École universitaire Paris-Saclay
Licence de Mathématiques 1ère année

Année 2020–2021

MEU152 – Algèbre linéaire

Anne Moreau & Valentin Hernandez

d'après les notes de Guy Henniart et Thierry Ramond

anne.moreau@universite-paris-saclay.fr
https://www.imo.universite-paris-saclay.fr/~moreau/

***Johann Carl Friedrich Gauss**, né le 30 avril 1777 à Brunswick et mort le 23 février 1855 à Göttingen, est un mathématicien, astronome et physicien allemand. Doté d'un grand génie, il apporte de très importantes contributions à ces trois sciences. Surnommé «le prince des mathématiciens», il est considéré comme l'un des plus grands mathématiciens de tous les temps.*

Table des matières

1

Espaces vectoriels

1. Définitions et exemples

Soit E un ensemble. On suppose que l'on sait additionner deux éléments de E, et qu'on obtient alors un élément de E. On note

$$a + b$$

la somme des éléments a et b de E. On suppose aussi que l'on sait multiplier un élément de E quelconque par un réel, et que l'on obtient ainsi un élément de E. On note $\lambda.a$ le produit du réel λ et de $a \in E$. Dans ces conditions, on dit que «+» est une *loi de composition interne sur* E, et que «.» est une *loi de composition externe sur* $\mathbb{R} \times E$.

Exemple 1. Soit $E = \mathbb{R}$. On sait ajouter deux éléments de $\mathbb{R}$, et multiplier un élément de $\mathbb{R}$ par un élément de $\mathbb{R}$.

Exemple 2. Soit $E = \mathbb{R}^n$, c'est-à-dire l'ensemble des n-uplets $(x_1, \ldots, x_n)$ d'éléments de $\mathbb{R}$. Les réels $x_1, \ldots, x_n$ sont appelées les *composantes* du n-uplets $(x_1, \ldots, x_n)$. Soient $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ deux n-uplets, et $\lambda \in \mathbb{R}$. On note

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_1 + y_n),$$

le n-uplet obtenu en ajoutant les deux n-uplets composante par composante, et

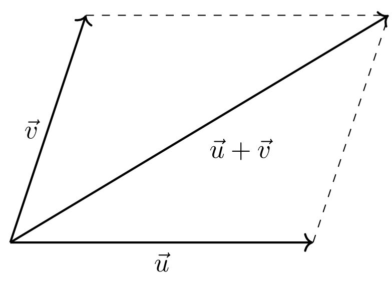
$$\lambda.(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n),$$

le n-uplet obtenu en multipliant chacune des composantes par λ.

Exemple 3. Soit $\mathscr{V}$ l'ensemble des vecteurs du plan ou de l'espace. On rappelle qu'un vecteur non nul est la donnée d'une direction, d'un sens et d'une longueur. Si $\vec{v}$ est le vecteur nul, $\lambda.\vec{v}$ est le vecteur nul pour tout $\lambda \in \mathbb{R}$. Pour $\lambda \in \mathbb{R}$ et $\vec{v} \in \mathscr{V}$ non nul, le vecteur $\lambda.\vec{v}$ est le vecteur nul si $\lambda = 0$, et pour $\lambda \neq 0$ le vecteur

— de même direction que $\vec{v}$,
— de même sens que $\vec{v}$ si $\lambda > 0$ et de sens opposé si $\lambda < 0$,
— de longueur $|\lambda|$ fois la longueur de $\vec{v}$.

On sait aussi ajouter deux vecteurs $\vec{u}$ et $\vec{v}$ de $\mathscr{V}$: on trace un représentant de $\vec{v}$ en partant de l'extrémité d'un représentant de $\vec{u}$.



Exemple 4. Soit X un ensemble quelconque, et $\mathscr{F}(X, \mathbb{R})$ l'ensemble des fonctions de X dans $\mathbb{R}$. La somme $f + g$ de deux éléments f et g de $\mathscr{F}(X, \mathbb{R})$ est la fonction de X dans $\mathbb{R}$ définie par

$$h\colon X \to \mathbb{R}, \quad x \mapsto f(x) + g(x).$$

Le produit $k = \lambda.f$ de la fonction f par un réel λ st la fonction de X dans $\mathbb{R}$ définie par

$$k\colon X \to \mathbb{R}, \quad x \mapsto \lambda f(x).$$

Ces définitions sont les définitions usuelles dans le cas $X = \mathbb{R}$.

Définition 1 – espace vectoriel

Soit $(E, +, .)$ un triplet constitué d'un ensemble E, d'une loi de composition interne «+» et d'une loi de composition externe «.» sur E. On dit que $(E, +, .)$ est un *espace vectoriel* lorsque les lois «+» et «.» vérifient les huit propriétés suivantes :

(1) pour tous $a, b \in E$, $a + b = b + a$,

(2) pour tous $a, b, c \in E$, $a + (b + c) = (a + b) + c$,

(3) il existe un élément e de E tel que pour tout $a \in E$, $a + e = e + a = a$,

(4) pour tout $a \in E$, il existe un élément b de E tel que $a + b = b + a = e$,

(i) pour tout $a \in E$, $1.a = a$,

(ii) pour tous réels λ, μ et tout $a \in E$, $\lambda.(\mu.a) = (\lambda\mu).a$,

(iii) pour tous réels λ, μ et tout $a \in E$, $(\lambda + \mu).a = \lambda.a + \mu.a$,

(iv) pour tous $a, b \in E$ et tout réel λ, $\lambda.(a + b) = \lambda.a + \lambda.b$.

Lorsque la propriété (1) est vérifiée, on dit que la loi «+» est *commutative*. Si la propriété (2) est vérifiée, on dit que la loi «+» est *associative*.

Exercice de cours 1 (unicité de l'élément neutre et du symétrique).

1. Démontrer que l'élément e de la propriété (3) est unique. On l'appelle l'*élément neutre* pour la loi «+». On le note la plupart du temps 0_E ou simplement 0, et on l'appelle aussi le *vecteur nul* de E.

2. Démontrer que l'élément b de la propriété (4) est lui aussi unique s'il existe. Le cas échéant, on l'appelle le *symétrique* ou l'*opposé* de a, et on le note $b = -a$.

Définition 2 – groupe commutatif

Lorsque les propriétés (1) à (4) sont vérifiées, on dit que $(E, +)$ est un *groupe commutatif*.

Lorsque la propriété (i) est vérifiée, on dit que le réel 1 est l'*élément neutre* pour la loi externe. On dit que la loi externe est *associative* lorsque (ii) est vraie. Les deux dernières propriétés (iii) et (iv) sont des propriétés de *distributivité*.

Proposition 3 – règles de calcul dans un espace vectoriel

Pour tout $x \in E$, notant 0_E l'élément neutre de l'addition dans $(E, +, .)$, on a

$$0.x = 0_E \qquad \text{et} \qquad (-1).x = -x.$$

Exercice de cours 2 (règles de calcul dans un espace vectoriel).

1. Démontrer la proposition 3.

2. En déduire que pour tous $x, y \in E$ et tous réels λ, μ,

$$\begin{aligned} &\lambda.(x - y) = \lambda.x - \lambda.y, \\ &(\lambda - \mu).x = \lambda.x - \mu.x, \\ &\lambda.x = 0_E \iff \lambda = 0 \text{ ou } x = 0_E, \end{aligned}$$

où l'on note pour $x, y \in E$, $x - y$ pour $x + (-y)$.

On retiendra donc que dans un espace vectoriel, on peut effectuer les calculs selon les mêmes règles que celles utilisées pour manipuler les vecteurs du plan ou de l'espace.

