Math II Algèbre
– Algèbre Linéaire et Polynôme –

Kenji IOHARA

Contents

Chapter 1

Qu'est-ce que l'algèbre linéaire ?(17/09)

Le but de ce premier cours est de montrer de la façon informelle ce que nous voulons faire dans le cours

1.0.1 L'espace $\mathbb{R}^n$

Pour un entier positif $n \in \mathbb{N}$, $\mathbb{R}^n$ désigne l'ensemble des n-uplets de réels, c'est-à-dire,

$$\mathbb{R}^n := \{(x_1, \cdots, x_n) | x_i \in \mathbb{R} \ (i = 1, \cdots, n)\}.$$

En particulier, on a $\mathbb{R}^1 = \mathbb{R}$. *Donner quelques exemples,* $n = 1, 2, 3$.

1.0.2 Additions et multiplications externe sur $\mathbb{R}^n$

Sur $\mathbb{R}^n$, on peut considérer une opération naturelle (dite *somme vectoriel* ou *loi interne*) et une multiplication externe (dite *multiplication par un scalaire* ou *composition externe*) données comme suit :

1. Pour $\mathbf{e} = (x_1, \cdots, x_n) \in \mathbb{R}^n$ et $\mathbf{f} = (y_1, \cdots, y_n) \in \mathbb{R}^n$, la **somme** de $\mathbf{e}$ et $\mathbf{f}$ est l'élément $(x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$ de $\mathbb{R}^n$, noté par $\mathbf{e} + \mathbf{f}$.

2. Pour $\mathbf{e} = (x_1, \cdots, x_n) \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$, le **produit externe** de $\mathbf{e}$ par λ est l'élément $(\lambda x_1, \cdots, \lambda x_n)$, noté par $\lambda \mathbf{e}$.

Proposition 1.0.1. $(\mathbb{R}^n, +)$ *est un groupe abelien.*

Preuve. À montrer en cours. (L'élément neutre est $(0, \cdots, 0)$, noté simplement par 0, et le symétrique de $\mathbf{e} = (x_1, \cdots, x_2)$ étant $-\mathbf{e} = (-x_1, \cdots, -x_n)$.) $\square$

Voici d'autres propriétés :

1. Pour $\lambda, \mu \in \mathbb{R}$ et $\mathbf{e} \in \mathbb{R}^n$, on a
$$(\lambda + \mu)\mathbf{e} = \lambda\mathbf{e} + \mu\mathbf{e}.$$
2. Pour $\lambda \in \mathbb{R}$ et $\mathbf{e}, \mathbf{f} \in \mathbb{R}^n$, on a
$$\lambda(\mathbf{e} + \mathbf{f}) = \lambda\mathbf{e} + \lambda\mathbf{f}.$$
3. Pour $\lambda, \mu \in \mathbb{R}$ et $\mathbf{e} \in \mathbb{R}^n$, on a
$$\lambda(\mu\mathbf{e}) = (\lambda\mu)\mathbf{e}.$$
4. Pour $\mathbf{e} \in \mathbb{R}^n$, on a
$$1\mathbf{e} = \mathbf{e}.$$

Toutes les propriétés discutées ci-dessus sont les bases de ce qu'on appelle espace vectoriel introduit prochainement.

Convention de vocabulaire : Les éléments de $\mathbb{R}^n$ sont dits **vecteurs**, et le mot **scalaire** signifie $\lambda \in \mathbb{R}$.
Convention de notation : On écrit $\mathbf{e}_1 + \cdots + \mathbf{e}_k$ au lieu de $(\cdots((\mathbf{e}_1 + \mathbf{e}_2) + \mathbf{e}_3) + \cdots) + \mathbf{e}_k$, car on peut déplacer les parenthèses à volonté grâce à l'associativité.

1.0.3 Applications linéaires de $\mathbb{R}^m$ vers $\mathbb{R}^n$

Etant donné deux groupes G et H, un morphisme f de G dans H est une application $f : G \longrightarrow H$ telle que $f(xy) = f(x)f(y)$, pour tous $x, y \in G$.

Maintenant, pour comparer $\mathbb{R}^m$ et $\mathbb{R}^n$, quel type d'application faut-il considérer ? Comme il y a deux opérations définies, il faut prendre ces deux opérations en compte. Voici ce qu'il faut :

Définition 1.0.1. *Une application f de $\mathbb{R}^m$ vers $\mathbb{R}^n$ est dite* ***application linéaire*** *lorsqu'elle satisfait*

1. *pour tous $\mathbf{e}, \mathbf{f} \in \mathbb{R}^m$, $u(\mathbf{e} + \mathbf{f}) = u(\mathbf{e}) + u(\mathbf{f})$, et*
2. *pour tout $\lambda \in \mathbb{R}$ et tout $\mathbf{e} \in \mathbb{R}^m$, $u(\lambda\mathbf{e}) = \lambda u(\mathbf{e})$.*

1. *Donner quelques exemples d'applications (non)linéaires.*
2. *Parler de l'image réciproque d'un point et le relier avec la notion de système d'équations linéaires.*

Chapter 2

Espaces vectoriels

2.1 Espace vectoriel (24/09)

Ici, on va considérer une généralisation de l'espace $\mathbb{R}^n$. On va donner plusieurs exemples et quelques propriétés générales qui seront utiles. Dans cette section, on supposera que $\mathbb{K}$ est un corps commutatif, e.g., $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, etc.

2.1.1 Espace vectoriel

Ici, on généralise $\mathbb{R}^n$ $(n \in \mathbb{N})$.

Définition 2.1.1. *Un ensemble non-vide E est dit un **espace vectoriel** sur le corps $\mathbb{K}$ (ou un $\mathbb{K}$-espace vectoriel) lorsque E est muni de deux opérations $+$ et $\cdot$ satisfaisant*

1. *$(E, +)$ est un groupe abelien (l'élément neutre est noté 0),*
2. *pour tout $\lambda \in \mathbb{K}$ et tous $u, v \in E$, $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$,*
3. *pour tous $\lambda, \mu \in \mathbb{K}$ et tout $u \in E$, $(\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u$,*
4. *pour tous $\lambda, \mu \in \mathbb{K}$ et tout $u \in E$, $\lambda \cdot (\mu \cdot u) = (\lambda\mu) \cdot u$, et*
5. *$1 \cdot u = u$, pour tout $u \in E$.*

Términologies :

- *Un élément de E est appelé **vecteur**.*
- *Un élément de $\mathbb{K}$ est appelé **scalaire**.*
- *L'opération $+$ est dite la **loi de composition interne** ou la somme vectorielle.*

- *L'opération $\cdot$ est dite la* ***loi de composition externe*** *ou la multiplication par un scalaire.*

Exemple 2.1.2. *1.* $\mathbb{K}^n$.

2. $\mathcal{F}([a,b]) := \{f : [a,b] \longrightarrow \mathbb{R} |\ f : \textit{une fonction}\}$.

3. Pour $r \in \mathbb{N} \cup \{\infty\}$, $\mathcal{C}^r(a,b) := \{f : (a,b) \longrightarrow \mathbb{R} |\ f : \textit{de classe } \mathcal{C}^r\}$.

Dans la suite, comme il est souvent d'usage, on écrira λu au lieu de $\lambda \cdot u$ pour $\lambda \in \mathbb{K}$ et $u \in E$.

2.1.2 Sous-espace vectoriel

Lorsque on souhaite définir un sous-groupe d'un groupe G, on considère un sous-ensemble H de G avec certaines propriétés, c'est-à-dire, H lui-même est un groupe par la restriction de la multiplication sur G.

Maintenant, pour un $\mathbb{K}$-espace vectoriel E, une partie F de E à laquelle on s'intéresse est munie d'une structure de sous-espace comme suit :

Définition 2.1.3. *On dit qu'une partie F de E est un* ***sous-espace vectoriel*** *de E lorsque les trois conditions suivantes sont vérifiées :*

1. *F n'est pas le vide.*

2. *Pour tous $u, v \in F$, la somme $u + v$ est aussi dans F.*

3. *Pour tout $u \in F$ et tout scalaire λ, λu est aussi dans F.*

Exemple 2.1.4. *Donner un exemple dans $\mathbb{R}^2$, $\mathbb{R}^3$ et $\mathcal{C}^0(a,b) \supset \mathcal{C}^1(a,b) \supset \cdots \supset \mathcal{C}^\infty(a,b)$.*

Les deux dernières conditions s'écrivent dans une seule condition de la façon suivante :

Lemme 2.1.1. *Une partie non-vide F d'un espace vectoriel E est un sous-espace vectoriel si et seulement si la condition suivante est vérifiée :*

$$\textit{Pour tout } u, v \in F \textit{ et } \lambda \in \mathbb{K},\ u + \lambda v \textit{ est aussi dans } F.$$

Preuve. Si $F \subset E$ est un sous-espace vectoriel, la condition est claire. Supposons que F satisfasse la condition du lemme. Alors, en posant $\lambda = 1 \in \mathbb{K}$, on a la condition 2. Posons $v = u$ et $\lambda = -1$, on voit que $u + (-u) = 0 \in F$. Par ailleurs, en posant $u = 0$ dans la condition du lemme, on obtient la condition 3. $\square$

2.1.3 Opérations sur les sous-espaces vectoriels*

Soit E un $\mathbb{K}$-espace vectoriel.

Proposition 2.1.2. *Soient F, F' deux sous-espaces vectoriels de E. Alors, $F \cap F'$ est également un sous-espace vectoriel de E.*

Preuve. Faire en TD. □

Attention ! En général, $F \cup F'$ n'est pas un sous-espace vectoriel de E.
Au lieu de la réunion, on considère une autre opération : pour deux parties A, B de E, on définit la **somme** de A et B, notée par $A + B$, comme suit

$$A + B := \{a + b | a \in A,\ b \in B\}.$$

Proposition 2.1.3. *Soient F, F' deux sous-espaces vectoriels de E. Alors, la somme $F + F'$ est un sous-espace vectoriel de E.*

Preuve. Faire en TD. □

Le cas particulier suivant est important :

Définition 2.1.5. *Soient F, F' deux sous-espaces vectoriels de E. On dit que la somme $F+F'$ est une* ***somme directe****, lorsque pour tout $u \in F+F'$, il existe unique couple $(v, v') \in F \times F'$ tel que $u = v + v'$. Dans ce cas, on la note par $F \oplus F'$.*

Un critère simple pour qu'une somme soit directe est le suivant :

Lemme 2.1.4. *Soient F, F' deux sous-espaces vectoriels de E. Alors, la somme $F + F'$ est une somme directe si et seulement si $F \cap F' = \{0\}$.*

Preuve. Faire en TD. □

Terminologie :

Définition 2.1.6. *Soient F, G deux sous-espaces vectoriels de E. G est dit un* ***supplémentaire*** *de F dans E si $E = F \oplus G$.*

Remarque 2.1.5. *Soit E un espace vectoriel et F un sous-espace vectoriel de E. Alors, un supplémentaire de F dans E n'est pas unique !*

Maintenant, une question se pose : *comment peut-on créer un sous-espace vectoriel d'un espace vectoriel ?*
On va répondre à cette question dans la sous-section suivante.

2.1.4 Système générateurs

Pour un groupe G et un sous-ensemble $S \subset G$, on a une notion de **sous-groupe engendré** par S. (Rappelons que c'est le plus petit sous-groupe qui contient S.) Ici, on va introduire une notion similaire :

Définition 2.1.7. *Soit S un sous-ensemble de E. Le sous-espace vectoriel **engendré** par S est le plus petit sous-espace vectoriel de E qui contient S.*

Ce sous-espace existe. En effet, il nous suffit de prendre

$$\bigcap_{\substack{F \subset E\text{: sous-espace vectoriel} \\ \text{t.q. } S \subset F}} F.$$

Mais, cette description n'est pas sympathique ! On ne voit pas concrètement ce que c'est. Ne vous inquiètes pas, voici une recette pour en trouver une.

Supposons que S soit une partie finie, posons donc $S = \{s_1, s_2, \cdots, s_k\}$. Pour chaque indice i ($1 \leq i \leq k$), on peut multiplier s_i par un scalaire $\alpha_i \in \mathbb{K}$ et obtient $\alpha_i s_i$. On peut les ajouter et obtient $\sum_{i=1}^k \alpha_i s_i$. Quand on a deux éléments de la même forme, i.e., $\sum_{i=1}^k \alpha_i s_i$ et $\sum_{i=1}^k \beta_i s_i$, où $\alpha_i, \beta_i \in \mathbb{K}$, on a

$$\sum_{i=1}^k \alpha_i s_i + \lambda \sum_{i=1}^k \beta_i s_i = \sum_{i=1}^k (\alpha_i + \lambda \beta_i) s_i.$$

D'après le lemme 2.1.1, on déduit que le sous-espace vectoriel engendré par S contient les vecteurs de la forme

$$\sum_{i=1}^k \alpha_i s_k \qquad \alpha_i \in \mathbb{K},$$

et l'ensemble des vecteurs de cette forme est un sous-espace vectoriel.

Définition 2.1.8. *On dit qu'un vecteur $u \in E$ est une **combinaison linéaire** de S, s'il existe des scalaires $\alpha_1, \cdots, \alpha_k \in \mathbb{K}$ tels que $u = \sum_{i=1}^k \alpha_i s_i$.*

On a montré le lemme suivant :

Lemme 2.1.6. *Soit S une partie finie de E. Le sous-espace vectoriel de E engendré par S est l'ensemble des combinaisons linéaires de S.*

On note le sous-espace vectoriel de E engendré par $S = \{s_1, \cdots, s_k\}$ par $\mathrm{Vect}(S)$, $\langle s_1, \cdots, s_k \rangle$, ou $\mathbb{K}s_1 + \mathbb{K}s_2 + \cdots + \mathbb{K}s_k$. La partie S est appelée un **système générateur**, ou une **famille génératrice** de $\mathrm{Vect}(S)$.

Remarque 2.1.7. *Dans le cas où S est une partie infinie, on pourra montrer que*

$$\mathrm{Vect}(S) = \{\sum_{s \in S} \alpha_s s | \alpha_s \in \mathbb{K} \ t.q. \ \mathrm{card}\{s | \alpha_s \neq 0\} < \infty\}.$$

Exemple 2.1.9. *Donner des exemples de systèmes générateurs dans* $\mathbb{R}^2$*,* $\mathbb{R}^3$*. En particulier, expliquer*

1. *qu'un système générateur d'un sous-espace vectoriel n'est pas unique,*
2. *que pour un système générateur* S*, l'expression d'un vecteur de* $\mathrm{Vect}(S)$ *comme combinaison linéaire n'est pas forcément unique.*

2.2 Base (01/10)

Ici, on va parler de systèmes générateurs économiques et d'une raison pour en considérer.

2.2.1 Système libre et base

Soit E un $\mathbb{K}$-espace vectoriel. Ici, on va parler d'un système générateur plus 'économique'.

Définition 2.2.1. *Un système de vecteurs* $S = \{s_1, \cdots, s_k\}$ *de* E *est dit*

1. ***libre****, lorsque* $\lambda_1 s_1 + \cdots + \lambda_k s_k = 0$*, avec* $\lambda_i \in \mathbb{K}$*, entraîne* $\lambda_1 = \cdots = \lambda_k = 0$*,*
2. *sinon, on dit que le système* S *est* ***lié****.*

Exemple 2.2.2. *Donner des exemples de familles de vecteurs qui sont libres ou liés dans* $\mathbb{R}^2$*.*

Voici une propriété utile :

Lemme 2.2.1. *Pour* $k \in \mathbb{N}^*$ *et* $s_1, \cdots, s_k \in E$*, la famille* $\{s_1, \cdots, s_k\}$ *est libre si et seulement si* $\{s_1, \cdots, s_{k-1}\}$ *est libre et* $s_k \notin \mathrm{Vect}\{s_1, \cdots, s_{k-1}\}$*.*

Preuve. Preuve de $\Longrightarrow$ (par contraposition) : supposons que $\{s_1, \cdots, s_{k-1}\}$ soit liée ou que $s_k \in \mathrm{Vect}(\{s_1, \cdots, s_{k-1}\})$.

- Dans le premier cas, il existe des scalaires $\lambda_1, \cdots, \lambda_{k-1} \in \mathbb{K}$ non tous nuls, tels que

  $$0 = \lambda_1 s_1 + \cdots + \lambda_{k-1} s_{k-1} = \lambda_1 s_1 + \cdots + \lambda_{k-1} s_{k-1} + 0 \cdot s_k = 0,$$

  i.e., la famille $\{s_1, \cdots, s_k\}$ est liée.

- Dans le deuxième cas, il existe $\lambda_1, \cdots, \lambda_{k-1} \in \mathbb{K}$ non tous nuls, tels que

  $$s_k = \lambda_1 s_1 + \cdots + \lambda_{k-1} s_{k-1} \quad \Longleftrightarrow \quad \lambda_1 s_1 + \cdots + \lambda_{k-1} s_{k-1} + (-1) \cdot s_k = 0,$$

  i.e., la famille $\{s_1, \cdots, s_k\}$ est liée.

