Faculté des sciences et techniques
DEUG MIAS – Math2
2002-2003

Cours d'algèbre

Table des matières

L'algèbre linéaire est présente dans tous les domaines des mathématiques (calcul différentiel, calcul intégral, théorie des nombres, équations différentielles, géométrie,...) et de ses applications (physique, analyse des données, modélisation,...).

La structure linéaire repose sur la définition d'ensemble munis de structure supplémentaire (les *espaces vectoriels*), qui sont mis en correspondance par des applications compatibles avec leur caractère vectoriel (les *applications linéaires*). La mise en œuvre de calculs linéaires donne lieu aux *matrices* et au calcul matriciel. Le problème particulier d'inversion des applications linéaires (ou, en termes matriciels, des systèmes linéaires) est résolu (partiellement) par le calcul des *déterminants*. Tels sont brièvement présentés les quatre chapitres de ce cours, suivis d'un appendice[1] qui en reprend quelques éléments en prenant un point de vue aisément qualifiable d'*algorithmique* et où les concepts précédemment développés donnent un éclairage utile.

En évitant l'abstraction, il convient de faire le choix explicite des nombres à la base des calculs linéaires, les *scalaires*. S'inspirant de la géométrie élémentaire (source de l'intuition géométrique de certains concepts linéaires et motivation à ces développements), un choix possible est celui de l'ensemble des nombres réels $\mathbb{R}$. Un autre choix (guidé par exemple par les équations différentielles de l'électricité avec les solutions complexes oscillantes $e^{i\omega t}$) est celui de l'ensemble des nombres complexes $\mathbb{C}$. C'est ce dernier choix qui a été la règle dans les notes présentes (au contraire de l'exposé oral qui se limitera aux espaces réels), à l'exception de l'appendice : le lecteur, incité à en faire une lecture *réelle*, n'y verra pas beaucoup de différence, tout en gagnant plus de proximité avec le monde linéaire (la lecture *complexe* de l'appendice achèvera de le convaincre du rôle véritable joué par le type de scalaires choisi).

Quelques courts éléments biographiques ont été introduits lorsqu'apparaît au détour d'une définition ou d'un exemple un mathématicien du passé. Ils témoignent du caractère ancien des préoccupations linéaires liées de manière naturelle à l'étude de systèmes d'équations linéaires, et qui aboutissent aux notions plus structurelles formalisées dans les années 1930 : la présentation de ces notes (à l'inverse du développement historique de la *linéarité*) s'est en fait imposée dans la seconde moitié du XXe siècle. Ces bribes de lecture historique inviteront le lecteur à d'autres lectures, tant les ouvrages d'algèbre linéaire remplissent toute bibliothèque (quelques références sont données en clôture, où exposés et exercices sont pareillement recommandables), sans parler de l'intervention récurrente dans toutes les mathématiques des quelques concepts linéaires introduits ici (le cours d'analyse qui suit énonce et prouve certaines propriétés de linéarité, réflexion faite déjà non triviales).

Aucune figure n'accompagne ces notes : le cours oral palliera à ce manque, rappelant l'origine géométrique des structures linéaires.

Laurent Guillopé
Nantes, le 1er janvier 2003

Quelques coquilles ont été corrigées à l'issue du cours oral de début 2003. Ce texte est disponible sur `www.math.sciences.univ-nantes.fr/~guillope/MIAS_Math2`. Merci de signaler toute remarque à `laurent.guillope@math.univ-nantes.fr`. Nantes, le 15 décembre 2003

[1] Cet appendice a été rédigé par X. Saint Raymond.

1. Espaces vectoriels

Généralisant des objets rencontrées précédemment (vecteurs du plan ou de l'espace, qui sont additionnés ou multipliés par un réel), la notion d'*espace vectoriel* offre un cadre conceptuel efficace, aux incarnations multiples. Ce cadre permet l'étude des phénomènes en première approximation (*linéaire* : une droite approche une courbe, un plan une surface) et est à la base de l'étude algébrico-géométrique des espaces à n dimensions.

1.1. Structure vectorielle

Les définitions générales (1.1 et 1.2) s'appliquent à des exemples connus

Exemples 1.1

1. le plan $\overrightarrow{E}_2$ (resp. l'espace $\overrightarrow{E}_3$) de la géométrie élémentaire vu comme ensemble de vecteurs $\overrightarrow{v}$;
2. l'ensemble $\mathcal{M}_{m,n}(\mathbb{C})$ des matrices à m lignes et n colonnes à coefficients complexes ;
3. l'ensemble $\mathbb{R}[X]$ des polynômes à coefficients réels.

Définition 1.1 *Un* espace vectoriel sur $\mathbb{C}$ *est un ensemble non vide E muni de deux opérations, à savoir une addition*

$$(e, f) \in E^2 \to e + f \in E$$

et une multiplication

$$(\alpha, e) \in \mathbb{C} \times E \to \alpha \cdot e \in E$$

vérifiant

(A1) Pour tout e, f, g dans E, $e + (f + g) = (e + f) + g$;

(A2) il existe un élément 0_E tel que pour tout e dans E, $e + 0_E = 0_E + e = e$, l'élément 0_E est dit vecteur nul de E *;*

(A3) pour tout $e \in E$, il existe un élément de E, appelé opposé de e *et noté $-e$ tel que $e + (-e) = (-e) + e = 0_E$;*

(A4) pour tout e, f de E, $e + f = f + e$;

(B) pour tout e, f de E et tout α, β de $\mathbb{C}$,

(B1) $\alpha \cdot (\beta \cdot e) = (\alpha\beta) \cdot e$,

(B2) $\alpha \cdot (e + f) = (\alpha \cdot e) + (\alpha \cdot f)$,

(B3) $(\alpha + \beta) \cdot e = (\alpha \cdot e) + (\beta \cdot e)$,

(B4) $1 \cdot e = e$.

Un élément de E est appelé vecteur *et un élément de $\mathbb{C}$ un* scalaire.

$\triangle$ Les axiomes (A1-4) font de $(E, +)$ un groupe commutatif (ou abélien[2]) : (A1) confère à l'addition de vecteurs la propriété d'associativité (qui permet de ne pas se soucier de parenthésages dans la somme de vecteurs $e_1 + e_2 + \ldots + e_p$), alors que (A4) énonce la commutativité de l'addition.

La notation $e - f$ signifie $e + (-f)$. On peut simplifier $e + g = f + g$ en $e = f$ en ajoutant à chaque membre de l'égalité $-g$ et en utilisant (A1) (associativité), (A3) (opposé) et (A2) (élément neutre). $\triangledown$

La notion d'*espace vectoriel sur* $\mathbb{C}$ est un cas particulier de la notion d'espace vectoriel sur un *ensemble de scalaires* $\mathbb{S}$ ayant la structure de corps. Ainsi, si $\mathbb{S} = \mathbb{R}$, on parlera d'espace vectoriel sur $\mathbb{R}$.

[2]Niels Henrik Abel, 5 août 1802, Frindoe, Norvège – 6 avril 1829, Froland, Norvège.

Définition 1.2 *L'ensemble non vide E est un* espace vectoriel sur $\mathbb{R}$ *s'il est muni d'une addition et d'une multiplication par un réel, opérations vérifiant les axiomes (A1-4) et (B1-4) de la définition 1.1 où α et β sont des nombres réels quelconques*

$\triangle$ Dans la suite, et sauf indication contraire explicite (comme lorsque la notion d'aire sera introduite comme vision géométrique du déterminant), toutes les notions introduites, et les résultats établis ci-dessous, pour les espaces vectoriels complexes (*i. e.* sur le corps des scalaires $\mathbb{S} = \mathbb{C}$) se transposent immédiatement aux espaces vectoriels réels : la raison en est que $\mathbb{R}$ et $\mathbb{C}$ sont munis d'une addition et d'une multiplication aux propriétés algébriques similaires (celles qui définissent une structure de *corps*, en particulier, tout scalaire, réel ou complexe, non nul est inversible). Un autre cas est fourni par $\mathbb{S} = \mathbb{Q}$ (et les espaces vectoriels sur $\mathbb{Q}$). À l'opposé, l'ensemble des entiers naturels $\mathbb{Z}$ ne peut être pris comme ensemble de scalaires : cela provient de la non inversibilité pour le produit de $\mathbb{Z}$ des entiers relatifs non nuls différents de 1 et -1. $\triangledown$

Exemples 1.2

1. Soit n un entier non nul. L'ensemble

$$\mathbb{C}^n = \{x = (x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{C}\}$$

muni des opérations

$$x + y = (x_1 + y_1, \dots, x_n + y_n), \quad \alpha \cdot x = (\alpha x_1, \dots, \alpha x_n)$$

est un espace vectoriel sur $\mathbb{C}$. On vérifie $0_{\mathbb{C}^n} = (0, \dots, 0)$ et $-x = (-x_1, \dots, -x_n)$.

2. Soit A un ensemble non vide et E un espace vectoriel. L'ensemble des applications

$$\mathcal{F}(A, E) = \{f \text{ application de } A \text{ dans } E\}$$

avec les opérations

$$(f + g)(a) = f(a) + g(a), \quad (\alpha f)(a) = \alpha f(a), \quad a \in A$$

est un espace vectoriel. On vérifie que l'application $0_{\mathcal{F}(A,E)}$ définie par $0_{\mathcal{F}(A,E)}(a) = 0_E, a \in A$ convient comme neutre et que l'opposé $-f$ de $f \in \mathcal{F}(A, E)$ est défini par $(-f)(a) = -f(a), a \in A$.
Comme cas particulier lorsque $E = \mathbb{C}$, on a, pour $A = I$ intervalle de $\mathbb{R}$, l'espace $\mathcal{F}(I, \mathbb{C})$ des applications de I dans $\mathbb{C}$ et pour $A = \mathbb{N}$, l'ensemble $\mathcal{F}(\mathbb{N}, \mathbb{C})$ des suites $u = (u_0, u_1, u_2, \dots)$ à valeurs scalaires. L'espace $\mathbb{C}^n$ est de ce type avec A l'intervalle d'entiers $[1, n]$: l'espace $\mathcal{F}(A, E)$ est noté parfois E^A (ce qui donne $\mathcal{F}(\mathbb{N}, \mathbb{C}) = \mathbb{C}^{\mathbb{N}}$).

Proposition 1.1 *Soit E un $\mathbb{C}$ espace vectoriel. Alors pour $e \in E$ et $\alpha \in \mathbb{C}$*

(i) $\alpha \cdot 0_E = 0_E$,
(ii) $0 \cdot e = 0_E$,
(iii) $(-\alpha) \cdot e = \alpha \cdot (-e) = -(\alpha \cdot e)$,
(iv) *si $\alpha \cdot e = 0_E$, alors $\alpha = 0$ ou $e = 0_E$.*

Preuve :

(i) $\alpha \cdot 0_E = \alpha \cdot (0_E + 0_E) = \alpha \cdot 0_E + \alpha \cdot 0_E$, soit $0_E = \alpha \cdot 0_E$.
(ii) $0 \cdot e = (0 + 0) \cdot e = 0 \cdot e + 0 \cdot e$ soit $0_E = 0 \cdot e$

(iii) $\alpha \cdot e + (-\alpha) \cdot e = (\alpha + (-\alpha)) \cdot e = 0 \cdot e = 0_E$ soit $(-\alpha) \cdot e = -(\alpha \cdot e)$
$\alpha \cdot (-e) + \alpha \cdot e = \alpha \cdot (-e + e) = \alpha \cdot 0_E = 0_E$ soit $\alpha \cdot (-e) = -(\alpha \cdot e)$

(iv) Si $\alpha \neq 0$ et $\alpha \cdot a = 0_E$, alors $0_E = \frac{1}{\alpha} \cdot 0_E = \frac{1}{\alpha} \cdot (\alpha \cdot e) = (\frac{1}{\alpha}\alpha) \cdot e = 1 \cdot e = e$.

□

Définition 1.3 *Soit* $e_1, \ldots, e_n$ *des vecteurs d'un espace vectoriel* E. *On appelle* combinaison linéaire *de* $e_1, \ldots, e_n$ *tout vecteur* e *de* E *qui s'écrit*

$$e = x_1 e_1 + x_2 e_2 \ldots + x_n e_n$$

avec les $x_1, x_2 \ldots, x_n$ *des scalaires.*

Exemples 1.3

1. Soit x un vecteur de $\mathbb{C}^n$. Il s'écrit

$$\begin{aligned} x &= (x_1, \ldots, x_n) \\ &= x_1(1, 0, \ldots, 0) + x_2(0, 1, 0, \ldots, 0) + \ldots + x_n(0, 0, \ldots, 0, 1). \end{aligned}$$

   Ainsi tout vecteur x de $\mathbb{C}^n$ est combinaison linéaire des $e_1, e_2, \ldots, e_n$ où e_i note le vecteur de $\mathbb{C}^n$ dont la seule coordonnée non nulle est la i-ème, qui vaut 1.

2. Soient $a_1 = (\alpha_{11}, \ldots, \alpha_{n1}), \ldots, a_p = (\alpha_{1p}, \ldots, \alpha_{np})$ p vecteurs de $\mathbb{C}^n$. Alors $b = (\beta_1, \ldots, \beta_n)$ est combinaison linéaire des $a_1, \ldots, a_p$ s'il existe des scalaires $(x_1, \ldots, x_p)$ tels que $b = x_1 a_1 + \ldots + x_p a_p$, c'est à dire

$$\begin{aligned} \beta_1 &= \alpha_{11} x_1 + \ldots + \alpha_{1p} x_p \\ \beta_2 &= \alpha_{21} x_1 + \ldots + \alpha_{2p} x_p \\ &\vdots \\ \beta_n &= \alpha_{n1} x_1 + \ldots + \alpha_{np} x_p \end{aligned}$$

   Cet exemple est à l'origine de la théorie *géométrique* (ou *vectorielle*) des systèmes d'équations linéaires.

Proposition/Définition 1.1 *Soient* E *et* F *des espaces vectoriels. Le produit cartésien* $E \times F = \{(e, f) : e \in E, f \in F\}$ *muni de l'addition*

$$(e, f) + (e', f') = (e + e', f + f'), \quad e, e' \in E, f, f' \in F$$

et de la multiplication par un scalaire

$$\alpha \cdot (e, f) = (\alpha \cdot e, \alpha \cdot f), \quad e \in E, f \in F, \alpha \in \mathbb{C}$$

est un espace vectoriel, appelé espace vectoriel produit de E et F.

Preuve : On a $0_{E \times F} = (0_E, 0_F)$ et $-(e, f) = (-e, -f)$. Il est aisé de vérifier explicitement les propriétés (A1-4, B1-4). □

△ On définit de manière analogue le produit $E_1 \times \ldots \times E_n$ de n espaces vectoriels $E_1, \ldots, E_n$. On retrouve $\mathbb{C}^n = \mathbb{C} \times \ldots \times \mathbb{C}$ (n facteurs). ▽

1.2. Sous-espaces vectoriels

Définition 1.4 *Une partie F de l'espace vectoriel E est appelée un* sous-espace vectoriel *de E si F vérifie*

(i) la partie F est non vide,
(ii) pour tous vecteurs e, f de F, la somme $e + f$ est dans F,
(iii) pour tout e de F et α de $\mathbb{C}$, le vecteur $\alpha \cdot e$ est dans F.

Ainsi, un sous-espace vectoriel F, partie non vide de E, est lui-même un espace vectoriel, avec ses opérations obtenues par restriction de celles de E.

Exemples 1.4

1. Si E est un espace vectoriel, les parties $\{0_E\}$ et E en sont des sous-espaces vectoriels.
2. Les parties $F = \{(x, 0, 0, 0) : x \in \mathbb{C}\}$, $G = \{(x, y, z, t) \in \mathbb{C}^4 : x+2x-\pi z+\sin(1)\, t = 0\}$ sont des sous-espaces vectoriels de $\mathbb{C}^4$.
3. Si E est un espace vectoriel, les parties $\mathcal{P}_E = \{f \in \mathcal{F}(\mathbb{R}, E) : f(-x) = f(x), x \in \mathbb{R}\}$ et $\mathcal{I}_E = \{f \in \mathcal{F}(\mathbb{R}, E) : f(-x) = -f(x), x \in \mathbb{R}\}$ sont des sous-espaces vectoriels de l'espace vectoriel $\mathcal{F}(\mathbb{R}, E)$.
4. L'ensemble $\mathbb{C}_n[X]$ des polynômes de degré au plus n est un sous-espace vectoriel de l'espace vectoriel des polynômes $\mathbb{C}[X]$.

Proposition 1.2 *Soit F une partie d'un espace vectoriel E. La partie F est un sous-espace vectoriel de E si et seulement si*

1. *la partie F contient le vecteur nul 0_E,*
2. *pour tout e, f de F et α, β de $\mathbb{C}$, le vecteur $\alpha e + \beta f$ est dans F.*

Preuve : Soit F sous-espace vectoriel de E. La partie F est non vide, donc si f est un de ses éléments, son opposé $-f = (-1) \cdot f$ est dans F d'après (iii), ainsi que $0_E = f + (-f)$ d'après (ii) : le vecteur nul 0_E est bien dans F. Si α, β sont des scalaires, e, f des vecteurs de F, alors les vecteurs αe et βf sont dans F d'après la propriété (iii) et leur somme aussi d'après (ii) : F est bien stable par combinaison linéaire $\alpha e + \beta f$.

Réciproquement, soit F une partie de E vérifiant les propriétés 1. et 2. Vu que le vecteur nul 0_E est dans F, la partie F est non vide. Par ailleurs, en prenant $\alpha = \beta = 1$, puis $\beta = 0$, les propriétés (ii) et (iii) résultent de la stabilité par combinaison linéaire. La partie F est bien un sous-espace vectoriel de E. □

Proposition/Définition 1.2 *Soient $e_1 \ldots, e_p$ des vecteurs de E. L'ensemble des combinaisons linéaires de $e_1, \ldots, e_p$ est un sous-espace vectoriel, noté* $\operatorname{Vect}(e_1, \ldots, e_p)$ *et appelé* sous-espace vectoriel engendré par $(e_1, \ldots, e_p)$.

Preuve : Le vecteur nul 0_E est une combinaison linéaire des $e_1, \ldots, e_p$:

$$0_E = 0 \cdot e_1 + \ldots + 0 \cdot e_p \,.$$

Ainsi le vecteur nul 0_E appartient à $\operatorname{Vect}(e_1, \ldots, e_p)$.

Par ailleurs, étant donné les deux combinaisons linéaires $x = x_1 e_1 + \ldots + x_p e_p, y = y_1 e_1 + \ldots + y_p e_p$, la combinaison linéaire $\alpha x + \beta y$ est combinaison linéaire des $e_1, \ldots, e_p$:

$$\alpha x + \beta y = (\alpha x_1 + \beta y_1) e_1 + \ldots + (\alpha x_p + \beta y_p) e_p \,.$$

La proposition résulte de la proposition 1.2. $\square$

$\triangle$ Le sous-espace $\mathrm{Vect}(e)$ est égal $\mathbb{C}e = \{\alpha e : \alpha \in \mathbb{C}\}$. Si e est non nul, c'est la *droite vectorielle* engendrée par e ; si $e = 0$, c'est le sous-espace $\{0_E\}$.

Si e est combinaison linéaire des $e_1, \ldots, e_p$, le sous-espace $\mathrm{Vect}(e, e_1, \ldots, e_p)$ coïncide avec le sous-espace $\mathrm{Vect}(e_1, \ldots, e_p)$. En effet, si $e = \alpha_1 e_1 + \ldots + \alpha_p e_p$,

$$\lambda e + \beta_1 e_1 + \ldots + \beta_p e_p = (\lambda\alpha_1 + \beta_1)e_1 + \ldots + (\lambda\alpha_p + \beta_p)e_p\,. \qquad \triangledown$$

Le sous-espace $\mathrm{Vect}(e_1, \ldots, e_p)$ est le plus petit (pour la relation d'inclusion) des sous-espaces vectoriels contenant les vecteurs $e_1, \ldots, e_p$ au sens suivant

Proposition 1.3 *Soit $F = \mathrm{Vect}(e_1, \ldots, e_p)$. Alors, pour $i = 1, \ldots, p$, le vecteur e_i appartient à F. Si G est un sous-espace vectoriel contenant les $e_i, i = 1, \ldots, p$, alors $G \supset F$.*

Preuve : Le vecteur e_k est combinaison linéaire $\sum_{i=1}^p \alpha_i e_i$ en prenant tous les scalaires α_i nuls, sauf celui d'indice k égal à 1.

Soit G un sous-espace contenant tous les e_i. Tout vecteur v de $F = \mathrm{Vect}(e_1, \ldots, e_p)$ est de la forme $v = \alpha_1 e_1 + \ldots + \alpha_p e_p$ d'après la Prop./Déf. 1.2 : le vecteur v appartient à G (qui est un sous-espace vectoriel) et donc $F \subset G$. $\square$

Proposition 1.4 *Si F et G sont des sous-espaces vectoriels de l'espace vectoriel E, alors l'intersection $F \cap G$ est un sous-espace vectoriel de E.*

Preuve : D'une part, par hypothèse le vecteur nul 0_E appartient aux sous-espaces F et G, et donc à leur intersection. D'autre part, si $e_1, e_2 \in F \cap G$ et α_1, α_2 sont des scalaires, la combinaison linéaire $\alpha_1 e_2 + \alpha_2 e_2$ appartient à F et à G et donc à leur intersection. Ainsi $F \cap G$ est un sous-espace vectoriel d'après la proposition 1.2. $\square$

Proposition/Définition 1.3 *Soient F et G des sous-espaces vectoriels de l'espace vectoriel E. La partie $F + G$ définie par*

$$F + G = \{f + g : f \in F, g \in G\}$$

est un sous-espace vectoriel de E, appelé somme de F et de G.

Preuve : On a $0_E = 0_F + 0_G$ (ces trois vecteurs nuls coïncident !) et donc 0_E est dans la partie $F + G$. Par ailleurs, soit $\alpha_1 e_1 + \alpha_2 e_2$ une combinaison linéaire avec e_1 et e_2 dans $F + G$. Le vecteur e_i $(i = 1, 2)$ est de la forme $e_i = f_i + g_i$ avec $f_i \in F$ et $g_i \in G$. Ainsi

$$\alpha_1 e_1 + \alpha_2 e_2 = \alpha_1(f_1 + g_1) + \alpha_2(f_2 + g_2) = (\alpha_1 f_1 + \alpha_2 f_2) + (\alpha_1 g_1 + \alpha_2 g_2)$$

est un vecteur de la partie $F + G$. Ainsi, d'après la proposition 1.2, $F + G$ est un sous-espace vectoriel de E. $\square$

Exemple 1.1 Soit pour α complexe le sous-espace $F_n(\alpha) = \{P \in \mathbb{C}_n[X], P(\alpha) = 0\}$. De l'écriture

$$P = \left(\frac{P - P(0)}{X} + P(0)\right) X + P(0)(1 - X)\,.$$

résulte $\mathbb{C}_n[X] = F_n(0) + F_n(1)$ pour $n \geq 1$, ainsi que la somme $\mathbb{C}[X] = F(0) + F(1)$ où $F(\alpha) = \{P \in \mathbb{C}[X], P(\alpha) = 0\}$.

$\triangle$ Outre l'inclusion de $F \cap G$ dans F et G, on a donc $F \subset F+G$ et $G \subset F+G$. En général, il n'est pas vrai que l'union $F \cup G$ soit un sous-espace vectoriel.