> EXERCICE DE COURS 3 (exemples fondamentaux). Reprendre les exemples précédents 1–4 et démontrer que ce sont bien des espaces vectoriels. Préciser pour chacun d'entre eux l'élément neutre et le symétrique de tout élément.

L'exemple suivant généralise l'exemple 4.

EXEMPLE 5. Si F est un espace vectoriel et X un ensemble, on peut définir sur l'ensemble $\mathscr{F}(X, F)$ des fonctions de X dans F les deux opérations sommes et produit par un réel comme dans l'exemple 4 en utilisant les opérations dans F. On montre facilement, comme dans le cas de $\mathscr{F}(X, \mathbb{R})$, que $\mathscr{F}(X, F)$ muni de ces opérations est un espace vectoriel.

> EXERCICE DE COURS 4. Vérifier les assertions de l'exemple précédent.

Voici un autre exemple d'espace vectoriel important.

EXEMPLE 6. Soit $\mathscr{S}$ l'ensemble des suites de nombres réels. On définit la somme $(u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}}$ de $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ comme étant la suite $(u_n + v_n)_{n\in\mathbb{N}}$, et le produit $\lambda.(u_n)_{n\in\mathbb{N}}$ de la suite $(u_n)_{n\in\mathbb{N}}$ par le réel λ comme étant la suite $(\lambda u_n)_{n\in\mathbb{N}}$. Comme dans les exemples précédents, il est facile de voir que $(\mathscr{S}, +, .)$ est un espace vectoriel en se ramenant aux propriétés de l'addition et de la multiplication dans $\mathbb{R}$. Son élément neutre est la suite nulle, i.e., la suite $(u_n)_{n\in\mathbb{N}}$ où $u_n = 0$ pour tout $n \in \mathbb{N}$. On peut voir $\mathscr{S}$ comme $\mathscr{F}(\mathbb{N}, \mathbb{R})$.

REMARQUE 1.

1. Un point de vocabulaire : en toute rigueur, il est abusif de dire ou d'écrire que E est un espace vectoriel; c'est le triplet $(E, +, .)$ qui est un espace vectoriel, E désignant seulement l'ensemble sous-jacent. Cependant, on s'autorisera bien souvent cet abus de langage, surtout dans les exercices.
2. Dans la définition 1, les éléments λ, μ qui opèrent dans E pour la loi externe sont des réels. On dit parfois que $(E, +, .)$ est un *espace vectoriel réel*, ou encore que $(E, +, .)$ est un $\mathbb{R}$-*espace vectoriel*. On définit de la même manière un *espace vectoriel complexe*, ou $\mathbb{C}$-*espace vectoriel*, $(E, +, .)$ en remplaçant dans la définition 1 «réel» par «complexe».

EXEMPLE 7. L'ensemble $\mathbb{C}^n$ des n-uplets $(x_1, \ldots, x_n)$ de nombres complexes muni des lois «+» et «.» définies comme pour $\mathbb{R}^n$ forme un $\mathbb{C}$-espace vectoriel.

2. Sous-espaces vectoriels

2.1. Généralités.

Définition 4 – sous-espace vectoriel

Soient $(E, +, .)$ un espace vectoriel, et F une partie de E. On dit que F est un *sous-espace vectoriel de E* lorsque les trois propriétés suivantes sont vérifiées :

1. le vecteur nul 0_E de E appartient à F,
2. pour tous éléments a et b de F, $a + b$ est aussi un élément de F,
3. pour tout réel λ et tout élément a de F, $\lambda.v$ est aussi un élément de F.

EXEMPLE 8. Dans $(\mathbb{R}, +, .)$ il n'y a pas beaucoup de sous-espaces vectoriels. Soit en effet F un sous-espace vectoriel de E. Si F contient un élément non nul a, alors pour tout $\lambda \in \mathbb{R}$, on sait que λa appartient à F. Comme n'importe quel élément x de $\mathbb{R}$ peut s'écrire λa (prendre $\lambda = x/a$), on voit que F contient $\mathbb{R}$ tout entier, donc $F = \mathbb{R}$.

D'autre part le singleton $(F = \{0\}, +, .)$ est un sous-espace vectoriel de $\mathbb{R}$ (les propriétés de la définition 4 sont faciles à vérifier). En conclusion, les seuls sous-espaces vectoriels de $\mathbb{R}$ sont $\{0\}$ et $\mathbb{R}$.

> EXERCICE DE COURS 5. Vérifier que l'ensemble $F = \{(x, y) \in \mathbb{R}^2 : x + y = 0\}$ est un sous-espace vectoriel de $\mathbb{R}^2$.

⚠ Attention, l'ensemble $G = \{(x, y) \in \mathbb{R}^2 : x + y = 1\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$. Il «ressemble» toutefois à un sous-espace vectoriel. En effet, le point $A = (1, 0)$ appartient à G, et l'ensemble

$$F = \{(x, y) \in \mathbb{R}^2 : A + (x, y) \in G\}$$

est un sous-espace vectoriel de $\mathbb{R}^2$.

Exercice de cours 6.

1. Représenter F et G sur une même figure et vérifier les assertions précédentes. Dans ces conditions, on dit que $G = A + F$ est un *espace affine de direction* F.
2. Donner d'autres exemples de sous-espaces vectoriels de $\mathbb{R}^2$, et d'autres exemples de sous-ensembles de $\mathbb{R}^2$ qui ne sont pas des sous-espaces vectoriels de $\mathbb{R}^2$.

Exemple 9. L'ensemble $\mathscr{C}^0(\mathbb{R}, \mathbb{R})$ des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ est un sous-espace vectoriel de $\mathscr{F}(\mathbb{R}, \mathbb{R})$.

Exemple 10. L'ensemble $\mathscr{S}_0$ des suites de nombres réels qui convergent est un sous-espace vectoriel de $\mathscr{S}$. En revanche, l'ensemble $\mathscr{S}_\infty$ des suites de nombres réels qui divergent n'est pas un sous-espace vectoriel de $\mathscr{S}$.

Exercice de cours 7. Vérifier les assertions des exemples précédents.

Remarque 2 (sous-espaces vectoriels triviaux). L'ensemble $F = \{0_E\}$ est toujours un sous-espace vectoriel de n'importe quel espace vectoriel E. En effet $0_E + 0_E = 0_E \in E$ et $\lambda.0_E = 0_E \in E$ pour tout $\lambda \in \mathbb{R}$. On parle du *sous-espace vectoriel nul*. L'ensemble E est lui aussi toujours un sous-espace vectoriel de E.

Proposition 5 – structure de l'ensemble des solutions d'un système linéaire homogène

Soient $n, p \in \mathbb{N}$. L'ensemble des solutions d'un système linéaire homogène (de n équations) à p inconnues est un sous-espace vectoriel de $\mathbb{R}^p$.

Exercice de cours 8. Démontrer la proposition précédente.

L'intérêt principal de la notion de sous-espace vectoriel réside dans la proposition suivante.

Proposition 6 – un sous-espace vectoriel est un espace vectoriel

Si F est un sous-espace vectoriel d'un espace vectoriel $(E, +, .)$, alors «+» est une loi interne sur F et «.» est une loi externe sur F. De plus $(F, +, .)$ est un espace vectoriel.

Exercice de cours 9. Démontrer la proposition précédente.

Remarque 3. En pratique, pour démontrer qu'un ensemble muni de deux lois est un espace vectoriel, on a toujours intérêt à démontrer que c'est un sous-espace vectoriel d'un espace vectoriel connu qui le contient. Il n'y a alors que trois propriétés à démontrer, au lieu des huit propriétés de la définition 1.

