Cours de mathématiques
Partie IV – Algèbre 2
MP2I

Alain TROESCH

Version du:

4 juillet 2024

Table des matières

23

Algèbre linéaire

Est-il déraisonnable d'imaginer que cette linéarité ne pourrait être que l'approximation d'une non-linéarité plus précise (mais aussi plus subtile) ?

(Roger Penrose)

Si une entité possède une grandeur et une direction, sa grandeur et sa direction prises ensemble constituent ce qui est appelé un vecteur. La description numérique d'un vecteur nécessite trois nombres, mais rien ne nous empêche d'utiliser une seule lettre pour sa désignation symbolique. Une algèbre ou une méthode analytique dans laquelle une seule lettre ou une autre expression est utilisée pour spécifier un vecteur peut être appelée une algèbre vectorielle, ou une analyse vectorielle.

(Josiah Gibbs)

La structure d'espace vectoriel est une structure rigide, généralisant le cadre géométrique usuel.

Il s'agit d'une structure algébrique liée à la donnée d'un corps, qui va constituer l'unité de la dimension : la droite réelle représente l'espace vectoriel typique sur $\mathbb{R}$ de dimension 1, alors que le $\mathbb{C}$-espace vectoriel typique de dimension 1 ressemblera au plan complexe, donc à un objet géométrique, qui, en tant que $\mathbb{R}$-espace vectoriel sera en fait de dimension 2.

La rigidité de la structure se traduit par le fait qu'on peut multiplier un élément par un scalaire (un élément du corps de base), ceci de façon injective (sauf pour 0) : ainsi, si un élément x est dans un espace vectoriel E, tous les éléments λx, $\lambda \in \mathbb{K}$ seront aussi dans E, et si $x \neq 0$, l'ensemble des λx, $\lambda \in \mathbb{K}$ « ressemble » à $\mathbb{K}$ (il y a une bijection entre les deux). On parle de la droite engendrée par x. Ainsi un espace vectoriel est une structure droite, qui, dès qu'elle contient un élément non nul, contient tout la droite (au sens du corps $\mathbb{K}$) contenant x.

La rigidité d'un espace vectoriel est même plus forte que cela : plus que la stabilité par multiplication par un scalaire, on a la stabilité par combinaison linéaire (et encore une fois, l'application qui à (λ, μ) de $\mathbb{K}^2$ associe $\lambda x + \mu y$ est bijective, sauf si x et y sont colinéaire). Ainsi, si E contient deux points (non colinéaires), il contient tout un $\mathbb{K}$-plan passant par ces deux points et par l'origine.

Cette structure rigide (plate pourrait-on dire) généralise la situation du plan réel usuel (approximation de la surface localement plate de la Terre sur laquelle nous faisons notre géométrie) ou de l'espace usuel, donc de la géométrie euclidienne classique. Elle ne permet en revanche pas de prendre en compte de façon implicite des phénomènes de courbure intrinsèque (géométrie sphérique définie intrinsèquement sur un objet de dimension 2, sans plongement dans un espace de dimension 3, ou propriétés de courbures de l'espace-temps) : ces structures courbes nécessitent l'introduction d'objets plus complexes (les variétés).

Dans tout le chapitre, on considère un corps $\mathbb{K}$.

I Espaces vectoriels

I.1 Définition

Définition 23.1.1 (Espace vectoriel)

Soit E un ensemble. On dit que E est un espace vectoriel sur $\mathbb{K}$ (en abrégé $\mathbb{K}$-ev) si E est muni de deux lois :

- une loi de composition interne $+ \ : \ E \times E \longrightarrow E$;
- une loi de composition externe $\cdot \ : \ \mathbb{K} \times E \longrightarrow E$;

telles que :

(i) $(E, +)$ soit un groupe abélien ;

(ii) pour tout $(\lambda, \mu) \in \mathbb{K}^2$, $(x, y) \in E^2$:
- $(\lambda\mu)x = \lambda(\mu x)$ (associativité externe de $\cdot$, ou pseudo-associativité) ;
- $1_{\mathbb{K}} \cdot x = x$ (compatibilité du neutre multiplicatif de $\mathbb{K}$) ;
- $\lambda(x_1 + x_2) = \lambda x_1 + \lambda x_2$ (distributivité de $\cdot$ sur la loi interne) ;
- $(\lambda + \mu)x = \lambda x + \mu x$ (distributivité de $\cdot$ sur la somme de $\mathbb{K}$).

Propriétés 23.1.2

Soit E un $\mathbb{K}$-ev. Pour tout $x \in E$:

1. $0 \cdot x = 0$, *c'est-à-dire* $0_{\mathbb{K}} \cdot x = 0_E$;
2. $\lambda \cdot 0_E = 0_E$;
3. $(-1) \cdot x = -x$.
4. *Si* $x \neq 0$, $\lambda \cdot x = 0 \Longrightarrow \lambda = 0$
5. *Si* $\lambda \neq 0$, $\lambda \cdot x = 0 \Longrightarrow x = 0$.

◁ **Éléments de preuve.**

1. Considérer $(0+0)x$, et utiliser la régularité additive (c'est un groupe aditif !
2. De même, considérer $\lambda(0+0)$
3. Développer $(1-1)x$
4. Si $x \neq 0$ et $\lambda \neq 0$, considérer $\lambda^{-1}(\lambda x) \neq 0$, donc $\lambda x \neq 0$.
5. C'est la même propriété que la précédente.

▷

Terminologie 23.1.3

- Les éléments de E sont appelés *vecteurs*.
- Les éléments de $\mathbb{K}$ sont appelés *scalaires*.
- Deux éléments x et y de E sont colinéaires s'il existe $(\lambda, \mu) \in \mathbb{K}^2$ tels que $(\lambda, \mu) \neq (0,0)$ et $\lambda x + \mu y = 0$.

I.2 Combinaisons linéaires

Une propriété cruciale d'un espace vectoriel E est la stabilité par combinaison linéaire : si $(\lambda, \mu) \in \mathbb{K}^2$ et $(x, y) \in \mathrm{E}^2$, alors $\lambda x + \mu y \in E$. La notion de combinaison linéaire étant centrale dans l'étude des espaces vectoriels, nous définissons une notion généralisée de combinaison linéaire.

Définition 23.1.4 (Famille à support fini)
Soit I un ensemble d'indices, et $(\lambda_i)_{i\in I}$ une famille d'éléments de $\mathbb{K}$. On dit que la famille $(\lambda_i)_{i\in I}$ est à support fini s'il existe $J \subset I$ un sous-ensemble fini de I tel que pour tout $i \in I \smallsetminus J$, $\lambda_i = 0$. Autrement dit, seul un nombre fini de λ_i sont non nuls.

Si I lui même est finie, toute famille indexée sur I est à support fini. Ainsi, cette notion est surtout pertinente lorsque I est infini.

Terminologie 23.1.5 (Famille de scalaire presque tous nuls)
Lorsque I est infini (et même parfois lorsque I est fini, mais dans ce cas, la terminologie n'est pas heureuse), une famille à support fini est parfois aussi appelé famille de scalaires presque tous nuls.

Définition 23.1.6 (Combinaison linéaire généralisée)
Soit E un $\mathbb{K}$-ev et $(x_i)_{i\in I}$ une famille de vecteurs de E. Une combinaison linéaire des $(x_i)_{i\in I}$ est un vecteur

$$x = \sum_{i\in I} \lambda_i x_i,$$

où $(\lambda_i)_{i\in I}$ est une famille à support fini de scalaires de $\mathbb{K}$.

Ainsi, toute combinaison linéaire sur une famille infinie est une combinaison linéaire d'un nombre fini de vecteurs de cette famille.

I.3 Un exemple important : espace de fonctions

Comme nous avons pu nous en rendre compte pour les groupes et les anneaux, on a des critères souvent rapides pour montrer qu'un ensemble est un sous-truc d'un truc plus gros. Nous verrons un peu plus loin que de la même façon, il est beaucoup plus commode de montrer qu'un ensemble est un sous-espace vectoriel d'un espace vectoriel connu plutôt que de montrer de façon directe qu'il s'agit d'un espace vectoriel, ce qui nécessite beaucoup de petites vérificiations, élémentaires mais fastidieuses. Pour cette raison, il est important de connaître un certain nombre d'espaces vectoriels de référence, qui seront suffisants pour justifier la structure d'espace vectoriel d'autres ensembles dans la plupart des cas rencontrés.

Proposition 23.1.7 (espace vectoriel de référence)

1. *$\mathbb{K}$ est un espace vectoriel sur lui-même.*
2. *Soit E un espace vectoriel sur $\mathbb{K}$ et F un ensemble quelconque. Alors l'ensemble des fonctions E^F est un espace vectoriel sur $\mathbb{K}$.*

◁ **Éléments de preuve.**

1. Immédiat d'après les propriétés des lois d'un corps
2. Vérifications longues et fastidieuses, mais sans difficulté. Bien avoir compris que les lois sur E^F sont obtenues de celles de E après évaluation point par point. Ainsi, $(f+g)(x) = f(x)+g(x)$ et $(\lambda f)(x) = \lambda f(x)$.

▷

Exemples 23.1.8 (Espaces vectoriels à bien connaître)

1. $\mathbb{K}^{\varnothing} = \{0\}$;

2. $\mathbb{K}^{[\![1,n]\!]} \simeq \mathbb{K}^n$; en particulier $\mathbb{C}$ est un $\mathbb{R}$-ev ;
3. $\mathbb{K}^{\mathbb{N}}$ l'ensemble des suites à valeurs dans $\mathbb{K}$;
4. $\mathbb{K}^{[\![0,n]\!]} \simeq \mathbb{K}_n[X]$ l'espace des polynômes de degré au plus n ;

I.4 Produits d'espaces vectoriels

Nous voyons maintenant deux façons de construire des espaces vectoriels à partir d'espaces vectoriels de référence : tout d'abord une construction externe (produit cartésien), puis dans la section suivante, une construction interne (sous-espaces vectoriels).

Proposition 23.1.9 (produit cartésien d'ev)

Soit $E_1, \cdots, E_n$ des espaces vectoriels sur un corps $\mathbb{K}$. Alors le produit cartésien $E_1 \times \cdots \times E_n$ est un espace vectoriel sur $\mathbb{K}$, lorsqu'on le munit des lois définies par :

- $\forall \lambda \in \mathbb{K},\ \forall (x_1, \cdots, x_n) \in E_1 \times \cdots \times E_n, \quad \lambda \cdot (x_1, \cdots, x_n) = (\lambda x_1, \cdots, \lambda x_n)$;
- $\forall (x_1, \cdots, x_n), (y_1, \cdots, y_n) \in E_1 \times \cdots \times E_n, \quad (x_1, \cdots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n)$.

◁ **Éléments de preuve.**
Vérifications simples ▷

On retrouve en particulier la structure d'espace vectoriel de $\mathbb{K}^n$, déjà obtenue en considérant $\mathbb{K}^{[\![1,n]\!]}$.

I.5 Sous-espaces vectoriels

Selon les définitions générales sur les structures, on définit :

Définition 23.1.10 (Sous-espace vectoriel, sev)

Soit E un espace vectoriel sur $\mathbb{K}$. Un sous-ensemble $F \subset E$ de E est un sous-espace vectoriel de E si F est stable par les lois $+$ et $\cdot$ et que les lois induites munissent F d'une structure d'espace vectoriel.

Comme dans le cas des groupes et des anneaux, nous disposons d'un critère simple permettant de court-circuiter un certain nombre de vérifications :

Théorème 23.1.11 (Caractérisation des sous-espace vectoriel)

Soit E un $\mathbb{K}$-espace vectoriel. Un ensemble F est un sous-espace vectoriel de E si et seulement si :

(i) $F \subset E$

(ii) $0 \in F$

(iii) *F est stable par combinaison linéaire, ce qui équivaut à :*

$$\forall x, y \in F^2,\ \forall \lambda \in \mathbb{K},\ \lambda x + y \in F.$$

◁ **Éléments de preuve.**
La stabilité permet de définir des lois sur F. Toutes les propriétés universelles sont préservées. ▷

Exemples 23.1.12

1. Étant donné un espace vectoriel E, le sous-espace vectoriel nul $\{0_E\}$ et le sous-espace vectoriel total E.

2. Étant donné un vecteur X de $\mathbb{K}^n$, la droite $\mathbb{K}X = \{\lambda X,\ \lambda \in \mathbb{K}\}$
3. Étant donné a, b et c trois réels, le plan de $\mathbb{R}^3$ d'équation $ax + by + cz = 0$.
4. $\mathbb{K}[X]$ espace des polynômes à coefficients dans le corps $\mathbb{K}$;
5. $\mathcal{C}(\mathbb{R}, \mathbb{R})$ ensemble des fonctions continues sur $\mathbb{R}$;
6. Plus généralement $\mathcal{C}(I, \mathbb{R})$, ensemble des fonctions continues sur un intervalle I ;
7. $\mathcal{C}^n(\mathbb{R}, \mathbb{R})$ ensemble des fonctions de classe $\mathcal{C}^n$ sur $\mathbb{R}$;
8. Plus généralement $\mathcal{C}^n(I, \mathbb{R})$, ensemble des fonctions de classe $\mathcal{C}^n$ sur un intervalle I ;
9. Les exemples d'espaces vectoriels de fonctions sont nombreux.

Vous remarquerez dans les premiers exemples les deux points de vue différents pour définir un sous-espace vectoriel : par l'intérieur (sous-espace engendré par un vecteur) ou par l'extérieur (sous-espace défini par une équation sur les coordonnées). On retrouvera souvent ces deux points de vue par la suite. L'exemple 2 ayant une importance particulière, on l'isole dans la définition et la propriété suivantes :

Définition 23.1.13 (Droite vectorielle)

Soit E un $\mathbb{K}$-espace vectoriel. Une droite vectorielle de E est un sous-ensemble D de E tel qu'il existe $x \in E$ non nul tel que $D = \mathbb{K}x$.

Proposition 23.1.14 (Stucture des droites vectorielles)

Les droites vectorielles d'un espace vectoriel E sur $\mathbb{K}$ sont des sous-espaces vectoriels de E.

Proposition 23.1.15

Soit E un $\mathbb{K}$-espace vectoriel, et D_1 et D_2 deux droites vectorielles. Alors soit $D_1 \cap D_2 = \{0_E\}$, soit $D_1 = D_2$.

◁ **Éléments de preuve.**

Écrire $D_1 = \mathbb{K}x_1$ et $D_2 = \mathbb{K}x_2$. Si $x \neq 0$ est dans $D_1 \cap D_2$, écrire $x = \lambda x_1$, $x = \mu x_2$ puis écrire x_2 on fonction de x_1 pour en déduire $D_2 \subset D_1$ (et inversement). ▷

Corollaire 23.1.16

Soit D une droite vectorielle d'un espace vectoriel E, et $x \in D$. Si $x \neq 0$, alors $D = \mathbb{K}x$.

◁ **Éléments de preuve.**

$D \cap \mathbb{K}x \neq \{0\}$. ▷

Exemple 23.1.17 (Sous-espaces vectoriels de $\mathbb{R}^2$)

Les sous-espaces vectoriels de $\mathbb{R}^2$ sont :

- le sous-espace vectoriel nul ;
- les droites vectorielles ;
- le sous-espace vectoriel total $\mathbb{R}^2$.

Remarque 23.1.18

L'aspect géométrique d'une droite vectorielle dépend du corps de base $\mathbb{K}$:

- Si le corps de base est $\mathbb{R}$, une droite vectorielle a l'aspect d'une droite géométrique usuelle.
- Si $\mathbb{K} = \mathbb{C}$, une droite a l'aspect d'un plan complexe : une droite complexe est donc un objet de dimension géométrique égale à 2.
- Si $\mathbb{K} = \mathbb{F}_p$, alors une droite est constituée d'un nombre fini de points « alignés circulairement » si on peut dire cela ainsi...
- Par exemple, si $\mathbb{K} = \mathbb{F}_2$, une droite est constituée de deux points : il y a dans ce cas autant de droites que de vecteurs non nuls de E, les droites étant les ensembles $\{0, x\}$, $x \neq 0$.

I.6 Intersections et unions de sev

Proposition 23.1.19 (Intersection de sev)

Soit E un $\mathbb{K}$-espace vectoriel, et $(E_i)_{i\in I}$ une famille de sous-espaces vectoriels de E. Alors $\bigcap_{i\in I} E_i$ est un sev de E.

◁ **Éléments de preuve.**

Utiliser la caractérisation des sev. ▷

Remarque 23.1.20

Que dire de l'union de deux sev ?

Proposition 23.1.21 (Union d'une chaîne de sev, HP)

Soit $I \neq \varnothing$, et $(E_i)_{i\in I}$ une chaîne de sous-espaces vectoriels de E, donc vérifiant, pour tout couple $(i, j) \in I^2$, $E_i \subset E_j$ ou $E_j \subset E_i$. Alors $\bigcup_{i\in I} E_i$ est un sous-espace vectoriel de E.

◁ **Éléments de preuve.**

Pour justifier la stabilité par combinaison linéaire, justifier, pour tous x et y de l'union, l'existence d'un indice commun i tel que $x \in E_i$ et $y \in E_i$. ▷

I.7 Sous-espace vectoriel engendré par un sous-ensemble

Définition 23.1.22 (Sous-espace vectoriel engendré par un sous-ensemble)

Soit E un $\mathbb{K}$-ev, et X un sous-ensemble de E. Le sous-espace vectoriel engendré par X est le plus petit sous-espace vectoriel de E contenant X. Il est noté $\mathrm{Vect}(X)$.

Remarque 23.1.23

Pourquoi un tel espace existe-t-il ?

Si X est exprimée sous forme d'une famille $(x_i)_{i\in I}$, on note $\mathrm{Vect}(X) = \mathrm{Vect}((x_i)_{i\in I})$. Si X est fini et enuméré par exemple $x = \{x_1, \dots, x_n\}$, on notera $\mathrm{Vect}(X) = \mathrm{Vect}(x_1, \dots, x_n)$.

Proposition 23.1.24 (Minimalité de Vect(X))
Par définition, tout sous-espace vectoriel de E contenant X contient aussi Vect(X).

On peut donner une description explicite de Vect(X) à l'aide de combinaisons linéaires :

Proposition 23.1.25 (Description de Vect(X))
Soit E un $\mathbb{K}$-ev et X un sous-ensemble de E. Alors Vect(X) *est l'ensemble des combinaisons linéaires d'éléments de X.*

◁ **Éléments de preuve.**
Ces CL appartiennent nécessairement à Vect(X), et forment un sous-espace vectoriel de E. ▷

Ainsi, $x \in \mathrm{Vect}(X)$ si et seulement s'il existe $(x_1, \dots x_n) \in X^n$ et $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ tels que

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

Exemples 23.1.26

1. $\mathrm{Vect}(x) = \{\lambda x, \quad \lambda \in \mathbb{K}\} = \mathbb{K}x$
2. $\mathrm{Vect}(x, y) = \{\lambda x + \mu y, \quad (\lambda, \mu) \in \mathbb{K}^2\}$
   - si $x = y = 0$, $\mathrm{Vect}(x, y) = \{0\}$
   - si x et y sont colinéaires, non tous deux nuls (disons $x \neq 0$), $\mathrm{Vect}(x, y) = \mathrm{Vect}(x)$
   - si x et y ne sont pas colinéaires, $\mathrm{Vect}(x, y)$ est un plan vectoriel.

I.8 Sommes de sev

Définition 23.1.27 (Somme de deux sev)
Soit G un $\mathbb{K}$-ev, et E et F deux sev de G. Alors la somme $E + F$ de E et F est le plus petit sous-espace vectoriel de G, contenant à la fois E et F :

$$E + F = \mathrm{Vect}(E \cup F).$$

Par définition de Vect, $E + F$ est donc la borne supérieure de $\{E, F\}$ dans l'ensemble des sous-espaces vectoriels de G muni de l'ordre d'inclusion.

Proposition 23.1.28 (Description d'une somme)
Soit G un $\mathbb{K}$-ev, et E et F deux sev de G. Alors :

$$E + F = \{x + y \mid x \in E, y \in F\} = \{z \in G \mid \exists (x, y) \in E \times F, \ z = x + y\}.$$

◁ **Éléments de preuve.**
Ces sommes sont obligatoirement dans Vect($E \cup F$) et forment un sev. ▷

Proposition 23.1.29
Soit X et Y deux sous-ensembles de G. Alors $\mathrm{Vect}(X \cup Y) = \mathrm{Vect}(X) + \mathrm{Vect}(Y)$.

◁ **Éléments de preuve.**

- $\mathrm{Vect}(X) + \mathrm{Vect}(Y)$ est un sev contenant $X \cup Y$
- $\mathrm{Vect}(X \cup Y)$ contient $\mathrm{Vect}(X)$ et $\mathrm{Vect}(Y)$, donc $\mathrm{Vect}(X) \cup \mathrm{Vect}(Y)$ donc l'espace engendré.

▷

Définition 23.1.30 (Somme d'un nombre fini de sous-espaces)

Plus généralement, on définit la somme des sous-espaces $E_1, \cdots, E_n$ par :

$$E_1 + \cdots + E_n = \mathrm{Vect}(E_1 \cup \cdots \cup E_n)$$

Proposition 23.1.31 (Description d'une somme d'un nombre fini de sev)

Soit $E_1, \dots, E_n$ et F des sev de E. Les propriétés suivantes sont équivalentes :

(i) $F = E_1 + \cdots + E_n$

(ii) $F = (((E_1 + E_2) + E_3) + \cdots + E_{n-1}) + E_n$

(iii) $F = \{x_1 + \cdots + x_n \mid (x_1, \dots, x_n) \in E_1 \times \cdots \times E_n\}$.

◁ **Éléments de preuve.**

Par récurrence en utilisant la propriété précédente. ▷

I.9 Sommes directes

Définition 23.1.32 (Somme directe)

Soit E et F deux sev de G. On dit que la somme $E + F$ est *directe*, et on note $E \oplus F$, si $E \cap F = \{0\}$.

Plus généralement, $E_1 + \cdots + E_n$ est directe si $E_1 \oplus E_2$, puis $(E_1 + E_2) \oplus E_3$, etc. On note $\bigoplus_{i=1}^{n} E_i$.

Ainsi, la somme $E_1 + \cdots + E_n$ est directe si et seulement si pour tout $k \in [\![2, n]\!]$,

$$E_k \cap \sum_{i<k} E_i = \{0\}.$$

Proposition 23.1.33 (Caractérisation de $\oplus$ par les intersections)

La somme $E_1 + \cdots + E_n$ est directe si et seulement si pour tout $k \in [\![1, n]\!]$,

$$E_k \cap \sum_{i \neq k} E_i = \{0\}.$$

◁ **Éléments de preuve.**

Le sens réciproque est immédiat. Pour le sens direct, raisonner par contraposée en considérant un élément non nul x dans l'intersection, en déduire une relation entre des x_i éléments des différents E_i, et isoler celui de plus grand indice pour obtenir une contradiction. ▷

Cette proposition assure notamment une totale symétrie des rôles de chacun des espaces dans la propriété de somme directe. Ainsi, la somme directe est une notion associative et commutative (la somme étant commutative). On peut donc utiliser la commutativité généralisée et permuter les termes à notre guise sans modifier le caractère direct de la somme.

Proposition 23.1.34 (Caractérisation de $\oplus$ par l'unicité)

La somme $\sum E_i$ de sous-espaces vectoriels de E est directe si et seulement si l'application ci-dessous est injective :

$$\begin{array}{rccl} \varphi : & E_1 \times \cdots \times E_n & \longrightarrow & E \\ & (x_1, \cdots, x_n) & \longmapsto & x_1 + \cdots + x_n, \end{array}$$

ce qui équivaut à dire que le vecteur 0_E admet une unique décomposition (la décomposition nulle) comme somme d'éléments des E_i.

◁ **Éléments de preuve.**

Pour simplifier, remarquez que φ est un morphisme de groupes additifs (avec la loi de groupe produit). On peut se contenter d'étudier le noyau. En considérant un élément non nul du noyau, isoler dans la somme n'importe quel terme non nul de la décomposition et contredire la proposition précédente. On peut aussi revenir à la définition en choisissant bien le terme qu'on isole. Lequel ? ▷

En d'autres termes, $E_1 \oplus \cdots \oplus E_n$ est directe si et seulement tout élément x de $E_1 \oplus \cdots \oplus E_n$ s'écrit de façon unique sous la forme $x = x_1 + \cdots + x_n$, $x_i \in E_i$.
Cette propriété également fournit la parfaite symétrie de la notion de somme directe.

Proposition 23.1.35

Soit $E_1, \dots, E_n, F_1, \dots F_n$ des sev de E tels que pour tout i, $F_i \subset E_i$. Alors, si $E_1 \oplus \cdots \oplus E_n$ est directe, il en est de même de $F_1 \oplus \cdots \oplus F_n$.

◁ **Éléments de preuve.**

Les hypothèses fournissent une inclusion sur les intersections. ▷

Définition 23.1.36 (Supplémentaire)

Soit E un espace vectoriel, et F et G deux sev de E. On dit que F et G sont supplémentaires dans E si $F \oplus G = E$.

Théorème 23.1.37 (Existence d'un supplémentaire, avec AC, HP sauf en dimension finie)

Soit E un espace vectoriel quelconque, et F un sev de E. Alors F admet au moins un supplémentaire G.

◁ **Éléments de preuve.**

La démonstration de ce théorème repose sur le lemme de Zorn, donc sur l'axiome du choix. Considérer l'ensemble $\mathcal{H} = \{H \text{ sev de } E \mid F \cap H = \{0\}\}$, ordonné par inclusion. Montrer qu'il est inductif. Si G est un élément maximal de $\mathcal{H}$, et que $F \oplus G \neq E$, obtenir une contradiction en trouvant $x \notin G$ tel que $G \oplus \mathbb{K}x \in \mathcal{H}$. ▷

On en verra une démonstration plus élémentaire, indépendante de l'axiome du choix, lorsque E est de dimension finie.

II Familles de vecteurs

Nous rappelons que toute combinaison linéaire d'une famille (finie ou infinie) s'exprime, par définition, comme somme *finie* d'éléments de cette famille, multipliés par des scalaires.

II.1 Familles libres

Proposition/Définition 23.2.1 (Famille libre)

Une famille $(x_i)_{i\in I}$ de vecteurs de E est *libre* si une des propriétés équivalentes suivantes est vérifiée :

(i) Pour toute famille $(\lambda_i)_{i\in I}$ de scalaires de $\mathbb{K}$, à support fini : $\sum_{i\in I} \lambda_i x_i = 0 \Longrightarrow \forall i \in I, \lambda_i = 0$;

(ii) Pour tout $x \in \mathrm{Vect}((x_i)_{i\in I})$, il existe une *unique* famille $(\lambda_i)_{i\in I}$ de scalaires de $\mathbb{K}$, à support fini tels que $x = \sum_{i\in I} \lambda_i x_i$.

◁ **Éléments de preuve.**

Dans le sens direct, faire la différence entre deux décompositions d'un même x. Dans le sens réciproque, que dire de l'unique décomposition de 0 ?

On peut aussi caractériser l'injectivité de l'application de $\mathbb{K}^n$ dans E qui aux coefficients associe la CL correspondante. ▷

Définition 23.2.2 (Famille liée)

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

Remarque 23.2.3

Une famille contenant 0 peut-elle être libre ?

Proposition 23.2.4 (Stabilité de la liberté par restriction)

Toute sous-famille d'une famille libre est libre.

La proposition suivante permet de ramener l'étude de la liberté des familles infinies à l'étude de la liberté de familles finies.

Proposition 23.2.5 (Caractérisation de la liberté pour des familles infinies)

Une famille $(x_i)_{i\in I}$ est libre si et seulement si toutes ses sous-familles finies sont libres. Une famille $(x_i)_{i\in \mathbb{N}}$ est libre si et seulement si pour tout $n \in \mathbb{N}$, la famille $(x_0, \cdots, x_n)$ est libre.

◁ **Éléments de preuve.**

- Le premier point provient du fait que par définition, les CL sont définies pour des familles de scalaires à support fini.
- Le second point provient du fait que tout sous-ensemble fini de $\mathbb{N}$ est inclus dans un $[\![0, n]\!]$; comment définir n ?

▷

Proposition 23.2.6 (Ajout d'un élément à une famille libre)

Soit $(x_i)_{i\in I}$ une famille libre de E et x_j $(j \notin I)$ un élément de E. Alors, la famille $(x_i)_{i\in I\cup\{j\}}$, obtenue en ajoutant x_j à la famille libre $(x_i)_{i\in I}$, est libre si et seulement si $x_j \notin \mathrm{Vect}((x_i)_{i\in I})$

◁ **Éléments de preuve.**

- Sens direct par contraposée, l'appartenance au Vect amenant l'existence d'une relation non triviale.

- Sens réciproque par contraposée aussi, une relation faisant nécessairement intervenir le terme x_j (pourquoi ?)

▷

Si un tel ajout est impossible, on dira que la famille libre est maximale :

Définition 23.2.7 (Famille libre maximale)

Une famille libre est maximale si et seulement s'il est impossible de lui rajouter un vecteur (quelconque) de E en préservant sa liberté.

Proposition 23.2.8

Une famille libre est maximale si et seulement si elle est génératrice dans le sens de la définition 23.2.12 (i.e. tout vecteur de E est CL de vecteurs de la famille)

◁ **Éléments de preuve.**

Sinon, on peut trouver $x_j \notin \mathrm{Vect}((x_i)_{i\in I})$. ▷

Proposition 23.2.9 (Caractérisation des sommes directes par la liberté)

Soit $E_1, \cdots, E_n$ des sev non triviaux de E. Alors la somme $E_1 \oplus \cdots \oplus E_n$ est directe si et seulement si tout n-uplet $(x_1, \cdots, x_n)$ d'éléments tous non nuls de $E_1 \times \cdots \times E_n$ est une famille libre dans E.

◁ **Éléments de preuve.**

Une relation entre les x_i donne une décomposition de 0 dans la somme directe. ▷

Corollaire 23.2.10 (Une caractérisation des familles libres finies, HP)

Soit $(x_1, \ldots, x_n)$ une famille de vecteurs non nuls de E. Les propriétés suivantes sont équivalentes :

(i) $(x_1, \ldots, x_n)$ *est une famille libre*

(ii) $\mathbb{K}x_1 \oplus \cdots \oplus \mathbb{K}x_n$ *est directe*

(iii) $\psi : (\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n \mapsto \lambda_1 x_1 + \cdots + \lambda_n x_n$ *est injective.*

◁ **Éléments de preuve.**

L'équivalence entre (i) et (ii) est une conséquence immédiate de la propriété précédente. L'équivalence entre (ii) et (iii) provient de la caractérisation de la somme directe par l'unicité de la décomposition. ▷

Proposition 23.2.11 (Liberté et sommes directes)

1. *Soit F et G deux sev de E tels que $F \oplus G$ soit directe. Alors la concaténation d'une famille libre de F et d'une famille libre de G est une famille libre de E.*
2. *Réciproquement, si $(b_1, \ldots, b_n)$ est une famille libre de E, alors $\mathrm{Vect}(b_1, \ldots, b_k) \oplus \mathrm{Vect}(b_{k+1}, \ldots, b_n)$ est directe.*

◁ **Éléments de preuve.**

1. Considérer une relation globale, utiliser d'abord la somme directe, puis la liberté de chacune des deux sous-familles.
2. Considérer une décomposition dans chacun des deux espaces d'une élément de l'intersection. Comment utiliser l'hypothèse de liberté ?

▷

II.2 Familles génératrices

Définition 23.2.12 (Familles génératrices)

Une famille $(x_i)_{i\in I}$ de vecteurs de E est une famille *génératrice* de E si l'une des propriétés équivalentes suivantes est satisfaite :

(i) tout $x \in E$ est une combinaison linéaire des x_i, $i \in I$;

(ii) $\mathrm{Vect}((x_i)_{i\in I}) = E$.

L'équivalence provient directement de la description explicite de l'espace vectoriel engendré par une famille. Pour une famille finie, on peut caractériser le caractère générateur à l'aide des sommes, ou à l'aide de la fonction ψ, comme pour la liberté ; les justifications des équivalences sont ici immédiates.

Proposition 23.2.13 (Caractérisation du caractère générateur d'une famille finie)

Soit $(x_1, \dots, x_n)$ une famille de vecteurs de E. Les propriétés suivantes sont équivalentes :

- *$(x_1, \dots, x_n)$ est une famille génératrice ;*
- *$E = \sum\limits_{i=1}^{n} \mathbb{K}x_i$;*
- *La fonction $\psi : \mathbb{K}^n \longrightarrow E$ définie par $\psi(\lambda_1, \dots, \lambda_n) = \lambda_1 x_1 + \cdots + \lambda_n x_n$ est surjective.*

Lorsque $(x_i)_{i\in I}$ est génératrice, on dit aussi que la famille $(x_i)_{i\in I}$ engendre E.
On obtient des propriétés duales de celles des familles libres :

Proposition 23.2.14 (Stabilité des familles génératrices par ajout)

Toute famille contenant une famille génératrice de E est une famille génératrice de E.

Proposition 23.2.15 (Restriction d'une famille génératrice)

La famille obtenue en retirant un élément x d'une famille génératrice de E est encore génératrice si et seulement si x est une combinaison linéaire des autres vecteurs de la famille.

◁ **Éléments de preuve.**
Sens direct immédiat. Sens réciproque : remplacer dans toutes les CL le terme x par son expression en fonction des autres vecteurs de la famille. Plus formellement, on peut utiliser la description de l'espace vectoriel d'une union, ainsi que la croissance de Vect pour se débarrasser du $\mathrm{Vect}(x)$ dans la somme obtenue. ▷

Lorsque cette situation n'est vérifiée pour aucun élément de la famille, on parle de famille génératrice minimale :

Définition 23.2.16 (Famille génératrice minimale)

Une famille génératrice est dite minimale, si et seulement s'il est impossible de lui retirer un élément en préservant son caractère générateur.

