Structures mathématiques - Partie I (ST3)

Muriel Boulakia[1]   Maxime Zavidovique[2]

2014 – 2015

1. boulakia@ann.jussieu.fr
2. maxime.zavidovique@imj-prg.fr

Table des matières

Ce polycopié a été réalisé par Clément Cancès et Frédérique Charles qui ont assuré le module de mise à niveau en ST3 les années précédentes.

Chapitre 1

Espaces vectoriels

1.1 Qu'est-ce qu'un vecteur ?

L'idée principale est la suivante : **toute chose que l'on peut additionner avec ses semblables et multiplier par un nombre sans en changer la nature est un vecteur**.

Exemple:

- En géométrie plane ou dans l'espace, les vecteurs peuvent être représentés comme des flèches (cf. collège et lycée)
- Des sons : On peut multiplier le son en augmentant le volume, et superposer (ajouter) un deuxième son.
- Deux tableaux de même taille : on peut additionner case par case, et multiplier toutes les cases par un même nombre.
- Deux fonctions continues, deux polynômes, deux suites,...

A partir de maintenant, $\mathbb{K}$ désigne soit $\mathbb{R}$, soit $\mathbb{C}$.

Définition 1.1 (espace vectoriel (e.v.))
Un ensemble E est un $\mathbb{K}$-espace vectoriel ($\mathbb{K}$-e.v.) si

1. *$E \neq \emptyset$: un espace vectoriel contient toujours au moins un élément ;*
2. *$\forall u, v \in E$, $u + v \in E$: un e.v. est stable par addition ;*
3. *$\forall u \in E$, $\forall \lambda \in \mathbb{K}$, $\lambda u \in E$: un e.v. est stable par multiplication par un élément de $\mathbb{K}$.*

Exemple:
Les ensembles suivants sont des espaces vectoriels

- $\{0\}$: cet ensemble ne contient que l'élément nul ;
- $\mathbb{R}^n$: les tableaux de réels à n entrées ;
- $\mathbb{C}^n$: les tableaux de complexes à n entrées ;
- $C([0,1];\mathbb{K})$: les fonctions continues sur $[0,1]$ à valeurs dans K ;
- $\mathbb{K}[X] = \left\{ P = \sum_{k\in\mathbb{N}} a_k X^k \mid \forall k,\ a_k \in \mathbb{K} \right\}$ les polynômes à coefficients dans $\mathbb{K}$;
- $\mathbb{K}_n[X] = \left\{ P = \sum_{k=0}^{n} a_k X^k \mid \forall n,\ a_k \in \mathbb{K} \right\}$ les polynômes de degré inférieur à n ;
- $\{(x,y) \in \mathbb{R}^2 \mid x + 2y = 0\}$.

En revanche, les ensembles suivants ne sont pas des e.v.
- $\mathbb{Z}$: si on multiplie un entier par un réel ou un complexe, on ne reste pas entier ;
- $\mathbb{R}_+$: si on multiplie un nombre positif par un nombre négatif, il devient négatif.
- $A = \{(x, y) \in \mathbb{R}^2 \mid x + 2y = 1\}$: A n'est pas stable par addition (ni multiplication)

Proposition 1.2 *Soit E un $\mathbb{K}$-espace vectoriel, alors il existe un élément de E, noté 0_E (ou parfois simplement 0 en l'absence d'ambiguité) tel que on ait*

$$u + 0_E = u, \qquad \forall u \in E.$$

De plus, pour tout élément u, il existe un élément noté $-u$ tel que

$$u + (-u) = 0_E.$$

Définition 1.3 (combinaison linéaire) *Soit E un $\mathbb{K}$-espace vectoriel, et soient $u_1, \ldots, u_n$ des éléments de E. On dit qu'un élément u est* combinaison linéaire *des $(u_k)_{k=1,\ldots,n}$ s'il existe des scalaires $\lambda_1, \ldots, \lambda_n$ dans $\mathbb{K}$ tels que*

$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n = \sum_{k=1}^{n} \lambda_k u_k. \tag{1.1}$$

Comme E est stable par multiplication par un scalaire et par addition, alors toute combinaison linéaire d'éléments de E est élément de E.

Définition 1.4 (Sous-espace vectoriel) *Soit E un $\mathbb{K}$-e.v. et soit F un ensemble tel que*

1. *$F \neq \emptyset$: F contient au moins un élément ;*
2. *$F \subset E$: tout élément de F appartient à E ($\forall u \in F,\ u \in E$) ;*
3. *$\forall u, v \in F,\ \forall \lambda \in \mathbb{K},\ u + \lambda v \in F$.*

Alors on dit que F est un sous-espace vectoriel *(s.e.v.) de E.*

Il est intéressant de remarquer le point 3 de la Définition 1.4 est équivalent à demander que

$$\forall u, v \in F, \quad u + v \in F \quad \text{ et que } \quad \forall u \in F,\ \forall \lambda \in \mathbb{K}, \quad \lambda u \in F.$$

On déduit alors des définitions 1.1 et 1.4 la Proposition suivante.

Proposition 1.5 *Un sous-espace vectoriel F de E est un $\mathbb{K}$-espace vectoriel.*

Exemple:
- Soit E un $\mathbb{K}$-e.v., alors E et $\{0_E\}$ sont deux s.e.v. de E.
- $E = \mathbb{R}^2$, $F = \{(x, y) \in E \mid\ x - 3y = 0\}$.
- $E = \mathbb{K}[X]$, $F = \mathbb{K}_n[X]$.

Proposition 1.6 *Soit E un $\mathbb{K}$-e.v., soit F un s.e.v. de E et G un s.e.v. de F, alors G est un s.e.v. de E.*

Définition 1.7 (espace engendré) *Soit E un $\mathbb{K}$-e.v. et soit $(\mathbf{u}_1, \dots, \mathbf{u}_m)$ une famille de vecteurs de E, alors on définit l'*espace engendré *par $\mathbf{u}_1, \dots, \mathbf{u}_m$, noté $\mathrm{Vect}(\mathbf{u}_1, \dots, \mathbf{u}_m)$ comme l'ensemble des combinaisons linéaires des $\mathbf{u}_k$:*

$$\mathrm{Vect}(\mathbf{u}_1, \dots, \mathbf{u}_m) = \left\{ \mathbf{u} \in E \,\middle|\, \quad \exists \lambda_1, \dots, \lambda_m \quad t.q. \quad \mathbf{u} = \sum_{k=1}^{m} \lambda_k \mathbf{u}_k \right\}.$$

Proposition 1.8 *Soit E un $\mathbb{K}$-e.v. et soit $(\mathbf{u}_1, \dots, \mathbf{u}_m)$ une famille de vecteurs de E, alors l'espace $\mathrm{Vect}(\mathbf{u}_1, \dots, \mathbf{u}_m)$ engendré par les $\mathbf{u}_k$ est un sous-espace vectoriel de E. C'est même le plus petit s.e.v. de E contenant tous les $\mathbf{u}_k$.*

1.2 Famille libre, famille génératrice, base, dimension

Définition 1.9 (famille libre, famille liée) *Soit E un $\mathbb{K}$-e.v. et soit $(u_k)_{k=1,\dots,n}$ une famille (finie) d'éléments de E. On dit que la famille est* libre *si toute combinaison linéaire nulle a tous ses coefficients nuls :*

$$\sum_{k=1}^{n} \lambda_k u_k = 0_E \quad \Leftrightarrow \quad \lambda_1 = \dots = \lambda_n = 0. \tag{1.2}$$

Une famille qui n'est pas libre est dite liée.

Il est à noter que dans l'équivalence (1.2), l'implication $\Leftarrow$ est toujours vraie. Pour montrer qu'une famille est libre, il suffit donc de montrer que

$$\sum_{k=1}^{n} \lambda_k u_k = 0_E \quad \Rightarrow \quad \lambda_1 = \dots = \lambda_n = 0.$$

Le fait que $(u_k)_{k=1,\dots,n}$ soit une famille libre signifie que chaque élément u_k de la famille est indépendant des autres éléments. Cela se traduit par cette caractérisation des familles liées : on peut exprimer un vecteur en fonction des autres.

Proposition 1.10 *Soit $(u_k)_{k=1,\dots,n}$ une famille liée d'éléments de E, alors il existe $k_0 \in \{1, \dots, n\}$ et des scalaires μ_k pour $k \neq k_0$ tels que $u_{k_0} = \sum_{k \neq k_0} \mu_k u_k$.*

Exemple:

- Toute famille d'éléments de E contenant le vecteur nul est liée.
- La famille $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ de $\mathbb{R}^3$ où

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

  est liée car $\mathbf{u}_1 = \mathbf{u}_2 + \mathbf{u}_3$.
- La famille $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ de $\mathbb{R}^3$ où

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

est libre car

$$\lambda_1\mathbf{v}_1+\lambda_2\mathbf{v}_2+\lambda_3\mathbf{v}_3=0 \quad \Leftrightarrow \quad \begin{cases}\lambda_1+\lambda_2=0\\ -2\lambda_1-\lambda_2+\lambda_3=0\\ 3\lambda_1+2\lambda_2+\lambda_3=0.\end{cases}$$

$$\Leftrightarrow \quad \begin{cases}\lambda_1=-\lambda_2\\ \lambda_2+\lambda_3=0\\ -\lambda_2+\lambda_3=0\end{cases} \quad \Leftrightarrow \quad \lambda_1=\lambda_2=\lambda_3=0.$$

Définition 1.11 (famille génératrice) *Soit E un $\mathbb{K}$-e.v., et soit $(u_k)_{k=1,\dots,n}$ une famille (finie) d'éléments de E. On dit que la famille est* génératrice *si tout élément v de E s'écrit comme combinaison linéaire des u_k :*

$$\forall v\in E, \quad \exists \lambda_1,\dots,\lambda_n\in\mathbb{K} \quad t.q. \quad v=\sum_{k=1}^{n}\lambda_k u_k.$$

Exemple:

1. La famille $(1, X+1, X-2, \frac{1}{2}X^2-X+1)$ est génératrice de l'espace $\mathbb{R}_2[X]$ des polynômes de degré inférieur ou égal à 2. En effet, pour $P\in\mathbb{R}_2[X]$, il existe a_0, a_1 et a_2 tels que $P=a_0+a_1X+a_2X^2$. On a aussi alors
$$P=2a_2\left(\frac{1}{2}X^2-X+1\right)+(a_1+a_2)(X+1)+a_2(X-2)+(a_0-a_1-a_2)1.$$
Ce choix n'est pas nécessairement unique, puisque
$$P=2a_2\left(\frac{1}{2}X^2-X+1\right)+(a_1+2a_2)(X+1)+(a_0-a_1-4a_2)1.$$
2. Soit E un $\mathbb{K}$-e.v. et soit $(\mathbf{u}_1,\dots,\mathbf{u}_m)$ une famille de vecteurs de E, alors la famille $(\mathbf{u}_1,\dots,\mathbf{u}_m)$ est génératrice du sous-espace $\mathrm{Vect}(\mathbf{u}_1;\dots,\mathbf{u}_m)$ de E.

Définition 1.12 (base de E) *Soit E un $\mathbb{K}$-e.v., et soit $(e_k)_{k=1,\dots,n}$ une famille (finie) d'éléments de E. On dit que la famille $(e_k)_{k=1,\dots,n}$ est une* base *de E si elle est à la fois libre et génératrice.*

Exemple: On considère le cas important où $E=\mathbb{K}^n$. Pour $j\in\{1,\dots,n\}$, on note

$$\mathbf{e}_1=\begin{pmatrix}1\\0\\ \vdots\\0\\ \vdots\\ \vdots\\0\end{pmatrix}, \qquad \mathbf{e}_j=\begin{pmatrix}0\\ \vdots\\0\\1\\0\\ \vdots\\0\end{pmatrix}\leftarrow j^{\text{ème}} \text{ ligne}, \qquad \mathbf{e}_n=\begin{pmatrix}0\\0\\ \vdots\\0\\ \vdots\\0\\1\end{pmatrix},$$

alors $(\mathbf{e}_1,\dots,\mathbf{e}_n)$ forme une base de $\mathbb{K}^n$ appelée **base canonique de** $\mathbb{K}^n$.

Théorème 1.13 (dimension de E) *Soit E un $\mathbb{K}$-e.v. admettant une base $(e_1, \ldots, e_n)$, alors toute base de E compte exactement n éléments. Ce nombre n est appelé dimension de E.*

Exemple: Soient $e_1 = (1, 0, -3)$ et $e_2 = (0, 1, 1)$ deux vecteurs de $\mathbb{R}^3$. La famille (e_1, e_2) forme une base du sous espace vectoriel de $\mathbb{R}^3$ constitué des vecteurs $v = (x, y, z)$ tels que

$$3x - y + z = 0.$$

Ce sous espace vectoriel est donc de dimension 2.

Théorème 1.14 *Soit E un $\mathbb{K}$-e.v. de dimension n.*

1. *Soit $(e_1, \ldots, e_p)$ une famille libre de E, alors $p \leq n$. De plus, il existe $(n-p)$ vecteurs $(e_{p+1}, \ldots, e_n)$ de E tels que $(e_1, \ldots e_n)$ soit une base.*
2. *Soit $(e_1, \ldots, e_p)$ une famille génératrice de E, alors $p \geq n$. De plus, on peut choisir n vecteurs $(e_{k_1}, \ldots, e_{k_n})$ parmi $(e_1, \ldots, e_p)$ tels que $(e_{k_1}, \ldots e_{k_n})$ soit une base de E.*

Corollaire 1.15 *Soit E un $\mathbb{K}$-e.v. de dimension n, alors*

1. *toute famille libre de n vecteurs de E est une base de E ;*
2. *toute famille génératrice de n vecteurs de E est une base de E.*

Théorème 1.16 *Soit E un $\mathbb{K}$-e.v. et soit $(e_1, \ldots, e_n)$ une base de E. Soit $u \in E$, alors il existe un unique n-uplet $(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$ tel que*

$$u = \lambda_1 e_1 + \cdots + \lambda_n e_n = \sum_{k=1}^{n} \lambda_k e_k. \tag{1.3}$$

Remarque: Le théorème précédent affirme que tout élément de E s'écrit de manière unique comme combinaison linéaire des éléments d'une base. Étant donné une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E, un élément u peut être caractérisé entièrement par ses coefficients λ_k de la formule (1.3). On peut donc identifier (via le choix d'une base) tout $\mathbb{K}$-e.v. de dimension n avec $\mathbb{K}^n$ (les coordonnées dans cette base).

Théorème 1.17 *Soit E un $\mathbb{K}$-e.v. de dimension finie, et soit F un s.e.v. de E, alors*

$$\dim F = \dim E \quad \Leftrightarrow \quad F = E.$$

1.3 Somme et intersection d'espaces vectoriels, sous-espaces supplémentaires

Définition 1.18 (intersection de deux s.e.v.) *Soient F et G deux s.e.v. de E, alors on définit*

$$F \cap G = \{u \in E \mid u \in F \text{ et } u \in G\}.$$

Proposition 1.19 *L'ensemble $F \cap G$ est un s.e.v. de E.*

Définition 1.20 (somme de deux s.e.v) *Soient F et G deux s.e.v. de E, alors on définit*

$$F + G = \{u \in E \mid \text{ il existe } u_F \in F \text{ et } u_G \in G \text{ tels que } u = u_F + u_G\}.$$

Proposition 1.21 *L'ensemble $F + G$ est un s.e.v. de E.*

Définition 1.22 *Deux s.e.v. F et G sont dits* supplémentaires *dans E si $F + G = E$ et $F \cap G = \{0_E\}$. On note $E = F \oplus G$*

Proposition 1.23 *Soit E un $\mathbb{K}$-e.v. et soient F, G deux s.e.v. de E, alors $E = F \oplus G$ si et seulement si, pour tout $u \in E$, il existe un unique $u_F \in F$ et un unique $u_G \in G$ tel que*

$$u = u_F + u_G.$$

Preuve : Pour prouver l'équivalence ci-dessus, on prouve les implications dans les deux sens.

"$\Rightarrow$" On suppose que $E = F \oplus G$, alors en particulier $E = F + G$. Soit $u \in E$, il existe $u_F \in F$ et $u_G \in G$ tels que $u = u_F + u_G$. On suppose que u s'écrit aussi $u = v_F + v_G$ avec $v_F \in F$ et $v_G \in G$, alors

$$\underbrace{u_F - v_F}_{\text{dans } F} = \underbrace{v_G - u_G}_{\text{dans } G} \in F \cap G = \{0_E\}.$$

On en déduit donc que $u_F = v_F$ et que $u_G = v_G$, la décomposition de u en un élément de F et un élément de G est donc unique.

"$\Leftarrow$" On suppose que : $\forall\ u \in E$, $\exists!\ (u_F, u_G) \in F \times G$ tel que $u = u_F + u_G$. En particulier, l'existence d'une telle décomposition assure que $E = F + G$. Il suffit donc de montrer que $F \cap G = \{0_E\}$. Procédons par l'absurde et supposons que cela soit faux, c'est à dire qu'il existe $v \neq 0_E$ dans $F \cap G$. Dans ce cas, 0_E peut se décomposer de deux façons :

$$0_E = \underbrace{0_E}_{\in F} + \underbrace{0_E}_{\in G} = \underbrace{v}_{\in F} - \underbrace{v}_{\in G},$$

ce qui contredit l'hypothèse de départ.

□

Proposition 1.24 *Soient F, G deux s.e.v. supplémentaires dans E. Soit $(f_1, \ldots, f_p)$ une base de F et $(g_1, \ldots, g_q)$ une base de G, alors $(f_1, \ldots, f_p, g_1, \ldots, g_q)$ est une base de E. En particulier, cela implique que*

$$\dim(F) + \dim(G) = \dim(E).$$

Preuve : Soit $u \in E$, alors il existe un unique couple $(u_F, u_G) \in F \times G$ tel que $u = u_F + u_G$. Comme $(f_1, \ldots, f_p)$ (resp. $(g_1, \ldots, g_q)$ est une base de F (resp. de G), il existe un unique $(a_1, \ldots, a_p) \in \mathbb{K}^p$ (resp. $(b_1, \ldots, b_q) \in \mathbb{K}^q$) tel que

$$u_F = \sum_{i=1}^{p} a_i f_i, \qquad u_G = \sum_{i=1}^{q} b_i g_i.$$

Dès lors, tout élément $u \in E$ se décompose de manière unique comme combinaison linéaire de $(f_1, \ldots, f_p, g_1, \ldots, g_q)$, donc la famille $(f_1, \ldots, f_p, g_1, \ldots, g_q)$ est une base de E. □

Théorème 1.25 (existence d'un supplémentaire) *Soit E un espace vectoriel de* dimension finie*, et soit F un s.e.v. de E, alors il existe un s.e.v. G de E tel que $E = F \oplus G$.*

Chapitre 2

Applications linéaires

Dans ce chapitre, on s'intéresse à une famille d'applications qui à un vecteur de E associe un vecteur de F : les applications linéaires. Ces dernières sont très utilisées en physique, traitement du signal, ...

Dans la suite de ce cours, on supposera toujours, sauf mention explicite du contraire, que les e.v. sont de dimension finie.