2.2. L'espace vectoriel des polynômes. On rappelle qu'une *fonction polynomiale*, ou *fonction polynôme*, est une fonction de la forme

$$\begin{array}{rccl} f : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n, \end{array}$$

où $n \in \mathbb{N}$ est un entier naturel et les coefficients $a_0, \ldots, a_n$ sont des réels éventuellement nuls.

Pour $k \in \mathbb{N}$, on note X^k la fonction polynomiale qui à x associe x^k. Pour $k = 0$, X^k est donc la fonction constante égale à 1 que l'on note simplement 1 (on adopte la convention $X^0 = 1$). Ainsi, par définition, une fonction polynomiale est une *combinaison linéaire* des X^k (voir la définition 9) : une fonction polynôme est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ telle qu'il existe $n \in \mathbb{N}$ et $a_0, \dots, a_n \in \mathbb{R}$ tels que $f = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$. Une telle combinaison linéaire $a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$ est également appelée un *polynôme*. On note $\mathbb{R}[X]$ l'ensemble des polynômes (à coefficients réels).

La distinction entre «polynôme» et «fonction polynomiale» sera faite dans les cours d'algèbres au cours des années années suivantes. Pour ce cours, on confondra les deux notions de sorte que $\mathbb{R}[X]$ est l'ensemble des fonctions polynômiales.

Définition 7 – degré et coefficient dominant d'un polynôme

Soit $P = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n \in \mathbb{R}[X]$ un polynôme non nul. On appelle *degré* de P, et on note $\deg P$, le maximum des $i \in \{0, \dots, n\}$ tels que $a_i \neq 0$. Le coefficient $a_{\deg P}$ est appelé le *coefficient dominant de P*.
Par convention, le degré du polynôme nul (où tous les a_i sont nuls) est $-\infty$.

Par exemple, le degré des fonctions polynomiales $x \mapsto 3x + 2$, $x \mapsto -x^2$, $x \mapsto -x^2 + 0 \times x^3$ est 1, 2, 2 respectivement. Le coefficient dominant des fonctions polynomiales $x \mapsto 3x + 2$, $x \mapsto -x^2$, $x \mapsto -x^2 + 0 \times x^3$ est 3, -1, et -1 respectivement.

Exercice de cours 10. Montrer que l'ensemble $\mathbb{R}[X]$ des polynômes à coefficients réels est un espace vectoriel. Préciser la loi interne «+» et la loi externe «.» ainsi que l'élément neutre.
Indication : identifier $\mathbb{R}[X]$ à un sous-espace vectoriel de l'espace vectoriel $\mathscr{F}(\mathbb{R}, \mathbb{R})$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

Proposition 8 – le degré de la somme de deux polynômes est inférieur au maximum des degrés

Pour tous $P, Q \in \mathbb{R}[X]$,
$$\deg(P + Q) \leqslant \max(\deg(P), \deg(Q)).$$
De plus, si $\deg(P) \neq \deg(Q)$ alors $\deg(P + Q) = \max(\deg(P), \deg(Q))$.

Exercice de cours 11. Démontrer la proposition.

⚠ Il se peut que $\deg(P + Q) < \max(\deg P, \deg Q)$ lorsque P et Q sont de même degré. Par exemple, si $P = X^2 + X$ et $Q = -X^2$, alors $\deg(P + Q) = \deg(X) = 1 < \max(\deg P, \deg Q) = 2$.

On peut aussi multiplier deux polynômes. Plus généralement, si $f, g \in \mathscr{F}(\mathbb{R}, \mathbb{R})$ sont deux fonctions de $\mathbb{R}$ dans $\mathbb{R}$, la fonction $h = f \times g$ de $\mathbb{R}$ dans $\mathbb{R}$ est définie par : $h(x) = f(x)g(x)$ pour tout $x \in \mathbb{R}$.

Exercice de cours 12 (produit de deux polynômes).

1. Montrer que le produit de deux polynômes est encore un polynôme. Précisément, si $P = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$ et $Q = b_0 + b_1X + b_2X^2 + \cdots + b_mX^m$, où $m, n \in \mathbb{N}$, $a_0, \dots, a_n, b_0, \dots, b_m \in \mathbb{R}$ sont deux polynômes, exprimer les coefficients du polynôme PQ en fonction des a_i et des b_j.
2. Que vaut $\deg(PQ)$?

⚠ Nous avons vu au cours de l'exercice précédent qu'il existe dans l'espace vectoriel $\mathbb{R}[X]$ des polynômes une autre loi interne, la multiplication entre polynômes (en plus de la loi interne «+»). Ceci est très spécifique à cet espace vectoriel : en général, on ne peut pas multiplier deux éléments d'un même espace vectoriel !

Remarque 4. Un espace vectoriel $(E, +, .)$ muni d'une (autre) loi interne, $E \times E \to E$, $(x, y) \mapsto x \times y$, notée «×» et appelée la *multiplication*, telle que la multiplication soit distributive par rapport à l'addition et compatible avec la loi externe «.» est appelé une *algèbre*. Nous rencontrerons d'autres exemples d'algèbres dans ce cours : l'espace vectoriel $\mathscr{M}_n(\mathbb{R})$ des matrices carrées réelles d'ordre n muni de la multiplication des matrices, et l'espace vectoriel $\mathscr{L}(E)$ des endomorphismes d'un espace vectoriel E muni de la composition des endomorphismes (voir le paragraphe 2.1).

3. Combinaisons linéaires

Nous allons généraliser la notion de combinaison linéaire vue à propos des polynômes à n'importe quel espace vectoriel.

Définition 9 – combinaison linéaire

Soient $u_1, u_2, \dots, u_n$ des vecteurs d'un espace vectoriel $(E, +, .)$. On dit que $v \in E$ est une *combinaison* linéaire de $u_1, u_2, \dots, u_n$ lorsqu'il existe des réels $\lambda_1, \lambda_2, \dots, \lambda_n$ tels que

$$\begin{aligned} v &= \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n \\ &= \sum_{j=1}^{n} \lambda_j u_j. \end{aligned}$$

Avec cette nouvelle terminologie, on a une version plus courte de la définition d'un sous-espace vectoriel :

Proposition 10 – sous-espace vectoriel et stabilité par combinaisons linéaires

Soit $(E, +, .)$ un espace vectoriel. Une partie F de E est un sous-espace vectoriel de E si et seulement si F contient 0_E et F est stable par combinaison linéaire, c'est-à-dire :

$$\forall\, u, v \in F,\ \forall\, \lambda, \mu \in \mathbb{R}, \qquad \lambda u + \mu v \in F.$$

Exemple 11. Dans $\mathbb{R}^3$, le vecteur $u = (3, 3, 1)$ est-il combinaison linéaire de $v_1 = (1, 1, 0)$ et $v_2 = (1, 1, 1)$? Il s'agit de savoir s'il existe deux réels λ_1 et λ_2 tels que $\lambda_1.(1, 1, 0) + \lambda_2.(1, 1, 1) = (3, 3, 1)$. On est donc amené à résoudre le système

$$\left\{ \begin{array}{ccccc} \lambda_1 & + & \lambda_2 & = & 3, \\ \lambda_1 & + & \lambda_2 & = & 3, \\ & & \lambda_2 & = & 1. \end{array} \right.$$

Ce système a une (unique) solution $(\lambda_1, \lambda_2) = (2, 1)$, donc $u = 2v_1 + v_2$ est bien combinaison linéaire de v_1 et v_2.

Exercice de cours 13. À quelle condition le vecteur $w = (a, b, c)$ est-il une combinaison linéaire de $u = (1, 2, -1)$ et $v = (6, 4, 2)$?