Preuve de $\Longleftarrow$ (par contraposition) : supposons que la famille $\{s_1, \cdots, s_k\}$ soit liée. Par hypothèse, il existe $\lambda_1, \cdots, \lambda_k \in \mathbb{K}$ non tous nuls, tels que

$$\lambda_1 s_1 + \cdots + \lambda_{k-1} s_{k-1} + \lambda_k s_k = 0.$$

- Si $\lambda_k = 0$, il existe un indice i $(1 \leq i < k)$ tel que $\lambda_i \neq 0$ et $\lambda_1 s_1 + \cdots + \lambda_{k-1} s_{k-1} = 0$, i.e., la famille $\{s_1, \cdots, s_{k-1}\}$ est liée.
- Si $\lambda_k \neq 0$, l'égalité ci-dessus récrit comme suit :

  $$s_k = -\frac{\lambda_1}{\lambda_k} s_1 - \cdots - \frac{\lambda_{k-1}}{\lambda_k} s_{k-1},$$

  i.e., $s_k \in \mathrm{Vect}(\{s_1, \cdots, s_{k-1}\})$.

□

Définition 2.2.3. *Soient F un sous-espace vectoriel de E et $\mathcal{B} = \{s_1, \cdots, s_k\} \subset F$ une famille de vecteurs. On dit que $\mathcal{B}$ est une **base** de F, lorsque $\mathcal{B}$ est libre et engendre F.*

Exemple 2.2.4. *1. Définir la **base canonique** de $\mathbb{R}^n$.*

2. Donner des exemples de bases de sous-espaces vectoriels de $\mathbb{R}^2$ et $\mathbb{R}^3$.

Par définition, le lemme suivant est facile à montrer :

Lemme 2.2.2. *Soient E un espace vectoriel et $S = \{s_1, \cdots, s_n\}$ une famille de vecteurs de E. Alors,*

1. *S engendre E si et seulement si, pour tout $v \in E$, il existe **au moins** un k-uplet de scalaires $(\alpha_1, \cdots, \alpha_k) \in \mathbb{K}^k$ tels que $v = \alpha_1 s_1 + \cdots + \alpha_k s_k$.*
2. *S est libre si et seulement si, pour tout $v \in E$, il existe **au plus** un k-uplet de scalaires $(\alpha_1, \cdots, \alpha_k) \in \mathbb{K}^k$ tels que $v = \alpha_1 s_1 + \cdots + \alpha_k s_k$.*
3. *S est une base de E si et seulement si, pour tout $v \in E$, il existe un **unique** k-uplet de scalaires $(\alpha_1, \cdots, \alpha_k) \in \mathbb{K}^k$ tels que $v = \alpha_1 s_1 + \cdots + \alpha_k s_k$.*

Preuve. 1. est évident par définition. 3. est un corollaire direct de 1. et 2. Il nous reste à montrer 2.
Preuve de $\Longrightarrow$: supposons que S soit libre. Pour un $v \in E$, supposons que les scalaires $\alpha_1, \cdots, \alpha_k$ et $\beta_1, \cdots, \beta_k$ permettent d'écrire

$$v = \alpha_1 s_1 + \cdots \alpha_k s_k = \beta_1 s_1 + \cdots + \beta_k s_k.$$

En soustrayant ces deux égalités, on obtient

$$0 = (\alpha_1 - \beta_1) s_1 + \cdots + (\alpha_k - \beta_k) s_k.$$

Par hypothèse, on a $\alpha_1 - \beta_1 = \cdots = \alpha_k - \beta_k = 0$, donc, l'unicité d'écriture de v.
Preuve de $\Longleftarrow$ (par contraposition) : supposons que le système S soit lié. Il existe donc des scalaires $\lambda_1, \cdots, \lambda_k \in \mathbb{K}$ non tous nuls, tels que $\lambda_1 s_1 + \cdots + \lambda_k s_k = 0$. Mais, on peut aussi écrire 0 d'une façon évidente : $0s_1 + \cdots + 0s_k = 0$. Donc, il existe un vecteur de E qui possède plus d'une écriture. □

Définition 2.2.5. *Soient E un espace vectoriel, $\mathcal{B} = \{b_1, \cdots, b_k\}$ une base de E, et v un vecteur de E. On appelle **coordonnées** de v dans la base $\mathcal{B}$ les scalaires $\alpha_1, \cdots, \alpha_k$ déterminés de façon unique, tels que*

$$v = \alpha_1 b_1 + \cdots + \alpha_k b_k.$$

On les note par

$$\mathrm{mat}_{\mathcal{B}}(v) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix}.$$

Une question se pose : *est-ce que le nombre de vecteurs d'une base d'un espace vectoriel dépend d'un choix d'une base ?*
La réponse est **NON**! On va le montrer dans la section suivante.

2.2.2 Le cardinal des bases

Soit E un $\mathbb{K}$-espace vectoriel. Ici, on va montrer que le cardinal d'une base de E ne dépend que E. Plus précisement, on va le monter quand E est de **type fini**, c.-à.-d., quand E possède une famille génératrice finie.

Lemme 2.2.3. *Soient $\{f_1, \cdots, f_p\}$ un système libre de E et $\{g_1, \cdots, g_q\}$ un système générateur de E. Alors, il existe un indice j $(1 \leq j \leq q)$ tel que $\{f_1, \cdots, f_{p-1}, g_j\}$ soit encore un système libre.*

Preuve. Montrons ce lemme par l'absurde. Supposons que la conclusion soit fausse, c.-à.-d., pour chaque j, $\{f_1, \cdots, f_{p-1}, g_j\}$ est un système lié.
Par le lemme 2.2.1, $\{f_1, \cdots, f_{p-1}\}$ est libre et g_j est une combinaison linéaire de $\{f_1, \cdots, f_{p-1}\}$. Donc, il existe des scalaires $a_{i,j} \in \mathbb{K}$ $(1 \leq i \leq p-1,\ 1 \leq j \leq q)$ tels que

$$g_1 = \sum_{i=1}^{p-1} a_{i,1} f_i, \quad g_2 = \sum_{i=1}^{p-1} a_{i,2} f_i, \quad \cdots \quad , g_q = \sum_{i=1}^{p-1} a_{i,q} f_i.$$

Comme $\{g_1, \cdots, g_q\}$ est un système générateur, tout vecteur de E, en particulier f_p, peut s'écrire comme combinaison linéaire de $\{g_1, \cdots, g_q\}$, c.-à.-d., il existe des scalaires $b_1, \cdots, b_q \in \mathbb{K}$ tels que

$$f_p = \sum_{j=1}^{q} b_j g_j.$$

Comme $g_1, \cdots, g_q$ sont combinaisons linéaires de $\{f_1, \cdots, f_{p-1}\}$, ceci entraîne que f_p est une combinaison linéaire de $\{f_1, \cdots, f_{p-1}\}$. C'est une contradiction. $\square$

Le corollaire suivant est un point important :

Corollaire 2.2.4. *Soient $\{f_1, \cdots, f_p\}$ un système libre de vecteurs de E et $\{g_1, \cdots, g_q\}$ un système générateur de E. Alors, $p \leq q$.*

Preuve. Si $p = 0$, le résultat est évident. Supposons que $p > 0$. Appliquant le lemme 2.2.3, il existe un indice j $(1 \leq j \leq q)$ tel que $\{f_1, \cdots f_{p-1}, g_j\}$ soit libre. Notons ce j par $\varphi(p)$.
Appliquons le même lemme à $\{g_{\varphi(p)}, f_1, \cdots, f_{p-1}\}$, il existe un indice j' $(1 \leq j' \leq q)$ tel que $\{g_{\varphi(p)}, f_1, \cdots, f_{p-2}, g_{j'}\}$ soit libre. Notons j' par $\varphi(p-1)$. Répétons cette idée, on voit qu'il existe une application φ de $\{1, \cdots, p\}$ dans $\{1, \cdots q\}$ telle que $\{g_{\varphi(1)}, \cdots, g_{\varphi(p)}\}$ soit libre. De plus, comme le système obtenu est libre, on a $\varphi(i) \neq \varphi(j)$ pour $i \neq j$, c'est-à-dire, φ est injective ce qui entraîne que $p \leq q$. $\square$

Voici une conséquence importante :

Théorème 2.2.5. *Soit E un espace vectoriel de type fini. Toutes les bases de E ont le même cardinal.*

Preuve. Soient $\mathcal{F} = \{f_1, \cdots, f_p\}$ et $\mathcal{G} = \{g_1, \cdots, g_q\}$ deux bases de E.
En appliquant le corollaire 2.2.4 au système libre $\mathcal{F}$ et au système générateur $\mathcal{G}$, on obtient $p \leq q$.
En appliquant le corollaire 2.2.4 au système libre $\mathcal{G}$ et au système générateur $\mathcal{F}$, on obtient $q \leq p$. $\square$

Corollaire 2.2.6. *Soit n un entier positif. Toutes les bases de $\mathbb{K}^n$ sont des n-uplets.*

Remarque 2.2.7. *Les arguments donnés ci-dessus montrent l'équivalence des notions suivantes :*

$$\begin{aligned} \textit{une base} &\sim \textit{un système générateur minimal} \\ &\sim \textit{un système libre maximal.} \end{aligned}$$

2.3 Dimension (08/10)

Ici, je vais discuter de quelques propriétés de la notion de dimension d'un espace vectoriel.

2.3.1 Existence des bases

Le théorème suivant est appelé le **théorème de la base incomplète** :

Théorème 2.3.1. *Soient E un espace vectoriel de type fini et $\{f_1, \cdots, f_p\}$ un système libre de E. Alors, il existe $l \in \mathbb{N}$ et $f_{p+1}, \cdots, f_{p+l} \in E$ tels que $\{f_1, \cdots, f_p, f_{p+1}, \cdots, f_{p+l}\}$ est une base de E.*

Preuve. Si le système $\{f_1, \cdots, f_p\}$ est déjà une base, il n'y a rien à faire ! Sinon, soit $f_{p+1} \in E \setminus \{0\}$ tel que $f_{p+1} \notin \mathrm{Vect}(\{f_1, \cdots, f_p\})$. D'après le lemme 2.2.1, $\{f_1, \cdots, f_{p+1}\}$ est un système libre. Si ce nouveau système est une base, c'est fini. Sinon, on peut lui adjoindre un nouveau vecteur f_{p+2}. On peut ensuite continuer ...
D'après le théorème 2.2.5, il est clair qu'on s'arrêtera un jour, car E est de type fini par hypothèse. □

En général, la démonstration de cette propriété nécessite ce que l'on appelle le **lemme de Zorn**. Voir, e.g., § A.1 pour plus de détails.

Bon, on arrive à conclure comme suit :

Théorème 2.3.2. *Soit E un espace vectoriel de type fini. Alors, E possède une base.*

Preuve. Si $E = \{0\}$, il n'y a rien à montrer. (Dans ce cas, il faut prendre $\emptyset$.) Sinon, soit $v \in E$ un vecteur non nul. Alors, le système $\{v\}$ est un système libre. En appliquant le théorème 2.3.1 à ce système libre, on obtient une base de E. □

2.3.2 Dimension

Maintenant, on a le droit de poser

Définition 2.3.1. *Soit E un espace vectoriel de type fini. On appelle* ***dimension*** *de E, notée par $\dim E$, le cardinal de n'importe quelle base de E.*

Voici quelques propriétés :

Lemme 2.3.3. *Soient E un espace vectoriel de dimension finie et F un sous-espace vectoriel de E. Alors,*

1. $\dim F \leq \dim E$.
2. *Si $F \neq E$, on a* $\dim F < \dim E$.

Preuve. 1. : Soit $\mathcal{F}$ une base de F. Alors, $\mathcal{F}$ est un système libre de E. Par le théorème de la base incomplète (Théorème 2.3.1), on obtient l'inégalité $\dim F \leq \dim E$.
2. (par contraposition) : supposons que $F \subset E$ et $\dim E = \dim F$. Prenons

une base $\mathcal{F}$ de F. C'est un système libre dans E qui possède $\dim E$ éléments, donc c'est aussi une base de E. L'espace engendré par ce système $\mathcal{F}$ est à la fois E et F, donc $E = F$. $\square$

2.3.3 Une formule de Grassmann

Soit E un $\mathbb{K}$-espace vectoriel de type fini.

Théorème 2.3.4 (Formule de Grassmann)**.** *Soient F, G deux sous-espaces vectoriels de E.*

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

Preuve. Notons $p = \dim F$, $q = \dim G$ et $r = \dim(F \cap G)$.
Partons d'une base $\{h_1, \cdots, h_r\}$ de $F \cap G$.
Par le théorème de la base incomplète appliqué à $\{h_1, \cdots, h_r\}$ et F, il existe des vecteurs $f_1, \cdots, f_a$ $(a = p - r)$ tels que $\{h_1, \cdots, h_r, f_1, \cdots, f_a\}$ soit une base de F. De même (en notant $b := q - r$), on fabrique une base $\{h_1, \cdots, h_r, g_1, \cdots, g_b\}$ de G. Il nous suffit de montrer que $\mathcal{B} := \{h_1, \cdots, h_r, f_1, \cdots, f_a, g_1, \cdots, g_b\}$ est une base de $F + G$. En effet, cela signifie que

$$\dim(F+G) = r+a+b = r+(p-r)+(q-r) = p+q-r = \dim F+\dim G-\dim(F\cap G).$$

Montrons que $\mathcal{B}$ est un système générateur de $F + G$. Notons H le sous-espace vectoriel engendré par $\mathcal{B}$.

- Montrons que $F + G \subset H$. Soit $v \in F + G$. Alors, il existe $f \in F$ et $g \in G$ tels que $v = f + g$. Il existe des scalaires $\gamma_1, \cdots, \gamma_r, \alpha_1, \cdots, \alpha_a$ tels que
$$f = \sum_{i=1}^{r} \gamma_i h_i + \sum_{i=1}^{a} \alpha_i f_i.$$
De même, il existe des scalaires $\delta_1, \cdots, \delta_r, \beta_1, \cdots, \beta_b$ tels que
$$g = \sum_{i=1}^{r} \delta_i h_i + \sum_{i=1}^{b} \beta_i g_i.$$
Dès lors
$$v = \sum_{i=1}^{r} (\gamma_i + \delta_i) h_i + \sum_{i=1}^{a} \alpha_i f_i + \sum_{i=1}^{b} \beta_i g_i,$$
donc $v \in G$.

- Montrons que $H \subset F + G$. Soit $v \in H$. Il existe des scalaires $\gamma_1, \cdots \gamma_r, \alpha_1, \cdots, \alpha_a, \beta_1, \cdots, \beta_b$ tels que
$$\begin{aligned} v &= \sum_{i=1}^{r} \gamma_i h_i + \sum_{i=1}^{a} \alpha_i f_i + \sum_{i=1}^{b} \beta_b g_b \\ &= \left(\sum_{i=1}^{r} \gamma_i h_i + \sum_{i=1}^{a} \alpha_i f_i \right) + \left(\sum_{i=1}^{b} \beta_i g_i \right). \end{aligned}$$
Dans ces parenthèses, on voit que v s'écrit comme somme d'un vecteur de F et d'un vecteur de G, donc $v \in F + G$.

Montrons que $\mathcal{B}$ est un système libre.
Soit $\lambda_1, \cdots, \lambda_r, \mu_1, \cdots, \mu_a, \nu_1, \cdots, \nu_b$ des scalaires tels que
$$\sum_{i=1}^{r} \lambda_i h_i + \sum_{i=1}^{a} \mu_i f_i + \sum_{i=1}^{b} \nu_i g_i = 0.$$
Regroupons les différemment :
$$\sum_{i=1}^{r} \lambda_i h_i + \sum_{i=1}^{a} \mu_i f_i = -\sum_{i=1}^{b} \nu_i g_i.$$
Appelons x ce nouveau vecteur.
Vu sa première expression, $x \in F$ et vu sa deuxième expression, $x \in G$, donc $x \in F \cap G$.
Comme $\{h_1, \cdots, h_r\}$ est une base de $F \cap G$, il existe des scalaires $\gamma_1, \cdots, \gamma_r$ tels que
$$x = \sum_{i=1}^{r} \gamma_i h_i.$$
Mettons côte à côte deux expressions :
$$\begin{aligned} x &= \sum_{i=1}^{r} \gamma_i h_i + \sum_{i=1}^{b} 0 g_i \\ &= \sum_{i=1}^{r} 0 h_i - \sum_{i=1}^{b} \nu_i g_i. \end{aligned}$$
Comme $\{h_1, \cdots, h_r, g_1, \cdots, g_b\}$ est libre dans G, ces deux écritures doivent coïncider. Donc, on obtient $\nu_1 = \cdots = \nu_b = 0$ et $x = 0$, c.-à.-d.,
$$\sum_{i=1}^{r} \lambda_i h_i + \sum_{i=1}^{a} \mu_i f_i.$$
Comme $\{h_1, \cdots, h_r, f_1, \cdots, f_a\}$ est libre dans F, on obtient $\lambda_1 = \cdots = \lambda_r = \mu_1 = \cdots = \mu_a = 0$, c.-à.-d., $\mathcal{B}$ est libre. $\square$

2.3.4 Somme directe*

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie.

Définition 2.3.2. *Soient $F_1, \cdots, F_k$ des sous-espaces vectoriels de E. Notant leur somme $F = F_1 + \cdots + F_k$, on dira que les sous-espaces $F_1, \cdots, F_k$ sont en **somme directe**, lorsque, pour tout $v \in F$, il existe un **unique** k-uplet de vecteurs $v_1 \in F_1, \cdots, v_k \in F_k$ tels que*

$$v = v_1 + \cdots + v_k.$$

Dans ce cas, on note cette somme par $F = F_1 \oplus \cdots \oplus F_k$.