On généralise à p sous-espaces $F_1, \ldots, F_p$, en introduisant la somme

$$F_1 + \ldots + F_p = \{f_1 + \ldots + f_p : f_i \in F_i, i = 1, \ldots, p\}.$$

On remarquera que $\mathrm{Vect}(e_1, \ldots, e_p) = \mathrm{Vect}(e_1) + \ldots + \mathrm{Vect}(e_p) = \mathbb{C}e_1 + \ldots + \mathbb{C}e_p$. $\triangledown$

Définition 1.5 *Soient F et G des sous-espaces vectoriels de E. Il est dit que E est* somme directe *de F et G, ou que F et G sont des* sous-espaces supplémentaires de E, *si*

(i) $F \cap G = \{0_E\}$,
(ii) $F + G = E$.

On notera $E = F \oplus G$ si F et G sont des sous-espaces de E en somme directe.

$\triangle$ Si F et G sont des sous-espaces tels que $F \cap G = \{0_E\}$, alors, si $S = F + G$, F et G sont des sous-espaces supplémentaires de l'espace vectoriel S. $\triangledown$

Exemple 1.2 Soient, dans l'espace vectoriel $\mathbb{C}^3$, les sous-espaces

$$F = \mathbb{C}(1,0,0), \quad G = \{(0, y, \sqrt{2}\, z) : y, z \in \mathbb{C}\}, \quad H = \{(x,y,z) \in \mathbb{C}^3 : y + z = 0\}.$$

On a $\mathbb{C}^3 = F \oplus G$ vu

$$(x,y,z) = x(1,0,0) + (0, y, \sqrt{2}\frac{z}{\sqrt{2}}),$$

mais les sous-espaces F et H ne sont pas supplémentaires : $F \cap H = F \neq \{0_{\mathbb{C}^3}\}$.

Reprenant l'exemple 1.2, on a $\mathcal{F}(\mathbb{R}, E) = \mathcal{P}_E \oplus \mathcal{I}_E$. En effet, une fonction f de l'intersection $\mathcal{P}_E \cap \mathcal{I}_E$ vérifie $f(x) = f(-x) = -f(x)$ soit $2f(x) = 0$ et donc $f = 0$. Par ailleurs f s'écrit comme somme $f = f_p + f_i$ avec

$$f_p(x) = \frac{f(x) + f(-x)}{2}, \quad f_i(x) = \frac{f(x) - f(-x)}{2}, \quad x \in \mathbb{R}$$

où $f_p \in \mathcal{P}_E$ et $f_i \in \mathcal{I}_E$.

Proposition 1.5 *Soient F et G des sous-espaces de l'espace vectoriel E. Alors $E = F \oplus G$ si et seulement si tout vecteur e s'écrit sous la forme $e = f + g$ avec $f \in F$ et $g \in G$, f et g étant définis de manière unique.*

Preuve : Supposons $E = F \oplus G$. Soit e un vecteur de E. Vu que $E = F + G$, alors il existe $f \in F$ et $g \in G$ tels que $e = f + g$. Supposons une autre décomposition $e = f' + g'$, avec $f' \in F, g' \in G$. Alors $f - f' = g' - g$, vecteur de E à la fois dans F et dans G, et donc égal à 0_E (unique vecteur de $F \cap G$). Ainsi $f = f'$, $g = g'$ et l'écriture $e = f + g$ est donc unique.

Réciproquement, supposons l'existence et l'unicité de l'écriture $e = f + g$ avec $f \in F, g \in G$ pour tout vecteur e de E. L'existence de cette écriture assure $E = F + G$. En outre, pour $e \in F \cap G$, on a $e = e + 0_G = 0_F + e$ et par unicité, il résulte $0_F = e = 0_G$ et donc $F \cap G = \{0_E\}$. Les sous-espaces F et G sont supplémentaires dans E. $\square$

$\triangle$ Les sous-espaces $F_1, \ldots, F_p$ sont dits en somme directe, et on écrit $F_1 \oplus \ldots \oplus F_p$, si tout vecteur e de E admet une décomposition unique $e = f_1 + \ldots + f_p$, avec f_i vecteur de F_i $(i = 1, \ldots, p)$. $\triangledown$

1.3. Familles libre, liée, génératrice et base

Par *famille* $\mathbf{f}$ d'éléments d'un ensemble A, on entend une énumération d'éléments de A, chacun de ces éléments étant étiqueté par un élément i d'un ensemble d'indices I. Souvent, I est l'ensemble des entiers de 1 à p et on écrit $\mathbf{f} = (a_1, a_2, \dots, a_p)$ (il peut avoir égalité entre des éléments de la famille $\mathbf{f}$, ce qui n'est pas le cas dans la description *ensembliste* d'une partie $\{\alpha_1, \dots, \alpha_q\}$) ; si $I = \mathbb{N}$, on écrit $\mathbf{f} = (a_0, a_1, a_2, \dots)$ et en général, on écrit $\mathbf{f} = (a_i)_{i \in I}$. Une sous-famille $\widetilde{\mathbf{f}}$ de $\mathbf{f} = (a_i)_{i \in I}$ est déterminée par une partie $\widetilde{I}$ de I : c'est la famille des éléments de $\mathbf{f}$ indexés par $\widetilde{I}$.

Définition 1.6

(i) La famille $\mathbf{e} = (e_1, \dots, e_p)$ *de vecteurs de l'espace vectoriel* E *est dite* liée *s'il existe des scalaires* $\alpha_1, \dots, \alpha_p$ *non tous nuls tels que* $\alpha_1 e_1 + \ldots + \alpha_p e_p = 0_E$. *La famille infinie* $\mathbf{e}$ *de vecteurs* E *est dite* liée *s'il elle admet une sous famille finie liée. Les vecteurs d'une famille liée sont dits* linéairement dépendants.

Deux vecteurs e_1, e_2 *sont dits* colinéaires *s'ils sont linéairement dépendants.*

(ii) Si la famille $\mathbf{e}$ *n'est pas liée, elle est dite* libre *et les vecteurs de la famille* $\mathbf{e}$ *sont dits* linéairement indépendants.

△ Si l'un des vecteurs, soit e_i, de la famille $\mathbf{e} = (e_1, \dots, e_p)$ est nul, alors cette famille est liée. Il suffit de prendre tous les $\alpha_1, \dots, \alpha_p$ nuls, à l'exception de α_i pris égal à 1 pour représenter $0_E = \sum_{j=1}^{p} \alpha_j e_j$.

La famille (e, e) est liée (il suffit de prendre $\alpha_1 = 1 = -\alpha_2$) ; toute sur-famille d'une famille liée est liée. Ainsi si deux vecteurs, soit e_i, e_j (avec $i < j$) de la famille $(e_1, \dots, e_p)$ sont égaux, la famille est liée.

La famille réduite à un vecteur (e) est libre si et seulement si e est non nul.

Si la famille $(e_1, \dots, e_p)$ est libre, alors toute égalité

$$\alpha_1 e_1 + \ldots + \alpha_p e_p = 0_E$$

implique la nullité de tous les α_i, et cette propriété est caractéristique : si pour la famille $\mathbf{e} = (e_1, \dots, e_p)$, la nullité de toute combinaison linéaire $\sum_{i=1}^{p} \alpha_i$ implique celle de tous les scalaires $\alpha_1, \dots, \alpha_p$, alors la famille $\mathbf{e}$ est libre. ▽

Proposition 1.6 *Soit une famille* $\mathbf{e} = (e_1, \dots, e_p)$, *avec* $p \geq 2$, *de vecteurs d'un espace vectoriel* E. *La famille est liée si et seulement si l'un des vecteurs est combinaison linéaire des autres.*

Preuve : Supposons la famille liée. Il existe une combinaison linéaire nulle

$$\alpha_1 e_1 + \ldots + \alpha_p e_p = 0$$

avec un scalaire, soit α_i, non nul. Alors

$$\alpha_i e_i = -\alpha_1 e_1 - \ldots - \alpha_{i-1} e_{i-1} - \alpha_{i+1} e_{i+1} - \ldots - \alpha_p e_p$$

soit

$$e_i = -\frac{\alpha_1}{\alpha_i} e_1 - \ldots - \frac{\alpha_{i-1}}{\alpha_i} e_{i-1} - \frac{\alpha_{i+1}}{\alpha_i} e_{i+1} - \ldots - \frac{\alpha_p}{\alpha_i} e_p$$

et donc e_i est combinaison linéaires des autres vecteurs de la famille $\mathbf{e}$.

Inversement, si

$$e_i = \lambda_1 e_1 + \ldots + \lambda_{i-1} e_{i-1} + \lambda_{i+1} e_{i+1} + \ldots + \lambda_p e_p\,,$$

alors $\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_p e_p = 0$ avec $\alpha_i = -1$ et $\alpha_j = \lambda_j$ si $i \neq j$: la famille $\mathbf{e}$ est liée. □

Définition 1.7 *La famille* $\mathbf{e}$ *de vecteurs de l'espace vectoriel* E *est dite* génératrice de E *si tout vecteur de* E *est combinaison linéaire de vecteurs de cette famille. On dit encore que la famille* $\mathbf{e}$ engendre *l'espace* E.

Exemple 1.3 La famille $(1, X, \ldots, X^n)$ est une famille génératrice de $\mathbb{C}_n[X]$.

$\triangle$ Si la famille $\mathbf{e}$ est finie, *i. e.* $\mathbf{e} = (e_1, \ldots, e_p)$, la famille $\mathbf{e}$ engendre E si et seulement si $\mathrm{Vect}(e_1, \ldots, e_p) = E$. $\triangledown$

Définition 1.8 *Une famille* $\mathbf{b} = (b_1, \ldots, b_p)$ *est appelée* base *de l'espace* E *si elle est libre et elle engendre* E.

$\triangle$ On ne parlera ici que de base qui soit une famille finie de vecteurs. $\triangledown$

Proposition/Définition 1.4 *La famille* $\mathbf{b} = (b_1, \ldots, b_p)$ *est une base de* E *si et seulement si pour tout vecteur* e *de* E *il existe un* p*-uplet scalaire* $(\alpha_1, \ldots, \alpha_p)$ *unique tel que* $e = \alpha_1 e_1 + \ldots + \alpha_p e_p$. *Les scalaires* $\alpha_1, \ldots, \alpha_p$ *s'appellent les* coordonnées du vecteur e dans la base $\mathbf{b}$.

Preuve : Supposons que $\mathbf{b}$ soit une base. Soit e un vecteur de E. Vu que $\mathbf{b}$ engendre E, il existe des scalaires $\alpha_1, \ldots, \alpha_p$ tels que $e = \alpha_1 b_1 + \ldots + \alpha_p b_p$. Supposons une autre écriture $e = \beta_1 b_1 + \ldots + \beta_p b_p$. Alors

$$(\alpha_1 - \beta_1) b_1 + \ldots + (\alpha_p - \beta_p) b_p = 0$$

et donc $\alpha_1 = \beta_1, \ldots, \alpha_p = \beta_p$ vu que la famille $\mathbf{b}$ est libre. L'écriture $e = \alpha_1 b_1 + \ldots + \alpha_p b_p$ est donc bien unique.

Réciproquement, que tout vecteur e s'écrive $e = \alpha_1 b_1 + \ldots + \alpha_p b_p$, énonce le caractère générateur de la famille $\mathbf{b}$. Par ailleurs, soit $\alpha_1, \ldots, \alpha_p$ tels que $\alpha_1 b_1 + \ldots + \alpha_p b_p = 0_E$. L'unicité de la représentation

$$0_E = \alpha_1 b_1 + \ldots + \alpha_p b_p = 0 \cdot b_1 + \ldots + 0 \cdot b_p$$

implique $\alpha_1 = 0, \ldots, \alpha_p = 0$, c'est à dire le caractère libre de la famille $\mathbf{b}$, qui est donc une base. $\square$

Exemple 1.4 L'espace $\mathbb{C}^n$ admet pour base la famille

$$\mathbf{b} = (b_1 = (1, 0, \ldots, 0), b_2 = (0, 1, 0, \ldots, 0), \ldots, b_n = (0, 0, \ldots, 1)),$$

dite *base canonique* de $\mathbb{C}^n$.

Proposition 1.7 *Soit* F *et* G *deux sous-espaces supplémentaires de l'espace* E. *Si* $\mathbf{f} = (f_1, \ldots, f_m)$ *(resp.* $\mathbf{g} = (g_1, \ldots, g_n)$*) est une base de* F *(resp* G*), alors*

$$\mathbf{f} \cup \mathbf{g} = (f_1, \ldots, f_m, g_1, \ldots, g_n)$$

est une base de E.

Preuve : On écrit tout vecteur de e comme somme d'un vecteur de F et d'un vecteur de G, vecteurs décomposés suivant les bases $\mathbf{f}$ et $\mathbf{g}$ resp., de telle sorte que

$$e = \alpha_1 f_1 + \ldots + \alpha_m f_m + \beta_1 g_1 + \ldots + \beta_n g_n .$$

Une décomposition de ce type est unique, du fait que F et G sont en somme directe, puis que $\mathbf{f}$ et $\mathbf{g}$ sont des bases de F et G resp. $\square$

1.4. Espaces vectoriels de dimension finie

Définition 1.9 *L'espace E est dit de* dimension finie *s'il admet une famille génératrice finie.*

Exemple 1.5 Si A est finie, l'espace $\mathcal{F}(A, \mathbb{C})$ est de dimension finie, puisque toute fonction f est combinaison linéaire $f = \sum_{a \in A} f(a)\delta_a$ des fonctions δ_a définies par

$$\delta_a(a) = 1\,, \quad \delta_a(b) = 0 \text{ si } b \neq a\,.$$

$\triangle$ Il existe des $\mathbb{C}$ espaces vectoriels non de dimension finie, ainsi de l'espace des polynômes $\mathbb{C}[X]$. En effet, si $\mathbf{P} = (P_1, \ldots, P_p)$ est une famille finie de polynômes et $m_{\mathbf{P}}$ désigne le plus grand des degrés des polynômes $P_1, \ldots, P_p$, toute combinaison linéaire de ces polynômes est de degré au plus $m_{\mathbf{P}}$: un polynôme de degré $m_{\mathbf{P}} + 1$ (par ex. $X^{m_{\mathbf{P}}+1}$) n'appartient pas à l'espace $\mathrm{Vect}(P_1, \ldots, P_p)$. L'espace des polynômes n'est pas de dimension finie.

On démontrera que, si l'ensemble A est infini, l'espace $\mathcal{F}(A, \mathbb{C})$ est de dimension infinie aussi. Remarquons que le calcul vectoriel en dimension finie est le cadre axiomatique naturel de la géométrie élémentaire, alors que les espaces non de dimension finie surviennent souvent en analyse. $\triangledown$

Proposition 1.8 *Soit E un espace vectoriel non nul. On peut extraire une base de toute famille génératrice finie.*

Preuve : Soit $\mathbf{g}(p) = (e_1, \ldots, e_p)$ une famille génératrice de E. Si elle est libre, c'est une base et la proposition est démontrée. Sinon, il existe un vecteur qui est combinaison linéaire des autres et donc, si $\mathbf{g}(p-1)$ est la famille $\mathbf{g}(p)$ privée de ce vecteur, la famille $\mathbf{g}(p-1)$ est encore génératrice (car $\mathrm{Vect}(\mathbf{g}(p)) = \mathrm{Vect}(\mathbf{g}(p-1))$). Continuant, on arrivera à un entier k (au plus égal à $p-1$ et au moins égal à 1, puisque E contient des vecteurs non nuls) tel que la famille $\mathbf{g}(k)$ de cardinal k, sous-famille de $\mathbf{g}(p)$, est génératrice et dont aucun vecteur n'est combinaison linéaire des autres, *i. e.* la famille $\mathbf{g}(k)$ est libre : la famille $\mathbf{g}(k)$ est donc une base. $\square$

Proposition 1.9 *Un espace non de dimension finie admet une famille libre infinie.*

Preuve : Soit E non de dimension finie. Si $\mathbf{f}$ est une famille libre finie, alors il existe un vecteur e tel que $\mathbf{f} \cup \{e\}$ soit libre. Sinon, tout vecteur e serait combinaison linéaire des vecteurs de $\mathbf{f}$ et E, engendré par la famille finie $\mathbf{f}$, serait de dimension finie. On construit ainsi par adjonction successive d'un vecteur une famille infinie libre. $\square$

Proposition 1.10 *Soit E un espace vectoriel. Soit $\mathbf{g} = (g_1, \ldots, g_q)$ une famille génératrice de E et $\ell = (\ell_1, \ldots, \ell_p)$ une famille libre de E. Alors, $p \leq q$ et il existe une sous-famille $\mathbf{g}'$ avec $q - p$ vecteurs de la famille $\mathbf{g}$ telle que la famille $\ell \cup \mathbf{g}'$ soit génératrice.*

Preuve : Soit $\mathcal{P}_{\mathbf{g}}(k)$ la propriété de récurrence avec $k \leq q$:

> $\mathcal{P}_{\mathbf{g}}(k)$: si $e_1, \ldots, e_k$ sont linéairement indépendants, il existe $q - k$ vecteurs $(b_1, \ldots, b_{q-k})$ de la famille $\mathbf{g}$ telle que la famille $(e_1, \ldots, e_k, b_1, \ldots, b_{q-k})$ soit génératrice.

On va montrer que $\mathcal{P}_{\mathbf{g}}(1)$ est vraie, puis que, si $k < q$, la validité de $\mathcal{P}_{\mathbf{g}}(k)$ entraîne celle de $\mathcal{P}_{\mathbf{g}}(k+1)$. Remarquons que $\mathcal{P}_{\mathbf{g}}(q)$ signifie qu'une famille libre de q éléments est génératrice.

Soit donc e un vecteur non nul. Il est combinaison linéaire des vecteurs de $\mathbf{g}$,

$$e = \beta_1 g_1 + \beta_2 g_2 + \ldots + \beta_q e_q \,.$$

L'un des scalaires $\beta_1, \beta_2, \ldots, \beta_q$ est non nul, soit β_i, ainsi

$$g_i = \frac{1}{\beta_i}\left(e - \beta_1 g_1 - \ldots - \beta_{i-1} g_{i-1} - \beta_{i+1} g_{i+1} - \ldots - \beta_q g_q\right) \,.$$

et

$$\mathrm{Vect}(e, g_1, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots, g_q) = \mathrm{Vect}(g_1, g_2, \ldots, g_q) = E$$

et la famille $(e, g_1, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots, g_q)$ est génératrice, ce qui prouve $\mathcal{P}_{\mathbf{g}}(1)$.

Supposons $\mathcal{P}_{\mathbf{g}}(k)$ vérifiée et soient $e_1, \ldots, e_{k+1}$ linéairement indépendants. D'après $\mathcal{P}_{\mathbf{g}}(k)$, il existe $q-k$ vecteurs $b_1, \ldots, b_{q-k}$ de la famille $\mathbf{g}$ telle que la famille $(e_1, \ldots, e_k, b_1, \ldots, b_{q-k})$ soit génératrice. Ainsi, e_{k+1} est combinaison linéaire

$$e_{k+1} = \alpha_1 e_1 + \ldots + \alpha_k e_k + \beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_{q-k} b_{q-k} \,.$$

où les $\beta_1, \ldots, \beta_{q-k}$ ne sont pas tous nuls (sinon, la famille $(e_1, e_2 \ldots, e_k, e_{k+1})$ ne serait pas libre), soit par exemple β_i. Alors

$$b_i = \frac{1}{\beta_i}\left(e_{k+1} - \alpha_1 e_1 - \ldots - \alpha_k e_k - \beta_1 b_1 - \ldots - \beta_{i-1} b_{i-1} - \beta_{i+1} b_{i+1} - \ldots - \beta_{q-k} b_{q-k}\right) \,.$$

et

$$\mathrm{Vect}(e_{k+1}, e_1, \ldots, e_k, b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_{q-k}) = \mathrm{Vect}(e_1, \ldots, e_k, b_1, \ldots, b_{q-k}) = E$$

et donc $(e_1, \ldots, e_k, e_{k+1}, b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_{q-k})$ est génératrice, ce qui prouve $\mathcal{P}_{\mathbf{g}}(k+1)$.

Si $p \leq q$, on applique la propriété de récurrence $\mathcal{P}_{\mathbf{g}}(p)$ à la famille libre ℓ.

Si $p > q$, d'après la propriété $\mathcal{P}_{\mathbf{g}}(q)$, la famille $(\ell_1, \ldots, \ell_q)$ est génératrice. Ainsi, le vecteur ℓ_{q+1} est combinaison linéaire des $\ell_1, \ldots, \ell_q$, et la famille ℓ n'est pas libre, ce qui est contraire à l'hypothèse. Ainsi $p \leq q$. □

Cette proposition a le corollaire suivant

Corollaire 1.1 *Si l'espace vectoriel E admet une famille génératrice $\mathbf{g}$ à q vecteurs, toute famille à $q+1$ vecteurs est liée.*

△ Si A est infini, l'espace $\mathcal{F}(A, \mathbb{C})$ n'est pas de dimension finie. En effet, d'après la proposition précédente, il suffit de trouver une famille libre infinie (dont toute sous-famille non vide est libre). La famille de fonctions $(\delta_a)_{a \in A}$ avec $\delta_a(a) = 1, \delta_a(b) = 0$ si $a \neq b$ convient.

L'espace $\mathcal{C}^0(\mathbb{R}, \mathbb{C})$ des fonctions continues de $\mathbb{R}$ dans $\mathbb{C}$, est aussi non de dimension finie. Soit $(p_n)_{n \geq 1}$ la suite de fonctions définie par

$$p_n : x \in \mathbb{R} \to (x-1)(x-2) \cdot \ldots \cdot (x-n) \in \mathbb{C} \,.$$

Si on a une relation linéaire

$$\alpha_1 p_1 + \alpha_2 p_2 + \ldots + \alpha_N p_N = 0 \,,$$

(qui provient éventuellement d'une relation $\sum_{i=1}^{n} \lambda_i p_{k_i} = 0$ avec $k_1 < k_2 < \ldots < k_n$, qu'on aura complétée de manière triviale en une combinaison linéaire sur $p_1, \ldots, p_N$ avec $N = k_n$), l'évaluation en $x = 2, x = 3, \ldots, x = N+1$ donne successivement $\alpha_1 = 0$, $2\alpha_2 = 0, \ldots, i!\alpha_i = 0, \ldots, N!\alpha_N = 0$: cette combinaison linéaire a tous ses coefficients nuls. La famille $(p_n)_{n \geq 1}$ est donc une famille libre infinie de l'espace $\mathcal{C}^0(\mathbb{R}, \mathbb{C})$. ▽

Théorème/Définition 1.1 *Soit E un espace vectoriel de dimension finie, non nul. L'espace E admet une base et toute base de E a le même nombre de vecteurs.*

Ce nombre est appelé la dimension *de E. Si l'espace E est nul, il est dit de* dimension zéro *ou nulle.*

Preuve : L'existence d'une base a été vue dans la prop. 1.8. Soient $\mathbf{b}$ et $\widetilde{\mathbf{b}}$ deux bases de l'espace E, avec n et $\widetilde{n}$ éléments resp. Vu que $\mathbf{b}$ est génératrice et $\widetilde{\mathbf{b}}$ libre, on a $\widetilde{n} \leq n$. L'inégalité inverse valant pareillement, on déduit l'égalité de n et $\widetilde{n}$. □

Exemples 1.5

1. L'espace $\mathbb{C}^n$ est de dimension n , avec pour base la base canonique
$$((1,0,\ldots,0),(0,1,\ldots,0),\ldots,(0,0,\ldots,1)).$$
2. Si A est fini de cardinal $\#A$, l'espace $\mathcal{F}(A,\mathbb{C})$ est de dimension finie $\#A$. La famille δ_a (introduite dans l'exemple 1.5) en est une base.
3. L'espace $\mathbb{C}_n[X]$ des polynômes de degré au plus n est de dimension $n+1$, la famille $(1, X, X^2, \ldots, X^n)$ étant une base.