4. Familles génératrices

On commence par un point de vocabulaire : une famille d'éléments d'un ensemble E est une liste, finie ou non, d'éléments de E. Puisque des éléments de la liste peuvent être égaux, on ne doit pas confondre cette notion de famille avec celle de sous-ensemble de E. On note

$$\mathscr{F} = (u_1, u_2, \dots u_n),$$

entre parenthèses, la famille constituée des éléments u_1, u_2, …, u_n. Dans le cadre des espaces vectoriels, on prendra garde à ne pas confondre une famille de n vecteurs avec un élément de $\mathbb{R}^n$ même si la notation est similaire.

Proposition 11 – les combinaisons linéaires de vecteurs donnés forment un sous-espace vectoriel

Soient $u_1, u_2, \dots, u_n$ des vecteurs d'un espace vectoriel $(E, +, .)$. L'ensemble de toutes les combinaisons linéaires de $u_1, u_2, \dots, u_n$ est un sous-espace vectoriel de E. On le note $\mathrm{Vect}(u_1, u_2, \dots, u_n)$ et on l'appelle le *sous-espace vectoriel engendré par la famille* $(u_1, u_2, \dots, u_n)$.

Exercice de cours 14. Démontrer la proposition précédente.

Remarque 5. Le sous-espace vectoriel $\mathrm{Vect}(u_1, u_2, \dots, u_n)$ est le plus petit (pour l'inclusion) des sous-espaces vectoriels de E qui contient $u_1,\dots, u_n$. En effet si F est un tel sous-espace vectoriel de E, comme il est stable par combinaison linéaire, il contient toutes les combinaisons linéaires de $u_1,\dots, u_n$, donc $\mathrm{Vect}(u_1, u_2, \dots, u_n)$.

Définition 12 – famille génératrice

Soient $(E, +, .)$ un espace vectoriel et G un sous-espace vectoriel de E. Soient $u_1, \dots, u_n$ des vecteurs de G. On dit que la famille de vecteurs $(u_1, u_2, \dots, u_n)$ *engendre* G, ou encore est une *famille génératrice* de G, lorsque $G = \mathrm{Vect}(u_1, u_2, \dots, u_n)$.

Exemple 12. Soient $n \in \mathbb{N}$ et $\mathbb{R}_n[X]$ l'ensemble des polynômes à coefficients réels de degré inférieur ou égal à n. Nous avons vu lors de l'exercice 10 que $\mathbb{R}[X]$ était un sous-espace vectoriel de $\mathscr{F}(\mathbb{R}, \mathbb{R})$. L'ensemble $\mathbb{R}_n[X]$ est lui aussi un sous-espace vectoriel de $\mathscr{F}(\mathbb{R}, \mathbb{R})$ (et de $\mathbb{R}[X]$). En effet la fonction nulle est par convention un polynôme de degré $-\infty$, donc (par abus de langage) inférieur ou égal à n, et une combinaison linéaire de polynômes de degré inférieur ou égal à n est bien un polynôme de degré inférieur ou égal à n (voir la proposition 8). Autrement dit, $\mathbb{R}_n[X]$ est le sous-espace vectoriel de $\mathscr{F}(\mathbb{R}, \mathbb{R})$ engendré par la famille $(1, X, \dots X^n)$:

$$\mathbb{R}_n[X] = \mathrm{Vect}(1, X, \dots X^n).$$

Exercice de cours 15.

1. La famille $((1, 1, 2), (-1, 0, 1), (2, 1/3, -1))$ engendre-t-elle $\mathbb{R}^3$?
2. Trouver une équation cartésienne de $F = \mathrm{Vect}((1, 1, 2), (-1, 0, 1), (2, 1/3, -1))$.

5. Familles libres

Définition 13 – famille libre

On dit que la famille $(u_1, u_2, \dots, u_n)$ de vecteurs d'un espace vectoriel $(E, +, .)$ est *libre* lorsque :

$$\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_n u_n = 0_E \quad \Longrightarrow \quad \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

Lorsqu'une famille de vecteurs d'un espace vectoriel $(E, +, .)$ n'est pas libre, on dit qu'elle est *liée*.

Exercice de cours 16. Dans $\mathbb{R}_n[X]$, montrer que la famille $(1, X, \dots, X^n)$ est libre.

Exercice de cours 17 (famille de polynômes non nuls de degrés deux à deux distincts). Montrer qu'une famille de polynômes non nuls de degrés deux à deux distincts est libre.

⚠ la réciproque de l'exercice précédent est fausse. Des polynômes peuvent parfaitement former une famille libre sans être de degrés deux à deux distincts ; ils peuvent même former une famille libre tout en étant tous de même degré, comme par exemple dans le cas de la famille $(X^n, X^n+1, X^n+X, X^n+X^2, \dots, X^n+X^{n-1})$ de $\mathbb{R}_n[X]$.

Exemple 13. Dans l'espace vectoriel $\mathscr{F}$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, on considère la famille $(\cos, \sin)$ constituée des deux fonctions trigonométriques cosinus et sinus bien connues, et on se demande si elle est libre. On suppose donc que pour deux réels λ_1, λ_2, on a

$$\lambda_1 . \cos + \lambda_2 . \sin = 0 \; (= 0_{\mathscr{F}}). \tag{1}$$

⚠ Cette égalité est une égalité dans $\mathscr{F}$, autrement dit entre fonctions ; en particulier le «0» qui figure dans le membre de droite est la fonction nulle, pas le nombre 0.

L'équation (1) signifie :

$$\forall\, x \in \mathbb{R}, \quad \lambda_1 \cos(x) + \lambda_2 \sin(x) = 0.$$

En particulier en prenant $x = 0$, on trouve $\lambda_1 = 0$, et pour $x = \pi/2$, on trouve $\lambda_2 = 0$. Donc la famille $(\cos, \sin)$ est libre dans $\mathscr{F}$.

Exercice de cours 18. Soit $a \in \mathbb{R}$. Dans $\mathbb{R}^3$, montrer que famille $((1,1,2),(-1,0,1),(2,a,-1))$ est libre si et seulement si $a \neq 1/3$.

Remarque 6. Il est important de noter que pour déterminer si une famille est génératrice, on s'est posé la question de *l'existence de solutions* pour un système linéaire, tandis que pour savoir si une famille est libre, il s'est agit de savoir si un système linéaire admet une *unique solution* ou bien plusieurs.

6. Bases d'un espace vectoriel

6.1. Définition, exemples.

Définition 14 – base

On dit que la famille finie $\mathscr{B} = (u_1, u_2, \ldots, u_n)$ de vecteurs d'un espace vectoriel $(E, +, .)$ est une *base* de E lorsque $\mathscr{B}$ est libre et génératrice.

Exemple 14. Dans $\mathbb{R}^n$ on note $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, ..., $e_n = (0, \ldots, 0, 1)$. La famille $(e_1, \ldots, e_n)$ est une base, qu'on appelle la *base canonique* de $\mathbb{R}^n$. Prouvons en effet que cette famille est génératrice. Soit $(x_1, x_2, \ldots, x_n)$ un vecteur de $\mathbb{R}^n$. On cherche $\lambda_1, \ldots, \lambda_n$ dans $\mathbb{R}$ tels que

$$\lambda_1.e_1 + \cdots + \lambda_n.e_n = (x_1, x_2, \ldots, x_n).$$

Puisque $\lambda_1.e_1 + \cdots + \lambda_n.e_n = (\lambda_1, \ldots, \lambda_n)$, il suffit de prendre $\lambda_j = x_j$ pour tout $j \in \{1, \ldots, n\}$. On montre maintenant que $(e_1, \ldots, e_n)$ est une famille libre. Supposons que

$$\lambda_1.e_1 + \cdots + \lambda_n.e_n = 0_{\mathbb{R}^n} = (0, \ldots, 0).$$

Encore une fois puisque $\lambda_1.e_1 + \cdots + \lambda_n.e_n = (\lambda_1, \ldots, \lambda_n)$, cette égalité entraine $\lambda_1 = \cdots = \lambda_n = 0$.

