ est libre (définition de p) et engendre F (définition de F). C'est donc une base de F. Nous venons de calculer les images des vecteurs de cette base par u donc la matrice de $\hat{u}$ dans cette base est la suivante :

$$A = \begin{pmatrix} 0 & & & & a_0 \\ & \ddots & & & \vdots \\ & & 0 & & \\ 1 & & & & a_1 \\ & & & & \\ & & & & \\ & & & 0 & a_{p-2} \\ & 0 & & & \\ & & & 1 & a_{p-1} \end{pmatrix}$$

C'est une matrice de Froebenius. Nous avions calculé son polynôme caractéristique dans le paragraphe 8.3.A. Nous avions trouvé :

$$P_{\hat{u}}(X) = P_A(X) = (-1)^p[X^p - a_{p-1}X^{p-1} - \cdots - a_1X - a_0].$$

A la proposition 8.3.6 page 121, nous avions prouvé que le polynôme caractéristique de u est un multiple de $P_{\hat{u}} = P_A$, c'est-à-dire qu'il est de la forme $P_u = Q\,P_A$ avec $Q \in \mathbb{K}[X]$. Par conséquent, en posant $w = Q(u) \in L(E)$, nous avons $v = w \circ P_A(u)$ et donc

$$v(x_0) = (-1)^p\, w\big(u^p(x_0) - a_{p-1}u^{p-1}(x_0) - \cdots - a_1u(x_0) - a_0x_0\big) = (-1)^p w(0) = 0.$$

C'est ce que nous voulions démontrer.

□

Exemple. Pour u un endomorphisme nilpotent, nous avions vu $P_u(X) = (-1)^n X^n$. Donc $P_u(u) = (-1)^n u^n = 0$. En particulier, l'indice de nilpotence est inférieur ou égal à n.

10.2 Polynôme minimal d'un endomorphisme

10.2.A Définition du polynôme minimal

Rappelons qu'un polynôme P est dit *unitaire* lorsque son coefficient dominant (le coefficient devant le terme de plus grand degré) vaut 1.

Théorème 10.2.1 (Définition du polynôme minimal) *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ un endomorphisme. Il existe un unique polynôme unitaire $\mathfrak{m}_u$ vérifiant les deux propriétés suivantes :*

1. *$\mathfrak{m}_u$ est un polynôme annulateur de u.*
2. *Tout polynôme annulateur de u est un multiple de $\mathfrak{m}_u$.*

En d'autres termes, pour un polynôme $P \in \mathbb{K}[X]$, on a :

$$P(u) = 0 \quad \Longleftrightarrow \quad \exists Q \in \mathbb{K}[X],\ P = \mathfrak{m}_u Q.$$

Le polynôme $\mathfrak{m}_u$ est appelé polynôme minimal *de u.*

Preuve. Soit I l'ensemble des polynômes annulateurs de u. On voit que $0 \in I$, que tout multiple d'un élément de I est encore dans I et que la somme de deux éléments de I est encore dans I. Par conséquent, I est un idéal de $\mathbb{K}[X]$. De plus, nous avons vu que I contient des polynômes non nuls. Le théorème 7.2.7 page 102 assure alors qu'il existe un unique générateur unitaire de I, noté $\mathfrak{m}_u$. C'est le polynôme recherché.

□

Exemple. Soit u un endomorphisme nilpotent d'indice p. On sait que $u^p = 0$ donc le polynôme X^p est un annulateur de u. Il s'ensuit que X^p est un multiple de $\mathfrak{m}_u$ donc on a forcément $\mathfrak{m}_u(X) = X^k$ avec $k \leq p$. Par ailleurs, $\mathfrak{m}_u(u) = 0$ donc $u^k = 0$, ce qui ne se produit que lorsque $k \geq p$. Par conséquent, il vient $\mathfrak{m}_u(X) = X^p$.

10.2.B Propriétés du polynôme minimal

Nous regroupons les principales propriétés du polynôme minimal dans la proposition suivante. Elles découlent toutes du théorème de Cayley-Hamilton.