On a le théorème dit *de la base incomplète*

Théorème 1.1 (de la base incomplète) *Soit E un espace vectoriel de dimension finie. Toute famille libre $\ell = (\ell_1, \ldots, \ell_p)$ peut être complétée en une base,* i. e. *il existe $q = \dim(E) - p$ vecteurs $f_1, \ldots, f_q$ tels que la famille $(\ell_1, \ldots, \ell_p, f_1, \ldots, f_q)$ soit une base de E.*

Preuve : Soit n la dimension de E et $\mathbf{b} = (b_1, \ldots, b_n)$ une base de E. C'est en particulier une famille génératrice. Ainsi d'après la proposition 1.10, il existe $n-p$ vecteurs $(\widetilde{b}_1, \ldots, \widetilde{b}_{n-p})$ pris dans la famille $\mathbf{b}$ tels que $(\ell_1, \ldots, \ell_p, \widetilde{b}_1, \ldots, \widetilde{b}_{n-p})$ soit une partie génératrice de E. Elle est libre, sinon on en aurait une sous-famille de cardinal $\leq n-1$ qui serait une base, ce qui ne peut être puisque toute base est de cardinal n. La famille $(\ell_1, \ldots, \ell_p, \widetilde{b}_1, \ldots, \widetilde{b}_{n-p})$ est une base de E. □

Proposition 1.11 *Soit F sous-espace vectoriel d'un espace vectoriel de dimension finie. Alors F est de dimension finie et sa dimension est au plus égale à celle de E.*

Le sous-espace F est de même dimension que E si et seulement si E et F coïncident.

Preuve : Soit n la dimension de E. Si F admet une famille libre de k vecteurs, une telle famille, libre aussi comme famille de vecteurs de E, ne peut exister dès que $k > n$ d'après la proposition 1.1 puisque E, de dimension finie n, admet une famille génératrice de cardinal n. Ainsi, F est de dimension finie (puisque F ne peut avoir de famille libre de cardinal arbitrairement grand) et sa dimension (cardinal de ses bases qui sont des familles de vecteurs de E libres) est au plus égale à $n = \dim E$.

Si F est de dimension $\dim E$, une base $\mathbf{b}$ de F engendre E, sinon, complétée en une base de E elle donnerait une base de E de cardinal strictement plus grand que $\dim E$, ce qui n'est pas ; ainsi $E = F$. La réciproque est claire. □

Définition 1.10 *Un espace vectoriel de dimension* 1 *(resp. 2) est appelé* droite vectorielle *(resp.* plan vectoriel*). Dans l'espace vectoriel E de dimension finie, un sous-espace de dimension* $\dim E - 1$ *est appelé* hyperplan vectoriel.

△ Dans un plan, une droite est aussi un hyperplan, ce qui n'est pas le cas dans un espace de dimension $n \geq 3$. ▽

Proposition 1.12 *Soit F un sous-espace vectoriel d'un espace vectoriel E de dimension finie. Alors F admet des sous-espaces supplémentaires. Si S est un supplémentaire de F,* $\dim E = \dim F + \dim S$.

Preuve : Soit $\mathbf{f} = (f_1, \ldots, f_p)$ une base de F et $\mathbf{e} = (e_1, \ldots, e_q)$ une famille libre de E complétant la base $\mathbf{f}$: $\mathbf{f} \cup \mathbf{e}$ est une base de E. Alors le sous-espace $S = \mathrm{Vect}(\mathbf{e})$ engendré par les vecteurs de la famille $\mathbf{e}$ est un sous-espace supplémentaire de F. En effet, tout vecteur v de E, combinaison linéaire des vecteurs de $\mathbf{f} \cup \mathbf{e}$

$$v = \alpha_1 f_1 + \cdots + \alpha_p f_p + \beta_1 e_1 + \cdots + \alpha_q e_q$$

est somme d'un vecteur $(\alpha_1 f_1 + \cdots + \alpha_p f_p)$ de F et d'un vecteur $(\beta_1 e_1 + \cdots + \beta_q e_q)$ de S, et ceci de manière unique puisque la famille $\mathbf{f} \cup \mathbf{e}$ est libre. La famille $\mathbf{e}$ est une base de S, il en résulte la relation sur les dimensions.

Si S est un sous-espace en somme directe avec F et $\mathbf{c}$ une base de S, la famille de $\mathbf{f} \cup \mathbf{c}$ est une base de E d'après la prop. 1.7. La relation sur les dimensions en résulte. □

Théorème 1.2 *Soient E et F des espaces vectoriels de dimension finie. Alors le produit $E \times F$ est de dimension finie et*

$$\dim(E \times F) = \dim E + \dim F \, .$$

Preuve : Si $\mathbf{b} = (b_1, \ldots, b_m)$ est une base de E, $\mathbf{c} = (c_1, \ldots, c_n)$ une base de F, alors la famille $((b_1, 0_F), \ldots, (b_m, 0_F), (0_E, c_1), \ldots, (0_E, c_n))$ est une base de $E \times F$. En effet, il suffit de décomposer tout vecteur (e, f) de $E \times F$ sous la forme $(e, f) = (e, 0_F) + (0_E, f)$ (les sous-espaces $E \times \{0_F\}$ et $\{0_E\} \times F$ de $E \times F$ sont en somme directe), puis de revenir aux définitions. □

△ On peut déduire le théorème précédent de la dernière proposition en remarquant

$$E \times F = E \times \{0_F\} \oplus \{0_E\} \times F \, .$$

Par ailleurs, la définition de somme directe de deux sous-espaces se généralise à la *somme directe d'une collection finie de sous-espaces* : $F_1, \ldots, F_k$ sont en somme directe de somme F, et on écrira $F = F_1 \oplus \ldots \oplus F_k$, si $F = F_1 + \ldots + F_k$ et la décomposition d'un vecteur de $e \in F$ en la somme $e = f_1 + \ldots + f_k$ avec $f_i \in F_i, i = 1, \ldots k$ est unique. Si F est somme directe de $F_1, \ldots, F_k$ et $\mathbf{b}_i$ est une base de F_i pour $i = 1, \ldots, k$, alors $\mathbf{b}_1 \cup \mathbf{b}_2 \cup \ldots \cup \mathbf{b}_k$ est une base de F. Ainsi

$$\dim(F_1 \oplus \ldots \oplus F_k) = \dim F_1 + \ldots + \dim F_k \, .$$ ▽

Théorème 1.3 *Soient F et G sous-espaces de l'espace E de dimension finie. Alors*

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G) \, .$$

Preuve : Soit S_F (resp. S_G) un supplémentaire de $F \cap G$ dans F (resp. G). Alors

$$F + G = (F \cap G) \oplus S_F \oplus S_G \, .$$

En effet, tout vecteur $v = f + g$ de $F + G$ s'écrit de manière unique $v = u + v_F + v_G$ avec $u \in F \cap G, v_F \in S_F, v_G \in S_G$. D'une part, on a $f = f_1 + f_2, g = g_1 + g_2$ avec $f_1, g_1 \in F \cap G$,

$f_2 \in S_F$, $g_2 \in G_F$: on prendra donc $u = f_1 + g_1, v_F = f_2$ et $v_G = g_2$. D'autre part, si $v = \widetilde{u} + \widetilde{v}_F + \widetilde{v}_G$ est une autre écriture, on obtient $\widetilde{u} + \widetilde{v}_F - u - v_F = v_G - \widetilde{v}_G$, vecteur de $F \cap G$ (membre de gauche dans F, de droite dans G) et de S_G (membre de gauche), donc nul vu $G = (F \cap G) \oplus S_G$. Par suite, $v_G = \widetilde{v}_G$, puis, de manière symétrique, $v_F = \widetilde{v}_F$ et enfin $u = \widetilde{u}$, ce qui donne bien l'unicité annoncée.

On en déduit

$$\dim(F + G) = \dim(F \cap G) + \dim S_F + \dim S_G$$

alors que l'écriture des sommes directes $F = (F \cap G) \oplus S_F$ et $G = (F \cap G) \oplus S_G$ donnent les relations entre dimensions

$$\dim F = \dim(F \cap G) + \dim S_F\,, \quad \dim G = \dim(F \cap G) + \dim S_G\,.$$

L'égalité du théorème en résulte. □

1.5. Rang d'une famille de vecteurs

Définition 1.11 *Le* rang de la famille de vecteurs $\mathbf{e} = (e_1, e_2, \ldots, e_p)$ *est la dimension de l'espace* $\mathrm{Vect}(e_1, e_2, \ldots, e_p)$ *qu'elle engendre.*

△ On n'a pas supposé l'espace E auquel appartiennent les vecteurs $e_1, \ldots, e_p$ de dimension finie : ces vecteurs (en nombre fini) engendrent un sous-espace de dimension finie. De manière plus générale, on parlera du rang d'une famille quelconque $\mathbf{e} = (e_i)_{i \in I}$ dans le cas où le sous-espace $\mathrm{Vect}(\mathbf{e})$ est de dimension finie et alors le rang de $\mathbf{e}$ sera la dimension de ce sous-espace.

Si $\#\mathbf{e}$ désigne le cardinal de la famille $\mathbf{e}$, on a donc

$$\mathrm{rg}(\mathbf{e}) = \dim \mathrm{Vect}(\mathbf{e}) \leq \#\mathbf{e}\,.$$ ▽

Terminons en répétant les propriétés suivantes valides dans un espace de dimension n :

- l'entier n est le cardinal maximal d'une famille libre de E,
- n est le cardinal minimal d'une famille génératrice de E ;
- une famille libre (resp. génératrice) de cardinal n est une base.

Exemple 1.6 Soit $x_0, x_1, \ldots, x_n$ $n+1$ scalaires distincts. Alors la famille de polynômes $(P_i)_{i=0}^n$, dits polynômes d'interpolation de Lagrange[3],

$$P_i(X) = (X - x_0)(X - x_1) \ldots (X - x_{i-1})(X - x_{i+1}) \ldots (X - x_n)\,, \quad i = 0, \ldots, n$$

est libre : c'est donc une base de l'espace $\mathbb{C}_n[X]$ (qui est de dimension $n+1$).

[3] Joseph-Louis Lagrange, 25 janvier 1736, Turin, Sardaigne-Piémont (désormais Italie) – 10 avril 1813, Paris, France.

2. Applications linéaires

Une fois les espaces vectoriels introduits, il s'agit d'étudier les relations qui peuvent s'établir entre eux. Les applications qui respectent les structures vectorielles (le réceptacle de la *morphologie linéaire*) sont celles qui importent : ce sont les *applications linéaires* (parfois nommés *morphismes linéaires*).

2.1. Application linéaire

Définition 2.1 *L'application f de l'espace vectoriel E dans l'espace vectoriel F est dite* linéaire *si elle vérifie*

(i) $f(e+e') = f(e) + f(e'), e, e' \in E$,

(ii) $f(\alpha e) = \alpha f(e), \alpha \in \mathbb{C}, e \in E$.

Une application linéaire de E dans E est appelée endomorphisme *de E.*

Exemples 2.1

1. L'application identité (notée Id ou Id_E) de E est linéaire, ainsi que l'injection naturelle $e \in F \to e \in E$ si F est un sous-espace vectoriel de E.
2. Soit $\lambda \in \mathbb{C}$. L'application $e \in E \to \lambda e$ est linéaire et est appelée l'*homothétie* de *rapport* λ.
3. L'application φ de $\mathbb{C}^3$ dans $\mathbb{C}$ définie par $\varphi(x, y, z) = x + e^{\pi} y - z$ est linéaire.
4. L'application de $\mathbb{C}[X]$ dans lui même qui au polynôme P associe son polynôme dérivé P' est linéaire.
5. Soit $(u_n)_{n \geq 0}$ une suite et Du la suite obtenue par décalage des indices : $(Du)_n = u_{n+1}, n \geq 0$. L'application $u \to Du$ est une application linéaire de $\mathbb{C}^{\mathbb{N}}$ dans lui-même.
6. L'application de $\mathbb{C}^k$ dans $\mathbb{C}$ qui à $x = (x_i)$ associe $\max_{i=1}^{k} \mathrm{Ré}\, x_i$ n'est pas linéaire.

Proposition 2.1 *Soit f application linéaire de E dans F. Alors*

(i) $f(0_E) = 0_F$,

(ii) $f(-e) = -f(e), e \in E$.

Preuve : On a simplement

$$\begin{aligned} f(0_E) &= f(0 \cdot 0_E) = 0 \cdot f(0_E) = 0_F \\ f(-e) &= f((-1) \cdot e) = (-1) \cdot f(e) = -f(e) . \end{aligned}$$

□

Proposition 2.2 *Soit f une application de E dans F. L'application f est linéaire si et seulement si pour $e, e' \in E, \alpha, \alpha' \in \mathbb{C}$*

$$f(\alpha e + \alpha' e') = \alpha f(e) + \alpha' f(e') .$$

Preuve : Si f est linéaire, les propriétés (i) et (ii) de la définition 2.1 sont appliquées successivement pour avoir l'égalité de la proposition.

Réciproquement, si f vérifie l'identité de la proposition, on prendra $\alpha = \alpha' = 1$, puis $\alpha' = 0$ pour obtenir les identités de la définition 2.1. □

△ Une application linéaire f conserve les combinaisons linéaires au sens où

$$f(\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_p e_p) = \alpha_1 f(e_1) + \alpha_2 f(e_2) + \ldots \alpha_p f(e_p) ,$$

pour $e_1, \ldots, e_p \in E, \alpha_1, \ldots, \alpha_p \in \mathbb{C}$. ▽

Proposition 2.3 *Soit E espace vectoriel de dimension n, $\mathbf{b} = (b_1, \ldots, b_n)$ une base de E et $(a_1, \ldots, a_n)$ une famille de vecteurs de F. Alors, il existe une unique application linéaire $f : E \to F$ telle que $f(b_i) = a_i$ pour $i = 1, \ldots, n$.*

Preuve : La fonction linéaire cherchée vérifie nécessairement, du fait de sa linéarité,

$$f(\lambda_1 b_1 + \ldots + \lambda_n b_n) = \lambda_1 f(b_1) + \ldots + \lambda_n f(b_n) = \lambda_1 a_1 + \ldots + \lambda_n a_n$$

et donc, si elle existe, est définie de manière unique. L'égalité précédente définit une fonction de manière unique, car un vecteur e a un unique n-uplet de coordonnées dans la base $\mathbf{b}$. Cette fonction est linéaire, en effet si $e = \lambda_1 b_1 + \ldots + \lambda_n b_n$ et $e' = \lambda'_1 b_1 + \ldots + \lambda'_n b_n$ sont deux vecteurs de E, α et α' deux scalaires

$$\begin{aligned}
f(\alpha e + \alpha' e') &= f(\alpha(\lambda_1 b_1 + \ldots + \lambda_n b_n) + \alpha'(\lambda'_1 b_1 + \ldots + \lambda'_n b_n)) \\
&= f((\alpha\lambda_1 + \alpha'\lambda'_1) b_1 + \ldots + (\alpha\lambda_n + \alpha'\lambda'_n) b_n) \\
&= (\alpha\lambda_1 + \alpha'\lambda'_1) a_1 + \ldots + (\alpha\lambda_n + \alpha'\lambda'_n) a_n \\
&= \alpha(\lambda_1 a_1 + \ldots + \lambda_n a_n) + \alpha'(\lambda'_1 a_1 + \ldots + \lambda'_n a_n) \\
&= \alpha f(e) + \alpha' f(e') .
\end{aligned}$$

et qui vérifie bien

$$f(b_i) = f(0 \cdot b_1 + \ldots + 1 \cdot b_i + \cdot + \ldots + 0 \cdot e_n) = a_i$$

pour $i = 1, \ldots, n$. $\square$

Exemple 2.1 Soit $(P_0, \ldots, P_n)$ des polynômes de degré au plus n. Il existe une et une seule application linéaire f de $\mathbb{C}_n[X]$ dans lui-même tel que $f(X^i) = P_i$ pour $i = 0, \ldots, n$.

Proposition 2.4 *Soient E, F et G des espaces vectoriels. Si f est une application linéaire de E dans F et g une application de F dans G, alors $g \circ f$ est une application linéaire de E dans G.*

Preuve : Il suffit de vérifier la définition de compatibilité avec les combinaisons linéaires

$$\begin{aligned}
(g \circ f)(\alpha e + \alpha' e') &= g[f(\alpha e + \alpha' e')] \\
&= g[\alpha f(e) + \alpha' f(e')] \\
&= \alpha g[f(e)] + \alpha' g[f(e')] \\
&= \alpha (g \circ f)(e) + \alpha' (g \circ f)(e') .
\end{aligned}$$

$\square$

2.2. Isomorphie, image et noyau

Rappelons qu'une application f de source A et de but B est
– bijective si tout élément de B est image d'un et d'un seul élément de A,
– surjective si tout élément de B est image par f d'au moins un élément de A,
– injective si les images de deux éléments distincts de A sont distinctes dans B.

Définition 2.2 *Une application linéaire de E dans F bijective est appelée un* isomorphisme *(de E dans F), un isomorphisme de E dans lui-même est appelé* automorphisme.

Proposition 2.5 *Soit f un isomorphisme de E dans F. Alors l'application inverse f^{-1} est une application linéaire de F dans E.*

Preuve : Soient v, v' des vecteurs de F, α, α' des scalaires. Pour montrer

$$f^{-1}(\alpha v + \alpha' v') = \alpha f^{-1}(v) + \alpha' f^{-1}(v')\,,$$

il suffit de montrer que l'image par f des deux membres coïncident, vu l'injectivité de f. On a d'une part

$$f[f^{-1}(\alpha v + \alpha' v')] = \alpha v + \alpha' v'$$

et d'autre part, par linéarité de f,

$$\begin{aligned} f[\alpha f^{-1}(v) + \alpha' f^{-1}(v')] &= \alpha f[f^{-1}(v)] + \alpha' f[f^{-1}(v')] \\ &= \alpha v + \alpha' v'\,, \end{aligned}$$

ce qui conclut la preuve. □

Définition 2.3 *S'il existe un isomorphisme de E dans F, les espaces E et F sont dits* isomorphes*, et on notera $E \simeq F$.*

Proposition 2.6 *Les espaces E et $\widetilde{E}$ de dimension finie sont isomorphes si et seulement si leurs dimensions sont égales.*

Preuve : Supposons E et $\widetilde{E}$ isomorphes. Soit $\mathbf{b} = (b_1, \ldots, b_n)$ une base de E. Alors $f(\mathbf{b}) = (f(b_1), \ldots, f(b_n))$ est une base de $\widetilde{E}$. En effet, d'une part $f(\mathbf{b})$ est libre : si $\lambda_1 f(b_1) + \ldots + \lambda_n f(b_n) = 0_F$, alors $f(\lambda_1 b_1 + \ldots + \lambda_n b_n) = 0_F$ et par unicité de l'image inverse 0_E du vecteur nul 0_F de F, on a $\lambda_1 b_1 + \ldots + \lambda_n b_n = 0_E$, soit tous les λ_i nuls puisque $\mathbf{b}$ est libre. D'autre part, $f(\mathbf{b})$ est génératrice : un vecteur v de F est image de $e \in E$ qui est combinaison linéaire des (b_i), $e = \alpha_1 b_1 + \ldots + \alpha_n b_n$, et par suite $v = f(e) = f(\alpha_1 b_1 + \ldots + \alpha_n b_n) = \alpha_1 f(b_1) + \ldots + \alpha_n f(b_n)$.

Réciproquement, si E et $\widetilde{E}$ sont de même dimension n, on introduit une base $\mathbf{b} = (b_1, \ldots, b_n)$ de E, une base $\widetilde{\mathbf{b}} = (\widetilde{b}_1, \ldots, \widetilde{b}_n)$ de F. Alors l'unique application linéaire f qui applique le vecteur b_i sur $\widetilde{b}_i$, $i = 1, \ldots, n$ (cf prop. 2.2), vérifiant

$$f(\alpha_1 b_1 + \ldots . \alpha_n b_n) = \alpha_1 \widetilde{b}_1 + \ldots + \alpha_n \widetilde{b}_n\,,$$

est bijective. En effet, tout vecteur v de $\widetilde{E}$ est de la forme $v = \alpha_1 \widetilde{b}_1 + \ldots + \alpha_n \widetilde{b}_n$, il est l'image du vecteur $\alpha_1 b_1 + \ldots + \alpha_n b_n$ de E, et, puisque $\mathbf{b}$ et $\widetilde{\mathbf{b}}$ sont des bases de E et $\widetilde{E}$ resp., ceci de manière unique (l'égalité $\alpha_1 \widetilde{b}_1 + \ldots + \alpha_n \widetilde{b}_n = \alpha'_1 \widetilde{b}_1 + \ldots + \alpha'_n \widetilde{b}_n$ implique l'égalité des coefficients $\alpha_i = \alpha'_i$). Ainsi les espaces E et $\widetilde{E}$ sont donc isomorphes. □

△ Un espace de dimension n est isomorphe à $\mathbb{C}^n$. ▽

Définition 2.4 *Soit f une application linéaire de E dans F.*

*(i) L'*image *de f est la partie de F, notée* Im f *ou $f(E)$, définie par*

$$\mathrm{Im}\, f = \{v \in F : \textit{ il existe } e \in E, v = f(e)\}\,.$$

(ii) Le noyau *de f est la partie de E, notée* Ker f*, définie par*

$$\mathrm{Ker}\, f = \{e \in E : f(e) = 0\}\,,$$

△ *Ker* est le début de *Kern*, le mot allemand signifiant noyau (et aussi de *kernel*, sans majuscule, l'équivalent anglais). ▽

Proposition 2.7 *Soit f une application linéaire de E dans F. Alors son noyau est un sous-espace vectoriel de E, son image un sous-espace vectoriel de F.*

Preuve : On a $f(0_E) = 0_F$, donc $0_E \in \operatorname{Ker} f$ et $0_F \in \operatorname{Im} f$. D'autre part, si $e_1, e_2 \in \operatorname{Ker} f$ et $\alpha_1, \alpha_2 \in \mathbb{C}$, alors $f(\alpha_1 e_1 + \alpha_2 e_2) = \alpha_1 f(e_1) + \alpha_2 f(e_2) = 0_F$, soit $\alpha_1 e_1 + \alpha_2 e_2$ dans $\operatorname{Ker} f$. De manière analogue, pour $e_1, e_2 \in E$ et $\alpha_1, \alpha_2 \in \mathbb{C}$, le vecteur $\alpha_1 f(e_1) + \alpha_2 f(e_2) = f(\alpha_1 e_1 + \alpha_2 e_2)$ est dans $\operatorname{Im} f$. Les parties $\operatorname{Ker} f$ et $\operatorname{Im} f$ sont des sous-espaces de E et F resp. d'après la caractérisation de la proposition 1.2. $\square$

Exemple 2.2 Soient U, V des sous-espaces supplémentaires dans E. L'application π qui au vecteur e de E associe le vecteur u de la décomposition $e = u + v, u \in U, v \in V$ est une application linéaire : il suffit de remarquer, avec les notations évidentes

$$\alpha e + \alpha' e' = \alpha(u + v) + \alpha'(u' + v') = (\alpha u + \alpha' u') + (\alpha v + \alpha v'),$$

où $\alpha u + \alpha' u' \in U$ et $\alpha v + \alpha v' \in V$, et donc $\pi(\alpha e + \alpha' e') = \alpha \pi(e) + \alpha' \pi(e')$. Elle est appelée *projection de E sur U parallèlement à V*. Elle vérifie $\operatorname{Ker} \pi = V$ et $\operatorname{Im} \pi = U$.