Proposition 23.2.17

Une famille génératrice est minimale si et seulement si elle est libre.

◁ **Éléments de preuve.**
Une relation non triviale permettrait de trouver un vecteur de la famille s'écrivant en fonction des autres. ▷

II.3 Bases

Définition 23.2.18 (Base d'un espace vectoriel)

Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un ev E. On dit que $(x_i)_{i\in I}$ est une *base* de E si elle est une famille à la fois libre et génératrice de E.

Ainsi :

Proposition/Définition 23.2.19 (Coordonnées d'un vecteur dans une base)

La famille $(b_i)_{i\in I}$ est une base de E si et seulement si tout élément x de E s'écrit de façon unique comme combinaison linéaire des éléments b_i :

$$x = \sum_{i\in I} \lambda_i b_i.$$

L'existence traduit le caractère générateur, l'unicité traduit la liberté. Les coefficients λ_i de cette combinaison linéaire sont appelés *coordonnées de x dans la base $(b_i)_{i\in I}$*.

Le choix d'une base de E permet donc de définir un système de coordonnées : la donnée d'un vecteur x équivaut à la donnée de ses coordonnées dans une base fixée.

Exemple 23.2.20

- Les coordonnées cartésiennes dans $\mathbb{R}^2$ correspondent aux coordonnées dans la base canonique $((1,0),(0,1))$.
- Un autre choix de base fournit d'autres coordonnées, par exemple $(2,3)$ dans la base $((1,0),(1,1))$

Notation 23.2.21 (Pour passer des coordonnées au vecteur)

On notera $\vec{v}_{\mathcal{B}}(X)$ l'unique vecteur de E donc les coordonnées dans la base $\mathcal{B}$ sont données par X, donc

$$\vec{v}_{\mathcal{B}}(X) = \sum x_i b_i,$$

les x_i étant les composantes de X.

Proposition 23.2.22 (Caractérisation des bases par minimalité/maximalité)

Les propriétés suivantes sont équivalentes :

(i) La famille $(x_i)_{i\in I}$ est une base de E ;

(ii) La famille $(x_i)_{i\in I}$ est une famille génératrice minimale de E ;

(iii) La famille $(x_i)_{i\in I}$ est une famille libre maximale de E.

◁ **Éléments de preuve.**

$(ii) \implies (i)$ et $(iii) \implies (i)$ découlent de ce qu'on a déjà fait. De plus, une famille génératrice non minimale ne peut pas être libre, une famille libre non maximale ne peut pas être génératrice (pourquoi ?) ▷

Exemples 23.2.23 (Exemples importants de bases, à connaître)

- Base canonique de $\mathbb{K}^n$.

- Base canonique $(1, X, X^2, \cdots)$ de $\mathbb{K}[X]$; base canonique de $\mathbb{K}_n[X]$.
- $((X - x_0)^k)_{k \in [\![0,n]\!]}$ est une base de $\mathbb{K}_n[X]$.

Ce dernier exemple a une généralisation importante, qui mérite d'être citée en tant que proposition :

Proposition 23.2.24 (Famille totalement échelonnée en degrés)

Si $(P_0, \dots, P_n)$ est une famille d'éléments de $\mathbb{K}_n[X]$ telle que pour tout $k \in [\![0,n]\!]$, $\deg(P_k) = k$, alors $(P_0, \dots, P_n)$ est une base de $\mathbb{K}_n[X]$.

◁ **Éléments de preuve.**

Nous verrons plus tard que cette proposition ne fait que traduire l'inversibilité des matrices triangulaires supérieures à termes diagonaux tous non nuls. En attendant ce point de vue, on peut procéder par récurrence sur n pour montrer que tout polynôme de $\mathbb{K}_n[X]$ a une décomposition unique, en se ramenant à l'hypothèse de récurrence par division euclidienne par P_n. ▷

Avertissement 23.2.25

La réciproque est fausse !

III Applications linéaires

Comme à chaque fois qu'on définit une nouvelle structure algébrique, il vient une notion de morphisme, adaptée à cette structure. La catégorie des espaces vectoriels sur un corps $\mathbb{K}$ est ainsi formée d'un ensemble d'objets (les $\mathbb{K}$-ev), et de flèches représentant les morphismes entre espaces vectoriels, c'est-à-dire, selon les définitions générales qu'on en a données, les applications respectant les deux lois définissant un espace vectoriel. La propriété définissant ces morphismes se traduit par une propriété de linéarité $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$. Pour cette raison, les morphismes d'espaces vectoriels sont plus fréquemment appelés *applications linéaires*.

Note Historique 23.3.1

Les phénomènes à caractère linéaire sont fréquents en physique, au moins à première approximation. La quasi-linéarité de certains phénomènes dans les conditions obervables font que de nombreuses théories physiques ont été considérées comme linéaires pendant longtemps, comme la théorie de la gravitation. ce n'est que plus tard qu'Einstein suggéra que cette linéarité apparente n'est qu'une approximation d'un phénomène non linéaire, approximation néanmoins bien suffisante pour traiter de façon raisonnable les situations de la vie courante. Ainsi, même si de nombreux phénomènes ne sont en réalité pas linéaires, sous certaines conditions, et à première approximation, on peut estimer qu'ils le sont presque. Les étudier sous cet angle permet déjà de comprendre grossièrement le phénomène, par des techniques balisées d'algèbre linéaire. C'est le principe de linéarisation, cher aux physiciens.

Cette linéarisation des phénomènes physiques, et la nécessité d'élaborer des techniques calculatoires adaptées, a eu un rôle déterminant dans le développement de l'algèbre linéaire, et l'introduction des applications linéaires.

Dans tout ce qui suit, $\mathbb{K}$ désigne un corps quelconque. Vous pouvez considérer, conformément au programme, que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, mais, sauf mention explicite du contraire, les résultats donnés sont valables pour tout corps.

III.1 Définitions et propriétés de stabilité

Définition 23.3.2 (Application linéaire, AL)

Une application $f : E \to F$ entre deux $\mathbb{K}$-ev est appelée *application $\mathbb{K}$-linéaire*, ou plus simplement *application linéaire* (en abrégé : AL), si :

$$\forall \lambda \in \mathbb{K},\ \forall x \in E,\ \ f(\lambda x) = \lambda f(x) \qquad \text{et} \qquad \forall (x,y) \in E^2,\ f(x+y) = f(x) + f(y).$$

Proposition 23.3.3 (Respect du neutre)

Soit $f : E \to F$ une AL. Alors $f(0_E) = 0_F$.

◁ **Éléments de preuve.**

C'est un morphisme de groupe ! ▷

Proposition 23.3.4 (Caractérisation des AL par respect des CL)

Une application $f : E \to F$ est une application linéaire si et seulement si :

$$\forall \lambda \in \mathbb{K},\ \forall (x,y) \in E^2,\ \ f(\lambda x + y) = \lambda f(x) + f(y).$$

◁ **Éléments de preuve.**

Sens direct évident. Sens réciproque en considérant des valeurs particulières de λ ou y. ▷

Exemples 23.3.5

1. L'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $f(x,y) = 2x + 3y$?
2. L'application de $\mathbb{R}^3$ dans $\mathbb{R}$ définie par $f(x,y,z) = 3x + 4y + 2z + 1$?
3. L'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $f(x,y) = xy$?
4. La somme $\sum$, de $\mathbb{C}^n$ dans $\mathbb{C}$.
5. L'intégrale, de $\mathcal{C}^0([a,b])$ dans $\mathbb{R}$.
6. L'opérateur de dérivation D de $\mathcal{C}^n$ dans ...
7. L'opérateur de dérivation D de $\mathbb{K}[X]$ dans $\mathbb{K}[X]$.
8. Étant donnés des sous-espaces vectoriels $E_1, \dots, E_n$ de E, l'application

$$\varphi : (x_1, \dots, x_n) \in E_1 \times \dots \times E_n \longmapsto x_1 + \dots + x_n.$$

9. Étant donnés $x_1, \dots, x_n$ dans E, l'application

$$\psi : (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n \longmapsto \lambda_1 x_1 + \dots + \lambda_n x_n.$$

10. Cas particulier : $\mathcal{B}$ étant une base de E, constituée d'un nombre fini n de vecteurs, l'application $v_{\mathcal{B}}$ qui à un élément X de $\mathbb{K}^n$ associe un vecteur x dont les coordonnées dans $\mathcal{B}$ sont X.
11. Étant donné une matrice $M \in \mathcal{M}_{n,p}(\mathbb{K})$, et en assimilant $\mathbb{K}^n$ à $\mathcal{M}_{n,1}(\mathbb{K})$, l'application de $\mathbb{K}^p$ dans $\mathbb{K}^n$ qui à une matrice colonne X associé la matrice colonne MX.

Nous verrons dans le chapitre suivant que ce dernier exemple fournit la description générique d'une application linéaire en dimension finie, après choix de bases de E et de F.

Proposition 23.3.6 (Linéarité généralisée)

Soit $f : E \to F$ une application linéaire, et $(x_i)_{i\in I}$ une famille de vecteurs de E et $(\lambda_i)_{i\in I}$ une famille de scalaires à support fini. Alors

$$f\left(\sum_{i\in I} \lambda_i x_i\right) = \sum_{i\in I} \lambda_i f(x_i).$$

◁ **Éléments de preuve.**

Récurrence facile. ▷

Nous obtenons ainsi l'ensemble des flèches entre deux objets de la catégorie des $\mathbb{K}$-espaces vectoriels.

Définition 23.3.7 (Ensemble des applications linéaires)

Soit E et F deux $\mathbb{K}$-ev. On note $\mathcal{L}(E,F)$ l'ensemble des applications linéaires de E vers F.

Ces ensembles de flèches possèdent eux-même une structure d'espace vectoriel :

Proposition 23.3.8 (Structure de $\mathcal{L}(E,F)$)

$\mathcal{L}(E,F)$ est un espace vectoriel sur $\mathbb{K}$.

◁ **Éléments de preuve.**

L'application nulle est dans $\mathcal{L}(E,F)$. Il suffit alors de justifier qu'une combinaison linéaire d'applications linéaires est linéaire, c'est-à-dire respecte les combinaisons linéaires de vecteurs. Sans difficulté, mais ne vous embrouillez pas dans les combinaisons linéaires : il y a celle relative aux applications et celle relative aux arguments. ▷

Autrement dit, une combinaison linéaire d'applications linéaires de E vers F est encore une application linéaire.
Étudions maintenant des propriétés liées à la composition.

Proposition 23.3.9 (Composée de deux applications linéaires)

Soit $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$. Alors $g \circ f$ est une application linéaire de E vers G.

◁ **Éléments de preuve.**

Déjà fait dans le contexte général. Peut être refait facilement dans ce cas particulier. ▷

De manière générale, la composition d'applications à valeurs dans un espace vectoriel est toujours linéaire à gauche, c'est-à-dire $(\lambda f + \mu g) \circ h = \lambda f \circ h + \mu g \circ h$. Si les applications considérés sont linéaires, on obtient aussi la linéarité à droite. On rappelle :

Définition 23.3.10 (Application bilinéaire)

Soit E, F et G trois $\mathbb{K}$-espaces vectoriels, et $\varphi : E \times F \to G$. On dit que φ est bilinéaire si elle est linéaire par rapport à chacune de ses deux variables, l'autre étant fixée, c'est-à-dire si pour tout $(x, x', y, y', \lambda) \in E \times E \times F \times F \times \mathbb{K}$,

- $\varphi(\lambda x + x', y) = \lambda\varphi(x,y) + \varphi(x', y)$
- $\varphi(x, \lambda y + y') = \lambda\varphi(x,y) + \varphi(x, y')$.

La propriété énoncée plus haut pour les compositions s'exprime alors de la manière suivante :

Proposition 23.3.11 (Bilinéarité de la composition)

La composition d'applications linéaires est bilinéaire. En termes plus précis, E, F et G étant trois $\mathbb{K}$-ev, l'application Φ de $\mathcal{L}(E,F) \times \mathcal{L}(F,G)$ dans $\mathcal{L}(E,G)$ définie par $\Phi(u,v) = v \circ u$ est une application bilinéaire.

◁ **Éléments de preuve.**

La linéarité à gauche ne pose pas de problème, et est vraie pour toute composition de fonctions, à condition que G soit un espace vectoriel (de sorte à pouvoir définir des combinaisons linéaires de fonctions)

La linéarité à droite provient de la linéarité de v. ▷

On pourrait de façon similaire définir une notion d'application n-linéaire (à n variables vectorielles). La composition de n AL est alors n-linéaire. On verra dans un chapitre ultérieur comment cette notion d'application multilinéaire est également liée à la notion de déterminant.

Remarque 23.3.12

Cette propriété signifie qu'on peut développer une composition d'applications linéaires comme un produit, par distributivité.

III.2 Image et noyau

Dans ce paragraphe, E et F sont deux espaces vectoriels, et $f \in \mathcal{L}(E,F)$. On étudie ici deux sous-espaces liés à une application linéaire : l'image (qui correspond à la notion usuelle d'image) et le noyau.

Définition 23.3.13 (Image et noyau)

1. L'image de f est $\mathrm{Im}(f) = \{y \in F \mid \exists x \in E,\ f(x) = y\} = f(E)$;
2. Le noyau de f est $\mathrm{Ker}(f) = \{x \in E \mid f(x) = 0\} = f^{-1}(\{0\})$

La structure algébrique de l'image et du noyau découle d'un résultat plus général de préservation de la structure par image directe et réciproque :

Lemme 23.3.14 (Structure des images directes et réciproques)

1. *Soit E' un sev de E. Alors $f(E')$ est un sev de F.*
2. *Soit F' un sev de F. Alors $f^{-1}(F')$ est un sev de E.*

◁ **Éléments de preuve.**

Vérifications faciles par caractérisation des sev. ▷

En appliquant ce lemme avec $E' = E$ et $F' = \{0\}$, on obtient :

Proposition 23.3.15 (Structure de l'image et du noyau)

$\mathrm{Im}(f)$ *est un sev de* F. $\mathrm{Ker}(f)$ *est un sev de* E.

De façon évidente, l'image mesure le défaut de surjectivité, De façon symétrique, le noyau mesure le défaut d'injectivité : si $\mathrm{Ker}(f)$ n'est pas un singleton, les définitions amènent de façon immédiate la non injectivité de f. La réciproque découle du fait que tout défaut d'injectivité peut être translaté en 0 par linéarité. Nous obtenons :

Théorème 23.3.16 (Caractérisation de la surjectivité et de l'injectivité)
Soit $f \in \mathcal{L}(E, F)$. *Alors :*
(i) f *est surjective ssi* $\mathrm{Im}(f) = F$ *;*
(ii) f *est injective ssi* $\mathrm{Ker}(f) = \{0\}$.

◁ **Éléments de preuve.**
La caractérisation de la surjectivité est évidente, celle de l'injectivité découle du fait qu'une application linéaire est en particulier un morphisme de groupes additifs. ▷

Connaissant une famille génératrice de E (par exemple une base), il n'est pas dur de déterminer l'image de f :

Proposition 23.3.17 (Famille génératrice de Im(f))
Soit $f \in \mathcal{L}(E, F)$ *et* $(e_i)_{i\in I}$ *une famille génératrice de* E. *Alors* $f(e_i)_{i\in I}$ *est une famille génératrice de* $\mathrm{Im}(f)$ *:*

$$\mathrm{Im}(f) = \mathrm{Vect}(f(e_i), i \in I).$$

En particulier, si f transforme une famille génératrice de E en une famille génératrice de F, alors f est surjective.
Si nous appliquons la propriété précédente à l'application de $\mathbb{K}^p$ dans $\mathbb{K}^n$ définie par $X \mapsto MX$, nous obtenons la description très simple d'une famille génératrice de $\mathrm{Im}(f)$, dans le cas où f est donnée matriciellement :

Corollaire 23.3.18 (Image d'une AL décrite matriciellement)
Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ *et* $f : \mathbb{K}^p \to \mathbb{K}^n$ *(ces ensembles étant vus comme ensemble de vecteurs colonnes), définie par* $f(X) = MX$. *Alors l'image de* f *est engendrée par la famille des colonnes de la matrice* M.

Exemple 23.3.19
Décrire l'image de $f : X \in \mathbb{R}^3 \mapsto \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{pmatrix} X$.

On en déduit :

Méthode 23.3.20 (Déterminer l'image et le noyau d'une AL)
- Pour déterminer l'image d'une AL f :
  - ⋆ Si f est donnée par une matrice (éventuellement après choix d'une base, voir chapitre suivant), considérer la famille des colonnes de cette matrice, qui engendre $\mathrm{Im}(f)$;
  - ⋆ Sinon, trouver une famille génératrice de E (par exemple une base) et considérer son image.
- Pour déterminer le noyau d'une AL f :
  - ⋆ Écrire l'équation $f(x) = 0$.
  - ⋆ Si nécessaire, décomposer x dans une base $\mathcal{B}$ de E. Grâce à la linéarité de f, l'équation précédente se ramène alors à un système linéaire d'équations portant sur les coordonnées de x dans la base $\mathcal{B}$.

Conformément à la terminologie générale, nous définissons :

Définition 23.3.21 (Isomorphisme)

1. Une application linéaire bijective de E vers F est appelée un *isomorphisme*.
2. On dit que deux espaces vectoriels E et F sont isomorphes s'il existe un isomorphisme $f : E \to F$.

Les résultats généraux sur les structures amènent directement :

Théorème 23.3.22 (Réciproque d'un isomorphisme)

Soit f un isomorphisme entre E et F. Alors f^{-1} est une application linéaire, et donc un isomorphisme de F vers E.

Exemple 23.3.23

Si $\mathcal{B}$ est une base de cardinal n, l'application $v_{\mathcal{B}} : \mathbb{K}^n \to E$ est un isomorphisme. Sa réciproque $v_{\mathcal{B}}^{-1}$ est l'application linéaire qui consiste à associer à un vecteur x ses coordonnées dans la base $\mathcal{B}$.

III.3 Endomorphismes

Conformément à la terminologie générale, un endomorphisme est une application linéaire d'un espace dans lui-même :

Définition 23.3.24 (Endomorphisme)

Soit E un espace vectoriel sur $\mathbb{K}$. Une application linéaire de E dans E est appelée *endomorphisme de E*. On note $\mathcal{L}(E)$ l'ensemble $\mathcal{L}(E, E)$ des endomorphismes de E.

Exemples 23.3.25

1. L'identité $\mathrm{Id}_E : x \mapsto x$ de E dans E.
2. L'homothétie (vectorielle) de rapport λ : $\lambda \mathrm{Id}_E : x \mapsto \lambda x$, de E dans E ($\lambda \in \mathbb{K}$).
3. La dérivation formelle $D : \mathbb{R}[X] \to \mathbb{R}[X]$
4. La dérivation analytique $D : \mathcal{C}^?([a, b]) \to \mathcal{C}^?([a, b])$
5. L'application matricielle $X \mapsto MX$ si M est...

La bilinéarité de la composition des applications linéaires permet de définir sur $\mathcal{L}(E)$ une loi de composition $\circ$, distributive sur la somme. On obtient alors la structure de l'ensemble $\mathcal{L}(E)$ des endomorphismes de E :

Proposition 23.3.26 (Structure de $\mathcal{L}(E)$)

L'ensemble $(\mathcal{L}(E), +, \cdot, \circ)$ est muni d'une structure de $\mathbb{K}$-algèbre.

Nous rappelons ci-dessous la définition d'une $\mathbb{K}$-algèbre :

Définition 23.3.27 ($\mathbb{K}$-algèbre)

Étant donné un corps $\mathbb{K}$, une $\mathbb{K}$-algèbre est un espace vectoriel A sur $\mathbb{K}$, muni d'une seconde loi de composition interne $\times$, compatible avec la loi externe dans le sens suivant :

$$\forall \lambda \in \mathbb{K}, \quad \forall (x, y) \in A^2, \quad \lambda \cdot (x \times y) = (\lambda \cdot x) \times y = x \times (\lambda \cdot y),$$

et telle que $(A, +, \times)$ soit un anneau.

La composition ayant les propriétés usuelles d'un produit, on utilise souvent les conventions de notation usuelles pour les produits. En particulier la notation vu désigne l'endomorphisme $v \circ u$ (omission du signe opératoire), et :

Notation 23.3.28 (Composition itérée)

Étant donné un endomorphisme u de E, et un entier $n \in \mathbb{N}$, on désigne par u^n la n-ième composée itérée de u, définie récursivement par $u^0 = \mathrm{Id}_E$ et $u^n = u \circ u^{n-1}$ pour $n \in \mathbb{N}^*$.

Remarquez que ces conventions de notation ne sont pas gênantes dans la mesure où dans l'espace vectoriel E, on ne dispose pas d'un produit. L'expression $v(x)u(x)$ n'ayant pas de sens, la notation vu ne peut désigner que la composition. De même pour f^n, l'expression $f(x)^n$ n'ayant pas de sens.

Remarque 23.3.29

1. L'anneau $(\mathcal{L}(E), +, \circ)$ est-il commutatif ? À quelle condition nécessaire et suffisante sur la dimension de E l'est-il ?
2. L'anneau $(\mathcal{L}(E), +, \circ)$ est-il intègre ?

Cette dernière remarque amène la définition suivante :

Définition 23.3.30 (Endomorphisme nilpotent)

Un endomorphisme nilpotent de E est un endomorphisme u tel qu'il existe $n \in \mathbb{N}$ tel que $u^n = 0_{\mathcal{L}(E)}$. La valeur minimale de n vérifiant cette propriété est appelée *indice de nilpotence de u*.

Exemple 23.3.31

1. $X \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} X$
2. La dérivation, vue comme endomorphisme de ...

Ainsi, d'une certaine façon, si u est nilpotent d'incide n, l'endormophisme u « annule » le polynôme X^n. On dira alors que X^n est polynôme annulateur de u. Un autre polynôme annulateur est X^{n+1}, ou encore $X^n(X-1)$. La valeur de n étant minimale, X^n est le polynôme unitaire de plus bas degré annulant u : on dira dans ce cas qu'il s'agit du polynôme minimal de u. Nous généralisons ci-dessous ces notions pour des endomorphismes quelconques.

Définition 23.3.32 (Polynôme d'endomorphisme)

Soit $P = \sum_{k=0}^{n} a_k X^k$ un polynôme de $\mathbb{K}[X]$ et u un endomorphisme d'un $\mathbb{K}$-ev E. On définit le polynôme $P(u)$ de l'endomorphisme u par :

$$P(u) = \sum_{k=0}^{n} a_k u^k : \ x \longmapsto \sum_{k=0}^{n} a_k u^k(x)$$

La structure d'algèbre de $\mathcal{L}(E)$ nous permet d'affirmer que $P(u) \in \mathcal{L}(E)$.

Remarquez qu'il s'agit d'un cas particulier de spécialisation d'un polynôme à des éléments d'une algèbre, ainsi que nous l'avons vu dans le chapitre sur les polynômes.
Deux polynômes d'un même endomorphisme commutent, ce qui découle de la commutativité dans $\mathbb{K}[X]$, via le lemme suivant :

Lemme 23.3.33 (Spécialisation d'un produit)

Soit u un endomorphisme de E et P et Q deux polynômes de $\mathbb{K}[X]$. Alors $(PQ)(u) = P(u) \circ Q(u)$.

◁ **Éléments de preuve.**
Noter $B_i = X^i$, et remarquer que $B_iB_j(u) = u^{i+j} = B_i(u) \circ B_j(u)$. Écrire alors P et Q comme combinaisons des B_i. ▷

Corollaire 23.3.34 (Commutation des polynômes d'endomorphisme)

Soit u un endormophisme de E et P et Q deux polynômes de $\mathbb{K}[X]$. Alors les deux endomorphismes $P(u)$ et $Q(u)$ commutent :

$$P(u) \circ Q(u) = Q(u) \circ P(u).$$

Cela permet en particulier d'exploiter une factorisation d'un polynôme P pour le calcul de $P(u)$.

Exemple 23.3.35

D désignant l'endomorphisme de dérivation de $\mathcal{C}^\infty(\mathbb{R})$, et P désignant le polynôme $P = (X-1)(X-2)(X-3) = X^3 - 6X^2 + 11X - 6$, on obtient, pour une fonction f de classe $\mathcal{C}^\infty$:

$$f^{(3)} - 6f^{(2)} + 11f' - 6f = (D - \mathrm{Id}) \circ (D - 2\mathrm{Id}) \circ (D - 3\mathrm{Id})(f),$$

ce qu'on vérifie aisément de façon directe (Id désigne l'identité de $\mathcal{C}^\infty(\mathbb{R})$).

Nous pouvons alors généraliser les notions entrevues lors de l'étude des endomorphismes nilpotents.

Définition 23.3.36 (Polynôme annulateur)

On dit que $P \in \mathbb{K}[X]$ est un polynôme annulateur de $u \in \mathcal{L}(E)$ si $P(u) = 0_{\mathcal{L}(E)}$.

Proposition 23.3.37 (Structure de l'ensemble des polynômes annulateurs)

L'ensemble des polynômes annulateurs de u forme un idéal de $\mathbb{K}[X]$.

◁ **Éléments de preuve.**
C'est une bonne occasion de revoir la définition d'un idéal. ▷

L'anneau $\mathbb{K}[X]$ étant principal, on peut alors définir :

Définition 23.3.38 (Polynôme minimal)

Si l'endomorphisme u admet au moins un polynôme annulateur non nul, on définit le polynôme minimal de u comme étant l'unique polynôme unitaire engendrant l'idéal des polynômes annulateurs de u.

III.4 Automorphisme

Définition 23.3.39 (Automophisme, GL(E))

Un automophisme de E est un endomorphisme bijectif, donc une application linéaire qui est à la fois un endomorphisme et un isomorphisme. On appelle *groupe linéaire de E*, et on note $\mathrm{GL}(E)$, l'ensemble des automorphismes de E

La terminologie est justifiée par le théorème suivant :

Théorème 23.3.40 (Structure de $\mathrm{GL}(E)$)
L'ensemble $\mathrm{GL}(E)$ *muni de la composition des endomorphismes, est un groupe.*

◁ **Éléments de preuve.**
Conséquence d'un résultat plus général sur les anneaux. ▷

Comme on le verra, dans le chapitre suivant, si E est de dimension finie n, il existe un isomophisme de groupe entre $\mathrm{GL}(E)$ et le groupe $\mathrm{GL}_n(\mathbb{K})$ des matrices inversibles de taille $n \times n$ à coefficients dans $\mathbb{K}$.

III.5 Projecteurs et symétries

Définition 23.3.41 (Projecteur, symétrie)

1. Soit $p \in \mathcal{L}(E)$. On dit que p est un projecteur de E ssi $p \circ p = p$.
2. Soit $s \in \mathcal{L}(E)$. On dit que s est une symétrie ssi $s \circ s = \mathrm{id}$.

Ainsi, un projecteur est par définition un endomorphisme dont $X^2 - X$ est polynôme annulateur, et une symétrie est un endomorphisme dont $X^2 - 1$ est polynôme annulateur.

Proposition 23.3.42 (Caractérisation de l'image d'un projecteur)
Soit p un projecteur de E. Alors $x \in \mathrm{Im}(p)$ si et seulement si $p(x) = x$, ce qui se traduit par l'égalité ensembliste :

$$\mathrm{Im}(p) = \mathrm{Ker}(p - \mathrm{Id}_E).$$

◁ **Éléments de preuve.**
Écrire $x = p(y)$, et appliquer p. Réciproque évidente. ▷

Théorème 23.3.43 (Diagonalisation d'un projecteur)
Soit p un projecteur de E. Alors :

$$E = \mathrm{Ker}(p) \oplus \mathrm{Ker}(p - \mathrm{Id}_E).$$

◁ **Éléments de preuve.**
Écrire $x = x - p(x) + p(x)$. Ne pas oublier aussi de justifier la somme directe. ▷

Cette propriété exprime le fait que p est « diagonalisable ».
De façon générale, dire qu'un endomorphisme f est diagonalisable signifie que l'espace E peut être décomposé en somme directe d'espaces stables par f sur lesquels p induit une homothétie. Ainsi, f est entièrement déterminé par un ensemble de vecteurs engendrant E et sur lesquels f est simplement une dilatation. En termes plus précis :

Définition 23.3.44 (Endomorphisme diagonalisable)

- Un endomorphisme f de E est diagonalisable s'il existe une base $(b_i)_{i \in I}$ de E et une famille $(\lambda_i)_{i \in I}$ de scalaires tels que pour tout $i \in I$, $f(b_i) = \lambda_i b_i$.
- Les λ_i sont appelées valeurs propres de f.

- Étant donné une valeur propre λ, un vecteur x non nul tel que $f(x) = \lambda x$ est appelé vecteur propre associé à λ
- Si λ est une valeur propre de f, $\mathrm{Ker}(f - \lambda \mathrm{Id})$ est appelé sous-espace propre de f associé à la valeur propre λ.

Ainsi, dire que f est diagonalisable revient à dire que E est somme directe des sous-espaces propres de f. Vous verrez l'année prochaine que cela équivaut en dimension finie à l'existence d'une base (par exemple la base donnée dans l'énoncé de la définition) relativement à laquelle la matrice de f est diagonale (voir chapitre suivant pour la notion de matrice associée à un endomorphisme)

Corollaire 23.3.45

Un projecteur distinct de 0 ou Id, admet exactement deux valeurs propres 0 et 1, et est diagonalisable.

Remarque 23.3.46

Les valeurs propres de p sont les racines du polynôme annulateur $X^2 - X$. Ce n'est pas anodin. L'ensemble des valeurs propres est toujours inclus dans l'ensemble des racines d'un polynôme annulateur, et on a l'égalité s'il s'agit du polynôme minimal.

Le théorème précédent est en fait une caractérisation des projecteurs

Théorème 23.3.47 (Caractérisation géométrique des projecteurs)

Soit $p \in \mathcal{L}(E)$.

- *Alors p est un projecteur si et seulement s'il existe deux sev F et G de E tels que $F \oplus G = E$, et :*
$$\forall u \in F,\ \forall v \in G,\ p(u+v) = u.$$
Cette dernière identité traduit le fait que p est la projection géométrique sur F parallèlement à G.
- *Dans ce cas, on a $F = \mathrm{Im}(p)$ et $G = \mathrm{Ker}(p)$.*
- *Ainsi, un projecteur est une projection géométrique sur $\mathrm{Im}(p)$ parallèlement à $\mathrm{Ker}(p)$.*

◁ **Éléments de preuve.**

On vient de faire le plus dur (le sens direct), provenant de la diagonalisation. La réciproque est facile. ▷

Définition 23.3.48 (Projecteurs associés)

Soit $E = F \oplus G$ et p le projecteur sur F parallèlement à G. Le projecteur associé à p est le projecteur q sur G parallèlement à F. Les projecteurs p et q sont donc reliés par la relation $p + q = \mathrm{Id}_E$. Ainsi, le projecteur associé à p est $\mathrm{id}_E - p$.

Une description géométrique similaire peut être obtenue pour les symétries.

Théorème 23.3.49 (Diagonalisation d'une symétrie)

Soit E un espace vectoriel sur un corps $\mathbb{K}$ de caractéristique différente de 2. Soit s une symétrie de E. Alors :
$$E = \mathrm{Ker}(s + \mathrm{Id}_E) \oplus \mathrm{Ker}(s - \mathrm{Id}_E).$$

◁ **Éléments de preuve.**
Quelle décomposition de x adopter cette fois ? Aidez-vous d'un dessin et de votre intuition géométrique de ce qu'est une symétrie. Si vraiment vous ne trouvez pas, vous pouvez vous lancer dans une analyse-synthèse. ▷

Ce dernier résultat traduit le fait que s est diagonalisable, et si s est distinct de Id_E et $-\mathrm{Id}_E$, alors les valeurs propres de s sont exactement 1 et -1. On peut à nouveau remarquer qu'il s'agit des racines du polynôme annulateur $X^2 - 1$.
Encore une fois, le théorème précédent est une caractérisation des symétries, et donne l'interprétation géométrique des symétries.

Théorème 23.3.50 (Caractérisation géométrique des symétries)
Soit $s \in \mathcal{L}(E)$.

- *Alors s est une symétrie si et seulement s'il existe deux sev F et G de E tels que $F \oplus G = E$, et :*
$$\forall u \in F,\ \forall v \in G,\ s(u+v) = u - v.$$
Cette dernière identité traduit le fait que s est la symétrie géométrique par rapport à F parallèlement à G.
- *Dans ce cas, on a $F = \mathrm{Ker}(s - \mathrm{Id}_E)$ et $G = \mathrm{Ker}(s + \mathrm{Id}_E)$.*
- *Ainsi, une symétrie au sens algébrique s est une symétrie géométrique par rapport à $\mathrm{Ker}(s - \mathrm{Id}_E)$ (l'ensemble des points fixes), parallèlement à $\mathrm{Ker}(s + \mathrm{Id}_E)$.*

IV Applications linéaires et familles de vecteurs

IV.1 Détermination d'une application linéaire

Nous commençons par un résultat de rigidité, exprimant le fait que l'image d'un nombre limité de vecteurs de E par une application linéaire u permet de déterminer entièrement l'application linéaire u. En effet, par linéarité, la connaissance de u sur une famille de vecteurs amène sa connaissance sur tout l'espace engendré par cette famille. On obtient donc le résultat suivant :

Proposition 23.4.1 (Détermination d'une AL par l'image d'une base, ou rigidité)
Étant donné $(b_i)_{i \in I}$ une base de E et $(f_i)_{i \in I}$ une famille quelconque de F, il existe une unique application linéaire $u \in \mathcal{L}(E, F)$ telle que pour tout $i \in I$, $u(b_i) = f_i$.