Proposition 10.2.2 *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ un endomorphisme non nul.*

1. *P_u est un multiple de $\mathfrak{m}_u$.*
2. *Le degré de $\mathfrak{m}_u$ vérifie $1 \leq \deg(\mathfrak{m}_u) \leq n$.*
3. *$\deg(\mathfrak{m}_u) = 1$ si et seulement si u est une homothétie.*
4. *$\deg(\mathfrak{m}_u) = n \iff P_u(X) = (-1)^n \mathfrak{m}_u(X)$.*
5. *Les polynômes P_u et $\mathfrak{m}_u$ ont les mêmes racines dans $\mathbb{K}$.*

Preuve.

1. C'est une simple reformulation du théorème de Cayley-Hamilton. Comme P_u est un polynôme annulateur de u, alors P_u est un multiple de $\mathfrak{m}_u$.
2. Nous savons que $\mathfrak{m}_u$ n'est pas le polynôme nul donc $\deg(\mathfrak{m}_u) \geq 0$. Si on avait $\deg(\mathfrak{m}_u) = 0$, alors $\mathfrak{m}_u$ serait une constante non nulle λ et alors $\mathfrak{m}_u(u) = \lambda \mathrm{id}_E \neq 0$, ce qui est absurde. Donc $\deg(\mathfrak{m}_u) \geq 1$.
   Comme P_u est un multiple de $\mathfrak{m}_u$, alors $n = \deg(P_u) \geq \deg(\mathfrak{m}_u)$.
3. Si $u = \lambda \mathrm{id}_E$, alors $X - \lambda$ est un polynôme unitaire annulateur de u. Comme il est de degré 1, qui est le plus petit degré possible pour un polynôme minimal, alors $\mathfrak{m}_u = X - \lambda$.
   Réciproquement, si $\mathfrak{m}_u = X - \lambda$ est de degré 1, alors $u - \lambda \mathrm{id}_E = 0$ donc u est bien une homothétie.
4. Si $\deg(\mathfrak{m}_u) = n$, alors P_u est un multiple de $\mathfrak{m}_u$ de même degré que $\mathfrak{m}_u$. On a donc $P_u(X) = a\mathfrak{m}_u(X)$ avec $a \in \mathbb{K}^*$. En regardant ce que valent les coefficients dominants, on voit que $a = (-1)^n$.
   Réciproquement, si $P_u(X) = (-1)^n \mathfrak{m}_u(X)$, alors $\deg(\mathfrak{m}_u) = \deg(P_u) = n$.
5. Comme P_u est un multiple de $\mathfrak{m}_u$, toute racine de $\mathfrak{m}_u$ est aussi une racine de P_u. Soit maintenant λ une racine de P_u. C'est une valeur propre de u. Soit x_0 un vecteur propre associé. Nous savons alors que, pour tout polynôme P, $P(u)(x_0) = P(\lambda)x_0$. En particulier, comme $\mathfrak{m}_u$ est annulateur de u, $0 = \mathfrak{m}_u(u)(x_0) = \mathfrak{m}_u(\lambda)x_0$. Etant donné que x_0 est non nul, cela implique que $\mathfrak{m}_u(\lambda) = 0$.

□

10.3 Diagonalisation et polynôme minimal

Le polynôme minimal permet de trouver une caractérisation simple du fait d'être diagonalisable.

Définition 10.3.1 Un polynôme $P \in \mathbb{K}[X]$ est dit *simplement scindé* sur $\mathbb{K}$ lorsqu'il se factorise sous la forme $P(X) = (X - \lambda_1) \dots (X - \lambda_p)$ avec les λ_i distincts deux à deux.

En d'autres termes, un polynôme est simplement scindé s'il est scindé et que ses racines sont toutes simples.

Théorème 10.3.2 (Diagonalisation et polynôme minimal) *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ un endomorphisme. Les propriétés suivantes sont équivalentes :*

1. *L'endomorphisme u est diagonalisable.*
2. *Il existe un polynôme P simplement scindé sur $\mathbb{K}$ tel que $P(u) = 0$.*
3. *Le polynôme minimal $\mathfrak{m}_u$ est simplement scindé sur $\mathbb{K}$.*

Bien entendu, un théorème identique est valable pour les matrices.