Corollaire 2.1 *Soit U un sous-espace de E. Alors $f(U)$ est un sous-espace de F et $f(U) \subset \operatorname{Im} f$.*

Preuve : L'application $f_{|U} : u \in U \to f(u) \in F$ est la restriction de $f : E \to F$ à U et est linéaire. Le résultat se lit dans la proposition précédente. $\square$

Proposition 2.8 *Soit f application linéaire de E dans F. Si la famille $\mathbf{e} = (e_1, \ldots, e_p)$ engendre E, alors la famille $f(\mathbf{e}) = (f(e_1), \ldots, f(e_p))$ engendre l'image $\operatorname{Im} f$.*

Preuve : Soit $f(e)$ dans $\operatorname{Im} f$, pour un certain vecteur e de E. Alors e est combinaison linéaire des $e_1, \ldots, e_p$

$$e = \alpha_1 e_1 + \ldots + \alpha_p e_p$$

et donc

$$f(e) = f(\alpha_1 e_1 + \ldots + \alpha_p e_p) = \alpha_1 f(e_1) + \ldots + \alpha_p f(e_p)$$

c'est-à-dire, $f(e)$ est combinaison linéaire de $f(e_1), \ldots, f(e_p)$. $\square$

Définition 2.5 *Soit f une application linéaire de E dans F. Si $\operatorname{Im} f$ est de dimension finie, alors le* rang *de f est défini comme étant la dimension de l'espace $\operatorname{Im} f$.*

$\triangle$ Si $(e_i)_{i \in I}$ engendre E, le rang $\operatorname{rg} f$ est celui de la famille $(f(e_i))_{i \in I}$ (supposée engendrer un espace de dimension finie), comme défini dans la définition 1.11. $\triangledown$

Proposition 2.9 *Soit $f : E \to F$ linéaire. L'application est injective si et seulement si $\operatorname{Ker} f = \{0_E\}$.*

Preuve : Si f est injective, 0_E est l'unique e de E tel que $f(e) = 0_F$. Ainsi $\operatorname{Ker} f = \{0_E\}$.

Réciproquement, supposons $\operatorname{Ker} f$ égal à l'espace vectoriel nul $\{0_E\}$. Si e_1 et e_2 ont même image, alors $f(e_1 - e_2) = f(e_1) + f(-e_2) = f(e_1) - f(e_2) = 0_F$ et donc $e_1 - e_2$ est un vecteur du noyau $\operatorname{Ker} f$, par suite c'est le vecteur nul. Ainsi $e_1 = e_2$ et f est injective. $\square$

Exemple 2.3 L'endomorphisme Φ de $\mathbb{C}[X]$ qui à P associe le polynôme $P(X+1) - P(X)$ a pour noyau le sous-espace des polynômes constants, il n'est pas injectif donc. En effet, si P est un polynôme non constant de ce noyau, si s est une de ses racines, $s+1, s+2, \ldots$ sont aussi racines, ce qui ne peut être (P a un nombre fini de racines).

Théorème 2.1 (du noyau et de l'image) *Soit E de dimension finie et f une application de E dans F. Alors, les espaces $\operatorname{Ker} f$ et $\operatorname{Im} f$ sont de dimension finie. De plus*

$$\dim E = \dim \operatorname{Ker} f + \dim \operatorname{Im} f .$$

Preuve : Soit S un supplémentaire (dont l'existence est assurée par la Prop. 1.12) de $\operatorname{Ker} f$ dans E. Alors l'application $f_{|S} : S \to \operatorname{Im} f$ est bijective : un vecteur de son noyau est dans $\operatorname{Ker} f$, et donc nul vu $S \cap \operatorname{Ker} f = \{0_E\}$; tout vecteur $e \in E$ s'écrit $e = k + s$ avec $k \in \operatorname{Ker} f$ et $s \in S$ et alors $f(e) = f(k) + f(s) = f(s)$, en résulte la surjectivité de $f_{|S} : S \to \operatorname{Im} f$. Ainsi, l'application $f_{|S} : S \to \operatorname{Im} f$ est un isomorphisme, donc S et l'image $\operatorname{Im} f$ sont de même dimension. En outre, vu que $\dim E = \dim \operatorname{Ker} f + \dim S$, l'égalité entre dimensions de la proposition est établie. □

Exemple 2.4 L'endomorphisme Φ_n de $\mathbb{C}_n[X]$ qui à P associe $P(X+1) - P(X)$ a pour noyau la droite vectorielle des polynômes constants. Son image, de dimension n, est contenue dans $\mathbb{C}_{n-1}[X]$: c'est donc ce sous-espace tout entier.

△ Vu que, si E de dimension finie, $\operatorname{rg} f = \dim \operatorname{Im} f = \dim E - \dim \operatorname{Ker} f$ et que $\operatorname{Im} f$ est un sous-espace de F, on a, si E et F sont de dimension finie, $\operatorname{rg} f \leq \inf(\dim E, \dim F)$. ▽

Corollaire 2.2 *Soit E et F deux espaces vectoriels de dimension finie, de mêmes dimensions. Si f est une application linéaire de E dans F, les propriétés suivantes sont équivalentes*

(i) f est injective,
(ii) f est surjective,
(iii) f est bijective.

Preuve : On a $n = \dim F = \dim E = \dim \operatorname{Ker} f + \dim \operatorname{Im} f$. La proposition énonce en fait l'équivalence de

1. $\operatorname{Ker} f = \{0_E\}$ (soit $\dim \operatorname{Ker} f = 0$),
2. $\operatorname{Im} f = F$ (soit $\dim \operatorname{Im} f = n$),
3. $\operatorname{Im} f = F$ et $\dim \operatorname{Ker} f = 0$ (soit $\dim \operatorname{Ker} f = 0$ et $\dim \operatorname{Im} f = n$).

ce qui est immédiat vu la relation entre les dimensions rappelée en début de preuve. □

Exemple 2.5 L'hypothèse *de dimension finie* est essentielle. Ainsi, l'endomorphisme Φ introduit dans l'exemple 2.3 n'est pas injectif, mais est surjectif puisque, avec les notations de l'exemple 2.4,

$$\mathbb{C}[X] = \cup_{n\geq 1}\mathbb{C}_n[X] = \cup_{n\geq 2}\Phi_n(\mathbb{C}_n[X]) = \cup_{n\geq 2}\Phi(\mathbb{C}_n[X]) = \Phi(\cup_{n\geq 2}\mathbb{C}_n[X]) = \Phi(\mathbb{C}[X]) .$$

À l'inverse, l'endomorphisme de $\mathbb{C}[X]$ d'intégration

$$\Psi : P \to \left(x \to \int_0^x P(t)dt \right)$$

est injectif, mais n'est pas surjectif (son image est le sous-espace des polynômes P s'annulant en 0).

2.3. Espaces d'applications linéaires

Proposition 2.10 *Soient E et F deux espaces vectoriels.*

L'ensemble $\mathcal{L}(E,F)$ des applications linéaires de E dans F muni de l'addition

$$f_1, f_2 \in \mathcal{L}(E,F) \to (f_1 + f_2 : e \in E \to f_1(e) + f_2(e) \in F)$$

et de la multiplication par un scalaire

$$\alpha \in \mathbb{C}, f \in \mathcal{L}(E,F) \to (\alpha \cdot f : e \in E \to \alpha \cdot f(e) \in F)$$

est un espace vectoriel.

Preuve : L'ensemble $\mathcal{L}(E,F)$ est une partie de l'espace vectoriel $\mathcal{F}(E,F)$. Cette partie contient l'application $0_{\mathcal{F}(E,F)} : e \in E \to 0_F$, vecteur nul de $\mathcal{F}(E,F)$. Ainsi, pour montrer que $\mathcal{L}(E,F)$ est un sous-espace vectoriel de $\mathcal{F}(E,F)$, il suffit de montrer que $\lambda_1 f_1 + \lambda_2 f_2$ est une application linéaire pour λ_1, λ_2 scalaires et f_1, f_2 applications linéaires. On vérifie, pour $e, e' \in E$ et $\alpha, \alpha' \in \mathbb{C}$

$$\begin{aligned}
[\lambda_1 f_1 + \lambda_2 f_2](\alpha e + \alpha' e') &= \lambda_1 f_1(\alpha e + \alpha' e') + \lambda_2 f_2(\alpha e + \alpha' e') \\
&= \lambda_1(\alpha f_1(e) + \alpha' f_1(e')) + \lambda_2(\alpha f_2(e) + \alpha' f_2(e')) \\
&= \lambda_1 \alpha f_1(e) + \lambda_1 \alpha' f_1(e') + \lambda_2 \alpha f_2(e) + \lambda_2 \alpha' f_2(e') \\
&= \alpha(\lambda_1 f_1(e) + \lambda_2 f_2(e)) + \alpha'(\lambda_1 f_1(e') + \lambda_2 f_2(e')) \\
&= \alpha(\lambda_1 f_1 + \lambda_2 f_2)(e) + \alpha'(\lambda_1 f_1 + \lambda_2 f_2)(e') .
\end{aligned}$$

□

Proposition 2.11 *Si E et F sont de dimension finie, l'espace $\mathcal{L}(E,F)$ est de dimension finie et* $\dim \mathcal{L}(E,F) = \dim E \dim F$.

Preuve : Soit $\mathbf{b} = (b_1, \ldots, b_p)$ une base de E, $\mathbf{c} = (c_1, \ldots, c_n)$ une base de F. Si $e = x_1 b_1 + \ldots + x_p b_p$ est un vecteur de E et f une application de $\mathcal{L}(E,F)$,

$$\begin{aligned}
f(e) &= x_1 f(b_1) + \ldots + x_p f(b_p) \\
&= x_1(f_{11} c_1 + \ldots + f_{n1} c_n) + \ldots + x_p(f_{1p} c_1 + \ldots + f_{np} c_n)
\end{aligned}$$

où on a décomposé $f(b_j), j = 1, \ldots, p$, dans la base $\mathbf{c}$

$$f(b_j) = f_{1j} c_1 + \ldots + f_{nj} c_n .$$

Soit pour i et j entiers avec $1 \leq i \leq n, 1 \leq j \leq p$ l'application linéaire Φ_{ij} de E dans F définie par $\Phi_{ij}(b_k) = 0$, $k = 1, \ldots, p$ sauf pour $k = j$ où $\Phi_{ij}(b_j) = c_i$, on a $f = \sum_{(i,j)} f_{ij} \Phi_{ij}$ où la sommation porte sur les couples (i,j) avec $i = 1, \ldots, p, j = 1, \ldots, n$. La famille Φ_{ij} est donc génératrice dans $\mathcal{L}(E,F)$ (qui est donc de dimension finie). Elle est aussi libre : si $\sum_{(i,j)} \lambda_{ij} \Phi_{ij} = 0$, évaluant sur le vecteur b_k, on obtient $0 = \sum_i \lambda_{ik} c_i$, soit tous les λ_{ik} nuls, puisque $\mathbf{c}$ est libre. Ainsi la famille Φ_{ij} est-elle une base de $\mathcal{L}(E,F)$, qui est donc de dimension $pn = \dim E \dim F$. □

Définition 2.6 *Une application linéaire de E dans $\mathbb{C}$ est appelée* forme linéaire. *L'espace $\mathcal{L}(E,\mathbb{C})$ des formes linéaires sur E est appelé l'*espace dual *de E. Il est souvent noté E^*.*

△ D'après la Prop. 2.11, si E est de dimension finie, son dual E^* est de même dimension. ▽

Exemple 2.6 La famille de formes $\ell_k, k = 0, \ldots, n$ définies sur $\mathbb{C}_n[X]$ par $\ell_k(P) = P(k)$ est libre : si $\lambda_0\ell_0 + \ldots + \lambda_n\ell_n = 0$, évaluant sur le polynôme P_k de degré n et ayant comme ensemble de racines $\{0, \ldots, n\} \setminus \{k\}$ et tel que $P_k(k) = 1$ (le polynôme de Lagrange de l'exemple 1.6 associé à la suite de réels $x_i = i, i = 0, \ldots, n$), on obtient $\lambda_k = 1$ pour $k = 0, \ldots, n$. C'est donc une base de l'espace des formes linéaires sur $\mathbb{C}_n[X]$. Si θ est une fonction continue sur $[0, 1]$, la forme linéaire $\Phi_\theta : P \to \int_0^1 P(t)\theta(t)dt$ est combinaison linéaire des $\ell_k, k = 0, \ldots n$, ainsi il existe des scalaires $\alpha_0, \ldots, \alpha_n$ tels que

$$\int_0^1 P(t)\theta(t)dt = \alpha_0 P(0) + \ldots + \alpha_n P(n)\,, \quad P \in \mathbb{C}_n[X]\,.$$

Les scalaires $\alpha_k, k = 0, \ldots, n$ dépendent de θ et de n (mais pas de P).

Proposition 2.12 *Soit E de dimension finie. Le noyau d'une forme linéaire non nulle sur E est un hyperplan de E et tout hyperplan de E apparaît comme le noyau d'une forme linéaire non nulle. Deux formes linéaires non nulles sur E définissent le même hyperplan de E si et seulement si elles sont colinéaires.*

Preuve : Soit φ une forme linéaire non nulle. D'après la proposition 2.1 et vu $\operatorname{Im}\varphi = \mathbb{C}$, on a $\dim \operatorname{Ker}\varphi = \dim E - 1$, ainsi $\operatorname{Ker}\varphi$ est un hyperplan. Inversement, soit $\mathcal{H}$ un hyperplan, $\mathbf{b} = (b_1, \ldots, b_{n-1})$ une base de $\mathcal{H}$ et v un vecteur hors de $\mathcal{H}$. Alors $(b_1, \ldots, b_{n-1}, v)$ est une base de E et l'application qui au vecteur e associe sa coordonnée y dans la décomposition $e = x_1b_1 + \ldots + x_{n-1}b_{n-1} + yv$ est une forme linéaire dont le noyau est $\mathcal{H}$.

Soient $\varphi_i, i = 1, 2$ deux formes dont le noyau coïncident avec l'hyperplan $\mathcal{H}$ et v un vecteur hors de $\mathcal{H}$, qui n'est pas annulé donc par les $\varphi_i, i = 1, 2$. Soit $\alpha = \varphi_1(v)/\varphi_2(v)$. Alors, vu que l'hyperplan $\mathcal{H}$ et la droite $\mathbb{C}v$ sont supplémentaires, tout vecteur e s'écrit $e = h + \lambda v$ avec $h \in \mathcal{H}, \lambda \in \mathbb{C}$ et donc

$$\varphi_1(e) = \varphi_1(h + \lambda v) = \lambda\varphi_1(v) = \lambda\alpha\varphi_2(v) = \alpha\varphi_2(\lambda v) = \alpha\varphi_2(e)\,,$$

soit $\varphi_1 = \alpha\varphi_2$. $\square$

$\triangle$ Avec les notation du Lemme précédent, l'hyperplan $\mathcal{H}$ a pour *équation cartésienne* $\varphi = 0$. Par ailleurs, on dit que $\mathcal{H}$ a pour *représentation paramétrique*

$$(x_1, \ldots, x_{n-1}) \in \mathbb{C}^{n-1} \to x_1b_1 + \ldots + x_{n-1}b_{n-1}\,.$$

Ceci peut être généralisé à tout sous-espace vectoriel U de E. Si $(b_1, \ldots, b_k)$ est une base de U, une représentation paramétrique de U est donnée par

$$(x_1, \ldots, x_k) \in \mathbb{C}^k \to x_1b_1 + \ldots + x_kb_k\,.$$

Le sous-espace U peut-être considéré comme l'intersection des hyperplans $\mathcal{H}_i, i = k+1, \ldots, n$ définis par

$$\mathcal{H}_i = \{y_1b_1 + \ldots + y_nb_n : y_i = 0\}\,,$$

où la base $(b_1, \ldots, b_k)$ de U a été complétée par la famille $(b_{k+1}, \ldots, b_n)$ en une base $(b_1, \ldots, b_n)$ de E. Le sous-espace U a

$$y_i = 0\,, \quad i = k + 1, \ldots, n$$

comme représentation cartésienne. $\triangledown$

3. Matrices

Les calculs explicites sur une application linéaire $f : E \to F$ sont en général tributaires de choix de bases dans E et F, choix qui permettent de représenter par une matrice l'application linéaire f. Toute opération impliquant une application linéaire se traduit en termes de matrices, l'effet des changements du choix de bases de références sur ces représentations étant spécialement étudié. Le point de vue inverse, qui associe à une matrice une application linéaire, éclaire de manière fructueuse certaines propriétés matricielles.

Dans cette partie, tout espace vectoriel sera supposé de dimension finie.

3.1. Tableaux matriciels

Définition 3.1 *Soit n, p des entiers positifs non nuls. Une matrice M à n lignes et p colonnes, est la donnée de np scalaires $(m_{ij})_{1\leq i\leq n, 1\leq j\leq p}$ (notés aussi $(M)_{ij}$ ou M_{ij}), représentée sous la forme d'un tableau*

$$M = \begin{pmatrix} m_{11} & \dots & m_{1p} \\ \vdots & & \vdots \\ m_{n1} & \dots & m_{np} \end{pmatrix},$$

à n lignes et p colonnes. L'ensemble des matrices à n colonnes et p lignes est noté $\mathcal{M}_{np}$.

$\triangle$ Si le corps des scalaires est $\mathbb{C}$ (resp. $\mathbb{R}$), on parle de matrices à coefficients complexes (resp. réels) et on précise la notation $\mathcal{M}_{np}$ en $\mathcal{M}_{np}(\mathbb{C})$ (resp. $\mathcal{M}_{np}(\mathbb{R})$). Dans la suite, le corps des scalaires sera implicitement celui des complexes (les résultats restant valables si le choix de $\mathbb{R}$ est fait). $\triangledown$

Si $n = 1$, on parle de *matrice ligne* ou *vecteur ligne*, si $p = 1$, on parle de *matrice colonne* ou *vecteur colonne*. En particulier, les *vecteurs lignes de la matrice M* sont les n vecteurs ligne

$$L_i = (m_{i1}, \dots, m_{ip}), \quad i = 1, \dots, n,$$

et les *vecteurs colonnes de la matrice M* sont les p vecteurs colonne

$$C_j = \begin{pmatrix} m_{1j} \\ \vdots \\ m_{nj} \end{pmatrix}, \quad j = 1, \dots, p.$$

Si $n = p$, les éléments de $\mathcal{M}_{nn}$ (noté aussi $\mathcal{M}_n$) sont dits *matrices carrées* d'ordre n. La matrice carrée d'ordre n, dont les coefficients diagonaux $m_{ii}, i = 1, \dots, n$ valent 1 et les autres coefficients sont nuls est notée I_n.

Définition 3.2 *Si M est une matrice de $\mathcal{M}_{np}$, sa matrice* transposée *${}^tM$ est la matrice de $\mathcal{M}_{pn}$ définie par*

$$({}^tM)_{k\ell} = (M)_{\ell k}, \quad k = 1, \dots, p, \ \ell = 1, \dots, n.$$

$\triangle$ La transposition induit une bijection de $\mathcal{M}_{np}$ sur $\mathcal{M}_{pn}$, dont l'inverse est induit par la transposition. La matrice identité I_n est laissée invariant par la transposition. $\triangledown$

La transposée d'une matrice ligne est une matrice colonne, et inversement

$${}^t(x_1, \dots, x_p) = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, \quad {}^t\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1, \dots, x_n).$$

Définition/Proposition 3.1 *La* somme $M+N$ *de deux matrices* M,N *de* $\mathcal{M}_{np}$ *est la matrice de* $\mathcal{M}_{np}$ *de coefficients*

$$(M+N)_{ij} = (M)_{ij} + (N)_{ij}\,, \quad i=1,\dots,n,\;\; j=1,\dots,p\,.$$

La matrice obtenue en multipliant la matrice $M \in \mathcal{M}_{np}$ *par le scalaire* λ *est la matrice* λM *de* $\mathcal{M}_{np}$ *de coefficients*

$$(\lambda M)_{ij} = \lambda (M)_{ij}\,, \quad i=1,\dots,n,\;\; j=1,\dots,p\,.$$

Muni de ces opérations, l'ensemble $\mathcal{M}_{np}$ *est un espace vectoriel.*

En fait, $\mathcal{M}_{np}$ est de la forme $\mathcal{F}(A,\mathbb{C})$ avec A l'ensemble produit $\{1,\dots,n\}\times\{1,\dots,p\}$. Ainsi, d'après l'exemple 1.5, l'espace $\mathcal{M}_{np}$ est de dimension finie np, avec pour base la famille de matrices $E_{ij}, i=1,\dots,n, j=1\dots,p$ où la matrice E_{ij} a tous ses coefficients nuls à l'exception de celui sur la ligne d'indice i et la colonne d'indice j valant 1.

Définition 3.3 *Le* produit MN *de la matrice* $M \in \mathcal{M}_{np}$ *et* $N \in \mathcal{M}_{pq}$ *est la matrice de* $\mathcal{M}_{nq}$ *de coefficients*

$$\begin{aligned}(MN)_{ik} &= (M)_{i1}(N)_{1k} + \ldots + (M)_{ij}(N)_{jk} + \ldots + (M)_{ip}(N)_{pk} \\ &= \sum_{j=1}^{p} (M)_{ij}(N)_{jk}\,, \quad i=1,\dots,n,\;\; k=1,\dots,q\,.\end{aligned}$$

Proposition 3.1 *Si les matrices* M,N,P *ont des dimensions compatibles de telle sorte que les produits* $(MN)P$ *et* $M(NP)$ *sont bien définis, alors on a égalité* $M(NP)=(MN)P$.

Preuve : Cela résulte d'un calcul explicite

$$\begin{aligned}(M(NP))_{i\ell} &= \sum_j M_{ij}(NP)_{j\ell} = \sum_j M_{ij}\left(\sum_k N_{jk}P_{k\ell}\right) \\ &= \sum_k \left(\sum_j M_{ij}N_{jk}\right) P_{k\ell} = \sum_k (MN)_{ik}P_{k\ell} = ((MN)P)_{i\ell}\,.\end{aligned}$$

où il n'a pas été précisé les tailles des diverses matrices. □

△ Les parenthèses otées, on écrira simplement MNP pour le produit $M(NP)=(MN)P$. On remarquera que si $M\in\mathcal{M}_{np}$, on a $I_nM = MI_p = M$. Si M et N sont carrées de même ordre, il n'est pas vrai en général que $MN=NM$, par exemple $M = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$ et $N = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$. ▽

Définition/Proposition 3.2 *Une matrice* M *carrée d'ordre* n *est dite* inversible *s'il existe une matrice* N *carrée d'ordre* n *telle que* $MN = NM = I_n$. *Si* M *est inversible, une matrice* N *est uniquement déterminée et sera appellé* inverse *de* M, *notée* M^{-1}.

Preuve : Si M est inversible, si N_1 et N_2 sont deux tels inverses, multipliant à gauche l'égalité $I_n = NN_2$ par N_1, on obtient

$$N_1 = N_1(NN_2) = (N_1N)N_2 = I_nN_2 = N_2\,.$$

Ainsi l' inverse est-il bien uniquement défini. $\square$

La transposition est compatible vis à vis du produit

$$^t(AB) = {}^tB\,{}^tA\,, \quad A \in \mathcal{M}_{np},\ B \in \mathcal{M}_{pq}\,,$$

puisque

$$({}^t(AB))_{ki} = (AB)_{ik} = \sum_j A_{ij}B_{jk}\,, \quad ({}^tB{}^tA)_{ki} = \sum_j ({}^tB)_{kj}{}^tA_{ji} = \sum_j B_{jk}A_{ij}$$

et de l'inverse

$$^t(A^{-1}) = ({}^tA)^{-1}\,,$$

si A est inversible, puisque ${}^t(A^{-1})\,{}^tA = {}^t(AA^{-1}) = {}^tI_n = I_n$ et pareillement ${}^tA\,{}^t(A^{-1}) = I_n$.