Voici un critère pratique pour déterminer qu'une somme est directe :

Théorème 2.3.5. *Soient $F_1, \cdots, F_k$ des sous-espaces vectoriels de E. Les sous-espaces sont en somme directe si et seulement si $(F_1+\cdots+F_{i-1}) \cap F_i = \{0\}$ pour tout $1 < i \leq k$.*

Preuve. Procédons par récurrence. Le cas $k = 2$ sera traité en TD. Supposons que l'énoncé soit vrai jusqu'à $k-1$. Montrons le cas k.
D'abord, on montre que la condition est suffisante. Soit $v \in F_1+\cdots+F_k$. Le cas $k = 2$ implique qu'il existe un unique $v' \in F_1 + \cdots + F_{k-1}$ et un unique $v_k \in F_k$ tels que $v = v' + v_k$. Par hypothèse, il existe des $v_i \in F_i$ $(1 \leq i < k)$ tels que $v' = v_1 + \cdots + v_{k-1}$. Donc, l'écriture $v = v_1 + \cdots + v_k$ est unique.
Ensuite, on montre que la condition est aussi nécessaire par contraposition. Supposons qu'il existe $1 < i \leq k$ tel que $(F_1 + \cdots + F_{i-1}) \cap F_i \neq \{0\}$. Le cas $k = 2$ implique que la somme $(F_1 + \cdots + F_{i-1}) + F_i$ n'est pas directe, d'où la conclusion. □

Donc, par la formule de Grassmann, on déduit aisément que, si la somme $F_1 + \cdots + F_k$ est directe, on a $\dim(F_1 + \cdots + F_k) = \dim F_1 + \cdots + \dim F_k$. Le théorème suivant nous assure que cette condition est aussi suffisante.

Théorème 2.3.6. *Soient $F_1, \cdots, F_k$ des sous-espaces vectoriels de E. Les sous-espaces sont en somme directe si et seulement si*

$$\dim(F_1 + \cdots + F_k) = \dim F_1 + \cdots + \dim F_k.$$

Preuve. On a déjà vu que si les sous-espaces $F_1, \cdots, F_k$ sont en somme directe, alors $\dim(F_1+\cdots+F_k) = \dim F_1+\cdots+\dim F_k$. Donc, il nous reste à vérifier le contraire. Supposons que $\dim(F_1+\cdots+F_k) = \dim F_1+\cdots+\dim F_k$. Notons $\mathcal{B}_i$ une base de F_i $(1 \leq i \leq k)$. Il est clair que $\mathcal{B} := \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ est une famille génératrice de $F := F_1+\cdots+F_k$ qui est minimale par hypothèse, c'est donc une base. Soit $v \in F$ et considérons deux écritures

$$\begin{aligned} v &= v_1 + \cdots + v_k \\ &= w_1 + \cdots + w_k \end{aligned}$$

de v, dans lesquelles $v_i, w_i \in F_i$ ($1 \leq i \leq k$). Soient α_b, β_b des scalaires tels que $v_i = \sum_{b \in \mathcal{B}_i} \alpha_b b$ et que $w_i = \sum_{b \in \mathcal{B}_i} \beta_b b$. Reportons ces écriture dans les deux écritures du vecteur v :

$$v = \sum_{b \in \mathcal{B}} \alpha_b b = \sum_{b \in \mathcal{B}} \beta_b b.$$

Comme $\mathcal{B}$ est une base de F, on voit bien que $\alpha_b = \beta_b$, pour tout $b \in \mathcal{B}$, c.-à.-d., $v_i = w_i$, pour tout $1 \leq i \leq k$. Donc, l'écriture de v est bien unique. $\square$

Chapter 3

Applications linéaires

3.1 Applications linéaires I (15/10)

Ici, on fixe un corps commutatif $\mathbb{K}$.

3.1.1 Définitions

On commence par

Définition 3.1.1. *Soient E, F deux $\mathbb{K}$-espaces vectoriels et u une application de E vers F. On dit que u est une **application linéaire**, lorsque*

1. *pour tous $x, y \in E$, $u(x+y) = u(x) + u(y)$,*

2. *Pour tout $\lambda \in \mathbb{K}$ et tout $x \in E$, $u(\lambda x) = \lambda u(x)$.*

L'ensemble des applications linéaires de E vers F sera noté $\mathcal{L}(E, F)$.

Voici un cas particulier :

Définition 3.1.2. *Soit E un $\mathbb{K}$-espace vectoriel.*

1. *Une application linéaire de E vers E est dite un **endomorphisme**.*

2. *Un endomorphisme bijectif est appelé un **automorphisme**.*

On notera $\mathcal{L}(E)$ pour $\mathcal{L}(E, E)$.

On peut facilement caractériser les applications linéaires comme on a fait pour les sous-espaces vectoriels :

Proposition 3.1.1. *Une application u d'un espace vectoriel E vers un espace vectoriel F est linéaire si et seulement si :*

$$\text{pour tous } x, y \in E \text{ et tout } \lambda \in \mathbb{K},\ u(x + \lambda y) = u(x) + \lambda u(y).$$

Preuve. Preuve de $\Longrightarrow$: Supposons que f soit linéaire. Soient $x, y \in E$ et $\lambda \in \mathbb{K}$. On a alors :

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

Preuve de $\Longleftarrow$: Supposons que f vérifie la caractérisation de l'enoncé de la proposition. Soient $x, y \in E$. On a alors, $u(x + y) = u(x + 1y) = u(x) + 1u(y) = u(x) + u(y)$. On en déduit ensuite que $u(0) = u(0 + 0) = u(0) + u(0)$, donc $u(0) = 0$. Alors, pour $x \in E$ et $\lambda \in \mathbb{K}$, on a alors $u(\lambda x) = u(0 + \lambda x) = u(0) + \lambda u(x) = \lambda u(x)$. Donc, u est bien linéaire. $\square$

3.1.2 Exemples

D'abord, on considère le cas $E = F = \mathbb{R}^2$. Soit $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2\}$ la base canonique.

Le premier exemple est une

1. **Dilatation** Soient $a, b \in \mathbb{R}$ des réels positifs. On définit l'application linéaire $u \in \mathcal{L}(E)$ par

$$u(\mathbf{e}_1) = a\mathbf{e}_1, \qquad u(\mathbf{e}_2) = b\mathbf{e}_2.$$

Calculons l'image du cercle $\{(x, y)|x^2 + y^2 = 1\}$, etc.

Le deuxième exemple est une

2. **Rotation** Soient $\alpha \in \mathbb{R}$ un réel positif et u la rotation de centre $(0, 0)$ et d'angle α. Montrons que $u \in \mathcal{L}(E)$.

Pour tout $(x, y) \in E \setminus \{(0, 0)\}$, il existe $r \in \mathbb{R}$ positif et $\theta \in [0, 2\pi[$ tels que $x = r\cos\theta$ et $y = r\sin\theta$. Par la formule d'addition des fonctions trigonométriques, on a

$$\begin{aligned}
&u(x\mathbf{e}_1 + y\mathbf{e}_y) = u((x, y)) \\
&= (r\cos(\theta + \alpha), r\sin(\theta + \alpha)) = r\cos\theta(\cos\alpha, \sin\alpha) + r\sin\theta(-\sin\alpha, \cos\alpha) \\
&= x(\cos\alpha\mathbf{e}_1 + \sin\alpha\mathbf{e}_2) + y(-\sin\alpha\mathbf{e}_1 + \cos\alpha\mathbf{e}_2),
\end{aligned}$$

c.-à.-d., u est une application linéaire.

Le troisième exemple est une

3. **Symétrie** Soient $\alpha \in \mathbb{R}$ un réel positif et u la symétrie par rapport à la droite $(\sin\alpha)x - (\cos\alpha)y = 0$. Montrons que $u \in \mathcal{L}(E)$.

Notons l'image $u((x, y))$ par (x', y'). Alors, on sait que

1. le milieu de (x, y) et (x', y') est sur l'axe de cette symétrie et que
2. la droite qui passe par (x, y) et (x', y') est perpendiculaire à cet axe.

Ces deux propriétés se traduisent sous la forme suivante :

$$\sin\alpha\left(\frac{x + x'}{2}\right) - \cos\alpha\left(\frac{y + y'}{2}\right) = 0, \qquad \cos\alpha(x' - x) + \sin\alpha(y' - y) = 0.$$

De ces deux égalités, on déduit

$$x' = (\cos 2\alpha)x + (\sin 2\alpha)y, \qquad y' = (\sin 2\alpha)x - (\cos 2\alpha)y,$$

c.-à.-d.,

$$u(x\mathbf{e}_1 + y\mathbf{e}_2) = x(\cos 2\alpha\mathbf{e}_1 + \sin 2\alpha\mathbf{e}_2) + y(\sin 2\alpha\mathbf{e}_1 - \cos 2\alpha\mathbf{e}_2).$$

Donc, u est une application linéaire.

Maintenant, on considère le cas $E = F = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$.

Le quatrième exemple est la

4. **Différentielle** Soit u l'application de E dans E définie par $u(f)(x) := \frac{df}{dx}(x)$. Alors, u est une application linéaire. (Le montrer.)

Le cinquième exemple est l'

5. **Intégrale** Soient $a \in \mathbb{R}$ et u l'application de E dans E définie par $u(f)(x) := \int_a^x f(t)dt$. Alors, u est une application linéaire. (Le montrer.)

3.1.3 Applications linéaires et bases

Ici, on discute des données minimales pour définir une application linéaire. Une proposition assez pratique est donnée comme suit :

Proposition 3.1.2. *Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et F un $\mathbb{K}$-espace vectoriel. Soient $\{\mathbf{e}_1, \cdots, \mathbf{e}_k\}$ une base de E et $\{\mathbf{f}_1, \cdots, \mathbf{f}_k\}$ une famille de vecteurs de F. Alors, il existe une et une seule application linéaire u de E vers F telle que*

$$u(\mathbf{e}_1) = \mathbf{f}_1, \ \cdots, \ u(\mathbf{e}_k) = \mathbf{f}_k.$$

Preuve. D'abord, on va montrer que si une telle application linéaire existe, elle est unique.

Soit donc u une application linéaire de E vers F avec les propriétés ci-dessus. Soit $x \in E$. Comme $\{\mathbf{e}_1, \cdots, \mathbf{e}_k\} \subset E$ est une base, il existe $\alpha_1, \cdots, \alpha_k \in \mathbb{K}$ tels que $x = \sum_{i=1}^k \alpha_i\mathbf{e}_i$. Alors, on a

$$u(x) = u(\sum_{i=1}^k \alpha_i\mathbf{e}_i) = \sum_{i=1}^n \alpha_i u(\mathbf{e}_i) = \sum_{i=1}^k \alpha_i\mathbf{f}_i.$$

Donc, u est déterminée sur chacun des vecteurs de E ; elle est donc unique si elle existe.

Pour l'existence, il nous suffit de montrer que l'application u définie par la formule ci-dessus est bien linéaire. On procède par un calcul direct. □

3.1.4 Opérations sur les applications linéaires*

Soient E, F deux $\mathbb{K}$-espaces vectoriels.

Définition 3.1.3. *Soient $f, g \in \mathcal{L}(E, F)$ et $\lambda \in \mathbb{K}$.*

1. *La **somme** de f et g est l'application, notée par $f + g$, définie par $(f + g)(x) := f(x) + g(x)$.*

2. *La **multiplication de** f par λ est l'application, notée par λf, définie par $(\lambda f)(x) := \lambda(f(x))$.*

On voit aisément que $f + g \in \mathcal{L}(E, F)$ et que $\lambda f \in \mathcal{L}(E, F)$.

Maintenant, il est naturel d'espérer

Proposition 3.1.3. *$\mathcal{L}(E, F)$ est muni d'une structure de $\mathbb{K}$-espace vectoriel.*

Preuve. Exercice. □

Soit G un espace vectoriel. On peut parler de la composition d'applications linéaires :

Lemme 3.1.4. *Soient $f, f_1, f_2 \in \mathcal{L}(E, F)$ et $g, g_1, g_2 \in \mathcal{L}(F, G)$.*

1. *La composée $g \circ f$ est une application linéaire de E dans G, c.-à.-d., $g \circ f \in \mathcal{L}(E, G)$.*

2. *$(g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f$.*

3. *$g \circ (f_1 + f_2) = g \circ f_1 + g \circ f_2$.*

Preuve. On montre le premier et on laisse le reste comme exercice.
Soient $x, y \in E$ et $\lambda \in \mathbb{K}$. Alors, on a $(g \circ f)(x + \lambda y) = g(f(x + \lambda y)) = g(f(x) + \lambda f(y)) = g(f(x)) + \lambda g(f(y)) = (g \circ f)(x) + \lambda(g \circ f)(y)$. □

Remarque 3.1.5. *Évidemment, la composition est associative. Donc, d'après le lemme ci-dessus, on déduit que $\mathcal{L}(E)$ est muni d'une structure d'anneau.*

Alors, on va parler de l'inverse d'une application linéaire :

Proposition 3.1.6. *Soit f une application linéaire bijective de E dans F. Alors, f^{-1} est aussi linéaire.*

Preuve. Soient $y_1, y_2 \in F$ et $\lambda \in \mathbb{K}$. Comme u est bijective, il existe $x_1, x_2 \in E$ tels que $y_1 = f(x_1)$ et $y_2 = f(x_2)$. Montrons que $f^{-1}(y_1 + \lambda y_2) = f^{-1}(y_1) + \lambda f^{-1}(y_2)$. Posons $x := f^{-1}(y_1 + \lambda y_2)$ et $x' := f^{-1}(y_1) + \lambda f^{-1}(y_2)$. Alors, on a $f(x) = f(f^{-1}(y_1 + \lambda y_2)) = y_1 + \lambda y_2$ et $f(x') = f(f^{-1}(y_1) + \lambda f^{-1}(y_2)) = f(f^{-1}(y_1)) + \lambda f(f^{-1}(y_2)) = y_1 + \lambda y_2$, d'où $f(x) = f(x')$. Comme f est bijective, on en déduit que $x = x'$. □

3.2 Applications linéaires II (22/10)

Ici, $\mathbb{K}$ désigne un corps commutatif. Je vais expliquer quelques propriétés générales des applications linéaires.

3.2.1 Injectivité et surjectivité

Ici, on va discuter de l'injectivité et de la sujectivité d'une application linéaire. Soient E et F deux espaces vectoriels.

On commence par un lemme préliminaire :

Lemme 3.2.1. *Soient u une application linéaire de E vers F et F' un sous-espace vectoriel de F. Alors, $u^{-1}(F')$ est un sous-espace vectoriel de E.*

Preuve. L'image réciproque $u^{-1}(F')$ est non vide, car il contient $0 \in E$. Soient $x, y \in u^{-1}(F')$ et $\lambda \in \mathbb{K}$. Montrons que $x + \lambda y \in u^{-1}(F')$. Par définition, on a $u(x), u(y) \in F'$. Comme F' est un sous-espace vectoriel, on a $u(x) + \lambda u(y) \in F'$, donc $u(x + \lambda y) \in F'$, car u est linéaire. Ceci implique que $x + \lambda y \in u^{-1}(F')$. □

D'abord, on cherche une condition pour qu'une application linéaire soit injective. Soit u une application linéaire de E vers F. Supposons que u soit injective, c.-à.-d., si $x, y \in E$ tels que $u(x) = u(y)$, alors on a $x = y$. Comme u est linéaire, $u(x) = u(y)$ implique que $0 = u(x) - u(y) = u(x - y)$. Par hypothèse, on obtient $u^{-1}(0) = 0$. Maintenant, supposons que $u^{-1}(0) = 0$. Soient $x, y \in E$ tels que $u(x) = u(y)$. Alors, comme u est linéaire, $u(x) = u(y)$ implique que $0 = u(x) - u(y) = u(x - y)$, donc, $x - y = 0$ c.-à.-d., u est injective.

Donc, il est utile d'introduire la notion suivante :

Définition 3.2.1. *Soit $u \in \mathcal{L}(E, F)$. Le sous-espace vectoriel $u^{-1}(\{0\})$ est appelé le **noyau** de u et noté par* $\operatorname{Ker} u$.

D'après ce que l'on a discuté, on obtient

Proposition 3.2.2. *Soit u une application linéaire de E vers F. Alors, u est injective si et seulement si* $\operatorname{Ker} u = \{0\}$.

Maintenant, on aborde la notion de surjectivité. Cette fois-ci, il n'y a rien de nouveau...

En effet, soit u une application linéaire de E vers F et supposons que u soit surjective. Alors, on a $u(E) = F$ par définition. C'est tout ! Il nous reste à définir

Définition 3.2.2. *Soit u une application linéaire de E vers F. Le sous-espace vectoriel $u(E)$ de F est appelé l'**image** de u et noté par* $\operatorname{Im} u$.