Preuve. $\boxed{1 \Longrightarrow 2}$ Supposons que u est diagonalisable. Nous nommons $\lambda_1, \dots, \lambda_p$ les valeurs propres distinctes de u. Nous introduisons le polynôme $P(X) = (X - \lambda_1) \dots (X - \lambda_p)$. Comme les λ_i sont distincts deux à deux, le polynôme P est simplement scindé sur $\mathbb{K}$. De plus, pour $x \in E_{\lambda_i}$, nous écrivons $P(X)$ en mettant le facteur $X - \lambda_i$ à la fin : $P(X) = Q(X)(X - \lambda_i)$. On voit alors que $P(u)(x) = Q(u)(u(x) - \lambda_i x) = Q(u)(0) = 0$. Ceci prouve que l'endomorphisme $P(u)$ est nul sur chacun des termes E_{λ_i} de la somme directe $E_{\lambda_1} \oplus \dots \oplus E_{\lambda_p}$. Ceci prouve que $P(u)$ est nul sur la somme $E_{\lambda_1} \oplus \dots \oplus E_{\lambda_p}$. Comme u est diagonalisable, cette somme vaut E donc $P(u) = 0$. On a bien trouvé un polynôme simplement scindé sur $\mathbb{K}$ qui annule u.

$\boxed{2 \Longrightarrow 3}$ Si P est un polynôme simplement scindé qui annule u, alors P est un multiple de $\mathfrak{m}_u$ donc il est clair que $\mathfrak{m}_u$ est également simplement scindé.

$\boxed{3 \Longrightarrow 1}$ Supposons que $\mathfrak{m}_u$ soit simplement scindé : $\mathfrak{m}_u(X) = \prod_{j=1}^{p} (X - \lambda_j)$. Les λ_i sont ici les valeurs propres de u car P_u et $\mathfrak{m}_u$ ont les mêmes racines. D'après le lemme des noyaux, comme $\mathfrak{m}_u(u) = 0$, il vient :

$$E = \operatorname{Ker}(\mathfrak{m}_u(u)) = \bigoplus_{j=1}^{p} \operatorname{Ker}(u - \lambda_j \operatorname{id}_E) = \bigoplus_{j=1}^{p} E_{\lambda_j}$$

On retrouve la caractérisation 3 du théorème 9.2.2 page 127. Par conséquent, u est diagonalisable.

□

Remarque importante : Comment appliquer ce théorème dans une situation concrète ? On calcule $P_u(X)$ et on trouve les valeurs propres en résolvant l'équation $P_u(X) = 0$ (bien penser à la résoudre dans $\mathbb{K}$). On introduit alors le polynôme $Q(X) = (X - \lambda_1) \dots (X - \lambda_p)$ où les λ_j sont les valeurs propres distinctes de u. Deux cas se présentent :

– Si $Q(u) = 0$, alors u est diagonalisable et Q est le polynôme minimal de u.
– Si $Q(u) \neq 0$, alors u n'est pas diagonalisable et Q n'est pas le polynôme minimal de u.

Corollaire 10.3.3 *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $u \in L(E)$ un endomorphisme. Soit F un sous-espace de E stable par u et soit $\hat{u}$ l'endomorphisme de F induit par u. Si u est diagonalisable, alors $\hat{u}$ l'est aussi.*

Preuve. Si u est diagonalisable, alors il existe un polynôme simplement scindé P tel que $P(u) = 0$. Mais alors, en se contentant de regarder la restriction sur F, on a $P(\hat{u}) = 0$. Comme $\hat{u}$ possède un polynôme annulateur simplement scindé, on en conclut que $\hat{u}$ est diagonalisable.