3.2. Matrice d'une application linéaire

Définition 3.4 *Soit E un espace vectoriel de dimension n et $\mathbf{b} = (b_1, \dots, b_n)$ une base de E. Si $e = x_1b_1 + \dots + x_nb_n$ est un vecteur de E, la matrice colonne $X = {}^t(x_1, \dots, x_n)$ est appelée* la matrice colonne des coordonnées ou composantes du vecteur e dans la base $\mathbf{b}$.

△ Relativement à la base canonique, le vecteur $(x_1, \dots, x_n)$ de $\mathbb{C}^n$ a ${}^t(x_1, \dots, x_n)$ pour vecteur colonne de coordonnées. ▽

Définition 3.5 *Soient E, F des espaces vectoriels de dimension finie, p, n leurs dimensions respectives, $\mathbf{b} = (b_1, \dots, b_p)$ une base de E, $\mathbf{c} = (c_1, \dots, c_n)$ une base de F. Si f est une application linéaire de E dans F, la* matrice de f relativement aux bases $\mathbf{b}$ et $\mathbf{c}$ *est la matrice $M = (m_{ij})_{1\le i\le n, 1\le j\le p}$ dont le j-ième vecteur colonne est le vecteur des coordonnées de $f(b_j)$ relativement à la base $\mathbf{c}$. On note la matrice M par $M_{\mathbf{b},\mathbf{c}}(f)$, ou simplement $M(f)$ s'il n'y a pas d'ambiguïté sur le choix des bases.*

△ La matrice $M_{\mathbf{b},\mathbf{c}}(f)$ de f dépend des bases de E et F choisies. Néanmoins, si $E = F$, auquel cas f est un endomorphisme de E, on prendra en général les bases $\mathbf{b}$ et $\mathbf{c}$ identiques : on parlera alors de la matrice $M_{\mathbf{b}}(f)$ de f dans la base $\mathbf{b}$ de E, matrice carrée d'ordre la dimension de E. ▽

Exemple 3.1 Dans un espace E de dimension n, l'application Id_E est représentée dans n'importe quelle base $\mathbf{b}$ de E par la matrice diagonale I_n : $M_{\mathbf{b}}(\mathrm{Id}_E) = I_n$.

Dans l'espace $\mathbb{C}_n[X]$ des polynômes de degré au plus n, soient l'application de dérivation $D : P \to P'$ et les bases

$$\mathbf{b} = (X^n, X^{n-1}, \dots, X, 1)\,, \quad \mathbf{c} = (nX^{n-1}, (n-1)X^{n-2}, \dots, 2X, 1, X^n)\,.$$

Alors

$$M_{\mathbf{b},\mathbf{c}}(D) = \begin{pmatrix} 1 & 0 & & & \\ 0 & 1 & & & \\ & 0 & \ddots & 0 & \\ & & 0 & 1 & 0 \\ & & & 0 & 0 \end{pmatrix}, \quad M_{\mathbf{b},\mathbf{b}}(D) = \begin{pmatrix} 0 & 0 & & & \\ n & 0 & & & \\ & n-1 & 0 & 0 & \\ & & \ddots & \ddots & 0 \\ & & & 1 & 0 \end{pmatrix}.$$

$\triangle$ Soient $\mathbb{C}^p$ et $\mathbb{C}^n$, munis de leurs bases canoniques $(e_1, \ldots, e_p)$ et $(\widetilde{e_1}, \ldots, \widetilde{e_n})$ respectives. La matrice $M = (m_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$ est la matrice de l'application linéaire de $\mathbb{C}^p$ dans $\mathbb{C}^n$ qui envoie e_j sur $m_{1j}\widetilde{e_1} + \ldots + m_{nj}\widetilde{e_n} = (m_{1j}, \ldots, m_{nj})$. $\triangledown$

Proposition 3.2 *Soit f une application linéaire de E dans F, $\mathbf{b} = (b_1, \ldots, b_p)$ et $\mathbf{c} = (c_1, \ldots, c_n)$ des bases de E et F respectivement. Si X (resp. Y) est le vecteur colonne des coordonnées du vecteur x de E (resp. $y = f(x)$ de F) relativement à la base $\mathbf{b}$ (resp. $\mathbf{c}$). alors*

$$Y = M_{\mathbf{b},\mathbf{c}}(f)X$$

Preuve : Soit $x = x_1 b_1 + \ldots + x_p b_p$. Alors

$$\begin{aligned} y = f(x) &= f(x_1 b_1 + \ldots + x_p b_p) \\ &= x_1 f(b_1) + \ldots + x_p f(b_p) \\ &= x_1(m_{11}c_1 + \ldots + m_{n1}c_n) + \ldots + x_p(m_{1p}c_1 + \ldots + m_{np}c_n) \\ &= (m_{11}x_1 + \ldots + m_{1p}x_p)c_1 + \ldots + (m_{n1}x_1 + \ldots + m_{np}x_p)c_n \end{aligned}$$

soit

$$Y = \begin{pmatrix} m_{11}x_1 + \ldots + m_{1p}x_p \\ \vdots \\ m_{n1}x_1 + \ldots + m_{np}x_p \end{pmatrix} = M_{\mathbf{b},\mathbf{c}}(f)X$$

$\square$

$\triangle$ L'écriture $Y = MX$, ou plutôt $MX = Y$ peut s'interpréter comme un système d'équations linéaires, où, Y étant donné, on cherche un X vérifiant $MX = Y$: l'étude d'un système linéaire (à savoir l'existence et l'unicité de la solution) est donc équivalente (par le choix de bases) à la considération des propriétés (*i. e.* surjectivité et injectivité) d'une application linéaire. $\triangledown$

Théorème 3.1 *Soient E et F des espaces vectoriels de dimension p et n resp., $\mathbf{b} = (b_1, \ldots, b_p)$ et $\mathbf{c} = (c_1, \ldots, c_n)$ des bases de E et F resp.*

L'application

$$M_{\mathbf{b},\mathbf{c}} : f \in \mathcal{L}(E, F) \to M_{\mathbf{b},\mathbf{c}}(f) \in \mathcal{M}_{np}$$

est un isomorphisme linéaire.

Preuve : Soient f, g des applications linéaires de E dans F de matrices resp. $M_{\mathbf{b},\mathbf{c}}(f) = (\alpha_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$ et $M_{\mathbf{b},\mathbf{c}}(g) = (\beta_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$ resp. On a donc, pour $j = 1, \ldots, p$,

$$f(b_j) = \sum_{i=1}^{n} \alpha_{ij} c_i \,, \quad g(b_j) = \sum_{i=1}^{n} \beta_{ij} c_i \,.$$

Ainsi, pour λ, μ scalaires,

$$(\lambda f + \mu g)(b_j) = \lambda f(b_j) + \mu g(b_j) = \lambda \sum_{i=1}^{n} \alpha_{ij} c_i + \mu \sum_{i=1}^{n} \beta_{ij} c_i = \sum_{i=1}^{n} (\lambda \alpha_{ij} + \mu \beta_{ij}) c_i$$

soit $M_{\mathbf{b},\mathbf{c}}(\lambda f + \mu g) = \lambda M_{\mathbf{b},\mathbf{c}}(f) + \mu M_{\mathbf{b},\mathbf{c}}(g)$. L'application $M_{\mathbf{b},\mathbf{c}}$ est donc linéaire. La proposition 2.3 énonce qu'une application linéaire est déterminée, et de manière unique, par les images des vecteurs d'une base : c'est simplement exprimer que l'application $M_{\mathbf{b},\mathbf{c}} : \mathcal{L}(E, F) \to \mathcal{M}_{\mathbf{b},\mathbf{c}}$ est un isomorphisme. $\square$

$\triangle$ On retrouve par cet isomorphisme que $\mathcal{L}(E, F)$ a pour dimension $\dim E \dim F$ (et donc l'espace des endomorphismes de E a pour dimension $(\dim E)^2$). $\triangledown$

Proposition 3.3 *Soient E, F, G des espaces vectoriels, de dimensions respectives p, n, m et munis de base $\mathbf{b}, \mathbf{c}, \mathbf{d}$ resp., f une application linéaire de E dans F et g une application linéaire de F dans G. Alors*

$$M_{\mathbf{b},\mathbf{d}}(g \circ f) = M_{\mathbf{c},\mathbf{d}}(g) M_{\mathbf{b},\mathbf{c}}(f) .$$

Preuve : Soient $M_{\mathbf{b},\mathbf{c}}(f) = (\alpha_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$ et et $M_{\mathbf{c},\mathbf{d}}(g) = (\beta_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ resp. Alors, pour $j = 1, \dots, p$

$$\begin{aligned}
g \circ f(b_j) &= g(f(b_j)) \\
&= g(\alpha_{1j} c_1 + \ldots + \alpha_{nj} c_n) \\
&= \alpha_{1j} g(c_1) + \ldots + \alpha_{nj} g(c_n) \\
&= \alpha_{1j}(\beta_{11} d_1 + \ldots + \beta_{m1} d_m) + \alpha_{nj}(\beta_{1n} d_1 + \ldots + \beta_{mn} d_m) \\
&= (\beta_{11}\alpha_{1j} + \ldots + \beta_{1n}\alpha_{nj}) d_1 + \ldots + (\beta_{m1}\alpha_{1j} + \ldots + \beta_{mn}\alpha_{nj}) d_m .
\end{aligned}$$

Ainsi, si $g \circ f$ a pour matrice $M_{\mathbf{b},\mathbf{d}}(g \circ f) = (\gamma_{kj})_{1 \leq k \leq m, 1 \leq j \leq p}$, on a

$$\gamma_{kj} = \beta_{k1}\alpha_{1j} + \ldots + \beta_{kn}\alpha_{nj} = \sum_{i=1}^{n} \beta_{ki}\alpha_{ij} ,$$

soit $M_{\mathbf{b},\mathbf{d}}(g \circ f) = M_{\mathbf{c},\mathbf{d}}(g) M_{\mathbf{b},\mathbf{c}}(f)$. □

△ Le produit des matrices a été construit en fait à partir de cette relation sur la matrice d'une composée d'applications linéaires. ▽

Proposition 3.4 *Soit E muni de la base $\mathbf{b}$, F de la base $\mathbf{c}$, E et F de même dimension n. L'application linéaire f de E dans F est un isomorphisme si et seulement si la matrice $M_{\mathbf{b},\mathbf{c}}(f)$ est inversible. Si f est un isomorphisme, alors $M_{\mathbf{c},\mathbf{b}}(f^{-1}) = M_{\mathbf{b},\mathbf{c}}(f)^{-1}$.*

Preuve : Si f est un isomorphisme de E dans F, son inverse f^{-1}, isomorphisme de F dans E vérifie $f^{-1} \circ f = f \circ f^{-1} = \mathrm{Id}_E$. Ainsi $M_{\mathbf{c},\mathbf{b}}(f^{-1}) M_{\mathbf{b},\mathbf{c}}(f) = M_{\mathbf{b}}(\mathrm{Id}_E) = I_n$ et $M_{\mathbf{b},\mathbf{c}}(f) M_{\mathbf{c},\mathbf{b}}(f^{-1}) = M_{\mathbf{c}}(\mathrm{Id}_F) = I_n$. Ainsi $M_{\mathbf{b},\mathbf{c}}(f)$ est inversible, d'inverse $M_{\mathbf{c},\mathbf{b}}(f^{-1})$: $(M_{\mathbf{b},\mathbf{c}}(f))^{-1} = M_{\mathbf{c},\mathbf{b}}(f^{-1})$.

Réciproquement, supposons $M_{\mathbf{b},\mathbf{c}}(f)$ inversible, d'inverse N. Alors, l'unique application linéaire g de F dans E ayant pour matrice N relativement aux bases $\mathbf{c}$ et $\mathbf{b}$, *i. e.* $N = M_{\mathbf{c},\mathbf{b}}(g)$, vérifie

$$\begin{aligned}
M_{\mathbf{b}}(g \circ f) &= M_{\mathbf{c},\mathbf{b}}(g) M_{\mathbf{b},\mathbf{c}}(f) = N M_{\mathbf{b},\mathbf{c}}(f) = I_n , \\
M_{\mathbf{c}}(f \circ g) &= M_{\mathbf{b},\mathbf{c}}(f) M_{\mathbf{c},\mathbf{b}}(g) = M_{\mathbf{b},\mathbf{c}}(f) N = I_n .
\end{aligned}$$

Par injectivité du morphisme $M_{\mathbf{b}}$ de $\mathcal{L}(E)$ dans $\mathcal{M}_n$ et vu que $M_{\mathbf{b}}(\mathrm{Id}_E) = I_n$, on a $g \circ f = \mathrm{Id}_E$ et, pareillement, $f \circ g = \mathrm{Id}_F$, soit l'isomorphie de f. □

3.3. Changement de base

Définition 3.6 *Soient $\mathbf{b} = (b_1, \dots, b_n)$, $\widetilde{\mathbf{b}} = (\widetilde{b_1}, \dots, \widetilde{b_n})$ deux bases de l'espace vectoriel E de dimension n. La* matrice de passage de la base $\mathbf{b}$ à la base $\widetilde{\mathbf{b}}$ *est définie comme étant la matrice $P_{\mathbf{b},\widetilde{\mathbf{b}}} = (p_{ij})$ de $\mathcal{M}_n$ avec*

$$\widetilde{b_j} = p_{1j} b_1 + \ldots + p_{nj} b_n , \quad j = 1, \dots, n.$$

La matrice $P = P_{\mathbf{b},\widetilde{\mathbf{b}}}$ est appelée aussi matrice de changement de base.

Exemple 3.2 Soient les bases

$$\begin{array}{rcl}\mathbf{b} & = & (X^n, X^{n-1}, \dots, X, 1)\,, \\ \mathbf{c} & = & (nX^{n-1}, (n-1)X^{n-2}, \dots, 2X, 1, X^n)\end{array}$$

de $\mathbb{C}_n[X]$. Alors

$$P_{\mathbf{b},\mathbf{c}} = \begin{pmatrix} 0 & & & & 1 \\ n & 0 & & & 0 \\ 0 & n-1 & & & \vdots \\ & & \ddots & & \\ & & & 1 & 0 \end{pmatrix}, \quad P_{\mathbf{c},\mathbf{b}} = \begin{pmatrix} 0 & n^{-1} & & & \\ 0 & 0 & (n-1)^{-1} & & 0 \\ & & & & \vdots \\ & & & \ddots & 1 \\ 1 & & & & 0 \end{pmatrix}$$

sont les matrices de passage de la base $\mathbf{b}$ à la base $\mathbf{c}$ et de $\mathbf{c}$ à $\mathbf{b}$ resp..

$\triangle$ On a deux visions d'une matrice de changement de base $P_{\mathbf{b},\widetilde{\mathbf{b}}}$:

- celle des coordonnées des vecteurs de $\widetilde{\mathbf{b}}$ dans la base $\mathbf{b}$, c'est donc la matrice de l'application Id_E relativement aux bases $\widetilde{\mathbf{b}}, \mathbf{b}$, *i. e.* $P_{\mathbf{b},\widetilde{\mathbf{b}}} = M_{\widetilde{\mathbf{b}},\mathbf{b}}(\mathrm{Id}_E)$;
- celle de l'endomorphisme $f_{\mathbf{b}\widetilde{\mathbf{b}}}$ de E qui à b_j associe le vecteur $\widetilde{b}_j$ relativement à la base $\mathbf{b}$, *i. e.* $P_{\mathbf{b},\widetilde{\mathbf{b}}} = M_{\mathbf{b}}(f_{\mathbf{b}\widetilde{\mathbf{b}}})$.

Dans la première version, on change le système de coordonnées, alors que dans la seconde on a un endomorphisme différent de l'identité représenté dans une seule base de E. $\triangledown$

Proposition 3.5 *Une matrice de changement de base est inversible et toute matrice inversible peut s'interpréter comme matrice de passage.*

La matrice de passage de la base $\mathbf{b}$ *à la base* $\widetilde{\mathbf{b}}$ *et celle de la base* $\widetilde{\mathbf{b}}$ *à la base* $\mathbf{b}$ *sont inverses l'une de l'autre.*

Preuve : Vu que $P_{\mathbf{b},\widetilde{\mathbf{b}}} = M_{\widetilde{\mathbf{b}},\mathbf{b}}(\mathrm{Id}_E)$, une matrice de changement de base est inversible. Inversement, soit M une matrice carrée d'ordre n inversible : si C_j note le $j$ème vecteur colonne de M, la matrice inversible M représente l'application linéaire de $\mathbb{C}^n$ qui associe au $j$ème vecteur de la base canonique le vecteur C_j, application qui est donc un isomorphisme d'après la proposition 3.4 : la famille (C_j), image d'une base par un isomorphisme est une base.

On a

$$\begin{array}{rcl} M_{\widetilde{\mathbf{b}},\mathbf{b}}(\mathrm{Id}_E) M_{\mathbf{b},\widetilde{\mathbf{b}}}(\mathrm{Id}_E) & = & M_{\mathbf{b},\mathbf{b}}(\mathrm{Id}_E) = I_n \\ M_{\mathbf{b},\widetilde{\mathbf{b}}}(\mathrm{Id}_E) M_{\widetilde{\mathbf{b}},\mathbf{b}}(\mathrm{Id}_E) & = & M_{\widetilde{\mathbf{b}},\widetilde{\mathbf{b}}}(\mathrm{Id}_E) = I_n \end{array}$$

soit

$$P_{\mathbf{b},\widetilde{\mathbf{b}}} P_{\widetilde{\mathbf{b}},\mathbf{b}} = P_{\widetilde{\mathbf{b}},\mathbf{b}} P_{\mathbf{b},\widetilde{\mathbf{b}}} = I_n$$

ainsi les matrices $P_{\mathbf{b},\widetilde{\mathbf{b}}}$ et $P_{\widetilde{\mathbf{b}},\mathbf{b}}$ sont inverses l'une de l'autre. $\square$

Proposition 3.6 *Soient* $\mathbf{b}$, $\widetilde{\mathbf{b}}$ *deux bases de l'espace vectoriel* E *et* P *la matrice de passage de* $\mathbf{b}$ *à* $\widetilde{\mathbf{b}}$. *Si le vecteur* v *de* E *a vecteur colonne de coordonnées* X *dans la base* $\mathbf{b}$ *et* $\widetilde{X}$ *dans la base* $\widetilde{\mathbf{b}}$, *on a* $X = P_{\mathbf{b},\widetilde{\mathbf{b}}}\widetilde{X}$.

Preuve : On a

$$\widetilde{X} = M_{\mathbf{b},\widetilde{\mathbf{b}}}(\mathrm{Id}_E) X = P_{\widetilde{\mathbf{b}},\mathbf{b}} X$$

soit,

$$X = P_{\mathbf{b},\widetilde{\mathbf{b}}} \widetilde{X}$$

en ayant multiplié par l'inverse $P_{\widetilde{\mathbf{b}},\mathbf{b}}^{-1} = P_{\mathbf{b},\widetilde{\mathbf{b}}}$ $\square$

Proposition 3.7 *Soient* $\mathbf{b}$, $\widetilde{\mathbf{b}}$ *(resp.* $\mathbf{c}$, $\widetilde{\mathbf{c}}$*) deux bases de l'espace vectoriel* E *(resp.* F*) et* $P = P_{\mathbf{b},\widetilde{\mathbf{b}}}$ *(*$Q = P_{\mathbf{c},\widetilde{\mathbf{c}}}$ *resp.) la matrice de passage de* $\mathbf{b}$ *à* $\widetilde{\mathbf{b}}$ *(*$\mathbf{c}$ *à* $\widetilde{\mathbf{c}}$ *resp.). Alors, pour* f *application linéaire de* E *dans* F*, on a*

$$M_{\widetilde{\mathbf{b}},\widetilde{\mathbf{c}}}(f) = Q^{-1} M_{\mathbf{b},\mathbf{c}}(f) P = P_{\widetilde{\mathbf{c}},\mathbf{c}} M_{\mathbf{b},\mathbf{c}}(f) P_{\mathbf{b},\widetilde{\mathbf{b}}}\,.$$

$\triangle$ En particulier, pour un endomorphisme f de E, on a $M_{\widetilde{\mathbf{b}}}(f) = P^{-1} M_{\mathbf{b}}(f) P$ avec $P = P_{\mathbf{b},\widetilde{\mathbf{b}}}$. $\triangledown$

Preuve : On écrit $f = \mathrm{Id}_E \circ f \circ \mathrm{Id}_E$, d'où

$$M_{\widetilde{\mathbf{b}},\widetilde{\mathbf{c}}}(f) = M_{\mathbf{c},\widetilde{\mathbf{c}}}(\mathrm{Id}_E) M_{\mathbf{b},\mathbf{c}}(f) M_{\widetilde{\mathbf{b}},\mathbf{b}}(\mathrm{Id}_E)$$

ce qui permet de conclure en invoquant $P = P_{\mathbf{b},\widetilde{\mathbf{b}}} = M_{\widetilde{\mathbf{b}},\mathbf{b}}(\mathrm{Id}_E)$ et $Q = P_{\mathbf{c},\widetilde{\mathbf{c}}} = P_{\widetilde{\mathbf{c}},\mathbf{c}}^{-1} = M_{\mathbf{c},\widetilde{\mathbf{c}}}(\mathrm{Id}_E)^{-1}$. $\square$

Exemple 3.3 On a donc $M_{\mathbf{b},\mathbf{c}}(f) = P_{\mathbf{c},\mathbf{b}} M_{\mathbf{b},\mathbf{b}}(f)$, soit en reprenant l'exemple de la dérivation D dans $\mathbb{C}_n[X]$

$$\begin{pmatrix} 1 & 0 & & & \\ 0 & 1 & & & \\ & 0 & \ddots & 0 & \\ & & 0 & 1 & 0 \\ & & & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & n^{-1} & & \\ 0 & 0 & (n-1)^{-1} & 0 \\ & & & \vdots \\ & & \ddots & 1 \\ 1 & & & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & & & \\ n & 0 & & & \\ & n-1 & 0 & 0 & \\ & & \ddots & \ddots & 0 \\ & & & 1 & 0 \end{pmatrix}.$$

3.4. Rang d'une matrice

Définition 3.7 *Le* rang d'une matrice M *de* $\mathcal{M}_{np}$ *est celui de la famille* $(C_1, \ldots, C_p)$ *de ses vecteurs colonnes, considérés comme des vecteurs de* $\mathbb{C}^n$*.*

Exemple 3.4 Soient p, n avec $p \leq n$ et M la matrice de $\mathcal{M}_{np}$

$$M = \begin{pmatrix} m_{11} & 0 & \ldots & 0 \\ \vdots & \ddots & & 0 \\ \vdots & & m_{kk} & \\ \vdots & & \ddots & 0 \\ m_{p1} & & \ldots & m_{pp} \\ \vdots & & & \vdots \\ m_{n1} & & \ldots & m_{np} \end{pmatrix}.$$

Si les $m_{ii}, i = 1, \ldots, p$ sont tous non nuls, le rang de la matrice M est p. En effet, si $(C_1, \ldots, C_p)$ sont les vecteurs colonnes de M, l'égalité $x_1 C_1 + \ldots + x_p C_p = 0$ est équivalente au système échelonné

$$\begin{array}{ccccccc} m_{11}x_1 & & & & & = & 0 \\ m_{21}x_1 & + & m_{22}x_2 & & & = & 0 \\ \vdots & & & & & & \\ m_{p1}x_1 & + & \ldots & + & m_{pp}x_p & = & 0\,, \\ m_{p+11}x_1 & + & \ldots & + & m_{p+1p}x_p & = & 0 \\ \vdots & & & & & & \\ m_{n1}x_1 & + & \ldots & + & m_{np}x_p & = & 0 \end{array}$$

dont les p premières équations résolues successivement donnent $x_1 = \ldots = x_p = 0$: les vecteurs $(C_1, \ldots, C_p)$ sont linéairement indépendants. Remarquons que, sans que ses coefficients diagonaux soient tous non nuls, la matrice

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

est de rang 2.