3.2.2 La formule du rang

Ici, on va montrer une formule importante :

Théorème 3.2.3. *Soient E un espace vectoriel de dimension finie et F un espace vectoriel. Soit u une application linéaire de E vers F. Alors, on a*

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

Preuve. Soit $\{\mathbf{e}_1, \cdots, \mathbf{e}_k\}$ une base de Ker u. Par le théorème de la base incomplète, il existe $\mathbf{e}_{k+1}, \cdots, \mathbf{e}_n \in E$ tels que $\{\mathbf{e}_1, \cdots, \mathbf{e}_k, \mathbf{e}_{k+1}, \cdots, \mathbf{e}_n\}$ forme une base de E. Avec ces notations, on a $\dim \operatorname{Ker}\ u = k$ et $\dim E = n$. Donc, il faut montrer que $\dim \operatorname{Im}\ u = n - k$. Pour cela, montrons que $\{u(\mathbf{e}_{k+1}), \cdots, u(\mathbf{e}_n)\}$ forme une base de Im u.
D'abord, on montre que cette famille est une famille génératrice de Im u. Soit $y \in \operatorname{Im}\ u$. Alors, il existe $x \in E$ tel que $y = u(x)$. On peut écrire x dans la base $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ de E, $x = \alpha_1\mathbf{e}_1 + \cdots + \alpha_k\mathbf{e}_k + \alpha_{k+1}\mathbf{e}_{k+1} + \cdots \alpha_n\mathbf{e}_n$. Donc,

$$\begin{aligned} y = u(x) &= u(\alpha_1\mathbf{e}_1 + \cdots + \alpha_k\mathbf{e}_k + \alpha_{k+1}\mathbf{e}_{k+1} + \cdots \alpha_n\mathbf{e}_n) \\ &= \alpha_1 u(\mathbf{e}_1) + \cdots + \alpha_k u(\mathbf{e}_k) + \alpha_{k+1}u(\mathbf{e}_{k+1}) + \cdots \alpha_n u(\mathbf{e}_n) \\ &= \alpha_{k+1}u(\mathbf{e}_{k+1}) + \cdots \alpha_n u(\mathbf{e}_n), \end{aligned}$$

car $\mathbf{e}_1, \cdots, \mathbf{e}_k \in \operatorname{Ker}\ u$. Donc, $y \in \langle u(\mathbf{e}_{k+1}), \cdots, u(\mathbf{e}_n)\rangle$.
Montrons maintenant que la famille $\{u(\mathbf{e}_{k+1}), \cdots, u(\mathbf{e}_n)\}$ est libre. Soient $\lambda_{k+1}, \cdots, \lambda_n \in \mathbb{K}$ tels que $\lambda_{k+1}u(\mathbf{e}_{k+1}) + \cdots + \lambda_n u(\mathbf{e}_n) = 0$, i.e., $u(\lambda_{k+1}\mathbf{e}_{k+1} + \cdots + \lambda_n\mathbf{e}_n) = 0$, i.e., $\lambda_{k+1}\mathbf{e}_{k+1} + \cdots + \lambda_n\mathbf{e}_n \in \operatorname{Ker}\ u$. Comme Ker u est engendré par $\{\mathbf{e}_1, \cdots, \mathbf{e}_k\}$, il existe $\lambda_1, \cdots, \lambda_k \in \mathbb{K}$ tels que $\lambda_{k+1}\mathbf{e}_{k+1} + \cdots + \lambda_n\mathbf{e}_n = \lambda_1\mathbf{e}_1 + \cdots + \lambda_k\mathbf{e}_k$. Cela signifie que $\lambda_{k+1} = \cdots = \lambda_n = 0$, car $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ est libre. □

Définition 3.2.3. *Soient E un espace vectoriel de dimension finie et F un espace vectoriel. Soit u une application linéaire de E vers F. La dimension* $\dim \operatorname{Im} u$ *est appelée le **rang** de u et noté par* rg u.

Voici une application directe de la formule du rang : supposons que E et F soient deux espaces vectoriels tels que $\dim E = \dim F$.

Théorème 3.2.4. *Soient E un espace vectoriel de dimension finie, F un espace vectoriel et u une application linéaire de E vers F. Alors, on a*

$$u \text{ est injective} \iff u \text{ est surjective} \iff u \text{ est bijective.}$$

Soient E, F deux $\mathbb{K}$-espaces vectoriels.

Définition 3.2.4. *Lorsqu'il existe une application linéaire bijective u de E dans F, on dit que E est **isomorphe à** F, noté $E \cong F$, et que u est un **isomorphisme**.*

Par définition, le lemme suivant est immédiat :

Lemme 3.2.5. *La relation $\cong$ est une relation d'équivalence.*

3.2.3 Premiers pas avec les matrices

Soient E, F deux $\mathbb{K}$-espaces vectoriels, $\mathcal{B} = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ et $\mathcal{C} = \{\mathbf{f}_1, \cdots, \mathbf{f}_m\}$ des bases de E et F, respectivement. Soit $u : E \longrightarrow F$ une application linéaire. On a vu que, pour déterminer u complétement, il suffit de se donner l'image $\{u(\mathbf{e}_i) | 1 \leq i \leq n\}$. Explicitement, si $u(\mathbf{e}_j) = \sum_{i=1}^m a_{i,j}\mathbf{f}_i$, alors, pour $x = \sum_{j=1}^n x_j\mathbf{e}_j \in E$, l'image $y = \sum_{i=1}^m y_i\mathbf{f}_i = u(x)$ est donnée par

$$y = u(x) = u(\sum_{j=1}^{n} x_j\mathbf{e}_j) = \sum_{j=1}^{n} x_j u(\mathbf{e}_j) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}x_j\mathbf{f}_i,$$

c.-à.-d., on a

$$\begin{cases} y_1 = & a_{1,1}x_1 + & \cdots + & a_{1,n}x_n, \\ y_2 = & a_{2,1}x_1 + & \cdots + & a_{2,n}x_n, \\ \vdots & \vdots & \vdots & \vdots \\ y_m = & a_{m,1}x_1 + & \cdots + & a_{m,n}x_n. \end{cases} \tag{3.1}$$

Notons le vecteur des coordonnées de x dans la base $\mathcal{B}$ par X et celui de y dans la base $\mathcal{C}$ par Y. Introduisons le tableau rectangulaire suivant

$$A := \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}.$$

On récrira (3.1) sous la forme $Y = AX$. Le tableau A est appelé une **matrice de taille** (m, n) **à coefficients dans** $\mathbb{K}$ et l'ensemble de telles matrices est notée par $M_{m,n}(\mathbb{K})$. On note aussi la matrice A ci-dessus comme $A = (a_{i,j})_{1\leq i\leq m, 1\leq j\leq n}$. Le scalaire $a_{i,j}$ est dit **le coefficient de la i-ème ligne, j-ème colonne**. Par rapport à l'application linéaire f, cette matrice est dite **la matrice de u dans les bases $\mathcal{B}$ et $\mathcal{C}$** et notée par $A = \mathrm{mat}_{\mathcal{B},\mathcal{C}}(u)$. On veut $\ll$ définir $\gg$ le produit des matrices de manière à ce que

$$\mathrm{mat}_{\mathcal{C}}(u(x)) = \mathrm{mat}_{\mathcal{B},\mathcal{C}}(u)\mathrm{mat}_{\mathcal{B}}(x). \tag{3.2}$$

En particulier, lorsque $m = n$, on note $M_{m,n}(\mathbb{K})$ par $M_n(\mathbb{K})$.

3.2.4 Structure d'espace vectoriel sur $M_{m,n}(\mathbb{K})$

Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimension n et m, respectivement. Comme $\mathcal{L}(E, F)$ est muni d'une structure d'espace vectoriel, il est naturel que $M_{m,n}(\mathbb{K})$ le soit aussi. Ceci est fait comme suit :

Définition 3.2.5. *1. La **somme** de deux matrices $A = (a_{i,j})$, $B = (b_{i,j}) \in M_{m,n}(\mathbb{K})$ est la matrice $C = (c_{i,j}) \in M_{m,n}(\mathbb{K})$ définie par $c_{i,j} := a_{i,j} + b_{i,j}$. On la note par $A + B$.*

2. *Soient* $A = (a_{i,j}) \in M_{m,n}(\mathbb{K})$ *et* $\lambda \in \mathbb{K}$. *La* ***multiplication de*** A ***par un scalaire*** λ *est la matrice* $C = (c_{i,j}) \in M_{m,n}(\mathbb{K})$ *définie par* $c_{i,j} = \lambda a_{i,j}$. *On la note par* λA.

Alors, on a

Lemme 3.2.6. $M_{m,n}(\mathbb{K})$ *est muni d'une structure de* $\mathbb{K}$*-espace vectoriel.*

Preuve. Exercice. □

Quelle est la dimension de $M_{m,n}(\mathbb{K})$? Pour répondre à cette question, on considère les matrices suivantes :

Définition 3.2.6. *Fixons* m *et* n. *On appelle* ***matrices élémentaires*** *les matrices* $E_{i,j}$ $(1 \leq i \leq m, 1 \leq j \leq n)$ *dont les coefficients sont nuls, sauf le coefficient de la* i*-ème ligne,* j*-ème colonne qui vaut* 1.

Alors, il est facile de montrer

Lemme 3.2.7. *La famille* $\{E_{i,j}\}_{1\leq i\leq m, 1\leq j\leq n}$ *est une base de* $M_{m,n}(\mathbb{K})$, *en particulier,* $\dim M_{m,n}(\mathbb{K}) = mn$.

Chapter 4

Matrices

4.1 Matrices I : aspect théorique (29/10)

Ici, on fixe un corps commutatif $\mathbb{K}$ comme d'habitude. On va définir une certaine structure sur $M_{m,n}(\mathbb{K})$.

4.1.1 Produit de matrices

Soient E, F, G des $\mathbb{K}$-espaces vectoriels de dimension l, m et n respectivement. Alors, on peut définir la composée de $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, G)$ qui est une application linéaire $v \circ u \in \mathcal{L}(E, G)$. Soient $\mathcal{B} = \{\mathbf{e}_1, \cdots, \mathbf{e}_l\}, \mathcal{C} = \{\mathbf{f}_1, \cdots, \mathbf{f}_m\}$ et $\mathcal{D} = \{\mathbf{g}_1, \cdots, \mathbf{g}_n\}$ des bases de E, F et G, respectivement. Au niveau des matrices, on veut donc définir ≪ le produit ≫ de matrices de telle sorte qu'il satisfasse

$$\mathrm{mat}_{\mathcal{B},\mathcal{D}}(v \circ u) = \mathrm{mat}_{\mathcal{C},\mathcal{D}}(v)\mathrm{mat}_{\mathcal{B},\mathcal{C}}(u). \tag{4.1}$$

Posons $A = (a_{i,j}) = \mathrm{mat}_{\mathcal{C},\mathcal{D}}(v)$ et $B = (b_{i,j}) = \mathrm{mat}_{\mathcal{B},\mathcal{C}}(u)$. On va calculer l'image de $x = \sum_{j=1}^{l} x_j \mathbf{e}_j \in E$ par $v \circ u$. Alors,

$$\begin{aligned}
v \circ u(x) = v \circ u(\sum_{j=1}^{l} x_j \mathbf{e}_j) &= \sum_{j=1}^{l} x_j v(u(\mathbf{e}_j)) \\
&= \sum_{j=1}^{l} x_j v(\sum_{r=1}^{m} b_{r,j} \mathbf{f}_r) = \sum_{j=1}^{l}\sum_{r=1}^{m} b_{r,j} x_j v(\mathbf{f}_r) = \sum_{j=1}^{l}\sum_{r=1}^{m} b_{r,j} x_j \sum_{i=1}^{n} a_{i,r}\mathbf{g}_i \\
&= \sum_{i=1}^{n}\left(\sum_{j=1}^{l}\left(\sum_{r=1}^{m} a_{i,r}b_{r,j}\right) x_j\right)\mathbf{g}_i.
\end{aligned}$$

Par conséquence, il est naturel de définir ≪le produit≫ de matrices comme suit :

Définition 4.1.1. *Soient* $A = (a_{i,j}) \in M_{n,m}(\mathbb{K})$ *et* $B = (b_{i,j}) \in M_{m,l}(\mathbb{K})$. *Le* ***produit*** *des matrices* A *et* B *est la matrice* $C = (c_{i,j}) \in M_{n,l}(\mathbb{K})$ *définie par* $c_{i,j} = \sum_{r=1}^{m} a_{i,r}b_{r,j}$.

Remarque 4.1.1. ***Attention !*** *Le produit de deux matrices* $A \in M_{k,l}(\mathbb{K})$ *et* $B \in M_{m,n}(\mathbb{K})$ *peut être défini si et seulement si* $l = m$.

Soient H un $\mathbb{K}$-espace vectoriel de dimension k et $w \in \mathcal{L}(G, H)$. Alors, on a l'associativité $w \circ (v \circ u) = (w \circ v) \circ u$. Cette propriété se traduit dans le langage matriciel comme suit :

Lemme 4.1.2. *Soient* $A \in M_{k,l}(\mathbb{K}), B \in M_{l,m}(\mathbb{K})$ *et* $C \in M_{m,n}(\mathbb{K})$. *Alors, on a l'associativité* $(AB)C = A(BC)$.

La distributivité, c.-à.-.d, le lemme 3.1.4 se traduit comme suit :

Lemme 4.1.3. *Soient* $A, A_1, A_2 \in M_{l,m}(\mathbb{K})$ *et* $B, B_1, B_2 \in M_{m,n}(\mathbb{K})$. *Alors,*

$$(A_1 + A_2)B = A_1B + A_2B, \qquad A(B_1 + B_2) = AB_1 + AB_2.$$

Remarque 4.1.4. *Soient* $A, B \in M_(\mathbb{K})$. *En général, on a* $AB \neq BA$. *Par exemple, on a* $E_{i,j}E_{j,i} = E_{i,i}$, *tandis que* $E_{j,i}E_{i,j} = E_{j,j}$.

4.1.2 Matrices inversibles

Ici, on considère le cas $E = F$, donc $\mathcal{L}(E)$. Rappelons que pour un automorphisme $u \in \mathcal{L}(E)$, l'inverse u^{-1} est aussi un automorphisme. Fixons une base $\mathcal{B}$ de E, ceci se traduit dans le langage matriciel comme suit.
D'abord, la matrice $\mathrm{mat}_{\mathcal{B}}(\mathrm{id}_E)$ est une matrice très particulière et importante :

Définition 4.1.2. *La* ***matrice identité*** *de taille* (n, n) *est la matrice, notée* $\mathbf{1}_n$, *dont le coefficient de la* i*-ième ligne* j*-ième colonne est* $\delta_{i,j}$, *le symbole dit de Kronecker.*

Évidemment, pour tout $A \in M_{m,n}(\mathbb{K})$, on a

$$A\mathbf{1}_n = \mathbf{1}_m A = A.$$

Lemme 4.1.5. $M_n(\mathbb{K})$ *est muni d'une structure d'anneau unitaire non commutatif.*

Maintenant, supposons que $\dim E = n$ et posons $A = \mathrm{mat}_{\mathcal{B}}(u)$ et $B = \mathrm{mat}_{\mathcal{B}}(u^{-1})$. Par (4.1), on a $\mathrm{mat}_{\mathcal{B}}(\mathrm{id}_E) = \mathrm{mat}_{\mathcal{B}}(u)\mathrm{mat}_{\mathcal{B}}(u^{-1}) = \mathrm{mat}_{\mathcal{B}}(u^{-1})\mathrm{mat}_{\mathcal{B}}(u)$, c.-à-d.,

$$\mathbf{1}_n = AB = BA.$$

Définition 4.1.3. *Soit* $A \in M_n(\mathbb{K})$. *La matrice* A *est dite* ***inversible*** *s'il existe une matrice* $B \in M_n(\mathbb{K})$ *telle que* $AB = BA = \mathbf{1}_n$. *La matrice* B *est appelée l'****inverse de*** A *et notée par* A^{-1}.

Il est clair que quand on a deux matrices inversibles A et B, alors le produit AB l'est aussi et l'inverse de AB est $B^{-1}A^{-1}$.

4.1.3 Changements de bases

Rappelons que la matrice d'une application linéaire dépend du choix des bases. Une question très naïve se pose : *qu'est-ce qu'il se passe sur la matrice quand on change de bases ?* On va aborder ce problème ici.