2.1 Définition et premières propriétés

Définition 2.1 (application linéaire) *Soient E et F deux $\mathbb{K}$-e.v., et soit $f : \begin{cases} E \to F \\ u \to f(u) \end{cases}$ une application vérifiant*

$$\forall u, v \in E,\ \forall \lambda \in \mathbb{K}, \qquad f(u + \lambda v) = f(u) + \lambda f(v), \tag{2.1}$$

alors on dit que f est une application linéaire *(ou* morphisme*) de E dans F.*

Remarque: On pourrait remplacer de manière équivalente (2.1) par

$$\forall u, v \in E, \quad f(u + v) = f(u) + f(v) \qquad \text{et} \qquad \forall u \in E,\ \forall \lambda \in \mathbb{K}, \quad f(\lambda u) = \lambda f(u).$$

Exemple: L'application $\varphi : \begin{array}{ccc} \mathcal{C}([0,1], \mathbb{R}) & \to & \mathbb{R} \\ f & \mapsto & \int_0^1 f(t)dt \end{array}$ est linéaire.

Proposition 2.2 *Soit f une application linéaire de E dans F, alors nécessairement $f(0_E) = 0_F$.*

Proposition 2.3 *On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F, alors $\mathcal{L}(E, F)$ est un espace vectoriel.*

Preuve : Tout d'abord, $\mathcal{L}(E, F)$ est non vide, car il contient l'application nulle $0_{\mathcal{L}(E,F)}$ définie par

$$\forall u \in E, \qquad 0_{\mathcal{L}(E,F)}(u) = 0_F,$$

dont la linéarité est facile à vérifier. Pour $f, g \in \mathcal{L}(E, F)$, on définit l'application $f+g$ par $(f+g)(u) = f(u) + g(u)$. On a alors

$$\begin{aligned}(f+g)(u+\lambda v) =& f(u+\lambda v) + g(u+\lambda v)\\ =& f(u) + g(u) + \lambda(f(v) + g(v))\\ =& (f+g)(u) + \lambda(f+g)(v),\end{aligned}$$

ce qui assure que $f + g$ est linéaire. Pour $\mu \in \mathbb{K}$, on définit μf par $(\mu f)(u) = \mu f(u)$. On obtient alors

$$(\mu f)(u+\lambda v) = \mu f(u+\lambda v) = \mu(f(u) + \lambda f(v)) = (\mu f)(u) + \lambda(\mu f)(v),$$

ce qui prouve que μf est linéaire et achève donc la preuve. □

Définition 2.4 (composition de morphismes) *Soient E, F et G trois $\mathbb{K}$-e.v., soit $v \in \mathcal{L}(E, F)$ et $u \in \mathcal{L}(F, G)$, alors on définit la* composée *$u \circ v$ de u et v par*

$$\forall x \in E, \qquad (u \circ v)(x) = u(v(x)).$$

Il faut faire attention que toutes les quantités écrites aient un sens. Ici, x appartient à E, donc $v(x)$ appartient à F car $v \in \mathcal{L}(E, F)$. On peut donc appliquer u à $v(x)$, et on obtient alors un élément de G puisque $u \in \mathcal{L}(F, G)$. On en déduit que $u \circ v$ envoie un élément $x \in E$ sur un élément $u \circ v(x) \in G$. La proposition ci-dessous dit que la composée de deux applications linéaires est aussi linéaire.

Proposition 2.5 *Soit $v \in \mathcal{L}(E, F)$ et $u \in \mathcal{L}(F, G)$, alors $u \circ v \in \mathcal{L}(E, G)$.*

2.2 Noyau, image d'un morphisme, Théorème du rang

Définition 2.6 (Noyau d'un application linéaire) *Soit E, F deux $\mathbb{K}$-e.v., et soit $f \in \mathcal{L}(E, F)$. On définit le sous-ensemble de E noté* $\mathrm{Ker} f$ *et appelé* noyau de f *par*

$$\mathrm{Ker} f = \{x \in E \mid f(x) = 0_F\}.$$

Proposition 2.7 *Soit $f \in \mathcal{L}(E, F)$, alors $\mathrm{Ker} f$ est un sous espace vectoriel de E.*

Définition 2.8 (Image d'un application linéaire) *Soit E, F deux $\mathbb{K}$-e.v., et soit $f \in \mathcal{L}(E, F)$. On définit le sous-ensemble de F noté* $\mathrm{Im} f$ *et appelé* image de f *par*

$$\mathrm{Im} f = \{y \in F \mid \exists\, x \in E \ \ t.\ q.\ f(x) = y\}.$$

Proposition 2.9 *Soit $f \in \mathcal{L}(E, F)$, alors $\mathrm{Im} f$ est un sous espace vectoriel de F.*

Définition 2.10 (rang d'un application) *On appelle* rang d'une application linéaire $f \in \mathcal{L}(E, F)$ *la quantité*

$$\mathrm{rg} f = \dim \mathrm{Im} f.$$

Théorème 2.11 (Théorème du rang) *Soit E, F des e.v. de dimension finie et soit $f \in \mathcal{L}(E, F)$, alors*

$$\dim \mathrm{Ker} f + \mathrm{rg} f = \dim E.$$

Définition 2.12 (application injective, surjective, bijective) *Une application* $f \in \mathcal{L}(E, F)$ *est dite :*

1. *injective si* $\mathrm{Ker} f = \{0_E\}$ *;*
2. *surjective si* $\mathrm{Im} f = F$ *;*
3. *bijective si elle est injective et surjective.*

Proposition 2.13 *On a les propriétés suivantes :*

1. *si* $f \in \mathcal{L}(E; F)$ *est injective, alors* $\dim F \geq \dim E$ *;*
2. *si* $f \in \mathcal{L}(E; F)$ *est surjective, alors* $\dim F \leq \dim E$ *;*
3. *si* $f \in \mathcal{L}(E; F)$ *est bijective, alors* $\dim F = \dim E$*.*

On déduit ensuite du théorème du rang (Théorème 2.11) le thórème suivant.

Théorème 2.14 *Soit* $f \in \mathcal{L}(E, F)$*, avec* E *et* F *de dimensions finies. Alors*

1. f *est injective si et seulement si* $rg(f) = \dim E$*.*
2. f *est surjective si et seulement si* $rg(f) = \dim F$*.*

En particulier, si $\dim E = \dim F$*,* f *est injective si et seulement* f *est surjective.*

Ainsi, lorsque les espaces E et F ont la même dimension, pour montrer qu'une application f est bijective il suffit de montrer qu'elle est injective ou surjective.

Exemple: Soient a_0, a_1, a_2 des réels distincts, f une application continue de $\mathbb{R}$ dans $\mathbb{R}$. Le théorème 2.14 permet de montrer qu'il existe un unique polnyôme $P \in \mathbb{R}_2[X]$ interpolant f aux points a_0, a_1 et a_2, c'est-à-dire tel que $P(a_0) = f(a_0)$, $P(a_1) = f(a_1)$, $P(a_2) = f(a_2)$.

Proposition 2.15 *Soit* $f \in \mathcal{L}(E; F)$ *une application injective, alors* $\tilde{f} : \begin{cases} E \to \mathrm{Im} f \\ u \mapsto f(u) = \tilde{f}(u) \end{cases}$ *est bijective.*

Proposition 2.16 *Soit* $f \in \mathcal{L}(E, F)$ *et* $g \in \mathcal{L}(F, G)$*, alors*

$$\mathrm{Ker} f \subset \mathrm{Ker}(g \circ f), \qquad \mathrm{Im}(g \circ f) \subset \mathrm{Im} g.$$

2.3 Cas particulier des endomorphismes $\mathcal{L}(E)$

On considère maintenant un cas particulier : le cas où l'espace de départ coïncide avec l'espace d'arrivée, c'est à dire $\mathcal{L}(E; E)$. Un application linéaire de E dans E est appelé un *endomorphisme* de E. Dans ce cas, on note $\mathcal{L}(E)$ (à la place de $\mathcal{L}(E; E)$) l'ensemble des endomorphismes de E.

On donne maintenant un théorème important qui est une conséquence directe des théorèmes 1.17 et 2.11.

Théorème 2.17 *Soit* f *un endomorphisme, alors les trois propriétés suivantes sont équivalentes :*

1. f *est injectif;*
2. f *est surjectif :*

3. f est bijectif.

Un endomorphisme joue un rôle particulier : l'identité de E.

Définition 2.18 (identité de E) *On définit l'identité de E, noté* Id_E *comme l'élément de $\mathcal{L}(E)$ tel que*

$$\forall x \in E, \qquad Id_E(x) = x.$$

Soit f un endomorphisme bijectif de E (parfois appelé *automorphisme* de E) alors chaque élément de E admet un unique antécédent :

$$\forall y \in E, \quad \exists\, !\, x \in E \quad \text{t.q.} \quad f(x) = y. \tag{2.2}$$

On peut donc définir l'application f^{-1} qui à y associe cet unique x dans (2.2).

Définition 2.19 (application inverse) *Soit f un endomorphisme bijectif de E, alors il existe un unique endomorphisme de E noté f^{-1} et appelé* inverse de f *tel que*

$$\forall x \in E, \qquad f \circ f^{-1}(x) = f^{-1} \circ f(x) = x,$$

ce qui s'écrit aussi $f \circ f^{-1} = f^{-1} \circ f = \mathrm{Id}_E$.

Proposition 2.20 *Soit f un endomorphisme bijectif de E, alors l'endomorphisme f^{-1} est aussi bijectif, et $\left(f^{-1}\right)^{-1} = f$.*

Définition 2.21 (groupe linéaire de E) *On appelle* groupe linéaire de E *l'ensemble noté $\mathcal{G}\ell(E)$ des endomorphismes bijectifs (ou automorphismes) de E.*

Remarque: $\mathcal{G}\ell(E)$ n'est pas un sous-espace vectoriel de $\mathcal{L}(E)$. En effet, $0_{\mathcal{L}(E)}$ n'est évidemment pas bijectif et n'appartient donc pas à $\mathcal{G}\ell(E)$. Le fait que $\mathcal{G}\ell(E)$ ne puisse pas être un espace vectoriel découle alors de la Proposition 1.2.

Si l'ensemble des automorphismes n'est pas stable par addition et par multiplication externe comme le souligne la remarque précédente, cet ensemble est en revanche stable par composition : la composée de deux automorphismes est un automorphisme. C'est le sujet de la proposition suivante.

Proposition 2.22 *Soient $f, g \in \mathcal{G}\ell(E)$, alors $f \circ g$ et $g \circ f$ appartiennent à $\mathcal{G}\ell(E)$.*

Un des aspects qui donne une structure particulière à l'ensemble $\mathcal{L}(E)$ vient du fait que la composée de deux endomorphismes est encore un endomorphisme. cela permet en particulier d'itérer plusieurs fois le même endomorphisme.

Définition 2.23 (puissance d'endomorphisme) *Soit $f \in \mathcal{L}(E)$, on définit, pour n entier strictement positif,*

$$f^n = \underbrace{f \circ f \circ \cdots \circ f}_{n \text{ fois}}.$$

On définit aussi $f^0 = \mathrm{Id}_E$. Dans le cas où $f \in \mathcal{G}\ell(E)$, on peut aussi définir les puissance négative de f par

$$f^{-n} = \underbrace{f^{-1} \circ f^{-1} \circ \cdots \circ f^{-1}}_{n \text{ fois}}.$$

On déduit directement de la proposition 2.16 que

$$\mathrm{Ker} f^n \subset \mathrm{Ker} f^p \text{ si } n \leq p, \qquad \mathrm{Im} f^n \subset \mathrm{Im} f^p \text{ si } n \geq p,$$

ce qui traduit que plus on itère un endomorpisme f, plus le noyau de l'endomorphisme obtenu après itération grandit, alors que l'image rétrécit. Cela est cohérent avec le théorème du rang, qui impose que si le noyau grandit, l'image doit rétrécir.

On peut définir alors la notion de polynôme d'endomorphisme qui sera utile plus tard dans ce cours (cf. Chapitre 6).

Définition 2.24 (polynôme d'endomorphisme) *Soit* $P := \sum_{k=0}^{N} a_k X^k \in \mathbb{K}[X]$ *un polynôme à coefficients dans* $\mathbb{K}$*, et soit* $f \in \mathcal{L}(E)$*, on définit* $P(f) \in \mathcal{L}(E)$ *par*

$$P(f) := \sum_{k=0}^{N} a_k f^k.$$

Chapitre 3

Matrices

3.1 Représentation des morphismes de $\mathcal{L}(\mathbb{K}^n;\mathbb{K}^m)$

On considère une application linéaire $f \in \mathcal{L}(E,F)$ quelconque. Soit $\mathcal{B} = (e_1,\dots,e_n)$ une base de $E \simeq \mathbb{K}^n$, et soit $\mathcal{B}' = (e'_1,\dots,e'_m)$ une base de $F \simeq \mathbb{K}^m$. Soit $j \in \{1,\dots,n\}$, alors $f(e_j) \in F$ peut se décomposer de manière unique dans la base $\mathcal{B}'$:

$$f(e_j) = \sum_{i=1}^{m} a_{i,j} e'_i. \tag{3.1}$$

Soit maintenant $x \in E$, alors comme $\mathcal{B}$ est une base de E, il existe un unique n-uplet $(x_1,\dots,x_n) \in \mathbb{K}^n$ tel que

$$x = \sum_{j=1}^{n} x_j e_j. \tag{3.2}$$

En utilisant la linéarité de f et la décomposition (3.2) de x, on obtient que

$$f(x) = f\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} x_j f(e_j).$$

En utilisant en plus (3.1), alors on obtient que

$$f(x) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{i,j} x_j\right) e'_i. \tag{3.3}$$

En particulier, pour connaitre complètement f, il suffit de se donner des bases $\mathcal{B}$ et $\mathcal{B}'$ de E et F et de connaitre tous les coefficients a_{ij} $(1 \le i \le n, 1 \le j \le m)$ de la relation (3.1).

Définition 3.1 (matrice associée à un morphisme) *Soit E un $\mathbb{K}$-e.v. de dimension n muni d'une base $\mathcal{B} = (e_1,\dots,e_n)$, soit F un $\mathbb{K}$-e.v. de dimension m muni d'une base $\mathcal{B}' = (e'_1,\dots,e'_m)$,* *et soit $f \in \mathcal{L}(E,F)$. On appelle matrice de f dans les bases $\mathcal{B}$ et $\mathcal{B}'$ le tableau à m lignes et n colonnes contenant à la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne la quantité a_{ij} donnée par l'équation (3.1) :*

$$\mathrm{mat}_{\mathcal{B}',\mathcal{B}}(f) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_{(m-1)n} \\ a_{m,1} & \cdots & \cdots & a_{m,(n-1)} & a_{m,n} \end{pmatrix} =: (a_{i,j})_{1\le i\le m,\, 1\le j\le n} \tag{3.4}$$

Exemple: Si $E = \mathbb{R}^2$ muni de sa base canonique, la rotation r_θ d'angle θ dans le plan orienté a pour matrice

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Définition 3.2 *On note $\mathcal{M}_{m,n}(\mathbb{K})$ l'ensemble des matrices à m lignes et n colonnes à coefficients dans $\mathbb{K}$. Dans le cas particulier où $m = n$, on note $\mathcal{M}_n(\mathbb{K}) = \mathcal{M}_{n,n}(\mathbb{K})$.*

Pour $x \in E$ (resp. $y \in F$), on note $\mathbf{x}_{\mathcal{B}} \in \mathbb{K}^n$ (resp. $\mathbf{y}_{\mathcal{B}'} \in \mathbb{K}^m$) le vecteur colonne constitué des coordonnées de x (resp. y) dans la base $\mathcal{B}$ (resp. $\mathcal{B}'$) :

$$\mathbf{x}_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \mathbf{y}_{\mathcal{B}'} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

La $j^{\text{ième}}$ colonne de la matrice $\mathrm{mat}_{\mathcal{B}',\mathcal{B}}(f)$ correspond alors à la décomposition $(\mathbf{c}_j)_{\mathcal{B}'}$ de $f(e_j)$ dans la base $\mathcal{B}'$. La matrice $\mathrm{mat}_{\mathcal{B}',\mathcal{B}}(f)$ contient donc l'image de la base $\mathcal{B}$, et donc l'image de tout élément de E.

On déduit de la formule (3.3) la proposition suivante.

Proposition 3.3 *Soit $x = \sum_{j=1}^{n} x_j e_j \in E$ et $y = f(x) \in F$, alors si $\mathrm{mat}_{\mathcal{B}',\mathcal{B}}(f)$ est donnée par (3.4), on a*

$$\mathbf{y}_{\mathcal{B}'} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \qquad \text{où} \quad y_i = \sum_{j=1}^{n} a_{i,j} x_j.$$

Remarque: Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$, $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ la base canonique de $\mathbb{K}^n$ et $\mathcal{B}' = (e'_1, e'_2, \ldots, e'_m)$ la base canonique de $\mathbb{K}^m$. On peut définir une application linéaire de $\mathbb{K}^n$ vers $\mathbb{K}^m$ en posant

$$\varphi : \left\{ \begin{array}{ccc} \mathbb{K}^n & \to & \mathbb{K}^m \\ \mathbf{x} & \to & A\mathbf{x} \end{array} \right. .$$

On vérifie alors facilement que $\mathrm{mat}_{\mathcal{B}',\mathcal{B}}(\varphi) = A$

Pour $A = (a_{i,j})$, $B = (b_{i,j})$ deux matrices de $\mathcal{M}_{m,n}(\mathbb{K})$, on définit

$$A + B = (a_{i,j} + b_{i,j})_{i,j} = \begin{pmatrix} a_{1,1} + b_{1,1} & \dots & \dots & a_{1,n} + b_{1,n} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a_{m,1} + b_{m,1} & \dots & \dots & a_{m,n} + b_{m,n} \end{pmatrix},$$

et pour $\lambda \in \mathbb{K}$,

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

Proposition 3.4 *L'ensemble des matrices à m lignes et n colonnes à coefficients dans $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel de dimension $m \times n$ dont l'élément neutre est donné par*

$$0_{\mathcal{M}_{m,n}(\mathbb{K})} = \begin{pmatrix} 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix},$$

Comme cela a été vu précédemment, la jème colonne de la matrice associée à un endomorphisme f de la base $\mathcal{B}$ vers la base $\mathcal{B}'$ est constituée des coefficients du vecteur $f(e_j)$ exprimé dans la base $\mathcal{B}'$. On suppose maintenant que l'on a la configuration suivante

$$f : (E, \mathcal{B}) \to (F, \mathcal{B}') \quad \text{et} \quad g : (F, \mathcal{B}') \to (G, \mathcal{B}'').$$

La matrice $\mathrm{mat}_{\mathcal{B}'',\mathcal{B}}(g \circ f)$ est donc constituée des colonnes représentant les $g \circ f(e_j)$ ($e_j \in \mathcal{B}$) dans la base $\mathcal{B}'' = (e''_1, \dots, e''_p)$. On suppose que

$$\mathrm{mat}_{\mathcal{B}'',\mathcal{B}'}(g) = \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & \dots & b_{1,m} \\ b_{2,1} & b_{2,2} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & b_{(p-1),m} \\ b_{p,1} & \dots & \dots & b_{p,(m-1)} & b_{p,m} \end{pmatrix},$$

c'est-à-dire que $g(e'_k) = \displaystyle\sum_{i=1}^{p} b_{i,k} e''_i$. On a alors

$$g \circ f(e_j) = g\left(\sum_{k=1}^{m} a_{k,j} e'_k\right) = \sum_{k=1}^{m} a_{k,j}\, g(e'_k) = \sum_{k=1}^{m} a_{k,j} \sum_{i=1}^{p} b_{i,k} e''_i = \sum_{i=1}^{p} \left(\sum_{k=1}^{m} b_{i,k}\, a_{k,j}\right) e''_i. \quad (3.5)$$

Définition 3.5 (Produit de matrices) *Soient* $m, n, p \in \mathbb{N}^*$, $B = (b_{i,j})_{1 \leq i \leq p,\, 1 \leq j \leq m} \in \mathcal{M}_{p,m}(\mathbb{K})$, $A = (a_{i,j})_{1 \leq i \leq m,\, 1 \leq j \leq n} \in \mathcal{M}_{m,n}(\mathbb{K})$. *On définit la matrice* $C = (c_{i,j})_{1 \leq i \leq p,\, 1 \leq j \leq n} \in \mathcal{M}_{p,n}(\mathbb{K})$ *par*

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

La matrice C *est appelée* produit des matrices B et A, *et on note* $C = BA$.