Le choix d'une base $\mathbf{b}$ de E de dimension n induit un isomorphisme de E avec $\mathbb{C}^n$: si $(v_1, \ldots, v_p)$ est une famille de vecteurs de l'espace vectoriel E, alors au vecteur v_j correspond dans $\mathbb{C}^n$ le vecteur colonne C_j de ses coordonnées relativement à la base $\mathbf{b}$ et les familles $(v_1, \ldots, v_p)$ et $(C_1, \ldots, C_p)$ ont même rang. La matrice constituée de ces vecteurs colonnes a même rang que la famille $(v_1, \ldots, v_p)$. Le rang d'une matrice est relié à la notion introduite précédemment pour les applications linéaires.

Proposition 3.8 *Le rang de la matrice M est le rang de toute application linéaire représentée par cette matrice.*

Preuve : Soit f l'application linéaire de E dans F représentée par la matrice M pour un choix de bases $\mathbf{b} = (b_1, \ldots, b_p)$ et $\mathbf{c} = (c_1, \ldots, c_n)$ de E et F. Les vecteurs colonnes de M correspondent aux vecteurs colonnes des coordonnées des vecteurs $f(b_1), \ldots, f(b_p)$ décomposés dans la base $\mathbf{c}$ de F. Ainsi le rang de la matrice M est celui de la famille $f(\mathbf{b}) = (f(b_1), \ldots, f(b_p))$ et donc celui de l'application linéaire f. $\square$

Proposition 3.9 *Soient E et F espaces vectoriels de dimensions respectives p et n. Soit f une application linéaire de E dans F de rang r. Il existe une base $\mathbf{b}$ de E, une base $\mathbf{c}$ de F telles que*

$$M_{\mathbf{b},\mathbf{c}}(f) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

Preuve : Soit S un supplémentaire de $\operatorname{Ker} f$, $\widetilde{\mathbf{b}}$ une base de S, $\widetilde{\widetilde{\mathbf{b}}}$ une base de $\operatorname{Ker} f$ et $\mathbf{b}$ la base $\mathbf{b} = \widetilde{\mathbf{b}} \cup \widetilde{\widetilde{\mathbf{b}}}$ de E. Vu que $f_{|S} : S \to \operatorname{Im} f$ est un isomorphisme (cf. preuve du Th. 2.1), la famille $f(\widetilde{\mathbf{b}})$ est une base de $\operatorname{Im} f$, qu'on complète avec une famille $\widetilde{\mathbf{c}}$ en une base $\mathbf{c}$ de F. la matrice $M_{\mathbf{b},\mathbf{c}}(f)$ est de la forme indiquée dans le théorème. $\square$

Soient p, q, r des entiers avec $r \leq \inf(p, q)$. La matrice $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ de $\mathcal{M}_{pq}$ est de rang r. Si P et Q sont des matrices inversibles d'ordre p et q resp., la matrice $Q \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P$ est aussi de rang r, puisque elle représente la même application linéaire que $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ dans des bases déduites des bases canoniques de $\mathbb{C}^p$ et $\mathbb{C}^n$ par des changements de bases décrits par les matrices P et Q^{-1}.

On peut reformuler la discussion précédente par le théorème

Théorème 3.2 *Soit M une matrice de $\mathcal{M}_{np}$. Il existe une matrice Q inversible d'ordre n, P inversible d'ordre p et un entier r tel que*

$$M = Q \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P.$$

L'entier r est le rang de M.

Preuve : On considère l'application linéaire f de $\mathbb{C}^p$ dans $\mathbb{C}^n$ représentée par la matrice M dans les bases canoniques. La proposition précédente assure l'existence de bases de $\mathbb{C}^p$ et $\mathbb{C}^n$ dans lesquelles f se représentent par $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. Ainsi, en considérant les matrices de passage $P_{\mathrm{can},\mathbf{b}}, P_{\mathrm{can},\mathbf{c}}$ de chacune de ces bases à la base canonique, on obtient

$$M = M_{\mathrm{can},\mathrm{can}}(f) = P_{\mathrm{can},\mathbf{c}}^{-1} M_{\mathbf{b},\mathbf{c}}(f) P_{\mathrm{can},\mathbf{b}} = P_{\mathrm{can},\mathbf{c}}^{-1} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P_{\mathrm{can},\mathbf{b}}$$

soit la forme indiquée par l'énoncé. □

Théorème 3.3 *Soit M une matrice de $\mathcal{M}_{np}$, ${}^tM$ sa matrice transposée. Les rangs de M et de ${}^tM$ sont égaux.*

Preuve : Si $M = Q \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P$, sa transposée vérifie ${}^tM = {}^tP \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} {}^tQ$. Ainsi M et ${}^tM$ ont même rang. □

Corollaire 3.1 *Le rang d'une matrice M de $\mathcal{M}_{np}$ est égal au rang de la famille de ses vecteurs ligne, considérés comme des vecteurs de $\mathbb{C}^p$.*

Preuve : Les vecteurs ligne de M sont les vecteurs colonnes de sa transposée ${}^tM$. □

Théorème 3.4 *Soit M une matrice de $\mathcal{M}_n$. Les assertions suivantes sont équivalentes :*

1. *la matrice M est inversible,*
2. *relativement à une base* $\mathbf{b}$ *de l'espace vectoriel E de dimension n, la matrice M représente un automorphisme f de E,*
3. *la matrice M est la matrice d'un changement de base dans un espace de dimension n,*
4. *le rang de la matrice M est n.*

Preuve : L'équivalence de 1 et 2 a été vue précédemment, celle de 3 et 2 a été vue lors de la définition des matrices de passage. Enfin, la dernière assertion affirme simplement qu'une application linéaire représentée par la matrice M est surjective, et donc un isomorphisme vu que les espaces de départ et d'arrivée ont même dimension. □

3.5. Matrices particulières

3.5.1. Matrice diagonale

Définition 3.8 *La matrice carrée $M = (m_{ij})_{1\leq i,j\leq n}$ est dite* diagonale *si m_{ij} est nul dès que $i \neq j$.*

△ Une matrice diagonale est donc de la forme

$$\begin{pmatrix} d_1 & 0 & \dots & \dots & 0 \\ 0 & d_2 & 0 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & 0 & d_{n-1} & 0 \\ 0 & \dots & \dots & 0 & d_n \end{pmatrix}$$

qu'on notera $\mathrm{Diag}(d_1, d_2, \dots, d_n)$. ▽

Proposition 3.10 (Interprétation vectorielle) *Soit E, muni d'une base $\mathbf{b}$. La matrice de l'endomorphisme f de E relativement à la base $\mathbf{b}$ est diagonale si et seulement si l'image par f de tout vecteur b_i de $\mathbf{b}$ est colinéaire à b_i.*

Preuve : Dire que $f(b_i)$ est colinéaire est équivalent à affirmer l'existence d'un scalaire d_i tel que $f(b_i) = d_i b_i$. □

Exemple 3.5 Soient U et V sous-espaces en somme directe de E. La projection p sur U parallèlement à V est représenté par une matrice diagonale dans une base constituée d'une base de U et d'une base de V.

L'homothétie de rapport λ est représentée par la matrice diagonale $\mathrm{Diag}(\lambda, \dots, \lambda)$ dans toute base de E.

La matrice diagonale avec des 1 sur la diagonale, sauf sur la i-ème ligne, est dite *affinité*.

Proposition 3.11 *(i) L'ensemble des matrices diagonales d'ordre n est un sous-espace vectoriel de $\mathcal{M}_n$ de dimension n, stable par multiplication et transposition.*

(ii) Une matrice diagonale est inversible si et seulement si tous ses coefficients diagonaux sont non nuls. L'inverse d'une matrice diagonale inversible M est diagonale, à coefficients inverses de ceux de M.

Preuve : La première partie résulte directement de la définition. Une base est par exemple donnée par $(E_{11}, \dots, E_{nn})$ où E_{ii} a été introduite dans le paragraphe 3.1. Si le coefficient d_i de la matrice $D = \mathrm{Diag}(d_1, \dots, d_n)$ est nul, alors le noyau de D contient au moins le i-ème vecteur colonne et la matrice n'est pas inversible. Si tous les d_i sont non nuls, la matrice diagonale $\mathrm{Diag}(d_1^{-1}, d_2^{-1}, \dots, d_n^{-1})$ est bien définie et est inverse de $\mathrm{Diag}(d_1, d_2, \dots, d_n)$. □

3.5.2. Matrice triangulaire

Définition 3.9 *La matrice carrée $M = (m_{ij})_{1 \le i,j \le n}$ est dite* triangulaire supérieure *(inférieure resp.) si m_{ij} est nul dès que $i > j$ (resp $i < j$).*

Une matrice triangulaire supérieure ou inférieure avec tous ses coefficients diagonaux nuls est dite nilpotente.

La définition des matrices triangulaires supérieure (U, comme *Upper*) et inférieure (L comme *Lower*) est facilement rendue par la représentation picturale des matrices :

$$U = \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ 0 & m_{22} & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & m_{nn} \end{pmatrix}, \quad L = \begin{pmatrix} m_{11} & 0 & \dots & 0 \\ m_{21} & m_{22} & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ m_{n1} & \dots & \dots & m_{nn} \end{pmatrix}$$

Proposition 3.12 (Interprétation vectorielle) *Soit E un espace vectoriel muni d'une base $\mathbf{b}$ et E_i la suite des sous-espaces $E_0 = \{0\}, E_1 = \mathrm{Vect}(b_1), E_2 = \mathrm{Vect}(b_1, b_2), \dots, E_{n-1} = \mathrm{Vect}(b_1, \dots, b_{n-1}), E_n = E$. La matrice d'un endomorphisme f de E relativement à la base $\mathbf{b}$ est triangulaire supérieure si et seulement si*

$$f(E_i) \subset E_i, \quad i = 1, \dots, n-1.$$

Preuve : Le i-ème vecteur colonne de la représentation matricielle M de l'endomorphisme f relativement à la base $\mathbf{b}$ est le vecteur des coordonnées de $f(b_i)$ dans la base $\mathbf{b}$. Ainsi, si M est triangulaire supérieure, $f(b_i)$ est combinaison linéaire de $b_1, \ldots, b_i$ et donc $f(E_i) \subset E_i$. La réciproque va de soi. □

Proposition 3.13 *L'ensemble des matrices triangulaires supérieures d'ordre n est un sous-espace vectoriel de dimension $n(n+1)/2$, stable par multiplication. La transposée d'une matrice triangulaire supérieure est une matrice triangulaire inférieure.*

Preuve : Commençons par remarquer que la transposition induit un isomorphisme de l'espace des matrices triangulaires supérieures sur celui des matrices triangulaires inférieures.

Une base de l'espace des matrice triangulaires supérieures (inférieures resp.) est la famille des E_{ij} avec $i \leq j$ (resp. $i \geq j$), d'où la dimension $n(n+1)/2$.

Si M_1, M_2 sont deux matrices triangulaires supérieures représentant les endomorphismes f_1, f_2 dans la base canonique de $\mathbb{C}^n$, on a, avec les notations précédentes,

$$f_1 \circ f_2(E_i) \subset f_1(E_i) \subset E_i$$

ce qui montre que $M_1 M_2$, matrice représentant $f_1 \circ f_2$ est triangulaire supérieure. Si M_1, M_2 sont triangulaires inférieures, leur produit $M_1 M_2$ (transposé de ${}^tM_2\,{}^tM_1$ où ${}^tM_2$ et ${}^tM_1$ sont triangulaires supérieures) l'est aussi. □

Proposition 3.14 *(i) Si M est une matrice triangulaire nilpotente d'ordre n, $M^k = 0$ pour $k \geq n$.*

(ii) Si M est une matrice triangulaire nilpotente d'ordre n, $I_n - M$ est inversible et

$$(I_n - M)^{-1} = I_n + M + M^2 + M^3 + \ldots + M^{n-1}\,.$$

(iii) Une matrice triangulaire supérieure (inférieure resp.) est inversible si et seulement si tous ses coefficients diagonaux sont non nuls et l'inverse d'une matrice triangulaire supérieure (inférieure resp.) inversible est une matrice triangulaire supérieure (inférieure resp.).

Preuve : Si M est nilpotente et si f est l'endomorphisme représenté par M dans la base canonique de $\mathbb{C}^n$, on a $f(E_i) \subset E_{i-1}$ et, par suite, $f^k(E_i) \subset E_{i-k}$, soit $f^k(E_i) \subset \{0\}$ dès que $k \geq i$. On en déduit $f^n = 0$, soit $M^n = 0$.

On a

$$(I_n - M)(I_n + M + M^2 + \ldots + M^k) = I_n - M^{k+1}\,,$$

ainsi pour $k \geq n-1$,

$$(I_n - M)(I_n + M + M^2 + \ldots + M^k) = I_n$$

et $I_n + M + M^2 + \ldots + M^k$ est l'inverse de $I_n - M$.

Si U triangulaire supérieure est à coefficients diagonaux tous non nuls, la matrice $D = \mathrm{Diag}(u_{11}, \ldots, u_{nn})$ est inversible et $D^{-1}U$ est de la forme $I_n - N$ avec N nilpotente. S'en déduit l'inversibilité de $U = D(I_n - N)$, dont l'inverse est donné par

$$(I_n - N)^{-1} D^{-1} = (I_n + N + N^2 + \ldots + N^n) D^{-1}$$

qui est triangulaire supérieure.

Si le coefficient diagonal u_{ii} est nul, alors

$$E_i = \mathrm{Vect}(C_1, \ldots, C_i) = \mathrm{Vect}(f(b_1), \ldots, f(b_i)) \subset E_{i-1}$$

et

$$\mathrm{Vect}(C_1,\dots,C_n) \subset E_{i-1} + \mathrm{Vect}(f(b_{i+1}),\dots,f(b_n))$$

est de rang au plus $n-1$ et ainsi M est non inversible, comme l'affirme la seconde partie de la seconde assertion.

Les affirmations concernant les matrices triangulaires inférieures se déduisent de celles sur les matrices triangulaires supérieures via la transposition. □

3.5.3. Matrice de permutation

Définition 3.10 *Soit σ une permutation de $\{1,\dots,n\}$. La matrice de permutation P_σ est la matrice représentant dans la base canonique $\mathbf{c}$ de $\mathbb{C}^n$ l'automorphisme transformant le vecteur c_i en le vecteur $c_{\sigma(i)}$ pour $i=1,\dots,n$.*

Une matrice de permutation a comme coefficients des 0 et des 1, chaque colonne (et chaque ligne) ne contenant qu'un seul 1.

Exemple 3.6 Si $\sigma : (1,2,3) \to (2,3,1)$, la matrice de permutation est

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

Proposition 3.15 *Soient σ, ψ des permutations. Alors $P_{\sigma\circ\psi} = P_\sigma P_\psi$ et $P_\sigma^{-1} = P_{\sigma^{-1}} = {}^tP_\sigma$.*

Preuve : On a $P_{\sigma\circ\psi}(c_i) = c_{\sigma\circ\psi(i)} = c_{\sigma(\psi(i))} = P_\sigma(c_{\psi(i)}) = P_\sigma(P_\psi c_i)$ □

Exemple 3.7 Si, pour $i \neq j$, τ_{ij} est la transposition qui échange i et j en laissant les autres entiers fixes, la matrice de permutation associée, notée aussi $P_{(i,j)}$, est sa propre inverse, vu que $\tau_{ij}^2 = 1$.

4. Déterminants

Les déterminants fournissent une méthode effective du test de l'inversibilité d'une matrice et de la résolution d'un système linéaire. Leur intérêt est malgré tout souvent de nature théorique, car leur calcul effectif requiert rapidement beaucoup d'opérations élémentaires (multiplications et additions de scalaires) : pour $n = 10$, on ainsi 3.6 millions ($\simeq 10\,!$) opérations nécessaires pour le calcul du déterminant d'une matrice d'ordre n générale.

4.1. Déterminants d'ordre 2 ou 3

Définition 4.1 *Soit* $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *une matrice carrée d'ordre 2. Le* déterminant de M *est le scalaire, noté* $\text{dét}\, M$*, défini par*

$$\text{dét}\, M = ad - bc\,.$$

On notera aussi

$$\text{dét} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

Soit $N = \begin{pmatrix} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \\ \alpha'' & \beta'' & \gamma'' \end{pmatrix}$ *une matrice carrée d'ordre 3. Le* déterminant de N *est le scalaire, noté* $\text{dét}\, N$*, défini par*

$$\text{dét}\, N = \alpha \begin{vmatrix} \beta' & \gamma' \\ \beta'' & \gamma'' \end{vmatrix} - \beta \begin{vmatrix} \alpha' & \gamma' \\ \alpha'' & \gamma'' \end{vmatrix} + \gamma \begin{vmatrix} \alpha' & \beta' \\ \alpha'' & \beta'' \end{vmatrix}\,.$$

On notera aussi

$$\text{dét} \begin{pmatrix} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \\ \alpha'' & \beta'' & \gamma'' \end{pmatrix} = \begin{vmatrix} \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \\ \alpha'' & \beta'' & \gamma'' \end{vmatrix}\,.$$

Proposition 4.1 *Soit la fonction* dét *définie sur* $\mathcal{M}_n$ *pour* $n = 2$ *ou* 3*.*

1. *Si on considère le déterminant comme fonction de la* j*-ème colonne ou de la* i*-ème ligne, on a une forme linéaire sur* $\mathbb{C}^n$*.*
2. *Si on considère le déterminant comme fonction des vecteurs colonnes* $C_1, C_2, \ldots, C_n$*, alors*

$$\text{dét}(C_1, \ldots, C_{j_0}, \ldots, C_{j_1}, \ldots, C_n) = -\,\text{dét}(C_1, \ldots, C_{j_1}, \ldots, C_{j_0}, \ldots, C_n)$$

$$\text{dét}(C_1, \ldots, C_{j_0}, \ldots, C_{j_1}, \ldots, C_n) = 0 \text{ si } C_{j_0} = C_{j_1}$$

   On dit que la fonction dét *est une* fonction alternée*.*

3. $\text{dét}\, {}^t M = \text{dét}\, M$*.*
4. $\text{dét}(MM') = \text{dét}\, M\, \text{dét}\, M'$*.*
5. $\text{dét}\, \text{Diag}(\lambda_1, \ldots, \lambda_n) = \lambda_1 \ldots \lambda_n$*. Si* M *est triangulaire,* $\text{dét}\, M = m_{11} \ldots m_{nn}$*.*

Preuve : On ne fait les preuves qu'en dimension 2.

Pour la linéarité, on vérifie que l'expression $ad - bc$ du déterminant a les linéarités partielles relativement aux colonnes et lignes : par ex., pour la deuxième colonne, il faut fixer a et c...

Le caractère alterné par rapport aux colonnes

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

est clair, de même que par rapport aux lignes. L'invariance par transposition est immédiate

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

Quant au déterminant du produit de deux matrices, il résulte d'une identité algébrique (polynômiale en 8 variables) aisément vérifiable

$$\begin{aligned} \text{dét}\, M \,\text{dét}\, M' &= (ad - bc)(a'd' - b'c') \\ &= [aa' + bc'][cb' + dd'] - [ab' + bd'][ca' + dc'] = \text{dét}\, MM' \,. \end{aligned}$$ □

On en déduit

$$\text{dét}(C_1, C_2 + \lambda C_1) = \text{dét}(C_1, C_2)$$

en utilisant la linéarité par rapport à la deuxième colonne et le caractère alterné

$$\text{dét}(C_1, C_2 + \lambda C_1) = \text{dét}(C_1, C_2) + \lambda\, \text{dét}(C_1, C_1) = \text{dét}(C_1, C_2)\,.$$

En fait, la matrice $(C_1, C_2 + \lambda C_1)$ s'obtient en multipliant à droite la matrice (C_1, C_2) par la matrice $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$.

On retiendra qu'on peut ajouter à un vecteur colonne d'une matrice une combinaison linéaire des autres vecteurs colonnes sans changer la valeur de son déterminant. Vu que la transposition laisse invariant le déterminant, il en est de même pour les lignes : ajouter à une ligne une combinaison linéaire des autres lignes ne change pas la valeur du déterminant (cet ajout de multiple scalaire de ligne, équivalent à la multiplication à gauche par la matrice $I_2 + \lambda E_{12}$ ou $I_2 + \lambda E_{21}$, correspond à une étape dans la réduction de Gauß des matrices ou des systèmes).

Exemple 4.1 Le calcul suivant du déterminant de Cauchy[4] d'ordre 2 est effectué en utilisant les propriétés du déterminant (soustraction d'une ligne à une autre, linéarité partielle, valeur du déterminant d'une matrice triangulaire), sans utiliser la formule initiale le définissant : c'est l'archétype de calculs de déterminants d'ordre n (pour des matrices particulières, au sens où ils aboutissent rapidement, le calcul d'un déterminant *général* demandant beaucoup d'opérations).

$$\begin{aligned}
\begin{vmatrix} \frac{1}{x_1+y_1} & \frac{1}{x_1+y_2} \\ \frac{1}{x_2+y_1} & \frac{1}{x_2+y_2} \end{vmatrix} &= \begin{vmatrix} \frac{1}{x_1+y_1} & \frac{1}{x_1+y_2} - \frac{1}{x_1+y_1} \\ \frac{1}{x_2+y_1} & \frac{1}{x_2+y_2} - \frac{1}{x_2+y_1} \end{vmatrix} \\
&= \begin{vmatrix} 1 & \frac{y_1-y_2}{(x_1+y_1)(x_2+y_1)(x_1+y_2)} \\ 1 & \frac{y_1-y_2}{(x_1+y_1)(x_2+y_1)(x_2+y_2)} \end{vmatrix} \\
&= \frac{y_1 - y_2}{(x_1+y_1)(x_2+y_1)} \begin{vmatrix} 1 & \frac{1}{x_1+y_2} \\ 1 & \frac{1}{x_2+y_2} \end{vmatrix} \\
&= \frac{y_1 - y_2}{(x_1+y_1)(x_2+y_1)} \begin{vmatrix} 1 & \frac{1}{x_1+y_2} \\ 0 & \frac{1}{x_2+y_2} - \frac{1}{x_1+y_2} \end{vmatrix} \\
&= \frac{(y_1 - y_2)(x_1 - x_2)}{(x_1+y_1)(x_2+y_1)(x_1+y_2)(x_2+y_2)}
\end{aligned}$$

[4] Augustin Louis Cauchy, 21 août 1789, Paris, France – 23 mai 1857, Sceaux, France.