D'abord, soient $\mathcal{B} = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$, $\mathcal{B}' = \{\mathbf{e}'_1, \cdots, \mathbf{e}'_n\}$ deux bases de E. Pour un $x \in E$, on pose

$$X = \mathrm{mat}_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{et} \qquad X' = \mathrm{mat}_{\mathcal{B}'}(x) = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$

Comme $\mathcal{B}$ et $\mathcal{B}'$ sont des bases de E, il existe des scalaires $p_{i,j} \in \mathbb{K}$ ($1 \leq i, j \leq n$) tels que $\mathbf{e}'_j = \sum_{i=1}^n p_{i,j}\mathbf{e}_i$, c.-à.-.d., $P = (p_{i,j}) = \mathrm{mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E)$. Alors, on a

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

c.-à.-d.,

$$X = PX' \qquad \Longleftrightarrow \qquad \mathrm{mat}_{\mathcal{B}}(x) = \mathrm{mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E)\mathrm{mat}_{\mathcal{B}'}(x). \tag{4.2}$$

Définition 4.1.4. *La matrice $P = \mathrm{mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E)$ est appelée la* ***matrice de passage*** *de $\mathcal{B}$ à $\mathcal{B}'$.*

Maintenant, soient $\mathcal{B}, \mathcal{B}'$ des bases de E et $\mathcal{C}, \mathcal{C}'$ des bases de F. Pour une application linéaire $u \in \mathcal{L}(E, F)$, on pose $A := \mathrm{mat}_{\mathcal{B},\mathcal{C}}(u)$ et $B := \mathrm{mat}_{\mathcal{B}',\mathcal{C}'}(u)$. On pose aussi $P := \mathrm{mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E)$ et $Q := \mathrm{mat}_{\mathcal{C}',\mathcal{C}}(\mathrm{id}_F)$. L'objectif est de trouver le lien entre les matrices A et B. Pour $x \in E$, par (3.2) et (4.2), on a

$$\begin{aligned} \mathrm{mat}_{\mathcal{C}'}(u(x)) &= \mathrm{mat}_{\mathcal{C},\mathcal{C}'}(\mathrm{id}_F)\mathrm{mat}_{\mathcal{C}}(u(x)) = \mathrm{mat}_{\mathcal{C}',\mathcal{C}}(\mathrm{id}_F)^{-1}\mathrm{mat}_{\mathcal{C}}(u(x)) \\ &= \mathrm{mat}_{\mathcal{C}',\mathcal{C}}(\mathrm{id}_F)^{-1}\mathrm{mat}_{\mathcal{B},\mathcal{C}}(u)\mathrm{mat}_{\mathcal{B}}(x) \\ &= \mathrm{mat}_{\mathcal{C},\mathcal{C}'}(\mathrm{id}_F)^{-1}\mathrm{mat}_{\mathcal{B},\mathcal{C}}(u)\mathrm{mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E)\mathrm{mat}_{\mathcal{B}'}(x), \end{aligned}$$

c.-à.-d., on obtient

$$\begin{aligned} & \mathrm{mat}_{\mathcal{B}',\mathcal{C}'}(u) = \mathrm{mat}_{\mathcal{C}',\mathcal{C}}(\mathrm{id}_F)^{-1}\mathrm{mat}_{\mathcal{B},\mathcal{C}}(u)\mathrm{mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E) \\ \Longleftrightarrow \quad & B = Q^{-1}AP. \end{aligned} \tag{4.3}$$

Remarque 4.1.6. *Soit $x \in E$ et posons $X := \mathrm{mat}_{\mathcal{B}}(x), X' := \mathrm{mat}_{\mathcal{B}'}(x)$ et $Y := \mathrm{mat}_{\mathcal{C}}(u(x))$, $Y' := \mathrm{mat}_{\mathcal{C}'}(u(x))$. Alors, par (3.2) et (4.2), on a*

$$Y = AX, \qquad Y' = BX', \qquad X = PX', \qquad Y = QY'.$$

Ceci implique que

$$QY' = Y = AX = APX' \qquad \Longrightarrow \qquad Y' = Q^{-1}APX',$$

c.-à-d., $B = Q^{-1}AP$.

En résumé, on a montré la proposition suivante :

Proposition 4.1.7. *Soient $\mathcal{B}, \mathcal{B}'$ des bases de E et $\mathcal{C}, \mathcal{C}'$ des bases de F. Pour une application linéaire $u \in \mathcal{L}(E, F)$, on pose $A := \mathrm{mat}_{\mathcal{B},\mathcal{C}}(u)$ et $B := \mathrm{mat}_{\mathcal{B}',\mathcal{C}'}(u)$. Notons les matrices de passage de $\mathcal{B}$ à $\mathcal{B}'$ par P et de $\mathcal{C}$ à $\mathcal{C}'$ par Q. Alors,*

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

4.1.4 Transposée d'une matrice*

Ici, on introduit la transposée d'une matrice et on donne une des propriétés de cette opération. La signification de la transposée en terme d'application linéaire sera donnée en appendice § A.2.

Définition 4.1.5. *Soit $A = (a_{i,j})$ une matrice de taille (m, n). La **transposée** de A est la matrice $B = (b_{i,j})$ de taille (n, m) définie par $b_{i,j} := a_{j,i}$, pour tous $1 \leq i \leq n$ et $1 \leq j \leq m$. La transposée de A est notée par ${}^tA$.*

Voici une propriété importante :

Proposition 4.1.8. *Pour toutes matrices A de taille (l, m), B de taille (m, n), on a ${}^t(AB) = {}^tB{}^tA$.*

Preuve. Notons $C = AB, D = {}^t(AB), E = {}^tA, F = {}^tB$ et $G = {}^tB{}^tA$. D'après les définitions de la transposition et du produit, pour tout i avec $1 \leq i \leq l$ et tout j avec $1 \leq j \leq n$, on a

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

donc, on obtient $D = G$. □

Voici quelques types particuliers de matrices :

Définition 4.1.6. *Soit $A \in M_n(\mathbb{K})$ une matrice carrée.*

1. *A est dite **symétrique**, lorsque ${}^tA = A$.*
2. *A est dite **anti-symétrique**, lorsque ${}^tA = -A$.*

4.2 Matrices II : aspect pratique (05/11)

Ici, on explique une méthode très efficace pour calculer l'inverse d'une matrice carrée. Comme d'habitude, on fixe un corps commutatif $\mathbb{K}$.

4.2.1 Élimination de Gauss

Ici, on explique ce que l'on appelle l'**élimination de Gauss** ou le **pivot de Gauss**. On fixe $n \in \mathbb{N}^*$.

Soit $A = (a_{i,j})$ une matrice carrée de taille (n, n). Alors, l'inverse de A, s'il existe, est la matrice $A^{-1} = (x_{i,j})$ de taille (n, n) qui satisfait $A \cdot A^{-1} = \mathbf{1}_n$. Regardons chaque colonne, on voit bien que ceci équivaut aux systèmes d'équations linéaires

$$A \begin{pmatrix} x_{1,j} \\ \vdots \\ x_{j,j} \\ \vdots \\ x_{n,j} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} j \qquad 1 \leq j \leq n.$$

Donc, il nous suffit de trouver une méthode pour résoudre un système d'équations linéaires :

$$AX = B, \qquad \text{où} \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

$$\iff \qquad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1, \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2, \\ \qquad\qquad \vdots \qquad\qquad\quad \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = b_n. \end{cases}$$

Rappelons comment on peut résoudre ce système d'équations. On appelle la i-ième ligne par L_i.

<u>**1^er^ étape**</u> Supposons qu'un des coefficients $a_{i,1}$ ne soit pas nul, i.e., $a_{i_1,1} \neq 0$. Alors, en permutant L_1 avec L_{i_1}, on suppose que $a_{1,1} \neq 0$. Maintenant,

1. Renommer $\dfrac{1}{a_{1,1}} L_1$ comme L_1.
2. Pour $1 < i \leq n$, renommer $L_i - a_{i,1}L_1$ comme L_i.

On exprime cette nouvelle ligne comme $L_i : a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i$, pour économiser les symboles. Alors, le coefficient de x_1 de la nouvelle ligne L_i est 1 pour $i = 1$ et 0 pour $1 < i \leq n$, i.e., $a_{i,1} = \delta_{i,1}$.

<u>**2^e^ étape**</u> On fait la même chose pour x_2 avec $L_2, \cdots, L_n$. C'est-à-dire, on suppose qu'un des coefficients $a_{i,2}$ ne soit pas nul, i.e., $a_{i_2,2} \neq 0$ $(2 \leq i_2 \leq n)$. Alors, en permutant L_2 avec L_{i_2}, on suppose que $a_{2,2} \neq 0$. Maintenant,

1. Renommer $\dfrac{1}{a_{2,2}} L_2$ comme L_2.

2. Pour $1 \leq i \leq n$ tel que $i \neq 2$, renommer $L_i - a_{i,2}L_2$ comme L_i.

On exprime cette nouvelle ligne comme $L_i : a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i$, pour économiser les symboles. Alors, le coefficient de x_2 de la nouvelle L_i est 1 pour $i = 2$ et 0 pour $1 < i \leq n$ tel que $i \neq 2$, i.e., $a_{i,2} = \delta_{i,2}$. Notons que les coefficients $a_{i,1}$ ne changent pas dans cette étape.

On répète la même étape jusqu'à x_n : on obtient la solution

$$\begin{array}{llllr} x_1 & & & & = b'_1, \\ & x_2 & & & = b'_2, \\ & & \ddots & & = \vdots \\ & & & x_n & = b'_n. \end{array}$$

En résumé, on a utilisé les trois types d'opérations suivantes :

(I) Permuter la i-ième ligne et j-ième ligne ($i \neq j$),

(II) Multiplier la i-ième ligne par un scalaire non nul, et

(III) Ajouter un multiple scalaire de la j-ième à i-ième ligne ($i \neq j$).

Ce que l'on a fait ci-dessus est interprété en terme de matrice

$$\widetilde{A} := \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} & b_n \end{pmatrix},$$

comme suit. Appliquant des opérations (I), (II) et (III) d'une manière appropriée, on arrive à la matrice de la forme

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & b'_1 \\ 0 & 1 & \ddots & \vdots & b'_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & b'_n \end{pmatrix}.$$

(La seule différence est de ne pas écrire $x_1, \cdots, x_n$!) Pour trouver l'inverse de A, on applique ces opérations à la matrice

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

on trouve l'inverse $X = (x_{i,j})$ de A dans la forme

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ 0 & 1 & \ddots & \vdots & x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \ddots & \ddots & 0 & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 1 & x_{n,1} & x_{n,2} & \cdots & x_{n,n} \end{pmatrix}.$$

Cette méthode qui est extrêmement pratique et puissante est appelée l'**élimination de Gauss** ou le **pivot de Gauss**.

4.2.2 Cas général

Dans la sous-section précédente, on était vraiment optimiste. C'est-à-dire, on a considéré un système d'équations linéaires

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1, \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2, \\ \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = b_n, \end{cases}$$

et on a supposé que l'on pouvait aller jusqu'au bout. En général, il se peut que l'on ne puisse pas aller jusqu'au bout. On considère un tel cas, ici.

Supposons que l'on ait pu appliquer le pivot de Gauss jusqu'à r-ième étape avec $0 < r < n$, en changeant l'ordre d'arrangement de $x_1, \cdots, x_n$ si nécessaire, mais pas plus c.-à-d., on est arrivé à la forme suivante :

$$\begin{pmatrix} 1 & & & a_{1,r+1} & \cdots & a_{1,n} & b_1 \\ & \ddots & & \vdots & & \vdots & \vdots \\ & & 1 & a_{r,r+1} & \cdots & a_{r,n} & b_r \\ 0 & \cdots & 0 & 0 & \cdots & 0 & b_{r+1} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & b_n \end{pmatrix}$$

qui correspond au système d'équations linéaires

$$\begin{cases} x_1 & & & +a_{1,r+1}x_{r+1} + \cdots + a_{1,n}x_n = b_1, \\ & \ddots & & \vdots \\ & & x_r & +a_{r,r+1}x_{r+1} + \cdots + a_{r,n}x_n = b_r, \\ & & & 0 = b_{r+1}, \\ & & & \vdots \\ & & & 0 = b_n. \end{cases}$$

Donc, ce système admet une solution si et seulement si $b_{r+1} = \cdots = b_n = 0$. Dans ce cas, la solution générale est donnée par

$$\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix} - t_{r+1} \begin{pmatrix} a_{1,r+1} \\ \vdots \\ a_{r,r+1} \end{pmatrix} \cdots - t_n \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{r,n} \end{pmatrix}, \qquad x_i = t_i \quad r < i \leq n.$$

4.2.3 Interprétation matricielle

Ici, on interprète les opérations que l'on a utilisées avec le pivot de Gauss en termes matriciels.

En effet, les opérations (I), (II) et (III) correspondent à la multiplication d'une matirce à gauche; précisément,

l'opération (I) correspond à la multiplication de la matrice

$$P(i,j) := \left(\sum_{k \neq i,j} E_{k,k} \right) + E_{i,j} + E_{j,i}$$

à gauche,

l'opération (II) correspond à la multiplication de la matrice

$$D_i(t) := \left(\sum_{k \neq i} E_{k,k} \right) + tE_{i,i}$$

à gauche,

l'opération (III) correspond à la multiplication de la matrice

$$x_{i,j}(t) := \mathbf{1}_n + tE_{i,j}$$

à gauche. Cette matrice est appelée une **transvection**.

Une question naïve se pose : *si on multiplie ces matrices à droite, que va-t-il se passer ?*

La réponse est que la multiplication de ces matrices à gauche correspond à certaines opérations sur les **lignes** tandis que celle à droite correspond à certaines opérations sur les **colonnes**. C'est-à-dire,

(I') la multiplication de la matrice $P(i,j)$ à droite fait permuter la i-ième et j-ième colonne,

(II') la multiplication de la matrice $D_i(t)$ à droite fait multiplier t sur la i-ième cologne, et

(III') la multiplication de la matrice $x_{i,j}(t)$ à droite fait ajouter t fois la j-ième colonne à i-ième colonne.

Pourtant, en ce qui concerne les systèmes d'équations linéaires, il est naturel de considérer les opérations sur les lignes. C'est la raison pour laquelle, on a traité les opérations sur les lignes en détail.

Remarque 4.2.1. *Par la formule*

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

et les arguments ci-dessus, on voit que toute matrice carrée inversible peut s'écrire comme le produit de matrices diagonales inversibles et des transvections.

4.2.4 Le rang d'une matrice

Rappelons la démonstration de la formule du rang. Soient E, F des $\mathbb{K}$-espaces vectoriels de dimension n et m respectivement et u une application linéaire. La clef est que l'on peut trouver des bases $\mathcal{B} = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ de E et $\mathcal{C} = \{\mathbf{f}_1, \cdots, \mathbf{f}_m\}$ de F telles que

1. $\{\mathbf{e}_{r+1}, \cdots, \mathbf{e}_n\}$ est une base de Ker u,
2. $u(\mathbf{e}_i) = \mathbf{f}_i$ pour $1 \leq i \leq r$.

Dans ces bases, la matrice de u a la forme simple suivante

$$\mathrm{mat}_{\mathcal{B},\mathcal{C}}(u) = \begin{pmatrix}1 & & & 0 & \cdots & 0\\ & \ddots & & \vdots & & \vdots\\ & & 1 & 0 & \cdots & 0\\ 0 & \cdots & 0 & 0 & \cdots & 0\\ \vdots & & \vdots & \vdots & & \vdots\\ 0 & \cdots & 0 & 0 & \cdots & 0\end{pmatrix}, \tag{4.4}$$

avec r nombres de 1. Ce nombre r est la dimension de $\mathrm{Im} u$, donc le rang $\mathrm{rg}(u)$.

En général, soient $u \in \mathcal{L}(E, F)$ et $\mathcal{B} \subset E$, $\mathcal{C} \subset F$ des bases. Posons $A := \mathrm{mat}_{\mathcal{B},\mathcal{C}}(u)$. L'argument ci-dessus avec la proposition 4.1.7 montre qu'il existe des matrices inversibles $P \in M_m(\mathbb{K})$ et $Q \in M_n(\mathbb{K})$ telles que PAQ soit une matrice de la forme (4.4). Dans ce cas, on dit que **le rang de la matrice A est r**. Le lemme suivant est un corollaire immédiate de la propsotion 4.1.8 :

Lemme 4.2.2. *Pour toute matrice $A \in M_{m,n}(\mathbb{K})$, on a $\mathrm{rg}({}^tA) = \mathrm{rg}(A)$.*

D'après la sous-section précédente, on voit que toute la matrice peut se transformer sous la forme (4.4) par les opérations (I) $\sim$ (III) et (I') $\sim$ (III'), c.-à-d., `le rang d'une matrice peut être calculé par le pivot de Gauss`.

Au point de vue pratique, cela demande un peu trop de calculs... En effet, on peut calculer le rang d'une matrice avec les opérations sur les lignes sans utiliser celles sur les colonnes. Par opérations sur les lignes, on peut toujours ramener une matrice $A \in M_{m,n}(\mathbb{K})$ à la forme

$$\begin{pmatrix} 1 & & & a_{1,r+1} & \cdots & a_{1,n} \\ & \ddots & & \vdots & & \vdots \\ & & 1 & a_{r,r+1} & \cdots & a_{r,n} \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Alors le rang $\mathrm{rg}(A)$ est r, car le rang est, par définition, le nombre de vecteurs en colonne (où en ligne, c'est la même chose d'après le lemme 4.2.2).

La remarque dans la sous-section précédente se généralise comme suit :

Remarque 4.2.3. *Toute matrice carrée peut s'écrire comme le produit de matrices diagonales et de transvections, car elle est le produit de matrices inversibles et d'une matrice de la forme (4.4).*

4.3 Déterminant I : aspect théorique (19/11)

Ici, on va parler du déterminant de matrices carrées et de ses applications.

4.3.1 Rappels sur les groupes symétriques

Soit $n \in \mathbb{N}$ tel que $n > 1$. Rappelons que le **groupe des permutations de degré** n, noté $\mathfrak{S}_n$, est l'ensemble

$$\{\sigma : \{1, 2, \cdots, n\} \longrightarrow \{1, 2, \cdots, n\} \mid \sigma \text{ bijective}\}$$

muni de la composition d'applications. Le groupe est aussi appelé le **groupe symétrique d'indice** n. L'ordre de ce groupe est $n!$, c.-à-d., il y a $n!$ applications bijectives de $\{1, 2, \cdots, n\}$ vers lui-même.