Exemple 15. Dans $\mathbb{R}_n[X]$, on a déjà vu que la famille $(1, X, \ldots, X^n)$ est génératrice (c'est la définition de $\mathbb{R}_n[X]$), et libre. C'est donc une base de $\mathbb{R}_n[X]$, appelée elle aussi la *base canonique*, de $\mathbb{R}_n[X]$ cette fois.

Proposition 15 – coordonnées dans une base

Soit $\mathscr{B} = (u_1, u_2, \ldots, u_n)$ une base de l'espace vectoriel $(E, +, .)$. Pour chaque $x \in E$, il existe un unique n-uplet $(x_1, \ldots, x_n) \in \mathbb{R}^n$ tel que

$$x = \sum_{j=1}^{n} x_j u_j.$$

On note alors

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{\mathscr{B}}, \text{ ou } x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

s'il n'y a pas d'ambiguïté, et l'on dit que $x_1, x_2, \ldots x_n$ sont les *coordonnées* de x dans la base $\mathscr{B}$.

Exercice de cours 19. Démontrer cette proposition.

Dans le cas de l'espace vectoriel $\mathbb{R}^n$, les coordonnées du n-uplet $x = (x_1, x_2, \ldots, x_n)$ dans la base canonique $\mathscr{E} = (e_1, e_2, \ldots, e_n)$ sont

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathscr{E}},$$

puisque l'on voit facilement que $x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$.

⚠ Il est important de bien distinguer les composantes $x_1, x_2, \ldots, x_n$ du n-uplet $x = (x_1, x_2, \ldots, x_n)$ de $\mathbb{R}^n$ et ses coordonnées dans une base donnée.

Par exemple, dans $\mathbb{R}^2$, on dispose de la base canonique $\mathscr{E} = (e_1, e_2)$ avec $e_1 = (1, 0)$ et $e_2 = (0, 1)$, mais aussi de la base $\mathscr{B} = (u_1, u_2)$ avec $u_1 = (1, 1)$ et $u_2 = (1, -1)$. Pour le vecteur $x = (2, 3)$, on a

$$x = \begin{pmatrix} 2 \\ 3 \end{pmatrix}_{\mathscr{E}} = \begin{pmatrix} \frac{5}{2} \\ -\frac{1}{2} \end{pmatrix}_{\mathscr{B}},$$

puisque $x = 2e_1 + 3e_2 = \frac{5}{2}u_1 - \frac{1}{2}u_2$.

Exercice de cours 20. Donner les coordonnées du polynôme $P(X) = 2X^2 - 3X + 1$ dans la base canonique de $\mathbb{R}_2[X]$, puis dans la base $\mathscr{B} = (Q_1, Q_2, Q_3)$ avec $Q_1(X) = X + 1$, $Q_2(X) = X^2$ et $Q_3(X) = X - 1$.

Exercice de cours 21. Donner les coordonnées des vecteurs de la base canonique $\mathscr{E}$ de $\mathbb{R}^n$ dans la base $\mathscr{E}$. Si $\mathscr{B}$ est une base quelconque, quelles sont les coordonnées des vecteurs de $\mathscr{B}$ dans $\mathscr{B}$?

6.2. Dimension d'un espace vectoriel. On commence par un résultat dont la preuve est un peu longue, mais qui a de nombreuses conséquences importantes.

Proposition 16 – lien entre le nombre d'éléments d'une famille libre et d'une famille génératrice

Soit $(E, +, .)$ un espace vectoriel. On suppose que E admet
- une famille génératrice $(u_1, u_2, \ldots, u_n)$,
- une famille libre $(v_1, v_2, \ldots v_p)$.

Alors nécessairement $p \leqslant n$. De plus, si $p = n$, alors $(u_1, u_2, \ldots, u_n)$ et $(v_1, v_2, \ldots v_p)$ sont des bases de E.

Exercice de cours 22. Le but de cet exercice est de démontrer la proposition 16.

1. Soit $v \in \mathrm{Vect}(v_1, v_2, \ldots v_p)$.

1.1 Montrer qu'il existe un unique p-uplet $(\lambda_1, \ldots, \lambda_p)$ de $\mathbb{R}^p$ tel que

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p = v. \tag{2}$$

1.2 Montrer qu'il existe des réels $a_{i,j}$, $i \in \{1, \ldots, p\}$, $j \in \{1, \ldots, n\}$, et des réels $x_1, \ldots x_n$ tels que

$$\left\{\begin{array}{lcl} v_1 & = & a_{1,1}u_1 + a_{1,2}u_2 + \cdots + a_{1,n}u_n, \\ v_2 & = & a_{2,1}u_1 + a_{2,2}u_2 + \cdots + a_{2,n}u_n, \\ & \vdots & \vdots \\ v_p & = & a_{p,1}u_1 + a_{p,2}u_2 + \cdots + a_{p,n}u_n, \end{array}\right.$$

et $v = x_1 u_1 + x_2 u_2 + \cdots + x_n u_n$.

1.3 Remarquer que si le système

$$\left\{\begin{array}{lcl} a_{1,1}\lambda_1 + a_{2,1}\lambda_2 + \cdots + a_{p,1}\lambda_p & = & x_1, \\ a_{1,2}\lambda_1 + a_{2,2}\lambda_2 + \cdots + a_{p,2}\lambda_p & = & x_2, \\ \vdots & & \vdots \\ a_{1,n}\lambda_1 + a_{2,n}\lambda_2 + \cdots + a_{p,n}\lambda_p & = & x_n, \end{array}\right. \tag{3}$$

admet une solution $(\lambda_1, \ldots, \lambda_p)$, c'est aussi une solution de l'équation (2). En déduire que le rang de ce système est p puis que $p \leqslant n$.

2. Si $p = n$, déduire des questions précédentes que $(v_1, v_2, \ldots v_p)$ est une famille génératrice, donc une base, et que $(u_1, u_2, \ldots, u_n)$ est une famille libre, donc une base.

Proposition 17 – toutes les bases ont le même nombre d'éléments

Si E admet une base, alors toutes les bases de E ont le même nombre d'éléments.

Exercice de cours 23. Démontrer la proposition à l'aide de la proposition 16.

Définition 18 – dimension

Soit $(E, +, .)$ un espace vectoriel non nul. S'il existe une base de E, on appelle *dimension* de E, et on note $\dim E$, le nombre commun d'éléments de toutes les bases de E. Dans ce cas, on dit que E est *de dimension finie*. Sinon on dit que E est de *dimension infinie*, et on note $\dim E = +\infty$.

Remarque 7. L'espace vectoriel $E = \{0_E\}$ n'admet pas de famille libre puisque toute famille de E contient 0_E. Par convention, on dit pourtant que E a pour dimension 0.

Par analogie avec la situation dans $E = \mathbb{R}^3$, les sous-espaces vectoriels de dimension 1 sont appelés «*droites*», et les sous-espaces vectoriels de dimension 2 sont appelés «*plans*». Les sous-espaces vectoriels de dimension $n-1$ d'un espace de dimension n sont appelés *hyperplans*.

Exercice de cours 24. Soit $(E, +, .)$ un espace vectoriel non nul. Montrer que les droites de E sont les sous-espaces vectoriels de la forme $\mathrm{Vect}(u)$ où u est un vecteur non nul de E.

Exercice de cours 25. Montrer que $\mathbb{R}_n[X]$ est de dimension finie. Quelle est sa dimension ?

⚠ Il existe des espaces vectoriels qui ne sont pas de dimension finie.