□

Exemple. Calculer le polynôme minimal de la matrice suivante

$$A = \begin{pmatrix} 0 & 1 & \cdots & \cdots & 1 \\ 1 & 0 & 1 & \vdots & 1 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & 1 \\ 1 & \cdots & \cdots & 1 & 0 \end{pmatrix}$$

Est-elle diagonalisable ? Est-elle inversible ? Calculer A^3 et, si possible, A^{-1}.

$$A = \begin{pmatrix} n-1 & n-2 & \cdots & \cdots & n-2 \\ n-2 & n-1 & n-2 & \vdots & n-2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & n-1 & n-2 \\ n-2 & \cdots & \cdots & n-2 & n-1 \end{pmatrix}$$

Calculons A^2. Nous voyons qu'il vient

Par conséquent, nous avons $A^2 = (n-1)I_n + (n-2)A$. Nous définissons alors le polynôme $Q(X) = X^2 - (n-2)X - (n-1)$ et nous constatons que $Q(A) = 0$. Par conséquent, Q est un polynôme annulateur de A.

Résolvons l'équation $Q(x) = 0$ dans $\mathbb{R}$. Son discriminant vaut $\Delta = (n-2)^2 + 4(n-1) = n^2$. Le discriminant est strictement positif donc Q a deux racines distinctes qui sont :

$$\frac{n-2+n}{2} = n-1 \text{ et } \frac{n-2-n}{2} = -1.$$

Q étant simplement scindé sur $\mathbb{R}$, la matrice A est diagonalisable.

$Q(A) = 0$ donc Q est un multiple de m_A. Par conséquent, il y a *a priori* trois possibilités pour $\mathfrak{m}_A$: $X + 1$, $X - (n - 1)$ et Q. Comme A n'est pas une matrice d'homothétie de la forme λI_n, on sait que $\deg(m_A) > 1$. Par conséquent, il vient $m_A = Q = X^2 - (n-2)X - (n-1)$. Les racines de m_A sont (-1) et $(n-1)$ donc ce sont les valeurs propres de A. Remarquons au passage qu'on peut constater que

$$A - (-1)I_n = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}.$$

C'est une matrice de rang 1 donc la multiplicité géométrique β_{-1} est égale à $n-1$. Comme la multiplicité géométrique β_{n-1} est au moins égale à 1, on voit que $\beta_{-1} = \alpha_{-1} = n-1$ et $\beta_{n-1} = \alpha_{n-1} = 1$. On retrouve donc par une autre méthode que A est diagonalisable.

Comme 0 n'est pas valeur propre de A, on en déduit que A est inversible. Plus précisément, il nous est même possible de donner la valeur du déterminant de A : c'est le produit des valeurs propres (répétées selon leurs multiplicités algébriques) donc

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

Calculons maintenant A^{-1}. Nous partons de la formule $A^2 = (n-1)I_n + (n-2)A$. Nous multiplions par A^{-1} et nous trouvons $A = (n-1)A^{-1} + (n-2)I_n$ soit

$$A^{-1} = \frac{1}{n-1}[A - (n-2)I_n] = \frac{1}{n-1}\begin{pmatrix} 2-n & 1 & \dots & 1 \\ 1 & 2-n & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cdots & 1 & 2-n \end{pmatrix}.$$

Pour calculer A^3, on procède de la même façon :

$$\begin{aligned} A^3 = A^2A &= [(n-1)I_n + (n-2)A]A \\ &= (n-1)A + (n-2)A^2 \\ &= (n-1)A + (n-2)[(n-1)I_n + (n-2)A] \\ &= [(n-2)^2 + n - 1]A + (n-2)(n-1)I_n \\ &= (n^2 - 3n + 3)A + (n^2 - 3n + 2)I_n \end{aligned}$$

On a

$$A^3 = \begin{pmatrix} n^2-3n+2 & n^2-3n+3 & \dots & n^2-3n+3 \\ n^2-3n+3 & n^2-3n+2 & \cdots & n^2-3n+3 \\ \vdots & \vdots & \ddots & \vdots \\ n^2-3n+3 & \cdots & n^2-3n+3 & n^2-3n+2 \end{pmatrix}.$$

□