◁ **Éléments de preuve.**
En décomposant x dans la base $\mathcal{B} = (b_i)$, on obtient une unique description possible de u. Vérifier que cette description définit bien une application linéaire. On pourra utiliser $v_{\mathcal{B}}^{-1}$ l'application qui à un vecteur associe ses coordonnées. ▷

Ainsi, une application linéaire de E dans F est entièrement déterminée par l'image d'une base de E.

Exemples 23.4.2

1. Déterminer l'expression générale de l'application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$ telle que $f(1,0) = (3,2)$ et $f(0,1) = (2,1)$.
2. Montrer que toute application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$ est de la forme $X \mapsto MX$, où $M \in \mathcal{M}_{n,p}(\mathbb{K})$. Cette matrice est notée
$$M = \mathrm{Mat}_{b.c.}(f),$$
le b.c. présent dans cette notation faisant référence à la base canonique. Ou plus précisément aux deux bases canoniques de $\mathbb{K}^p$ et $\mathbb{K}^n$ respectivement.

3. Soit $(b_j)_{j\in J}$ une base de E et $(c_i)_{i\in I}$ une base de F. Alors pour tout $(i,j)\in I\times J$, il existe une unique application linéaire $u_{i,j}$ telle que $u_{i,j}(b_j)=c_i$ et pour tout $k\neq j$, $u_{i,j}(b_k)=0$.

Proposition 23.4.3 (Base de $\mathcal{L}(E,F)$)

Avec les notations précédentes, si J est fini (ce qui traduit le fait que E est de dimension finie), la famille $(u_{i,j})_{(i,j)\in I\times J}$ décrite dans l'exemple ci-dessus est une base de $\mathcal{L}(E,F)$.

◁ **Éléments de preuve.**

- Pour la liberté, évaluer une combinaison linéaire en b_i puis exploiter la liberté de (c_j).
- Pour le caractère générateur, les décompositions des $f(b_i)$ dans la base (c_j) définissent les coefficients à attribuer aux $u_{i,j}$.

▷

Proposition 23.4.4 (Détermination d'une AL sur une somme directe)

Soit E et F deux espaces vectoriels et E_1 et E_2 deux sous-espaces supplémentaires de E, c'est-à-dire tels que $E_1\oplus E_2=E$. Alors l'application de $\mathcal{L}(E,F)$ dans $\mathcal{L}(E_1,F)\times\mathcal{L}(E_2,F)$ qui à u associe $(u_{|E_1},u_{|E_2})$ est un isomorphisme.

Ainsi, une application linéaire est entièrement déterminée, de façon unique et non équivoque, par ses restrictions à deux sous-espaces supplémentaires de l'espace de départ.

IV.2 Caractérisations de l'injectivité et de la surjectivité par l'image de bases

L'image d'une base par un endomorphisme déterminant entièrement l'application linéaire, toutes les propriétés telles que l'injectivité et la surjectivité peuvent se voir déjà dans la description de l'image d'une base. Ainsi, ces propriétés peuvent être caractérisées par l'image d'une base.

Proposition 23.4.5 (Caractérisation de l'injectivité par l'image d'une base)

Soit $f\in\mathcal{L}(E,F)$. Les propriétés suivantes sont équivalentes :

(i) f est injective ;

(ii) l'image de toute famille libre de E par f est une famille libre de F ;

Si de plus, E admet au moins une base (par exemple si E est de dimension finie, ou bien en supposant l'axiome du choix), elles sont aussi équivalentes à :

(iii) l'image de toute base de E par f est une famille libre de F ;

(iv) il existe une base de E dont l'image par f est une famille libre de F.

◁ **Éléments de preuve.**

- $(i)\Longrightarrow(ii)$: par injectivité ramener une relation sur les images à une relation sur les antécédents.
- $(ii)\Longrightarrow(i)$: étudier le noyau, en remarquant qu'une famille constitué d'un unique vecteur x est libre si et seulement si $x\neq 0$.
- $(ii)\Longrightarrow(iii)\Longrightarrow(iv)$: évident
- $(iv)\Longrightarrow(i)$: décomposer un vecteur du noyau dans la base donnée par (iv) et appliquer f.

Où utilise-t-on l'hypothèse supplémentaire donnée avant (iii) ? ▷

Proposition 23.4.6 (Caractérisation de la surjectivité par l'image d'une base)
Soit $f \in \mathcal{L}(E,F)$. Les propriétés suivantes sont équivalentes :

(i) f est surjective ;

(ii) l'image de toute famille génératrice de E par f est une famille génératrice de F.

Si de plus, E admet au moins une base, elles sont aussi équivalentes à :

(iii) l'image de toute base de E par f est une famille génératrice de F ;

(iv) il existe une base de E dont l'image par f est une famille génératrice de F.

◁ **Éléments de preuve.**

- $(i) \Longrightarrow (ii)$: Relever y de F dans E, décomposer dans la famille génératrice donnée et appliquer f.
- $(ii) \Longrightarrow (i)$. Décomposer y dans la famille génératrice image, cela fournit un antécédent de y.
- $(ii) \Longrightarrow (iii) \Longrightarrow (iv)$: évident
- $(iv) \Longrightarrow (i)$: comme $(ii) \Longrightarrow (i)$.

▷

En combinant ces deux caractérisations, nous obtenons :

Proposition 23.4.7 (Caractérisation de la bijectivité par l'image d'une base)
Soit $f \in \mathcal{L}(E,F)$. Si E admet au moins une base (ce qui est toujours vrai en dimension finie, et vrai en dimension quelconque avec l'AC), les propriétés suivantes sont équivalentes :

(i) f est un isomorphisme ;

(ii) l'image de toute base de E par f est une base de F ;

(iii) il existe une base de E dont l'image par f est une base de F.

V Sous-espaces affines d'un espace vectoriel

De nombreux contextes linéaires (suites récurrentes linéaires, équations différentielles linéaires, systèmes linéaires) amènent des résultats un peu similaires. Nous avons vu en particulier dans divers contextes liés à des équations linéaires, l'ensemble des solutions a une structure bien particulière : la somme d'une solution particulière et de l'ensemble des solutions de l'équation homogène.
Dans ce paragraphe, nous étudions rapidement cette structure de sous-espace affine, et nous voyons que les différents exemples rencontrés sont issus de la même description comme fibre d'une application linéaires.

Définition 23.5.1 (Sous-espace affine d'un espace vectoriel E)
Soit E un espace vectoriel. Un sous-ensemble $F \subset E$ est un sous-espace affine de E s'il existe un point $A \in E$ et un sous-espace vectoriel V de E tels que

$$F = A + V = \{A + x \mid x \in V\}.$$

Remarque 23.5.2
Les éléments de E jouent à la fois le rôle de points et le rôle de vecteurs. Les éléments de F (dont A) sont à voir comme des points, alors que les éléments de V sont à comprendre comme des vecteurs de translation : l'ensemble F est donc l'ensemble des points de E obtenus en translatant A par un vecteur quelconque de V.

Notation 23.5.3

Soit A et $B \in E$ (interprétés comme des points). On définit le vecteur $\overrightarrow{AB} = B - A$.

Lemme 23.5.4 (Caractérisation de l'appartenance à un sous-espace affine)

Soit F le sous-espace affine $A + V$. Alors $B \in F$ si et seulement si $\overrightarrow{AB} \in V$.

◁ **Éléments de preuve.**

C'est une réexpression de la définition. ▷

Proposition 23.5.5 (Unicité de la direction)

Soit A et B deux points de E et V et W deux sous-espaces vectoriels de E. Alors

$$A + V = B + W \iff V = W \text{ et } \overrightarrow{AB} \in V.$$

◁ **Éléments de preuve.**

Sens direct : utiliser le lemme précédent pour montrer que $V \subset W$

Sens réciproque facile par double inclusion. ▷

Définition 23.5.6 (Direction)

Soit F un sous-espace affine de E. L'unique sous-espace vectoriel V tel que F puisse s'écrire $A + V$, pour un certain $A \in E$, est appelé direction de F.

Proposition 23.5.7 (Représentations de F)

Soit F un sous-espace affine de E, de direction V. Alors

$$F = A + V \iff A \in F.$$

◁ **Éléments de preuve.**

C'est la combinaison de deux résultats précédents. ▷

Théorème 23.5.8 (Intersection de sous-espaces affines)

L'intersection de sous-espaces affines est soit vide, soit égale à un sous-espace affine. Si cette intersection est non vide, sa direction est l'intersection des directions de chacun des sous-espaces affines.

◁ **Éléments de preuve.**

Utiliser la caractérisation 23.5.7. ▷

Certains auteurs définissent parfois la notion de sous-variété affine : il s'agit d'un sous-espace affine, ou de l'ensemble vide. Ainsi, l'intersection de sous-variétés affines est toujours une sous-variété affine.

Un théorème important fournissant des sous-espaces affines (et on en a déjà vu des cas particuliers) est le suivant :

Théorème 23.5.9 (Fibres d'une application linéaire)

Soit $u \in \mathcal{L}(E,F)$, et $a \in F$. Alors l'image réciproque $u^{-1}(\{a\})$ (appelée fibre en a de u, ou ligne de niveau), est soit l'ensemble vide, soit un sous-espace affine dirigé par $\mathrm{Ker}(u)$ *(donc toujours une sous-variété affine).*

◁ **Éléments de preuve.**

Manipulations simples, par équivalences successives, en utilisant la linéarité de u. On avait déjà vu une propriété similaire dans le cadre des groupes. ▷

Les exemples que nous avons déjà rencontrés sont :

Exemple 23.5.10 (Sous-espaces affines obtenus comme fibres)

1. Ensemble des solutions d'un système linéaire
2. Résolution des équations différentielles linéaires non homogènes de degré 1 ou 2.
3. L'ensemble des polynômes interpolateurs en un certain nombre de points.
4. Équations arithmético-géométriques et autres.

24
Dimension finie

La première impulsion est venue de considérations sur la signification des nombres négatifs en géométrie. Habitué à voir AB comme une longueur, j'étais néanmoins convaincu que $AB = AC + CB$, quelle que soit la position de A, B et C sur une droite.

(Herrmann Günther Grassmann)

Il semble que ce soit le destin de Grassmann d'être redécouvert de temps en temps, à chaque fois comme s'il avait été pratiquement oublié.

(A.C. Lewis)

Si l'espace vectoriel E admet une base finie, tout vecteur va pouvoir être décrit à l'aide d'un nombre fini de scalaires (ses coordonnées dans la base). Si F également admet une base finie, les applications linéaires de E dans F également pourront être décrites par un nombre fini de scalaires, puisqu'elles sont déterminées par l'image des vecteurs de la base de E, chacune de ces images étant déterminée par la donnée des coordonnées dans la base de $\mathcal{F}$: ainsi, une application linéaire sera dans ce contexte déterminée par une famille de $n \times m$ scalaires, où n et m sont les cardinaux des bases respectives de E et F, et ces scalaires pourront être rangés dans une matrice à m lignes et n colonnes, chaque colonne représentant les coordonnées d'un vecteur image de la base de E. Ainsi, ce contexte nous donne un accès simple, explicite et concret à la fois aux vecteurs et aux applications linéaires, idéal dans tous les contextes calculatoires, et permet de ramener le calcul sur les vecteurs et les applications linéaires à du calcul matriciel.
Pour commencer, nous relions le cardinal des bases à la notion de dimension. Ce cardinal est en fait la définition de la dimension d'un espace, mais pour que la notion soit bien définie, il faut commencer par montrer que ce cardinal est indépendant de la base, donc que si E admet au moins une base finie, alors toutes ses bases auront le même cardinal (théorème de la dimension). Ce cardinal commun est alors la dimension de E.

I Espaces vectoriels de dimension finie

I.1 Théorie de la dimension

Définition 24.1.1 (Espace vectoriel de dimension finie)

Un espace vectoriel E sur $\mathbb{K}$ est dit *de dimension finie* s'il existe une famille génératrice de cardinal fini $(x_i)_{i \in I}$ de E. Une famille génératrice finie est souvent appelée *système de générateurs.*

Proposition 24.1.2

Soit E un espace vectoriel de dimension finie. Alors de toute famille génératrice de E, on peut extraire une famille génératrice finie.

◁ **Éléments de preuve.**

Soit $\mathcal{G}$ une famille génératrice, et $\mathcal{G}_0$ une famille génératrice finie. Considérer les éléments de $\mathcal{G}$ intervenant de façon effective dans des décompositions de tous les éléments de $\mathcal{G}_0$. Justifier qu'ils sont en nombre fini, et qu'ils forment une famille génératrice. ▷

Corollaire 24.1.3 (Finitude des bases)

Toute base d'un espace vectoriel de dimension finie est finie.

◁ **Éléments de preuve.**

On ne peut pas extraire strictement une sous-famille génératrice d'une base ! ▷

Ce résultat n'affirme pas l'existence d'une base ! Nous allons maintenant montrer l'existence d'une base, qu'on va obtenir par complétion d'une famille libre quelconque (par exemple de la famille vide). Il s'agit du théorème de la base incomplète, affirmant que toute famille libre peut être vue comme le début d'une base.

Théorème 24.1.4 (Théorème de la base incomplète)

Soit E un espace vectoriel de dimension finie.

(i) Soit $\mathcal{L}$ une famille libre de E, et $\mathcal{G}$ une famille génératrice de E. Alors on peut compléter $\mathcal{L}$ en une base de E par ajout de vecteurs de $\mathcal{G}$.

(ii) Toute famille libre peut être complétée en une base.

(iii) Tout espace de dimension finie admet au moins une base.

◁ **Éléments de preuve.**

1. Considérer $\mathcal{G}_0$ sous-famille génératrice finie de $\mathcal{G}$. Considérer une sous-famille $\mathcal{H}$ maximale de vecteurs de $\mathcal{G}_0$ à ajouter à $\mathcal{L}$ tout en préservant la liberté (pourquoi une telle famille existe-t-elle ?). Justifier que $\mathcal{L} \cup \mathcal{H}$ est génératrice.
2. Considérer une certaine famille génératrice $\mathcal{G}$ bien choisie.
3. Quelle famille libre $\mathcal{L}$ considérer ?

▷

En particulier, si $(x_i)_{1\leqslant i\leqslant n}$ est génératrice de E, et $(x_i)_{i\in I}$ est libre, pour $I \subset [\![1,n]\!]$, il existe J tel que $I \subset J \subset [\![1,n]\!]$, tel que $(x_j)_{j\in J}$ soit une base de E.

Remarque 24.1.5

La preuve donnée du théorème de la base incomplète s'adapte bien à la dimension infinie, avec le lemme de Zorn. Essayez de mettre les détails en place.

Corollaire 24.1.6 (Caractérisation de la dimension finie par le cardinal des familles libres)

Soit E un espace vectoriel. Alors E est de dimension finie si et seulement si toute famille libre de E est de cardinal fini.

◁ **Éléments de preuve.**

Compléter en une base. Que dire de son cardinal ? ▷

Corollaire 24.1.7 (Sous-espace d'un espace de dimension finie)

Soit E un espace vectoriel de dimension finie. Alors tout sous-espace vectoriel de E est aussi de dimension finie.

◁ **Éléments de preuve.**

Si ce n'est pas le cas, construire une famille libre infinie. La construction ne s'arrête pas car à aucun moment la famille obtenue est génératrice. ▷

Corollaire 24.1.8 (Théorème de la base extraite)

Soit E un espace vectoriel de dimension finie. De toute partie génératrice de E on peut extraire une base de E.

◁ **Éléments de preuve.**

Il s'agit encore une fois de revenir à la première version du théorème de la base incomplète, en choisissant bien $\mathcal{L}$. ▷

Nous établissons maintenant le théorème de la dimension, qui à la base de la théorie de la dimension. Pour le démontrer, nous utilisons le lemme suivant :

Lemme 24.1.9 (Théorème d'échange)

Soit E un espace vectoriel. Soit F une famille de vecteurs de E, et x et y dans E tels que $x \notin \mathrm{Vect}(F)$ et $x \in \mathrm{Vect}(F \cup \{y\})$. Alors $\mathrm{Vect}(F \cup \{x\}) = \mathrm{Vect}(F \cup \{y\})$.

◁ **Éléments de preuve.**

- L'inclusion directe est immédiate.
- Justifier que $y \notin \mathrm{Vect}(F)$ et que $y \in \mathrm{Vect}(F \cup \{x\})$. Cela symétrise donc les hypothèses, donc aussi la conclusion.

▷

Théorème 24.1.10 (Théorème de la dimension)

Soit E un espace vectoriel de dimension finie. Alors toutes les bases de E sont finies et de même cardinal.

◁ **Éléments de preuve.**

Se fixer une base de référence $\mathcal{C}$ de cardinal n. Montrer par récurrence descendante sur $k \in [\![0, n]\!]$ que toute base $\mathcal{B}$ telle que $|\mathcal{B} \cap \mathcal{C}| = k$ est de cardinal n. Pour l'hérédité, utiliser le theorème d'échange pour remplacer dans $\mathcal{B}$ un vecteur de $\mathcal{B} \setminus \mathcal{C}$ par un vecteur de $\mathcal{C}$: pourquoi existe-t-il un vecteur c de $\mathcal{C}$ respectant les conditions du lemme d'échange ? Pourquoi la famille obtenue est-elle encore une base ? ▷

L'importance de ce théorème provient du fait qu'il permet de définir la notion fondamentale suivante :

Définition 24.1.11 (Dimension d'un espace vectoriel)

Soit E est espace vectoriel de dimension finie. Le cardinal commun de toutes les bases de E est appelé *dimension de E*, et est noté $\dim E$. Si E n'est pas de dimension finie, on dira que E est de dimension infinie.

La notion de dimension est la formalisation de la notion de nombre de degrés de liberté dont on dispose pour construire un objet. Par exemple, pour définir entièrement une suite par une récurrence linéaire homogène d'ordre 3, on dispose de 3 degrés de liberté, à savoir le choix des trois premiers termes de la suite. L'espace des suites vérifiant une telle relation de récurrence est *de facto* de dimension 3.

Exemples 24.1.12

1. Dimension de $\mathbb{K}^n$.
2. Dimension de $\mathbb{K}_n[X]$.
3. Dimension de l'ensemble des suites vérifiant une récurrence linéaire homogène d'ordre k.
4. Dimension de l'ensemble des solutions d'une équation différentielle linéaire homogène d'ordre 1, d'ordre 2 à coefficients constants.
5. Dimension de $\mathbb{C}$ en tant que...
6. Dimension de $\mathcal{M}_{n,m}(\mathbb{R})$, de $\mathcal{M}_n(\mathbb{R})$.

Proposition 24.1.13 (Dimension d'un sous-espace vectoriel)

Soit E un espace vectoriel de dimension finie, et F un sous-espace vectoriel de E. Alors

(i) $\dim(F) \leqslant \dim(E)$,

(ii) l'égalité est obtenue si et seulement si $E = F$.

◁ **Éléments de preuve.**

Compléter une base de F. ▷

Par ailleurs, puisqu'un isomorphisme envoie une base sur une base, on en déduit la propriété suivante importante :

Théorème 24.1.14 (Comparaison des dimensions de deux espaces isomorphes)

Soit E et F deux espaces vectoriels et $f : E \to F$ un isomorphisme. Si l'un des deux espaces E ou F est de dimension finie, alors l'autre aussi, et dans ce cas, $\dim(E) = \dim(F)$.

Nous terminons cette section par deux résultats précisant la dimension de certains espaces vectoriels construits en fonction d'autres :

Proposition 24.1.15 (Dimension d'un produit cartésien)

Soit E et F deux espaces vectoriels sur $\mathbb{K}$. Si E et F sont de dimension finie, l'espace vectoriel produit $E \times F$ est de dimension finie, égale à :

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

Plus précisément, si $(b_1, \dots, b_m)$ et $(c_1, \dots, c_n)$ sont des bases de E et F, une base de $E \times F$ est $((b_1, 0), \dots, (b_m, 0), (0, c_1), \dots, (0, c_n))$.

Proposition 24.1.16 (Dimension de $\mathcal{L}(E, F)$)

Si E et F sont de dimension finie, alors :

$$\dim \mathcal{L}(E, F) = \dim(E) \times \dim(F).$$

◁ **Éléments de preuve.**

On en avait décrit une base dans le chapitre précédent. ▷

I.2 Dimension, liberté et rang

Proposition 24.1.17 (Majoration du cardinal d'une famille libre)

Soit E un espace vectoriel de dimension n. Alors toute famille libre de E est de cardinal au plus n, avec égalité si et seulement si c'est une base.

◁ **Éléments de preuve.**

Compléter en une base ; à quoi correspond le cas d'égalité pour cette complétion ? ▷

En particulier, toute famille de strictement plus de n éléments est liée. Une conséquence en est par exemple l'existence de polynômes annulateurs des endormophismes en dimension finie ainsi que des matrices carrées, les deux résultats étant liés, d'après les résultats démontrés dans la suite de ce chapitre

Proposition 24.1.18 (Existence de polynômes annulateurs et minimaux)

1. *Soit $M \in \mathcal{M}_n(\mathbb{K})$. Alors M admet un polynôme annulateur non nul de degré au plus n^2.*
2. *L'ensemble des polynômes annulateurs de M est un idéal non nul de $\mathbb{K}[X]$, engendré par un polynôme unitaire μ_M, appelé polynôme minimal de M. Le polynôme μ_M est de degré minimal parmi les polynômes annulateurs non nuls de M.*
3. *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, et $u \in \mathcal{L}(E)$. Alors u admet un polynôme annulateur non nul de degré au plus n^2.*
4. *L'ensemble des polynômes annulateurs de u est un idéal non nul de $\mathbb{K}[X]$, engendré par un polynôme unitaire μ_u, appelé polynôme minimal de u. Le polynôme μ_u est de degré minimal parmi les polynômes annulateurs non nuls de u.*

Proposition 24.1.19 (Minoration du cardinal d'une famille génératrice)

Soit E un espace vectoriel de dimension n. Alors toute famille génératrice de E est de cardinal au moins n, avec égalité si et seulement si c'est une base.

◁ **Éléments de preuve.**

Extraire une base ; à quoi correspond le cas d'égalité pour cette extraction ? ▷

On en déduit en particulier :

Corollaire 24.1.20 (Caractérisation par le cardinal des familles libres maximales)

Soit E un espace vectoriel de dimension n.

- *Une famille libre est maximale dans le sens donné plus haut si et seulement si elle est de cardinal n.*
- *Une famille génératrice est minimale si et seulement si elle est de cardinal n.*

◁ **Éléments de preuve.**

- Une famille libre est de cardinal au plus n. Si elle est de cardinal n, la majoration précédente affirme qu'on ne peut plus lui ajouter de vecteur sans perdre la liberté. Si elle est de cardinal strictement plus petit, elle ne peut pas être génératrice, donc n'est pas maximale.
- Raisonnement similaire pour les familles génératrices minimales.

▷

Définition 24.1.21 (Rang d'une famille de vecteurs)

Soit E un espace vectoriel, $k \in \mathbb{N}^*$, et $(x_1, \cdots, x_k)$ une famille de vecteurs de E. Le *rang* de la famille $(x_1, \cdots, x_k)$ est la dimension de $\mathrm{Vect}(x_1, \cdots, x_k)$ (cet espace est de dimension finie, puisque engendré par une famille finie). On note :

$$\mathrm{rg}(x_1, \cdots, x_k) = \dim \mathrm{Vect}(x_1, \cdots, x_k).$$

Proposition 24.1.22 (Majoration du rang et cas d'égalité)

$\mathrm{rg}(x_1, \cdots, x_k) \leqslant k$, *avec égalité si et seulement si la famille* $(x_1, \cdots, x_k)$ *est libre.*

◁ **Éléments de preuve.**

On peut extraire une base de la famille génératrice $(x_1, \dots, x_k)$ de $\mathrm{Vect}(x_1, \dots, x_k)$. Le cas d'égalité est obtenu ssi cette base a même cardinal que la famille initiale, donc est égale à la famille initiale. ▷

Proposition 24.1.23 (Conservation du rang par un isomorphisme)

Soit E et F deux $\mathbb{K}$-espaces vectoriels, et $\varphi : E \to F$ un isomorphisme de E dans F. Soit $(x_1, \dots, x_k)$ une famille de E. Alors

$$\mathrm{rg}(x_1, \dots, x_k) = \mathrm{rg}(\varphi(x_1), \dots, \varphi(x_k)).$$

◁ **Éléments de preuve.**

Justifier que φ induit un isomorphisme de $\mathrm{Vect}(e_1, \dots, e_k)$ sur $\mathrm{Vect}(\varphi(e_1), \dots, \varphi(e_k))$. ▷

I.3 Dimension d'une somme

Étant donnés deux sous-espaces vectoriels de E, la dimension de leur somme sera facile à déterminer s'il n'y a pas de redondance (autrement dit, si on a unicité des écritures), ce qui se traduit par le fait que la somme est directe. Cette absence de redondances correspond aussi intuitivement au cas où la dimension de la somme est maximale par rapport aux sommes des deux sous-espaces, ce que l'on justifiera rigoureusement plus tard.

Théorème 24.1.24 (Dimension d'une somme directe)

Soit E un espace vectoriel, et F et G des sous-espaces vectoriels de dimension finie de E, en somme directe. Alors $F \oplus G$ est de dimension finie, et :

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

◁ **Éléments de preuve.**

Juxtaposer deux bases. ▷

Le principe de la démonstration est aussi important que le résultat lui-même, car il donne une façon de construire une base de la somme directe à partir d'une base de chaque membre, par concaténation. Une base ainsi construite est dite *adaptée* à la décomposition en somme directe :

Définition 24.1.25 (Base adaptée à une somme directe)

Soit F et G deux sous-espaces supplémentaires de E, de dimensions respectives k et $n - k$. Une base $(b_1, \dots, b_n)$ de E adaptée à la décomposition $E = F \oplus G$ est une base telle que :

- $(b_1, \dots, b_k)$ est une base de F ;
- $(b_{k+1}, \dots, b_n)$ est une base de G.

On obtient facilement, par récurrence :

Corollaire 24.1.26 (Dimension d'une somme directe de n espaces)
Soit E un espace vectoriel, et $(E_1, \cdots, E_n)$ une famille de sous-espaces vectoriels de dimension finie de E, en somme directe. Alors $\bigoplus_{i=1}^{n} E_i$ est de dimension finie, et :

$$\dim \bigoplus_{i=1}^{n} E_i = \sum_{i=1}^{n} \dim E_i.$$

De même, une base adaptée à cette décomposition en somme directe sera une base obtenue par concaténation de bases des E_i.
On en déduit en particulier la dimension des supplémentaires, après avoir justifié leur existence de façon plus élémentaire qu'en début de chapitre, dans le cadre de la dimension finie :

Théorème 24.1.27 (Existence et dimension d'un supplémentaire, en dimension finie)
Soit E un espace vectoriel de dimension finie, et F un sous-espace vectoriel de E. Alors il existe un supplémentaire S de F dans E, et :

$$\dim S = \dim E - \dim F.$$

◁ **Éléments de preuve.**
Pour l'existence de S, considérer l'espace engendré par les vecteurs ajoutés pour compléter une base de F en une base de E. ▷

Pour terminer, nous donnons une formule générale exprimant la dimension d'une somme quelconque :

Théorème 24.1.28 (Formule de Grassmann)
Soit E un espace vectoriel, et F et G deux sous-espaces de dimension finie de E. Alors :

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

◁ **Éléments de preuve.**
Considérer un supplémentaire S de $F \cap G$ dans G. Justifier que c'est aussi un supplémentaire de F dans $F+G$. Que déduit-on de ces deux propriétés sur les dimensions ? ▷

Théorème 24.1.29 (Majoration de la dimension d'une somme)
Soit E un espace vectoriel, et $E_1, \dots, E_n$ des sous-espaces vectoriels de dimension finie de E. Alors

$$\dim(E_1 + \cdots + E_n) \leqslant \dim(E_1) + \cdots + \dim(E_n),$$

avec égalité si et seulement si la somme est directe.

◁ **Éléments de preuve.**
Le cas $n=1$ est trivial, le cas $n=2$ provient de la formule de Grassmann. Continuer par récurrence. ▷

Comme évoqué plus haut, ces deux derniers résultats affirment que moins il y a de redondances dans l'écriture d'une somme, plus la dimension de la somme est importante ; elle est maximale lorsqu'il n'y a aucune redondance, ce qui s'exprime par le fait que la somme est directe.

Remarque 24.1.30

- Comparez cette formule à $|A \cup B| = |A| + |B| - |A \cap B|$
- Peut-on généraliser, en trouvant pour les dimensions une formule analogue à la formule du crible de Poincaré :

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{j=1}^{n} (-1)^{j+1} \sum_{\substack{J \subset [\![1,n]\!] \\ |J|=j}} \left|\bigcap_{i \in J} A_i\right| ?$$

Proposition 24.1.31 (Caractérisation des couples de sous-espaces supplémentaires)

Soit E un espace de dimension finie, et F et G deux sous-espaces vectoriels de E. Alors F et G sont supplémentaires l'un de l'autre si et seulement si

$$F \cap G = \{0\} \qquad \text{et} \qquad \dim F + \dim G = \dim E.$$

◁ **Éléments de preuve.**

Le sens direct est évident. Le sens réciproque s'obtient en remarquant qu'alors, $\dim(F+G) = \dim E$. ▷

Note Historique 24.1.32

Herrmann Günther Grassmann (1809-1877) est le précurseur incompris de toute la théorie de l'algèbre linéaire. Il expose les fondements de cette théorie dans sa thèse *Théorie des flots et des marées*, thèse qui ne sera pas lue par son examinateur, et qui sera publiée uniquement au début du XX-ième siècle. Ce n'est qu'une vingtaine d'année plus tard (vers 1860) que certains mathématiciens se rendent comptent de l'importance des travaux de Grassmann. Entre temps, Grassmann s'est contenté d'un poste d'enseignant en lycée. Détourné de la recherche mathématique, c'est dans des études linguistiques (sanscrit, gotique) qu'il finit par trouver la consécration.

II Applications linéaires en dimension finie

II.1 Rang d'une application linéaire

Définition 24.2.1 (Rang d'une application linéaire)

Soit $u \in \mathcal{L}(E,F)$ une application linéaire. Si $\mathrm{Im}(u)$ est de dimension finie, on définit le rang de u par :

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

Proposition 24.2.2

Soit $(x_i)_{i\in I}$ une famille génératrice de E. Le rang de u, s'il existe, est égal au rang de la famille $(u(x_i))_{i\in I}$.

◁ **Éléments de preuve.**

Où a-t-on une famille génératrice de l'image ? ▷

Corollaire 24.2.3 (Existence du rang en dimension finie)

Si E ou F est de dimension finie et $u \in \mathcal{L}(E,F)$, alors $\mathrm{rg}(u)$ est fini et

$$\mathrm{rg}(u) \leqslant \min(\dim(E), \dim(F)).$$

II.2 Théorème du rang

Plus l'image est grosse, plus les redondances sont petites (et donc plus le noyau est petit). C'est ce qu'énoncé le théorème du rang : tout ce qui se gagne à l'image se perd dans le noyau, et inversement.
Le cas extrême (qui est aussi la clé de la preuve du théorème du rang) est le cas d'une fonction injective : dans ce cas, la dimension de l'image est égale à celle de l'espace source. C'est ce que nous avons déjà constaté en montrant que deux espaces de dimension finie isomorphes ont même dimension.
Afin de démontrer le théorème du rang, nous restreindrons l'application linéaire étudiée de sorte à nous ramener à la situation ci-dessus. Pour cela, nous décrivons le noyau et l'image d'une restriction.

Lemme 24.2.4 (Noyau et image d'une restriction)

Soit $u \in \mathcal{L}(E,F)$, *et* E' *un sous-espace de* E. *Soit* $v \in \mathcal{L}(E',F)$ *la restriction de* u *à* E'. *Alors :*

- $\mathrm{Ker}(v) = \mathrm{Ker}(u) \cap E'$
- *Si* $\mathrm{Ker}(u) + E' = E$, $\mathrm{Im}(v) = \mathrm{Im}(u)$.

◁ **Éléments de preuve.**
Vérification facile. ▷

Proposition 24.2.5 (Restriction de u à un supplémentaire de Ker(u))

Soit S *un supplémentaire de* $\mathrm{Ker}(u)$ *dans* E. *Alors* u *induit un isomorphisme de* S *sur* $\mathrm{Im}(u)$.