Le produit de B par A n'est donc possible que si le nombre de lignes de A est égal au nombre de colonnes de B !

Exemple: Si $A = \begin{pmatrix} 2 & 3 & 5 \end{pmatrix}$, $B = \begin{pmatrix} 7 \\ -1 \\ 2 \end{pmatrix}$, et $C = \begin{pmatrix} 7 & \pi & -2 \\ 0 & -1 & 1.4 \\ 2 & e^1 & -1 \end{pmatrix}$, les produits AB, BA, AC et CB sont possibles, mais pas les produits CA et BC.

On déduit de la formule (3.5) la proposition suivante.

Proposition 3.6 *Soit* E, F *et* G *des* $\mathbb{K}$*-e.v. de dimension finis. Soit* $\mathcal{B}$ *une base de* E, $\mathcal{B}'$ *une base de* F *et* $\mathcal{B}''$ *une base de* G. *Si* $f \in \mathcal{L}(E, F)$ *et* $g \in \mathcal{L}(F, G)$ *on a*

$$\operatorname{mat}_{\mathcal{B}'',\mathcal{B}}(g \circ f) = \operatorname{mat}_{\mathcal{B}'',\mathcal{B}'}(g) \operatorname{mat}_{\mathcal{B}',\mathcal{B}}(f).$$

Remarque:

1. Il faut faire très attention lorsque l'on multiplie des matrices : pour deux matrices carrées A et B, on a en général
$$AB \neq BA.$$
C'est par exemple le cas pour
$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$
On dit que le **produit de matrices n'est pas commutatif**. On distinguera alors la multiplication à gauche et la multiplication à droite pour le produit de matrices. Cependant, dans certains cas, on a $AB = BA$, et on dit alors que les matrice A et B commutent.
2. On peut avoir $AB = 0$ mais $A \neq 0$ et $B \neq 0$. Par exemple, on peut prendre
$$A = B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$
On dit que **le produit de matrices n'est pas intègre**.

3.2 Matrices surjectives, injectives, bijectives

Définition 3.7 *On dit que* $A \in \mathcal{M}_{m,n}(\mathbb{K})$ *est injective (resp. surjective, bijective) s'il existe un morphisme* $f \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ *et des bases* $\mathcal{B}, \mathcal{B}'$ *de* $\mathbb{K}^n, \mathbb{K}^m$ *pour lesquels* $A = \operatorname{mat}_{\mathcal{B}',\mathcal{B}}(f)$ *tel que* f *soit injectif (resp. surjectif, bijectif).*

Proposition 3.8 *Le fait que* $A = \mathrm{mat}_{\mathcal{B}',\mathcal{B}}(f)$ *soit injective (resp. surjective, bijective) ne dépend pas du choix des bases* $\mathcal{B}$ *et* $\mathcal{B}'$*. On peut donc se contenter de considérer les bases canoniques respectives de* $\mathbb{K}^n, \mathbb{K}^m$*.*

Définition 3.9 (noyau, image et rang d'un matrice) *Soit* $A \in \mathcal{M}_{m,n}(\mathbb{K})$*, alors on appelle* noyau de A *le sous-espace vectoriel de* $\mathbb{K}^n$ *noté* $\mathrm{Ker}(A)$ *défini par*

$$\mathrm{Ker}(A) = \{\mathbf{x} \in \mathbb{K}^n \mid A\mathbf{x} = 0_{\mathbb{K}^m}\}.$$

On appelle image de A *le sous-espace vectoriel de* $\mathbb{K}^m$ *noté* $\mathrm{Im}(A)$ *défini par*

$$\mathrm{Im}(A) = \{\mathbf{y} \in \mathbb{K}^m \mid \exists\, \mathbf{x} \in \mathbb{K}^n \ t.\ q.\ A\mathbf{x} = \mathbf{y}\}.$$

On appelle rang de A *l'entier, noté* $\mathrm{rg}(A)$ *défini comme la dimension de* $\mathrm{Im}(A)$*.*

Théorème 3.10 (Théorème du rang) *Soit* $A \in \mathcal{M}_{m,n}(\mathbb{K})$*, alors*

$$\dim \mathrm{Ker}(A) + \mathrm{rg}(A) = n.$$

Théorème 3.11 *Une matrice* $A \in \mathcal{M}_{m,n}(\mathbb{K})$ *est injective si et seulement si* $\dim \mathrm{Ker}(A) = 0$*. Elle est surjective si et seulement si* $\mathrm{rg}(A) = m$*.*

Proposition 3.12 *Une matrice* $A \in \mathcal{M}_{m,n}(\mathbb{K})$ *est*

1. *injective si et seulement si la famille engendrée par ses colonnes*

$$\left(C_1 = \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{pmatrix}, \ldots, C_n = \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{pmatrix} \right)$$

   est une famille libre de $\mathbb{K}^n$*. En particulier,* $m \geq n$ *;*
2. *surjective si et seulement si la famille* $(C_1, \ldots, C_n)$ *engendrée par ses colonnes est génératrice de* $\mathbb{K}^m$*. En particulier,* $m \leq n$ *;*
3. *bijective si et seulement si la famille* $(C_1, \ldots, C_n)$ *engendrée par ses colonnes est une base de* $\mathbb{K}^m$*. En particulier,* $n = m$*.*

Le théorème suivant est un conséquence du Théorème 2.17, ou encore de la proposition précédente et du Corollaire 1.15

Théorème 3.13 *Soit* $A \in \mathcal{M}_n(\mathbb{K})$*, alors les trois assertions suivantes sont équivalentes :*

1. A *est injective ;*
2. A *est surjective ;*
3. A *est bijective.*

Définition 3.14 *On appelle* matrice identité de $\mathbb{K}^n$ *la matrice* $I_n \in \mathcal{M}_n(\mathbb{K})$ *contenant des* 1 *sur la diagonale et des* 0 *en dehors :*

$$I_n = \begin{pmatrix} 1 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \ldots & \ldots & 0 & 1 \end{pmatrix}$$

La matrice identité tient son nom de la propriété suivante.

Proposition 3.15 *Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$, alors*

$$AI_n = A, \qquad I_m A = A.$$

Proposition 3.16 *Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$, alors A est*

1. *injective si et seulement si il existe une matrice $B \in \mathcal{M}_{n,m}(\mathbb{K})$ telle que $BA = I_n$*
2. *surjective si et seulement si il existe une matrice $B \in \mathcal{M}_{n,m}(\mathbb{K})$ telle que $AB = I_m$*

Théorème 3.17 *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice bijective, alors il existe une unique matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que*

$$AB = BA = I_n.$$

*On note alors $B = A^{-1}$, et l'on appelle A^{-1} l'*inverse *de A.*

Remarque: Le fait qu'une matrice A soit bijective est équivalent au fait que A possède un inverse. On dit souvent qu'une matrice est *inversible* pour dire qu'elle est bijective. L'ensemble des matrices inversibles est parfois appelé *groupe linéaire de* $\mathbb{K}^n$ et noté $GL_n(\mathbb{K})$.

Proposition 3.18 *Soit A une matrice inversible, alors $\left(A^{-1}\right)^{-1} = A$.*

Proposition 3.19 *Soient A, B deux matrices de $\mathcal{M}_n(\mathbb{K})$.*

1. *si A et B sont injectives, alors AB est injective.*
2. *si A et B sont surjectives, alors AB est surjective.*

En particulier, le produit de deux matrices inversibles est inversible. On a alors

$$(AB)^{-1} = B^{-1}A^{-1}.$$

3.3 Résolution de systèmes linéaires

On s'intéresse maintenant à la résolution de système linéaires de n équations à n inconnues.

Problème : Pour $A = (a_{i,j})_{1\le i,j\le n} \in \mathcal{M}_n(\mathbb{K})$ et $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{K}^n$, trouver $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$ tel que

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1, \\ \vdots \\ a_{n,1}x_1 + \cdots + a_{n,n}x_n = b_n. \end{cases} \tag{3.6}$$

Au niveau matriciel, le système (3.6) s'écrit

$$A\mathbf{x} = \mathbf{b}. \tag{3.7}$$

Proposition 3.20 *Le système* (3.6) *admet une unique solution si et seulement si A est inversible.*

Preuve : Supposons tout d'abord que A est inversible, alors A est injective et surjective. Comme A est surjective, pour tout $\mathbf{b} \in \mathbb{K}^n$, il existe $\mathbf{x}$ dans $\mathbb{K}^n$ tel que $A\mathbf{x} = \mathbf{b}$, donc le systeme (3.6) admet au moins une solution. Soient maintenant $\mathbf{x}_1$ et $\mathbf{x}_2$ deux solutions du système, i.e. telles que

$$A\mathbf{x}_1 = A\mathbf{x}_2 = \mathbf{b},$$

alors

$$A(\mathbf{x}_1 - \mathbf{x}_2) = 0_{\mathbb{K}_n},$$

et comme A est injective, $\mathbf{x}_1 - \mathbf{x}_2 = 0$. On en déduit que $\mathbf{x}_1 = \mathbf{x}_2$ et que la solution est unique.

Supposons maintenant que A n'est pas inversible, alors A n'est ni bijective, ni injective d'après le Théorème 3.13. En particulier, $\mathrm{Im}(A)$ est strictement inclus dans $\mathbb{K}^n$, et si $\mathbf{b} \notin \mathrm{Im}(A)$, le système (3.6) n'admet pas de solution. Maintenant, si $\mathbf{b} \in \mathrm{Im}(A)$, alors le système admet au moins une solution $\mathbf{x}$, mais celle-là n'est pas unique. En effet, comme A n'est pas injective, $\mathrm{Ker}(A) \neq \{0_{\mathbb{K}^n}\}$ et il existe un élément non nul $\mathbf{x}_0$ tel que $A\mathbf{x}_0 = 0_{\mathbb{K}^n}$. Il est facile alors de vérifier que

$$A(\mathbf{x} + \mathbf{x}_0) = A\mathbf{x} + A\mathbf{x}_0 = b + 0_{\mathbb{K}^n},$$

et donc que $\mathbf{x} + \mathbf{x}_0$ est aussi solution. La solution ne peut donc pas être unique. □

À partir de maintenant, on ne considèrera plus que le cas où la matrice A est inversible. En multipliant la relation (3.7) à gauche par A^{-1} (qui existe puisque A est inversible), on remarque que

$$\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}. \tag{3.8}$$

Donc si l'on connait la matrice A^{-1}, on obtient une formule explicite pour la solution $\mathbf{x}$ du système. Cependant, inverser rapidement une matrice de grande taille est un problème très compliqué qui revient à résoudre n systèmes linéaires différents. Donc pour résoudre le système linéaire (3.6), on n'inverse pas la matrice A.

Algorithme du pivot de Gauss Pour résoudre un système linéaire, on peut utiliser l'algorithme dit du pivot de Gauss qui consiste à faire des combinaisons des lignes du système (3.6).

— **Initialisation :** On pose $A^{(0)} = \left(a_{i,j}^{(0)}\right) = A \in \mathcal{M}_n\mathbb{K}$, et $\mathbf{b}^{(0)} = \left(b_i^{(0)}\right) = \mathbf{b}$. Le problème devient donc de résoudre le système de taille n :

$$A^{(0)}\mathbf{x}^{(0)} = \mathbf{b}^{(0)}.$$

— **Itération :** Pour $k \geq 0$, on suppose que l'on a un système de taille $n-k$ à résoudre :

$$A^{(k)}\mathbf{x}^{(k)} = \mathbf{b}^{(k)}, \qquad \text{où } A^{(k)} \in \mathcal{M}_{n-k}(\mathbb{K}) \text{ et } \ \mathbf{b}^{(k)} \in \mathbb{K}^{n-k}, \tag{3.9}$$

et l'on veut se ramener à la résolution d'un système de taille $n-k-1$. Clairement, le processus s'arrête au bout de $n-1$ itérations, puisque la résolution d'un système de taille 1 est évidente.

1. choix d'un pivot : On suppose que la matrice $A^{(k)}$ est inversible (c'est la cas d'après le Lemme 3.21), il existe au moins un coefficient $a_{i_1,1}^{(k)} \neq 0$ dans la première colonne de la matrice. Si $a_{1,1}^{(k)} = 0$, on échange la ligne i_1 avec la première ligne. De manière à se ramener à une matrice $\tilde{A}^{(k)}$ donc le premier coefficient $\tilde{a}_{1,1}^{(k)}$ est non nul. Dans la suite, pour alléger les notations, on ne mettra pas le ˜ sur les coefficients, mais on supposera que $a_{1,1}^{(k)} \neq 0$.

2. <u>élimination des coefficients de la première colonne de $A^{(k)}$</u> : Il s'agit de remplacer la ligne i ($i \geq 2$) du système (3.9) par "ligne $i - \frac{a_{i,1}^{(k)}}{a_{1,1}^{(k)}}$ligne 1". Le nouveau système se réécrit alors

$$\left\{\begin{array}{llll} a_{1,1}^{(k)}x_1^{(k)} & + a_{1,2}^{(k)}x_2^{(k)} & + \cdots + a_{1,n-k}^{(k)}x_{n-k}^{(k)} & = b_1^{(k)}, \\ 0 & + \bar{a}_{2,2}^{(k)}x_2^{(k)} & + \cdots + \bar{a}_{2,n-k}^{(k)}x_{n-k}^{(k)} & = \bar{b}_2^{(k)}, \\ \vdots & & & \vdots \\ 0 & + \bar{a}_{n-k,2}^{(k)}x_1^{(k)} & + \cdots + \bar{a}_{n-k,n-k}^{(k)}x_{n-k}^{(k)} & = \bar{b}_{n-k}^{(k)}, \end{array}\right. \tag{3.10}$$

où, pour tout $i \geq 2$, on a

$$\bar{a}_{i,j}^{(k)} = a_{i,j}^{(k)} - \frac{a_{i,1}^{(k)}}{a_{1,1}^{(k)}} a_{1,j}^{(k)} \qquad \text{et} \qquad \bar{b}_i^{(k)} = b_j^{(k)} - \frac{a_{i,1}^{(k)}}{a_{1,1}^{(k)}} b_1^{(k)}.$$

En particulier, on a $\bar{a}_{i,1}^{(k)} = 0$ pour tout $i \geq 2$. Il suffit alors de poser

$$a_{i,j}^{(k+1)} = \bar{a}_{i+1,j+1}^{(k)}, \qquad b_i^{(k+1)} = \bar{b}_{i+1}^{(k)} \qquad \text{et} \qquad x_j^{(k+1)} = x_{j+1}^{(k)}$$

pour se retrouver avec un système $A^{(k+1)}\mathbf{x}^{(k+1)} = \mathbf{b}^{(k+1)}$ de taille $n - k - 1$.

— **Remontée :** On trouve x_n grâce à la formule

$$x_n = x_1^{(n-1)} = \frac{b_1^{(n-1)}}{a_{1,1}^{(n-1)}}.$$

Supposons que l'on connaisse $x_p, \ldots, x_n$ pour $p \in \{2, \ldots, n\}$, on détermine alors x_{p-1} grâce à la relation

$$x_{p-1} = x_1^{(p-2)} = b_1^{(p-2)} - \frac{1}{a_{1,1}^{(p-2)}} \sum_{j=2}^{n-p+2} a_{1,j}^{(p-2)} x_j^{(p-2)}.$$

En identifiant $x_k^{(\ell)} = x_{k+\ell}$, on obtient la formule

$$x_{p-1} = b_1^{(p-2)} - \frac{1}{a_{1,1}^{(p-2)}} \sum_{k=p}^{n} a_{1,k-p+2}^{(p-2)} x_k$$

où x_{p-1} ne dépend que des termes $x_p, \ldots, x_n$ déjà calculés.

On donne maintenant un lemme qui garantit que si la matrice de départ est inversible, alors on pourra effectuer l'algorithme du pivot de Gauss

Lemme 3.21 *Si A est une matrice inversible de $\mathcal{M}_n(\mathbb{K})$, alors toutes les matrices $A^{(k)} \in \mathcal{M}_{n-k}(\mathbb{K})$ construites dans l'algorithme du pivot de Gauss sont aussi inversibles.*

3.4 Inversion de matrices

Comme cela a été mis en avant dans le Théorème 3.17, toute matrice bijective $A \in \mathcal{M}_n(\mathbb{K})$ admet un inverse $A^{-1} \in \mathcal{M}_n(\mathbb{K})$. On se propose dans cette partie de donner des méthodes de calcul de A^{-1}.

Résolution de n-systèmes linéaires. Soit $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ la base canonique de $\mathbb{K}^n$, alors la $j^{\text{ème}}$ colonne de A^{-1}, notée $\mathbf{c}_j$, est donnée par

$$\mathbf{c}_j = A^{-1}\mathbf{e}_j.$$

En multipliant à gauche par A, on obtient que $\mathbf{c}_j$ est solution du système linéaire

$$A\mathbf{c}_j = \mathbf{e}_j, \qquad j \in \{1, \ldots, n\}. \tag{3.11}$$

Déterminer A^{-1} revient à calculer toutes ses colonnes, et donc à résoudre n systèmes linéaires (3.11).

Une méthode à l'aide d'opérations sur les lignes. La méthode proposée ci-dessous est une conséquence du fait que lorsque l'on applique la méthode du pivot de Gauss, on fait les mêmes opérations sur les lignes de la matrice indépendamment du second membre. On peut donc a priori traiter toutes les colonnes $\mathbf{c_j}$ simultanément.

On commence par mettre les matrices A et l'identité l'une à côté de l'autre

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

1. **Étape de descente :** On suppose que $a_{1,1}$ est non nul sinon, on échange deux lignes comme dans l'algorithme du pivot de Gauss. On fait alors $L_1 \leftarrow \frac{1}{a_{1,1}} L_1$ pour ramener $a_{1,1}$ à 1. On fait simultanément la même opération dans la colonne de droite :

$$\begin{pmatrix} 1 & \frac{a_{1,2}}{a_{1,1}} & \ldots & \ldots & \frac{a_{1,n}}{a_{1,1}} \\ a_{2,1} & a_{2,2} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n,1} & \ldots & \ldots & a_{n,n-1} & a_{n,n} \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{a_{1,1}} & 0 & \ldots & \ldots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \ldots & \ldots & 0 & 1 \end{pmatrix}$$

Pour $i \in \{2, \ldots, n\}$, on fait $L_i = L_i - a_{i,1}L_1$ dans chacune des colonnes, ce qui donne

$$\begin{pmatrix} 1 & \frac{a_{1,2}}{a_{1,1}} & \dots & \dots & \frac{a_{1,n}}{a_{1,1}} \\ 0 & (*) & \dots & \dots & (*) \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & (*) & \dots & \dots & (*) \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{a_{1,1}} & 0 & \dots & \dots & 0 \\ (*) & 1 & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 & 0 \\ (*) & 0 & \dots & 0 & 1 \end{pmatrix}$$

On refait alors le même travail à partir de la deuxième ligne, puis ainsi de suite, jusqu'à obtenir

$$\begin{pmatrix} 1 & \frac{a_{1,2}}{a_{1,1}} & \dots & \dots & \alpha_{1,n} \\ 0 & 1 & (*) & \dots & (*) \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & \alpha_{n-1,n} \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} \frac{1}{a_{1,1}} & 0 & \dots & \dots & 0 \\ (*) & (*) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & (*) & 0 \\ (*) & \dots & \dots & (*) & (*) \end{pmatrix}$$

2. **Étape de remontée :** Pour toutes les lignes i entre 1 et $(n-1)$, on fait $L_i \leftarrow L_i - \alpha_{i,n} L_n$ dans les deux colonnes. Cela donne :

$$\begin{pmatrix} 1 & (*) & \dots & (*) & 0 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & (*) & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} (*) & (*) & \dots & \dots & (*) \\ (*) & (*) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & (*) & (*) \\ (*) & \dots & \dots & (*) & (*) \end{pmatrix}$$

On réitère le processus pour les colonnes de $n-1$ à 2, jusqu'à obtenir l'identité à gauche. La matrice de la colonne de droite est alors A^{-1}.