Un élément de $\mathfrak{S}_n$ est dit une **permutation** et, pour $1 \leq i \neq j \leq n$, la permutation σ définie par

$$\sigma(k) := \begin{cases} j & \text{si } k = i, \\ i & \text{si } k = j, \\ k & \text{sinon,} \end{cases}$$

est appellée aussi une **transposition** et notée par (i, j). Alors,

Lemme 4.3.1. $\mathfrak{S}_n$ *est engendré par les transpositions, autrement dit, toute permutation se décompose en produit de transpositions.*

Preuve. Procédons par récurrence. Dans le caas $n = 2$, l'ennoncé est clair (car $\mathfrak{S}_2 = \{\mathrm{id}, (1,2)\}$). Supposons le cas du rang $n-1$. Montrons pour le rang n. Soit $\sigma \in \mathfrak{S}_n$. Si $\sigma(n) = n$, σ permute $\{1, 2, \cdots, n-1\}$, donc par récurrence, elle se décompose en produit de transpositions. Sinon, posons $i := \sigma(n)$, la permutation $(i, n) \circ \sigma$ fixe n, donc, ce cas peut se déduire du cas où $\sigma(n) = n$. □

Remarque 4.3.2. *Pour $1 \leq i < n$, posons $\sigma_i = (i, i+1)$. Alors, pour $1 \leq i < j \leq n$, on peut vérifier (exercice !)*

$$\begin{aligned}(i,j) =& \sigma_i \circ \sigma_{i+1} \circ \cdots \circ \sigma_{j-2} \circ \sigma_{j-1} \circ \sigma_{j-2} \circ \cdots \circ \sigma_{i+1} \circ \sigma_i \\ =& \sigma_{j-1} \circ \sigma_{j-2} \circ \cdots \circ \sigma_{i+1} \circ \sigma_i \circ \sigma_{i+1} \circ \cdots \circ \sigma_{j-2} \circ \sigma_{j-1}.\end{aligned}$$

Avec le lemme ci-dessus, on voit que $\mathfrak{S}_n$ est engendré par $\sigma_1, \sigma_2, \cdots, \sigma_{n-1}$.

Maintenant, on considère l'application sgn de $\mathfrak{S}_n$ dans $\mathbb{R}^*$ définie par

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

Alors, on peut montrer que (exercice !)

1. $\mathrm{Im}\,\mathrm{sgn} = \{\pm 1\}$,
2. pour $\sigma, \tau \in \mathfrak{S}_n$, $\mathrm{sgn}(\sigma \circ \tau) = \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau)$.

C'est-à-dire, sgn est un morphisme de groupes de $\mathfrak{S}_n$ dans $\{\pm 1\}$.

Remarque 4.3.3. *Le noyau de ce morphisme* $\mathrm{Ker}\,\mathrm{sgn} := \{\sigma \in \mathfrak{S}_n | \mathrm{sgn}(\sigma) = 1\}$ *est le sous-groupe de $\mathfrak{S}_n$ d'indice 2 appelé le* ***groupe alterné de degré*** *n et noté $\mathfrak{A}_n$.*

4.3.2 Définition et quelques propriétés

Fixons $n \in \mathbb{N}$ tel que $n > 1$. Pour $A = (a_{i,j})_{1 \leq i,j \leq n} \in M_n(\mathbb{K})$, on aussi note $A = (\mathbf{a}_1, \cdots, \mathbf{a}_n)$ où $\mathbf{a}_j = {}^t(a_{1,j}, a_{2,j}, \cdots, a_{n,j})$. Alors,

Lemme 4.3.4. *Il existe une application F non triviale de $M_n(\mathbb{K})$ dans $\mathbb{K}$ telle que*

1. *(Multilinéaire) pour $1 \leq j \leq n$ et $\lambda \in \mathbb{K}$,*

$$\begin{aligned}& F(\mathbf{a}_1, \cdots, \overbrace{\mathbf{a}_j + \lambda \mathbf{a}'_j}^{j}, \cdots, \mathbf{a}_n) \\ =& F(\mathbf{a}_1, \cdots, \overbrace{\mathbf{a}_j}^{j}, \cdots, \mathbf{a}_n) + \lambda F(\mathbf{a}_1, \cdots, \overbrace{\mathbf{a}'_j}^{j}, \cdots, \mathbf{a}_n),\end{aligned}$$

2. *(Antisymétrique)* *pour* $\sigma \in \mathfrak{S}_n$,

$$F(\mathbf{a}_{\sigma(1)}, \cdots, \mathbf{a}_{\sigma(n)}) = (\mathrm{sgn}\sigma)F(\mathbf{a}_1, \cdots, \mathbf{a}_n).$$

L'antisymétrie pour $\sigma = (i, j)$ implique que

$$F(\mathbf{a}_1, \cdots, \underbrace{\mathbf{a}_j}^{i}, \cdots, \underbrace{\mathbf{a}_i}^{j}, \cdots, \mathbf{a}_n) = -F(\mathbf{a}_1, \cdots, \underbrace{\mathbf{a}_i}^{i}, \cdots, \underbrace{\mathbf{a}_j}^{j}, \cdots, \mathbf{a}_n).$$

(D'après le lemme 4.3.1, l'antisymétrie pour tout (i, j) $(i \neq j)$ implique l'antisymétrie pour tout $\sigma \in \mathfrak{S}_n$.) En particulier, on a

$$F(\mathbf{a}_1, \cdots, \underbrace{\mathbf{a}}^{i}, \cdots, \underbrace{\mathbf{a}}^{j}, \cdots, \mathbf{a}_n) = 0. \tag{4.5}$$

Preuve. Notons $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ la base canonique de $\mathbb{K}^n$. Alors, on a

$$\begin{aligned} F(\mathbf{a}_1, \cdots, \mathbf{a}_n) =& F(\sum_{i=1}^{n} a_{i,1}\mathbf{e}_i, \cdots, \sum_{i=1}^{n} a_{i,n}\mathbf{e}_i) \\ =& \sum_{i_1,\cdots,i_n} a_{i_1,1}a_{i_2,2}\cdots, a_{i_n,n}F(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \cdots, \mathbf{e}_{i_n}). \end{aligned}$$

Par (4.5), s'il existe $j \neq k$ tel que $i_j = i_k$, alors $F(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \cdots, \mathbf{e}_{i_n}) = 0$, donc on peut supposer que $i_1, i_2, \cdots, i_n$ sont tous différents. Donc,

$$\begin{aligned} F(\mathbf{a}_1, \cdots, \mathbf{a}_n) =& \sum_{\{i_1,\cdots,i_n\}=\{1,\cdots,n\}} a_{i_1,1}a_{i_2,2}\cdots, a_{i_n,n}F(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \cdots, \mathbf{e}_{i_n}) \\ =& \sum_{\sigma\in\mathfrak{S}_n} a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n}F(\mathbf{e}_{\sigma(1)}, \mathbf{e}_{\sigma(2)}, \cdots, \mathbf{e}_{\sigma(n)}). \end{aligned}$$

Par l'antisymétrie, on a $F(\mathbf{e}_{\sigma(1)}, \cdots, \mathbf{e}_{\sigma(n)}) = (\mathrm{sgn}\sigma)F(\mathbf{e}_1, \cdots, \mathbf{e}_n)$, ce qui implique que

$$F(\mathbf{a}_1, \cdots, \mathbf{a}_n) = \sum_{\sigma\in\mathfrak{S}_n} (\mathrm{sgn}\sigma)a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n}F(\mathbf{e}_1, \cdots, \mathbf{e}_n).$$

Posons $c := F(\mathbf{e}_1, \cdots, \mathbf{e}_n)$, on obtient une application non triviale. Si l'application F est donnée par cette formule, il est facile de vérifier qu'une telle application est multilinéaire et antisymétrique. □

Définition 4.3.1. *La valeur de* $A \in M_n(\mathbb{K})$ *par* F *du lemme 4.3.4 normalisée par* $F(\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n) = 1$ *est dit le* ***déterminant*** *de* A*, noté* $\det A$*, ou encore* $|A|$*. Pour une matrice* $A = (a_{i,j}) \in M_n(\mathbb{K})$*, il est aussi noté comme suit*

$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}.$$

Remarque 4.3.5. *D'après la démonstration du lemme 4.3.4, on a pour* $A = (a_{i,j})$

$$\begin{aligned}\det A &= \sum_{\sigma\in\mathfrak{S}_n} (\mathrm{sgn}\sigma) a_{\sigma(1),1}a_{\sigma(2),2}\cdots a_{\sigma(n),n} \\ &= \sum_{\sigma\in\mathfrak{S}_n} (\mathrm{sgn}\sigma) a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{n,\sigma(n)}.\end{aligned}$$

La remarque ci-dessus implique que

Lemme 4.3.6. *Pour tout* $A \in M_n(\mathbb{K})$*, on a* $\det({}^tA) = \det(A)$.

Ce lemme implique que `le déterminant est non seulement multilinéaire et antisymétrique pour les colonnes mais aussi pour les lignes !`
Le lemme suivant est aussi important :

Lemme 4.3.7. *Pour* $A, B \in M_n(\mathbb{K})$*, on a* $\det(AB) = \det(A)\det(B)$.

Preuve. Notons $A = (\mathbf{a}_1, \cdots, \mathbf{a}_n)$ et $B = (b_{i,j})$. Alors, il est clair que

$$AB = (\sum_{i=1}^n b_{i,1}\mathbf{a}_i, \cdots, \sum_{i=1}^n b_{i,n}\mathbf{a}_i).$$

Donc, par un argument similaire à la démonstration du lemme 4.3.4, on a

$$\begin{aligned}\det(AB) &= \det(\sum_{i=1}^n b_{i,1}\mathbf{a}_i, \cdots, \sum_{i=1}^n b_{i,n}\mathbf{a}_i) \\ &= \sum_{\sigma\in\mathfrak{S}_n} b_{\sigma(1),1}b_{\sigma(2),2}\cdots b_{\sigma(n),n}\det(\mathbf{a}_{\sigma(1)}, \cdots, \mathbf{a}_{\sigma(n)}) \\ &= \det(\mathbf{a}_1, \cdots, \mathbf{a}_n)\left(\sum_{n\in\mathfrak{S}_n} (\mathrm{sgn}\sigma) b_{\sigma(1),1}b_{\sigma(2),2}\cdots b_{\sigma(n),n}\right) \\ &= \det(A)\det(B).\end{aligned}$$

□

Ce lemme implique un corollaire important :

Corollaire 4.3.8. *Si une matrice carrée* $A \in M_n(\mathbb{K})$ *est inversible, on a* $\det(A) \neq 0$.

Preuve. En effet, si A est inversible, il existe une matrice $B \in M_n(\mathbb{K})$ telle que $AB = BA = \mathbf{1}_n$. Donc, par le lemme ci-dessus, on a $\det(A)\det(B) = \det(\mathbf{1}_n) = 1$. □

Plus tard, on montrera que cette condition est aussi suffisante.

Exemple 4.3.2. *Soit* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{K})$*. Alors, on a*

$$\begin{aligned}\det(A) &= \det(a\mathbf{e}_1 + c\mathbf{e}_2, b\mathbf{e}_1 + d\mathbf{e}_2) = \det(a\mathbf{e}_1, d\mathbf{e}_2) + \det(c\mathbf{e}_2, b\mathbf{e}_1) \\ &= (ad - bc)\det(\mathbf{e}_1, \mathbf{e}_2) = ad - bc.\end{aligned}$$

4.3.3 Interprétation géométrique*

Pour une matrice carrée $A = (\mathbf{a}_1, \cdots, \mathbf{a}_n) \in M_n(\mathbb{K})$, soit

$$\Gamma_A := \{\sum_{i=1}^{n} t_i \mathbf{a}_i | 0 \leq t_i \leq 1\}$$

le parallélotope engendré par $\mathbf{a}_1, \ldots, \mathbf{a}_n$. En particulier, dans le cas $n = 2$, ce n'est que le parallélogramme engendré par $\mathbf{a}_1$ et $\mathbf{a}_2$.

Considérons le cas $n = 2$. Par (4.5), on a

$$\det(\mathbf{a}_1, \mathbf{a}_2) = \det(\mathbf{a}_1, \mathbf{a}_2 + t\mathbf{a}_1) = \det(\mathbf{a}_1 + t\mathbf{a}_2, \mathbf{a}_2), \qquad t \in \mathbb{K}.$$

Alors, la forme des parallelogrammes changent mais ses aires sont toujours les mêmes, car ce sont des transvections ! Avec les transvections, on peut se ramener au cas où Γ_A devient un rectangle, en particulier, A est une matrice diagonale $\begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix}$ qui a le déterminant égal à pq. Evidement, l'aire du rectangle engendré par $(p, 0)$ et $(0, q)$ est $|pq|$, ceci implique que $|\det(A)| = \mathrm{vol}(\Gamma_A)$, où $\mathrm{vol}(\Gamma_A)$ signifie le volume de Γ_A.
Cet argument se généralise pour n général, et on en conclut que

$$|\det A| = \mathrm{vol}(\Gamma_A).$$

Remarque 4.3.9. *Soit $u \in \mathcal{L}(\mathbb{K}^n)$ et $\mathcal{B}$ la base canonique de $\mathbb{K}^n$. Posons $A = \mathrm{mat}_{\mathcal{B}}(u)$. Par définition, on a $\Gamma_A = u(\Gamma_{\mathbf{1}_n})$. Découpant la carte cartésienne en petits morceaux, on voit que, pour tout $D \subset \mathbb{K}^n$ dont cela a du sens de parler de son volume, la formule suivante est valable :*

$$\mathrm{vol}(u(D)) = |\det A| \mathrm{vol}(D).$$

Les détails seront traités plus tard en analyse.

4.4 Déterminant II : aspect pratique (03/12)

Ici, on va présenter une méthode pratique pour calculer le déterminant d'une matrice carrée. Comme d'habitude, $\mathbb{K}$ est un corps commutatif.

4.4.1 Cas simple

Soit $n \in \mathbb{N}$ tel que $n > 1$. On commence par une observation

$$\begin{vmatrix} t_1 & 0 & \cdots & 0 \\ 0 & t_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & t_n \end{vmatrix} = \prod_{i=1}^{n} t_i.$$

Soient $p, q \in \mathbb{N}$ tels que $p, q > 1$ et $A = (a_{i,j}) \in M_p(\mathbb{K})$, $B = (b_{i,j}) \in M_q(\mathbb{K})$. Le but est de calculer le déterminant de la matrice $M := \begin{pmatrix} A & \mathbf{0}_{p,q} \\ \mathbf{0}_{q,p} & B \end{pmatrix} \in M_{p+q}(\mathbb{K})$, où $\mathbf{0}_{k,l}$ est la matrice nulle dans $M_{k,l}(\mathbb{K})$. Notons que $M = \begin{pmatrix} A & \mathbf{0}_{p,q} \\ \mathbf{0}_{q,p} & \mathbf{1}_q \end{pmatrix} \begin{pmatrix} \mathbf{1}_p & \mathbf{0}_{p,q} \\ \mathbf{0}_{q,p} & B \end{pmatrix}$, d'après le lemme 4.3.7, il nous suffit de calculer le déterminant de chaque matrice. Calculons celui de la première d'abord. Par définition, on a

$$\begin{aligned}
\det \begin{pmatrix} A & \mathbf{0}_{p,q} \\ \mathbf{0}_{q,p} & \mathbf{1}_q \end{pmatrix} &= \det(\sum_{i=1}^{p} a_{i,1}\mathbf{e}_i, \cdots, \sum_{i=1}^{p} a_{i,p}\mathbf{e}_i, \mathbf{e}_{p+1}, \cdots, \mathbf{e}_{p+q}) \\
&= \sum_{1 \le i_1, \cdots, i_p \le p} a_{i_1,1} \cdots a_{i_p,p} \det(\mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_p}, \mathbf{e}_{p+1}, \cdots, \mathbf{e}_{p+q}) \\
&= \sum_{\sigma \in \mathfrak{S}_p} a_{\sigma(1),1} \cdots a_{\sigma(p),p} \det(\mathbf{e}_{\sigma(1)}, \cdots, \mathbf{e}_{\sigma(p)}, \mathbf{e}_{p+1}, \cdots, \mathbf{e}_{p+q}) \\
&= \sum_{\sigma \in \mathfrak{S}_p} (\mathrm{sgn}\sigma) a_{\sigma(1),1} \cdots a_{\sigma(p),p} \det(\mathbf{e}_1, \cdots, \mathbf{e}_p, \mathbf{e}_{p+1}, \cdots, \mathbf{e}_{p+q}) \\
&= \det(A).
\end{aligned}$$

Maintenant, calculons celui de la deuxième. Pour le cas $p = 1$, on a

$$\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & b_{1,1} & \cdots & b_{1,q} \\ \vdots & \vdots & & \vdots \\ 0 & b_{q,1} & \cdots & b_{q,q} \end{vmatrix} = (-1)^q \begin{vmatrix} 0 & \cdots & 0 & 1 \\ b_{1,1} & \cdots & b_{1,q} & 0 \\ \vdots & & \vdots & \vdots \\ b_{q,1} & \cdots & b_{q,q} & 0 \end{vmatrix} = \begin{vmatrix} b_{1,1} & \cdots & b_{1,q} & 0 \\ \vdots & & \vdots & \vdots \\ b_{q,1} & \cdots & b_{q,q} & 0 \\ 0 & \cdots & 0 & 1 \end{vmatrix},$$

par l'antisymétrie. Pour un p quelconque, c'est pareil et on obtient que

$$\det \begin{pmatrix} \mathbf{1}_p & \mathbf{0}_{p,q} \\ \mathbf{0}_{q,p} & B \end{pmatrix} = \det(B),$$

on en déduit que $\det(M) = \det(A)\det(B)$. Pour le cas général, on peut montrer par récurrence :

Lemme 4.4.1. *Soit M une matrice diagonale par blocs, i.e., il existe $A_i \in M_{n_i}(\mathbb{K})$ $(1 \le i \le k)$ telles que*

$$M = \begin{pmatrix} A_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & A_k \end{pmatrix}.$$

Alors, on a $\det(M) = \prod_{i=1}^{k} \det(A_i)$.

4.4.2 Développement en cofacteurs

Ici, on va expliquer une méthode pour calculer le déterminant.