◁ **Éléments de preuve.**
Obtenir l'injectivité et la surjectivité par étude du noyau et de l'image, à l'aide du lemme précédent. ▷

Remarquez que ce dernier résultat ne nécessite pas d'hypothèse de finitude. Cependant, si E n'est pas de dimension finie, il est nécessaire de supposer l'axiome du choix, afin d'assurer l'existence du supplémentaire S.
On déduit du corollaire précédent le très important :

Théorème 24.2.6 (Théorème du rang)

Soit E *un espace vectoriel de dimension finie, et* F *un espace vectoriel quelconque. Soit* $f \in \mathcal{L}(E,F)$. *Alors :*

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

◁ **Éléments de preuve.**
Dans la propriété précédente, comparer la dimension de S et la dimension de $\mathrm{Im}(u)$. Par ailleurs, trouver une relation entre la dimension de S, celle de E et celle de $\mathrm{Ker}(u)$. ▷

On en déduit en particulier :

Corollaire 24.2.7 (Effet d'une application linéaire sur la dimension)

Une application linéaire diminue toujours la dimension : si $u \in \mathcal{L}(E,F)$ *et* $E' \subset E$, *alors*

$$\dim(u(E')) \leqslant \dim(E').$$

II.3 Caractérisation de l'injectivité et la surjectivité par le rang

Proposition 24.2.8 (Caractérisation de l'injectivité, de la surjectivité)

Soit $u \in \mathcal{L}(E,F)$. *Comme démontré plus haut,* $\mathrm{rg}(u) \leqslant \dim(E)$ *et* $\mathrm{rg}(u) \leqslant \dim(F)$. *les cas d'égalité caractérisent l'injectivité et la surjectivité :*

- *Si* $\dim(E)$ *est finie,* $\mathrm{rg}(u) = \dim(E)$ *si et seulement si* u *est injective ;*
- *Si* $\dim(F)$ *est finie,* $\mathrm{rg}(u) = \dim(F)$ *si et seulement si* u *est surjective.*

◁ **Éléments de preuve.**

Le premier point découle du théorème du rang. Le second découle du cas d'égalité dans l'inégalité des dimensions pour les sous-espaces. ▷

Théorème 24.2.9 (Caractérisation des isomorphismes en dimension finie)

Soit E *et* F *deux espaces vectoriels de* ***même dimension*** *finie* n, *et soit* $f \in \mathcal{L}(E,F)$. *Alors les propositions suivantes sont équivalentes :*

(i) f *est un isomorphisme ;*

(ii) $\mathrm{rg}(f) = n$ *;*

(iii) f *est injective ;*

(iv) f *est surjective.*

◁ **Éléments de preuve.**

On peut le déduire de la proposition précédente. Ou alors refaire le tour en se servant du théorème du rang. ▷

En particulier, si E est de dimension finie, un endomorphisme $f \in \mathcal{L}(E)$ est un automorphisme si et seulement si f est injective si et seulement si f est surjective. Comme l'injectivité et la surjectivité caractérisent l'inversibilité à gauche et à droite, on obtient :

Proposition 24.2.10 (Caractérisation des automorphismes en dimension finie)

Soit E *un espace de dimension finie. Soit* $u \in \mathcal{L}(E)$. *Alors* u *est un automorphisme si et seulement s'il est inversible à droite ou à gauche.*

◁ **Éléments de preuve.**

L'injectivité équivaut à l'inversibilité à gauche, la surjectivité à l'inversibilité à droite. Et ceci sans axiome du choix (pourquoi ?) ▷

II.4 Rang d'une composée

Théorème 24.2.11 (Effet d'une composition sur le rang)

Soit E *un espace vectoriel de dimension finie, et soit* $u \in \mathcal{L}(E,F)$ *et* $v \in \mathcal{L}(F,G)$.

1. $\mathrm{rg}(v \circ u) \leqslant \min(\mathrm{rg}(u), \mathrm{rg}(v))$.
2. *Si* v *est injective,* $\mathrm{rg}(v \circ u) = \mathrm{rg}(u)$.
3. *Si* u *est surjective,* $\mathrm{rg}(v \circ u) = \mathrm{rg}(v)$

◁ **Éléments de preuve.**

Remarquer que $\mathrm{Im}(v \circ u) = \mathrm{Im}\left(v_{|\mathrm{Im}(u)}^{|\mathrm{Im}(v)}\right)$, et utiliser la proposition 24.2.8. ▷

En particulier :

Corollaire 24.2.12 (Invariance du rang par composition par un isomophisme)

Tout est dit dans le titre, la composition pouvant se faire à droite ou à gauche.

III Applications linéaires et matrices

Soit E et F deux espaces de dimensions finies p et n respectiviement, et $\mathcal{B}$ une base de E, $\mathcal{C}$ une base de F. Toute l'étude de cette section particulièrement importante est basée sur l'observation suivante : une application linéaire $f \in \mathcal{L}(E,F)$ est entièrement déterminée par les images $f(b_i)$ des p vecteurs de la base $\mathcal{B}$, chacune de ces images pouvant être décrite par la donnée de n scalaires (les coordonnées dans la base $\mathcal{C}$). Ainsi, f est complètement décrite par une famille de $p \times n$ scalaires, qu'on peut ranger dans une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$, dont la i-ième colonne représente les coordonnées dans $\mathcal{C}$ de l'image $f(b_i)$ du i-ième vecteur de la base $\mathcal{B}$.
Après choix de bases, une application linéaire se décrit par une matrice, et de nombreuses constructions vectorielles peuvent se décrire matriciellement.

III.1 Matrice d'une application linéaire

Soit E et F deux espaces vectoriels de dimensions finies respectives p et n.

Définition 24.3.1 (Coordonnées d'un vecteur)

Soit $\mathcal{C} = (c_1, \cdots, c_n)$ une base de F, et $x \in F$. Alors il existe d'uniques scalaires $x_1, \cdots, x_n$ tels que $x = x_1c_1 + \cdots + x_nc_n$. On définit la matrice colonne des coordonnées de x dans la base $\mathcal{C}$ par :

$$[x]_{\mathcal{C}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Réciproquement, étant donné $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$, il existe un unique vecteur x tel que $X = [x]_{\mathcal{C}}$. Nous noterons par commodité :

$$x = \overrightarrow{v_{\mathcal{C}}}(X)$$

Les notations introduites dans cette définition sont personnelles.

Définition 24.3.2 (Matrice associée à une famille)

Sous les mêmes hypothèses, soit $(x_1, \cdots, x_k)$ une famille de vecteurs de F. Alors la matrice de cette famille dans la base $\mathcal{C}$ est :

$$[x_1, \cdots, x_k]_{\mathcal{C}} = \left(\ [x_1]_{\mathcal{C}} \middle| \cdots \middle| [x_k]_{\mathcal{C}} \ \right).$$

Ainsi, il s'agit de la matrice dont la i-ème colonne comporte les coordonnées dans la base $\mathcal{C}$ du vecteur x_i.

Définition 24.3.3 (Matrice d'une AL relativement à des bases)

Soit $f \in \mathcal{L}(E,F)$, et soit $\mathcal{B} = (b_1, \cdots, b_p)$ une base de E et $\mathcal{C} = (c_1, \cdots, c_n)$ une base de F. Alors la matrice de f relativement aux bases $\mathcal{B}$ et $\mathcal{C}$ est la matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ de la famille $(f(b_1), \cdots, f(b_p))$ dans la base $\mathcal{C}$:

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f) = [f(b_1), \cdots, f(b_p)]_{\mathcal{C}} = \left(\ [f(b_1)]_{\mathcal{C}} \middle| \cdots \middle| [f(b_p)]_{\mathcal{C}} \ \right).$$

Ainsi, il s'agit de la matrice de type (n,p) dont la i-ème colonne donne les coordonnées du vecteur $f(b_i)$ dans la base $\mathcal{C}$.

Réciproquement, étant données des bases $\mathcal{B}$ et $\mathcal{C}$ de E et F, et M une matrice de taille adaptée, la propriété de rigidité des applications linéaires nous assure qu'il existe une unique application linéaire envoyant pour tout i le i-ième vecteur de la base $\mathcal{B}$ sur le vecteur dont les coordonnées dans $\mathcal{C}$ sont données par la i-ème colonne de M. En d'autres termes, il existe une unique application linéaire de $\mathcal{L}(E,F)$ dont la matrice relativement aux bases $\mathcal{B}$ et $\mathcal{C}$ est M. On énonce :

Proposition 24.3.4 (Correspondance AL-matrice)

Soit E un espace vectoriel de dimension p et F un espace vectoriel de dimension n. L'application $\mathrm{Mat}_{\mathcal{B},\mathcal{C}} : \mathcal{L}(E,F) \to \mathcal{M}_{n,p}(\mathbb{K})$ *est un isomorphisme. Sa réciproque est l'application linéaire qui à* $M = (C_1|\dots|C_p)$ *associe* $u \in \mathcal{L}(E,F)$ *telle que* $u(b_i) = \overrightarrow{v_{\mathcal{C}}}(C_i)$.

◁ **Éléments de preuve.**
La linéarité est évidente. Par construction, les deux applications décrites sont bien réciproques l'une de l'autre. Qu'est-ce qui justifie la bonne définition de la réciproque ? On peut aussi étudier la bijectivité en étudiant le noyau et en comparant les dimensions. ▷

Remarque 24.3.5

Si on prend $E = \mathbb{K}^p$, $F = \mathbb{K}^n$, et pour $\mathcal{B}$ et $\mathcal{C}$ les bases canoniques de ces espaces, on retrouve la matrice $\mathrm{Mat}_{b.c.}(f)$. On dit dans ce cas que $\mathrm{Mat}_{b.c.}(f)$ est la *matrice canoniquement associée* à $f \in \mathcal{L}(\mathbb{K}^p,\mathbb{K}^n)$. Réciproquement, étant donné une matrice $M \in \mathcal{M}_{n,p}(\mathbb{K})$, l'unique application linéaire $f \in \mathcal{L}(\mathbb{K}^p,\mathbb{K}^n)$ dont la matrice relativement aux bases canoniques est M est appelée *application linéaire canoniquement associée* à la matrice M.

Exemple 24.3.6

1. Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x,y) = 2x + 3y$. Déterminer $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ lorsque :
   - $\mathcal{B} = b.c.$, $\mathcal{C} = (1)$
   - $\mathcal{B} = \left(\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}\right)$, $\mathcal{C} = (2)$
2. Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ définie par $f(x,y) = (2x - y, x + 2y)$.
   - $\mathrm{Mat}_{b.c.}(f)$?
   - $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ lorsque $\mathcal{B} = \mathcal{C} = \left(\begin{pmatrix}1\\-1\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}\right)$
3. Matrice de l'opérateur de dérivation $D \in \mathcal{L}(\mathbb{R}_n[X])$ relativement à la base canonique (au départ et à l'arrivée)
4. Trouver une base $\mathcal{B}$ de $\mathbb{R}_n[X]$ telle que
$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(D) = J_{n+1} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}.$$
   La matrice J_{n+1} est appelée matrice de Jordan d'ordre $n+1$.
5. Plus généralement, soit E un espace de dimension n, et u un endomorphisme nilpotent, d'indice de nilpotence égal à n. Montrer qu'il existe une base relativement à laquelle la matrice de u est J_n.

Proposition 24.3.7 (Compatibilité du produit matriciel avec l'évaluation)
Soit $f \in \mathcal{L}(E,F)$, et $X \in E$. Soit $\mathcal{B}$ une base de E et $\mathcal{C}$ une base de F. On a alors

$$[f(X)]_{\mathcal{C}} = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f) \times [X]_{\mathcal{B}}.$$

◁ **Éléments de preuve.**
Décomposer X dans la base $\mathcal{B}$, appliquer f, et prendre les coordonnées dans $\mathcal{C}$. On obtient un combinaison linéaire des " colonnes de $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$, nous ramenant à la description du produit matriciel par colonnes. ▷

Proposition 24.3.8 (Caractérisation de la matrice d'une AL dans des bases)
Soit $f \in \mathcal{L}(E,F)$. Soit $p = \dim(E)$, $n = \dim(F)$. Soit $\mathcal{B}$ une base de E et $\mathcal{C}$ une base de F. Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors les deux propositions suivantes sont équivalentes :

(i) $M = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$

(ii) *pour tout* $X \in E$, $[f(X)]_{\mathcal{C}} = M \times [X]_{\mathcal{B}}$

◁ **Éléments de preuve.**
Le sens direct résulte de la proposition précédente. Le sens réciproque peut s'obtenir en évaluant (ii) en des vecteurs X bien choisis. ▷

La formule suivante est d'une importance capitale, et a été historiquement la motivation principale (sous une forme légèrement différente) de la définition du produit matriciel :

Théorème 24.3.9 (Matrice associée à une composition)
Soit E, F et G trois espaces vectoriels de dimension finie, munis respectivement des bases $\mathcal{B}$, $\mathcal{C}$ et $\mathcal{D}$. Soit $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$. Alors :

$$\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(g \circ f) = \mathrm{Mat}_{\mathcal{C},\mathcal{D}}(g) \cdot \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f).$$

◁ **Éléments de preuve.**
Appliquer la propriété précédente pour le calcul des coordonnées dans la base $\mathcal{D}$ d'un vecteur $g \circ f(X)$, en procédant de deux façons. ▷

Puisque les bases varient, on ne peut pas évoquer à proprement parler de propriété de morphisme. En revanche, dans $\mathcal{L}(E)$, en se fixant une base $\mathcal{B}$ unique, et en considérant l'application $\mathcal{M}_{\mathcal{B}} = \mathcal{M}_{\mathcal{B},\mathcal{B}}$, la propriété ci-dessus dit que $\mathrm{Mat}_{\mathcal{B}}$ (dont on connaît aussi la bijectivité) est un isomorphisme de monoïde de $(\mathcal{L}(E), \circ)$ à $(\mathcal{M}_n(\mathbb{K}), \times)$. Comme elle respecte aussi la somme (et même les combinaisons), il s'agit en fait d'un isomorphisme d'anneaux (et même d'algèbre).

III.2 Image, noyau et rang d'une matrice

Soit $(n,p) \in (\mathbb{N}^*)^2$.

Définition 24.3.10 (AL canoniquement associée à une matrice)
Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$. L'application linéaire canoniquement associée à M est l'application $f : \mathbb{K}^p \to \mathbb{K}^n$ définie par $f(X) = MX$, les vecteurs de $\mathbb{K}^p$ et $\mathbb{K}^n$ étant identifiés à des matrices-colonnes.

Proposition 24.3.11

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ *et soit* $\mathcal{B}$ *et* $\mathcal{C}$ *les bases canoniques de* $\mathbb{K}^p$ *et* $\mathbb{K}^n$ *respectivement. Alors l'application linéaire* f *canoniquement associée à* M *est l'unique application* $f \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ *telle que* $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f) = M$.

◁ **Éléments de preuve.**
Cela provient de la caractérisation de la matrice d'une AL, et le fait que le vecteur X est ici identifié au vecteur colonne de ses coordonnées dans la base canonique. ▷

Définition 24.3.12 (Image et noyau d'une matrice)

Par définition, l'image et le noyau d'une matrice M sont l'image et le noyau de l'application linéaire canoniquement associée. Plus précisément, soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ et $f : X \mapsto MX$ l'application de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ canoniquement associé. Alors

- $\mathrm{Im}(M) = \mathrm{Im}(f) = \{Y \in \mathbb{K}^n \simeq \mathcal{M}_{n,1}(\mathbb{K}) \mid \exists X \in \mathbb{K}^p,\ MX = Y\}$
- $\mathrm{Ker}(M) = \mathrm{Ker}(f) = \{X \in \mathbb{K}^p \simeq \mathcal{M}_{p,1}(\mathbb{K}) \mid MX = 0\}$.

En particulier, puisque les colonnes de M sont les images par f des éléments de la base canonique, nous obtenons :

Proposition 24.3.13 (Description de l'image d'une matrice)

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ *une matrice de colonnes* $C_1, \dots, C_p$. *Alors*

$$\mathrm{Im}(M) = \mathrm{Vect}(C_1, \dots, C_p).$$

Définition 24.3.14 (Rang d'une matrice)

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ une matrice. Son rang est égal au rang de l'application linéaire canoniquement associée, c'est-à-dire :

$$\mathrm{rg}(M) = \mathrm{rg}(f) = \dim(\mathrm{Im}(f)) = \dim(\mathrm{Im}(M)) = \dim(\mathrm{Vect}(C_1, \dots, C_p)) = \mathrm{rg}(C_1, \dots, C_p),$$

où les C_i sont les colonnes de M.

Lemme 24.3.15

Soit $(x_1, \dots, x_n)$ *une famille de vecteurs de* E, *et* $\mathcal{B}$ *une base de* E. *Alors*

$$\mathrm{rg}(x_1, \dots, x_n) = \mathrm{rg}([x_1]_{\mathcal{B}}, \dots, [x_n]_{\mathcal{B}}) = \mathrm{rg}([x_1, \dots, x_k]_{\mathcal{B}}).$$

◁ **Éléments de preuve.**
Considérer l'application $f : E \longrightarrow \mathbb{K}^n$ définie par $f(x) = [x]_{\mathcal{B}}$ et utiliser le fait que les isomorphismes préservent le rang. ▷

Le rang d'une application linéaire peut alors se relier au rang de sa matrice, quelles que soient les bases choisies :

Proposition 24.3.16 (Lien entre rang d'une AL et rang de sa matrice)

Soit $u \in \mathcal{L}(E, F)$. *Soit* $\mathcal{B}$ *une base de* E *et* F *une base de* F. *Alors*

$$\mathrm{rg}(u) = \mathrm{rg}(\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)).$$

◁ **Éléments de preuve.**
C'est le lemme précédent appliqué à la famille $(u(b_1), \dots, u(b_n))$. ▷

Théorème 24.3.17 (Théorème du rang)
Soit $M = \mathcal{M}_{n,p}(\mathbb{K})$. Alors
$$\dim(\mathrm{Ker}(M)) + \mathrm{rg}(M) = p.$$

Méthode 24.3.18 (Déterminer le noyau d'une matrice)

- On peut bien sûr résoudre le système linéaire $AX = 0$
- On peut aussi remarquer que $A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$ se traduit par la relation $x_1C_1 + \dots x_nC_n = 0$ entre les colonnes de A. Pour trouver un vecteur du noyau, il suffit de trouver une relation entre les colonnes. Pour être sûr d'avoir une description complète du noyau, il faut alors trouver des relations entre les colonnes donnant des vecteurs linéairement indépendants en nombre égal à la dimension de $\mathrm{Ker}(f)$. Cette dimension peut souvent se calculer au préalable grâce au théorème du rang. Ce type de calcul est particulièrement efficace pour les petites matrices (nombre de lignes et/ou colonnes inférieur à 3 ou 4), car dans ce contexte, le rang se calcule facilement, parfois de visu, également en se servant de relations entre les colonnes.
- Il peut être avantageux de penser à cette méthode lorsqu'on cherche à résoudre un système linéaire.

III.3 Méthodes matricielles pour l'étude d'une famille

Soit E un espace vectoriel de dimension finie n, muni d'une base $\mathcal{B}$ L'application $\mathrm{Mat}_{\mathcal{B}}$ étant un isomorphisme de l'anneau $\mathcal{L}(E)$ sur $\mathrm{Mat}_n(\mathbb{K})$, elle préserve l'inversibilité, dans un sens comme dans l'autre. Ainsi :

Proposition 24.3.19
Soit $u \in \mathcal{L}(E)$. Alors $\mathrm{Mat}_{\mathcal{B}}(u)$ est inversible si et seulement si $u \in \mathrm{GL}(E)$.

En particulier, une matrice carrée M est inversible ssi l'endomorphisme $u \in \mathcal{L}(E)$ canoniquement associé à M est un automophisme.
On peut alors enfin démontrer une propriété qu'on avait jusqu'ici laissé en suspens :

Proposition 24.3.20 (Caractérisation de l'inversibilité)
Soit $M \in \mathcal{M}_n\mathbb{K})$. Alors M est inversible si et seulement si elle est inversible à droite ou à gauche.

◁ **Éléments de preuve.**
L'hypothèse d'inversibilité à droite ou gauche est équivalente à l'inversibilité à droite ou gauche de l'endomorphisme canoniquement associé, et on termine par la caractérisation des automorphismes en dimension finie. ▷

Proposition 24.3.21 (Caractérisation matricielle des bases)
Soit E un espace vectoriel muni d'une base $\mathcal{B}$, et $(x_1, \dots, x_n)$ une famille de E. Alors $(x_1, \dots, x_n)$ est une base ssi $[x_1, \dots, x_n]_{\mathcal{B}}$ est inversible.

◁ **Éléments de preuve.**
Cette matrice est la matrice de l'application linéaire qui envoie les vecteurs de $\mathcal{B}$ sur les x_i. La caractérisation des automorphismes par l'image d'une base permet de conclure. ▷

Définition 24.3.22 (Matrice échelonnée en colonnes)
Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$. On dit que M est échelonnée en colonnes si $M^\top$ est échelonnée (en lignes).

Définition 24.3.23 (Famille échelonnée en coordonnées)
Soit E un espace vectoriel de dimension n, muni d'une base $\mathcal{B} = (b_1, \dots, b_n)$, et soit $(x_1, \dots, x_k)$ une famille de E. On note $\overline{\mathcal{B}} = (b_n, \dots, b_1)$ obtenue en inversant l'ordre des élements de la base $\mathcal{B}$. On dit que :

- $(x_1, \dots, x_k)$ est échelonnée en coordonnée (par le haut) sur $\mathcal{B}$ si $[x_1, \dots, x_k]_{\mathcal{B}}$ est une matrice échelonnée en colonnes
- $(x_1, \dots, x_k)$ est totalement échelonnée en coordonnée (par le haut) si $k = n$ et sur $\mathcal{B}$ si $[x_1, \dots, x_k]_{\mathcal{B}}$ est une matrice triangulaire inférieure à coefficients diagonaux non nuls.
- $(x_1, \dots, x_k)$ est échelonnée en coordonnée (par le bas) sur $\mathcal{B}$ si $(x_1, \dots, x_k)$ est échelonnée en coordonnée (par le haut) sur $\overline{\mathcal{B}}$
- $(x_1, \dots, x_k)$ est totalement échelonnée en coordonnée (par le bas) sur $\mathcal{B}$ si $(x_1, \dots, x_k)$ est totalement échelonnée en coordonnée (par le haut) sur $\overline{\mathcal{B}}$

Proposition 24.3.24 (Liberté des familles échelonnées)
Avec les notations précédentes :

1. *Si $(x_1, \dots, x_n)$ est totalement échelonnée (par le bas ou le haut), alors c'est une base.*
2. *Si $(x_1, \dots, x_k)$ est échelonnée (par le bas ou le haut), et si les x_i sont tous non nuls alors c'est une famille libre.*

◁ **Éléments de preuve.**
Dans le premier cas, la matrice de la famille dans la base $\mathcal{B}$ est triangulaire à coefficients diagonaux non nuls, donc inversible. Le deuxième cas peut se ramener au premier par ajout de certains vecteurs de $\mathcal{B}$. ▷

Remarque 24.3.25
À inversion près des vecteurs de la famille, les familles (finies) de polynôme totalement échelonnées en degré correspondent à la définition générale donnée ici de familles totalement échelonnées par le bas dans $\mathbb{R}_n[X]$, relativement à la base canonique.

Corollaire 24.3.26 (Rang des matrices échelonnées)

1. *Le rang d'une matrice échelonnée en colonnes est le nombre de ses colonnes non nulles*
2. *Le rang d'une matrice échelonnée en lignes est le nombre de ses lignes non nulles*
3. *En particulier, si M est échelonnée (en lignes ou colonnes), alors $\mathrm{rg}(M) = \mathrm{rg}(M^\top)$.*

◁ **Éléments de preuve.**

1. Conséquence directe de la description de l'image d'une matrice et de la liberté d'une famille échelonnée.

2. Si k est le nombre de lignes non nulles de M, la description de l'image de M dit que l'image est incluse dans l'espace engendré par les k premiers vecteurs de la base. Il n'est pas dur ensuite de sortir une famille échelonnée de cardinal k de la famille des colonnes.

▷

III.4 Calcul du rang

Une propriété importante, permettant le calcul du rang par la méthode du pivot, est la conservation du rang par multiplication par une matrice inversible, donc en particulier par les matrices de codage des opérations. Ainsi, le pivot de Gauss conserve le rang.

Théorème 24.3.27 (Conservation de l'image et du noyau)
Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ *une matrice quelconque, et* $P \in \mathrm{GL}_n(\mathbb{K})$, $Q \in \mathrm{GL}_p(\mathbb{K})$. *Alors :*

(i) $\mathrm{Ker}(PM) = \mathrm{Ker}(M)$

(ii) $\mathrm{Im}(MQ) = \mathrm{Im}(M)$.

Ainsi, la multiplication à gauche par une matrice inversible conserve le noyau et la multiplication à droite conserve l'image.

◁ **Éléments de preuve.**
Immédiat par équivalences. ▷

Corollaire 24.3.28 (Conservation du rang)
Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ *une matrice quelconque, et* $P \in \mathrm{GL}_n(\mathbb{K})$, $Q \in \mathrm{GL}_p(\mathbb{K})$. *Alors :*

$$\mathrm{rg}(PMQ) = \mathrm{rg}(M)$$

◁ **Éléments de preuve.**
On y a répondu lors de la démontration précédente. ▷

Corollaire 24.3.29 (Conservation de l'image et du noyau par opérations élémentaires)

(i) Les opérations élémentaires sur les lignes conservent le noyau.

(ii) Les opérations élémentaires sur les colonnes conservent l'image.

(iii) Les opérations élémentaires sur les lignes et les colonnes conservent le rang.

Méthode 24.3.30 (Calcul du rang d'une matrice)

- Effectuer un pivot pour se ramener à une matrice échelonnée.
- Le rang est égal au rang de la matrice échelonnée, donc au nombre de ses lignes non nulles.
- Si les opérations sur les colonnes semblent plus simples, c'est possible aussi.
- On peut même combiner les deux.

Méthode 24.3.31 (Trouver une base de l'image)

- Cette fois, on veut conservation de l'image, on travaille donc sur les colonnes.
- On effectue un pivot sur les colonnes jusqu'à obtenir une matrice échelonnée sur les colonnes, dont les colonnes engendrent l'image.
- Les colonnes non nulles de cette matrice forment une base de l'image.

Méthode 24.3.32 (Trouver un supplémentaire)

- On suppose donné un sous-espace E de $\mathbb{K}^n$ dont on connaît une famille génératrice de p vecteurs.
- On écrit la matrice de cette famille (chaque vecteur est écrit en colonne) : il s'agit de la matrice de l'application linéaire qui envoie les vecteurs de la base canonique de $\mathbb{K}^p$ sur les vecteurs de la famille génératrice. L'image de cette application est E.
- On effectue un pivot en travaillant sur les colonnes, pour conserver l'image.
- À la matrice échelonnée sur les colonnes obtenue au bout, on ajoute les colonnes correspondant aux vecteurs de la base canonique de sorte à obtenir une matrice triangulaire supérieure.
- Les vecteurs ainsi ajoutés forment une base d'un supplémentaire de E dans $\mathbb{K}^n$.

On généralise le résultat déjà vu, comparant le rang d'une matrice échelonnée à celui de sa transposée :

Théorème 24.3.33 (Rang d'une transposée)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. On a :

$$\operatorname{rg}(A) = \operatorname{rg}(A^\top).$$

◁ **Éléments de preuve.**

Effectuer sur la matrice $A^\top$ les opérations sur les colonnes correspondant aux opérations sur les lignes effectuées sur A pour se ramener à une matrice échelonnée M par la méthode du pivot. ▷

Proposition 24.3.34 (Dimension de l'ensemble des solutions d'un système linéaire)

Soit $AX = B$ un système linéaire, de matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ de rang r.

1. *Le système $AX = B$ admet au moins une solution ssi $B \in \operatorname{Im}(A)$.*
2. *Dans ce cas, l'ensemble des solutions est un espace affine dirigé par* $\operatorname{Ker}(A)$.
3. *De plus* $\dim(\operatorname{Ker}(A)) = p - r$
4. *Si A est une matrice carrée, alors le système admet une unique solution si et seulement si A est inversible (système de Cramer)*

III.5 Caractérisation du rang par les matrices extraites

Définition 24.3.35 (Matrice extraite)

Soit $A = (a_{i,j})$ une matrice de $\mathcal{M}_{n,p}$. Une matrice B est extraite de A si et seulement s'il existe $1 \leqslant i_1 < \dots < i_q \leqslant n$ et $1 \leqslant j_1 < \dots < j_r \leqslant p$ tels que $B = (a_{i_k,j_\ell})_{(k,\ell)\in[\![1,q]\!]\times[\![1,r]\!]}$.

Ainsi, la matrice B est obtenue en ne conservant de A que les lignes d'indice $i_1, \dots, i_q$ et les colonnes d'indice $j_1, \dots, j_r$.

Proposition 24.3.36 (Rang d'une matrice extraite)

Soit B une matrice extraite de A. Alors $\operatorname{rg}(B) \leqslant \operatorname{rg}(A)$.

◁ **Éléments de preuve.**

L'extraction de colonnes baisse le rang. Transposer et recommencer. ▷

Théorème 24.3.37 (Caractérisation du rang par les matrices carrées extraites)
Le rang de A est l'ordre maximal d'une matrice carrée inversible extraite de A.

◁ **Éléments de preuve.**
Avec ce qui précède, il suffit de montrer qu'il existe une matrice inversible extraite d'ordre $r = \mathrm{rg}(A)$. Extraire d'abord r colonnes, transposer et recommencer. ▷

IV Changements de base

La correspondance entre applications linéaires et matrice dépend du choix de bases de l'espace de départ E et de l'espace d'arrivée F. Le but de ce paragraphe est d'étudier l'effet d'un changement de base sur E ou sur F : la matrice relativement aux nouvelles bases se déduit-elle d'une certaine manière de la matrice relativement aux anciennes bases ?

IV.1 Changements de base pour des applications linéaires

Définition 24.4.1 (Matrice de passage)
Soit E un espace vectoriel, et $\mathcal{B}_1$ et $\mathcal{B}_2$ deux bases de E. Alors la matrice de passage de la base $\mathcal{B}_1$ à la base $\mathcal{B}_2$ est la matrice :

$$P_{\mathcal{B}_1}^{\mathcal{B}_2} = \mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(\mathrm{Id}_E) = [\mathcal{B}_2]_{\mathcal{B}_1}.$$

Ainsi, la i-ème colonne de cette matrice est constituée des coordonnées du i-ème vecteur de la base $\mathcal{B}_2$ dans la base $\mathcal{B}_1$. Il s'agit donc de la matrice de la famille des vecteurs de la seconde base $\mathcal{B}_2$ dans la première base $\mathcal{B}_1$.

La formule de composition amène facilement :

Proposition 24.4.2 (Inversibilité des matrices de passage)
Toute matrice de passage $P_{\mathcal{B}_1}^{\mathcal{B}_2}$ est inversible. Son inverse est la matrice de passage $P_{\mathcal{B}_2}^{\mathcal{B}_1}$

Toute matrice inversible peut être interprétée comme une matrice de passage :

Proposition 24.4.3
Soit $P \in \mathrm{GL}_n(\mathbb{K})$ une matrice inversible, et E un espace vectoriel de dimension n.
- *Pour toute base $\mathcal{B}$ de E, il existe une base $\mathcal{C}$ telle que $P = P_{\mathcal{B}}^{\mathcal{C}}$.*
- *Pour toute base $\mathcal{C}$ de E, il existe une base $\mathcal{B}$ telle que $P = P_{\mathcal{B}}^{\mathcal{C}}$.*

◁ **Éléments de preuve.**
Définir les vecteurs de $\mathcal{C}$ par leurs coordonnées dans la base $\mathcal{B}$. Pour obtenir le deuxième point, à quelle matrice appliquer le premier point ? ▷

Puisque $[X]_{\mathcal{B}_1} = \mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(\mathrm{Id})[X]_{\mathcal{B}_2}$, on obtient l'expression de l'effet d'un changement de base sur les coordonnées d'un vecteur :

Proposition 24.4.4 (Effet d'un changement de base sur les coordonnées d'un vecteur)
Soit $X \in E$. Alors $[X]_{\mathcal{B}_1} = P_{\mathcal{B}_1}^{\mathcal{B}_2} \cdot [X]_{\mathcal{B}_2}$.

En composant à gauche et à droite par l'identité, avec bases différentes, on obtient, toujours d'après la formule de composition, la très importante formule de changement de base.

Théorème 24.4.5 (Formule de changement de base)

Soit E un espace vectoriel de dimension finie, muni de deux bases $\mathcal{B}_1$ et $\mathcal{B}_2$, et F un espace vectoriel de dimension finie, muni de deux bases $\mathcal{C}_1$ et $\mathcal{C}_2$. Soit $f \in \mathcal{L}(E,F)$. Alors :

$$\mathrm{Mat}_{\mathcal{B}_2,\mathcal{C}_2}(f) = P_{\mathcal{C}_2}^{\mathcal{C}_1} \cdot \mathrm{Mat}_{\mathcal{B}_1,\mathcal{C}_1}(f) \cdot P_{\mathcal{B}_1}^{\mathcal{B}_2} = (P_{\mathcal{C}_1}^{\mathcal{C}_2})^{-1} \cdot \mathrm{Mat}_{\mathcal{B}_1,\mathcal{C}_1}(f) \cdot P_{\mathcal{B}_1}^{\mathcal{B}_2}$$

Ainsi, cette formule s'écrit $M' = Q^{-1}MP$, où M est la matrice dans les bases initiales, M' la matrice dans les nouvelles bases, P et Q les matrices de passage de la première base vers la seconde, respectivement dans E et dans F.

Exemple 24.4.6

1. Retrouver les matrices des exemples précédents.
2. Donner une relation entre la matrice de D dans la base canonique de $\mathbb{R}_n[X]$ (au départ et à l'arrivée) et la matrice de Jordan J_{n+1}.

IV.2 Matrices équivalentes

On remarque que deux matrices associées à une même application linéaire avec des choix différents de bases, s'obtiennent l'une de l'autre par multiplication à gauche et à droite par des matrices inversibles. Cela motive la définition suivante :

Définition 24.4.7 (Matrices équivalentes)

Soit $(M,N) \in \mathcal{M}_{n,p}(\mathbb{K})$. On dit que M et N sont équivalentes si et seulement s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ et $Q \in \mathrm{GL}_p(\mathbb{K})$ tels que $N = PMQ$.

Proposition 24.4.8

Cela définit une relation d'équivalence.

Ainsi, on obtient :

Théorème 24.4.9

Soit $f \in \mathcal{L}(E,F)$.

(i) Deux matrices M et N représentant f dans des choix différents de bases sont équivalentes ;

(ii) Réciproquement, si $M = \mathrm{Mat}_{\mathcal{B}_1,\mathcal{C}_1}(f)$ est la matrice de f relativement à deux bases $\mathcal{B}_1$ et $\mathcal{C}_1$, et si N est équivalente à M, alors il existe deux bases $\mathcal{B}_2$ et $\mathcal{C}_2$ de E et F telles que $N = \mathrm{Mat}_{\mathcal{B}_2,\mathcal{C}_2}(f)$.