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} \qquad A^{-1} = \begin{pmatrix} (*) & (*) & \dots & \dots & (*) \\ (*) & (*) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & (*) & (*) \\ (*) & \dots & \dots & (*) & (*) \end{pmatrix}.$$

3.5 Changement de base

Soit E, un $\mathbb{K}$-e.v. de dimension finie, $\mathcal{B} = (e_1, \ldots, e_n)$ et $\mathcal{B}' = (e'_1, \ldots, e'_n)$ deux bases de E. Pour tout j, $1 \leq j \leq n$, le vecteur e'_j s'écrit de façon unique en fonction des vecteurs $(e_1, \ldots, e_n)$:

$$e'_j = \sum_{i=1}^{n} p_{i,j} e_i. \tag{3.12}$$

La matrice $P = (p_{i,j})_{1 \leq i \leq n, 1 \leq j \leq n}$, dont les colonnes correspondent aux coordonnées des vecteurs de $\mathcal{B}'$ dans la base $\mathcal{B}$ est appellée *matrice de passage* de $\mathcal{B}$ à $\mathcal{B}'$.

Proposition 3.22 *Soient $\mathcal{B}, \mathcal{B}'$ deux bases de E, et soit P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, alors*

$$P = \mathrm{mat}_{\mathcal{B},\mathcal{B}'}(Id).$$

La matrice P est inversible, son inverse P^{-1} est la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$.

Soit $x \in E$, qui peut s'écrire de façon unique dans chacune des bases $\mathcal{B}$ et $\mathcal{B}'$:

$$x = \sum_{i=1}^{n} x_i e_i = \sum_{j=1}^{n} x'_j e'_j. \tag{3.13}$$

En utilisant (3.12), on peut alors écrire

$$x = \sum_{j=1}^{n} x'_j \left(\sum_{i=1}^{n} p_{i,j}\, e_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{i,j}\, x'_j \right) e_i,$$

d'où l'on obtient, par identification avec (3.13) :

$$x_i = \sum_{j=1}^{n} p_{i,j}\, x'_j, \quad \text{pour } 1 \leq j \leq n.$$

On en déduit alors la Proposition suivante.

Proposition 3.23 *Soit E, un $\mathbb{K}$-e.v. de dimension finie, $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Soit $x \in E$. Si X (respectivement X') désigne le vecteur colonne dont les éléments sont les coordonnées de x dans la base $\mathcal{B}$ (respectivement $\mathcal{B}'$), on a alors*

$$X = P\, X'.$$

Proposition 3.24 *Soit E et F deux $\mathbb{K}$-e.v. de dimension finie, $\mathcal{B}_E$ et $\mathcal{B}'_E$ deux bases de E, P la matrice de passage de $\mathcal{B}_E$ à $\mathcal{B}'_E$, $\mathcal{B}_F$ et $\mathcal{B}'_F$ deux bases de F, Q la matrice de passage de $\mathcal{B}_F$ à $\mathcal{B}'_F$. Soit $f \in \mathcal{L}(E, F)$. On a alors*

$$\mathrm{mat}_{\mathcal{B}'_F,\mathcal{B}'_E}(f) = Q^{-1}\, \mathrm{mat}_{\mathcal{B}_F,\mathcal{B}_E}(f)\, P.$$

Le résultat de la Proposition 3.24 est illustré par le diagramme suivant.

$$\begin{tikzcd}
(E,\mathcal{B}_E) \arrow[r, "f", "\mathrm{mat}_{\mathcal{B}_F,\mathcal{B}_E}(f)"'] & (F,\mathcal{B}_F) \arrow[d, "Id"', "Q^{-1}"] \\
(E,\mathcal{B}'_E) \arrow[u, "Id", "P"'] \arrow[r, "f", "\mathrm{mat}_{\mathcal{B}'_F,\mathcal{B}'_E}(f)"'] & (F,\mathcal{B}'_F)
\end{tikzcd}$$

On s'intéresse maintenant au cas particulier des endomorphismes de E, où $\dim(E) = n$. Notons $\mathcal{B}$ une base de E. Comme vu précédemment, cet espace s'identifie à $\mathcal{M}_n(\mathbb{K})$ via

$$A = (a_{i,j}) = \mathrm{mat}_{\mathcal{B},\mathcal{B}}(f) = \mathrm{mat}_{\mathcal{B}}(f) \quad \text{si} \quad f(\mathbf{e}_j) = \sum_{i=1}^{n} a_{i,j}\ \mathbf{e}_i.$$

Soit $\mathcal{B}'$ une autre base de E, et on note P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, et $B = \mathrm{mat}_{\mathcal{B}',\mathcal{B}'}(f)$ alors d'après la proposition 3.24, on a

$$B = P^{-1}AP. \tag{3.14}$$

Réciproquement, si A et B sont deux matrices de $\mathcal{M}_n(\mathbb{K})$ qui vérifient la relation (3.14) pour une matrice P inversible, alors il existe deux bases $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ et $\mathcal{B}' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ de $\mathbb{K}^n$ et un endomorphisme $f \in \mathcal{L}(\mathbb{K}^n)$ tels que

$$A = \mathrm{mat}_{\mathcal{B}}(f) \quad \text{et} \quad B = \mathrm{mat}_{\mathcal{B}'}(f). \tag{3.15}$$

Définition 3.25 (matrices semblables) *Deux matrices A et B de $\mathcal{M}_n(\mathbb{K})$ sont dites* semblables *s'il existe $P \in GL_n(\mathbb{K})$ telle que*

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

D'après la relation (3.15), deux matrices sont semblables si et seulement si elle peuvent représenter le même endomorphisme pour deux choix de base différents.

Remarque: Il est à noter que dans (3.15), A (resp. B) est une matrice représentant f pour la même base de départ et d'arrivée $\mathcal{B}$ (resp. $\mathcal{B}'$).

Définition 3.26 (trace d'une matrice) *Soit $A = (a_{i,j})$ une matrice carrée de $\mathcal{M}_n(\mathbb{K})$, alors on définit la* trace *de A, notée* $\mathrm{tr}(A)$ *par*

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

Proposition 3.27 *Soient A, B deux matrices de $\mathcal{M}_n(\mathbb{K})$, alors $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. En particulier, deux matrices semblables ont la même trace.*

Chapitre 4

Matrices et orthogonalité

4.1 Produit scalaire et orthogonalité

Définition 4.1 (produit scalaire canonique de $\mathbb{K}^n$) *Soit* $\mathbf{x} = (x_1, \ldots, x_n)^T$*, et* $\mathbf{y} = (y_1, \ldots, y_n)^T$ *deux vecteurs de* $\mathbb{K}^n$*, alors on définit le* produit scalaire *de* $\mathbf{x}$ *avec* $\mathbf{y}$ *par*

$$< \mathbf{x} \,|\, \mathbf{y} >_{\mathbb{K}^n} = \sum_{i=1}^{n} x_i \bar{y}_i, \tag{4.1}$$

où $\bar{y}$ *désigne le complexe conjugué du complexe* y*. La norme associée au produit scalaire, notée* $\|\cdot\|_2$*, est définie par*

$$\|\mathbf{x}\|_2 = \sqrt{< \mathbf{x} \,|\, \mathbf{x} >_{\mathbb{K}^n}} = \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

Pour alléger les notations dans la suite de ce chapitre, nous noterons $\|\cdot\| = \|\cdot\|_2$, et lorsqu'il n'y a pas d'ambiguité, $< \cdot \,|\, \cdot > = < \cdot \,|\, \cdot >_{\mathbb{K}^n}$.

Proposition 4.2 *Soit* $\mathbf{x} \in \mathbb{K}^n$*, on a* $\|x\| = 0$ *si et seulement si* $x = 0_{\mathbb{K}^n}$*. En particulier tout vecteur non nul a une norme strictement positive.*

Proposition 4.3

1. *Soient* $\mathbf{x}, \mathbf{y}$ *dans* $\mathbb{K}^n$*, alors*
$$< \mathbf{x} \,|\, \mathbf{y} > = \overline{< \mathbf{y} \,|\, \mathbf{x} >}.$$
2. *Soient* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{K}^n$*, on a*
$$< \mathbf{x} \,|\, \mathbf{y} + \mathbf{z} > = < \mathbf{x} \,|\, \mathbf{y} > + < \mathbf{x} \,|\, \mathbf{z} >, \qquad < \mathbf{x} + \mathbf{y} \,|\, \mathbf{z} > = < \mathbf{x} \,|\, \mathbf{z} > + < \mathbf{y} \,|\, \mathbf{z} > .$$
3. *Soient* $\mathbf{x}, \mathbf{y}$ *dans* $\mathbb{K}^n$*, soit* $\lambda \in \mathbb{K}$*, on a*
$$< \lambda\mathbf{x} \,|\, \mathbf{y} > = \lambda < \mathbf{x} \,|\, \mathbf{y} >, \qquad < \mathbf{x} \,|\, \lambda\mathbf{y} > = \overline{\lambda} < \mathbf{x} \,|\, \mathbf{y} > .$$

Définition 4.4 (orthogonalité)

1. *Deux vecteurs* $\mathbf{x}, \mathbf{y}$ *de* $\mathbb{K}^n$ *sont dits orthogonaux si* $< \mathbf{x} \,|\, \mathbf{y} >_{\mathbb{K}^n} = 0$.
2. *Deux s.e.v.* E *et* F *de* $\mathbb{K}^n$ *sont dits orthogonaux si*

$$\forall \mathbf{x} \in E, \forall \mathbf{y} \in F, \qquad < \mathbf{x} \,|\, \mathbf{y} >_{\mathbb{K}^n} = 0.$$

Remarque:
— Le vecteur $0_{\mathbb{K}^n}$ est orthogonal à tous les vecteurs de $\mathbb{K}^n$.
— Dans $\mathbb{R}^2$, la notion d'orthogonalité décrite dans la Définition 4.4 coïncide avec la notion d'orthogonalité géométrique usuelle.

Proposition 4.5 *On a* $\mathbf{x} = 0_{\mathbb{K}^n}$ *si et seulement si, pour tout* $\mathbf{y} \in \mathbb{K}^n$, *on a* $< \mathbf{x} \,|\, \mathbf{y} >= 0$.

Preuve :
" $\Rightarrow$ " C'est une conséquence de la formule (4.1).
" $\Leftarrow$ " On choisit $\mathbf{y} = \mathbf{x}$, ce qui donne $\|\mathbf{x}\| = 0$, et donc $\mathbf{x} = 0$ d'après la Proposition 4.2.

□

Définition 4.6 (famille orthonormée) *Soit* $\mathcal{F} = (\mathbf{e}_1, \ldots, \mathbf{e}_m)$ *une famille de* $\mathbb{K}^n$, *alors on dit que* $\mathcal{F}$ *est* orthonormée *si*

1. *pour tout* $i, j \in \{1, \ldots, m\}$ *avec* $i \neq j$, *on a* $< \mathbf{e}_i \,|\, \mathbf{e}_j >= 0$. *On dit alors que la famille est* orthogonale.
2. *pour tout* $i \in \{1, \ldots, m\}$, *on a* $< \mathbf{e}_i \,|\, \mathbf{e}_i >= \|\mathbf{e}_i\|^2 = 1$ *On dit alors que la famille est* normée.

Proposition 4.7 *Une famille orthonormée de* $\mathbb{K}^n$ *est nécessairement libre. En particulier,* $m \leq n$.

Preuve : Soit $\mathcal{F} = (\mathbf{e}_1, \ldots, \mathbf{e}_m)$ une famille orthonormée de $\mathbb{K}^n$. On suppose par l'absurde que $\mathcal{F}$ est liée, alors il existe $i \in \{1, \ldots, m\}$ tel que

$$\mathbf{e}_i = \sum_{j \neq i} \lambda_j \mathbf{e}_j. \tag{4.2}$$

On prend le produit scalaire de (4.2) avec $\mathbf{e}_i$, ce qui donne, d'après la Proposition 4.3 et la Définition 4.6,

$$1 = \|\mathbf{e}_i\|^2 = \left\langle \mathbf{e}_i \,\middle|\, \sum_{j \neq i} \lambda_j \mathbf{e}_j \right\rangle = \sum_{j \neq i} \lambda_j < \mathbf{e}_i \,|\, \mathbf{e}_j >= 0.$$

On obtient donc une contradiction, et la famille $\mathcal{F}$ est nécessairement libre. □

Remarque: D'après la proposition précédente et le Corollaire 1.15, on voit que toute famille orthonormée de n éléments est une base de $\mathbb{K}^n$. Une telle base est appelée **base orthonormée** de $\mathbb{K}^n$.

Exemple: La base canonique est une base orthonormée de $\mathbb{K}^n$.

Définition 4.8 (supplémentaire orthogonal) *Soit E un s.e.v. de $\mathbb{K}^n$, alors on appelle* supplémentaire orthogonal *(ou orthogonal) de E l'ensemble noté $E^\perp$ défini par*

$$E^\perp = \{\mathbf{x} \in \mathbb{K}^n \,|\, \forall \mathbf{y} \in E, \; < \mathbf{x} \,|\, \mathbf{y} >= 0\}.$$

La proposition ci-dessous justifie la dénomination *supplémentaires orthogonaux*.

Proposition 4.9 *Soit E un s.e.v. de dimension m de $\mathbb{K}^n$, alors $E^\perp$ est un s.e.v. de $\mathbb{K}^n$ tel que*

$$E \oplus E^\perp = \mathbb{K}^n.$$

En particulier, $\dim E^\perp = n - m$.

Remarque: Les s.e.v. E et $E^\perp$ sont orthogonaux. $E^\perp$ est même le plus grand s.e.v. de $\mathbb{K}^n$ orthogonal à E.

4.2 Matrice transposée, matrice adjointe

Définition 4.10 (matrice transposée, matrice adjointe) *Soit $A = (a_{i,j}) \in \mathcal{M}_{m,n}(\mathbb{K})$, alors on appelle* matrice transposée *de A la matrice de $\mathcal{M}_{n,m}(\mathbb{K})$ notée A^T définie par*

$$A^T = (a_{j,i}).$$

On appelle matrice adjointe *de A la matrice de $\mathcal{M}_{n,m}(\mathbb{K})$ notée A^* définie par*

$$A^* = (\bar{a}_{j,i}).$$

Il est clair que dans le cas où $A \in \mathcal{M}_{m,n}(\mathbb{R})$, on a $A^T = A^*$.

Exemple:

$$A = \begin{pmatrix} 1+i & -1 & 2-i \\ -1 & i & 3 \end{pmatrix}, \qquad A^T = \begin{pmatrix} 1+i & -1 \\ -1 & i \\ 2-i & 3 \end{pmatrix}, \qquad A^* = \begin{pmatrix} 1-i & -1 \\ -1 & -i \\ 2+i & 3 \end{pmatrix},$$

Proposition 4.11 *On a $\left(A^T\right)^T = A$ et $(A^*)^* = A$.*

Proposition 4.12 *Soient $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $B \in \mathcal{M}_{n,p}(\mathbb{K})$, on a*

$$(AB)^T = B^T A^T, \qquad (AB)^* = B^* A^*.$$

Définition 4.13 (matrice symétrique, matrice hermitienne) *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite symétrique si $A = A^T$. Elle est dite hermitienne au auto-adjointe si $A = A^*$.*

Le théorème suivant donne le lien entre matrice adjointe et produit scalaire.

Théorème 4.14 *Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$, alors A^* est l'unique matrice telle que*

$$< A\mathbf{x} \,|\, \mathbf{y} >_{\mathbb{K}^m} = < \mathbf{x} \,|\, A^*\mathbf{y} >_{\mathbb{K}^n}, \qquad \forall \mathbf{x} \in \mathbb{K}^n, \forall \mathbf{y} \in \mathbb{K}^m.$$

Remarque: On peut identifier le vecteur colonne $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$ avec une matrice à n lignes et 1 colonnes $\mathbf{x}, \mathbf{y} \in \mathcal{M}_{n,1}(\mathbb{K})$. Avec cette convention, on a

$$< \mathbf{x} \,|\, \mathbf{y} >= \mathbf{y}^* \mathbf{x}.$$

Proposition 4.15 *Soit* $A \in \mathcal{M}_{m,n}(\mathbb{K})$, *alors*

1. $\mathrm{Ker}(A^*) = \mathrm{Im}(A)^{\perp}$.
2. $\mathrm{Im}(A^*) = \mathrm{Ker}(A)^{\perp}$.

En particulier, $\mathrm{rg}(A) = \mathrm{rg}(A^*)$ *et* A^* *est inversible si et seulement si* A *est inversible.*

4.3 Matrices orthogonales, unitaires

Définition 4.16 (matrices orthogonales, unitaires)

1. *Une matrice* $O \in \mathcal{M}_n(\mathbb{R})$ *est dite orthogonale si* $O^T = O^{-1}$. *On note* $\mathcal{O}_n(\mathbb{R})$ *l'ensemble des matrices orthogonales de* $\mathcal{M}_n(\mathbb{R})$.
2. *Une matrice* $U \in \mathcal{M}_n(\mathbb{C})$ *est dite unitaire si* $U^* = U^{-1}$. *On note* $\mathcal{U}_n(\mathbb{R})$ *l'ensemble des matrices unitaires de* $\mathcal{M}_n(\mathbb{C})$.

Proposition 4.17 *Si* O *est unitaire (resp. orthogonale dans le cas réel), alors* O^{-1} *l'est aussi.*

Proposition 4.18 *Soient* O, P *deux matrices orthogonales (resp. unitaires) de* $\mathcal{M}_n(\mathbb{R})$ *(resp.* $\mathcal{M}_n(\mathbb{C})$*), alors* OP *est orthogonale (resp. unitaire).*

Preuve : Soient O, P deux matrices orthogonales, alors $(OP)^{-1} = P^{-1}O^{-1} = P^T O^T = (OP)^T$. $\square$

Proposition 4.19 *Soit* $U \in \mathcal{U}_n(\mathbb{K})$, *alors*

$$< U\mathbf{x} \,|\, U\mathbf{y} >=< \mathbf{x} \,|\, \mathbf{y} >, \qquad \forall \mathbf{x}, \mathbf{y} \in \mathbb{K}^n.$$

En particulier,

$$\|U\mathbf{x}\| = \|\mathbf{x}\|, \qquad \forall \mathbf{x} \in \mathbb{K}^n. \tag{4.3}$$

Remarque: La relation (4.3) signifie qu'une matrice unitaire (ou orthogonale dans le cas réel) ne modifie pas la norme des vecteurs. Ce sont donc des **isométries**.