Par exemple l'espace $\mathbb{R}[X]$ des polynômes à coefficients réels. Supposons en effet qu'il existe une base $\mathscr{B} = (P_1, P_2, \dots, P_k)$ de $\mathbb{R}[X]$, et notons n le maximum des degrés des polynômes de $\mathscr{B}$. Puisque $X^{n+1} \in \mathbb{R}[X]$, il existe des réels $\lambda_1, \dots, \lambda_k$ tel que

$$X^{n+1} = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_k P_k.$$

Mais le polynôme du membre de droite est de degré inférieur ou égal à n, alors que celui du membre de gauche est de degré $n+1$, ce qui est absurde. On a donc

$$\dim \mathbb{R}[X] = +\infty.$$

Exercice de cours 26 (un lemme utile sur les familles libres). Soient $(v_1, v_2, \dots, v_k)$ une famille libre de E et v un vecteur de E n'appartenant pas à $\mathrm{Vect}(v_1, v_2, \dots, v_k)$. Montrer que la famille $(v_1, v_2, \dots, v_k, v)$ est encore libre.

Voici une autre conséquence importante de la proposition 16.

Proposition 19 – inclusion et dimension

Soit E un espace vectoriel de dimension finie. Soient V et W deux sous-espaces vectoriels de E, tels que $V \subset W$. Alors V et W sont de dimension finie et $\dim V \leqslant \dim W$. De plus $V = W$ si et seulement si $\dim V = \dim W$.

Exercice de cours 27. Le but de l'exercice est de démontrer la proposition 19.

1. Démontrer qu'un sous-espace vectoriel V d'un espace E de dimension finie est de dimension finie. *Indication : remarquer que si $V \neq \{0_E\}$, V admet des familles libres, et considérer une famille libre avec un nombre maximal d'éléments.*
2. Supposons maintenant que $V \subset W$ soient des sous-espaces vectoriels de E. Montrer que $\dim V \leqslant \dim W$.
3. Montrer que si $\dim V = \dim W$, alors $V = W$. Conclure.

6.3. Le théorème de la base incomplète. On montre maintenant deux résultats importants pour les espaces vectoriels de dimension finie. D'une part, on peut toujours extraire une base d'une famille génératrice, et d'autre part, on peut toujours compléter une famille libre de manière à obtenir une base. C'est cette deuxième propriété qui porte le nom de «théorème de la base incomplète».

On commence par un exercice.

Exercice de cours 28. Soient $v_1 = (1, 2, 3, 0)$, $v_2 = (0, 2, 2, 4)$ et $v_3 = (-1, a, 2a-3, 4a)$ trois vecteurs de $\mathbb{R}^4$, et $F = \mathrm{Vect}(v_1, v_2, v_3)$. Par définition, la famille (v_1, v_2, v_3) est une famille génératrice de F. On veut trouver une base de F.

1. Montrer que (v_1, v_2, v_3) est une base de F si et seulement si $a \neq 2$.

2. Si $a = 2$, montrer que (v_1, v_2) est une base de F.

Dans l'exercice précédent, on a obtenu une base à partir d'un système générateur, en gardant les vecteurs correspondants aux inconnues principales. C'est un fait général :

Proposition 20 – de toute famille génératrice on peut extraire une base

Soient V un sous-espace vectoriel de $\mathbb{R}^n$, et $\mathscr{F} = (v_1, \ldots, v_p)$ une famille génératrice de V qui n'est pas une base de V. On peut extraire de $\mathscr{F}$ une base de V : il suffit d'échelonner le système correspondant à l'équation

$$\lambda_1 v_1 + \cdots + \lambda_p v_p = 0, \tag{4}$$

d'inconnues $\lambda_1, \ldots, \lambda_p$. Les vecteurs $v_{j_1}, \ldots, v_{j_r}$ correspondant aux inconnues principales constituent alors une base de V. En particulier la dimension de V est égale au rang r du système.

On veut maintenant compléter une famille libre en une base. Le procédé vu ci-dessus permet de prouver le résultat suivant, très important :

Théorème 21 – théorème de la base incomplète

Soient $(E, +, .)$ un espace vectoriel de dimension finie et $\mathscr{B}$ une base de E. Toute famille libre de E qui n'est pas une base peut être complétée en une base de E en lui adjoignant des vecteurs de $\mathscr{B}$.

Exercice de cours 29. Le but de l'exercice est de démontrer le théorème de la base incomplète. Soient $(v_1, v_2, \ldots, v_k)$ une famille libre de E, et $\mathscr{B} = (w_1, \ldots, w_n)$ une base de E.

1. Pourquoi la famille $\mathscr{F} = (v_1, v_2, \ldots v_k, w_1, \ldots, w_n)$ engendre-t-elle E ?

2. En considérant le système homogène de n équations à $k + n$ inconnues dont les coefficients sont les coordonnées des vecteurs de $\mathscr{F}$ dans la base $\mathscr{B}$, associé à l'équation

$$x_1 v_1 + x_2 v_2 + \cdots + x_k v_k + y_1 w_1 + \cdots + y_n w_n = 0_E,$$

montrer que l'on peut trouver une base de E qui commence par les v_j. Conclure.

Exercice de cours 30 (Une autre démonstration du théorème de la base incomplète). Démontrer le théorème de la base incomplète par récurrence sur $n-k$, où k et n sont comme dans l'exercice précédent.

Voici une conséquence utile des deux résultats de ce paragraphe.

Proposition 22 – une condition suffisante pour qu'une famille libre/génératrice soit une base

Dans un espace vectoriel de dimension finie n, toute famille libre qui a n éléments est une base. Toute famille génératrice qui a n éléments est une base.

Exercice de cours 31. Démontrer cette proposition à l'aide des propositions 16 et 20.

Exercice de cours 32. Soit n un entier naturel. Pour tout $k \in \{0, \ldots, n\}$, on suppose que P_k est un polynôme à coefficients réels de degré k. Justifier que $(P_0, \ldots, P_n)$ est une base de $\mathbb{R}_n[X]$.

7. Bases et équations d'un sous-espace vectoriel

Dans la pratique, un sous-espace vectoriel F peut être donné de deux manières : par la donnée d'un système linéaire homogène dont l'ensemble des solutions est F, ou bien par la donnée d'une base de F ou même seulement d'une famille génératrice. Il est vraiment essentiel de savoir passer d'un mode de définition à l'autre.

7.1. Comment trouver un système d'équations d'un sous-espace vectoriel donné par une famille génératrice ? Soit E un espace vectoriel de dimension n. Supposons qu'un sous-espace vectoriel V de E soit donné comme l'espace engendré par une famille $(v_1, v_2, \ldots, v_p)$ de vecteurs de E. Par définition, un vecteur $v \in E$ appartient à V si et seulement s'il existe des réels $\lambda_1, \lambda_2, \ldots, \lambda_p$ tels que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_p v_p.$$