Soit $A = (a_{i,j}) \in M_n(\mathbb{K})$. On va calculer $\det(A)$. Pour $1 \leq i, j \leq n$, posons

$$A_{i,j} := (-1)^{i+j} \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}.$$

$A_{i,j}$ est apellé le **cofacteur associé à** (i,j) de la matrice A et $(-1)^{i+j}A_{i,j}$ est appelé la **mineur associé à** (i,j) de la matrice A.

Par multilinéarité, pour la i-ième ligne, on a

$$\det(A) = \sum_{j=1}^{n} a_{i,j} \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}.$$

Dans cette formule, le coefficient de $a_{i,j}$ est égal à $A_{i,j}$ pour la raison suivante. En effet, il est égal à

$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix},$$

par (4.5) et ce dernier est égal à $A_{i,j}$ par antisymétrie et le lemme 4.4.1. D'après le lemme 4.3.6, ceci implique que l'on peut aussi développer par rapport à la j-ième colonne ($1 \leq j \leq n$). En résumé, on obtient

Proposition 4.4.2. *Pour $1 \leq i, j \leq n$, on a:*

1. *développement par rapport à la i-ième ligne*

$$\det A = \sum_{j=1}^{n} a_{i,j} A_{i,j},$$

2. *développement par rapport à la j-ième colonne*

$$\det A = \sum_{i=1}^{n} a_{i,j} A_{i,j}.$$

Par cette proposition et (4.5), on voit bien que

$$\sum_{j=1} a_{i,j} A_{i',j} = \delta_{i,i'} \det(A), \qquad \sum_{i=1}^{n} a_{i,j} A_{i,j'} = \delta_{j,j'} \det(A).$$

C'est-à-dire, posons $\mathrm{com}(A) := (A_{i,j})_{1 \leq i,j \leq n} \in M_n(\mathbb{K})$, on a

$$A \cdot {}^t\mathrm{com}(A) = {}^t\mathrm{com}(A) \cdot A = (\det(A))\mathbf{1}_n.$$

La matrice $\mathrm{com}(A)$ est appelée la **comatrice** de A. Cette formule implique la récirpoque du corollaire 4.3.8. Donc, on obtient la proposition suivante :

Proposition 4.4.3. *Soit $A \in M_n(\mathbb{K})$.*

1. *La matrice A admet un inverse si et seulement si* $\det(A) \neq 0$.

2. *Lorsque* $\det(A) \neq 0$, *on a*

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

Remarque 4.4.4. *Attention! En général, le deuxième énoncé de cette proposition n'est pas utile, même s'il montre théoriquement qu'une matrice carrée A est inversible si* $\det(A) \neq 0$.

4.4.3 Exemples

On a déjà considéré le déterminant d'une matrice diagonale, ainsi que d'une matrice diagonale par blocs. Ici, on présente quelques autres cas qu'il faut connaître.

Le premier cas est celui des **matrices triangulaires**. Une matrice carrée $A = (a_{i,j}) \in M_n(\mathbb{K})$ est dite **triangulaire supérieure** si $a_{i,j} = 0$ pour $i > j$ et **triangulaire inférieure** si $a_{i,j} = 0$ pour $i < j$. Notons que la transposée d'une matrice triangulaire supérieure est une matrice triangulaire inférieure et vice-versa. Donc, d'après le lemme 4.3.6, il nous suffit de considérer, par exemple, une matrice triangulaire. Par le développement par rapport à la première colonne (cf. proposition 4.4.2), on a

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

Donc, par récurrence, on obtient que $\det(A) = \prod_{i=1}^n a_{i,i}$, c.-à-d., le déterminant de A est le produit des composantes diagonales de A.

Le deuxième cas est celui des matrices de **Vandermonde** de taille (n, n). Soient $\lambda_1, \cdots, \lambda_n \in \mathbb{K}^*$. La matrice dont on s'intéresse ici est

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}.$$

Calculons $\det(V)$. Par (4.5) et l'antisymétrie, ajoutant $-\lambda_1$ fois la i-ième ligne à $(i+1)$-ième ligne, dans l'ordre de $i = n-1, n-2, \cdots, 1$, on voit que

$$\det(V) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & (\lambda_2 - \lambda_1) & \cdots & (\lambda_n - \lambda_1) \\ \vdots & \vdots & & \vdots \\ 0 & \lambda_2^{n-2}(\lambda_2 - \lambda_1) & \cdots & \lambda_n^{n-2}(\lambda_n - \lambda_1) \end{vmatrix}.$$

Appliquant le développement par rapport à la première colonne, par la multilinéarité, on a

$$\det(V) = \prod_{i=2}^n (\lambda_i - \lambda_1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_2 & \lambda_3 & \cdots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_2^{n-2} & \lambda_3^{n-2} & \cdots & \lambda_n^{n-2} \end{vmatrix}.$$

Donc, par récurrence, on obtient que

$$\det(V) = \prod_{1 \le i < j \le n} (\lambda_j - \lambda_i).$$

Ce déterminant est appelé le **déterminant de Vandermonde**. Cette formule est très utile.

Le troisième cas est celui des **matrices circulaires**. Soient $a_0, a_1, \cdots, a_{n-1} \in \mathbb{K}$, on considère la matrice

$$C := \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ a_1 & \cdots & a_{n-1} & a_0 \end{pmatrix}.$$

Soit ω une n-ième racine de l'unité, e.g., $\omega = \exp\left(\frac{2\pi\sqrt{-1}}{n}\right)$. Pour $0 \le i < n$, on pose $\mathbf{v}_i := {}^t(1, \omega^i, \cdots, \omega^{(n-1)i})$. Posons $Q(\lambda) := \sum_{k=0}^{n-1} a_k \lambda^k$, on voit que

$C\mathbf{v}_i = Q(\omega^i)\mathbf{v}_i$ pour $0 \leq i < n$. Donc, posons $P = (\mathbf{v}_0, \mathbf{v}_1, \cdots, \mathbf{v}_{n-1})$ et

$$D := \begin{pmatrix} Q(1) & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & Q(\omega^{n-1}) \end{pmatrix},$$

on obtient $CP = PD$. Par le déterminant de Vandermonde, on a $\det(P) = \prod_{0 \leq i \leq j < n}(\omega^j - \omega^i) \neq 0$, donc P est inversible d'après la proposition 4.4.3. Donc, les lemmes 4.3.7 et 4.4.1 impliquent que

$$\begin{aligned} \det(C) &= \det(PDP^{-1}) = \det(D) = \prod_{i=0}^{n-1} Q(\omega^i) \\ &= \prod_{i=0}^{n-1}(a_0 + a_1\omega^i + \cdots + a_{n-1}\omega^{(n-1)i}). \end{aligned}$$

Chapter 5

Polynômes et Fractions rationelles

5.1 Polynôme (10/12)

Soit A un anneau commutatif et unitaire.

5.1.1 Définition

L'anneau des polynômes en une indéterminée, on notera $A[X]$ cet anneau, est défini de la façon suivante. Un monôme est une expression de la forme

$$\lambda X^m$$

où $\lambda \in A$ et $m \in \mathbb{N}$. Ici, par convention, on pose $X^0 := 1$. Un **polynôme** est une somme d'un nombre fini de monômes. L'addition et la multiplication s'éffectuent $\ll$ comme on l'imagine $\gg$. C'est-à-dire, pour $P = \sum_{i=0}^{m} a_i X^i, Q = \sum_{j=0}^{n} b_j X^j \in A[X]$, l'addition est définie par

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

où on pose $a_k = 0$, pour $k > m$, et $b_k = 0$, pour $k > n$. La multiplication est définie par

$$PQ := \sum_{k=0}^{m+n} \left(\sum_{\substack{i,j \geq 0 \\ i+j=k}} a_i b_j \right) X^k.$$

Avec ces deux opérations,

Lemme 5.1.1. *$A[X]$ est un anneau commutatif et unitaire.*

Définition 5.1.1. *Soit* $P = \sum_{i=0}^{n} a_i X^i \in A[X]$ *tel que* $a_n \neq 0$. *Le* ***degré*** *de* P *est l'entier* n, *noté* $\deg(P)$ *ou* $d^\circ(P)$. *Par convention, on pose* $\deg(0) := -\infty$.

Remarque 5.1.2. *La définition d'un polynôme donnée ci-dessus est imprécise. Si cela vous pose problème, voici une autre définition. Considérons l'ensemble* $A^{\mathbb{N}}$ *des éléments de la forme* $(a_i)_{i\in\mathbb{N}}$ *tels que* $a_i \in A$ *et que* a_i *soit nul pour* i *suffisament grand. Soient* $a = (a_i)_{i\in\mathbb{N}}$ *et* $b = (b_i)_{i\in\mathbb{N}}$ *deux éléments de* $A^{\mathbb{N}}$. *On pose*

$$a + b := (a_i + b_i)_{i\in\mathbb{N}},$$

et

$$a \cdot b := c, \qquad \textit{avec} \qquad c_m = \sum_{k=0}^{m} a_k b_{m-k}.$$

Alors, $A^{\mathbb{N}}$ *est isomorphe à* $A[X]$ *en associant* $(0, \ldots, 0, \overset{i}{\smile}{1}, 0, \ldots)$ *à* X^i.

5.1.2 Division euclidienne

Ici, on considère le cas où A est un corps commutatif $\mathbb{K}$. Dans ce cas, on a

Lemme 5.1.3. *L'anneau* $\mathbb{K}[X]$ *est intègre.*

Preuve. Soient $P = \sum_{i=0}^{m} a_i X^i$ et $Q = \sum_{j=0}^{n} b_j X^j$ deux polynômes dans $\mathbb{K}[X]$ tels que $a_m, b_n \in \mathbb{K}^*$. Alors, le terme de plus haut degré de PQ est $a_m b_n X^{m+n}$ qui est non nul car $a_m b_n \neq 0$ par hypothèse, d'où $PQ \neq 0$. □

Remarque 5.1.4. *Par la démonstration ci-dessus, on peut généraliser ce résultat de la façon suivante: pour un anneau intègre unitaire* A, *l'anneau des polynômes* $A[X]$ *est aussi intègre.*

De plus, on a l'additivité du degré :

Lemme 5.1.5. *Soient* P, Q *deux polynômes dans* $\mathbb{K}[X]$. *Alors,*

$$\deg(PQ) = \deg(P) + \deg(Q).$$

Preuve. Exercice. □

D'abord, rappelons la division euclidienne pour l'anneau $\mathbb{Z}$ des entiers relatifs. Soient a un entier et $b \geq 1$ un entier strictement positif. Alors, il existe un et un seul couple (q, r) d'entiers vérifiant les conditions suivantes :

$$a = bq + r \qquad 0 \leq r < b.$$

Ici, on montre que cette division se généralise au cas de l'anneau $\mathbb{K}[X]$, c.-à-d.,

Théorème 5.1.6. *Soient P, Q deux polynômes de $\mathbb{K}[X]$. On suppose que $\deg(Q) > 0$. Alors, il existe un unique couple de polynômes (R, S) dans $\mathbb{K}[X]$ vérifiant les conditions suivantes :*

1. $P = QS + R$,
2. $\deg(R) < \deg(Q)$.

Lorsque $R = 0$ dans ce théorème, on dit que Q **divise** P.

Preuve. Montrons d'abord l'unicité. Si $P = QS + R = QS' + R'$, alors $Q(S' - S) = R - R'$ et de degré au plus $\max\{\deg(R), \deg(R')\} < \deg(Q)$. Supposons que $S \neq S'$, i.e., $S' - S \neq 0$. Alors, on a

$$\deg(Q) > \deg(R - R') = \deg(Q(S' - S)) = \deg(Q) + \deg(S' - S) \geq \deg(Q),$$

ce qui contredit, c.-à-d., $R' = R$ et $S' = S$.
Montrons maintenant l'existence du couple (R, S) comme dans le théorème. Notons $c_Q X^{\deg(Q)}$ le terme de plus haut degré de Q. On raisonne par récurrence sur le degré de P. Si $\deg(P) < \deg(Q)$, il suffit de poser $R = 0$ et $S = P$. Sinon, soit $c_P X^{\deg(P)}$ le terme de plus haut degré de P. Alors, $P' := P - c_P c_Q^{-1} X^{\deg(P) - \deg(Q)} Q$ est un polynôme de degré au plus $\deg(P)$ dont le coefficient du terme de degré $\deg(P)$ égal à $c_P - c_P c_Q^{-1} c_Q = 0$. Ainsi, $\deg(P') < \deg(P)$. Par récurrence, il existe un couple de polynômes (R', S') tels que

$$P' = QS' + R', \qquad \deg(R') < \deg(Q).$$

Alors, on a

$$P = P' + c_P c_Q^{-1} X^{\deg(P) - \deg(Q)} Q = (c_P c_Q^{-1} X^{\deg(P) - \deg(Q)} + S')Q + R'.$$

Donc, il suffit de poser $S = S' + c_P c_Q^{-1} X^{\deg(P) - \deg(Q)}$ et $R = R'$. □

Comme dans le cas de $\mathbb{Z}$, une conséquence importante du théorème 5.1.6 est ce que l'on appelle l'**identité de Bézout** :

Corollaire 5.1.7. *Soient P, Q deux polynômes dans $\mathbb{K}[X]$ non nuls tels que $\deg(P) \geq \deg(Q) > 0$. Alors, il existe un couple de polynômes (A, B) dans $\mathbb{K}[X]$ tel que*

$$AP + BQ = D,$$

où $D := PGCD(P, Q)$.

Une autre conséquence utile concerne la factorisation d'un polynôme. Soit $P = \sum_{i=1}^{n} a_i X^i$ un polynôme de degré strictement positif. Pour tout $\alpha \in \mathbb{K}$, on pourra considérer l'évaluation de P en α, c-à-d., $P(\alpha) = \sum_{i=0}^{n} a_i \alpha^i$.

On dit aussi que l'on substitue α à X. Il est clair que, pour tous $P, Q \in \mathbb{K}[X]$ et tout $\alpha \in \mathbb{K}$, on a

$$(P+Q)(\alpha) = P(\alpha) + Q(\alpha), \qquad (PQ)(\alpha) = P(\alpha)Q(\alpha).$$

D'après le théorème 5.1.6, il existe un polynôme Q dans $\mathbb{K}[X]$ et un scalaire $R \in \mathbb{K}$ tels que

$$P(X) = Q(X)(X-\alpha) + R.$$

En spécialisant X à α, on obtient $P(\alpha) = R$. En particulier, on a

Corollaire 5.1.8. *Un polynôme P dans $\mathbb{K}[X]$ est divisible par $X - \alpha$, avec un scalaire $\alpha \in \mathbb{K}$, si et seulement si $P(\alpha) = 0$.*

Définition 5.1.2. *Soit P un polynôme dans $\mathbb{K}[X]$.*

1. *P est dit **scindé** s'il peut s'écrire comme produit de polynômes du premier degré.*

2. *P est dit **irréductible** s'il est ni nul, ni constant, ni produit de deux polynômes non-constants.*

Le théorème suivant est connu sous le nom de **théorème fondamental de l'algèbre** :

Théorème 5.1.9 (D'Alembert-Gauss)**.** *Tout polynôme dans $\mathbb{C}[X]$ est scindé.*

Une démonstration de ce théorème utilise certaine théorie d'analyse complexe. On admettra ce théorème.

5.2 Fraction rationnelle (17/12)

A partir de $\mathbb{Z}$, on peut construire le corps $\mathbb{Q}$ des nombres rationnelles. Ici, on va construire le corps $\mathbb{K}(X)$ à partir de l'anneau $\mathbb{K}[X]$ en imitant la construction de $\mathbb{Q}$.

5.2.1 Définition

On commence par une définition formelle. Considérons $\mathcal{T} := \mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\})$. On définit une relation $\sim$ sur $\mathcal{T}$ comme suit: pour $(P,Q), (P',Q') \in \mathcal{T}$,

$$(P,Q) \sim (P',Q') \qquad \Longleftrightarrow \qquad PQ' = P'Q.$$

C'est une relation d'équivalence sur $\mathcal{T}$ (montrer-le). L'ensemble des classes d'équivalence $\mathcal{T}/\sim$ est noté $\mathbb{K}(X)$. L'élément de $\mathbb{K}(X)$ représenté par $(P,Q) \in \mathcal{T}$ est noté $\dfrac{P}{Q}$, appelé **fraction rationnelle**. Pour $\dfrac{P}{Q}, \dfrac{P'}{Q'} \in \mathbb{K}(X)$, on définit une addition sur $\mathbb{K}(X)$ par

$$\frac{P}{Q} + \frac{P'}{Q'} := \frac{PQ' + P'Q}{QQ'},$$

et une multiplication sur $\mathbb{K}(X)$ par

$$\frac{P}{Q} \cdot \frac{P'}{Q'} := \frac{PP'}{QQ'}.$$

Comme dans le cas du corps $\mathbb{Q}$, on peut montrer

Lemme 5.2.1. *Ces opérations sont bien définies. De plus, pour ces opérations, $\mathbb{K}(X)$ est un corps, appelé **corps des fractions rationnelles**.*

Preuve. Exercice. □

5.2.2 Décomposition en éléments simples

Soit $\dfrac{P}{Q} \in \mathbb{K}(X)$ un élément non nul. Quitte à simplifier la fraction, on suppose que P et Q sont premiers entre eux. Par la division euclidienne (cf. le théorème 5.1.6), il existe un couple de polynômes (R, S) tels que $P = QS + R$ et que $\deg(R) < \deg(Q)$. Alors, on a

$$\frac{P}{Q} = \frac{QS + R}{Q} = S + \frac{R}{Q}.$$

Une question se pose : *est-ce que la fraction $\frac{R}{Q}$ se décompose dans une forme encore plus simple ?*
La réponse est **OUI** ! On va répondre à la question ci-dessus.