Proposition 4.20 *Si* $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ *une base orthonormée de* $\mathbb{K}^n$ *et soit* U *une matrice unitaire d'ordre* n *(*$U \in \mathcal{U}_n(\mathbb{K})$*), alors* $U\mathcal{B} = (U\mathbf{e}_1, \ldots, U\mathbf{e}_n)$ *est aussi une base orthonormée de* $\mathbb{K}^n$.

Le corollaire suivant est une conséquence du fait que les colonnes d'une matrice sont les images par cette matrice des vecteurs de la base canonique, elle même orthonormée.

Corollaire 4.21 *Les colonnes* $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$ *d'une matrice orthogonale (resp. unitaire) forment une base orthonormée de* $\mathbb{R}^n$ *resp.* $\mathbb{C}^n$.

Corollaire 4.22 *Les matrices de passage entre bases orthonormées sont les matrices unitaires (ou orthogonales dans le cas réel).*

4.4 Représentation des formes linéaires

Définition 4.23 (formes linéraires sur $\mathbb{K}^n$) *On appelle ensemble des formes linéaires sur $\mathbb{K}^n$ l'ensemble $\mathcal{L}(\mathbb{K}^n, \mathbb{K})$ des applications linéaires définies sur $\mathbb{K}^n$ à valeurs dans $\mathbb{K}$.*

L'ensemble $\mathcal{L}(\mathbb{K}^n, \mathbb{K})$ s'identifie via le choix d'une base $\mathcal{B}$ de $\mathbb{K}^n$ à l'ensemble des matrices $\mathcal{M}_{1,n}(\mathbb{K})$ à une ligne et n colonnes. C'est un espace vectoriel de dimension n sur $\mathbb{K}$, que l'on peut donc identifier à $\mathbb{K}^n$. On va voir qu'un moyen de faire cette identification consiste à utiliser le produit scalaire.

Théorème 4.24 (Théorème de Riesz) *Soit $\phi \in \mathcal{L}(\mathbb{K}^n, \mathbb{K})$, alors il existe un unique $\mathbf{x}_\phi \in \mathbb{K}^n$ tel que*

$$\phi(\mathbf{y}) =< \mathbf{y} \,|\, \mathbf{x}_\phi >, \qquad \forall \mathbf{y} \in \mathbb{K}^n.$$

On déduit du théorème précédent par exemple que $\mathrm{Ker}(\phi) = \mathrm{Vect}(\mathbf{x}_\phi)^\perp$.

Chapitre 5

Déterminants

5.1 Interprétation volumique

On considère une matrice $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, alors il est facile de voir en résolvant les systèmes linéaires $A\mathbf{x}_1 = \mathbf{e}_1$ et $A\mathbf{x}_2 = \mathbf{e}_2$ (ici $(\mathbf{e}_1, \mathbf{e}_2)$ désigne la base canonique de $\mathbb{K}^2$) que la matrice est inversible si et seulement si

$$\mathcal{A}(A) := a_{11}\ a_{22}\ - a_{12}a_{21} \neq 0. \tag{5.1}$$

La quantité $\mathcal{A}(A)$ est exactement égale à l'aire algébrique du parallélogramme de côtés $\mathbf{c}_1 = (a_{11}, a_{21})^T$ et $\mathbf{c}_2 = (a_{12}, a_{22})^T$ les vecteurs colonnes de A. Cette aire est positive si $\mathbf{c}_1, \mathbf{c}_2$ forment un angle de moins de π dans le sens trigo, et négative si cet angle est compris entre π et 2π. On a la formule

$$\mathcal{A}(A) = \|\mathbf{c}_1\|\|\mathbf{c}_2\| \sin(\mathbf{c}_1, \mathbf{c}_2). \tag{5.2}$$

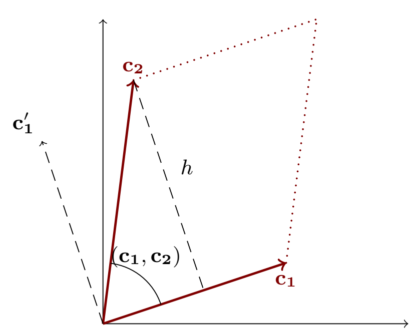

FIGURE 5.1 – Aire du parallélogramme engendré par ses vecteurs $\mathbf{c_1}$ et $\mathbf{c_2}$.

On a $\mathcal{A}(A) = 0$ si et seulement si les vecteurs $\mathbf{c}_1, \mathbf{c}_2$ sont colinéaires, donc si et seulement si $\mathrm{rg}(A) \leq 1$, ce qui équivaut au fait A n'est pas surjective et donc pas inversible d'après le théorème 3.13.

Exemple: Sur la Figure 5.1, on a

$$|\det(\mathbf{c_1}, \mathbf{c_2})| = |\det(6\mathbf{i} + 2\mathbf{j}, \mathbf{i} + 8\mathbf{j})| = \begin{vmatrix} 6 & 1 \\ 2 & 8 \end{vmatrix} = 46.$$

On peut appliquer le même raisonnement dans le cas où $A = (a_{i,j}) = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3) \in \mathcal{M}_3(\mathbb{K})$. Dans ce cas, le volume algébrique $\mathcal{V}(A)$ du parallélépipède engendré par les colonnes $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$ de A est donné par la formule

$$\begin{aligned} \mathcal{V}(A) \quad = \quad & a_{11}\ a_{22}\ a_{33}\ + a_{12}a_{23}\ a_{31} + a_{13}a_{21}a_{32} \\ & -a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}\ a_{21}a_{12}. \end{aligned} \tag{5.3}$$

La quantité $\mathcal{V}(A)$ est positive si le trièdre formé par $\mathbf{c}_1, \mathbf{c}_2$ et $\mathbf{c}_3$ est direct, et négative dans le cas où ce trièdre est indirect. On a alors la proposition suivante.

Proposition 5.1 *Soit $A \in \mathcal{M}_3(\mathbb{K})$, alors A est inversible si et seulement si la quantité $\mathcal{V}(A)$ définie par* (5.3) *est non nulle.*

Remarque: Si on échange deux colonnes, alors on change le signe de $\mathcal{A}(A)$ (resp. $\mathcal{V}(A)$) si $A \in \mathcal{M}_2(\mathbb{K})$ (resp. $A \in \mathcal{M}_3(\mathbb{K})$).

5.2 Définition générale

On cherche maintenant à étendre la notion du paragraphe précédent au cas de la dimension n quelconque. Dans la suite, $A = (\mathbf{c}_1, \cdots, \mathbf{c}_n)$, où $\mathbf{c}_i \in \mathbb{K}^n$ (i = 1,..., n), et $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ la base canonique de $\mathbb{K}^n$.

Théorème 5.2 *Il existe une unique application $\mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$, appelée* déterminant *et notée* det, *telle que les propriétés suivantes soient vérifiées.*

1. *En échangeant deux colonnes, on change le signe :*

$$\det(\mathbf{c}_1, \ldots, \mathbf{c}_i, \ldots, \mathbf{c}_j, \ldots, \mathbf{c}_n) = -\det(\mathbf{c}_1, \ldots, \mathbf{c}_j, \ldots, \mathbf{c}_i, \ldots, \mathbf{c}_n).$$

2. *Le déterminant est linéaire par rapport à chaque colonne :*

$$\det(\mathbf{c}_1, \ldots, \lambda\mathbf{c}_i + \mu\mathbf{c}'_i, \ldots, \mathbf{c}_n) = \lambda\det(\mathbf{c}_1, \ldots, \mathbf{c}_i, \ldots \mathbf{c}_n) + \mu\det(\mathbf{c}_1, \ldots, \mathbf{c}'_i, \ldots, \mathbf{c}_n).$$

3. *Le déterminant vaut* 1 *pour la matrice identité*

$$\det(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \det(I_n) = 1.$$

On note aussi

$$\det(A) = \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n,1} & \cdots & \cdots & a_{n,n-1} & a_{n,n} \end{vmatrix}$$

On déduit des propriétés fondamentales du déterminant données au Théorème 5.2 les propriétés suivantes :

Proposition 5.3 *1. Si la matrice A a deux colonnes égales, alors* $\det(A) = 0$ *:*

$$\det(\mathbf{c}_1, \ldots, \mathbf{c}_i, \ldots, \mathbf{c}_i, \ldots, \mathbf{c}_n) = 0.$$

2. Si la matrice A a une colonne nulle, alors $\det(A) = 0$ *:*

$$\det(\mathbf{c}_1, \ldots, 0_{\mathbb{K}^n}, \ldots, \mathbf{c}_n) = 0.$$

3. Pour tout $\lambda \in \mathbb{K}$, $\det(\lambda A) = \lambda^n \det(A)$ *:*

$$\det(\lambda \mathbf{c}_1, \ldots, \lambda \mathbf{c}_n) = \lambda^n \det(\mathbf{c}_1, \ldots, \mathbf{c}_n).$$

4. Dans le cas $n = 2$, $\det(A) = \mathcal{A}(A)$*, et dans le cas $n = 3$,* $\det(A) = \mathcal{V}(A)$*.*

La principale motivation pour l'introduction du déterminant est la suivante.

Théorème 5.4 *Soit $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$ une famille de n vecteurs de $\mathbb{K}^n$, alors*

$$\det(\mathbf{c}_1, \ldots, \mathbf{c}_n) = 0 \quad \Leftrightarrow \quad \textit{la famille } (\mathbf{c}_1, \ldots, \mathbf{c}_n) \textit{ est liée.}$$

En particulier, si A est une matrice de $\mathbb{K}^n$, alors

$$\det(A) \neq 0 \quad \Leftrightarrow \quad A \textit{ est inversible.}$$

Donnons encore quelques propriétés importantes du déterminant.

Proposition 5.5 *Soient $A, B \in \mathcal{M}_n(\mathbb{K})$, alors on a*

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

Corollaire 5.6 *Soit A une matrice inversible. Alors*

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

Corollaire 5.7 *Soient A, B deux matrices semblables (i.e. $B = P^{-1}AP$, cf. Définition 3.25) alors* $\det(A) = \det(B)$*.*

Proposition 5.8 *Soit $A \in \mathcal{M}_n(\mathbb{R})$, alors* $\det(A) = \det(A^T)$*. En particulier, toutes les règles de calcul sur les colonnes de la matrice sont valables pour des calculs sur les lignes de la matrice.*

5.3 Développements par rapport à une ligne ou colonne

Nous n'avons pour l'instant pas donné de méthode pour le calcul explicite de déterminants d'ordre n quelconque, mais seulement dans les cas $n = 2$ ou $n = 3$ (le cas $n = 1$ est trivial). Une des méthodes pour calculer un déterminant est de développer par rapport à une ligne ou une colonne pour pouvoir passer d'un déterminant de taille n à n déterminants de taille $n - 1$.

Théorème 5.9 (développement par rapport à une colonne) *Soit $A = (a_{i,j})$ une matrice de $\mathcal{M}_n(\mathbb{K})$ avec $n \geq 2$, et pour tout $(i, j) \in \{1, \dots, n\}$on note $M_{i,j}$ la matrice de $\mathcal{M}_{n-1}(\mathbb{K})$ obtenue à partir de A en enlevant la ligne i et la colonne j. Alors*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(M_{i,j}), \qquad \forall j \in \{1, \dots, n\}.$$

Or comme d'après la Proposition 5.8, $\det(A) = \det(A^T)$, on déduit directement du Théorème 5.9 que l'on peut aussi développer par rapport à une ligne.

Corollaire 5.10 (développement par rapport à une ligne) *Avec les notations du Théorème 5.9, on a*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det(M_{i,j}), \qquad \forall i \in \{1, \dots, n\}.$$

Exemple: On considère

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

En développant par rapport à la première colonne, on obtient

$$\det(A) = 4(-1)^{1+1} \begin{vmatrix} -1 & -1 & 2 \\ 1 & 0 & 1 \\ 3 & 0 & 1 \end{vmatrix} + 2(-1)^{1+4} \begin{vmatrix} -2 & 3 & 4 \\ -1 & -1 & 2 \\ 1 & 0 & 1 \end{vmatrix}$$

En développant le premier déterminant par rapport à la deuxième colonne, et le deuxième déterminant par rapport à la dernière ligne, on obtient que

$$\det(A) = 4 \begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} - 2 \left(\begin{vmatrix} 3 & 4 \\ -1 & 2 \end{vmatrix} + \begin{vmatrix} -2 & 3 \\ -1 & -1 \end{vmatrix} \right) = 4 \times (-2) - 2(10 + 5) = -38.$$

Définition 5.11 (matrices triangulaires) *Une matrice $A = (a_{i,j})$ est dite triangulaire supérieure (resp. inférieure) si $a_{i,j} = 0$ pour $i > j$ (resp $i < j$).*

On déduit du Théorème 5.9 et du corollaire 5.10 la proposition importante suivante, qui dit que le déterminant d'une matrice triangulaire est donné par le produit des termes diagonaux.

Proposition 5.12 *Soit $A = (a_{i,j})$ une matrice triangulaire supérieure ou inférieure, alors*

$$\det(A) = \prod_{i=1}^{n} a_{i,i}.$$

5.4 Calcul à l'aide d'opérations élémentaires sur les lignes et colonnes

Une des méthodes pratiques pour calculer un déterminant consiste à effectuer des opérations élémentaires sur les lignes et les colonnes pour se ramener à une matrice triangulaire. Le calcul du déterminant est alors trivial grâce à la Proposition 5.12. L'algorithme est très proche de celui du pivot de Gauss détaillé dans la partie 3.3. Illustrons-la sur un exemple. On cherche à calculer

$$\det(A) = \begin{vmatrix} 4 & -2 & 3 & 4 \\ 1 & -1 & -1 & 2 \\ 3 & 1 & 0 & 1 \\ 2 & 3 & 0 & 1 \end{vmatrix}.$$

On échange les deux premières lignes, ce qui change le signe du résultat :

$$\det(A) = -\begin{vmatrix} 1 & -1 & -1 & 2 \\ 4 & -2 & 3 & 4 \\ 3 & 1 & 0 & 1 \\ 2 & 3 & 0 & 1 \end{vmatrix} = -\det(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4).$$

On remplace la deuxième colonne par deuxième plus première, ce qui ne change pas le déterminant :

$$\det(\mathbf{c}_1, \mathbf{c}_2 + \mathbf{c}_1, \mathbf{c}_3, \mathbf{c}_4) = \underbrace{\det(\mathbf{c}_1, \mathbf{c}_1, \mathbf{c}_3, \mathbf{c}_4)}_{=0} + \det(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4) = \det(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4).$$

De même, on remplace $\mathbf{c}_3$ par $\mathbf{c}_3 + \mathbf{c}_1$ et $\mathbf{c}_4$ par $\mathbf{c}_4 - 2\mathbf{c}_1$, ce qui donne à l'aide d'un développement par rapport à la première ligne

$$\det(A) = -\begin{vmatrix} 1 & 0 & 0 & 0 \\ 4 & 2 & 7 & -4 \\ 3 & 4 & 3 & -5 \\ 2 & 5 & 2 & -3 \end{vmatrix} = -\begin{vmatrix} 2 & 7 & -4 \\ 4 & 3 & -5 \\ 5 & 2 & -3 \end{vmatrix} = -\det(\mathbf{c}'_1, \mathbf{c}'_2, \mathbf{c}'_3).$$

En remplaçant $\mathbf{c}'_3$ par $\mathbf{c}'_3 + 2\mathbf{c}'_1$ et $\mathbf{c}'_2$ par $\mathbf{c}'_2 - \frac{7}{2}\mathbf{c}'_1$ (ce qui ne change pas le déterminant), on obtient

$$\det(A) = -\begin{vmatrix} 2 & 0 & 0 \\ 4 & -11 & 3 \\ 5 & -\frac{31}{2} & 7 \end{vmatrix} = -2\begin{vmatrix} -11 & 3 \\ -\frac{31}{2} & 7 \end{vmatrix} = \begin{vmatrix} 22 & 3 \\ 31 & 7 \end{vmatrix} = 22 \times 7 - 31 \times 3 = 61.$$

Chapitre 6

Réduction des matrices carrées

6.1 Notion de valeur propre, vecteur propre

Définition 6.1 *Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$. On dit que λ est valeur propre de A s'il existe un vecteur $\mathbf{x} \in \mathbb{K}^n$, $\mathbf{x} \neq 0_{\mathbb{K}^n}$, tel que $A\mathbf{x} = \lambda\mathbf{x}$. Le vecteur $\mathbf{x}$ est appelé vecteur propre de A associé à la valeur propre λ.*

Définition 6.2 *On appelle spectre d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ l'ensemble de ses valeurs propres de A dans le corps $\mathbb{K}$ considéré. On le note $\mathrm{Sp}(A)$.*

Comme

$$A\mathbf{x} = \lambda\mathbf{x} \quad \Leftrightarrow (A - \lambda I_n)\mathbf{x} = 0_{\mathbb{K}^n} \Leftrightarrow \mathbf{x} \in \mathrm{Ker}(A - \lambda I_n)$$

on en déduit que $\mathbf{x}$ est un vecteur propre de A associé à la valeur propre λ si $\mathbf{x}$ appartient au l'ensemble

$$V_\lambda := \mathrm{Ker}(A - \lambda I_n)$$

et que λ est une valeur propre de A si il existe $\mathbf{x} \neq 0_{\mathbb{K}^n}$ appartenant à V_λ, c'est-à dire si $V_\lambda \neq \{0_{\mathbb{K}_n}\}$.

Proposition 6.3 *Soit λ une valeur propre de A. Alors V_λ est un sous-espace vectoriel de $\mathbb{K}^n$ et est stable par A, c'est-à-dire que*

$$\forall \mathbf{x} \in V_\lambda, \quad A\mathbf{x} \in V_\lambda.$$

Remarque:

— La proposition précédente signifie que si $\mathbf{x}_1$ et $\mathbf{x}_2$ sont des vecteurs propres associés à la valeur propre λ, alors $\mathbf{x}_1 + \mathbf{x}_2$, $\alpha\mathbf{x}_1$ (pour tout scalaire $\alpha \neq 0$) et $A\mathbf{x}_1$ sont encore des vecteurs propres associés à la valeur propre λ. Si λ est une valeur propre de A, il n'y a donc pas unicité du vecteur propre associé.

— Un vecteur $\mathbf{x}$ est un vecteur propre de la matrice A s'il existe $\lambda \in \mathbb{K}$ tel que $A\mathbf{x} = \lambda\mathbf{x}$. Si ce scalaire λ existe, il est unique.

Définition 6.4 *Le sous-espace vectoriel V_λ est appelé sous-espace propre de A associé à la valeur propre λ.*

Remarque: En pratique, pour déterminer une base V_λ lorsque λ est une valeur propre d'une matrice $A = (a_{ij})_{1\le i\le n, 1\le j\le n}$, on résout le système linéaire $AX = \lambda X$.