Fixons une base $\mathscr{B}$ de E. Pour $j = 1, \ldots, p$, on note $a_{1,j}, a_{2,j}, \ldots, a_{n,j}$ les coordonnées de v_j dans la base $\mathscr{B}$, et $x_1, x_2, \ldots, x_n$ celles de v. L'équation ci-dessus s'écrit alors

$$\left\{\begin{array}{ccc} a_{1,1}\lambda_1 + a_{1,2}\lambda_2 + \cdots + a_{1,p}\lambda_p & = & x_1, \\ a_{2,1}\lambda_1 + a_{2,2}\lambda_2 + \cdots + a_{2,p}\lambda_p & = & x_2, \\ \vdots & & \vdots \\ a_{n,1}\lambda_1 + a_{n,2}\lambda_2 + \cdots + a_{n,p}\lambda_p & = & x_n, \end{array}\right.$$

et le vecteur $v \in E$ appartient à V si et seulement si ce système admet une solution. Procédons comme d'habitude : pour savoir si c'est le cas, on écrit le système sous forme échelonnée réduite. On obtient un système équivalent de la forme

$$\left\{\begin{array}{cccccccccccc} \lambda_{k(1)} & & & & + & a'_{1,k(r)+1}\lambda_{k(r)+1} & + & \ldots & + & a'_{1,p}\lambda_p & = & x'_1, \\ & \lambda_{k(2)} & & & + & a'_{2,k(r)+1}\lambda_{k(r)+1} & + & \ldots & + & a'_{2,p}\lambda_p & = & x'_2, \\ & & \ddots & & & \vdots & & & & \vdots & & \vdots \\ & & & \lambda_{k(r)} & + & a'_{r,k(r)+1}\lambda_{k(r)+1} & + & \ldots & + & a'_{r,p}\lambda_p & = & x'_r, \\ & & & & & & & & & 0 & = & x'_{r+1}, \\ & & & & & & & & & \vdots & & \vdots \\ & & & & & & & & & 0 & = & x'_n, \end{array}\right.$$

où r est le rang du système, et les x'_j des combinaisons linéaires des x_j. Notons d'ailleurs $b_{i,j}$, pour $i \in \{r+1, r+2, \ldots, n\}$, les réels tels que

$$x'_i = \sum_{j=1}^{n} b_{i,j} x_j.$$

Les $b_{i,j}$ sont parfaitement déterminés et ne dépendent que des $a_{i,j}$. On a déjà vu que le système ci-dessus admet des solutions si et seulement si les $n-r$ dernières lignes, qui s'écrivent

$$\left\{\begin{array}{cccccccc} b_{r+1,1}x_1 & + & b_{r+1,2}x_2 & + & \ldots & + & b_{r+1,n}x_n & = & 0, \\ b_{r+2,1}x_1 & + & b_{r+2,2}x_2 & + & \ldots & + & b_{r+2,n}x_n & = & 0, \\ & & & \vdots & & & & \vdots & \\ b_{n,1}x_1 & + & b_{n,2}x_2 & + & \ldots & + & b_{n,n}x_n & = & 0, \end{array}\right.$$

sont vraies. Autrement dit, $x = (x_1, \ldots, x_n)$ appartient à V si et seulement si $(x_1, \ldots, x_n)$ est une solution du système homogène ci-dessus, qui est donc ce que l'on cherche.

On peut remarquer que, compte tenu de la proposition 20 ou comme on le verra dans la proposition 23, la dimension de V est $r = n-(n-r)$, le rang du système initial, qui d'ailleurs est inférieur ou égal à n comme il se doit pour un sous-espace vectoriel de E.

Exercice de cours 33. On reprend l'exercice 28. Soit $a \in \mathbb{R}$. On considère $v_1 = (1, 2, 3, 0)$, $v_2 = (0, 2, 2, 4)$ et $v_3 = (-1, a, 2a-3, 4a)$, trois vecteurs de $\mathbb{R}^4$, et $V = \mathrm{Vect}(v_1, v_2, v_3)$. Trouver un système d'équations pour V.

7.2. Comment trouver une base d'un sous-espace vectoriel donné par un système d'équations ? On sait que l'ensemble des solutions d'un système homogène de n équations à p inconnues est un sous-espace vectoriel de $\mathbb{R}^p$. On se pose la question d'en trouver une base, et, ainsi, de déterminer sa dimension.

Exemple 16. Soit $V = \{(x_1, x_2, x_3) \in \mathbb{R}^3,\ 2x_1 + 3x_2 = 0, x_1 + x_2 + x_3 = 0\}$. L'ensemble V est l'ensemble des solutions du système

$$\left\{\begin{array}{ccccccc} 2x_1 & + & 3x_2 & & & = & 0, \\ x_1 & + & x_2 & + & x_3 & = & 0. \end{array}\right.$$

On l'échelonne avec la méthode du pivot. On obtient ($L_2 \leftarrow L_2 - \frac{1}{2}L_1$) le système équivalent

$$\left\{\begin{array}{ccccccc} 2x_1 & + & 3x_2 & & & = & 0, \\ & - & \frac{1}{2}x_2 & + & x_3 & = & 0. \end{array}\right.$$

Les inconnues principales sont x_1 et x_2. On peut les exprimer en fonction de l'inconnue secondaire x_3 en écrivant le système sous forme échelonnée réduite ($L_1 \leftarrow L_1 + 6L_2$, $L_1 \leftarrow \frac{1}{2}L_1$, $L_2 \leftarrow -2L_2$) :

$$\left\{\begin{array}{ccccccc} x_1 & & & + & 3x_3 & = & 0, \\ & & x_2 & - & 2x_3 & = & 0. \end{array}\right.$$

Le système est de rang 2, ses inconnues principales sont x_1 et x_2, et la seule inconnue secondaire est x_3. Le solutions s'obtiennent en donnant une valeur quelconque à x_3, disons s_3. En fonction de cette valeur, les solutions s'écrivent

$$(x_1, x_2, x_3) = (-3s_3, 2s_3, s_3) = s_3(-3, 2, 1).$$

Autrement dit $V = \mathrm{Vect}((-3, 2, 1))$, et V est de dimension $1 = p - r$ où $p = 3$ est le nombre d'inconnues et $r = 2$ le rang du système.

Exemple 17. Soit $V \subset \mathbb{R}^5$ le sous-espace vectoriel des solutions du système

$$\left\{\begin{array}{ccccccccc} x_2 & & & - & x_4 & - & x_5 & = & 0, \\ & & x_3 & - & 2x_4 & + & x_5 & = & 0. \end{array}\right.$$

Ce sytème étant déjà sous forme échelonnée réduite, on voit que son rang est $r = 2$, que ses inconnues principales sont x_2 et x_3 et que ses inconnues secondaires sont x_1, x_4 et x_5. Les solutions s'obtiennent en donnant des valeurs quelconques à x_1, x_4 et x_5, disons s_1, s_4 et s_5 respectivement. En fonction de ces valeurs, les solutions s'écrivent

$$\begin{aligned}(x_1, x_2, x_3, x_4, x_5) &= (s_1, s_4 + s_5, 2s_4 - s_5, s_4, s_5) \\ &= s_1(1, 0, 0, 0, 0) + s_4(0, 1, 2, 1, 0) + s_5(0, 1, -1, 0, 1).\end{aligned}$$

Autrement dit, les vecteurs $v_1 = (1, 0, 0, 0, 0)$, $v_4 = (0, 1, 2, 1, 0)$ et $v_5 = (0, 1, -1, 0, 1)$ forment une base de V, qui est de dimension $3 = p - r$, où $p = 5$ est le nombre d'inconnues du système et $r = 2$ son rang.

Tout cela est complètement général :

Proposition 23 – obtention d'une base de l'ensemble des solutions d'un système homogène

Soit V un sous-espace vectoriel de $\mathbb{R}^p$ donné comme l'ensemble des solutions d'un système homogène à p inconnues. Pour trouver une base de V, il suffit d'écrire le système sous forme échelonnée réduite. En particulier, la dimension de V est $p - r$, où r est le rang du système.

La table 2 fournit un récapitulatif des différents savoir-faire vus dans les sections 6 et 7.

8. Intersection et somme de sous-espaces vectoriels

Pour fixer les idées, on rappelle quelques définitions.