◁ **Éléments de preuve.**

Le premier point résulte de la formule de changement de base, le second du fait que toute matrice inversible peut être vue comme matrice de passage, l'une des deux bases étant imposée. ▷

Soit, pour tout $(n,p) \in (\mathbb{N}^*)^2$, et pour tout $r \in [\![0, \min(n,p)]\!]$, soit

$$I_{n,p,r}=\begin{pmatrix}1&0&\cdots&\cdots&0&\cdots&\cdots&0\\0&1&\ddots&&\vdots&&&\vdots\\\vdots&\ddots&\ddots&\ddots&\vdots&&&\vdots\\\vdots&&\ddots&1&0&\cdots&\cdots&0\\0&\cdots&\cdots&0&0&\cdots&\cdots&0\\\vdots&&&\vdots&\vdots&&&\vdots\\0&\cdots&\cdots&0&0&\cdots&\cdots&0\end{pmatrix}=\left(\begin{array}{c|c}I_r&0_{r,p-r}\\\hline 0_{n-r,r}&0_{n-r,p-r}\end{array}\right)$$

la matrice de type (n,p), constituée de 1 sur ses r premiers coefficients diagonaux, et de 0 partout ailleurs.

Théorème 24.4.10

Soit E et F des espaces de dimensions respectives p et n. Soit $f\in\mathcal{L}(E,F)$, de rang r. Il existe deux bases $\mathcal{B}$ et $\mathcal{C}$, respectivement de E et de F, telles que

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)=I_{n,p,r}.$$

Corollaire 24.4.11

Toute matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ de rang r est équivalente à $I_{n,p,r}$.

◁ **Éléments de preuve.**

Deux points de vue, importants l'un et l'autre :

- il s'agit d'un échelonnement en lignes de la matrice par le pivot, suivi d'un échelonnement en colonnes par un pivot sur les colonnes (autrement dit, on transpose et on refait un pivot).
- Vectoriellement, construire au départ une base adaptée à $S\oplus\mathrm{Ker}(u)$ où S est un supplémentaire de $\mathrm{Ker}(u)$, et à l'arrivée, construire une base dont les r premiers vecteurs sont des vecteurs (bien choisis) de $\mathrm{Im}(u)$. Exprimer la matrice de u relativement à ces bases.

▷

Le premier point de vue, purement matriciel, a l'avantage d'insister sur le fait que ce résultat est une formalisation du pivot de Gauss, et permet souvent de rédiger de façon très efficace des arguments abstraits basés sur le pivot. Le deuxième point de vue donne un éclairage plus géométrique.

Corollaire 24.4.12 (Classification des matrices équivalentes par le rang)

Deux matrices de même format sont équivalentes si et seulement si elles ont même rang.

◁ **Éléments de preuve.**

Elles sont alors équivalentes à la même matrice $I_{n,p,r}$. ▷

Ainsi, dans $\mathcal{M}_{n,p}(\mathbb{R})$, les classes d'équivalence, pour la relation d'équivalence des matrices, sont les sous-ensembles de matrices de même rang $r\in[\![0,\min(n,p)]\!]$. En particulier, l'espace quotient est en bijection avec $[\![0,\min(n,p)]\!]$.

IV.3 Matrice d'un endomorphisme, matrices semblables

Dan le cas où f est un endomorphisme, il est fréquent de choisir sur E la même base au départ et à l'arrivée (même si ceci n'est en théorie par strictement nécessaire). Dans ce cas, on allège un peu les notations, en notant simplement $\mathrm{Mat}_{\mathcal{B}}(f)$ au lieu de $\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(f)$. Par ailleurs, un changement de base sur un endomorphisme s'effectue dans ce cas en faisant le même changement de variable au départ et à l'arrivée. Ainsi :

Théorème 24.4.13 (Changement de base pour un endomorphisme)
Soit $f \in \mathcal{L}(E)$, E étant un espace vectoriel de dimension finie, muni de deux bases $\mathcal{B}_1$ et $\mathcal{B}_2$. Alors :

$$\mathrm{Mat}_{\mathcal{B}_2}(f) = (P_{\mathcal{B}_1}^{\mathcal{B}_2})^{-1}\mathrm{Mat}_{\mathcal{B}_1}(f)P_{\mathcal{B}_1}^{\mathcal{B}_2}.$$

Ainsi, cette relation s'écrit : $M' = P^{-1}MP$, où M est la matrice de f dans l'ancienne base, M' la matrice de f dans la nouvelle base, et P la matrice de passage de l'ancienne base vers la nouvelle base.

Exemple 24.4.14

1. Soit p un projecteur de rang r d'un espace de dimension n. Montrer qu'il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(p) = I_{n,n,r}$.
2. Décrire de même une matrice simple représentant une symétrie dans un certain choix de base.
3. Déterminer une base relativement à laquelle la matrice de $f : (x, y) \mapsto (3x - 2y, x)$ est diagonale. En déduire une relation entre cette matrice diagonale et la matrice $\begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}$.

On dit qu'on a diagonalisé $\begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}$, ou de façon équivalente, qu'on a diagonalisé f.
La notion de diagonalisation est liée à l'étude des classes d'équivalence de la relation fournie par les changements de base pour les endomorphismes. Cela motive la définition de la relation d'équivalence idoine, définissant la notion de matrices semblables.

IV.4 Matrices semblables

Définition 24.4.15 (Matrices semblables)
On dit que deux matrices $A, B \in \mathcal{M}_n(\mathbb{K})$ sont semblables s'il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Proposition 24.4.16 (Relation de similitude)
La relation ainsi définie, appelée relation de similitude, est une relation d'équivalence.

Ainsi :

Corollaire 24.4.17
Les matrices d'un endomorphisme dans différentes bases de E sont semblables.

La classification des matrices semblables est beaucoup plus compliquée que la classification des matrices équivalentes. Cette classification est à l'origine du problème de la réduction des endomorphismes. Le problème de la réduction des endomophismes consiste à trouver un système simple de représentants des classes de similitude, donc de trouver une matrice simple canonique équivalente à une matrice donnée. La diagonalisation des matrices (ou des endomorphismes) est une des branches de ce problème, mais les matrices diagonales ne suffisent pas à donner un système de représentants des classes de similitude. Par ailleurs, il est important de noter que cette notion dépend beaucoup du corps de base. En effet, toute matrice de $\mathcal{M}_n(\mathbb{C})$ est $\mathbb{C}$-semblable à une matrice triangulaire supérieure (et même à une matrice de forme assez particulière, avec seulement deux diagonales non nulles), donc toute matrice de $\mathcal{M}_n(\mathbb{R})$ aussi ; en revanche elles ne sont pas toutes $\mathbb{R}$-semblables à une matrice triangulaire supérieure. Trouver un système de représentants des classes de $\mathbb{R}$-similitude est un problème plus compliqué que trouver un système de représentants de $\mathbb{C}$-similitude.

La classification des classes de similitude étant loin d'être triviale, une étape importante est la recherche d'invariants de similitude, c'est-à-dire de propriétés ou quantités préservées par similitude. Ces invariants permettent de distinguer certaines matrices non semblables (si elles n'ont pas même invariant) mais sont en général insuffisants pour prouver qu'elles sont semblable (sauf si l'invariant caractérise la similitude).

Nous avons déjà étudié un invariant de similitude : le rang. En effet, deux matrices semblables sont équivalentes. Cet invariant est assez faible, puisqu'il ne permet de classer les matrices de $\mathcal{M}_n(\mathbb{K})$ qu'en $n+1$ catégories.

Un deuxième invariant, encore plus facile à calculer, est la trace, que nous étudions dans le paragraphe suivant.

IV.5 Trace d'une matrice, trace d'un endomorphisme

Définition 24.4.18 (Trace d'une matrice)

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée d'ordre n. La trace de A est la somme de ses coefficients diagonaux :

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

De façon assez immédiate, on obtient :

Proposition 24.4.19 (Linéarité de la trace)

L'application $\operatorname{tr} : \mathcal{M}_n(\mathbb{K}) \longrightarrow \mathbb{K}$ *est une forme linéaire sur* $\mathcal{M}_n(\mathbb{K})$.

Proposition 24.4.20 (Invariance par transposition)

Soit $A \in \mathcal{M}_n(\mathbb{K})$. *Alors* $\operatorname{tr}(A) = \operatorname{tr}(A^\top)$.

Théorème 24.4.21 (Invariance de la trace par commutation interne)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ *et* $B \in \mathcal{M}_{p,n}(\mathbb{K})$. *Alors* $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

◁ **Éléments de preuve.**

Il s'agit essentiellement d'une interversion de deux signes somme. ▷

Remarquez que A et B ne sont pas des matrices carrées, mais de format transposé. Ainsi, AB est une matrice carrée d'ordre n alors que BA est une matrice carrée d'ordre p.

Avertissement 24.4.22

Le théorème affirme qu'on peut inverser l'ordre d'un produit dans la trace, lorsque la matrice dont on cherche la trace s'exprime comme produit de 2 termes. Cela ne signifie pas qu'on peut commuter comme on veut les termes d'un produit de n termes à l'intérieur de la trace. Ainsi, on peut écrire :

$$\operatorname{tr}(ABC) = \operatorname{tr}(CAB) = \operatorname{tr}(BCA),$$

mais en général, **on n'a pas** :

$$\operatorname{tr}(ABC) = \operatorname{tr}(ACB),$$

l'expression ACB ne pouvant pas se déduire de l'interversion (globale) de 2 termes de ABC.

Exemple 24.4.23

Comparer $\mathrm{tr}\left(\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\right)$ et $\mathrm{tr}\left(\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\right)$.

Corollaire 24.4.24 (Invariance de la trace par similitude)

La trace est un invariant de similitude. Autrement dit, si M et N sont semblables, alors $\mathrm{tr}(M) = \mathrm{tr}(N)$.

◁ **Éléments de preuve.**
Intervertir P^{-1} et MP. ▷

Cela permet de définir :

Définition 24.4.25 (Trace d'un endomorphisme)

La trace d'un endomorphisme est la valeur commune de la trace des matrices de f relativement à un choix quelconque d'une base.

On en déduit notamment que, comme pour les matrices, la trace est une forme linéaire sur $\mathcal{L}(E)$, et que pour tout $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, E)$, on a $\mathrm{tr}(uv) = \mathrm{tr}(vu)$.
En particulier, on a :

Proposition 24.4.26 (Trace d'un projecteur, d'une symétrie)

1. *Soit p un projecteur de E, alors $\mathrm{tr}(p) = \mathrm{rg}(p)$.*
2. *Soit s une symétrie de E. Alors $\mathrm{tr}(s) = n - 2\,\mathrm{rg}(s - \mathrm{id})$*

◁ **Éléments de preuve.**
Cela provient de la diagonalisation de ces endomorphismes, la diagonalisation étant basée sur la notion de similitude. ▷

Méthode 24.4.27 (Dimension des éléments géométriques d'une projection/symétrie)

- S'assurer que l'endomorphisme considéré est une projection ou une symétrie, en calculant u^2.
- Déterminer la matrice de u dans une base quelconque
- Déterminer la trace de u grâce à cette matrice.
- la propriété précédente donne la dimension des éléments propres $\mathrm{Ker}(u)$ et $\mathrm{Im}(u)$ de u si u est un projecteur, et de $\mathrm{Ker}(u - \mathrm{id})$ et $\mathrm{Ker}(u + \mathrm{id})$ si u est une symétrie (ces deux espaces étant supplémentaires pour une symétrie).

IV.6 Introduction à la réduction des endomorphismes (Spé)

Ces invariants ne représentent qu'un tout petit pas vers la classification des classes de similitude. Vous définirez l'année prochaine les valeurs propres d'une matrice. L'ensemble des valeurs propres d'une matrice est également un invariant de similitude.

Nous donnons sans preuve la description complète d'un système de représentant des classes de similitude dans $\mathcal{M}_n(\mathbb{C})$. Pour cela, nous définissons, pour tout $\ell \in [\![1, n]\!]$ et tout $\lambda \in \mathbb{C}$, la matrice suivante de $\mathcal{M}_\ell(\mathbb{C})$.

$$J_\ell(\lambda) = \lambda I_\ell + J_\ell = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

Dans le cas où $n = 1$, la matrice $J_1(\lambda)$ est réduite à la matrice à un seul coefficient (λ).
Alors, toute matrice de $\mathcal{M}_n(\mathbb{C})$ est semblable, à permutation près des blocs diagonaux, à une unique matrice

$$\begin{pmatrix} J_{\ell_1}(\lambda_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_{\ell_k}(\lambda_k) \end{pmatrix}, \qquad \ell_1 + \cdots + \ell_k = n, \quad (\lambda_1, \dots, \lambda_k) \in \mathbb{C}^k,$$

matrice constituée de blocs diagonaux carrés égaux aux matrices $J_{\ell_i}(\lambda_i)$.
Le problème de la réduction d'un endomorphisme u, ou de façon équivalente, d'une matrice M, est de trouver une base relativement à laquelle la matrice de u est de la forme ci-dessus, ou de façon équivalente, de trouver une matrice P inversible telle que $P^{-1}MP$ soit de la forme ci-dessus.
Si les ℓ_k sont tous égaux à 1, la matrice obtenue est diagonale. Ainsi, les matrices diagonales (à permutation près des facteurs) font partie de la famille décrite ci-dessus. Le problème de la diagonalisation est donc un sous-problème du problème plus vaste de la réduction, et si elle n'aboutit pas, il faudra chercher un représentant de la classe de M sous le format plus général donné ci-dessus. Nous définissons :

Définition 24.4.28 (Endomorphisme diagonalisable, matrice diagonalisable)

(i) Un endomophisme est diagonalisable s'il existe une base dans laquelle sa matrice est diagonale.

(ii) Une matrice est diagonalisable si elle est semblable à une matrice diagonale.

Trouver une telle matrice diagonale (et la matrice de passage associée) s'appelle « diagonaliser » l'endomorphisme ou la matrice. Vous verrez l'année prochaine des techniques efficaces pour diagonaliser un endomorphisme ou une matrice.

De façon plus générale, trouver un représentant sous la forme générale exposée ci-dessus (et la matrice de passage correspondante) s'appelle « réduire » ou « jordaniser » l'endomorphisme ou la matrice.
Pour la culture, le résultat qui est à la base du théorème de jordanisation, disant que toute matrice à coefficients complexes est jordanisable, est le théorème suivant :

Théorème 24.4.29 (Théorème des noyaux itérés)

Soit u un endomorphisme de E de dimension finie, on a les inclusions :

$$\mathrm{Ker}(u^0) \subset \mathrm{Ker}(u^1) \subset \mathrm{Ker}(u^2) \subset \cdots \subset \mathrm{Ker}(u^k) \subset \mathrm{Ker}(u^{k+1}) \subset \cdots.$$

De plus, cette suite est stationnaire, et « s'essouffle », dans le sens où les sauts de dimension forment une suite décroissante, ultimement nulle.

◁ **Éléments de preuve.**

Nous verrons la démonstration complète de ce théorème en exercice. L'idée générale consiste à raisonner plutôt sur les images et à remarquer qu'on passe d'une image à la suivante restreignant u à des espaces de plus en plus petits. ▷

V Formes linéaires et Hyperplans

Nous terminons ce chapitre par l'étude d'une famille importante d'applications linéaires : les formes linéaires, qui correspondent aux applications linéaires à valeurs dans le corps de base. Les formes linéaires sont intimement liées à la notion d'hyperplan (généralisant la notion de plan en dimension 3). Elles sont aussi à la base des théories de dualité (qui ne sont pas au programme). Nous définissons ces notions sans hypothèse de finitude de dimension, mais nous observerons qu'en dimension finie, les hyperplans se caractérisent facilement par leur dimension.

Définition 24.5.1 (Forme linéaire)

Une forme linéaire sur un $\mathbb{K}$-espace vectoriel E est une application linéaire de E vers $\mathbb{K}$, donc un élément de $\mathcal{L}(E,\mathbb{K})$.

Exemples 24.5.2

1. $f \mapsto \int_a^b f(t)\,\mathrm{d}t$ sur ...
2. La trace sur $\mathcal{M}_n(\mathbb{R})$.
3. L'évaluation des polynômes en a.

Définition 24.5.3 (Dual)

Soit E un espace vectoriel. On appelle *dual* de E, et on note E^*, l'espace vectoriel $\mathcal{L}(E,\mathbb{K})$ constitué des formes linéaires.

On remarquera que si E est de dimension finie, $\dim(E) = \dim(E^*)$.
On définit alors les hyperplans de la manière suivante :

Définition 24.5.4 (Hyperplan)

Soit H un sous-espace vectoriel de E. On dit que H est un *hyperplan* de E s'il existe une forme linéaire non nulle $\varphi \in E^*$ telle que $H = \mathrm{Ker}(\varphi)$. L'équation $\varphi(x) = 0$, caractérisant l'appartenance à H, est appelée *équation de* H.

Proposition 24.5.5 (Caractérisation des hyperplans en dimension finie)

Si E est de dimension finie n, les hyperplans de E sont exactement les sous-espaces vectoriels de E de dimension $n-1$.

◁ **Éléments de preuve.**

Dans un sens, utiliser le théorème du rang. Dans l'autre, considérer un supplémentaire S, et définir explicitement à l'aide de H et S une forme linéaire dont H est le noyau. ▷

En appelant *codimension du sous-espace F de E* la quantité $\mathrm{codim}_E(F) = \dim(E) - \dim(F)$, les hyperplans sont donc les sous-espaces de E de codimension 1.

Exemples 24.5.6

1. Les plans vectoriels de $\mathbb{R}^3$.
2. Les droites vectorielles de $\mathbb{R}^2$.
3. Les polynômes sans terme constant.

Théorème 24.5.7 (Caractérisation par les supplémentaires)

Soit E un espace vectoriel quelconque (de dimension finie ou non). Soit H un sous-espace de E. Alors H est un hyperplan de E si et seulement si H admet un supplémentaire égal à une droite de E.

◁ **Éléments de preuve.**

Pour la réciproque, c'est la même chose que dans le cas de la dimension finie. Pour le sens direct, pour ne pas utiliser la propriété (HP et dépendant de l'AC) d'existence d'un supplémentaire, considérer

x tel que $\varphi(x) \neq 0$, et montrer que $\mathbb{K}x$ est un supplémentaire de H. Si ce n'est pas le cas, montrer l'existence d'une injection de $\mathrm{Vect}(x,y)$ dans $\mathbb{K}$, où (x,y) est une famille libre. ▷

Malgré l'utilisation de supplémentaires en dimension non nécessairement finie, ce théorème ne nécessite pas l'axiome du choix.

Proposition 24.5.8 (Comparaison de deux équations de H)

Soit H un hyperplan de E, d'équation $\varphi \in \mathrm{E}^*$. Alors pour tout $\psi \in E^*$, $\psi(x) = 0$ est une équation de H si et seulement si $\psi \neq 0$ et $\psi \in \mathrm{Vect}(\varphi)$.

◁ **Éléments de preuve.**

Écrire $E = H \oplus \mathbb{K}x_0$, et comparer $\psi(x_0)$ et $\varphi(x_0)$. le coefficient de colinéarité sera alors le même pour tout vecteur y de E. ▷

Ainsi, à tout hyperplan de E correspond une droite du dual E^*.

Théorème 24.5.9 (Intersection d'hyperplans)

Soit E un espace de dimension finie n.

1. *L'intersection de m hyperplans de E est un sous-espace vectoriel de dimension au moins $n-m$.*
2. *Réciproquement, tout sous-espace vectoriel F de E de dimension $n-m$ peut s'écrire comme l'intersection de m hyperplans.*

◁ **Éléments de preuve.**

1. Par récurrence sur m, en utilisant la formule de Grassmann.
2. Construire une base de E à partir d'une base de H et d'une base d'un supplémentaire S. Considérer les hyperplans définis par l'annulation d'une des composantes sur S.

▷

Remarquez qu'en dimension finie, après choix d'une base $(b_1,\cdots,b_n)$, une forme linéaire sera décrite par une expression du type

$$\varphi(x) = \sum_{i=1}^{n} a_i x_i,$$

où les x_i sont les coordonnées de x dans la base $(b_1,\ldots,b_n)$. Ainsi, l'équation d'un hyperplan est de la forme

$$a_1x_1 + \cdots + a_nx_n = 0.$$

On retrouve les équations usuelles d'une droite dans $\mathbb{R}^2$ ($ax+by=0$) ou d'un plan dans $\mathbb{R}^3$ ($ax+by+cz=0$), les x, y et z correspondant ici aux coordonnées dans la base canonique.

Ce que dit le dernier théorème est alors simplement le fait qu'un sous-espace de dimension $n-m$ (donc de codimension m) peut être décrit par un système de m équations de ce type.

25
Groupes symétriques

L'examen de ces Formules fournit cette Règle générale. Le nombre des équations & des inconnues étant n, on trouvera la valeur de chaque inconnue en formant n fractions dont le dénominateur commun a autant de termes qu'il y a de divers arrangements de n choses différentes. Chaque terme est composé des lettres ZYXV etc. toujours écrites dans le même ordre, mais auquelles on distribue, comme exposants, les n premiers chiffres rangés en toutes les manières possibles [les Z^i, Y^i... représentent les coefficients du système d'équations]. Ainsi, lorsqu'on a trois inconnues, le dénominateur a ($1 \times 2 \times 3 =$) 6 termes, composés des trois lettres ZYX qui recoivent successivement les exposants 123, 132, 213, 231, 312, 321. On donne à ces termes les signes + ou −, selon la règle suivante. Quand un exposant est suivi, dans le même terme, médiatement ou immédiatement, d'un exposant plus petit que lui, j'appellerai cela un dérangement. Qu'on compte, pour chaque terme, le nombre de dérangements : s'il est pair ou nul, le terme aura le signe + ; s'il est impair, le terme aura le signe −. Par ex. dans le terme $Z^1Y^2V^3$ [sic] il n'y a aucun dérangement : ce terme aura donc le signe +. Le terme $Z^3Y^1X^2$ aura le signe + parce qu'il y a deux dérangements, 3 avant 1 & 3 avant 2. Mais le terme $Z^3Y^2X^1$ qui a trois dérangements, 3 avant 2, 3 avant 1 & 2 avant 1, aura le signe −.
Le dénominateur commun étant ainsi formé, on aura la valeur de z en donnant à ce dénominateur le numérateur qui se forme en changeant, dans tous ses termes, Z en A [(les A^i forment le second membre)]. Et la valeur d'y est la fraction qui a le même dénominateur & pour numérateur la quantité qui résulte quand on change Y en A dans tous les termes du dénominateur. Et on trouve de manière semblable la valeur des autres inconnues.

(Gabriel Cramer)

Le texte ci-dessus décrit la règle de Cramer pour la résolution des systèmes linéaires. Au passage, Cramer donne une description des numérateurs et dénominateurs correspondant à la définition actuelle des déterminants. La signature des permutations y est défini par le nombre d'inversions (terminogie actuelle pour désigner ce que Cramer appelait des dérangements ; la notion de dérangement désigne actuellement plutôt des permutations sans point fixe ; on peut retenir au passage qu'avant de se fixer, la terminologie mathématique peut être fluctuante).
Nous rappelons que le groupe symétrique $\mathfrak{S}_n$ ou S_n est le groupe des permutations de $[\![1, n]\!]$, la loi du groupe étant la composition. L'étude algébrique des groupes $\mathfrak{S}_n$ est très intéressante, et aboutit à des résultats aussi importants que la non résolubilité par radicaux des équations de degré au moins 5. Cette étude va bien au-delà des objectifs du programme, qui sont uniquement d'introduire les outils nécessaires pour pouvoir définir correctement les déterminants.

Notre but est d'introduire les outils nécessaires à la définition de la signature d'une permutation, notion utilisée ensuite dans l'étude des déterminants. La signature d'une permutation est définie comme l'unique morphisme de groupes $\varepsilon : \mathfrak{S}_n \mapsto \{-1,1\}$ non constant. Notre objectif est donc de montrer l'existence et l'unicité de ce morphisme.

I Notations et cycles

Nous rencontrerons deux façons de décrire des permutations. La première, la plus naturelle, est d'associer, dans un tableau, à chaque valeur $i \in [\![1,n]\!]$, la valeur de $\sigma(i)$, en rangeant les valeurs de i dans l'ordre croissant. Nous disposerons ces données sous la forme matricielle suivante :

Notation 25.1.1 (Permutation)

Soit $\sigma \in \mathfrak{S}_n$. Nous désignerons explicitement σ par le tableau :

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

Exemple 25.1.2

1. Décrire de la sorte tous les éléments de $\mathfrak{S}_2$ et $\mathfrak{S}_3$.
2. Décrire de la sorte la permutation de $\mathfrak{S}_n$ inversant l'ordre : $\sigma(k) = n+1-k$.

Certaines permutations ont un comportement particulier : elles effectuent une rotation sur un ensemble d'éléments (qu'on peut ranger en cercle), et laissent les autres invariants. On appelle *cycle* ces permutations. Pour rendre le caractère cyclique plus visible sur les notations on adopte pour ces permutations une notation plus adaptée. Nous définissons de façon plus précise :

Définition 25.1.3 (Cycle)

Un cycle est une permutation σ telle qu'il existe $k \in n\mathbb{N}^*$, et k éléments deux à deux distincts $i_1, \dots, i_k$ de $[\![1,n]\!]$, tels que

(i) $\forall \ell \in [\![1,k-1]\!], \quad \sigma(i_\ell) = i_{\ell+1}$ et $\sigma(i_k) = i_1$

(ii) $\forall i \in [\![1,n]\!] \smallsetminus I, \quad \sigma(i) = i$.

Définition 25.1.4 (Support d'un cycle)

- L'ensemble $I = \{i_1, \dots, i_k\}$ de la définition précédente est appelé support de σ.
- Le support est entièrement défini par le cycle σ, sauf si $k = 1$. En effet, sauf dans ce cas-là, il s'agit du supplémentaire de l'ensemble des points fixes.

Il est fréquent de ne pas considérer les cycles de longueur 1 dans la définition précédente (et donc d'imposer $k \geqslant 2$). Cela se justifie par le fait qu'ils sont tous égaux (à l'identité). Pour simplifier les énoncés de certains résultats faisant intervenir des cycles, nous laissons le cas $k = 1$ dans la définition. Remarquez qu'alors, un cycle de support $\{x_i\}$ ne peut être autre chose que l'identité.

Exemple 25.1.5

Sont-ce des cycles : $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 5 & 3 & 2 & 6 \end{pmatrix}$? $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 2 & 6 & 1 \end{pmatrix}$?

Notation 25.1.6 (Cycles)
Avec les conditions de la définition 25.1.3, on note

$$\sigma = (i_1\ i_2\ \ldots\ i_k).$$

Son support est donc l'ensemble $I = \{i_1, \ldots, i_k\}$. La longueur du cycle est l'entier $k = \mathrm{Card}(I)$.

Cette notation signifie que tout élément de la liste est envoyé sur l'élément suivant, le dernier étant envoyé sur le premier.

Exemple 25.1.7

1. Écrire le cycle de l'exemple 25.1.5 avec les notations 25.1.6.
2. Écrire le cycle $(1\ 5\ 8\ 2\ 9\ 7\ 3)$ avec les notations 25.1.1.

Avertissement 25.1.8
Attention, l'ordre des éléments $i_1, \ldots, i_n$ est important.

Remarque 25.1.9
La représentation d'un cycle sous la forme 25.1.6 est-elle unique ?

On rencontre parfois :

Définition 25.1.10 (Grand cycle, ou permutation circulaire)
Un grand cycle, ou permutation circulaire, de $\mathfrak{S}_n$ est un cycle de support $[\![1, n]\!]$.

Ainsi, un grand cycle effectue une rotation, dans un certain ordre, entre les n éléments de $[\![1, n]\!]$. Un exemple important de grand cycle est :

Définition 25.1.11 (Permutations circulaires directe et indirecte)

(i) On appelle permutation circulaire directe de $\mathfrak{S}_n$ le grand cycle $(1\ 2\ \cdots\ n)$.

(ii) On appelle permutation circulaire indirecte de $\mathfrak{S}_n$ le grand cycle $(n \ldots\ 2\ 1)$.

Exemple 25.1.12
Écrire les permutations circulaires directes et indirectes avec les notations 25.1.1.

Une famille importante de cycles est constituée des cycles de longueur 2. En effet, il s'agit d'une famille génératrice de $\mathfrak{S}_n$, comme on le verra plus tard.

Définition 25.1.13 (Transpositions)
On appelle transposition de $\mathfrak{S}_n$ un cycle de longueur 2 : $\tau = (i\ j)$.

Ainsi, la transposition $\tau = (i\ j)$ est la permutation consistant en l'échange des valeurs i et j.

Exemple 25.1.14

Écrire la transposition $\tau = (2\ 5)$ de $\mathfrak{S}_6$ avec les notations 25.1.1.

II Signature d'une permutation

Lemme 25.2.1

Soit $\sigma \in \mathfrak{S}_n$. Alors σ induit une bijection sur $\mathcal{P}_2(\llbracket 1, n \rrbracket)$, l'ensemble des sous-ensembles de cardinal 2 de $\llbracket 1, n \rrbracket$.

◁ **Éléments de preuve.**

La décrire explicitement et décrire sa réciproque. ▷

En notant, pour tout $X = \{i, j\}$, $\delta_\sigma(X) = |\sigma(i) - \sigma(j)|$, on en déduit :

Lemme 25.2.2

Soit $\sigma \in \mathfrak{S}_n$. On a :

$$\left| \prod_{1 \leqslant i < j \leqslant n} (\sigma(j) - \sigma(i)) \right| = \prod_{X \in \mathcal{P}_2(\llbracket 1, n \rrbracket)} \delta_\sigma(X) = \prod_{1 \leqslant i < j \leqslant n} (j - i)$$

◁ **Éléments de preuve.**

C'est juste un changement de variable sur le produit, défini par la bijection du lemme précédent. ▷

Puisque $\dfrac{\sigma(j) - \sigma(i)}{j - i} = \dfrac{\sigma(i) - \sigma(j)}{i - j}$, cette quantité ne dépend pas de l'ordre respectif entre i et j, mais uniquement de l'ensemble $\{i, j\}$. On peut donc définir une application τ_σ sur $\mathcal{P}_2(\llbracket 1, n \rrbracket)$ par :

$$X = \{i, j\} \mapsto \tau_\sigma(X) = \frac{\sigma(i) - \sigma(j)}{i - j}.$$

Définition 25.2.3 (Signature)

On définit la signature $\varepsilon : \mathfrak{S}_n \longrightarrow \{-1, 1\}$ par :

$$\varepsilon(\sigma) = \frac{\displaystyle\prod_{1 \leqslant i < j \leqslant n} (\sigma(j) - \sigma(i))}{\displaystyle\prod_{1 \leqslant i < j \leqslant n} (j - i)} = \prod_{X \in \mathcal{P}_2(\llbracket 1, n \rrbracket)} \tau_\sigma(X).$$

Théorème 25.2.4

La signature ε est un morphisme de groupes de $(\mathfrak{S}_n, \circ)$ dans $(\{-1, 1\}, \times)$.

◁ **Éléments de preuve.**

La bonne définition provient du lemme 25.2.2. La propriété de morphisme s'obtient en introduisant une « étape » dans le quotient, en coupant le produit en 2, et en utilisant encore un changement de variable du même type sur l'un des deux produits. ▷

De l'étude du signe de l'expression définissant $\varepsilon(\sigma)$, il découle que celui-ci-dépend du nombre de couples (i, j) avec $i < j$ tel que $\sigma(i) > \sigma(j)$.

Définition 25.2.5 (Inversion)
Soit $\sigma \in \mathfrak{S}_n$. On appelle inversion de σ un couple (i,j) tel que $i < j$ et $\sigma(i) > \sigma(j)$, c'est-à-dire $\tau_\sigma(\{i,j\}) < 0$.

Ainsi, la signature peut être décrite à l'aide du nombre d'inversions :

Théorème 25.2.6 (Description de la signature par les inversions)
En notant $\mathrm{Inv}(\sigma)$ *le nombre d'inversions de* σ*, on a :*

$$\varepsilon(\sigma) = (-1)^{\mathrm{Inv}(\sigma)}.$$

◁ **Éléments de preuve.**
Il n'y a qu'à déterminer le signe ! ▷

Une transposition $(i\ i+1)$ n'admettant que le couple $(i, i+1)$ comme inversion, il vient alors :

Proposition 25.2.7
Notons, pour tout $i \in [\![1, n-1]\!]$, $\tau_i = (i\ i+1)$*. On a alors* $\varepsilon(\tau_i) = -1$.

◁ **Éléments de preuve.**
Compter les inversions ! ▷

Or, toute transposition peut s'écrire à l'aide de ces transpositions entre éléments consécutifs :

Proposition 25.2.8 (Décomposition d'une transposition à l'aide des τ_i)
Soit $1 \leqslant i < j \leqslant n$*, et* $\tau = (i\ j)$*. Alors*

$$\tau = \tau_{j-1} \circ \cdots \circ \tau_{i+1} \circ \tau_i \circ \cdots \circ \tau_{j-2} \circ \tau_{j-1}.$$

◁ **Éléments de preuve.**
Vérification facile (suivre les éléments, peut se faire graphiquement sur des diagammes sagittaux) ▷

En utilisant le fait que ε est un morphisme, on en déduit :

Théorème 25.2.9 (Signature d'une transposition)
Soit τ *une transposition. Alors* $\varepsilon(\tau) = -1$.

◁ **Éléments de preuve.**
Provient des résultats précédents. On peut aussi le voir directement en comptant les croisements dans le diagramme sagittal (pourquoi est-ce la même chose ?) ▷

Nous avons donc répondu au problème de l'existence d'un morphisme $\varepsilon : \mathfrak{S}_n \to \{-1, 1\}$, prenant la valeur -1 sur les transpositions.
Pour étudier son unicité, nous allons établir que toute permutation s'écrit comme composition de transposition. Ainsi, la valeur d'un morphisme ε sur une permutation est entièrement déterminée par la valeur de ε sur les transpositions. Imposer la valeur sur les transpositions fournit dans ce cas l'unicité.