Proposition 6.5 *Si λ_1 et λ_2 sont des valeurs propres distinctes d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$, alors l'intersection des sous-espaces propres correspondant est réduite à $0_{\mathbb{K}^n}$:*

$$V_{\lambda_1} \cap V_{\lambda_2} = 0_{\mathbb{K}^n}$$

Preuve : Soit $\mathbf{x} \in V_{\lambda_1} \cap V_{\lambda_2}$. Alors $\mathbf{x} = 0_{\mathbb{K}^n}$ ou $\mathbf{x}$ est un vecteur propre de A pour les valeurs propres λ_1 et λ_2. On a donc $A\mathbf{x} = \lambda_1\mathbf{x}$ et $A\mathbf{x} = \lambda_2\mathbf{x}$, d'où l'on déduit que $(\lambda_1 - \lambda_2)\mathbf{x} = 0_{\mathbb{K}^n}$, puis, comme $\lambda_1 \neq \lambda_2$, que $\mathbf{x} = 0_{\mathbb{K}^n}$. □

Proposition 6.6 *Soit p un entier, $p \ge 2$, λ_1, ..., λ_p des valeurs propres distinctes d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ et $(\mathbf{x}_1, \ldots, \mathbf{x}_p)$ une famille de vecteurs telles que $\mathbf{x}_i$ soit un vecteur propre de A associé à la valeur propre λ_i. Alors cette famille est libre.*

Preuve : On montre ceci par récurrence sur l'entier p.

- Pour $p = 2$. Si $\mu_1\mathbf{x}_1 + \mu_2\mathbf{x}_2 = 0_{\mathbb{K}^n}$ alors
  - si $\mu_1 \neq 0$ alors $\mathbf{x}_1 = -\frac{\lambda_2}{\lambda_1}\mathbf{x}_2$ et donc $\mathbf{x}_1 \in V_{\lambda_1} \cap V_{\lambda_2}$. Or d'après la proposition 6.5, $V_{\lambda_1} \cap V_{\lambda_2} = \{0_{\mathbb{K}^n}\}$ et donc $\mathbf{x}_1 = 0_{\mathbb{K}^n}$. C'est impossible car $\mathbf{x}_1$ est un vecteur propre (et un vecteur propre est non nul par définition).
  - on a donc $\mu_1 = 0$; il en résulte que $\mu_2 = 0$ aussi car $\mathbf{x}_2 \neq 0_{\mathbb{K}^n}$.

  Ainsi $\mu_1\mathbf{x}_1 + \mu_2\mathbf{x}_2 = 0_{\mathbb{K}^n} \Rightarrow \mu_1 = \mu_2 = 0$: la famille $(\mathbf{x}_1, \mathbf{x}_2)$ est donc libre.
- On suppose la propriété [la famille $(\mathbf{x}_1, \ldots, \mathbf{x}_p)$ est libre] vraie pour l'entier p.
  Soit $\lambda_1, \ldots, \lambda_p, \lambda_{p+1}$ des valeurs propres distinctes d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ et $(\mathbf{x}_1, \ldots, \mathbf{x}_p, \mathbf{x}_{p+1})$ une famille de vecteurs propres associés. Si

  $$\mu_1\mathbf{x}_1 + \ldots \mu_p\mathbf{x}_P + \mu_{p+1}\mathbf{x}_{p+1} = 0_{\mathbb{K}^n} \tag{6.1}$$

  alors
  - si $\mu_{p+1} \neq 0$, alors

    $$\mathbf{x}_{p+1} = -\sum_{k=1}^{p} \frac{\mu_k}{\mu_{p+1}}\mathbf{x}_k \tag{6.2}$$

    d'où

    $$A\mathbf{x}_{p+1} = -\sum_{k=1}^{p} \lambda_k \frac{\mu_k}{\mu_{p+1}}\mathbf{x}_k \tag{6.3}$$

    car $\mathbf{x}_k$ est un vecteur propre de A pour la valeur propre λ_k. Par ailleurs

    $$A\mathbf{x}_{p+1} = \lambda_{p+1}\mathbf{x}_{p+1} = -\lambda_{p+1}\sum_{k=1}^{p} \frac{\mu_k}{\mu_{p+1}}\mathbf{x}_k \quad \text{d'après (6.3).} \tag{6.4}$$

    On déduit alors de (6.3) et (6.4) que

    $$-\sum_{k=1}^{p} \lambda_{p+1}\frac{\mu_k}{\mu_{p+1}}\mathbf{x}_k = -\sum_{k=1}^{p} \lambda_k \frac{\mu_k}{\mu_{p+1}}\mathbf{x}_k.$$

La famille $(\mathbf{x}_1, \dots, \mathbf{x}_p)$ étant libre, on en déduit que

$$\lambda_{p+1} \frac{\mu_k}{\mu_{p+1}} = \lambda_k \frac{\mu_k}{\mu_{p+1}} \quad \forall k \in \{1, \dots, p\}.$$

Les valeurs propres $\lambda_1, \dots, \lambda_{k+1}$ étant deux à deux distinctes on a alors nécessairement $\mu_k = 0$ pour tout $k \in \{1, \dots, p\}$. Alors (6.1) implique $\mu_{p+1}\mathbf{x}_{p+1} = 0_{\mathbb{K}^n}$ ce qui est impossible car $\mathbf{x}_{p+1} \neq 0_{\mathbb{K}^n}$ et car on a supposé $\mu_{p+1} \neq 0$.

— on a donc $\mu_{p+1} = 0$, et (6.1) devient

$$\mu_1 \mathbf{x}_1 + \dots \mu_p \mathbf{x}_p = 0_{\mathbb{K}^n}$$

ce qui implique que $\lambda_1 = \dots = \lambda_p = 0$ car la famille $(\mathbf{x}_1, \dots, \mathbf{x}_p)$ est libre.

Finalement $\mu_1\mathbf{x}_1 + \dots \mu_p\mathbf{x}_p + \mu_{p+1}\mathbf{x}_{p+1} = 0_{\mathbb{K}^n}$ implique $\lambda_1 = \dots = \lambda_p = \lambda_{p+1} = 0$: la famille $(\mathbf{x}_1, \dots, \mathbf{x}_p, \mathbf{x}_{p+1})$ est libre. L'hypothèse de récurrence est ainsi vérifiée pour l'entier $p+1$.

□

Comme une famille libre de vecteurs de $\mathbb{K}^n$ ne peut comporter qu'au plus n éléments car $\dim(\mathbb{K}^n) = n$, on déduit en particulier de la Proposition 6.6 la propriété suivante.

Corollaire 6.7 *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ possède au plus n valeurs propres disctinctes.*

6.2 Matrices diagonales, diagonalisables

On commence par introduire les notions de matrices diagonales et diagonalisables.

Définition 6.8 (Matrice diagonale) *Une matrice $D \in \mathcal{M}_n(\mathbb{K})$ est dite diagonale si tous les coefficients situés en dehors de la diagonale de celle-ci sont nuls, c'est-à-dire si elle est de la forme*

$$\begin{pmatrix} d_{11} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & d_{nn} \end{pmatrix}.$$

Remarque: Si D est une matrice diagonale et si on note $\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ le i-ème vecteur de la base canonique de $\mathbb{K}^n$, alors

$$D\mathbf{e}_i = \begin{pmatrix} d_{11} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & d_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ d_{ii} \\ \vdots \\ 0 \end{pmatrix} = d_{ii}\mathbf{e}_i.$$

Les valeurs propres de la matrice D correspondent donc aux termes diagonaux de celle-ci.

Définition 6.9 (Matrice diagonalisable) *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite diagonalisable si elle est semblable à une matrice diagonale, c'est-à-dire s'il existe une matrice $P \in \mathcal{M}_n(\mathbb{K})$ inversible telle que*

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

soit une matrice diagonale.

La Proposition suivante nous montre ensuite que lorsqu'une matrice est diagonalisable, la matrice D de la Définition 6.9 possède les même valeurs propres que la matrice A.

Proposition 6.10 *Deux matrices de $\mathcal{M}_n(\mathbb{K})$ semblables ont les même valeurs propres.*

Preuve : Soient A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$ semblables : il existe donc $P \in \mathcal{M}_n(\mathbb{K})$ inversible telle que $B = P^{-1}AP$. Soit λ une valeur propre de B, $\mathbf{x}$ un vecteur propre de B associé à cette valeur propre λ. Alors

$$\begin{aligned} B\mathbf{x} = \lambda\mathbf{x} \quad &\Rightarrow PB\mathbf{x} = \lambda P\mathbf{x} \\ &\Rightarrow AP\mathbf{x} = \lambda P\mathbf{x} \end{aligned}$$

car $PB = AP$. De plus, comme $\mathbf{x} \neq 0$ et que P est inversible (donc injective), $P\mathbf{x} \neq 0$. Ainsi $P\mathbf{x}$ est un vecteur propre et λ une valeur propre de la matrice A. Toute valeur propre de B est donc une valeur propre de A ; de la même façon on montre que toute valeur propre de A est une valeur propre de B. □

Ainsi une matrice est diagonalisable si on peut la factoriser sous la forme

$$A = P \underbrace{\begin{pmatrix} \lambda_1 & & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & & \dots & \dots & \lambda_n \end{pmatrix}}_{D} P^{-1}$$

les valeurs λ_1, …, λ_n correspondant aux valeurs propres de la matrice A. De plus, la matrice $P \in \mathcal{M}_n(\mathbb{K})$ est inversible et correspond à la matrice de passage de la base canonique à la base formé de vecteurs propres de A. On peut donc reformuler la définition 6.9 de la façon suivante :

Définition 6.11 (Matrice diagonalisable, bis) *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable s'il existe une base de $\mathbb{K}^n$ formé de vecteurs propres de A.*

6.3 Polynôme caractéristique

Donnons maintenant une caractérisation équivalente des valeurs propres en utilisant le déterminant d'une matrice.

Proposition 6.12 *Le scalaire λ est une valeur propre de A si et seulement si det$(A - \lambda I_n) = 0$, c'est à dire si et seulement si*

$$\begin{vmatrix} a_{11}-\lambda & & \dots & \dots & a_{1,n} \\ & a_{22}-\lambda & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \vdots \\ a_{n,1} & & \dots & \dots & a_{nn}-\lambda \end{vmatrix} = 0$$

Preuve : D'après ce que l'on a vu avant, λ est une valeur propre de A si $V_\lambda = \mathrm{Ker}(A - \lambda I_n) \neq \{0_{\mathbb{K}_n}\}$. Or $\mathrm{Ker}(A - \lambda I_n) \neq \{0_{\mathbb{K}_n}\}$ signifie que A est non injective, et A n'est pas injective si et seulement si A n'est pas bijective d'après le Théorème 3.13. Donc $\mathrm{Ker}(A - \lambda I_n) \neq \{0_{\mathbb{K}_n}\}$ si et seulement si $\det(A - \lambda I_n) = 0$ d'après le théorème 5.4. $\square$

Remarque: En particulier, 0 est une valeur propre de A si et seulement $\det(A) = 0$.

La matrice A étant fixée, on peut définir l'application

$$\begin{aligned} \mathbb{K} &\to \mathbb{K} \\ t &\mapsto \det(A - tI_n) = \begin{vmatrix} a_{11}-t & & \dots & \dots & a_{1,n} \\ & a_{22}-t & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \vdots \\ a_{n,1} & & \dots & \dots & a_{nn}-t \end{vmatrix} \end{aligned}$$

En développant $\det(A - tI_n)$, on obtient une expression polynômiale en t. Cela nous amène à introduire la définition suivante :

Définition 6.13 *On appelle polynôme caractéristique de la matrice A le polynôme de $\mathbb{K}[X]$ défini par*

$$\chi_A(X) := \det(A - XI_n).$$

Remarque: Le coefficient dominant de χ_A est égal à $(-1)^n$, et le terme de degré 0 est égal à $\chi_A(0) = \det(A)$. Le polynôme $\chi_A(X)$ a donc la forme suivante

$$\chi_A(X) = (-1)^n X^n + \beta_{n-1} X^{n-1} + \cdots + \det(A).$$

Proposition 6.14 *Si A et B sont deux matrices semblables, alors $\chi_A = \chi_B$.*

La Proposition 6.12 signifie alors que λ est valeur propre de A si et seulement si c'est une racine de son polynôme caractéristique. On peut la reformuler de la façon suivante.

Proposition 6.15 *λ est valeur propre de $A \in \mathcal{M}_n(\mathbb{K})$ si et seulement si $\chi_A(\lambda) = 0$.*

Remarque: Si $\mathbb{K} = \mathbb{C}$, tout polynôme de $\mathbb{C}[X]$ possède au moins une racine dans $\mathbb{C}$. On en déduit donc que toute matrice de $\mathcal{M}_n(\mathbb{C})$ possède au moins une valeur propre dans $\mathbb{C}$.

Pour déterminer les valeurs propres d'une matrice, on peut donc calculer le polynôme caractéristique de celle-ci et en déterminer les racines.

Exemple: On veut déterminer les valeurs propres de la matrice $A = \begin{pmatrix} 3 & 1 \\ -2 & 0 \end{pmatrix}$ ainsi qu'une base des sous-espaces propres associés. On calcule

$$\det(A - tI_3) = \begin{vmatrix} 3-t & 1 \\ -2 & -t \end{vmatrix} = (t-1)(t-2).$$

La matrice A possède donc deux valeurs propres $\lambda_1 = 1$ et $\lambda_2 = 2$. On détermine ensuite une base de chacun des sous-espaces propres associés V_1 et V_2 en résolvant deux systèmes linéaires.

$$AX = X \quad \Leftrightarrow \quad \begin{pmatrix} 3 & 1 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \Leftrightarrow \begin{pmatrix} 2x+y \\ -2x-y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow 2x+y=0.$$

Le sous espace propre V_1 correspond donc à la droite d'équation $2x+y=0$ (il est donc de dimension 1) donc une base est donnée par exemple par le vecteur $\mathbf{e}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. On détermine ensuite V_2 :

$$AX = 2X \quad \Leftrightarrow \quad \begin{pmatrix} 3 & 1 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix} \quad \Leftrightarrow \begin{pmatrix} x+y \\ -2x-2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow x+y=0$$

Le sous espace propre V_2 correspond donc à la droite d'équation $x+y=0$ donc une base est donnée par exemple par le vecteur $\mathbf{e}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Remarque: Si A est une matrice triangulaire supérieure, c'est-à-dire de la forme

$$\begin{pmatrix} a_{11} & \dots & \dots & a_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}$$

alors un calcul simple montre que

$$\chi_A(X) = \det(A - XI_n) = \prod_{i=1}^{n} (a_{ii} - X).$$

Les valeurs propres de la matrice A correspondent donc aux éléments diagonaux de A.

Remarque: Lorsque $n > 3$, on ne sait pas trouver de manière systématique les racines d'un polynôme ! On ne peut donc pas toujours trouver toutes les valeurs propres d'une matrice.

Remarque: Le calcul du polynôme caractéristique de A n'est pas la seule façon de déterminer ses valeurs propres. En particulier, on sait qu'une matrice de $\mathcal{M}_n(\mathbb{K})$ est non inversible si et seulement si son rang est strictement inférieur à n. Pour montrer que λ est valeur propre de A il suffit alors de montrer $\mathrm{rg}(A - \lambda I_n) < n$, ce qui est parfois plus simple que de calculer un déterminant. Par exemple, si on considère la matrice $A = \begin{pmatrix} a & 1 & \dots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & a \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$, on remarque que

$$A - (a-1)I_n = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & 1 \end{pmatrix}$$

qui est une matrice de rang 1. On peut donc en déduire que $a-1$ est valeur propre de A. De plus, d'après le théorème du rang (Théorème 2.11) on a

$$\dim(Ker(A - (a-1)I_n)) = n - \mathrm{rg}(Ker(A - (a-1)I_n)) = n - 1.$$

La dimension de du sous-espace propre $V_{a-1} = Ker(A - (a-1)I_n)$ est donc égale à $n-1$.

Définition 6.16 *On dit qu'un polynôme P de $\mathbb{K}[X]$ est scindé s'il s'écrit sous la forme*

$$P(X) = \prod_{i=1}^{m} (X - \alpha_i)^{\mu_i}$$

avec $\alpha_1, \dots, \alpha_m$ des élements de $\mathbb{K}$ deux à deux distincts, $\mu_i \in \mathbb{N}^*$ pour tout $i \in \{1, \dots, m\}$. L'entier μ_i est appelé ordre de multiplicité de la racine α_i.

Remarque: On a alors $\displaystyle\sum_{i=1}^{m} \mu_i = deg(P)$.

Remarque: Tout polynôme de $\mathbb{C}[X]$ est scindé.

Définition 6.17 *Soit $A \in \mathcal{M}_n(\mathbb{K})$, λ une valeur propre de A. On appelle multiplicité de λ la multiplicité μ de celle-ci en tant que racine du polynôme caractérique de A χ_A. Si $\mu = 1$, on dit que la valeur propre λ est simple. Si $\mu > 1$, on dit que la valeur propre λ est multiple.*

Proposition 6.18 *Soit $A \in \mathcal{M}_n(\mathbb{K})$ dont le polynôme caractéristique est scindé et $\lambda_1, \dots, \lambda_n$ ses valeurs propres (non nécessairement disctinctes). Alors*

$$\sum_{i=1}^{n} \lambda_i = \mathrm{tr}(A), \quad \prod_{i=1}^{n} \lambda_i = \det(A).$$

Remarque: Si $A \in \mathcal{M}_2(\mathbb{C})$, alors son polynôme caractéristique s'écrit

$$\chi_A(X) = X^2 - \mathrm{tr}(A)X + \det(A).$$

Théorème 6.19 *Soit $A \in \mathcal{M}_n(\mathbb{K})$, λ une valeur propre de A de multiplicité μ. Alors*

$$1 \leq \dim(V_\lambda) \leq \mu.$$

6.4 Théorèmes de diagonalisabilité

6.4.1 Premiers théorèmes

On commence par montrer une condition suffisante simple de diagonalisabilité.

Proposition 6.20 *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ dont le polynôme caractéristique est scindé et possède n racines simples est diagonalisable.*

Preuve : Si la matrice A possède n valeurs propres distinctes, elle posséde alors n vecteurs propres associés à des valeurs propres distinctes. D'après la Proposition 6.6, ces vecteurs propres forment une famille libre. Puisqu'elle est composée de n éléments de $\mathbb{K}^n$ qui est de dimension n, cette famille de vecteurs propres est une base de $\mathbb{K}^n$. D'après la Définition 6.11, A est donc diagonalisable. $\square$

Exemple: On a vu que la matrice $A = \begin{pmatrix} 3 & 1 \\ -2 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$ a pour polynôme caractéristique $\chi_A(X) = (X-1)(X-2)$, qui est scindé et a comme racines 1 et 2. La matrice A est donc diagonalisable. De plus, $\mathbf{e}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ étant un vecteur propre associé à la valeur propre 1 et $\mathbf{e}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ un vecteur propre associé à la valeur propre 2, la matrice de passage de la base canonique à la base $(\mathbf{e}_1, \mathbf{e}_2)$ est

$$P = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}.$$

Dans le cas général, on peut ensuite montrer le résultat suivant.

Théorème 6.21 *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable si et seulement si la somme des dimensions de ses sous-espaces propres est égale à n.*

En pratique, on utilise souvent le polynôme caractéristique. On a alors la caratérisation fondamentale de la diagonalisabilité des matrices suivante.