Le déterminant de Vandermonde[5] d'ordre 3 se calcule tout aussi aisément.

$$\begin{aligned}\begin{vmatrix}1 & x & x^2\\ 1 & y & y^2\\ 1 & z & z^2\end{vmatrix} &= \begin{vmatrix}1 & x & x^2\\ 0 & y-x & y^2-x^2\\ 0 & z-x & z^2-x^2\end{vmatrix} = (y-x)(z-x)\begin{vmatrix}1 & x & x^2\\ 0 & 1 & y+x\\ 0 & 1 & z+x\end{vmatrix}\\ &= (y-x)(z-x)\begin{vmatrix}1 & x & x^2\\ 0 & 1 & y+x\\ 0 & 0 & z-y\end{vmatrix} = (y-x)(z-x)(z-y)\end{aligned}$$

Proposition 4.2 *La matrice M d'ordre 2 est inversible si et seulement si* $\text{dét}\, M$ *est non nul.*

Si la matrice $M = \begin{pmatrix} a & b\\ c & d\end{pmatrix}$ est inversible,

$$\begin{pmatrix} a & b\\ c & d\end{pmatrix}^{-1} = \frac{1}{\text{dét}\, M}\begin{pmatrix} d & -b\\ -c & a\end{pmatrix}.$$

Preuve : Si la matrice M est inversible, alors $\text{dét}\, M\, \text{dét}\, M^{-1} = \text{dét}(MM^{-1}) = \text{dét}\, 1_2 = 1$ et par suite $\text{dét}\, M$ est non nul. Réciproquement, si le déterminant $\text{dét}\, M$ est non nul, on vérifie que la matrice de l'énoncé est un inverse (et donc l'inverse) pour M. □

4.2. Déterminants d'ordre n

Définition 4.2 *Soit M une matrice carrée d'ordre n. La* matrice cofacteur du coefficient *m_{ij} est la matrice M_{ij} carrée d'ordre $n-1$ obtenue à partir de M en supprimant la ligne i et la colonne d'indice j (auxquelles appartient le coefficient m_{ij}).*

$$M_{ij} = \begin{pmatrix} m_{11} & \dots & m_{1j-1} & m_{1j+1} & \dots & m_{1n}\\ \vdots & & \vdots & \vdots & & \vdots\\ m_{i-11} & \dots & m_{i-1j-1} & m_{i-1j+1} & \dots & m_{i-1n}\\ m_{i+11} & \dots & m_{i+1j-1} & m_{i+1j+1} & \dots & m_{i+1n}\\ \vdots & & \vdots & \vdots & & \vdots\\ m_{n1} & \dots & m_{nj-1} & m_{nj+1} & \dots & m_{nn}\end{pmatrix}.$$

Définition 4.3 *Le déterminant* $\text{dét}_n\, M$ *d'une matrice carrée M d'ordre n, noté encore*

$$\text{dét}\, M = \begin{vmatrix} m_{11} & \dots & m_{1n}\\ \vdots & & \vdots\\ m_{n1} & \dots & m_{nn}\end{vmatrix}$$

est l'évaluation sur M de la fonction $\text{dét}_n : \mathcal{M}_n \to \mathbb{C}$ *(notée par abus simplement par* dét*) définie par récurrence de la manière suivante*

- *si $n = 1$,* $\text{dét}_1\, M = m_{11}$.
- *si $n > 1$,*

$$\text{dét}_n\, M = \sum_{j=1}^{n} (-1)^{1+j} m_{1j}\, \text{dét}_{n-1}\, M_{1j}.$$

Cette définition (qui prolonge les définitions données précédemment pour $n = 2$ et 3) est justifiée par le théorème suivant qui sera admis

[5] Alexandre Vandermonde, 28 Février 1735, Paris, France – 1er janvier 1796, Paris, France.

Théorème 4.1 *Soit M une matrice d'ordre n. La valeur du déterminant de M s'obtient en développant par rapport à n'importe quelle ligne*

$$\text{dét}_n M = \sum_{j=1}^{n} (-1)^{i+j} m_{ij} \, \text{dét}_{n-1} M_{ij} \,, \quad i = 1, \dots, n \,,$$

ou en développant par rapport à n'importe quelle colonne

$$\text{dét}_n M = \sum_{i=1}^{n} (-1)^{i+j} m_{ij} \, \text{dét}_{n-1} M_{ij} \,, \quad j = 1, \dots, n \,.$$

Les propriétés du déterminant énoncées précédemment (et démontrées en dimension 2) sont valides en toute dimension :

Proposition 4.3 *Soit* dét *la fonction définie sur $\mathcal{M}_n$.*

0. $\text{dét}\, I_n = 1$.

1. Si on considère le déterminant comme fonction de la j-ème colonne ou de la i-ème ligne, on a une forme linéaire sur $\mathbb{C}^n$.

2. Si on considère le déterminant comme fonction des vecteurs colonnes $C_1, C_2, \dots, C_n$, alors

$$\text{dét}(C_1, \dots, C_{j_0}, \dots, C_{j_1}, \dots, C_n) = -\, \text{dét}(C_1, \dots, C_{j_1}, \dots, C_{j_0}, \dots, C_n) \,,$$

$$\text{dét}(C_1, \dots, C_{j_0}, \dots, C_{j_1}, \dots, C_n) = 0 \ \textit{si} \ C_{j_0} = C_{j_1} \,.$$

Comme fonction des vecteurs ligne,

$$\text{dét}\, {}^t(L_1, \dots, L_{i_0}, \dots, L_{i_1}, \dots, L_n) = -\, \text{dét}(L_1, \dots, L_{i_1}, \dots, L_{i_0}, \dots, L_n) \,,$$

$$\text{dét}(L_1, \dots, L_{i_0}, \dots, L_{i_1}, \dots, L_n) = 0 \ \textit{si} \ L_{i_0} = L_{i_1} \,.$$

On dit que la fonction dét *est* alternée*.*

3. $\text{dét}\, {}^t M = \text{dét}\, M$.

4. $\text{dét}(MM') = \text{dét}\, M \, \text{dét}\, M'$.

5. $\text{dét}\, \text{Diag}(\lambda_1, \dots, \lambda_n) = \lambda_1 \dots \lambda_n$. *Si M est triangulaire,* $\text{dét}\, M = m_{11} \dots m_{nn}$.

Les propriétés 0, 1 et 2 caractérisent la fonction déterminant.

4.3. Inversion d'une matrice

Définition 4.4 *Soit M une matrice d'ordre n. La* matrice des cofacteurs *de M est la matrice, notée* $\text{Cof}\, M$, *définie par*

$$(\text{Cof}\, M)_{ij} = (-1)^{i+j} \, \text{dét}\, M_{ij}$$

où M_{ij} est le cofacteur du coefficient m_{ij}.

La comatrice de M *est la matrice, notée* $\text{Com}\, M$, *définie par* $\text{Com}\, M = {}^t\, \text{Cof}\, M$.

Proposition 4.4 *Soit M une matrice d'ordre n. Alors*

$$M(\text{Com}\, M) = (\text{Com}\, M)M = (\text{dét}\, M) I_n \,.$$

Preuve : C'est en fait une reformulation du développement du déterminant en lignes ou colonnes (cf Thm. 4.1). La relation $M(\operatorname{Com} M) = (\operatorname{dét} M)I_n$ est équivalente aux égalités

$$\sum_{k=1}^{n} m_{ik}(-1)^{k+j} \operatorname{dét} M_{jk} = \begin{cases} \operatorname{dét} M \text{ si } i = j, \\ 0 \text{ sinon.} \end{cases}$$

Pour $i = j$, cette égalité est le développement du déterminant $\operatorname{dét} M$ par rapport la i-ème ligne. Pour $i \neq j$, on a le développement par rapport la j-ème ligne de la matrice $M(i,j)$ déduite de M en remplaçant la j-ème ligne par la i-ème ligne, dont le déterminant est nul puisque la matrice $M(i,j)$ a deux lignes (la i-ème et la j-ème) identiques.

L'égalité $(\operatorname{Com} M)M = (\operatorname{dét} M)I_n$ s'obtient de manière analogue en invoquant des développements de déterminant par rapport à des colonnes. □

Théorème 4.2 *Soit M matrice carrée d'ordre n. La matrice M est inversible si et seulement si son déterminant* $\operatorname{dét} M$ *est non nul.*

Si M est inversible, alors

$$M^{-1} = (\operatorname{dét} M)^{-1} \operatorname{Com} M\,, \quad \operatorname{dét}(M^{-1}) = (\operatorname{dét} M)^{-1}\,.$$

Preuve : Si M est inversible, d'inverse $M' = M^{-1}$, alors $\operatorname{dét} M \operatorname{dét} M' = \operatorname{dét} I_n = 1$ et donc $\operatorname{dét} M$ est non nul. Inversement, si M a un déterminant non nul, alors la formule de l'énoncé convient pour l'inverse, d'après la proposition précédente. □

4.4. Déterminants d'une famille de vecteurs et d'un endomorphisme

Définition 4.5 *Soit E un espace vectoriel de dimension n, muni d'une base* $\mathbf{b}$. *Le* déterminant $\operatorname{dét}_{\mathbf{b}}(v_1, \ldots, v_n)$ de la famille ordonnée de vecteurs $(v_1, \ldots, v_n)$ relativement à la base $\mathbf{b}$ *est défini comme étant le déterminant de la matrice $M = (m_{ij})$ des vecteurs colonne C_j des coordonnées des vecteurs v_j dans la base* $\mathbf{b}$, i. e. $C_j = {}^t(m_{1j}, \ldots, m_{nj})$ *avec* $v_j = m_{1j}b_1 + \ldots + m_{nj}b_n$ *pour* $j = 1, \ldots, n$.

△ Dans $\mathbb{C}^n$ muni de sa base canonique $\mathbf{c}$, le déterminant $\operatorname{dét}_{\mathbf{c}}(v_1, \ldots, v_n)$ des n vecteurs $v_1, \ldots, v_n$ coïncide avec celui du tableau matriciel $V(v_1, \ldots, v_n)$ des coordonnées de ces vecteurs dans la base canonique.

Si A est une matrice d'ordre n représentant un endomorphisme de $\mathbb{C}^n$ noté pareillement et $v_1, \ldots, v_n$ n vecteurs de $\mathbb{C}^n$, vu que $V(Av_1, \ldots, Av_n) = AV(v_1, \ldots, v_n)$, il résulte de la propriété 4 de la Prop. 4.3 que

$$\operatorname{dét}_{\mathbf{c}}(Av_1, \ldots, Av_n) = \operatorname{dét} A \operatorname{dét}_{\mathbf{c}}(v_1, \ldots, v_n)\,. \quad ▽$$

Théorème 4.3 *Soit E un espace vectoriel de dimension n, muni d'une base* $\mathbf{b}$. *La famille de vecteurs $(v_1, \ldots, v_n)$ est une base de E si et seulement si* $\operatorname{dét}_{\mathbf{b}}(v_1, \ldots, v_n)$ *est non nul.*

Preuve : Le déterminant $\operatorname{dét}_{\mathbf{b}}(v_1, \ldots, v_n)$ peut être interprété comme celui de la matrice M représentant l'endomorphisme f de $\mathbb{C}^n$ qui au vecteur e_i de la base canonique associe le vecteur colonne des coordonnées de v_i dans la base $\mathbf{b}$. La matrice M est inversible si et seulement si l'application linéaire f est un isomorphisme, ainsi le déterminant $\operatorname{dét} M = \operatorname{dét}_{\mathbf{b}}(v_1, \ldots, v_n)$ est non nul si et seulement si la famille (v_i) est une base. □

Proposition/Définition 4.1 *Soit f un endomorphisme de E. Le déterminant* $\text{dét}\, M_{\mathbf{b}}(f)$ *de la matrice $M_{\mathbf{b}}(f)$ représentant f dans la base $\mathbf{b}$ de E ne dépend pas de la base $\mathbf{b}$. Sa valeur est appelée le déterminant de l'endomorphisme f.*

Preuve : Si $P_{\mathbf{b},\widetilde{\mathbf{b}}}$ note la matrice de passage de la base $\mathbf{b}$ à la base $\widetilde{\mathbf{b}}$,

$$M_{\mathbf{b}}(f) = P^{-1}_{\mathbf{b},\widetilde{\mathbf{b}}} M_{\widetilde{\mathbf{b}}}(f) P_{\mathbf{b},\widetilde{\mathbf{b}}}$$

et par suite

$$\text{dét}\, M_{\mathbf{b}}(f) = \text{dét}\, P^{-1}_{\mathbf{b},\widetilde{\mathbf{b}}}\, \text{dét}\, M_{\widetilde{\mathbf{b}}}(f)\, \text{dét}\, P_{\mathbf{b},\widetilde{\mathbf{b}}} = \text{dét}\, M_{\widetilde{\mathbf{b}}}(f)\,,$$

ce qui assure l'indépendance de $\text{dét}\, M_{\mathbf{b}}(f)$ vis à vis du choix de la base $\mathbf{b}$. □

4.5. Rang

Proposition 4.5 *Soit A une matrice carrée d'ordre r et $\widetilde{A}$ obtenue de A par échange de lignes et colonnes. Alors* $\text{dét}\, A = \varepsilon\, \text{dét}\, \widetilde{A}$ *avec $\varepsilon^2 = 1$.*

Preuve : L'échange de deux lignes ou deux colonnes transforme le déterminant en son opposé, ainsi le résultat final est la multiplication par 1 ou -1. □

△ Une matrice de permutation a pour déterminant 1 ou -1. ▽

Définition 4.6 *Soit $I = \{i_1, \ldots, i_m\} \subset \{1, \ldots, n\}$ et $J = \{j_1, \ldots, j_q\} \subset \{1, \ldots, p\}$. Si A est une matrice de $\mathcal{M}_{np}$, la matrice $A_{IJ} = (a_{i_k j_l})_{1\leq k\leq m, 1\leq l\leq q}$ est dite* matrice extraite *de A. Si A_{IJ} est carrée, son déterminant est dit* déterminant extrait *de A.*

△ La matrice A_{IJ} est obtenue à partir de A en enlevant toutes les lignes (et colonnes) d'indice hors de I (resp. J). ▽

Théorème 4.4 *Le rang r d'une matrice M de $\mathcal{M}_{np}$ est l'ordre maximal d'un déterminant extrait de M qui soit non nul ou d'une matrice carrée extraite inversible.*

Preuve : Supposons que M ait un déterminant $\Delta_{\widetilde{r}}$ extrait d'ordre $\widetilde{r}$ non nul. À permutation des lignes et colonnes près, on peut supposer que $\Delta_{\widetilde{r}}$ est le déterminant de la matrice déduite de M en considérant les $\widetilde{r}$ premières lignes et $\widetilde{r}$ premières colonnes, soit $M_{\widetilde{r}} = (m_{ij})_{1\leq i,j\leq \widetilde{r}}$. La matrice $M_{\widetilde{r}}$ de déterminant $\text{dét}\, M_{\widetilde{r}} = \Delta_{\widetilde{r}}$ non nul est inversible, ses vecteurs colonnes $\widetilde{C_j} = {}^t(a_{1j}, \ldots, a_{\widetilde{r}j}), j = 1, \ldots, \widetilde{r}$ sont indépendants, et donc aussi par suite les vecteurs $C_j = {}^t(a_{1j}, \ldots, a_{nj}), j = 1, \ldots, \widetilde{r}$: M est de rang au moins $\widetilde{r}$, soit $r \geq \sup_{\Delta_{\widetilde{r}}\neq 0} \widetilde{r}$.

Supposons M de rang r. À permutation des colonnes près, on peut supposer que les vecteurs colonnes $(C_1, \ldots, C_r)$ (vecteurs de $\mathbb{C}^n$) sont indépendants. Le rang de la matrice

$$\begin{pmatrix} a_{11} & \ldots & a_{1r} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nr} \end{pmatrix}$$

est aussi celui de ses lignes, donc à permutation des lignes près, on peut supposer que les r premières lignes (vecteurs de $\mathbb{C}^r$) sont indépendantes, ainsi on a une matrice carrée

$$\begin{pmatrix} a_{11} & \ldots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \ldots & a_{rr} \end{pmatrix}$$

dont les r lignes sont indépendantes : ces r lignes, vecteurs de $\mathbb{C}^r$, constituent une base de $\mathbb{C}^r$ et donc la matrice A_r est inversible, donc de déterminant non nul. La matrice A a donc un déterminant extrait d'ordre r non nul, soit, avec ce qui précède, $r = \sup_{\Delta_{\widetilde{r}}\neq 0} \widetilde{r}$. □

4.6. Aire et volume

Définition 4.7 *Le parallélotope P engendré par n vecteurs $(v_1, \ldots, v_n)$ d'un espace vectoriel E est l'ensemble $P = \{\lambda_1 v_1 + \ldots + \lambda_n v_n, \lambda_i \in [0,1]\}$.*

Soient n vecteurs $v_1, \ldots, v_n$ indépendants de $\mathbb{R}^n$. Le volume $\mathrm{Vol}(P)$ du parallélotope P engendré par ces n vecteurs est défini comme la valeur absolue $|\,\mathrm{dét}(v_1, \ldots, v_n)|$ du déterminant de ces n vecteurs relativement à la base canonique.

$\triangle$ Si $n = 1$, le vecteur $v_1 \in \mathbb{R}$ détermine le segment de longueur $|v_1|$, le scalaire-vecteur 1 représentant la longueur unité.

Si $n = 2$, un parallélotope est l'habituel *parallélogramme* et le volume, nommé *aire*, du parallélogramme engendré par (ae_1, be_2) (qui est un rectangle dans une représentation où les vecteurs (e_1, e_2) sont considérés comme orthogonaux, de longueur unité) est ab si (e_1, e_2) est la base canonique. Les vecteurs (v_1, v_2) déterminent le parallélogramme d'aire $|\,\mathrm{dét}(v_1, v_2)|$, de même aire que le parallélotope déterminé par $(v_1, v_2 + \alpha v_1)$, qui est un “rectangle” pour α convenable.

Si $n = 3$, on obtient un parallélépipède.

En général, si $(v_1, \ldots, v_n)$ sont dépendants, le déterminant $\mathrm{dét}(v_1, \ldots, v_n)$ est nul d'une part, le parallélogramme engendré par ces vecteurs est dans un sous-espace strict de $\mathbb{R}^n$ et donc a un volume n-dimensionnel nul d'autre part. $\triangledown$

Proposition 4.6 *Soit AP le parallélotope image du parallélotope P par l'endomorphisme A de $\mathbb{R}^n$. Alors $\mathrm{Vol}(AP) = |\,\mathrm{dét}\, A|\,\mathrm{Vol}(P)$.*

Preuve : L'image AP est un parallélotope

$$AP = A\{\lambda_1 v_1 + \ldots + \lambda_n v_n, \lambda_i \in [0,1]\} = \{\lambda_1 A v_1 + \ldots + \lambda_n A v_n, \lambda_i \in [0,1]\}$$

de volume

$$\mathrm{Vol}(AP) = |\,\mathrm{dét}(Av_1, \ldots, Av_n)| = |\,\mathrm{dét}\, A\, \mathrm{dét}(v_1, \ldots, v_n)| = |\,\mathrm{dét}\, A|\,\mathrm{Vol}(P)\,. \qquad \square$$

Exemple 4.2 Si $A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, le parallélogramme AP a même aire que le parallélogramme P.

4.7. Systèmes linéaires

Le système linéaire

$$(S) \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2p}x_p = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{np}x_p = b_n \end{cases}$$

peut être interprété sous forme matricielle

$$(S) \qquad AX = B\,, \quad X \in \mathbb{C}^p\,,$$

avec $A = (a_{ij}) \in \mathcal{M}_{np}$ et $B = {}^t(b_1, \ldots, b_n) \in \mathbb{C}^n$ donnés et $X = {}^t(x_1, \ldots, x_p)$ le vecteur colonne des inconnues, ou en termes vectoriels

$$(S) \qquad f(v) = b\,, \quad v \in E\,,$$

avec E espace vectoriel isomorphe à $\mathbb{C}^p$, F espace vectoriel isomorphe à $\mathbb{C}^n$ et f application linéaire de E dans F (f, b, v correspondant à A, B, X par le dictionnaire induit par le choix de bases dans E et F).

Définition 4.8 *Le système* (S) $\quad AX = B$ *est dit* homogène *si* $B = 0$.

Proposition 4.7 *L'ensemble des solutions du système homogène*

$$(H) \qquad AX = 0\,, \quad X \in \mathbb{C}^p\,,$$

est un sous-espace vectoriel de dimension $p - \operatorname{rg} A$.

Preuve : Cela résulte aisément de l'interprétation opérationnelle, où l'ensemble des solutions est vu comme un noyau, de dimension $p - \dim \operatorname{Im} A = p - \operatorname{rg} A$. □

Proposition 4.8 *Si le système*

$$(S) \qquad AX = B\,, \quad X \in \mathbb{C}^p\,,$$

a une solution, l'ensemble des solutions du système (S) *s'obtient en ajoutant une solution particulière de* (S) *à l'ensemble des solutions du système homogène*

$$(H) \qquad AX = 0\,, \quad X \in \mathbb{C}^p\,,$$

associé.

Le cas particuler des systèmes à même nombre d'inconnues que d'équations est spécialement important.

Définition 4.9 *Soient* $A \in \mathcal{M}_{np}$ *et* $B \in \mathcal{M}_{n1}$. *Le système de* n *équations linéaires à* p *inconnues* $(x_1, \ldots, x_p)$

$$AX = B\,, \quad X = {}^t(x_1, \ldots, x_p)\,.$$

est dit système de Cramer[6] *si*

- *le nombre* n *d'équations est égal au nombre* p *d'inconnues,*
- *la matrice* $A = (a_{ij})_{1 \leq i,j \leq n}$ *est inversible.*

Théorème 4.5 *Soit* (S) $\quad AX = B$ *un système de Cramer.*
(i) Le système (S) *a une solution unique donnée par* $X = A^{-1}B$.
(ii) La solution X *du système* (S) *a pour coordonnées*

$$x_i = (\text{dét}\, A)^{-1}\, \text{dét}\, A_i\,, \quad i = 1, \ldots, n \tag{4.1}$$

où A_i *est la matrice d'ordre* n *obtenue à partir de* A *en y remplaçant la colonne d'indice* i *par le vecteur colonne* B.

Preuve : Vu que A est inversible, le système (S) a une solution et une seule donnée par

$$X = A^{-1}B = (\text{dét}\, A)^{-1}\, \text{Com}\, A\ B$$

dont la i-ème coordonnée vaut

$$\begin{aligned} x_i &= (\text{dét}\, A)^{-1} \sum_{k=1}^{n} (\text{Com}\, A)_{ik} b_k \\ &= (\text{dét}\, A)^{-1} \sum_{k=1}^{n} b_k (-1)^{i+k}\, \text{dét}\, A_{ki}\,. \end{aligned}$$

[6]Gabriel Cramer, 31 juillet 1704, Genève, Suisse – 4 janvier 1752, Bagnols-sur-Cèze, France.

où on reconnaît le développement par rapport à la i-ème colonne du déterminant

$$\begin{vmatrix} a_{11} & \dots & a_{1i-1} & b_1 & a_{1i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots & \\ a_{k1} & \dots & a_{ki-1} & b_k & a_{ki+1} & \dots & a_{k-1n} \\ \vdots & & \vdots & \vdots & & \vdots & \\ a_{n1} & \dots & a_{ni-1} & b_n & a_{ni+1} & \dots & a_{nn} \end{vmatrix}.$$

□

En général, le rang r du système (S) est celui de la matrice associée

$$r = \operatorname{rg} A = p - \dim\ker A = \operatorname{rg} f = \dim E - \dim \operatorname{Ker} f\,,$$

où p est le nombre d'inconnues.

Ce rang est égal à la dimension du sous-espace des B tels que le système (S) ait une solution : il est appelé *nombre d'inconnues principales*. L'entier $p-r$ *nombre d'inconnues libres* : une solution générale du système dépend de $p-r$ variables.