Théorème 25.2.10 (Caractère générateur des transpositions)

Toute permutation $\sigma \in \mathfrak{S}_n$ est un produit de transpositions.

◁ **Éléments de preuve.**

Peut se démontrer par l'étude de la correction de l'algorithme du tri à bulle. On peut aussi procéder par récurrence, en composant par une transposition bien choisie pour rendre n fixe, et se ramener ainsi à l'hypothèse de récurrence. ▷

Les transpositions étant elles mêmes engendrées par les τ_i, on se rend compte que toute permutation s'écrit comme produit des τ_i. Cela ne soit pas nous étonner, on l'avait déjà prouvé en étudiant la correction de certains algorithmes de tri (lesquels ?)

Remarque 25.2.11

Le rapport obtenu entre signature et décomposition en produit de transpositions permet d'affirmer que la parité du nombre de terme de toute décomposition en produits de transpositions d'une permutation donnée est la même.

Pour montrer l'unicité du morphisme non trivial ε, nous montrons qu'un morphisme $\varphi : \mathfrak{S}_n \to \{-1, 1\}$ attribue la même valeur à toute transposition.

Lemme 25.2.12 (Effet de la conjugaison sur un cycle)

Soit $\sigma \in \mathfrak{S}_n$ et $(a_1 \ \dots \ a_k)$ un cycle. Alors

$$\sigma \circ (a_1 \ \dots \ a_k) \circ \sigma^{-1} = (\sigma(a_1) \ \dots \ \sigma(a_k))$$

◁ **Éléments de preuve.**

Vérification élément par élément en distinguant deux cas. Se comprend bien sur un dessin, en mettant les 2 cycles sur deux étages ; σ est l'ascenseur permettant de monter, σ^{-1} va dans l'autre sens. ▷

Corollaire 25.2.13

Soit $\alpha \in \{-1, 1\}$. S'il existe une transposition τ_0 telle que $\varphi(\tau_0) = \alpha$, alors pour toute transposition τ, $\varphi(\tau) = \alpha$.

◁ **Éléments de preuve.**

Toutes les transpositions sont conjuguées. ▷

On en déduit enfin notre théorème d'unicité :

Théorème 25.2.14 (propriété d'unicité de la signature)

La signature est le seul morphisme de groupes non trivial (i.e. non égal au morphisme constant $\sigma \mapsto 1$) de $\mathfrak{S}_n$ dans $\{-1, 1\}$.

◁ **Éléments de preuve.**

On a deux cas à étudier : φ prend la valeur 1 sur toutes les transpositions, ou φ prend la valeur -1 sur toutes les transpositions. ▷

Terminologie 25.2.15 (Permutation paire, permutation impaire)

On dira qu'une permutation est paire si $\varepsilon(\sigma) = 1$, et impaire si $\varepsilon(\sigma) = -1$.

Définition 25.2.16 (Groupe alterné)

Le groupe alterné $\mathfrak{A}_n$ est le noyau de la signature, c'est-à-dire l'ensemble des permutations paires. C'est un sous-groupe distingué de $\mathfrak{S}_n$.

III Décomposition cyclique d'une permutation

Décomposer une permutation en produit de transpositions, ou compter le nombre d'inversions n'est pas quelque chose d'immédiat. Nous donnons dans ce paragraphe une autre façon, plus rapide, de calculer la signature d'une permutation, en s'aidant du type cyclique de cette permutation.

Lemme 25.3.1 (Commutation des cycles à supports disjoints)

Soit C_1 et C_2 deux cycles à supports disjoints. Alors $C_1 \circ C_2 = C_2 \circ C_1$.

◁ **Éléments de preuve.**

Le vérifier élément par élément. ▷

Théorème 25.3.2 (Décomposition en cycles d'une permutation)

Soit σ une permutation de $\mathfrak{S}_n$. À permutation près des facteurs, il existe une unique décomposition de σ en produits de cycles :

$$\sigma = C_1 \circ \cdots \circ C_k,$$

telle que les supports des cycles forment une partition de $[\![1, n]\!]$.

◁ **Éléments de preuve.**

Existence : Étant donné x, il existe $j > 0$ minimal tel que $\sigma^j(x) = x$. Les images successives définissent un cycle. Composer par l'inverse de ce cycle pour augmenter le nombre de points fixes et faire eun récurrence.

On peut aussi remarquer que les supports des cycles correspondent aux orbites de l'action de $\mathcal{S}_n$ sur $[\![1, n]\!]$ par l'évaluation.

Unicité : deux décompositions étant données, un unique cycle de chaque décomposition contient x : montrer que ce sont les mêmes. ▷

Remarque 25.3.3

- Les cycles C_i sont les cycles associés à un système de représentants modulo la relation $\equiv_\sigma$.
- L'ordre dans lequel on effectue cette composition importe peu, du fait du lemme qui suit.
- Certains auteurs ne considèrent que des cycles de longueur au moins égale à deux (cas où la donnée du support est déterminée par la donnée du cycle). Dans ce cas, le résultat se réénonce en disant que toute permutation est de façon unique (à permutation près) composée de cycles à supports disjoints.

Méthode 25.3.4 (Comment déterminer la décomposition en produits de cycles)

Soit σ une permutation de $\mathfrak{S}_n$.

- Partir de 1 et suivre ses images successives jusqu'à retomber sur 1. Cela donne le premier cycle.
- Recommencer en partant du plus petit élément de $[\![1, n]\!]$ n'appartenant pas au premier cycle trouvé.
- Recommencer ainsi jusqu'à épuisement des éléments de $[\![1, n]\!]$.

Exemple 25.3.5

Trouver la décomposition en cycles à supports disjoints de :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 5 & 2 & 8 & 10 & 7 & 9 & 1 & 3 & 4 & 6 \end{pmatrix}.$$

Définition 25.3.6 (Support cyclique d'une permutation)

Le support cyclique d'une permutation σ est la partition formée des supports des cycles formant la décomposition en cycles de σ. Ainsi, il s'agit de la partition formée des classes d'équivalence de la relation $\equiv_\sigma$.

Définition 25.3.7 (Type cyclique d'une permutation)

Le type cyclique d'une permutation σ est la suite croissante des tailles des parts du support cyclique. Une telle suite croissante de somme n est appelée partition de l'entier n.

Dans ces deux dernières définitions, il est nécessaire de tenir compte aussi des supports de taille 1.

IV Cycles et signature

Nous voyons enfin comment utiliser la décomposition en cycles pour déterminer la signature. Pour cela nous remarquons qu'il est facile d'écrire un cycle comme produit de transpositions :

Lemme 25.4.1 (Décomposition d'un cycle en transpositions)

Soit $\{i_1, \dots, i_k\}$ des entiers 2 à 2 distincts de $[\![1, n]\!]$. Alors :

$$(\ i_1\ i_2\ \cdots\ i_k\ ) = (i_1\ i_k) \circ (i_1\ i_{k-1}) \circ \cdots \circ (i_1\ i_2).$$

◁ **Éléments de preuve.**

Suivre les éléments. ▷

Dans le cas où $k = 1$, le terme de droite est réduit à un produit vide de transpositions (égal à l'identité)

Proposition 25.4.2 (Signature d'un cycle)

Soit C un cycle et $\ell(C)$ sa longueur. Alors

$$\varepsilon(C) = (-1)^{\ell(C)-1}.$$

Théorème 25.4.3 (Détermination de ε par le type cyclique)
Soit σ une permutation de $\mathfrak{S}_n$, et $c(\sigma)$ le nombre de parts dans son support cyclique (ou de façon équivalente dans son type cyclique). Alors :

$$\varepsilon(\sigma) = (-1)^{n-c(\sigma)}.$$

◁ **Éléments de preuve.**
Par propriété de morphisme, et en utilisant le fait que les supports des cycles forment une partition.
▷

26

Déterminants

M. Gauss s'en est servi avec avantage dans ses Recherches analytiques pour découvrir les propriétés générales des formes du second degré, c'est-à-dire des polynômes du second degré à deux ou plusieurs variables, et il a désigné ces mêmes fonctions sous le nom de déterminants. Je conserverai cette dénomination qui fournit un moyen facile d'énoncer les résultats ; j'observerai seulement qu'on donne aussi quelquefois aux fonctions dont il s'agit le nom de résultantes à deux ou à plusieurs lettres. Ainsi les deux expressions suivantes, déterminant et résultante, devront être regardées comme synonymes.

(Augustin de Cauchy)

La grande notoriété n'est assurée en Mathématiques qu'aux noms associés à une méthode, à un théorème, à une notation. Peu importe d'ailleurs que l'attribution soit fondée ou non, et le nom de Vandermonde serait ignoré de l'immense majorité des mathématiciens si on ne lui avait attribué ce déterminant que vous connaissez bien, et qui n'est pas de lui !

(Henri Lebesgue)

Dans ce chapitre, nous étudions les déterminants. Nous verrons cette notion tout d'abord vue comme un objet défini sur une famille de vecteurs, puis nous en déduirons une notion de déterminant d'un endomorphisme, puis d'une matrice carrée. Notre but étant en grande partie de caractériser l'inversibilité d'une matrice grâce à cet objet, nous introduisons ensuite des techniques calculatoires efficaces. Nous voyons enfin comment donner une expression de l'inverse d'une matrice à l'aide des déterminants, même si en pratique, la formule obtenue manque d'efficacité.

Note Historique 26.0.1

- Cardan est le premier à introduire un objet correspondant au déterminant d'ordre 2, dans le cadre de résolutions de systèmes de deux équations à deux inconnues. Il s'agit d'un cas particulier de la règle établie plus tard par Cramer. Il appelle cette règle la *regula de modo*.
- Les déterminants d'ordre supérieur (mais toujours pour des petites valeurs) doivent attendre Leibniz (ordre 3 et presque ordre 4, avec quelques erreurs de signes ; comme quoi même les grands peuvent en faire !) et indépendamment le japonais Kowa Seki. Leibniz montre au passage la formule de développement suivant une colonne, pour les déterminants d'ordre 3. C'est MacLaurin en 1748 qui donnera l'expression du déterminant d'ordre 4 avec les bons signes.
- Cramer décrit de façon complète les déterminants de tout ordre en 1750, et à l'occasion, établit les formules de résolution de systèmes pour lesquelles il est devenu célèbre (même si ces formules ne constituent qu'un petit appendice technique d'un ouvrage beaucoup plus vaste).
- Gauss introduit le mot déterminant.
- Cayley introduit la notation matricielle et l'écriture par barres. Cayley et Sylvester sont à la base de l'étude algébrique des déterminants.

I Définition des déterminants

I.1 Formes multilinéaires

Soit $\mathbb{K}$ un corps.

Définition 26.1.1 (Application multilinéaire)

Soit $E_1, \dots, E_n$ et F des $\mathbb{K}$-espaces vectoriels. Une application n-linéaire est une application

$$\varphi : E_1 \times \cdots \times E_n \longrightarrow F$$

telle que pour tout $i \in [\![1, n]\!]$, φ soit linéaire par rapport à sa i-ième variable, les autres étant fixées quelconques, donc si pour tout $x_1, \dots, x_n$, x'_i et λ,

$$\begin{aligned}&\varphi(x_1, \dots, x_{i-1}, \lambda x_i + x'_i, x_{i+1}, \dots, x_n)\\ &\qquad = \lambda\varphi(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) + \varphi(x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_n).\end{aligned}$$

Une application est une application multilinéaire si elle est n-linéaire pour un certain $n \geqslant 2$.

Proposition 26.1.2

Soit φ une application n-linéaire. Alors $\varphi(x_1, \dots, x_n) = 0$ dès lors qu'une des variables x_i est nulle.

◁ **Éléments de preuve.**

Considérer l'application linéaire obtenue en fixant toutes les autres variables. ▷

Définition 26.1.3 (Forme n-linéaire)

Une forme n-linéaire est une application n-linéaire à valeurs dans $\mathbb{K}$.

Exemple 26.1.4

1. Voici des exemples de formes ou applications bilinéaires ($n = 2$) :
   - $(x, y) \in (\mathbb{R}^n)^2 \mapsto \langle x, y \rangle$, le produit scalaire canonique
   - $(x, y) \in \mathbb{C}^2 \mapsto xy$
   - $(f, g) \in \mathcal{L}(E) \mapsto f \circ g$
   - $(X, Y) \in (\mathbb{R}^n)^2 \mapsto Y^\top MX$
2. $(x_1, \dots, x_n) \mapsto x_1 \cdots x_n$
3. $(f_1, \dots, f_n) \mapsto \displaystyle\int_0^1 f_1(t) \cdots f_n(t)\ \mathrm{d}t$
4. L'aire (signée) du parallélogramme défini par deux vecteurs x et y
5. Le volume du parallélépipède défini par trois vecteurs.

Théorème 26.1.5 (n-linéarité généralisée)

Soit $\varphi : E_1 \times \cdots \times E_n \to F$ une application n-linéaire. Alors pour tout $(k_1, \dots, k_n) \in \mathbb{N}^*$, pour tout $x_{i,j} \in E_i$ et $\lambda_{i,j} \in \mathbb{K}$, ($i \in [\![1, n]\!]$, $j \in [\![1, k_i]\!]$), on a :

$$\varphi\left(\sum_{i_1=1}^{k_1} \lambda_{1,i_1} a_{1,i_1}, \dots, \sum_{i_n=1}^{k_n} \lambda_{n,i_n} a_{n,i_n}\right) = \sum_{i_1=1}^{k_1} \cdots \sum_{i_n=1}^{k_n} \lambda_{1,i_1} \cdots \lambda_{n,i_n} \varphi(a_{1,i_1}, \dots, a_{n,i_n}).$$

◁ **Éléments de preuve.**
Sortir les sommes les unes après les autres, en utilisant la linéarité généralisée par rapport à la k-ième variable, les autres étant fixées, en faisant varier k de 1 à n. Une mise en forme propre peut se faire par récurrence. ▷

Avertissement 26.1.6
Les indices des sommes doivent être indépendants !

Exemple 26.1.7
Dans le cas de la bilinéarité, on obtient :

$$\varphi\left(\sum_{i=1}^{k}\lambda_i a_i, \sum_{j=1}^{\ell}\mu_j b_j\right) = \sum_{i=1}^{k}\sum_{j=1}^{\ell}\lambda_i\mu_j\varphi(a_i, b_j).$$

Notation 26.1.8 (Applications et formes n-linéaires)

- On note $\mathcal{L}(E_1,\dots,E_n;F)$ l'ensemble des applications n-linéaires sur $E_1\times\cdots\times E_n$, à valeurs dans F.
- Si $E_1=\cdots=E_n=E$, on notera plus simplement $\mathcal{L}_n(E;F)$.
- Si de plus, $F=\mathbb{K}$, on notera $\mathcal{L}_n(E)$: il s'agit donc des formes n-linéaires sur E.

Proposition 26.1.9 (L'espace des applications n-linéaires)
L'ensemble $\mathcal{L}(E_1,\dots,E_n;F)$ des applications multilinéaires est un espace vectoriel sur $\mathbb{K}$.

◁ **Éléments de preuve.**
Vérifier qu'il s'agit d'un sev de $F^{E_1\times\cdots\times E_n}$. ▷

Comme pour les applications linéaires, pour déterminer entièrement une application n-linéaire (et donc en particulier une forme n-linéaire), il suffit d'en connaître l'image sur les vecteurs d'une base.

Proposition 26.1.10 (Détermination d'une application n-linéaire sur une base)
Soit pour tout $i\in[\![1,n]\!]$, $(e_{i,j})_{1\leqslant j\leqslant d_i}$ une base de E_i et pour tout $(j_1,\dots,j_n)\in[\![1,d_1]\!]\times\cdots\times[\![1,d_n]\!]$, $f_{j_1,\dots,j_n}$ un élément de F. Alors il existe une unique application n-linéaire $\varphi : E_1\times\cdots\times E_n\longrightarrow F$ telle que

$$\forall (j_1,\dots,j_n)\in[\![1,d_1]\!]\times\cdots\times[\![1,d_n]\!],\quad \varphi(e_{1,j_1},\dots,e_{n,j_n}) = f_{j_1,\dots,j_n}.$$

◁ **Éléments de preuve.**
Cette application est donnée explicitement par :

$$\varphi\left(\sum_{j_1=1}^{d_1}\lambda_{1,j_1}e_{1,j_1},\dots,\sum_{j_n=1}^{d_n}\lambda_{n,j_n}e_{n,j_n}\right) = \sum_{j_1=1}^{d_1}\cdots\sum_{j_n=1}^{d_n}\lambda_{1,j_1}\dots\lambda_{n,j_n}f_{j_1,\dots,j_n}.$$

▷

En particulier, si $E_1=\cdots=E_n=E$, muni d'une unique base $(e_1,\dots,e_d)$ de E, et si on se donne des éléments $f_{i_1,\dots,i_n}$ de F, pour $1\leqslant i_1,\dots,i_n\leqslant d$, il existe une unique application n-linéaire φ de $\mathcal{L}_n(E;F)$ telle que pour tout $(i_1,\dots,i_n)\in[\![1,d]\!]^n$,

$$\varphi(e_{i_1},\dots,e_{i_n}) = f_{i_1,\dots,i_n}.$$

Exemple 26.1.11

Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{K}^n$. Une forme bilinéaire φ sur $\mathbb{K}^n$ est entièrement déterminée par la donnée de n^2 scalaires $\varphi(e_i, e_j)$. En notant M la matrice carrée définie par ces scalaires ($M = (\varphi(e_i, e_j))_{1 \leqslant i,j \leqslant n}$), on peut vérifier que pour tout $X, Y \in \mathbb{K}^n$ (qu'on identifie au vecteur colonne des coordonnées dans la base canonique), on a :

$$\varphi(X, Y) = X^\top M Y.$$

Plus généralement, toute forme bilinéaire sur un espace vectoriel E de dimension finie s'écrit de la sorte après choix d'une base de E (voir un chapitre ultérieur).

I.2 Formes n-linéaires antisymétriques, alternées

Nous définissons maintenant deux types importants de formes n-linéaires prenant leurs variables dans un même espace. Nous en verrons un troisième type dans un chapitre ultérieur.

Définition 26.1.12 (Formes n-linéaires antisymétriques, alternées)

Soit $\varphi \in \mathcal{L}_n(E)$ une forme n-linéaire. On dit que :

1. φ est antisymétrique si pour tout $\sigma \in \mathfrak{S}_n$,

$$\varphi(x_1, \dots, x_n) = \varepsilon(\sigma)\varphi(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

2. φ est alternée si $\varphi(x_1, \dots, x_n) = 0$ dès lors qu'il existe $i \neq j$ tels que $x_i = x_j$.

Exemple 26.1.13

1. L'aire orientée du parallélogramme formé par deux vecteurs de $\mathbb{R}^2$ est une forme bilinéaire alternée.
2. Le volume orienté du parallélépipède formé par trois vecteurs de $\mathbb{R}^3$ est une forme trilinéaire alternée.

Lemme 26.1.14 (Caractérisation par les transpositions)

Pour qu'une forme f soit antisymétrique, (il faut et) il suffit que l'échange de deux quelconques de ses variables provoque un changement de signe.

◁ **Éléments de preuve.**

Utiliser une décomposition d'une permutation σ. ▷

Lemme 26.1.15

Soit φ une forme alternée. Alors pour tout $(x_1, \dots, x_n) \in E^n$, et tout $(i, j) \in [\![1, n]\!]^2$ tel que $i \neq j$:

$$\varphi(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = -\varphi(x_1, \dots, x_j, \dots, x_i, \dots, x_n).$$

Réciproquement, si cette condition est satisfaite, alors, si $\mathbb{K}$ n'est pas de caractéristique 2, φ est alternée.

◁ **Éléments de preuve.**

- En fixant les autres variables, on peut se contenter de formes à 2 variables ; développer $\varphi(x + y, x + y)$.

- Réciproquement, qu'obtient-on si on prend $x_i = x_j$?

▷

Des deux lemmes ci-dessus, on déduit immédiatement le sens direct ci-dessous. La réciproque nécessite une hypothèse supplémentaire sur $\mathbb{K}$.

Théorème 26.1.16 (Antisymétrie des formes alternées)

Toute forme n-linéaire alternée est antisymétrique. Si $\mathbb{K}$ n'est pas de caractéristique 2, toute forme antisymétrique est alternée.

Proposition 26.1.17 (Image d'une famille liée par une forme alternée)

Soit $(x_1, \dots, x_n)$ une famille liée, et φ une forme alternée. Alors $\varphi(x_1, \dots, x_n) = 0$.

◁ **Éléments de preuve.**

Quitte à permuter les variables (ce qui ne fait que changer le signe) on peut supposer que x_n s'écrit en fonction des autres. ▷

Nous aurons l'occasion de parler de formes n-linéaires symétriques (définies comme les formes antisymétriques, mais sans signe) (et plus particulièrement pour $n = 2$), lorsque nous étudierons les produits scalaires (la symétrie étant une des 3 propriétés requises pour qu'une forme bilinéaire définisse un produit scalaire). Pour l'heure, nous nous concentrons sur la notion de forme alternée, fortement liée, d'après les exemples, aux problèmes de calculs d'aires et de volumes. Nous allons voir qu'à un scalaire multiplicatif près, en dimension n, la mesure des hypervolumes est la seule forme n-linéaire.

I.3 Déterminant d'une famille de vecteurs

Théorème 26.1.18 (Formes n-linéaires d'un espace de dimension n)

Soit E un espace vectoriel de dimension n, et $(e_1, \dots, e_n)$ une base de E.

1. *Il existe une unique forme n-linéaire alternée φ sur E telle que $\varphi(e_1, \dots, e_n) = 1$.*
2. *Cette forme n-linéaire est entièrement décrite sur les vecteurs de la base par :*

$$\begin{cases} \varphi(e_{i_1}, \dots, e_{i_n}) = 0 & \text{s'il existe } j \neq k \text{ tel que } i_j = i_k \\ \varphi(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \varepsilon(\sigma) & \text{où } \sigma \in \mathfrak{S}_n \end{cases}$$

3. *Toute autre forme n-linéaire alternée sur E est de la forme $\lambda\varphi$, $\lambda \in \mathbb{K}$.*

◁ **Éléments de preuve.**

L'unicité provient du fait que $\varphi(e_1, \dots, e_n)$ impose 2 (soit par le caractère alternée, soit par antisymétrie). Réciproquement, considérer l'unique forme n-linéaire définie par 2 (d'où provient son existence et son unicité ?), et vérifier qu'elle est alternée. Si $x_i = x_j$ se ramener au cas où les autres x_k sont des vecteurs de la base, puis développer en exprimant x_i dans la base. ▷

Remarque 26.1.19

Si la première condition n'est pas satisfaite, les e_{i_k} sont deux à deux distincts, et en nombre n, donc sont constitués des éléments e_i pris chacun une et une seule fois. Aussi la famille $(e_{i,k})$ peut-elle être vue comme une permutation de la famille (e_i), ce qui permet de l'écrire sous la forme $(e_{\sigma(1)}, \dots, e_{\sigma(n)})$.

Cette unique forme n-linéaire alternée va être notre définition du déterminant (par rapport à une base $\mathcal{B}$).

Définition 26.1.20 (Déterminant d'une famille de vecteurs)

Soit E un espace de dimension n, $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et $(x_1, \dots, x_n)$ une famille de n vecteurs de E. Soit $\det_{\mathcal{B}}$ l'unique forme n-linéaire alternée telle que $\det_{\mathcal{B}}(e_1, \dots, e_n) = 1$. Le *déterminant de la famille* $(x_1, \dots, x_n)$ *par rapport à* $\mathcal{B}$ est le scalaire $\det_{\mathcal{B}}(x_1, \dots, x_n)$.

Ainsi, le déterminant par rapport à $\mathcal{B}$ est l'unique forme n-linéaire alternée prenant la valeur 1 sur la famille $\mathcal{B}$.

Remarque 26.1.21

Si E est un $\mathbb{R}$-espace vectoriel, cette notion est à relier à la notion d'hypervolume relative à une base : le déterminant de n vecteurs par rapport à $\mathcal{B}$ est le volume (orienté) du parallélépipède défini par les n vecteurs, l'unité de volume étant le parallélépipède défini par les vecteurs de la base $\mathcal{B}$.

Exemples 26.1.22

1. Voir l'interprétation géométrique dans $\mathbb{R}^2$ et $\mathbb{R}^3$, muni de la base canonique, puis dans $\mathbb{R}^2$ muni d'une base quelconque.
2. Déterminant par rapport à la base canonique de $\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right)$.
3. Déterminant par rapport à la base canonique de $\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}\right)$.

Nous obtenons plus généralement la description suivante :

Théorème 26.1.23 (Description du déterminant par les coordonnées)

Soit E un espace vectoriel de dimension n et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. Soit $(x_1, \dots, x_n)$ une famille d'éléments de E, dont les coordonnées sont :

$$\forall j \in [\![1, n]\!], \quad [x_j]_{\mathcal{B}} = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix}, \quad \textit{soit: } x_j = \sum_{i=1}^{n} a_{i,j} e_i.$$

On a alors :

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

◁ **Éléments de preuve.**

La première expression provient de l'explicitation de de l'unique forme bilinéaire définie par les relations du 2 du théorème 26.1.18, d'après la remarque qui suit le théorème.

La seconde expression s'obtient par le changement d'indice $\tau = \sigma^{-1}$, en remarquant qu'en réordonnant les termes a_i (par commutativité),

$$a_{\tau^{-1}(1),1} \cdots a_{\tau^{-1}(n),n} = a_{1,\tau(1)} \cdots a_{n,\tau(n)}.$$

Cette manipulation peut être vue formellement comme un changement d'indice dans le produit. Lequel ? ▷

Corollaire 26.1.24 (Effet d'un isomophisme sur le déterminant)
Soit $\Phi : E \to F$ un isomophisme, et $\mathcal{B}$ une base de E. Alors, pour tout $(x_1, \dots, x_n) \in E^n$:

$$\det_{\Phi(\mathcal{B})}(\Phi(x_1), \dots \Phi(x_n)) = \det_{\mathcal{B}}(x_1, \dots, x_n).$$

◁ **Éléments de preuve.**
En effet, les coordonnées des $\Phi(x_i)$ dans $\Phi(\mathcal{B})$ sont égales aux coordonnées des x_i dans $\mathcal{B}$. ▷

En particulier, avec l'isomorphisme $u : x \mapsto [x]_{\mathcal{B}}$, on obtient :

Corollaire 26.1.25 (Déterminant d'une famille versus déterminant des vecteurs coordonnées)
Soit $(x_1, \dots, x_n)$ une famille de vecteurs de E (de dimension n), et $\mathcal{B}$ une base de E. Alors

$$\det_{\mathcal{B}}(x_1, \dots, x_n) = \det_{b.c.}([x_1]_{\mathcal{B}}, \dots, [x_n]_{\mathcal{B}}).$$

Le point 3 du théorème 26.1.18 se réexprime :

Corollaire 26.1.26 (Formes n-linéaires alternées)
Soit E de dimension n, et $\mathcal{B}$ une base de E. Alors l'ensemble des formes n-linéaires alternées est $\mathrm{Vect}(\det_{\mathcal{B}})$.

Ainsi, changer de base ne nous fait pas sortir de la droite $\mathrm{Vect}(\det_{\mathcal{B}})$. Comme par ailleurs, une autre base $\mathcal{B}'$, définira également une droite $\mathrm{Vect}(\det_{\mathcal{B}'})$, aussi égale à l'ensemble des formes n-linéaires alternées, on peut donc affirmer que $\det_{\mathcal{B}}$ et $\det_{\mathcal{B}'}$ diffèrent d'une constante multiplicative. En évaluant en $\mathcal{B}'$, on obtient alors :

Proposition 26.1.27 (Effet d'un changement de base sur le déterminant)
Soit $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E. On a alors :

$$\det_{\mathcal{B}} = \det_{\mathcal{B}}(\mathcal{B}') \times \det_{\mathcal{B}'}.$$

En particulier, en évaluant en $\mathcal{B}$, on obtient :

$$1 = \det_{\mathcal{B}}(\mathcal{B}') \cdot \det_{\mathcal{B}'}(\mathcal{B}).$$

Notamment, si $\mathcal{B}'$ est une base, $\det_{\mathcal{B}}(\mathcal{B}')$ est non nul.
Ayant déjà vu l'image par une forme alternée d'une famille liée, on obtient la caractérisation suivante :

Proposition 26.1.28 (Caractérisation des bases)
Soit E de dimension n, muni d'une base $\mathcal{B}$. Une famille $\mathcal{B}'$ de cardinal n est une base de E si et seulement si $\det_{\mathcal{B}}(\mathcal{B}') \neq 0$.

◁ **Éléments de preuve.**
Si la famille n'est pas une base, elle est liée (pourquoi ?) ▷

Nous verrons un peu plus loin que cette caractérisation équivaut à l'inversibilité de la matrice de la famille $\mathcal{B}'$ dans la base $\mathcal{B}$, cette matrice étant alors la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$.

I.4 Déterminant d'un endomorphisme

On aimerait définir la notion de déterminant d'un endomorphisme indépendamment du choix d'une base. L'idée naturelle qui vient à l'esprit pourrait être de se fixer une base $\mathcal{B} = (b_1, \dots, b_n)$ et de définir le déterminant d'un endomorphisme u comme le déterminant de $(u(b_1), \dots, u(b_n))$ par rapport à $\mathcal{B}$. Cette définition est correcte (elle sera donnée en propriété), mais a l'inconvénient d'introduire une base $\mathcal{B}$, dont nous aimerions nous affranchir. Prenant une autre base, par exemple $(\lambda b_1, \dots, \lambda b_n)$, nous allons alors multiplier le déterminant de la famille par λ^n, mais également le déterminant de la base (toujours relativement à la base initiale). Pour s'affranchir de la notion de base, on peut constater que cela reste vrai pour toute forme n-linéaire alternée (puisqu'elles différent du déterminant d'une constante multiplicative seulement). Cela nous motive la définition suivante :

Lemme 26.1.29

Soit $u \in \mathcal{L}(E)$, E de dimension n. Soit, pour toute forme n-linéaire alternée φ sur E, φ_u définie par

$$\varphi_u(x_1, \dots, x_n) = \varphi(u(x_1), \dots, u(x_n)).$$

Alors φ_u est une forme n-linéaire alternée.

◁ **Éléments de preuve.**
Vérification facile ▷

Proposition/Définition 26.1.30 (Déterminant d'un endomorphisme)

Avec les notations du lemme précédent, il existe un unique scalaire $\det(u)$ tel que pour toute forme n-linéaire alternée φ, on ait :

$$\varphi_u = \det(u) \cdot \varphi.$$

Ce scalaire $\det(u)$ est appelé déterminant de u.

◁ **Éléments de preuve.**
Rappelez-moi quelle est la dimension de l'espace des formes n-linéaires alternées ? ▷

On a alors la propriété attendue :

Proposition 26.1.31 (Caractérisation du déterminant par l'image d'une base)

Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de E. Alors

$$\det(u) = \det_{\mathcal{B}}(u(b_1), \dots, u(b_n)).$$

◁ **Éléments de preuve.**
Prendre $\varphi = \det_{\mathcal{B}}$. ▷

Corollaire 26.1.32 (Déterminant de l'identité)

On a $\det(\mathrm{id}) = 1$.

Proposition 26.1.33

Soit n la dimension de E. Soit $u \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. On a $\det(\lambda u) = \lambda^n \det(u)$.

◁ **Éléments de preuve.**

Passer par une base. ▷

Théorème 26.1.34 (Déterminant d'une composée)

Soit u et v dans $\mathcal{L}(E)$. Alors $\det(v \circ u) = \det(v)\det(u)$. *En d'autres termes,* det *est un morphisme de monoïdes de* $(\mathcal{L}(E), \circ)$ *dans* $(\mathbb{K}, \times)$.

◁ **Éléments de preuve.**

Revenir à la définition. ▷

Enfin, notre but étant de caractériser les matrices inversibles (donc les automorphismes), on obtient :

Théorème 26.1.35 (Caractérisation des automophismes par le déterminant)

Soit $u \in \mathcal{L}(E)$. Alors u est un automorphime de E si et seulement si $\det(u) \neq 0$.

◁ **Éléments de preuve.**

C'est l'association de la caractérisation de la bijectivité par l'image d'une base, et de la caractérisation des bases par le déterminant. ▷

Proposition/Définition 26.1.36 (Groupe spécial linéaire)

$\det : (\mathrm{GL}(E), \times) \to (\mathbb{K}^*, \times)$ est un morphisme de groupe. Son noyau est donc un sous-groupe de $\mathrm{GL}(E)$, appelé groupe spécial linéaire, et noté $\mathrm{SL}(E)$. Ainsi

$$SL(E) = \{u \in \mathcal{L}(E) \mid \det(u) = 1\}.$$

Enfin, si E et F sont de même dimension, la donnée d'un isomorphisme $\Phi : E \to F$ permet de relier les déterminants d'endomorphismes de E et les déterminants d'endomorphismes de F :

Proposition 26.1.37 (Déterminant et conjugaison)

Soit $\Phi : E \to F$ un isomorphisme et $u \in \mathcal{L}(E)$. Alors :

$$\det(\Phi \circ u \circ \Phi^{-1}) = \det(u).$$

◁ **Éléments de preuve.**

Soit $\mathcal{B}$ une base de E et $\mathcal{C} = \Phi(\mathcal{B})$. Exprimer les déterminants par les coordonnées, et remarquer que les coordonnées de $u(b_i)$ dans $\mathcal{B}$ sont les coordonnées de $\Phi \circ u \circ \Phi^{-1}(c_i)$ dans $\mathcal{C}$. ▷

Corollaire 26.1.38 (Expressions d'un déterminant par les coordonnées)

Soit $u \in \mathcal{L}(E)$, et $\mathcal{B} = (b_1, \dots, b_n)$ une base de E. Alors

$$\det(u) = \det_{b.c.}([u(b_1)]_{\mathcal{B}}, \dots, [u(b_n)]_{\mathcal{B}}).$$

I.5 Déterminant d'une matrice carrée

Comme souvent, passer des endomorphismes aux matrices est assez automatique, en utilisant la correspondance usuelle entre des endomorphismes et leur matrice dans un choix de base. La seule difficulté peut résider dans l'indépendance vis-à-vis du choix effectué des bases.