Théorème 6.22 *Une matrice est diagonalisable si et seulement si son polynôme caractéristique est scindé et la dimension de chaque sous espace propre est égale à l'ordre de multiplicité de la valeur propre correspondante.*

Exemple: Soit $A = \begin{pmatrix} 1 & 4 & -2 \\ 0 & 6 & -3 \\ -1 & 4 & 0 \end{pmatrix}$. Calculons son polynôme caractéristique.

$$\begin{aligned}
\det(A - tI_3) &= \begin{vmatrix} 1-t & 4 & -2 \\ 0 & 6-t & -3 \\ -1 & 4 & -t \end{vmatrix} \\
&= \begin{vmatrix} 3-t & 4 & -2 \\ 3-t & 6-t & -3 \\ 3-t & 4 & -t \end{vmatrix} \qquad\qquad (C_1' = C_1 + C_2 + C3) \\
&= (3-t) \begin{vmatrix} 1 & 4 & -2 \\ 1 & 6-t & -3 \\ 1 & 4 & -t \end{vmatrix} \\
&= (3-t) \begin{vmatrix} 1 & 4 & -2 \\ 1 & 6-t & -3 \\ 0 & 0 & 2-t \end{vmatrix} \qquad\qquad (L_3' = L_3 - L_1) \\
&= (3-t)(2-t) \begin{vmatrix} 1 & 4 \\ 1 & 6-t \end{vmatrix} \\
&= (3-t)(2-t)^2.
\end{aligned}$$

Le polynôme χ_A est donc scindé et possède 3 comme racine simple et 2 comme racine double. Pour savoir si A est diagonalisable, il faut donc regarder la dimension du sous espace propre associé à la valeur propre 2. On résout pour cela le système $AX = 2X$.

$$\begin{aligned}
AX = 2X \quad &\Leftrightarrow \begin{pmatrix} 1 & 4 & -2 \\ 0 & 6 & -3 \\ -1 & 4 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&\Leftrightarrow \left\{ \begin{array}{rcl} x_1 & = & x_3 \\ 4x_2 & = & 3x_3 \end{array} \right.
\end{aligned}$$

Ainsi le sous espace propre V_2 est engendré par le vecteur $\begin{pmatrix} 4 \\ 3 \\ 4 \end{pmatrix}$ et il est de dimension 1. La multiplicité de la valeur propre 2 n'est donc pas égal à la dimension de V_2 : par conséquent A n'est pas diagonalisable.

6.4.2 Polynômes annulateurs

De la même façon qu'on a défini les polynômes d'endomorphismes au Chapitre 2, on peut définir les polômes de matrices

Définition 6.23 (polynôme de matrice) *Soit* $P := \sum_{k=0}^{N} a_k X^k \in \mathbb{K}[X]$ *un polynôme à coeffi-*

cients dans $\mathbb{K}$, et soit $A \in \mathcal{M}_n(\mathbb{K})$. On définit $P(A) \in \mathcal{M}_n(\mathbb{K})$ par $P(A) := \sum_{k=0}^{N} a_k A^k$.

Définition 6.24 (polynôme annulateur) *Soit $A \in \mathcal{M}_n(\mathbb{K})$, $P \in \mathbb{K}[X]$. On dit que P est un polynôme annulateur de A si $P(A) = [0]$.*

Exemple: Soit $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. alors $A^2 = I_2$, et donc $P(X) = X^2 - 1$ est un polynôme annulateur de A.

Donnons maintenant quelques théorèmes (admis dans ce cours) qui permettent de déterminer si une matrice est diagonalisable en utilisant cette notion.

Lemme 6.25 *Soit $A \in \mathcal{M}_n(\mathbb{K})$, $P \in \mathbb{K}[X]$ un polynôme annulateur de A. Alors pour toute valeur propre λ de A on a $P(\lambda) = 0$. Ainsi, le spectre de A est inclus dans l'ensemble des racines de n'importe quel polynôme annulateur de A.*

Remarque: Attention, la réciproque est fausse. Par exemple, $P(X) = X(X-1)$ est un polynôme annulateur de la matrice identité I_n mais 0 n'est pas valeur propre de I_n.

Théorème 6.26 *Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors A est diagonalisable si et seulement si A possède un polynôme annulateur $P \in \mathbb{K}[X]$ scindé sur $\mathbb{K}$ et à racines simples.*

Exemple: Si $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice vérifiant $A^2 = A$, alors A est diagonalisable.

Théorème 6.27 (Théorème de Cayley-Hamilton) *Soit $A \in \mathcal{M}_n(\mathbb{K})$, χ_A son polynôme caractéristique. Alors $\chi_A(A) = 0$, autrement dit, le polynôme caractéristique d'une matrice est un polynôme annulateur de celle-ci.*

Exemple: Soit A une matrice strictement triangulaire (c'est-à-dire triangulaire dont les élements diagonaux sont nuls). Alors le théorème de Cayley-Hamilton permet d'obtenir que $A^n = [0]$ (on dit que A est *nilpotente*).

6.5 Matrices remarquables

Dans toute la suite, $\mathbb{K} = \mathbb{C}$ ou $\mathbb{R}$. On s'intéresse maintenant au cas particulier des matrices hermitiennes (respectivement symétriques) et unitaires (respectivement orthogonales).

6.5.1 Réduction des matrices auto-adjointes

Définition 6.28 *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est autoadjointe si $A = A^*$*

Lemme 6.29 *Soit $A \in \mathcal{M}_n(\mathbb{K})$, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, une matrice autoadjointe, c'est-à-dire telle que $A = A^*$. Alors le polynôme caractéristique de A est à coefficients réels et il est scindé dans $\mathbb{R}[X]$.*

Ainsi, les valeurs propres d'une matrice autoadjointe dans $\mathbb{K} = \mathbb{C}$ ou $\mathbb{R}$ sont nécessairement réelles.

Lemme 6.30 *Soit $A \in \mathcal{M}_n(\mathbb{K})$, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ une matrice autoadjointe, soit λ et μ deux valeurs propres distinctes de A. Alors les sous-espaces propres V_λ et V_μ sont orthogonaux.*

Autrement dit, si $\mathbf{x}_1$ et $\mathbf{x}_2$ sont des vecteurs propres d'une matrice autoadjointe associés à des valeurs propres distinctes, on aura $< \mathbf{x}_1, \mathbf{x}_2 >= 0$.

Théorème 6.31 (Diagonalisation des matrices hermitiennes) *Soit $A \in \mathcal{M}_n(\mathbb{C})$. Les assertions suivantes sont équivalentes :*

i) A est hermitienne.

ii) Il existe une base othonormale de $\mathbb{C}^n$ composé de vecteurs propres de A et les valeurs propres associées sont réelles.

iii) A est semblable à une matrice diagonale réelle avec une matrice de passage unitaire, autrement dit A peut s'écrire sous la forme

$$A = U \begin{pmatrix} \lambda_1 & & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & & \dots & \dots & \lambda_n \end{pmatrix} U^*$$

avec $\lambda_i \in \mathbb{R}$ pour tout $i \in \{1, \dots, n\}$ et $U \in \mathcal{M}_n(\mathbb{C})$ vérifiant $U^* = U^{-1}$.

Théorème 6.32 (Diagonalisation des matrices symétriques réelles) *Soit $A \in \mathcal{M}_n(\mathbb{R})$. Les assertions suivantes sont équivalentes :*

i) A est symétrique.

ii) Il existe une base othonormale de $\mathbb{R}^n$ composé de vecteurs propres de A.

iii) A est semblable à une matrice diagonale réelle avec une matrice de passage orthogonale, autrement dit, A peut s'écrire sous la forme

$$A = O \begin{pmatrix} \lambda_1 & & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & & \dots & \dots & \lambda_n \end{pmatrix} O^t$$

avec $\lambda_i \in \mathbb{R}$ pour tout $i \in \{1, \dots, n\}$ et $O \in \mathcal{M}_n(\mathbb{R})$ vérifiant $O^t = O^{-1}$.

6.5.2 Réduction des matrices unitaires et orthogonales

Lemme 6.33 *Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice unitaire. Alors*

— *si λ est une valeur propre de A, $|\lambda| = 1$.*

— *si λ et μ sont des valeurs propres distinctes de A, les sous-espaces propres V_λ et V_μ sont orthogonaux.*

Théorème 6.34 (Diagonalisation des matrices unitaires) *Soit $A \in \mathcal{M}_n(\mathbb{C})$. Les assertions suivantes sont équivalentes :*

i) A est unitaire.

ii) Il existe une base othonormale de $\mathbb{C}^n$ composé de vecteurs propres de A et les valeurs propres associées sont de module égal à 1.

iii) A est semblable à une matrice diagonale avec des termes diagonaux de module 1 et une matrice de passage unitaire, autrement dit A peut s'écrire sous la forme

$$A = U \begin{pmatrix} \lambda_1 & & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & & \dots & \dots & \lambda_n \end{pmatrix} U^*$$

avec $\lambda_i \in \mathbb{C}$, $|\lambda_i| = 1$ pour tout $i \in \{1, \dots, n\}$ et $U \in \mathcal{M}_n(\mathbb{C})$ vérifiant $U^{-1} = U^*$.

Lemme 6.35 *Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice orthogonale. Alors*

- *si λ est une valeur propre réelle de A, $\lambda \in \{-1, 1\}$.*
- *si -1 et 1 sont valeurs propres de A, les sous-espaces propres V_{-1} et V_1 sont orthogonaux.*
- *son polynôme caractéristique s'écrit sous la forme suivante :*

$$\chi_A(X) = (1 - X)^p (1 - X)^q \prod_{i=1}^{m} (X^2 - 2\cos(\theta_i) X + 1)$$

avec $\theta_i \in \mathbb{R}$, $\theta_i \notin \pi\mathbb{Z}$, $m \in \mathbb{N}$, $p \in \mathbb{N}$, $q \in \mathbb{N}$.

Contrairement aux matrices unitaires, les matrices orthogonales ne sont pas toujours diagonalisables. On peut cependant les écrire sous une forme diagonale par blocs.

Théorème 6.36 (Réduction des matrices orthogonales) *Soit $A \in \mathcal{M}_n(\mathbb{R})$. Les assertions suivantes sont équivalentes :*

i) A est orthogonale.

ii) A peut s'écrire sous la forme

$$A = O \begin{pmatrix} 1 & 0 & & & \dots & & & & & 0 \\ 0 & \ddots & \ddots & & & & & & & \vdots \\ \vdots & \ddots & -1 & & & & & & & \\ & & & \ddots & & & & & & \\ & & & & -1 & 0 & 0 & & & \\ & & & & 0 & \cos\theta_1 & -\sin\theta_1 & & & \\ & & & & & \sin\theta_1 & \cos\theta_1 & & & \vdots \\ & & & & & & 0 & \ddots & & 0 \\ \vdots & & & & & & & & \cos\theta_m & -\sin\theta_m \\ 0 & & & \dots & & & & & \sin\theta_m & \cos\theta_m \end{pmatrix} O^t, \quad \text{avec } \theta_i \in \mathbb{R},\ m \in \mathbb{N}.$$

Chapitre 7

Équations différentielles ordinaires

7.1 Exponentielle de matrices

7.1.1 Définition et premières propriétés

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

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

De la même facon que nous avions défini des polynômes de matrices carrées, nous pouvons définir leur exponentielle par la formule suivante :

Définition 7.1 *Soit $A \in \mathcal{M}_n(\mathbb{K})$. On pose*

$$\exp A = I_n + A + \frac{A^2}{2!} + \ldots + \frac{A^i}{i!} + \ldots = \sum_{i=0}^{\infty} \frac{A^i}{i!}. \tag{7.1}$$

Bien entendu, cette définition n'a de sens que si la suite précédente converge, au sens ou chacun des éléments de la matrice possède une limite : nous admettrons ce résultat. Notons également que $\exp A$ est une limite de polynômes de A.

Donnons maintenant quelques propriétés.

Proposition 7.2 *Soient A et B deux éléments de $\mathcal{M}_n(\mathbb{K})$.*
i) Si A et B commutent, alors $\exp A \exp B = \exp(A + B)$.
ii) $\exp A$ est inversible d'inverse $\exp(-A)$.
iii) Si P est inversible alors $\exp(P^{-1}AP) = P^{-1}(\exp A)P$.

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

$$\exp(A + (-A)) = \exp 0 = I_n.$$

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

7.1.2 Calculs d'exponentielles

Le cas des matrices diagonalisables. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice diagonalisable, et soit $\lambda_1, \lambda_2, \dots \lambda_n)$, la suite des valeurs propres, répétées avec leur multiplicité. Dans une base de $\mathbb{K}^n$ de vecteurs propres $(\mathbf{u}_1, \dots, \mathbf{u}_n)$, la matrice associée dans cette base est alors

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

Rappelons que si p est un polynôme, alors

$$p(D) = \operatorname{diag}(p(\lambda_1), \dots, p(\lambda_n)).$$

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

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

c'est-à-dire

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

Si P est la matrice de passage de la base canonique $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ vers la base $(\mathbf{u}_1, \dots, \mathbf{u}_n)$, on a

$$A = PDP^{-1},$$

d'où l'on déduit d'après la Proposition 7.2 que

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

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

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

Si on revient à la formule (7.1), on s'aperçoit alors que cette dernière ne contient qu'un nombre fini de termes. on a donc

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

Exemple: Pour $N = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$, on obtient

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

S'il existe c tel que $N = A - cI_n$ soit nilpotente. Dans ce cas, $\exp(tA)$ est donné par

$$\exp(tA) = \exp(tc)\left(\sum_{k=0}^{p-1} \frac{t^k}{k!} N^k\right), \quad \text{où } p \text{ est l'indice de nilpotence de } N.$$

7.1.3 Dérivation

Proposition 7.3 *Soit $A \in \mathcal{M}_n(\mathbb{K})$, et φ la fonction définie par*

$$\begin{array}{rccl} \varphi: & \mathbb{R}^+ & \to & \mathcal{M}_n(\mathbb{K}) \\ & t & \to & \exp(tA). \end{array}$$

Alors la fonction φ est dérivable, et sa dérivée est donnée par

$$\varphi'(t) = A\exp(tA) = \exp(tA)A.$$

7.2 Équations différentielles linéaires d'ordre 1

7.2.1 Définitions et propriétés

Définition 7.4 *Soit I un intervalle de $\mathbb{R}$. On appelle équation différentielle linéaire (ou système différentiel linéaire) l'équation différentielle*

$$Y'(t) = A(t)Y(t) + B(t), \tag{E}$$

où

$$A: \left\{ \begin{array}{ccl} I & \to & \mathcal{M}_n(\mathbb{K}) \\ t & \to & A(t) = (a_{i,j}(t))_{1\le i,j\le n} \end{array} \right. \qquad et \qquad B: \left\{ \begin{array}{ccl} I & \to & \mathbb{K}^n \\ t & \to & B(t) = (b_1(t), \ldots, b_n(t))^T \end{array} \right.$$

sont des fonctions continues données.

On se place dans la suite dans le cas où $I = \mathbb{R}^+$ pour simplifier. Le théorème suivant, que nous admettrons et qui est un cas particulier du théorème général de Cauchy-Lipschitz, garantit l'existence de solutions à (E), et l'unicité de la solution si l'on se fixe la condition initiale (c'est-à-dire la valeur de la solution à $t = 0$).

Théorème 7.5 (Théorème de Cauchy-Lipschitz) *Soit $Y_0 \in \mathbb{K}^n$, A et B deux fonctions définies et continues sur $\mathbb{R}^+$ et à valeurs dans $\mathcal{M}_n(\mathbb{K})$ et $\mathbb{K}^n$ respectivement. Alors l'équation différentielle avec condition initiale (appelée aussi* problème de Cauchy*)*

$$\left\{ \begin{array}{l} Y'(t) = A(t)Y(t) + B(t), \\ Y(0) = Y_0 \end{array} \right.$$

admet une unique solution définie sur $\mathbb{R}^+$ tout entier, et cette solution est de classe $\mathcal{C}^1$ sur $\mathbb{R}^+$.

Définition 7.6 *On appelle équation différentielle linéaire homogène (ou "sans second membre") l'équation différentielle*

$$Y'(t) = A(t)Y(t), \tag{H}$$

où A est une fonction définie et continue sur $\mathbb{R}^+$ à valeurs dans $\mathcal{M}_n(\mathbb{K})$.

On a alors la propriété fondamentale suivante :

Proposition 7.7 *L'ensemble $\mathcal{S}$ des solutions définies sur $\mathbb{R}^+$ de l'équation différentielle linéaire homogène (H) est un $\mathbb{K}$-espace vectoriel de dimension n.*

Preuve : On vérifie tout d'abord facilement que si Y_1 et Y_2 sont dans $\mathcal{S}$, alors $Y_1 + \lambda Y_2$ est encore dans $\mathcal{S}$, pour tout scalaire $\lambda \in \mathbb{K}$. Considérons ensuite l'application

$$\phi : \left\{ \begin{array}{ccc} \mathcal{S} & \rightarrow & \mathbb{K}^n \\ Y & \rightarrow & Y(0). \end{array} \right.$$

Cette application est linéraire entre les espaces vectoriels $\mathcal{S}$ et $\mathbb{K}^n$. De plus, le thérorème 7.5 montre qu'elle est bijective. Ainsi $\dim(\mathcal{S}) = \dim(\mathbb{K}^n) = n$. □

On en déduit ensuite la Proposition suivante, qui permet en pratique de déterminer toutes les solutions de l'équation différentielle (E).

Proposition 7.8 *Soit Y_p une solution de l'équation différentielle linéaire (E). Alors l'ensemble des solutions de (E) est donné par*

$$\{Y_p + Z, Z \in \mathcal{S}\}.$$

Remarque: L'ensemble des solutions de l'équation différentielle (E) est donc un espace affine de dimension n, admettant $\mathcal{S}$ comme direction vectorielle.

On se place dans la suite de cette partie dans le cas où $A(t)$ est une matrice à coefficients indépendants de t, notée A.

7.2.2 Équation homogène

Théorème 7.9 *Soit $A \in \mathcal{M}_n(\mathbb{K})$. Les solutions de l'équation*

$$Y'(t) = AY(t), \quad t \in \mathbb{R} \tag{H_c}$$

sont données par

$$Y(t) = \exp(tA)\, U, \quad U \in \mathbb{K}^n.$$

En particulier, la solution de (H_c) telle que $Y(0) = Y_0$ est $Y(t) = \exp(At)\, Y_0$.

Cas de la dimension 1

Proposition 7.10 *L'équation différentielle homogène*

$$y'(t) = a(t)\, y(t) \tag{H_1}$$

a pour solution générale

$$y_h(t) = \alpha_0 \exp\left(\int_0^t a(s)ds\right), \quad \alpha_0 \in \mathbb{K}.$$

La fonction $F : t \mapsto \exp\left(\int_0^t a(s)ds\right)$ est donc une base de l'espace vectoriel (de dimension 1 sur $\mathbb{K}$) des solutions de (H_1).

Lorsque A est à coefficients constants et est diagonalisable, on peut se passer de calculer $\exp(tA)$.

Proposition 7.11 *Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice à coefficients constants diagonalisable, $\lambda_1 \ldots \lambda_n$ ses valeurs propres (non nécessairement distinctes) et $\mathbf{v}_1, \ldots, \mathbf{v}_n$ une base de vecteurs propres associés. On note*

$$\begin{array}{rcl} \mathbf{f}_i : \mathbb{R} & \mapsto & \mathbb{K}^n \\ t & \mapsto & e^{\lambda_i t}\mathbf{v}_i \end{array}.$$

Alors la famille de fonctions $(\mathbf{f}_1, \ldots, \mathbf{f}_n)$ est une base de l'espace des solutions de l'équation

$$Y'(t) = AY(t), \quad t \in \mathbb{R}. \tag{H_c}$$

En particulier, si $Y_0 \in \mathbb{K}^n$ se décompose dans la base $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ sous la forme $Y_0 = \sum_{i=1}^{n} y_i\mathbf{v}_i$, alors la solution de (H_c) telle que $Y(0) = Y_0$ s'écrit

$$Y(t) = \sum_{i=1}^{n} y_i\mathbf{f}_i.$$

Remarque: Si A n'est pas diagonalisable mais qu'il existe une matrice T triangulaire et une matrice P inversible telles que $A = PTP^{-1}$, la méthode décrite ci-dessus peut s'adapter de la faon suivante. Le système (S) devient

$$Y'(t) = PTP^{-1}Y(t).$$

On effectue alors le changement de fonction inconnue

$$U(t) = P^{-1}Y(t)$$

et on obtient le système différentiel d'inconnue U suivant :

$$U'(t) = TU(t).$$

On est alors ramené à résoudre un système d'équations différentielles linéaires du 1er ordre en cascade (que l'on résout à partir de la dernière équation si T est triangulaire supérieure).