Un système général (S) de rang r peut se réduire à un système de Cramer. On permute tout d'abord les lignes de telle manière que les r premières $L_1,\dots,L_r$ (avec $L_i = (a_{i1},\dots,a_{ip})$) soient de rang r. Si $I > r$, pour des scalaires λ_{iI} convenables,

$$L_I = \sum_{i=1}^{r} \lambda_{iI} L_i\,.$$

Une condition nécessaire pour que (S) ait une solution est

$$(C) \qquad b_I = \sum_{i=1}^{r} \lambda_{iI} b_i\,, \quad I = r+1,\dots,n\,.$$

Si ces conditions sont vérifiées, le système (S) est équivalent à

$$(S)' \qquad \sum_{j=1}^{p} a_{ij}x_j = b_i\,, \quad i = 1,\dots,r\,.$$

On permute les colonnes de telle sorte que la matrice $A_r = (a_{ij})_{1\le i,j\le r}$ soit de rang r. Le système $(S)'$ se réécrit

$$(S)'' \qquad A_r\,{}^t(x_1,\dots,x_r) = B_r - \widetilde{A}\,{}^t(x_{r+1},\dots,x_p)\,,$$

avec $\widetilde{A} = (a_{ij})_{1\le i\le r, r+1\le j\le p}$. Ce système est de Cramer avec solution

$${}^t(x_1,\dots,x_r) = A_r^{-1}(B_r - \widetilde{A}\,{}^t(x_{r+1},\dots,x_p))\,.$$

À permutation des variables près, et sous réserve des conditions (C), le système (S) a comme solution $(x_1,\dots,x_p)$ dont les variables $(x_1,\dots,x_r)$ (les *variables principales*) sont déterminées de manière unique une fois les variables $(x_{r+1},\dots,x_p)$ (les *variables libres*) fixées.

L'expression des solutions d'un système à la Cramer (4.1) n'est pas toujours la meilleure façon de le résoudre. L'élimination de Gauss est une méthode systématique (implémentable sur machine) souvent plus efficace : cette réduction à un système échelonné est, en termes matriciels, une mise sous forme triangulaire d'une matrice par des opérations élémentaires qui sont décrites dans l'appendice. Il ne faut pas oublier que des résolutions au cas par cas sont aussi possibles (voire nécessaires dans le cas où les nombres des variables et d'équations diffèrent) ainsi du deuxième exemple de la section suivante.

4.8. Deux exemples de systèmes

Le premier exemple donne lieu à interprétation géométrique.

Exemple 4.3 Soient $u_1, v_1, w_1, u_2, v_2, w_2$ des scalaires (qu'on pourra supposer réels afin de se placer dans l'espace réel de dimension 3) non tous nuls et le système

$$(D) \quad \begin{cases} u_1x & + & v_1y & + & w_1z & = & 0 \\ u_2x & + & v_2y & + & w_2z & = & 0 \end{cases},$$

considérons les déterminants

$$d_1 = \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix}, \quad d_2 = \begin{vmatrix} w_1 & u_1 \\ w_2 & u_2 \end{vmatrix}, \quad d_3 = \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}.$$

La nullité du déterminant

$$\Delta_1 = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \end{vmatrix}$$

donne, par développement par rapport à la première ligne et la deuxième

$$u_1d_1 + v_1d_2 + w_1d_3 = 0\,.$$

En considérant le déterminant

$$\Delta_2 = \begin{vmatrix} u_2 & v_2 & w_2 \\ u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \end{vmatrix},$$

on obtient pareillement

$$u_2d_1 + v_2d_2 + w_2d_3 = 0\,.$$

Si (d_1, d_2, d_3) n'est pas le vecteur nul, alors le système (D) est de rang 2 et son espace de solutions (de dimension $3 - 2 = 1$) est la droite $\mathbb{C}(d_1, d_2, d_3)$, intersection des plans P_i, $i = 1, 2$ d'équation

$$(P_i) \quad u_ix + v_iy + w_iz = 0\,.$$

Sinon, le système est de rang 1, avec espace de solutions de dimension 2 : les deux plans P_1 et P_2 sont confondus, égaux à cet espace de solutions.

Dans le deuxième exemple, des transformations simples permettent de résoudre le système donné.

Exemple 4.4 Considérons le système

$$(A) \quad \begin{cases} (1+a_1)x_1 & + & x_2 & + & \ldots & + & x_n & = & b_1 \\ x_1 & + & (1+a_2)x_2 & + & \ldots & + & x_n & = & b_2 \\ \vdots \\ x_1 & + & x_2 & + & \ldots & + & (1+a_n)x_n & = & b_n \end{cases}$$

En posant $s = x_1 + \ldots + x_n$, le système (A) se réécrit en le système équivalent

$$(A)' \quad \begin{cases} s & + & a_1x_1 & = & b_1 \\ s & + & a_2x_2 & = & b_2 \\ \vdots \\ s & + & a_nx_n & = & b_n \\ s & - & x_1 & - & \ldots & - & x_n & = & 0 \end{cases}$$

Soient $A_i = \prod_{j\neq i} a_j$ et $P = \prod_i a_i$.

Supposons tous les a_i non nuls. En multipliant, pour $i = 1$ à n, la i-ème ligne du système précédent par A_i et la dernière ligne par P, on obtient le système équivalent

$$(A)'' \quad \begin{cases} A_1 s + Px_1 = A_1 b_1 \\ A_2 s + Px_2 = A_2 b_2 \\ \vdots \\ A_n s + Px_n = A_n b_n \\ Ps - Px_1 - \ldots - Px_n = 0 \end{cases}$$

En additionnant les n première lignes à la dernière, on remplace $(A)''$ par le système $(A)'''$ équivalent avec dernière équation $Ds = \sum_{i=1}^n A_i b_i$ où $D = P + \sum_{i=1}^n A_i$. Ainsi, si D est non nul, on obtient $s = D^{-1}\sum_{i=1}^n A_i b_i$ et par suite

$$x_i = \frac{A_i b_i - A_i s}{P} = \frac{b_i - D^{-1}\sum_{i=1}^n A_i b_i}{a_i}.$$

Si D est nul, pour que le système (A) ait une solution, il est nécessaire que $\sum_{i=1}^n A_i b_i$ soit nul, auquel cas (A) a en terme de la variable libre s les solutions

$$x_i = (b_i - s)/a_i, \quad i = 1, \ldots, n.$$

Si I des a_i sont nuls (par exemple $a_1 = \ldots = a_I = 0$), la condition $b_1 = \ldots = b_I$ est nécessaire pour que $(A)'$ ait des solutions et dans ce cas, les solutions dépendent de $n-I+1$ variables principales (soit $x_I, \ldots, x_n$ et une $\widetilde{x}$ parmi les $x_1, \ldots, x_I$) et $I-1$ variables libres (les $x_1, \ldots, x_I$ à l'exception de $\widetilde{x}$)

$$\begin{aligned} x_j &= (b_j - b_1)/a_j, \quad j = I+1, \ldots, n, \\ \widetilde{x} &= b_1 - x_{I+1} - \ldots - x_n - \sum_{i=1, x_i\neq\widetilde{x}}^{I} x_i, \end{aligned}$$

Si $I = 1$, la solution est unique : si a_1 est nul, le déterminant de (A) est calculé aisément comme étant D. C'est vrai en général, puisque si tous les a_i sont non nuls, on a

$$\begin{aligned} \begin{vmatrix} 1+a_1 & 1 & \ldots & 1 \\ 1 & 1+a_2 & \ldots & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \ldots & 1+a_n \end{vmatrix} &= A_1^{-1} \begin{vmatrix} A_1+P & A_1 & \ldots & A_1 \\ 1 & 1+a_2 & \ldots & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \ldots & 1+a_n \end{vmatrix} \\ &= A_1^{-1} D \begin{vmatrix} 1 & 1 & \ldots & 1 \\ 1 & 1+a_2 & \ldots & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \ldots & 1+a_n \end{vmatrix} \\ &= A_1^{-1} D \begin{vmatrix} 1 & 0 & \ldots & 0 \\ 1 & a_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 1 & 0 & \ldots & a_n \end{vmatrix} = D \end{aligned}$$

Dans la deuxième égalité, on aura ajouté à la première ligne L_1 la combinaison linéaire $\sum_{i=2}^n A_i L_i$ et mis en facteur D, puis pour la troisième égalité retranché la première colonne à chacune des autres pour obtenir une matrice triangulaire.

5. Appendice : opérations élémentaires sur les lignes

Il s'agit dans cet appendice[7] de décrire un moyen pratique algorithmique qui remplace le calcul explicite des matrices inversibles Q et P permettant de passer d'une matrice A à la matrice équivalente de la forme $Q^{-1}A\,P = \left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Il est facile de voir que la multiplication d'une matrice A à gauche par une matrice carrée Q^{-1} revient à remplacer les lignes de A par des combinaisons linéaires de ces lignes, tandis que sa multiplication à droite par une matrice carrée P revient à remplacer les colonnes de A par des combinaisons linéaires de ces colonnes. Nous allons donc considérer ici seulement l'un de ces deux types d'opérations, les opérations sur les lignes, qui correspondent à des multiplications à gauche. La description des opérations sur les colonnes serait bien entendu tout à fait analogue.

Les formes réduites auxquelles on peut espérer parvenir en multipliant une matrice A seulement à gauche sont décrites dans la définition suivante.

Définition 5.1 *Soit $A \in \mathcal{M}_{p,n}$ une matrice dont on note a_{kj} les coefficients. On dit que cette matrice est* échelonnée *s'il existe un entier $r \leq p$ et des indices $1 \leq j_1 < j_2 < \cdots < j_r \leq n$ tels que pour $k \leq r$ on ait $a_{kj} = 0$ si $j < j_k$ et $a_{kj_k} \neq 0$, et que pour $k > r$ on ait $a_{kj} = 0$ si $j \leq n$. En outre, une matrice échelonnée est dite* réduite *si de plus pour tout $k \leq r$ on a $a_{\ell j_k} = 0$ si $\ell \neq k$ et $a_{kj_k} = 1$.*

Exemple 5.1 Considérons les matrices

$$A = \begin{pmatrix} 2 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

La matrice A n'est pas échelonnée car $j_3 = 3 \not> 3 = j_2$, mais les deux autres matrices sont échelonnées. La matrice échelonnée B n'est pas réduite car $b_{14} = 1$ alors que b_{34} devrait être le seul coefficient non nul de la colonne 4. Seule la matrice C est échelonnée réduite.

Bien entendu, pour obtenir des matrices équivalentes, il ne faut effectuer que des opérations réversibles. C'est pourquoi nous nous bornerons à considérer les *opérations élémentaires* suivantes :

(a) échange de deux lignes ;

(b) multiplication d'une ligne par un réel non nul ;

(c) addition à une ligne d'un multiple d'une *autre* ligne.

Il est important d'observer que ces opérations correspondent à des multiplications à gauche par des matrices carrées inversibles.

Et en effet, l'échange de la j-ème ligne et de la k-ème ligne s'obtient en multipliant à gauche par la matrice de permutation $P_{(j,k)}$ (qui est sa propre inverse).

De même, la multiplication de la j-ème ligne par le réel $x \neq 0$ s'obtient en multipliant à gauche par la matrice diagonale $D_{j,x} = D(1, \ldots, 1, x, 1, \ldots, 1)$ où x est le j-ème coefficient diagonal : on a $D_{j,x}^{-1} = D_{j,x^{-1}}$.

Et enfin, l'addition à la j-ème ligne de x fois la k-ème ligne (avec $j \neq k$) s'obtient en multipliant à gauche par la matrice triangulaire $T_{j,k,x}$ avec des 1 sur la diagonale et le coefficient sur la j-ème ligne et k-ème colonne non nul valant x, d'inverse $T_{j,k,-x}$.

Ces opérations élémentaires suffisent à effectuer toutes les transformations réversibles puisque, comme on le verra plus loin, toute matrice carrée inversible $Q^{-1} \in \mathcal{M}_p$ peut s'écrire comme un produit de matrices d'opérations élémentaires.

[7] Sans restreindre la généralité, on suppose dans cet appendice les scalaires *réels*.

Proposition 5.1 *Toute matrice $A \in \mathcal{M}_{p,n}$ peut être transformée en une matrice échelonnée réduite par une suite finie d'opérations élémentaires, et la matrice échelonnée réduite ainsi obtenue ne dépend pas du choix de ces opérations élémentaires (unicité de la matrice échelonnée réduite obtenue par opérations élémentaires, aussi appelée* équivalent canonique *de A).*

Preuve : L'existence de la réduction provient simplement de l'*algorithme de Gauss*, que nous rappelons maintenant.

Notons j_1 le plus petit indice d'une colonne non nulle. On peut obtenir $a_{1j_1} \neq 0$ en échangeant (au besoin) deux lignes, puis $a_{1j_1} = 1$ en divisant la première ligne par a_{1j_1}. En soustrayant alors a_{kj_1} fois la première ligne de la k-ème ligne, on fait ainsi apparaître des zéros dans tout le reste de la j_1-ème colonne.

On répète maintenant cette opération en cherchant dans les colonnes formées par les lignes 2 à p le plus petit indice j_2 d'une colonne non nulle, et comme la j_1-ème colonne (réduite aux lignes 2 à p) n'a que des zéros, on a donc $j_2 > j_1$. Par les mêmes opérations élémentaires que ci-dessus, on obtient sans modifier les a_{kj} pour $j < j_2$ une nouvelle matrice où $a_{2j_2} = 1$ est le seul coefficient non nul de la j_2-ème colonne, puis on continue en considérant les colonnes formées par les lignes 3 à p.

On poursuit ainsi la réduction tant que l'on trouve des colonnes non nulles dans les lignes restantes, et on note $r(\leq p)$ le dernier indice d'une ligne non entièrement nulle. On obtient par ce procédé une matrice échelonnée réduite.

Comme nous n'utiliserons pas explicitement la propriété d'unicité de l'équivalent canonique, nous ne ferons qu'en ébaucher la démonstration. Elle résulte simplement de ce que les opérations élémentaires sont réversibles, si bien que si une même matrice A peut être transformée en deux matrices B et C, alors les lignes de B sont des combinaisons linéaires des lignes de C et réciproquement. Si donc B et C sont des matrices échelonnées réduites obtenues à partir d'une même matrice A, on en déduit que les lignes de B sont des lignes de C et réciproquement, et donc que $B = C$. □

Suivant les applications que l'on a en vue, on pourra n'effectuer qu'une partie de la réduction, par exemple en obtenant une matrice échelonnée sans poursuivre jusqu'à la forme réduite. Nous décrivons maintenant quelques-unes des applications de cet algorithme.

5.1. Calcul du rang et construction d'une base

Il est facile de voir qu'avec les notations de la définition donnée plus haut, les r premières lignes d'une matrice échelonnée de $\mathcal{M}_{p,n}$ forment une famille libre de vecteurs de $\mathbb{R}^n$. En effet, si l'on note $L_1, \ldots, L_r$ ces r premières lignes, et si $x_1 L_1 + \cdots + x_r L_r = 0$,alors le j_1-ème coefficient de ce vecteur est $x_1 a_{1j_1} = 0$, d'où $x_1 = 0$ puisque $a_{1j_1} \neq 0$. On voit ensuite que le j_2-ème coefficient vaut $x_2 a_{2j_2} = 0$ d'où $x_2 = 0$, et on montre ainsi successivement que tous les x_k sont nuls.

L'entier r est donc le rang de la matrice échelonnée, et comme deux matrices équivalentes ont même rang, il suffit donc pour déterminer le rang d'une matrice A donnée d'effectuer une réduction de cette matrice à une forme échelonnée équivalente.

De même, si on se donne par leurs coefficients dans une même base $\mathbf{b}$ une famille $\mathcal{C}$ de p vecteurs d'un espace de dimension n, la méthode de réduction permet de déterminer le rang de cette famille ainsi qu'une base du sous-espace $\operatorname{Vect} \mathcal{C}$. Pour cela, disposons les coefficients dans $\mathbf{b}$ des p vecteurs donnés comme les p lignes d'une matrice $A \in \mathcal{M}_{p,n}$. La réduction de cette matrice par des opérations élémentaires sur les lignes remplace les lignes de A par des combinaisons linéaires de ces lignes sans modifier le rang. Les lignes de la matrice modifiée engendrent donc toujours le même sous-espace $\operatorname{Vect} \mathcal{C}$ à chaque étape. En parvenant à une matrice échelonnée, on obtient donc le rang r de la famille $\mathcal{C}$, et les r premières lignes de cette dernière matrice fournissent les coefficients dans $\mathbf{b}$ d'une base de $\operatorname{Vect}\mathcal{C}$.

Si par exemple $\mathcal{C} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ avec $\mathbf{v}_1 = (1,2,-1,2)$, $\mathbf{v}_2 = (2,5,-1,4)$ et $\mathbf{v}_3 = (0,3,3,0)$, on effectue la réduction

$$\begin{pmatrix} 1 & 2 & -1 & 2 \\ 2 & 5 & -1 & 4 \\ 0 & 3 & 3 & 0 \end{pmatrix} \quad \to \quad [\ldots] \quad \to \quad \begin{pmatrix} 1 & 2 & -1 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

qui permet de conclure que $\operatorname{rg}\mathcal{C} = 2$, et que $(\mathbf{v}_1, \frac{1}{3}\,\mathbf{v}_3)$ est une base de $\operatorname{Vect}\mathcal{C}$.

5.2. Test d'inversibilité et calcul de l'inverse

Soit $A \in \mathcal{M}_p$ une matrice carrée. Cette matrice est inversible si et seulement si elle est de rang p. Or une matrice échelonnée $B \in \mathcal{M}_p$ est de rang p si et seulement si $j_k = k$ pour tout k puisque l'on doit avoir $1 \le j_1 < j_2 < \cdots < j_p \le p$, et donc parmi les matrices carrées d'ordre p, I_p est la seule matrice échelonnée réduite de rang p. On obtient ainsi que A est inversible si et seulement si son équivalent canonique est I_p. Ce résultat a deux conséquences importantes.

La première, c'est que toute opération réversible sur les lignes peut s'obtenir par une suite finie d'opérations élémentaires. En effet, une opération réversible sur les lignes peut se représenter par une multiplication à gauche par une matrice inversible Q^{-1}. Or la réduction de la matrice Q par opérations élémentaires revient à multiplier à gauche cette matrice par les matrices Q_ℓ^{-1} des opérations élémentaires, de la forme $P_{(j,k)}$, $D_{j,x}$ $T_{j,k,x}$. Comme Q est inversible, son équivalent canonique est I_p, et la réduction complète s'écrit donc $Q_N^{-1} Q_{N-1}^{-1} \ldots Q_1^{-1} Q = I_p$, soit $Q^{-1} = Q_N^{-1} \ldots Q_1^{-1}$, ce qui traduit que l'opération réversible Q^{-1} peut être réalisée par la suite Q_1^{-1}, …, Q_N^{-1} d'opérations élémentaires.

L'autre conséquence, c'est que la même formule $Q^{-1} = Q_N^{-1} \ldots Q_1^{-1}$ que nous venons d'écrire permet aussi le calcul pratique de l'inverse en disposant à côté de la matrice A à inverser la matrice identité et en lui faisant subir les mêmes transformations qu'à la matrice A. En effet, cette deuxième matrice sert alors de mémoire des opérations effectuées, et prend la valeur A^{-1} lorsque l'on parvient à l'équivalent canonique I_p de A.

Exemple 5.2

$$A \;=\; \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3 \quad \to \quad [\ldots] \quad \to$$

$$\to \quad [\ldots] \quad \to \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix} = A^{-1}.$$

5.3. Résolution de systèmes linéaires

Un système linéaire

$$\left\{ \begin{array}{ccccccccc} a_{11}\,x_1 & + & a_{12}\,x_2 & + & \ldots & + & a_{1n}\,x_n & = & b_1, \\ a_{21}\,x_1 & + & a_{22}\,x_2 & + & \ldots & + & a_{2n}\,x_n & = & b_2, \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{p1}\,x_1 & + & a_{p2}\,x_2 & + & \ldots & + & a_{pn}\,x_n & = & b_p, \end{array} \right.$$

de p équations à n inconnues peut s'écrire matriciellement sous la forme $A\,X = B$ en notant $X \in \mathcal{M}_{n,1}$ le vecteur colonne des inconnues x_j, $B \in \mathcal{M}_{p,1}$ le vecteur colonne des seconds membres b_k et $A \in \mathcal{M}_{p,n}$ la matrice des coefficients a_{kj}. La réduction des matrices par opérations élémentaires sur les lignes va nous permettre d'obtenir un moyen algorithmique de résoudre un tel système.

En effet, si on effectue des opérations élémentaires simultanément sur la matrice A et le vecteur colonne B disposé à côté de A, alors chaque nouvelle forme obtenue est un "codage" d'un nouveau système linéaire équivalent au système d'origine. On peut donc chercher à résoudre ce système une fois que la matrice A a été transformée en son équivalent canonique. En supposant donc que la matrice A est échelonnée réduite, il y a alors deux situations possibles :

- Si pour un indice $k > r$ on a $b_k \neq 0$, la k-ème équation s'écrit $0\,x_1 + \cdots + 0\,x_n = b_k$, et le système n'a pas de solution.

- Si on a $b_k = 0$ pour tout indice $k > r$, le système a des solutions, que l'on décrit de la façon suivante : pour tout indice j différent des valeurs j_k, l'inconnue x_j est appelée *inconnue libre* et peut prendre n'importe quelle valeur ; les autres inconnues qui sont les x_{j_k} sont appelées *inconnues principales*, et s'expriment en fonction des inconnues libres et des seconds membres directement en utilisant la forme échelonnée réduite.

Exemple 5.3 Si m désigne un paramètre réel, on considère le système

$$\begin{cases} x + 3y - z = 2 \\ 2x + 7y - z = 5 \\ x + 2y - 2z = m \end{cases}$$

que l'on représente et réduit sous la forme

$$\begin{pmatrix} 1 & 3 & -1 \\ 2 & 7 & -1 \\ 1 & 2 & -2 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \\ m \end{pmatrix} \quad \to \quad \begin{pmatrix} 1 & 3 & -1 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ m-2 \end{pmatrix} \quad \to \quad \begin{pmatrix} 1 & 0 & -4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ m-1 \end{pmatrix},$$

ce qui signifie que le système d'origine est équivalent au système

$$\begin{cases} x - 4z = -1 \\ y + z = 1 \\ 0 = m-1\,. \end{cases}$$

Ce système se discute donc de la façon suivante :

- Si $m \neq 1$, il n'y a pas de solution.

- Si $m = 1$, la dernière équation devient $0 = 0$ et peut être ignorée. L'inconnue z est alors une inconnue libre tandis que les inconnues x et y sont les inconnues principales : la première équation se résout sous la forme $x = 4z - 1$ et la deuxième sous la forme $y = 1 - z$. Par conséquent, les solutions du système sont données dans ce cas par

$$\begin{cases} x = 4z - 1 \\ y = 1 - z \\ z = z \end{cases}$$

pour tout $z \in \mathbb{R}$.

6. Orientations bibliographiques et index

Quelques titres récents

G. Choquet, *Cours de mathématiques de Gustave Choquet.* Ellipses, Paris (2002).
A. Denmat, F. Héaulme, *Algèbre linéaire.* Dunod, Paris (1999).
P. Gabriel, *Matrices, géométrie, algèbre linéaire.* Cassini, Paris (2001).
R. Goblot, *Algèbre linéaire.* Masson, Paris (1995).
H. Roudier, *Algèbre linéaire.* Vuibert, Paris (1998).

Des ouvrages d'avant et d'ailleurs

G. Birkhoff, S. Mac Lane, *Algèbre 1. Structures fondamentales.* Gauthier-Villars, Paris (1983).
N. Bourbaki, *Algèbre. 2. Algèbre linéaire.* Hermann, Paris (1962).
N. Bourbaki, *Éléments d'histoire des mathématiques.* Hermann, Paris (1960).
R. Godement, *Cours d'algèbre.* Hermann (1980).
P. R. Halmos, *Linear algebra problem book.* Math. Assoc. America (1995).
P. Jänich, *Lineare Algebra.* Springer, Heidelberg (2000).

Index