Définition 26.1.39 (Déterminant d'une matrice)

Soit $M \in \mathcal{M}_n(\mathbb{K})$, et $f \in \mathcal{L}(\mathbb{K}^n)$ canonique associé. Le déterminant de M, noté $\det(M)$, est par définition égal au déterminant de f : $\det(M) = \det(f)$.

On obtient alors directement des définitions :

Proposition 26.1.40 (Caractérisation du déterminant de M par les colonnes)

Soit M une matrice de $\mathcal{M}_n(\mathbb{K})$, de colonnes $C_1, \dots, C_n$. Alors

$$\det(M) = \det_{b.c.}(C_1, \dots, C_n).$$

◁ **Éléments de preuve.**
Prendre la description de $\det(f)$ par une base, en choisissant une base convenable. ▷

Corollaire 26.1.41 (Cohérence relative au choix des bases)

Soit f un endormorphisme de E et $\mathcal{B}$ une base de E. Alors

$$\det(f) = \det(\mathrm{Mat}_{\mathcal{B}}(f)).$$

◁ **Éléments de preuve.**
Exprimer le déterminant en base $\mathcal{B}$ par les coordonnées. ▷

Les résultats obtenus pour les endomorphismes se transfèrent alors de façon immédiate aux matrices :

Corollaire 26.1.42 (Déterminant de I_n)

On a $\det(I_n) = 1$.

Proposition 26.1.43 (Effet de la multiplication par un scalaire)

Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$. On a $\det(\lambda A) = \lambda^n \det(A)$.

Théorème 26.1.44 (Déterminant d'un produit)

Soit A et B dans $\mathcal{M}_n(\mathbb{K})$. Alors $\det(AB) = \det(A)\det(B)$.

Et pour terminer, la caractérisation annoncée depuis le début :

Théorème 26.1.45 (Caractérisation des matrices inversibles)

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors $A \in \mathrm{GL}_n(\mathbb{K})$ si et seulement si $\det(A) \neq 0$. Si cette condition est satisfaite, on a alors :

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

◁ **Éléments de preuve.**
La caractérisation découle de la caractérisation des automorphismes. La formule découle de la formule du produit, appliquée à AA^{-1}. ▷

Définition 26.1.46 (Groupe spécial linéaire)
Comme dans le cas des endomorphismes, on définit le n-ième groupe spécial linéaire :

$$\mathrm{SL}_n(\mathbb{K}) = \mathrm{Ker}(\det) = \{A \in \mathrm{GL}_n(\mathbb{K}) \mid \det(A) = 1\}.$$

La condition $A \in \mathrm{GL}_n(\mathbb{K})$ est aussi une conséquence de $\det(A) = 1$, donc on peut aussi définir le groupe spécial par :

$$\mathrm{SL}_n(\mathbb{K}) = \{A \in \mathcal{M}_n(\mathbb{K}) \mid \det(A) = 1\}.$$

Pour des matrices décrites par la donnée tabulaire de leurs coefficients, on utilise souvent la notation suivante :

Notation 26.1.47 (Notation tabulaire du déterminant, Cayley)
Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$. On note

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

Nous transcrivons aussi la description initiale que nous avions donnée pour les familles de vecteurs grâce aux permutations :

Théorème 26.1.48 (Expression du déterminant par les coefficients, def. de Cramer)
Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$. Alors

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

Corollaire 26.1.49 (Expression des déterminant 2×2 et 3×3)

1. *On a :* $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$
2. *(Règle de Sarrus) On a :*

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - gec - hfa - idb,$$

c'est à dire « diagonales descendantes moins diagonales montantes ».

Avertissement 26.1.50
Prenez garde à ne pas généraliser trop vite la règle de Sarrus à des matrices d'ordre supérieur ! Pour $n \geqslant 4$, une règle aussi simpliste est fausse ! Il faut considérer toutes les permutations possibles !

La description explicite par les coefficients permet également d'établir de façon quasi-immédiate :

Théorème 26.1.51 (Invariance du déterminant par transposée)
Soit $A \in \mathcal{M}_n(\mathbb{R})$*. On a alors* $\det(A^\top) = \det(A)$.

II Calcul des déterminants

Nous abordons maintenant les techniques calculatoires. Nous verrons essentiellement quatre techniques, la première étant une adaptation de la méthode du pivot, la seconde étant une façon de se ramener à des déterminants plus petits lorsque la matrice est triangulaire par blocs, l'importante troisième méthode est le développement suivant une ligne, ramenant le calcul d'un déterminant d'ordre n au calcul de n déterminants d'ordre $n-1$, coefficientés par les coefficients de la ligne (intéressant surtout lorsque la plupart de ces coefficients sont nuls), et la dernière est l'utilisation du caractère polynomial du déterminant en chacune des coordonnées de la matrice.

II.1 Opérations sur les lignes et colonnes

Proposition 26.2.1 (Déterminant d'une matrice triangulaire)
Le déterminant d'une matrice triangulaire est égal au produit de ses coefficients diagonaux :

$$\begin{vmatrix} \lambda_1 & \bullet & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bullet \\ 0 & \cdots & 0 & \lambda_n \end{vmatrix} = \lambda_1\lambda_2\cdots\lambda_n.$$

◁ **Éléments de preuve.**
Dans la somme explicite, seul un terme est non nul. Lequel ? ▷

Lemme 26.2.2 (Déterminant des matrices de codage des opérations)
On a :

$$\det(E(i,j)) = -1, \qquad \det(E_i(\lambda)) = \lambda \qquad \det(E_{i,j}(\lambda)) = 1.$$

◁ **Éléments de preuve.**
Le calcul de $\det(E(i,j))$ correspond au calcul du determinant de la base canonique sur laquelle on a opéré une transposition.
Les autres sont triangulaires. ▷

Corollaire 26.2.3 (Effet des opérations élémentaires sur le déterminant)

(i) Échanger les lignes i et j change le signe du déterminant ;

(ii) Multiplier une ligne par un scalaire λ multiplie le déterminant par ce scalaire

(iii) Faire une combinaison $L_i \leftarrow L_i - \lambda L_j$ ne modifie pas le déterminant (attention à ne pas mettre de coefficient devant L_i).

(iv) De même pour les opérations sur les colonnes.

◁ **Éléments de preuve.**
Ces opérations correspondent à des multiplications par des matrices de codage d'opérations. ▷

Méthode 26.2.4 (Calcul du déterminant par la méthode du pivot)

Ainsi, on peut adapter la méthode du pivot pour le calcul du déterminant :

- Comme usuellement, faire des opérations pour échelonner la matrice, en écrivant des égalités entre les déterminants des différentes matrices, mais :
  - $\ast$ Changer le signe lorsqu'on fait un échange de lignes
  - $\ast$ Compenser en divisant à l'extérieur par λ lorsqu'on multiplie une ligne par λ.
- Une fois ramené à une matrice triangulaire, on calcule son déterminant en effectuant le produit des coefficients diagonaux.

Exemple 26.2.5

1. Déterminant de $\begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$
2. Déterminant de $\begin{pmatrix} a & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & a \end{pmatrix}$

II.2 Calcul par blocs

Dans certaines configurations, on peut se ramener à des déterminants plus petits :

Proposition 26.2.6 (Déterminant d'une matrice triangulaire par blocs)

Soit T une matrice triangulaire par blocs, c'est-à-dire de la forme

$$T = \begin{pmatrix} A_1 & \bullet & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bullet \\ 0 & \cdots & 0 & A_k \end{pmatrix}$$

où les A_k sont des matrices carrées. Alors

$$\det(T) = \det(A_1)\cdots\det(A_k).$$

◁ **Éléments de preuve.**

Effectuer un pivot pour échelonner chaque A_i (en agissant sur toute la ligne, donc sur ce qu'il y a derrière les A_i). Ces pivots sont indépendants. Comptabiliser ensuite les différents types d'opération et conclure. ▷

Exemples 26.2.7

1. Déterminant de $\begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 2 & 2 & 8 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -1 & 2 \end{pmatrix}$
2. Déterminant de $\begin{pmatrix} A_1 & \bullet & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bullet \\ 0 & \cdots & 0 & A_k \end{pmatrix}$, où $A_i = \begin{pmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{pmatrix}$

3. On peut combiner avec les opérations du pivot, pour éviter la dernière opération (ou les deux dernières), en se ramenant ainsi au calcul d'un déterminant d'ordre 2 ou 3 (par Sarrus).

II.3 Développements suivant une ligne ou une colonne

Sans doute une des techniques les plus importantes, lorsqu'une ligne ou une colonne contient beaucoup de 0. Elle est efficace lorsqu'il y a beaucoup de symétries dans l'expression, pour construire une récurrence, et si ce n'est pas possible, elle peut être combinée à d'autres méthodes pour les calculs des déterminants plus petits obtenus.
Pour exprimer la formule, nous définissons :

Définition 26.2.8 (Mineurs, cofacteurs, comatrice)

Soit $M \in \mathcal{M}_n(\mathbb{K})$.

- Le mineur de position (i,j) de M est le déterminant de la matrice $\widetilde{M}_{i,j}$ obtenue en supprimant de M la i-ième ligne et la j-ième colonne. On note $\Delta_{i,j}(M)$ ce mineur
- Le cofacteur de position (i,j) de M est le scalaire $(-1)^{i+j}\Delta_{i,j}(M)$.
- La comatrice de M est la matrice $\mathrm{Com}(M) = ((-1)^{i+j}\Delta_{i,j}(M))_{1\leqslant i,j\leqslant n}$, c'est à dire la matrice des cofacteurs de M.

Nous obtenons alors la formule de développement :

Théorème 26.2.9 (Développement suivant une colonne)

Soit $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, et $j \in [\![1,n]\!]$. Alors

$$\det(M) = \sum_{i=1}^{n} (-1)^{i+j} m_{i,j}\Delta_{i,j}(M).$$

◁ **Éléments de preuve.**

- Par linéarité, se ramener à une somme de déterminants n'ayant qu'un coefficient non nul (qu'on peut prendre égal à 1) sur la colonne j.
- Ramener ce coefficient en haut à gauche par des permutations de lignes et colonnes
- La matrice obtenue est triangulaire par bloc, avec un tout petit bloc et un gros bloc. À quoi correspond le gros bloc ?

▷

Théorème 26.2.10 (Développement suivant une ligne)

Soit $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, et $i \in [\![1,n]\!]$. Alors

$$\det(M) = \sum_{j=1}^{n} (-1)^{i+j} m_{i,j}\Delta_{i,j}(M).$$

◁ **Éléments de preuve.**

Transposer ! ▷

Corollaire 26.2.11 (Expression de l'inverse par la comatrice, Cayley)
Soit $M \in \mathcal{M}_n(\mathbb{K})$. *Alors*

$$M\mathrm{Com}(M)^\top = \mathrm{Com}(M)^\top M = \det(M)I_n.$$

En particulier, M *est inversible si et seulement si* $\det(M) \neq 0$, *et dans ce cas,*

$$M^{-1} = \frac{1}{\det(M)}\mathrm{Com}(M)^\top.$$

◁ **Éléments de preuve.**
Les coefficients diagonaux de $M\mathrm{Com}(M)^\top$ traduisent la formule de développement par une ligne
Les autres traduisent la formule de développement par une ligne d'une matrice N obtenue de M en remplaçant l'une de ses lignes par une autre : il y a donc une ligne en double. Qu'en conclure pour N ? ▷

Si la formule elle-même s'avère assez inefficace pour les calculs pratiques pour $n > 2$, la caractérisation de l'inversibilité par la non nullité du déterminant est d'une très grande utilité.
Les formules de Cramer (décrites par ce dernier avant la formule d'inversion) sont une conséquence de la formule d'inversion :

Corollaire 26.2.12 (Cramer)
Soit à résoudre le système $AX = B$, *de l'inconnue* $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, *la matrice* A *étant supposée inversible.*
Soit $A_1, \dots, A_n$ *les colonnes de* A. *On a alors, pour tout* $k \in [\![1, n]\!]$:

$$x_k = \frac{\det(A_1, \dots, A_{k-1}, B, A_{k+1}, \dots A_n)}{\det(A)}.$$

◁ **Éléments de preuve.**
Écrire $X = A^{-1}B$, et exprimer A^{-1} à l'aide de la comatrice, faire le produit avec B, et reconnaître pour chaque coefficient un développement suivant une colonne d'un déterminant. ▷

Voyons maintenant quelques exemples de calculs de déterminant par la méthode de développement suivant une colonne.

Exemples 26.2.13

1. Déterminant de la matrice tridiagonale $\begin{pmatrix} a+b & a & 0 & \cdots & 0 \\ b & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & a \\ 0 & \cdots & 0 & b & a+b \end{pmatrix}$.
2. On peut combiner pivot de Gauss et développement suivant les lignes ou colonnes, en annulant d'abord un grand nombre de termes par les opérations élémentaires avant de développer. Nous illustrons ceci sur le calcul du déterminant de Vandermonde :

$$V(x_1, \dots, x_n) = \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ x_1 & x_2 & \cdots & x_{n-1} & x_n \\ x_1^2 & x_2^2 & \cdots & x_{n-1}^2 & x_n^2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_{n-1}^{n-1} & x_n^{n-1}. \end{vmatrix}$$

Cet exemple est un peu plus qu'un exemple, nous le consignons dans la proposition suivante :

Proposition 26.2.14 (Déterminant de Vandermonde)

Le déterminant de Vandermonde, défini dans l'exemple ci-dessus, est donné par :

$$V(x_1,\dots,x_n) = \prod_{1\leqslant i<j\leqslant n}(x_j - x_i).$$

◁ **Éléments de preuve.**
Effectuer les opérations $L_k \leftarrow L_k - x_1 L_{k-1}$. ▷

Note Historique 26.2.15

Malgré son nom, ce déterminant n'apparaît à aucun moment dans l'œuvre du mathématicien français Alexandre-Théophile Vandermonde.

Nous voyons dans la section suivante une autre façon de calculer ce déterminant de Vandermonde.

II.4 Caractère polynomial du déterminant

La formule explicite du déterminant nous permet d'affirmer :

Proposition 26.2.16 (Polynomialité du déterminant)

L'application $M \mapsto \det(M)$ est polynomiale en chacune des coordonnées de M. Il s'agit plus précisément d'une fonction polynomiale en les n^2 variables définies par les coordonnées de M, vérifiant :

- $\det$ *est homogène de degré total n en ses n^2 coordonnées ;*
- *les degrés partiels par rapport à chacune des variables $m_{i,j}$ sont* 1 *;*
- *les degrés partiels par rapport à l'ensemble des variables constituant une même ligne ou une même colonne sont* 1

◁ **Éléments de preuve.**
Cela provient de la formule de développement par le groupe symétrique. ▷

Exemple 26.2.17

Calcul de $V_n(x_1,\dots,x_n)$.

◁ **Éléments de preuve.**
On retrouve le calcul du déterminant de Vandermonde : remplacer la variable x_n par une indéterminée formelle X. On obtient un polynôme de degré au plus $n-1$, dont on connaît $n-1$ racines. Il ne reste qu'à trouver le coefficient dominant, qu'on obtient en faisant un développement (suivant quoi ?) ▷

Remarque 26.2.18

La matrice de Vandermonde correspond à la matrice du système permettant de trouver les coefficients dans la base canonique du polynôme d'interpolation de Lagrange aux points $x_1,\dots,x_n$, associés à des valeurs $y_1,\dots,y_n$ arbitraires.
Ainsi, la non nullité de ce déterminant peut aussi être vue comme conséquence de l'existence et l'unicité du polynôme d'interpolation, traduisant le fait que le système associé est un système de Cramer.

27 Espaces préhilbertiens réels

Soit une multiplicité vectorielle
Un corps opère, seul, abstrait, commutatif
Le dual reste loin, solitaire et plaintif
Cherchant l'isomorphie et la trouvant rebelle.

Soudain bilinéaire a jailli l'étincelle
D'ou naît l'opérateur deux fois distributif.

Dans les rets du produit tous les vecteurs captifs
Ont célébré sans fin la structure plus belle.

Mais la base a troublé cet hymne aérien
Les vecteurs éperdus ont des coordonnées
Cartan ne sait que faire et n'y comprend plus rien

Et c'est la fin. Vecteurs, opérateurs foutus
Une matrice immonde expire. Le corps nu
Fuit en lui-même, au sein de lois qu'il s'est données.

(André Weil)

Un espace préhilbertien réel est un espace vectoriel muni d'un produit scalaire. Cet outil permet de parler d'orthogonalité, et donc d'introduire un certain nombre de concepts permettant de généraliser la géométrie euclidienne du plan à des situations plus abstraites.
Pour commencer nous introduisons donc cette notion abstraite de produit scalaire, ce qui nous oblige à faire quelques rappels sur les formes bilinéaires, déjà rencontrées dans le chapitre précédent.

Tous les espaces vectoriels considérés dans ce chapitre sont des espaces vectoriels sur $\mathbb{R}$.

I Produits scalaires

I.1 Formes bilinéaires

Les résultats de cette sous-section sont valables dans un cadre plus général que les espaces sur $\mathbb{R}$, mais toute la suite du chapitre nécessitant de travailler avec des espaces vectoriels sur $\mathbb{R}$, nous nous limitons également à ce cadre dans les rappels qui suivent.

On rappelle la définition d'une forme bilinéaire :

Définition 27.1.1 (Forme bilinéaire)

Soit E un espace vectoriel sur $\mathbb{R}$. Une forme bilinéaire φ sur E est une application $\varphi : E \times E \to \mathbb{R}$, linéaire par rapport à chaque facteur, l'autre étant fixé, c'est-à-dire :

(i) $\forall (x,y,z) \in E^3, \quad \forall \lambda \in \mathbb{R}, \quad \varphi(\lambda x + y, z) = \lambda \varphi(x,z) + \varphi(y,z)$

(ii) $\forall (x,y,z) \in E^3, \quad \forall \lambda \in \mathbb{R}, \quad \varphi(x, \lambda y + z) = \lambda \varphi(x,y) + \varphi(x,z)$.

On rappelle la propriété de bilinéarité généralisée

Lemme 27.1.2 (Bilinéarité généralisée)

Soit φ une forme bilinéaire sur E. Soit $(k,\ell) \in (\mathbb{N}^*)^2$, et soit $(x_1,\dots,x_k,y_1,\dots,y_\ell) \in E^{k+\ell}$, et $(\lambda_1,\dots,\lambda_k,\mu_1,\dots,\mu_\ell) \in \mathbb{K}^{k+\ell}$. Alors

$$\varphi\left(\sum_{i=1}^{k} \lambda_i x_i, \sum_{j=1}^{\ell} \mu_j y_j\right) = \sum_{i=1}^{k}\sum_{j=1}^{\ell} \lambda_i \mu_j \varphi(x_i, y_j).$$

Définition 27.1.3 (Ensemble des formes bilinéaires)

On note $\mathcal{B}(E)$ l'ensemble des formes bilinéaires de E

Nous avons déjà vu le résultat suivant :

Proposition 27.1.4 (Structure de $\mathcal{B}(E)$)

L'ensemble $\mathcal{B}(E)$ est un espace vectoriel sur $\mathbb{R}$.

Voici des exemples particulièrement importants :

Exemples 27.1.5 (Formes bilinéaires)

1. $\mathrm{cov} : (X,Y) \mapsto \mathrm{cov}(X,Y)$ sur le $\mathbb{R}$-espace vectoriel $\mathcal{L}^2$ des variables aléatoires réelles discrètes admettant un moment d'ordre 2 ?
2. $\varphi : (P,Q) \mapsto \displaystyle\int_a^b P(x)Q(x)\,\mathrm{d}x$ sur $\mathbb{R}[X]$, ou sur $\mathcal{C}^0([a,b])$?
3. $\varphi : \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) \mapsto \displaystyle\sum_{i=1}^{n} x_i y_i$ sur $\mathbb{R}^n$?
4. $p_1 : (x,y) \mapsto x$ sur $\mathbb{R}$?
5. $\varphi : (X,Y) \mapsto X^\top A Y$, sur $\mathbb{R}^n$, $A \in \mathcal{M}_n(\mathbb{R})$?

Ce dernier exemple est très important, car si E est de dimension finie, toute forme bilinéaire peut être représenté de cette forme après choix d'une base de E. C'est ce que nous étudions maintenant.

Définition 27.1.6 (Forme quadratique)

Une forme quadratique q sur E est une application $q : E \to \mathbb{R}$ telle qu'il existe φ bilinéaire sur E telle que pour tout $x \in \mathbb{R}$, $\varphi(x,x) = q(x)$. On notera q_φ la forme quadratique associée à φ.

La forme quadratique q est en général associée à plusieurs formes linéaires. Par exemple les formes linéaires sur $\mathbb{R}^2$ définies par

$$\varphi((x,y),(x',y')) = xx' + yy' + 2xy' \qquad \text{et} \qquad \psi((x,y),(x',y')) = xx' + yy' + xy' + x'y$$

définissent la même forme quadratique q.

I.2 Matrice d'une forme bilinéaire

Nous supposons jusqu'à la fin de ce paragraphe que E est de dimension finie n.

Définition 27.1.7 (Matrice associée à une forme bilinéaire)

Soit φ une forme bilinéaire sur E, et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. Alors on définit la matrice de φ relativement à la base $\mathcal{B}$ par :

$$\mathrm{Mat}_{\mathcal{B}}(\varphi) = (\varphi(e_i, e_j))_{1 \leqslant i,j \leqslant n}.$$

Exemples 27.1.8

1. Matrice des variances-covariances si $X_1, \dots, X_n$ sont linéairement indépendants.
2. Matrice dans la base canonique de $\varphi : (P, Q) \mapsto \displaystyle\int_0^1 P(x)Q(x)\,\mathrm{d}x$ sur $\mathbb{R}_n[X]$.
3. Matrice du produit scalaire canonique de $\mathbb{R}^n$ dans la base canonique; dans la base $\mathcal{B} = (b_1, \dots, b_n)$, où b_i est constitué de 1 sur les coordonnées 1 à i, et de 0 ailleurs.

Théorème 27.1.9 (Expression matricielle de $\varphi(x, y)$)

Soit φ une forme bilinéaire sur E, et $\mathcal{B}$ une base de E. Alors, pour tout $(x, y) \in E^2$,

$$\varphi(x, y) = [x]_{\mathcal{B}}^{\top} \mathrm{Mat}_{\mathcal{B}}(\varphi)[y]_{\mathcal{B}} = X^{\top} M Y,$$

où X et Y sont les vecteurs colonnes représentant les coordonnées de x et y dans la base $\mathcal{B}$, et M est la matrice de φ relativement à cette même base $\mathcal{B}$.

◁ **Éléments de preuve.**
On l'a vu au détour d'un exemple dans le chapitre précédent. ▷

Une base $\mathcal{B}$ étant fixée, cette relation caractérise d'ailleurs la matrice $\mathrm{Mat}_{\mathcal{B}}(\varphi)$:

Théorème 27.1.10 (Caractérisation de $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ par la relation $\varphi(x, y) = X^{\top} M Y$)

Soit $M \in \mathcal{M}_n(\mathbb{K})$. La relation $\varphi(x, y) = X^{\top} M Y$ caractérise la matrice de φ relativement à la base $\mathcal{B}$: si

$$\forall (x, y) \in E^2, \quad \varphi(x, y) = [x]_{\mathcal{B}}^{\top} M [y]_{\mathcal{B}},$$

alors $M = \mathrm{Mat}_{\mathcal{B}}(\varphi)$.

◁ **Éléments de preuve.**
Prendre pour X et Y des vecteurs de la base canonique, afin de montrer que $m_{i,j} = \varphi(b_i, b_j)$. ▷

Exemple 27.1.11

L'égalité $\mathrm{Mat}_{bc}(\varphi) = I_n$ se traduit par $\varphi(X, Y) = X^{\top} I_n Y = X^{\top} Y$; c'est le produit scalaire usuel.

Remarque 27.1.12

À quelle condition nécessaire et suffisante sur leur représentation matricielle relativement à une base donnée deux formes bilinéaires φ et ψ définissent-elles la même forme quadratique q ?

Corollaire 27.1.13

L'application

$$\Phi : \begin{array}{ccc} \mathcal{B}(E) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ \varphi & \longmapsto & \mathrm{Mat}_{\mathcal{B}}(\varphi) \end{array}$$

est un isomorphisme.

◁ **Éléments de preuve.**

Vérifications simples pour la linéarité. L'injectivité découle de 1.10, la surjectivité s'obtient en définissant φ par rigidité. ▷

Corollaire 27.1.14 (dimension de $\mathcal{B}(E)$)

Si E est de dimension finie n, alors $\mathcal{B}(E)$ est de dimension finie, et

$$\dim(\mathcal{B}(E)) = n^2.$$

Comme pour les applications linéaires, les changements de base s'expriment facilement par des opérations matricielles.

Théorème 27.1.15 (Formule de changement de base pour les formes bilinéaires)

Soit E un espace vectoriel sur $\mathbb{R}$ de dimension finie. Soit φ une forme bilinéaire sur E, et soit $\mathcal{C}$ et $\mathcal{D}$ deux bases de E. Soit P la matrice de passage de $\mathcal{C}$ à $\mathcal{D}$. Alors

$$\mathrm{Mat}_{\mathcal{D}}(\varphi) = P^{\top}\mathrm{Mat}_{\mathcal{C}}(\varphi)P.$$

◁ **Éléments de preuve.**

Calculer $\varphi(x,y)$ matriciellement dans $\mathcal{C}$, puis opérer le changement de base sur les vecteurs. ▷

Exemple 27.1.16

Explicitation du changement de la base canonique à la base $\mathcal{B}$ dans l'exemple 27.1.8, pour le produit scalaire canonique.

I.3 Formes bilinéaires symétriques, définies, positives

Définition 27.1.17 (Symétrie, positivité, caractère défini)

Soit φ une forme bilinéaire sur un espace vectoriel E.

1. On dit que φ est symétrique si : $\forall (x,y) \in E^2, \quad \varphi(x,y) = \varphi(y,x)$.
2. On dit que φ est positive si : $\forall x \in E, \quad \varphi(x,x) \geqslant 0$, c'est-à-dire $\mathrm{Im}(q_\varphi) \subset \mathbb{R}_+$
3. On dit que φ est négative si : $\forall x \in E, \quad \varphi(x,x) \leqslant 0$, c'est-à-dire $\mathrm{Im}(q_\varphi) \subset \mathbb{R}_-$
4. On dit que φ est définie si : $\forall x \in E, \quad \varphi(x,x) = 0 \Longleftrightarrow x = 0$, i.e. $q_\varphi(x) \neq 0$ pour tout $x \neq 0$.

Proposition 27.1.18 (positivité des formes définies)

Soit φ une forme définie. Alors φ est soit positive soit négative.

◁ **Éléments de preuve.**
Si q_φ change de signe entre x et y, paramétrer le segment $[x,y]$ et utiliser le TVI. ▷

Pour cette raison, on parle assez rarement de forme définie, cette propriété étant indissociable d'une propriété de positivité ou de négativité. On parlera alors de forme définie positive, ou définie négative.

Proposition 27.1.19 (Caractérisation matricielle de la symétrie)
Soit φ une forme bilinéaire sur un espace vectoriel E de dimension finie. Les propriétés suivantes sont équivalentes :

(i) φ est symétrique ;

(ii) Il existe une base de E telle que $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ est une matrice symétrique ;

(iii) Pour tout base $\mathcal{B}$ de E, $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ est une matrice symétrique.

◁ **Éléments de preuve.**
Vérifications faciles, en tournant en sens inverse. ▷

Exemples 27.1.20

1. $\varphi : (x,y) \mapsto xy$ sur $\mathbb{R}$ est symétrique, définie, positive.
2. cov est symétrique positive, mais pas définie.
3. Le produit scalaire canonique sur $\mathbb{R}^n$ est symétrique, défini, positif.
4. $\varphi : (f,g) \mapsto \displaystyle\int_a^b f(t)g(t)\,\mathrm{d}t$ sur $\mathcal{C}^0([a,b])$ est symétrique, définie et positive.

Proposition 27.1.21 (Symétrie et forme quadratique ; formule de polarisation)
Soit q une forme quadratique sur E. Il existe une unique forme bilinéaire symétrique sur E telle que $q = q_\varphi$, explicitement donnée par l'expression suivante, dite formule de polarisation :

$$\varphi(x,y) = \frac{1}{2}\big(q(x+y) - q(x) - q(y)\big).$$

◁ **Éléments de preuve.**
Montrer la bilinéarité de φ en remarquant que $\varphi(x,y)$ se décrit de façon symétrique en fonction de $\psi(x,y)$ et $\psi(y,x)$.
Si θ est symétrique et associée à q, elle vérifie forcément la formule de polarisation, d'où l'unicité. ▷

I.4 Produits scalaires

Dans cette section, E désigne un espace vectoriel sur $\mathbb{R}$.

Définition 27.1.22 (Produit scalaire)
Un produit scalaire sur E est une forme bilinéaire symétrique, définie et positive.

Remarque 27.1.23
La symétrie et la linéarité par rapport à la première variable entraînent la linéarité par rapport à la seconde variable.

Voici les exemples usuels (les 3 premiers), desquels on peut dériver d'autres exemples. Ces exemples sont à considérer comme des exemples du cours.

Exemples 27.1.24 (produits scalaires)

1. $(f,g) \mapsto \int_a^b f(t)g(t)\,\mathrm{d}t$ est un produit scalaire sur $\mathcal{C}^0([a,b])$ (ou sur $\mathbb{R}[X]$)
2. Le produit scalaire canonique de $\mathbb{R}^n$ est un produit scalaire au sens de cette définition.
3. Sur $\mathcal{M}_{n,p}(\mathbb{R})$: $(A,B) \mapsto \operatorname{tr}(A^\top B)$.
4. Un autre produit scalaire sur $\mathbb{R}^n$: $(X,Y) \mapsto X^\top MY$, où $M = (\min(i,j))_{1\leqslant i,j\leqslant n}$.

Ainsi, on n'a pas unicité d'un produit scalaire sur un espace vectoriel E.

Notation 27.1.25 (Notations fréquentes pour le produit scalaire)

Soit φ un produit scalaire sur E on note souvent :

$$\varphi(x,y) = \langle x,y\rangle \qquad \text{ou} \qquad \varphi(x,y) = (x|y).$$

Par ailleurs, on note $\|x\| = \sqrt{\varphi(x,x)} = \sqrt{\langle x,x\rangle}$ la racine de la forme quadratique associée, cette expression étant bien définie par positivité de φ.

Soit φ un produit scalaire, noté $\varphi(x,y) = \langle x,y\rangle$. Alors en particulier, φ est une forme bilinéaire, et si $\mathcal{B} = (e_1,\dots,e_n)$ est une base de E :

$$\operatorname{Mat}_{\mathcal{B}}(\varphi) = (\langle e_i,e_j\rangle)_{1\leqslant i,j\leqslant n}$$

Proposition 27.1.26

La matrice d'un produit scalaire dans une base quelconque est symétrique. La réciproque est fausse.

◁ **Éléments de preuve.**

C'est la symétrie du produit scalaire ! ▷

Un produit scalaire étant une forme bilinéaire symétrique, on dispose aussi de la formule de polarisation, s'écrivant ici :

$$\langle x,y\rangle = \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right).$$

Cette formule n'est autre que le développement d'un carré :

$$\|x+y\|^2 = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2.$$

On généralise facilement par la formule de bilinéarité généralisée :

Proposition 27.1.27 (Développement d'un carré)

Plus généralement, étant donné $(x_1,\dots,x_n) \in \mathrm{E}^n$,

$$\|x_1 + \cdots + x_n\|^2 = \sum_{i=1}^n \|x_i\|^2 + 2\sum_{1\leqslant i<j\leqslant n} \langle x_i,x_j\rangle.$$

◁ **Éléments de preuve.**

Utiliser la formule de multilinéarité généralisée, puis la symétrie pour regrouper les termes 2 par 2. ▷

L'inégalité de Cauchy-Schwarz vue dans le cadre scalaire se généralise à tout produit scalaire.

Théorème 27.1.28 (Inégalité de Cauchy-Schwarz pour un produit scalaire)
Soit φ une forme symétrique positive E, et $q : x \mapsto \varphi(x,x)$ la forme quadratique associée. Alors pour tout $(x,y) \in E^2$,

$$\varphi(x,y)^2 \leqslant q(x)q(y),$$

avec égalité si et seulement s'il existe des scalaires λ et μ non tous deux nuls tels que $q(\lambda x + \mu y) = 0$.
Si φ est un produit scalaire noté $\langle .,. \rangle$, l'inégalité de Cauchy-Schwarz se réécrit :

$$|\langle x,y \rangle| \leqslant \|x\| \cdot \|y\|,$$

avec égalité si et seulement si x et y sont colinéaires.