7.2.3 Recherche d'une solution particulière de l'équation avec second membre

Si aucune solution évidence n'apparaît, on peut utiliser la *méthode de la variation des constantes.* Dans le cas où la solution de l'équation homogène $Y'(t) = AY(t)$ est exprimée avec l'exponentielle de matrice, on cherche une solution particulière sous la forme

$$Y_p(t) = \exp(At)V(t),$$

où $V : I \to \mathbb{K}^n$ est la fonction à trouver. Alors

$$Y_p'(t) - AY_p(t) = \exp(tA)\,V'(t)$$

et il suffit donc de choisir V telle que $\exp(tA))V'(t) = B(t)$, soit par exemple

$$V(t) = \int_0^t \exp(-uA)\,B(u)du.$$

On obtient alors la Proposition suivante

Théorème 7.12 *L'ensemble des solutions de l'équation*

$$Y'(t) = AY(t) + B(t) \tag{E_c}$$

est

$$\left\{ Y(t) = \exp(tA)\, V + \int_0^t \exp((t-u)\, A)\; B(u)du, \quad V \in \mathbb{K}^n \right\}$$

En particulier, la solution de l'équation (E_c) telle que $Y(0) = Y_0$ est

$$Y(t) = \exp(tA)\, Y_0 + \int_0^t \exp((t-u)\, A)\; B(u)du.$$

On peut aussi utiliser la méthode de variation des constantes à partir d'une base de solutions de l'équation homogène. Soit $\mathbf{f}_1, \ldots, \mathbf{f}_n$ une solution de l'équation homogène $Y'(t) = AY(t)$. On cherche alors une solution sous la forme

$$Y_p(t) = \sum_{i=1}^{n} \mu_i(t)\mathbf{f}_i(t).$$

On obtient alors

$$\begin{aligned} Y_p(t)' - AY_p(t) &= \sum_{i=1}^{n} \mu_i'(t)\mathbf{f}_i(t) + \sum_{i=1}^{n} \mu_i(t)\mathbf{f}_i'(t) - A\sum_{i=1}^{n} \mu_i(t)\mathbf{f}_i(t) \\ &= \sum_{i=1}^{n} \mu_i'(t)\mathbf{f}_i(t). \end{aligned}$$

Pour que Y_p soit une solution de (E_c), il faut et il suffit donc que $\displaystyle\sum_{i=1}^{n} \mu_i'(t)\mathbf{f}_i(t) = B(t)$, soit

$$F(t)X(t) = B(t), \quad \text{avec} \quad F(t) = (\mathbf{f}_1(t), \ldots, \mathbf{f}_n(t)), \quad X(t) = (\mu_i'(t))_{1 \le i \le n}.$$

On aboutit alors à un système de taille $n * n$, qui possède une unique solution car la matrice $F(t)$ est inversible (ses vecteurs colonnes forment une famille libre).

Proposition 7.13 *L'ensemble des solutions de l'équation*

$$Y'(t) = AY(t) + B(t) \tag{E_c}$$

est

$$\left\{ Y(t) = \sum_{i=1}^{n} \left(\mu_i^0 + \int_0^t x_i(s)ds \right) \mathbf{f}_i(t), \quad (y_i)_{1 \le i \le n} \in \mathbb{K}^n, \right\}$$

où pour tout $t \in \mathbb{R}^+$, $X(t) = (x_i)_{1 \le i \le n}$ est la solution du système

$$F(t)X(t) = B(t), \quad \textit{avec} \quad F(t) = (\mathbf{f}_1(t), \ldots, \mathbf{f}_n(t)).$$

En particulier, si Y_0 se décompose dans la base $\{\mathbf{f}_1(0), \ldots, \mathbf{f}_n(0)\}$ de $\mathbb{K}^n$ sous la forme $Y_0 = \displaystyle\sum_{i=1}^{n} y_i \mathbf{f}_i(0)$, alors la solution de l'équation (E_c) telle que $Y(0) = Y_0$ est

$$Y(t) = \sum_{i=1}^{n} \left(y_i + \int_0^t x_i(s)ds \right) \mathbf{f}_i(t).$$

7.2.4 Recollement de solutions

Dans le cas où l'équation différentielle s'écrit sous la forme

$$c(t)Y'(t) = A(t)y(t) + B(t),$$

où c est une fonction s'annulant sur $\mathbb{R}^+$ en un point t^*, on est amené à résoudre l'équation différentielle

$$Y'(t) = \frac{1}{c(t)}A(t)yYt) + \frac{1}{c(t)}B(t). \tag{$\bar{E}_1$}$$

$]0, t^*[$ et $]t^*, +\infty[$ à gauche et à droite de t^*.

Présentons ici dans le cas où $n = 1$ une façon d'obtenir une solution sur $\mathbb{R}^+$ tout entier en "recollant" des solutions obtenues sur chacun des intervalles $[0, t^*[$ et $]t^*, +\infty[$.

— sur $[0, t^*[$, y s'écrit de la forme

$$y_1(t) = \alpha_1 + \int_0^t \frac{b(\tau)}{c(\tau)f_1(\tau)}d\tau, \quad \text{avec} \quad f_1(t) = \exp\left(\int_{t_1}^t \frac{a(s)}{c(s)}ds\right)$$

— sur $]t^*, +\infty[$, étant donné $t_2 \in]t^*, +\infty[$, y s'écrit de la forme

$$y_2(t) = \alpha_2 + \int_{t_2}^t \frac{b(\tau)}{c(\tau)f_2(\tau)}d\tau, \quad \text{avec} \quad f_2(t) = \exp\left(\int_{t_2}^t \frac{a(s)}{c(s)}ds\right).$$

Il est alors parfois possible d'obtenir une solution sur $\mathbb{R}^+$ tout entier en recollant deux solutions par continuité, s'il l'on peut choisir α_1 et α_2 tels que

$$\begin{cases} \lim\limits_{\substack{t \to t^* \\ <}} y_1(t) & = & \lim\limits_{\substack{t \to t^* \\ >}} y_2(t) \\ \lim\limits_{\substack{t \to t^* \\ <}} y_1'(t) & = & \lim\limits_{\substack{t \to t^* \\ >}} y_2'(t). \end{cases}$$

7.3 Équations différentielles d'ordre m

7.3.1 Des équations d'ordre m aux systèmes d'ordre 1

Dans cette partie, on s'intéresse à la résolution des équations de la forme

$$\begin{cases} y^{(m)}(t) + a_{m-1}(t)y^{(m-1)}(t) + \cdots + a_1 y'(t) + a_0 y(t) = f(t), \quad \forall t > 0, \\ y^{(k)}(0) = y_{k,0}, \quad \forall k \in \{0, \ldots, m-1\}. \end{cases} \tag{7.2}$$

L'idée est de se ramener à l'étude d'un système d'ordre 1. Pour cela, on pose

$$Y(t) = \begin{pmatrix} y(t) \\ y't) \\ \vdots \\ y^{(m-2)}(t) \\ y^{(m-1)}(t) \end{pmatrix}, \qquad \forall t \in \mathbb{R}^+.$$

On a alors

$$Y'(t) = \begin{pmatrix} y'(t) \\ y''(t) \\ \vdots \\ y^{(m-1)}(t) \\ y^{(m)}(t) \end{pmatrix} = \begin{pmatrix} y'(t) \\ y''(t) \\ \vdots \\ y^{(m-1)}(t) \\ f(t) - a_{m-1}(t)y^{(m-1)}(t) - \cdots - a_1 y'(t) - a_0 y(t) \end{pmatrix}, \qquad \forall t > 0.$$

On en déduit que Y est solution du système différentiel

$$\begin{cases} Y'(t) = A(t)Y(t) + F(t), & \forall t > 0, \\ Y(0) = Y_0 \end{cases} \tag{7.3}$$

où

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & 0 & 1 & 0 \\ 0 & & & & 0 & 1 \\ -a_0 & -a_1 & \dots & \dots & \dots & -a_{m-1} \end{pmatrix}, \qquad F = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f \end{pmatrix} \qquad \text{et} \quad Y_0 = \begin{pmatrix} y_{0,0} \\ y_{1,0} \\ \vdots \\ y_{m-1,0} \end{pmatrix}.$$

Résoudre l'équation d'ordre m (7.2) se ramène à la résolution du système d'ordre 1 donné par (7.3). Dans le cas où le système est à coefficients constants, on peut alors utiliser les techniques proposées dans la partie 7.2.1.

7.3.2 Cas des équations scalaires d'ordre 2

Considérons ici une équation différentielle scalaire d'ordre 2 de la forme

$$(E_2): \; y''(t) + c_1(t)y'(t) + c_0(t)y(t) = b(t),$$

où $b, c, f \, : \, I \to \mathbb{K}$ sont continues. L'équation (E_2) est équivalente à un système différentiel d'ordre 1 dans $\mathbb{K}^2$

$$(S): \; Y' = A(t)X + B(t), \quad \text{avec} \;\; Y = \begin{pmatrix} y \\ y' \end{pmatrix} \qquad A(t) = \begin{pmatrix} 0 & 1 \\ -c_0(t) & -c_1(t) \end{pmatrix} \quad \text{et} \;\; B(t) = \begin{pmatrix} 0 \\ b(t) \end{pmatrix}$$

Cas de coefficients constants

L'utilisation de la section 7.2.1 permet alors d'obtenir le résultat suivant.

Proposition 7.14 *L'ensemble des solutions de l'équation homogène*

$$y''(t) + c_1 y'(t) + c_0 y(t) = 0 \tag{H_{2c}}$$

est un $\mathbb{K}$-espace vectoriel de dimension 2. Soit

$$P(\lambda) = \lambda^2 + c_1\lambda + c_0$$

l'équation caractéristique de cette équation

— *Si P possède deux racines distinctes λ_1 et λ_2, alors une base de l'espace des solutions de (H_{2c}) est donnée par*
$$\mathcal{B} = \left\{ f_1 : t \mapsto e^{\lambda_1 t},\, f_2 : t \mapsto e^{\lambda_2 t} \right\}.$$
— *Si P possède une racine double λ_0, alors une base de l'espace des solutions de (H_{2c}) est donnée par*
$$\mathcal{B} = \left\{ f_1 : t \mapsto e^{\lambda_0 t},\, f_2 : t \mapsto te^{\lambda_0 t} \right\}.$$

Remarque: Le polynôme P correspond au polynôme caractéristique de la matrice A.

Remarque: Si l'équation (H_{2c}) est à coefficients dans $\mathbb{R}$ ($c_1 \in \mathbb{R}$, $c_2 \in \mathbb{R}$), les solutions sont alors à coefficients réels aussi. Dans le cas où P possède deux racines distinctes λ_1 et λ_2, ces racines sont soit réelles, soit complexes conjuguées $\lambda_1 = \mu + i\omega$ et $\lambda_2 = \mu - i\omega$, $(\mu, \omega) \in \mathbb{R}^2$: alors dans ce dernier cas la base de l'espace des solutions peut s'écrire
$$\mathcal{B} = \left\{ t \mapsto e^{\mu t}\cos(\omega t),\, t \mapsto e^{\mu t}\sin(\omega t) \right\}.$$

On peut ensuite utiliser la méthode de variation de la constante décrite précedemment. On obtient alors le résultat suivant.

Proposition 7.15 *Soit (f_1, f_2) une base de l'espace des solutions de l'équation homogène (H_{2c}). L'ensemble des solutions de l'équation*
$$y''(t) + c_1 y'(t) + c_0 y(t) = b(t) \tag{E_{2c}}$$
est de la forme
$$\left\{ y(t) = \left(\mu_1^0 + \int_0^t x_1(s)ds\right) f_1(t) + \left(\mu_2^0 + \int_0^t x_2(s)ds\right) f_2(t) \quad (\mu_i^0)_{1\le i\le 2} \in \mathbb{K}^2, \right\}$$
où pour tout $s \in \mathbb{R}^+$, $(x_1(s), x_2(s))^T$ est la solution du système
$$\begin{pmatrix} f_1(s) & f_2(s) \\ f_1'(s) & f_2'(s) \end{pmatrix} \begin{pmatrix} x_1(s) \\ x_2(s) \end{pmatrix} = \begin{pmatrix} 0 \\ b(s) \end{pmatrix}.$$

Lorsque le second membre $b(t)$ possède une forme particulière, on peut dans certains cas utiliser d'autres résultats.

Théorème 7.16 *Si le second membre de (E_{2c}) est de la forme $b(t) = Q(t)e^{\lambda t}$, où Q est un polynôme, alors (E_{2c}) admet une solution de la forme*
$$y_p(t) = t^m R(t) e^{\lambda t},$$
où R est un polynôme de même degré que Q et
— *$m = 0$ si λ n'est pas racine du polynôme $P(\lambda) = \lambda^2 + c_1\lambda + c_0$,*
— *$m = 1$ si λ est racine simple du polynôme P,*
— *$m = 2$ si λ est racine double du polynôme P.*

En pratique, on peut alors trouver les coefficients du polynôme R par identification. A noter deux cas particuliers du théorème précédent :

Corollaire 7.17

— *si* $b(t) = Q(t)$ *est un polynôme de degré* d *et si* $c_0 \neq 0$*, alors* (E_{2c}) *admet une solution particulière polynomiale de degré* d.
— *si* $b(t) = Q_1(t)\sin(\omega t) + Q_2(t)\cos(\omega t)$*, où* Q_1 *et* Q_2 *sont deux polynômes de* $\mathbb{R}[X]$ *(ou des constantes), alors on cherche une solution particulière de la forme*

$$y_p(t) = t^m \left(R_1(t)\sin(\omega t) + R_2(t)\cos(\omega t)\right),$$

où $d^\circ R = d^\circ Q$*, et* m *est la multiplicité (éventuellement nulle) de* $i\omega$ *comme racine du polynôme* P.

7.3.3 Cas de coefficients non constants

Revenons au cas de l'équation (E_2) à coefficients non constants. Il n'y a pas de méthode systématique pour ce genre d'équation, et on ne sait donc pas toujours la résoudre explicitement. Présentons ici quelques situations particulières.

Méthode de variation de la constante L'utilisation de la méthode suppose que l'on connaisse une base (f_1, f_2) de l'équation homogène associée à (E_2). La Proposition 7.18 s'applique alors encore ici

Proposition 7.18 *Soit* (f_1, f_2) *une base de l'espace des solutions de l'équation homogène*

$$y''(t) + c_1(t)y'(t) + c_0(t)y(t) = 0. \tag{H_2}$$

L'ensemble des solutions de l'équation

$$y''(t) + c_1(t)y'(t) + c_0(t)y(t) = b(t) \tag{E_2}$$

est de la forme

$$\left\{ y(t) = \left(\mu_1^0 + \int_0^t x_1(s)ds\right) f_1(t) + \left(\mu_2^0 + \int_0^t x_2(s)ds\right) f_2(t) \quad (\mu_i^0)_{1\leq i\leq 2} \in \mathbb{K}^2, \right\}$$

où pour tout $s \in \mathbb{R}^+$*,* $(x_1(s), x_2(s))^T$ *est la solution du système*

$$\begin{pmatrix} f_1(s) & f_2(s) \\ f_1'(s) & f_2'(s) \end{pmatrix} \begin{pmatrix} x_1(s) \\ x_2(s) \end{pmatrix} = \begin{pmatrix} 0 \\ b(s) \end{pmatrix}.$$

Cas où on connait une seule solution de (H_2) Supposons que l'on connaisse une solution f_1 de l'équation homogène (H_2). On peut alors chercher une autre solution f_2 indépendante de f_1 de l'équation (H_2) (ce qui permet ensuite d'utiliser la méthode de variation de la constante pour résoudre (E_2)) ou directement une solution particulière de (E_2) de la façon suivante. On cherche f_2 sous la forme $f_2(t) = f_1(t)z(t)$. On montre alors la Proposition suivante.

Proposition 7.19 *Soit* f_1 *une solution de l'équation homogène*

$$y''(t) + c_1(t)y'(t) + c_0(t)y(t) = 0. \tag{H_2}$$

Alors $f_2 = f_1 z$ est solution de l'équation

$$y''(t) + c_1(t)y'(t) + c_0(t)y(t) = b(t). \tag{E_2}$$

si et seulement si z est solution de l'équation différentielle

$$f_1(t)z''(t) + (2\, f_1'(t) + c_1(t)f_1(t))\, z'(t) = b(t).$$

On obtient alors une équation différentielle linéaire du premier ordre sur z' que l'on résout.

Un cas particulier : l'équation d'Euler On appelle équation d'Euler l'équation différentielle linéaire à coefficients non constants

$$at^2\, y''(t) + bt\, y'(t) + cy(t) = 0.$$

Comme $a(t) = at^2$ s'annule en 0, on résout cette équation sur $]-\infty, 0[$ et sur $]0, +\infty[$. Présentons le cas de la résolution sur $]0, +\infty[$. On se ramène en fait à une équation différentielle linéaire du second ordre à coefficients constants grâce au changement de variable $t = e^u$ (on ferait $t = -e^u$ sur $]-\infty, 0[$), et le changement de fonction inconnue

$$z(u) = y(e^u).$$

La fonction z vérifie alors une équation différentielle linéaire à coefficients constants

$$az''(u) + (b-a)z'(u) + cz(u) = 0.$$

7.4 Avatars d'équations linéaires

On s'intéresse ici à quelques exemples d'équations non linéaires qui peuvent en fait se ramener à la résolution d'une équation linéaire du premier ou second ordre.

7.4.1 Équation de Bernouilli

On appelle équation de Bernouilli une équation différentielle du type :

$$x'(t) + b(t)x(t) + c(t)x^{\alpha}(t) = 0, \quad \alpha \in \mathbb{R},\, \alpha \neq 1.$$

Il suffit alors d'effectuer le changement de fonction inconnue $y(t) = x(t)^{1-\alpha}$ pour obtenir une équation différentielle linéaire vérifiée par y :

$$y'(t) + (1-\alpha)b(t)\, y(t) = -(1-\alpha)c(t).$$

7.4.2 Équation de Ricatti

On appelle équation de Ricatti une équation différentielle du type

$$z'(t) + b(t)z(t) + c(t)z^2(t) = f(t). \tag{R}$$

Supposons que l'on connaisse une solution particulière z_p de l'équation (R). En posant

$$x(t) = z(t) - z_p(t),$$

on voit que z est solution de (R) si et seulement x est solution de l'équation différentielle

$$x'(t) + [2c(t)z_p(t) + b(t)]\, x(t) + c(t)x^2(t) = 0.$$

On est ramené au cas d'une équation de Bernouilli avec $\alpha = 2$. On effectue donc ensuite le changement de fonction inconnue $y(t) = x(t)^{-1}$.