Ici, on consèdere la fraction $\dfrac{P}{Q}$. Soit $Q = \prod_{i=1}^{r} Q_i^{m_i}$ une décomposition en produit de polynômes irréductibles, où Q_i et Q_j sont premiers entre eux pour $i \neq j$.

1. D'abord, considérons le cas $r = 1$!

Si $\deg(P) < \deg(Q_1)$, on ne peut plus décomposer. Supposons que $\deg(P) \geq \deg(Q_1)$. Alors, il existe un entier $n \in \mathbb{N}^*$ tel que $n \deg(Q_1) \leq \deg(P) < (n+1) \deg(Q_1)$. Par la division euclidienne, il existe un couple de polynômes (A_n, P_{n-1}) de $\mathbb{K}[X]$ tel que

$$P = A_n Q_1^n + P_{n-1}, \qquad 0 \leq \deg(P_{n-1}) < \deg(Q_1).$$

Par la définition de n, on a $0 \leq \deg(A_n) < \deg(Q_1)$. Par la division euclidienne de P_{n-1} par Q_1^{n-1}, il existe un couple de polynômes (A_{n-1}, P_{n-2}) de $\mathbb{K}[X]$ tel que

$$P_{n-1} = A_{n-1} Q_1^{n-1} + P_{n-2}, \qquad 0 \leq \deg(P_{n-2}) < \deg(Q_1).$$

Vu le degré de P_{n-1}, on a $\deg(A_{n-1}) < \deg(Q_1)$. En répétant cette opération, on déduit qu'il existe des polynômes $A_0, A_1, \cdots, A_n \in \mathbb{K}[X]$ tels que

$$P = A_n Q_1^n + A_{n-1} Q_1^{n-1} + \cdots + A_1 Q_1 + A_0, \qquad \deg(A_i) < \deg(Q_1).$$

Donc, on obtient l'expression

$$\frac{P}{Q} = \frac{A_n}{Q_1^{m_1-n}} + \frac{A_{n-1}}{Q_1^{m_1-(n-1)}} + \cdots + \frac{A_1}{Q_1^{m_1-1}} + \frac{A_0}{Q_1^{m_1}}.$$

Bien sûr, si $m < 0$, on interprète $\frac{1}{Q^m} = Q^{-m}$.

2. Maintenant, on considère le cas $r = 2$.

Comme $Q_1^{m_1}$ et $Q_2^{m_2}$ sont premiers entre eux, d'après l'identité de Bézout, il existe un couple de polynômes (P_1, P_2) de $\mathbb{K}[X]$ tel que

$$P_1 Q_2^{m_2} + P_2 Q_1^{m_1} = P.$$

(**Attention aux indices !**) Alors, on a

$$\frac{P}{Q} = \frac{P_1 Q_2^{m_2} + P_2 Q_1^{m_1}}{Q_1^{m_1} Q_2^{m_2}} = \frac{P_1}{Q_1^{m_1}} + \frac{P_2}{Q_2^{m_2}},$$

et on peut appliquer la méthode expliquée pour le cas $r = 1$. Donc, il existe deux entiers n_1, n_2 positifs et des polynômes $A_{i_1}^{(1)}, A_{i_2}^{(2)}$ $(0 \leq i_k \leq n_k, k = 1, 2)$ tels que

$$\frac{P}{Q} = \sum_{i=0}^{n_1} \frac{A_i^{(1)}}{Q_1^{m_1-i}} + \sum_{i=0}^{n_2} \frac{A_i^{(2)}}{Q_2^{m_2-i}}.$$

3. Finalement, considérons le cas général !

Comme $Q_1^{m_1}$ et $\prod_{i=2}^r Q_i^{m_i}$ sont premiers entre eux, d'après l'identité de Bézout, il existe un couple de polynômes (P_1, P_2') de $\mathbb{K}[X]$ tel que

$$P_1 \prod_{i=2}^{r} Q_i^{m_i} + P_2' Q_1^{m_1} = P.$$

Alors, on a

$$\frac{P}{Q} = \frac{P_1}{Q_1^{m_1}} + \frac{P_2'}{\prod_{i=2}^r Q_i^{m_i}}.$$

Par récurrence de r, on déduit qu'il existe des polynômes P_i $(1 \leq i \leq r)$ de $\mathbb{K}[X]$ tels que

$$\frac{P}{Q} = \sum_{i=1}^{r} \frac{P_i}{Q_i^{m_i}},$$

et on peut appliquer la méthode expliquée pour le cas $r = 1$. Donc, il existe des entiers positifs n_i $(1 \leq i \leq r)$ et des polynômes $A_{i_k}^{(k)}$ $(0 \leq i_k \leq n_k, 1 \leq k \leq r)$ tels que

$$\frac{P}{Q} = \sum_{k=1}^{r} \sum_{i_k=0}^{n_k} \frac{A_{i_k}^{(k)}}{Q_k^{m_k-i_k}}.$$

Cette décomposition est appelée **décomposition en éléments simples**.

Appendix A

Appendices

A.1 Lemme de Zorn

Ici, on énonce ce que l'on appelle le lemme de Zorn et on donne une application.

A.1.1 L'énoncé

Un ensemble ordonné $(E, \leq)$ est dit **inductif** si tout sous-ensemble totalement ordonné possède un majorant.

Lemme A.1.1 (Zorn)**.** *Tout ensemble inductif admet au moins un élément maximal.*

C'est un lemme assez délicat est ici on l'admettra. Par exemple, c'est équivalent à

L'axiome du choix Soit $\{E_\lambda\}_{\lambda\in\Lambda}$ une famille d'ensembles sur un ensemble Λ. S'il existe un $\lambda \in \Lambda$ tel que $E_\lambda = \emptyset$, alors par définition du produit cartésien d'ensembles, on a $\prod_{\lambda\in\Lambda} E_\lambda = \emptyset$. Mais, si $E_\lambda \neq \emptyset$ pour tout $\lambda \in \Lambda$, il n'est pas clair que $\prod_{\lambda\in\Lambda} E_\lambda \neq \emptyset$ pour la raison suivante : il faut $\ll$ choisir d'une infinité d'éléments $x_\lambda \in E_\lambda$ $\gg$ pour dire que $(x_\lambda)_{\lambda\in\Lambda} \in \prod_{\lambda\in\Lambda} E_\lambda$, donc ce dernier n'est pas vide. L'axiome du choix est formulé comme suit :

Soit $\{E_\lambda\}_{\lambda\in\Lambda}$ une famille d'ensembles tels que $E_\lambda \neq \emptyset$ pour tout $\lambda \in \Lambda$.
Alors, $\prod_{\lambda\in\Lambda} E_\lambda \neq \emptyset$.

A.1.2 Existence d'une base

Soit $\mathbb{K}$ un corps commutatif.

Théorème A.1.2. *Soit E un $\mathbb{K}$-espace vectoriel (de dimension quelconque). Alors, E admet une base.*

Preuve. Si $E = \{0\}$, il n'y a rien à montrer, donc on suppose que $E \neq \{0\}$. Posons

$$\mathcal{S} := \{S \subset E \mid S \text{ est une famille libre}\}.$$

Comme $\mathcal{S}$ contient $\{x\}$ pour un $x \in E \setminus \{0\}$, $\mathcal{S}$ n'est pas vide. Regardant $\mathcal{S}$ comme un ensemble ordonné par rapport à l'inclusion, montrons que $(\mathcal{S}, \subset)$ est un ensemble inductif.

Soit $\{S_\alpha\}_{\alpha \in A}$ un sous-ensemble totalement ordonné, c.-à.-d., pour tous $\alpha, \beta \in A$, soit $S_\alpha \subset S_\beta$ ou bien $S_\beta \subset S_\alpha$ est valable. Alors, $S_\infty := \cup_{\alpha \in A} S_\alpha$ est une famille libre. En effet, pour tout sous-ensemble fini $S' \subset S_\infty$, il existe $\alpha \in A$ tel que $S' \subset S_\alpha$. Comme S_α est une famille libre, S' l'est, donc, S_∞ l'est aussi. (Rappelons que l'on considère que les combinaisons linéaires finies.)

Donc, d'après le lemme de Zorn, il existe un élmément maximal $\mathcal{B}$ de $\mathcal{S}$. Montrons que $\mathcal{B}$ est une famille génératrice.

Supposons que $\mathrm{Vect}(\mathcal{B}) \subsetneq E$. Alors, pour un $x \in E \setminus \mathrm{Vect}(\mathcal{B})$, $\mathcal{B} \cup \{x\} \notin \mathcal{S}$ par maximalité de $\mathcal{B}$, c.-à.-d., la famille $\mathcal{B} \cup \{x\}$ ne peut pas être libre. Donc, il existe des scalaires non tous nuls $\alpha, \alpha_1, \cdots, \alpha_n \in \mathbb{K}$ et $x_1, \cdots, x_n \in \mathcal{B}$ tels que

$$\alpha x + \sum_{i=1}^{n} \alpha_i x_i = 0.$$

Si $\alpha = 0$, alors $x_1, \cdots, x_n$ deviennent liés, ce qui est une contradiction. Donc, $\alpha \neq 0$ et comme $\mathbb{K}$ est un corps, on a

$$x = -\alpha^{-1} \left(\sum_{i=1}^{n} \alpha_i x_i\right) \in \mathrm{Vect}(\mathcal{B}).$$

Ceci contredit l'hypothèse que $x \notin \mathrm{Vect}(\mathcal{B})$. On en conclut que $\mathrm{Vect}(\mathcal{B}) = E$, c.-à.d., $\mathcal{B}$ est une famille génératrice de E. Comme $\mathcal{B} \in \mathcal{S}$, elle est libre, c'est donc une base de E. □

A.2 Espace dual

Le but de cette section est d'expliquer l'interprétation de la transposée en terme d'application linéaire. Pour cela, on introduit la notion du dual d'un espace vectoriel. Fixons un corps commutatif $\mathbb{K}$ comme d'habitude.

A.2.1 Définition

Soit E un $\mathbb{K}$-espace vectoriel.

Définition A.2.1. *Le* ***dual*** *de E est le $\mathbb{K}$-espace vectoriel $\mathcal{L}(E, \mathbb{K})$, noté E^*. Un élément $f \in E^*$ est appelé une* ***forme linéaire****.*

Par exemple, supposons que E soit un $\mathbb{K}$-espace vectoriel de dimension n avec $n \in \mathbb{N}^*$ et $\mathcal{B} = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ une base de E. D'après la proposition 3.1.2, un élément $f \in E^*$ est déterminé par les données $\{f(\mathbf{e}_i) | 1 \leq i \leq n\}$. En effet, pour $x = \sum_{i=1}^n x_i \mathbf{e}_i \in E$, on a

$$f(x) = f(\sum_{i=1}^n x_i \mathbf{e}_i) = \sum_{i=1}^n x_i f(\mathbf{e}_i).$$

En particulier, pour $1 \leq i \leq n$, soit $\mathbf{e}_i^* \in E^*$ la forme linéaire telle que $\mathbf{e}_i^*(\mathbf{e}_j) = \delta_{i,j}$. Alors,

Lemme A.2.1. $\mathcal{B}^* = \{\mathbf{e}_1^*, \cdots, \mathbf{e}_n^*\}$ *forme une base de* E^*. *En particulier, on a* $\dim E^* = \dim E$.

$\mathcal{B}^*$ est appelée la **base duale** de $\mathcal{B}$. Pour $x = \sum_{i=1}^n x_i \mathbf{e}_i \in E$, on a

$$\mathbf{e}_i^*(x) = \mathbf{e}_i^*(\sum_{j=1}^n x_j \mathbf{e}_j) = \sum_{j=1}^n x_j \mathbf{e}_i^*(\mathbf{e}_j) = \sum_{j=1}^n x_j \delta_{i,j} = x_i,$$

c.-à-d., cela signifie que

$$x = \sum_{i=1}^n \mathbf{e}_i^*(x) \mathbf{e}_i.$$

Donc, $\mathcal{B}^*$ joue un rôle de coordonnées.

A.2.2 Transposée d'une application linéaire

Maintenant, soient E, F deux $\mathbb{K}$-espaces vectoriels.

Définition A.2.2. *Pour* $u \in \mathcal{L}(E, F)$, *on définit l'application linéaire* ${}^t u \in \mathcal{L}(F^*, E^*)$ *par* ${}^t u(f) := f \circ u$. *L'application* ${}^t u$ *est appelée la* ***transposée de*** u.

Supposons que E, F soient de dimension n et m, respectivement, avec $m, n \in \mathbb{N}^*$. Soient $\mathcal{B} = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ une base de E et $\mathcal{C} := \{\mathbf{f}_1, \cdots, \mathbf{f}_m\}$ une base de F. Posons $A = (a_{i,j}) := \mathrm{mat}_{\mathcal{B},\mathcal{C}}(u) \in M_{m,n}(\mathbb{K})$. Soit $\mathcal{B}^* = \{\mathbf{e}_1^*, \cdots, \mathbf{e}_n^*\}$ la base duale de $\mathcal{B}$ et $\mathcal{C}^* = \{\mathbf{f}_1^*, \cdots, \mathbf{f}_m^*\}$ la base duale de $\mathcal{C}$.

Calculons la matrice $\mathrm{mat}_{\mathcal{C}^*,\mathcal{B}^*}({}^t u)$. Pour $1 \leq j \leq n$, on a $u(\mathbf{e}_j) = \sum_{i=1}^m a_{i,j} \mathbf{f}_i$ par définition, donc on obtient, pour $1 \leq i \leq m$,

$$\begin{aligned} {}^t u(\mathbf{f}_i^*)(\mathbf{e}_j) &= \mathbf{f}_i^* \circ u(\mathbf{e}_j) = \mathbf{f}_i^*(u(\mathbf{e}_j)) = \mathbf{f}_i^* \left(\sum_{k=1}^m a_{k,j} \mathbf{f}_k \right) \\ &= \sum_{k=1}^m a_{k,j} \mathbf{f}_i^*(\mathbf{f}_k) = \sum_{k=1}^m a_{k,j} \delta_{i,k} = a_{i,j}, \end{aligned}$$

i.e., ${}^t u(\mathbf{f}_i^*) = \sum_{j=1}^n a_{i,j} \mathbf{e}_j^*$. On en déduit que

Proposition A.2.2. $\mathrm{mat}_{\mathcal{C}^*,\mathcal{B}^*}({}^t u) = {}^t\mathrm{mat}_{\mathcal{B},\mathcal{C}}(u)$.

Ceci explique le sens de la transposée d'une matrice. Ce lemme a une implication intéressante.
Soient E, F, G des $\mathbb{K}$-espaces vectoriels et $u \in \mathcal{L}(E,F), v \in \mathcal{L}(F,G)$ des applications linéaires. Alors, on a

Proposition A.2.3. ${}^t(v \circ u) = {}^t u \circ {}^t v$.

En effet, pour $f \in G^*$, on a

$$ {}^t(v \circ u)(f) = f \circ (v \circ u) = (f \circ v) \circ u = {}^t u(f \circ v) = {}^t u({}^t v(f)) = ({}^t u \circ {}^t v)(f). $$

Cette proposition explique la nature de la proposition 4.1.8. D'après le lemme 4.2.2, on a :

Proposition A.2.4. $\mathrm{rg}({}^t u) = \mathrm{rg}(u)$.

A.2.3 Injectivité et surjectivité

Soient E, F deux $\mathbb{K}$-espaces vectoriels. Ici, j'explique la dualité entre injectivité et surjectivité, c.-à-d.,

Proposition A.2.5. *Soit $u \in \mathcal{L}(E,F)$ une application linéaire.*

1. *u est injective si et seulement si ${}^t u$ est surjective.*

2. *u est surjective si et seulement si ${}^t u$ est injective.*

Preuve. Montrons 1. D'après la formule du rang, l'injectivité de u implique $\mathrm{rg}(u) = \dim E$. Donc, d'après la proposition A.2.4, on obtient $\mathrm{rg}({}^t u) = \mathrm{rg}(u) = \dim E = \dim E^*$, c.-à-d., ${}^t u$ est surjective. Montrons 2. L'application u étant surjective, on a $\mathrm{rg}(u) = \dim F$. Donc, d'après la proposition A.2.4, on obtient $\mathrm{rg}({}^t u) = \mathrm{rg}(u) = \dim F = \dim F^*$. D'après la formule du rang, ceci dit que ${}^t u$ est injective. □

A.2.4 Bidual

Da'près le lemme A.2.1, on voit que E et E^* sont isomorphes lorsque E est de dimension finie. En effet, en choisissant une base $\mathcal{B} = \{u_1, \cdots, u_n\}$ de E et une base $\mathcal{B}' = \{v_1, \cdots, v_n\}$ de E^*, on peut définir un isomorphisme φ par $\varphi(u_i) = v_i$ $(1 \leq i \leq n)$. Mais, il n'y a pas d'isomorphisme canonique, c.-à-d., on ne peut pas donner un isomorphisme de E vers E^* sans choisir des bases.

En revanche, on peut définir l'isomorphisme Φ de E dans le **bidual** $E^{**} := (E^*)^*$ de E, par

$$ \Phi(x)(f) := f(x). $$

La définition du morphisme Φ ne dépend pas d'un choix de base, Φ est alors appelé l'**isomorphisme canonique**.

Théorème A.2.6. *Le bidual E^{**} de E est canoniquement isomorphe à E.*