◁ **Éléments de preuve.**
Considérer $\lambda \mapsto q(\lambda x + y)$, polynôme de degré 2 (presque toujours, ne pas oublier le cas particulier dans lequel ce n'est pas vrai). Quel est son signe ? Que dire de son discriminant ?
Cas d'égalité : cela signifie que $\Delta = 0$, que peut-on en déduire ? ▷

Exemples 27.1.29

1. Inégalité de Cauchy-Schwarz numérique :
$$\forall X = (x_1, \dots, x_n),\ Y = (y_1, \dots, y_n) \in \mathbb{R}^n, \quad |\langle X,Y \rangle| \leqslant \|X\| \cdot \|Y\|,$$
à savoir :
$$\left|\sum_{i=1}^n x_i y_i\right| \leqslant \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^n y_i^2\right)^{\frac{1}{2}},$$
ou encore :
$$\left(\sum_{i=1}^n x_i y_i\right)^2 \leqslant \left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n y_i^2\right),$$
avec égalité si et seulement si X et Y sont colinéaires.
2. Inégalité de Cauchy-Schwarz intégrale :
$$\forall (f,g) \in \mathcal{C}^0([a,b]), \quad \left|\int_a^b f(t)g(t)\ \mathrm{d}t\right| \leqslant \left(\int_a^b f(t)^2\ \mathrm{d}t\right)^{\frac{1}{2}} \left(\int_a^b g(t)^2\ \mathrm{d}t\right)^{\frac{1}{2}},$$
avec égalité si et seulement si $g = 0$, ou s'il existe λ tel que $f = \lambda g$.
3. Si X et Y sont deux variables aléatoires admettant une variance, on obtient l'inégalité de Cauchy-Schwarz pour les variables aléatoires :
$$|\mathrm{cov}(X,Y)| \leqslant \sigma(X)\sigma(Y),$$
avec égalité si et seulement s'il existe λ et μ tels que $V(\lambda X + \mu Y) = 0$, donc tels que $\lambda X + \mu Y$ soit constante presque sûrement, donc si et seulement si X et Y sont reliés par une relation affine. Ce cas d'égalité correspond au cas où le coefficient de corrélation vérifie $|\rho(X,Y)| = 1$. On a ainsi prouvé une propriété admise lors du cours de probabilité.

I.5 Normes euclidiennes

Le but de cette section est de montrer que $x \mapsto \|x\|$ définit une norme vectorielle.

Définition 27.1.30 (Norme)

Une norme sur un espace vectoriel E est une application $N : E \to \mathbb{R}$ telle que :

(i) $\forall x \in E, \quad N(x) = 0 \Longleftrightarrow x = 0$ (séparation)

(ii) $\forall \lambda \in \mathbb{R}, \quad \forall x \in E, \quad N(\lambda x) = |\lambda| N(x)$ (absolue homogénéité)

(iii) $\forall (x,y) \in E^2, \quad N(x+y) \leqslant N(x) + N(y)$ (inégalité triangulaire)

Proposition 27.1.31 (Positivité des normes)

Si N est une norme sur E, alors pour tout $x \in E$, $N(x) \geqslant 0$.

◁ **Éléments de preuve.**

Utiliser tous les points de la définition, en considérant $N(x + (-x))$. ▷

Proposition 27.1.32

L'application $x \mapsto \|x\|$ est une norme sur E.

◁ **Éléments de preuve.**

L'inégalité triangulaire s'obtient avec Cauchy-Schwarz, en comparant les carrés. ▷

Définition 27.1.33 (Norme euclidienne associée au produit scalaire)

Soit $\langle \cdot, \cdot \rangle$ un produit scalaire sur E. L'application $\| \cdot \|$ est appelée norme euclidienne associée au produit scalaire $\langle \cdot, \cdot \rangle$.

Définition 27.1.34 (Norme euclidienne)

Soit N une norme sur E. La norme N est euclidienne si et seulement s'il existe un produit scalaire $\langle \cdot, \cdot \rangle$, dont N est la norme euclidienne associée

La formule de polarisation donne l'unique candidat possible pour ce produit scalaire :

Méthode 27.1.35 (Montrer qu'une norme est euclidienne)

On définit l'application φ, unique candidat possible :

$$\varphi(x,y) = \frac{1}{2}(N(x+y)^2 - N(x)^2 - N(y)^2).$$

On vérifie ensuite que φ est une application bilinéaire. En cas de bilinéarité, le caractère symétrique défini positif est immédiat par définition de N. Le point à étudier de plus près est donc la bilinéarité : c'est là que se trouve l'éventuelle obstruction.

Exemple 27.1.36

Soit $n \geqslant 2$, et $N : (x_1, \dots, x_n) \mapsto \max(|x_1|, \dots, |x_n|)$ est une norme sur $\mathbb{R}^n$, mais n'est pas une norme euclidienne.

Ainsi, toutes les normes ne sont pas euclidiennes !

I.6 Espaces préhilbertiens réels, espaces euclidiens

Définition 27.1.37 (Espace préhilbertien réel)
Un espace préhilbertien réel $(E, \langle\cdot,\cdot\rangle)$ est un espace vectoriel E sur $\mathbb{R}$, muni d'un produit scalaire $\langle\cdot,\cdot\rangle$.

S'il n'y a pas d'ambiguïté sur le produit scalaire, on parlera plus simplement de l'espace préhilbertien E, au lieu de $(E, \langle\cdot,\cdot\rangle)$.

Définition 27.1.38 (Espace euclidien)
Un espace euclidien est un espace préhilbertien réel de dimension finie.

Par exemple $\mathbb{R}^n$ muni du produit scalaire canonique est un espace euclidien. $\mathbb{R}_n[X]$ muni d'un produit scalaire intégral est un espace euclidien. En revanche, $\mathbb{R}[X]$ muni du même produit scalaire n'est qu'un espace préhilbertien réel, ainsi que $\mathcal{C}^0([a,b])$ muni du produit scalaire intégral.

II Orthogonalité

Dans toute cette section, on considère $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. Il sera précisé euclidien dans certains cas.

II.1 Vecteurs orthogonaux

Définition 27.2.1 (Vecteurs orthogonaux)
Soit $(x,y) \in E$ deux vecteurs de E. On dit que x et y sont orthogonaux, et on note $x \perp y$, si $\langle x, y\rangle = 0$.

Exemples 27.2.2

1. $\forall x \in E, \quad x \perp 0$.
2. Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. Alors pour tout $i \neq j$, $e_i \perp e_j$.
3. $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \perp \begin{pmatrix} -4 \\ -3 \\ 2 \\ 1 \end{pmatrix}$, pour le produit scalaire usuel.
4. $x \mapsto \sin \pi x$ et $x \mapsto \cos \pi x$ sont orthogonaux pour le produit scalaire $\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\, \mathrm{d}t$ sur $\mathcal{C}^0([-1,1])$.

Vous savez depuis longtemps déterminer si deux vecteurs du plan euclidien sont orthogonaux, par exemple en vérifiant qu'ils forment un triangle rectangle. Cette caractérisation des triangles rectangles se généralise :

Théorème 27.2.3 (Théorème de Pythagore)
Soit E un espace préhilbertien réel. Soit $(x,y) \in E^2$. Alors

$$x \perp y \iff \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

◁ **Éléments de preuve.**
Utiliser la formule de développement d'un carré, et comparer. ▷

On peut définir plus généralement :

Définition 27.2.4 (Famille orthogonale, orthonormale)

1. Une famille $(x_i)_{i\in I}$ est orthogonale si pour tout $i \neq j$ de I, $x_i \perp x_j$.
2. On dit que la famille $(x_i)_{i\in I}$ est orthonormale (ou orthonormée) si et seulement si elle est orthogonale, et que pour tout $i \in I$, $\|x_i\| = 1$.

Le théoèrme de Pythagore se généralise alors ainsi (pour une famille finie), mais on perd ici l'équivalence (ce n'est pas une caractérisaion de l'orthogonalité) :

Théorème 27.2.5 (Théorème de Pythagore généralisé)

Soit $(x_1, \dots, x_n)$ une famille orthogonale de E. Alors

$$\left\|\sum_{i=1}^{n} x_i\right\|^2 = \sum_{i=1}^{n} \|x_i\|^2.$$

◁ **Éléments de preuve.**
Encore un carré à développer avec la formule idoine. ▷

Une propriété qui facilite souvent la justification de la liberté :

Théorème 27.2.6 (Liberté des familles orthogonales)

Soit $\mathcal{F}$ une famille orthogonale ne contenant pas le vecteur nul. Alors $\mathcal{F}$ est une famille libre.

◁ **Éléments de preuve.**
Si une CL est nulle, en prendre le produit scalaire avec chaque vecteur de la famille. ▷

En particulier, toute famille orthonormale est libre.

Corollaire 27.2.7 (base orthonormale)

*On suppose E de dimension finie n (*i.e.* E euclidien). Soit $(e_1, \dots, e_n)$ une famille orthonormale de E. Alors $(e_1, \dots, e_n)$ est une base de E. On dit qu'il s'agit d'une base orthonormale, et on abrège en b.o.n..*

◁ **Éléments de preuve.**
Elle est libre de cardinal maximal. ▷

Théorème 27.2.8 (Existence d'une base orthonormale)

Soit E un espace euclidien. Alors E admet une b.o.n.

◁ **Éléments de preuve.**
Par récurrence, en considérant l'hyperplan $\mathrm{Ker}(\varphi_{b_1})$, où $\varphi(b_1(x)) = \langle b_1, x\rangle$. ▷

II.2 Sous-espaces orthogonaux

Définition 27.2.9 (Sous-espaces orthogonaux)

Soit E un espace préhilbertien réel.

1. Soit $x \in E$, et F un sous-espace vectoriel de E. On dit que x est orthogonal à F si pour tout $y \in F$, $x \perp y$. On note $x \perp F$, ou $x \in F^\perp$ comme on le verra plus tard.
2. Soit F et G deux sous-espaces vectoriels de E. On dit que F et G sont orthogonaux si et seulement si :
$$\forall x \in F, \quad \forall y \in G, \quad x \perp y.$$

Exemples 27.2.10

1. Dans $\mathbb{R}^3$, le plan (Oxy) et la droite (Oz).
2. Dans $\mathbb{R}^3$, le plan d'équation $x+y+z=0$, et la droite engendrée par $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.
3. Dans $\mathcal{C}^0([-1,1])$ muni du produit scalaire usuel, le sous-espace P des fonctions paires, et le sous-espace I des fonctions impaires.

Proposition 27.2.11 (orthogonalité et somme directe)

Soit F, G et H des sous-espaces vectoriels de E.

(i) Si $F \perp G$, alors la somme $F+G$ est directe. On écrit $F \overset{\perp}{\oplus} G$.

(ii) Si $H \perp F$ et $H \perp G$, alors $H \perp F+G$.

(iii) Si $(H_i)_{i \in I}$ est une famille finie de sous-espaces 2 à 2 orthogonaux, alors la somme $\bigoplus_{i \in I} H_i$ est directe.

◁ **Éléments de preuve.**

Un vecteur de l'intersection vérifie $x \perp x = 0$. De même dans le dernier point (il faut juste savoir de quelle intersection on parle au juste !) ▷

Du fait de la bilinéarité du produit scalaire, il n'est pas nécessaire d'établir l'orthogonalité de toutes les paires de vecteurs de E et F pour obtenir l'orthogonalité des deux espaces. Il suffit en fait de verifier cette orthogonalité sur des vecteurs de familles génératrices.

Proposition 27.2.12 (Caractérisation de l'orthogonalité par les familles génératrices)

Soit $(x_i)_{i \in I}$ et $(y_j)_{j \in J}$ deux familles de E. Alors

$$\mathrm{Vect}(x_i, i \in I) \perp \mathrm{Vect}(y_j, j \in J) \text{ si et seulement si } \forall i \in I, \forall j \in J, \quad x_i \perp y_j.$$

◁ **Éléments de preuve.**

Dans un sens, c'est clair. Dans l'autre, écrire des vecteurs de ces espaces dans les familles génératrices donnés, et calculer leur produit scalaire par bilinéarité généralisée. ▷

Définition 27.2.13 (orthogonal d'une partie de E)

Soit X une partie de E. On note $X^\perp = \{x \in E \mid x \perp X\}$, l'ensemble des vecteurs orthogonaux à tout vecteur de X. L'ensemble $X^\perp$ est appelé l'orthogonal de X.

Proposition 27.2.14 (orthogonal d'une union)

(i) *Soit X et Y deux parties de E. Alors $(X \cup Y)^\perp = X^\perp \cap Y^\perp$.*

(ii) *En particulier, si $X \subset Y$, $Y^\perp \subset X^\perp$*

◁ **Éléments de preuve.**

Par double inclusion, en revenant aux éléments. Le deuxième point s'en déduit : comment caractériser une inclusion avec des unions ? ▷

Proposition 27.2.15 (Structure de l'orthogonal)

1. *Soit $x \in E$. Alors $x^\perp$ est un sev de E*
2. *Soit $X \subset E$. Alors $X^\perp$ est un sev de E.*

◁ **Éléments de preuve.**

Vérifications faciles. Le 2 se ramène au 1, en le décrivant sous forme d'intersection. ▷

Proposition 27.2.16 (Stabilité de l'orthogonal par Vect)

Soit X une partie de E. Alors $X^\perp = \mathrm{Vect}(X)^\perp$.

◁ **Éléments de preuve.**

Remarquer que $\mathrm{Vect}(X) \perp \mathrm{Vect}(X^\perp)$ et que $\mathrm{Vect}(X^\perp) = X^\perp$. ▷

Lemme 27.2.17 (Double orthogonal)

Soit F un sous-espace vectoriel d'un espace préhilbertien réel E. Alors $F \subset F^{\perp\perp}$.

Attention au fait qu'en dimension infinie, ce n'est pas nécessairement une égalité (voir des exemples en TD).

Proposition 27.2.18 (Orthogonal d'une somme, d'une intersection)

Soit F et G des sous-espaces vectoriels de l'espace préhilbertien E. On a alors :

(i) $(F + G)^\perp = F^\perp \cap G^\perp$;

(ii) $F^\perp + G^\perp \subset (F \cap G)^\perp$

◁ **Éléments de preuve.**

(i) Revenir à la définition de la somme par Vect.

(ii) Utiliser le comportement du passage à l'orthogonal sur les inclusions.

▷

Proposition 27.2.19

Soit F un sev de E. Alors $F \perp F^\perp$, donc en particulier, la somme $F \oplus F^\perp$ est directe.

Avertissement 27.2.20

Attention, contrairement à l'idée intuitive qu'on se fait en considérant l'orthogonalité dans $\mathbb{R}^3$, $F^\perp$ n'est pas forcément un supplémentaire de F dans E. On montrera que c'est le cas si E est de dimension finie, mais que cette propriété peut entrer en défaut si E est de dimension infinie.

Exemple 27.2.21

Déterminer l'orthogonal dans $\mathbb{R}^4$ muni de la base canonique de

$$F = \mathrm{Vect}\left(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\1\end{pmatrix}\right).$$

II.3 Projeté orthogonal

Définition 27.2.22 (Projeté orthogonal sur un sev)

Soit F un sous-espace vectoriel de E, et $y \in E$. On dit que $z \in E$ est le projeté orthogonal de y sur F si et seulement si :

(i) $z \in F$

(ii) $(y - z) \perp F$.

Théorème 27.2.23 (Existence du projeté orthogonal sur un sev de dimension finie)

Soit $y \in E$, et F un sous-espace vectoriel ***de dimension finie*** *de E, et $\mathcal{B} = (b_1, \dots, b_m)$ une b.o.n. de F. Alors le projeté orthogonal de y sur F existe, est unique, et vaut :*

$$z = \sum_{i=1}^{m} \langle y, b_i \rangle\, b_i.$$

◁ **Éléments de preuve.**

Par analyse synthèse : pour l'analyse, décomposer le projeté dans la b.o.n. fournie, puis exploiter l'orthogonalité définissant le projeté, en la traduisant par une orthogonalité avec chacun des b_i. Synthèse facile. ▷

En particulier, en considérant la b.o.n. $\left(\frac{x}{\|x\|}\right)$ de $\mathrm{Vect}(x)$, on obtient :

Proposition 27.2.24 (Existence et expression du projeté orthogonal sur une droite)

Soit x et y deux éléments de E, $x \neq 0$. Alors, le projeté orthogonal de y sur $\mathrm{Vect}(x)$ existe, est unique, et vaut :

$$z = \langle y, x \rangle \cdot \frac{x}{\|x\|^2} = \left\langle y, \frac{x}{\|x\|} \right\rangle \cdot \frac{x}{\|x\|}.$$

Remarques 27.2.25

1. Si $(b_1, \dots, b_m)$ est une base orthogonale, mais non orthonormale, on obtient la formule suivante pour le projeté orthogonale de y sur F :

$$z = \sum_{i=1}^{m} \left\langle y, \frac{b_i}{\|b_i\|} \right\rangle \cdot \frac{b_i}{\|b_i\|}.$$

2. Si $y \in F$, son projeté est bien entendu lui-même, et on obtient, pour une b.o.n. $(b_1, \dots, b_m)$ de F :
$$y = \sum_{i=1}^{m} \langle y, b_i \rangle b_i.$$
Ainsi, on trouve l'expression des coordonnées d'un vecteur y dans une base orthonormale.

II.4 Orthonormalisation de Gram-Schmidt

Motivation : étant donné une famille libre $(e_1, \dots, e_n)$ de E, trouver un moyen concret de construire une famille libre orthonormée $(f_1, \dots f_n)$ engendrant le même espace que $(e_1, \dots, e_n)$, c'est-à-dire tel que $(f_1, \dots, f_n)$ est une b.o.n. de $\mathrm{Vect}(e_1, \dots, e_n)$.
En particulier, si $(e_1, \dots, e_n)$ est initialement une base de E, on décrit une façon canonique de construire une b.o.n. de E à partir de cette base.
On fait cette construction étape par étape, de manière à avoir, pour tout $k \in [\![1, n]\!]$,
$$\mathrm{Vect}(e_1, \dots, e_k) = \mathrm{Vect}(f_1, \dots, f_k).$$

Théorème 27.2.26 (Procédé d'orthonormalisation de Gram-Schmidt)
Soit E un espace préhilbertien réel. Soit $(e_1, \dots, e_n)$ une famille libre de E. Il existe une unique famille orthonormale $(f_1, \dots, f_n)$ telle que, pour tout $k \in [\![1, n]\!]$, on ait
$$\mathrm{Vect}(f_1, \dots, f_k) = \mathrm{Vect}(e_1, \dots, e_k), \qquad \textit{et} \qquad \langle e_k, f_k \rangle \geqslant 0.$$
Cette famille peut être construite explicitement par la description récursive suivante :
$$f_1 = \frac{e_1}{\|e_1\|} \qquad \textit{et} \qquad \forall k \in [\![2, n]\!], \quad f_k = \frac{u_k}{\|u_k\|}, \ \textit{où } u_k = e_k - \sum_{i=1}^{k-1} \langle e_k, f_i \rangle f_i.$$

◁ **Éléments de preuve.**
Les k premiers termes ayant été orthonormalisé, on trouve un vecteur orthogonal à F_k en projetant orthogonalement e_{k+1} sur $F_k = \mathrm{Vect}(e_1, \dots, e_k)$. On connaît, d'après les résultats précédents, la décomposition du projeté orthogonal dans la b.o.n. de F_k définie par les vecteurs $f_1, \dots f_k$. ▷

Ainsi, pour tout $k \in [\![1, n]\!]$, $(f_1, \dots, f_k)$ est une b.o.n. de $\mathrm{Vect}(e_1, \dots, e_k)$. En particulier, si initialement, $(e_1, \dots, e_n)$ est une base de E, alors $(f_1, \dots, f_n)$ en est une base orthonormale.

Remarque 27.2.27
Le procédé de Gram-Schmidt construit donc à partir d'une base quelconque une base orthonormale qui préserve le drapeau (c'est-à-dire la famille croissante $(\mathrm{Vect}(e_1, \dots, e_k))_{k \in [\![0, n]\!]}$), et les orientations intermédiaires.

Exemples 27.2.28

1. Base orthonormale de $F = \mathrm{Vect}\left(\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}\right)$
2. Orthonormalisée de Schmidt de $(1, X, X^2)$ dans $\mathbb{R}_n[X]$ muni de :
$$\langle P, Q \rangle = \int_0^2 P(t)Q(t) \ \mathrm{d}t.$$

III Espaces euclidiens

On suppose ici que E est un espace euclidien (donc de dimension finie). On commence par montrer l'existence de bases orthonormales. On en déduit en particulier la possibilité de projeter orthogonalement sur tout sous-espace vectoriel.

III.1 Bases orthonormales d'un espace euclidien

Nous avons déjà vu que tout espace euclidien admet une base orthonormale. À l'aide du procédé d'orthonormalisation de Gram-Schmidt, très constructif et explicite, on peut donner quelques propriétés supplémentaires, permettant d'étendre certaines propriétés usuelles sur les bases, en imposant des conditions d'orthogonalité supplémentiares.

Théorème 27.3.1 (théorème de la base orthonormale incomplète)

Soit E un espace euclidien.

1. *Toute famille orthonormale de E peut être complétée en une base orthonormale de E.*
2. *Toute famille libre orthogonale de E peut être complétée en une base orthogonale de E*

◁ **Éléments de preuve.**

Remarquer que l'orthonormalisation ne modifie pas une famille qui est déjà orthonormale. Donc elle ne modifie pas non plus les premiers vecteurs d'une base, si ceux-ci forment une famille orthonormale. Ainsi, construire une b.o.n. en orthonormalisant une base obtenu par théorème de la base incomplète. ▷

Corollaire 27.3.2 (Existence et unicité d'un supplémentaire orthogonal)

Soit E un espace euclidien. Tout sous-espace F de E admet un unique supplémentaire G tel que $F \overset{\perp}{\oplus} G$. De plus, on a $G = F^\perp$. On dit que $F^\perp$ est le supplémentaire orthogonal de F.

◁ **Éléments de preuve.**

Considérer l'espace engendré par les vecteurs qu'on rajoute dans le théorème de la b.o.n. incomplète. Pour l'unicité, montrer (par inclusion et dimension) que $G = F^\perp$. ▷

Proposition 27.3.3

Soit E un espace euclidien, et F et G deux sous-espaces.

1. $(F^\perp)^\perp = F$
2. $(F \cap G)^\perp = F^\perp + G^\perp$.

◁ **Éléments de preuve.**

1. La décomposition $F \overset{\perp}{\oplus} F^\perp = E$ donne l'unique supplémentaire orthogonal de $F^\perp$.
2. Appliquer 27.2.17(i) avec $F^\perp$ et $G^\perp$ et passer à l'orthogonal. Simplifier les double-orthogonaux.

▷

Proposition 27.3.4 (base orthonormale d'une somme orthogonale)

Soit E un espace euclidien, et F et G deux sous-espaces vectoriels de E tels que $F \perp G$. Soit $(b_1, \dots, b_p)$ une base orthonormale de F, et $(c_1, \dots, c_q)$ une base orthonormale de G. Alors $(b_1, \dots, b_p, c_1, \dots, c_q)$ est une base orthonormale de $F \oplus G$; en particulier, c'est une famille orthonormale de E.

◁ **Éléments de preuve.**
Les vecteurs étant unitaires, il suffit de vérifier qu'ils sont aussi 2 à 2 orthogonaux. ▷

Corollaire 27.3.5
Soit E un espace euclidien, et $F_1, \dots, F_p$ des sous-espaces vectoriels de E, deux à deux orthogonaux. Alors la somme $F_1 + \dots + F_p$ est directe, et on obtient une base orthonormale de $F_1 \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} F_p$ en juxtaposant des bases orthonormales des espaces $F_1, \dots, F_p$. On dira qu'il sagit d'une b.o.n. adaptée à la décomposition en somme directe orthogonale.

◁ **Éléments de preuve.**
Soit par récurrence, soit vérification directe sur le même principe. ▷

Corollaire 27.3.6
Soit E un espace euclidien, F un sous-espace vectoriel de E, et $F^\perp$ son supplémentaire orthogonal. Soit $(b_1, \dots, b_p)$ une base orthonormale de F et $(c_1, \dots, c_q)$ une base orthonormale de $F^\perp$. Alors $(b_1, \dots, b_p, c_1, \dots, c_q)$ est une base orthonormale de E.

◁ **Éléments de preuve.**
Conséquence immédiate du précédent. ▷

III.2 Coordonnées en base orthonormale

Théorème 27.3.7 (Coordonnées d'un vecteur dans une b.o.n. et norme)
Soit E un espace euclidien, et $\mathcal{B} = (b_1, \dots, b_n)$ une b.o.n. de E. Soit $x, y \in E$. Alors :

(i) $x = \sum\limits_{i=1}^{n} \langle x, b_i \rangle b_i$, *c'est-à-dire* $[x]_{\mathcal{B}} = \begin{pmatrix} \langle x, b_1 \rangle \\ \vdots \\ \langle x, b_n \rangle \end{pmatrix}$.

(ii) $\langle x, y \rangle = \sum\limits_{i=1}^{n} \langle x, b_i \rangle \langle y, b_i \rangle$

(iii) $\|x\|^2 = \sum\limits_{i=1}^{n} \langle x, b_i \rangle^2$

◁ **Éléments de preuve.**
Projeter. ▷

Ainsi, si $\mathcal{B}$ est une b.o.n. et si

$$[x]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{et} \qquad [y]_{\mathcal{B}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

sont les coordonnées de x et y dans la b.o.n. $\mathcal{B}$, le produit scalaire et la norme s'expriment par les formules usuelles :

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i \qquad \text{et} \qquad \|x\|^2 = \sum_{i=1}^{n} x_i^2.$$

La matrice d'un endomorphisme admet également une description très simple :

Théorème 27.3.8 (Matrice d'un endomorphisme relativement à une b.o.n.)
Soit E un espace euclidien muni d'une b.o.n. $\mathcal{B} = (b_1, \dots, b_n)$. Soit $u \in \mathcal{L}(E)$. Alors :

$$\mathrm{Mat}_{\mathcal{B}}(u) = (\langle b_i, u(b_j)\rangle)_{1\leqslant i,j\leqslant n} = \begin{pmatrix} \langle b_1, u(b_1)\rangle & \cdots & \langle b_1, u(b_n)\rangle \\ \vdots & & \vdots \\ \langle b_n, u(b_1)\rangle & \cdots & \langle b_n, u(b_n)\rangle \end{pmatrix}.$$

◁ **Éléments de preuve.**
Conséquence immédiate du précédent, la colonne j étant alors le vecteur des coordonnées de $u(b_i)$ dans la b.o.n. $\mathcal{B}$. ▷

Enfin, l'expression de la matrice du produit scalaire dans la base $\mathcal{B}$ permet de caractériser facilement les bases orthonormales.

Théorème 27.3.9 (Matrice du produit scalaire relativement à une b.o.n.)
Soit E un espace euclidien, et $\mathcal{B}$ une base de E. Alors la base $\mathcal{B}$ est orthonormale si et seulement si la matrice du produit scalaire dans $\mathcal{B}$ est I_n.
Dans ce cas, pour tout (x, y) dans E^2, on a :

$$\langle x, y\rangle = X^\top Y,$$

où $X = [x]_{\mathcal{B}}$ et $Y = [y]_{\mathcal{B}}$ sont les vecteurs des coordonnées de x et y dans la base $\mathcal{B}$.

◁ **Éléments de preuve.**
On a déjà justifié l'expression d'un ps avec les coordonnées dans une b.o.n. Cela donne directement l'expression de la matrice. ▷

III.3 Changements de base et matrices orthogonales

Les matrices de passage d'une b.o.n. à une autre vérifient une propriété très forte :

Théorème 27.3.10 (propriété des matrices de passage d'une b.o.n. à une autre)
Soit $\mathcal{B}$ et $\mathcal{C}$ deux b.o.n. de E. Soit $P = P_{\mathcal{B}}^{\mathcal{C}}$ la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{C}$. Alors :

$$P^\top P = I_n = PP^\top, \qquad \text{donc:} \qquad P^{-1} = P^\top.$$

◁ **Éléments de preuve.**
Utiliser la formule de changement de base pour la matrice du ps dans les bases $\mathcal{B}$ et $\mathcal{C}$. ▷

Cette propriété définit la notion de matrice orthogonale :

Définition 27.3.11 (Matrice orthogonale)
Soit $P \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée d'ordre n. On dit que P est une matrice orthogonale si et seulement si $P^\top P = I_n$.

De la définition même découle de façon immédiate :

Proposition 27.3.12 (Inverse d'une matrice orthogonale)

Soit P une matrice orthogonale. Alors P est inversible, et $P^{-1} = P^\top$. De plus P^{-1} est aussi orthogonale.

Une matrice orthogonale peut se caractériser également par l'orthonormalité de ses colonnes :

Théorème 27.3.13 (Caractérisation d'une matrice orthogonale par ses colonnes)

Une matrice P est orthogonale si et seulement si ses colonnes forment une base orthonormale de $\mathbb{R}^n = \mathcal{M}_{n,1}(\mathbb{R})$, muni du produit scalaire canonique.

◁ **Éléments de preuve.**

Exprimer le produit matriciel $P^\top P = I_n$ coefficient par coefficient (l'exprimer avec des produits scalaires). ▷

Théorème 27.3.14 (Caractérisation des matrices de passage entre b.o.n. par orthogonalité)

Soit E un espace euclidien.

1. *Toute matrice de passage d'une b.o.n. de E à une autre b.o.n. de E est une matrice orthogonale.*
2. *Réciproquement, soit $\mathcal{B}$ une b.o.n. de E, et P une matrice orthogonale. Alors il existe une unique base $\mathcal{C}$ telle que P soit la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, et cette base $\mathcal{C}$ est une b.o.n. de E.*
3. *De même, si $\mathcal{C}$ est une b.o.n. de E, et P une martice orthogonale, il existe une unique base $\mathcal{B}$ telle que P soit la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, et cette base $\mathcal{B}$ est une b.o.n. de E.*

◁ **Éléments de preuve.**

1. Résulte de la définition.
2. Remarquer que si x et y sont deux vecteurs de coordonnées X et Y dans une b.o.n, et si φ est le psc, $\langle x, y\rangle = \varphi(X, Y)$. Justifier que ceci permet de définir une b.o.n. dont les coordonnées dans $\mathcal{B}$ sont les colonnes de P.
3. Appliquer ce qui précède à P^{-1} qui est encore orthogonal.

▷

Définition 27.3.15 (Groupe orthogonal)

On note $\mathrm{O}_n(\mathbb{R})$, ou $\mathrm{O}(n)$ l'ensemble des matrices orthogonales de $\mathcal{M}_n(\mathbb{R})$. Cet ensemble $\mathrm{O}_n(\mathbb{R})$ est appelé groupe orthogonal.

Théorème 27.3.16 (Structure de $\mathrm{O}_n(\mathbb{R})$)

L'ensemble $\mathrm{O}_n(\mathbb{R})$ est un groupe multiplicatif, ce qui est cohérent avec la terminologie introduite dans la définition précédente.

◁ **Éléments de preuve.**

Montrer que c'est un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$; les vérifications sont simples. ▷

Théorème 27.3.17 (Déterminant d'une matrice orthogonale)

Soit $P \in \mathrm{O}_n(\mathbb{R})$. Alors $\det(P) \in \{-1, 1\}$. Plus précisément, $\det$ est un morphisme de groupe de $\mathrm{O}_n(\mathbb{R})$ dans $(\{-1, 1\}, \times)$.

◁ **Éléments de preuve.**
Appliquer le det à l'égalité $P^\top P = I_n$. ▷

Le noyau de ce morphisme est donc un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Définition 27.3.18 (Groupe spécial orthogonal)
Le noyau du déterminant défini sur $\mathrm{O}_n(\mathbb{R})$ est appelé groupe spécial orthogonal, et noté $\mathrm{SO}_n(\mathbb{R})$ ou $\mathrm{SO}(n)$. Ainsi, les éléments de $\mathrm{SO}_n(\mathbb{R})$ sont les matrices orthogonales P telles que $\det(P) = 1$.

Le choix d'une matrice orthogonale P telle que $\det(P) = -1$ définit alors une bijection de $\mathrm{SO}(n)$ dans $\mathrm{O}^-(n) = \mathrm{O}(n) \setminus \mathrm{SO}(n)$ par $Q \mapsto PQ$.

III.4 Projecteurs orthogonaux et distance à un sous-espace

Dans un espace euclidien, l'existence de b.o.n. permet de projeter orthogonalement sur tout sous-espace F. Puisqu'il s'agit de trouver la composante sur F de la décomposition d'un vecteur dans la somme directe $F \oplus F^\perp$, la projection orthogonale est un projecteur. La description gémétrique et l'involutivité de l'orthogonal nous assure que le projecteur associé est la projection orthogonale sur $F^\perp$.

Méthode 27.3.19 (Comment déterminer la matrice d'un projecteur othogonal de $\mathbb{R}^n$)
Soit dans $\mathbb{R}^3$ le plan F d'équation $x + 2y - z = 0$. Déterminer la matrice de p_F, le projecteur orthogonal sur le plan F.
(Réponse : $\frac{1}{6}\begin{pmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \\ 1 & 2 & 5 \end{pmatrix}$)

- Première méthode : trouver une b.o.n. de F, et utiliser la formule donnant le projeté à l'aide d'une b.o.n.
- Deuxième méthode : trouver une b.o.n. de $F^\perp$ et projeter sur $F^\perp$ puis passer au projecteur associé

Les méthodes sont symétriques. Les calculs seront d'autant plus longs que la dimension de l'espace sur lequel on projette est grande. Ainsi, on adoptera la première méthode plutôt lorsque $\dim(F) \leqslant \dim(F^\perp)$, et la seconde dans le cas contraire.

Théorème 27.3.20 (Distance d'un point à un sous-espace)
Soit E un espace euclidien, F un sous-espace de E, et $x \in E$. Alors :

$$\forall y \in F, \quad \|x - p_F(x)\| \leqslant \|x - y\|,$$

l'égalité étant réalisée si et seulement si $y = p_F(x)$.
Autrement dit, $p_F(x)$ est l'unique vecteur de F minimisant la distance de x à un point de F.

◁ **Éléments de preuve.**
Je n'ai qu'un mot à dire : Pythagore. ▷

Définition 27.3.21
On dit que $p_F(x)$ est la meilleure approximation de x dans F. On appelle distance de x à F le réel suivant :

$$d(x, F) = \|x - p_F(x)\| = \min_{y \in F} \|x - y\|.$$