	Données	Problème	Méthode
1	Famille génératrice $(u_1,\ldots,u_n)$	Extraire une base de $\mathrm{Vect}(u_1,\ldots,u_n)$	Échelonner le système correspondant à $\lambda_1 u_1+\cdots+\lambda_n u_n=0$. Identifier les inconnues principales.
2	Famille génératrice $(u_1,\ldots,u_n)$	Trouver un système d'équations pour $\mathrm{Vect}(u_1,\ldots,u_n)$	Échelonner le système correspondant à $\lambda_1 u_1+\cdots+\lambda_n u_n=v$. Trouver les conditions de compatibilité.
3	Système homogène	Trouver une base de l'espace des solutions	Échelonner et réduire le système. Identifier les inconnues secondaires.
4	Famille libre $\mathscr{F}$ et base $\mathscr{B}$	Compléter $\mathscr{F}$ avec des vecteurs de $\mathscr{B}$ pour obtenir une base	Appliquer le (1) à la famille génératrice $\mathscr{F}\cup\mathscr{B}$.

TABLE 2. Savoir-faire

Définition 24 – réunion et intersection de deux parties

Soient A et B deux parties d'un ensemble E.
— L'*intersection de A et B* est la partie $A\cap B$ de E définie par
$$A\cap B=\{x\in E\colon x\in A \text{ et } x\in B\}.$$
— La *réunion de A et B* est la partie $A\cup B$ de E définie par
$$A\cup B=\{x\in E\colon x\in A \text{ ou } x\in B\}.$$

8.1. Intersection de sous-espaces. On sait que l'ensemble des solutions d'un système linéaire homogène à p inconnues est un sous-espace vectoriel de $\mathbb{R}^p$. Cela vaut quel que soit le nombre d'équations : en particulier l'ensemble des solutions d'une seule équation linéaire homogène à p inconnues est un sous-espace vectoriel de $\mathbb{R}^p$. La résolution d'un système de plusieurs équations consiste donc à trouver l'intersection des espaces vectoriels définis par chacune de ces équations.

Dans le cas de $\mathbb{R}^p$, on a donc pratiquement déjà prouvé la proposition suivante.

Proposition 25 – l'intersection de deux sous-espaces vectoriels est un sous-espace vectoriel

Soient $(E,+,.)$ un espace vectoriel, et V, W deux sous-espaces vectoriels de E. Leur intersection de $V\cap W$ est un sous-espace vectoriel de E.

EXERCICE DE COURS 34. Démontrer la proposition de façon «directe». Donner une interprétation géométrique de ce résultat dans $\mathbb{R}^3$.

Il n'est pas plus difficile de montrer qu'une intersection quelconque de sous-espaces vectoriels est encore un sous-espace vectoriel.

EXERCICE DE COURS 35 (une caractérisation du sous-espace vectoriel engendré par une famille finie). Soient $(E,+,.)$ un sous-espace vectoriel, n un entier naturel non nul et $(u_1,u_2,\ldots,u_n)$ une famille de vecteurs de E. Montrer que le sous-espace vectoriel $\mathrm{Vect}(u_1,u_2,\ldots,u_n)$ est l'intersection de tous les sous-espaces vectoriels de E contenant $u_1,u_2,\ldots,u_n$. *On retrouve ainsi que* $\mathrm{Vect}(u_1,u_2,\ldots,u_n)$ *est le plus petit (pour l'inclusion) des sous-espaces vectoriels de E qui contient u_1,..., u_n (voir la remarque 5).*

Proposition 26 – majoration de la dimension de l'intersection de deux sous-espaces vectoriels

Soient $(E,+,.)$ un espace vectoriel de dimension finie, et V, W deux sous-espaces vectoriels de E. On a
$$\dim(V\cap W)\leqslant\min(\dim V,\dim W).$$
De plus s'il y a égalité, c'est que l'un des sous-espaces V ou W est inclus dans l'autre.

EXERCICE DE COURS 36. Démontrer la proposition à l'aide de la proposition 19.

8.2. Somme de sous-espaces.

⚠ La réunion deux deux sous-espaces vectoriels n'est en général pas un sous-espace vectoriel.

Par exemple si D_1 et D_2 sont deux droites distinctes de l'espace des vecteurs du plan, u_1 un vecteur non nul de D_1 et u_2 un vecteur non nul de D_2, le vecteur $u_1 + u_2$ n'est pas dans $D_1 \cup D_2$: cet ensemble n'est donc pas stable pour l'addition (voir la figure 1).

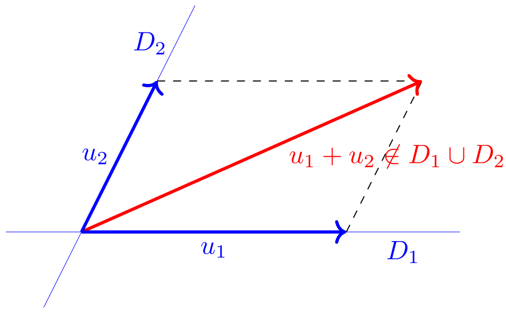


FIGURE 1. La réunion de deux droites vectorielles n'est pas stable pour l'addition en général

On définit maintenant le plus petit sous-espace vectoriel qui contient la réunion de deux sous-espaces vectoriels donnés.

Définition 27 – somme de deux sous-espaces vectoriels

Soient $(E, +, .)$ un espace vectoriel, et V, W deux sous-espaces vectoriels de E. On appelle somme de V et W, et on note $V + W$ le sous-ensemble de E défini par

$$V + W = \{x + y \colon x \in V,\, y \in W\}.$$

Proposition 28 – la somme de deux sous-espaces vectoriels est un sous-espace vectoriel

Si V et W sont deux sous-espaces vectoriels de E, alors $V + W$ est aussi un sous-espace vectoriel de E. De plus, si F est un sous-espace vectoriel qui contient V et W, alors F contient $V + W$.

EXERCICE DE COURS 37. Démontrer cette proposition.

EXEMPLE 18. Dans $E = \mathbb{R}^3$, on note $a = (0, 1, 2)$, $b = (1, 0, 1)$ et $c = (1, 1, 1)$. Soient aussi $V = \mathrm{Vect}(a, b)$ et $W = \mathrm{Vect}(c)$. L'espace vectoriel $V + W$ est constitué des vecteurs v qui s'écrivent

$$v = (\lambda a + \mu b) + \nu c = \lambda a + \mu b + \nu c.$$

Donc $V + W$ est engendré par la famille (a, b, c). D'ailleurs puisque cette famille est libre, $V + W = \mathbb{R}^3$: l'espace $\mathbb{R}^3$ s'écrit comme somme du plan V et de la droite W.

Proposition 29 – majoration de la dimension de la somme de deux sous-espaces vectoriels

Si $V = \mathrm{Vect}(v_1, v_2, \dots v_k)$ et $W = \mathrm{Vect}(w_1, w_2, \dots w_\ell)$, alors $V + W = \mathrm{Vect}(v_1, v_2, \dots v_k, w_1, w_2, \dots w_\ell)$. En particulier, $\dim(V + W) \leqslant \dim V + \dim W$.

EXERCICE DE COURS 38. Notons $Y = \mathrm{Vect}(v_1, v_2, \dots, v_k, w_1, w_2, \dots, w_\ell)$.

1. À l'aide de la proposition 28, montrer l'inclusion $V + W \subset Y$.
2. Montrer l'inclusion $Y \subset V + W$.

3. Démontrer la proposition 29.

Puisque $(v_1, \dots v_k, w_1, \dots w_\ell)$ est un système générateur de $V+W$, la proposition 20, permet de déterminer une base de $V+W$ (et sa dimension) :

Proposition 30 – construction d'une base de la somme de deux sous-espaces vectoriels

Soient $(v_1, \dots v_k)$ une base du sous-espace vectoriel V, et $(w_1, \dots, w_\ell)$ une base du sous-espace vectoriel W. Les vecteurs correspondant aux inconnues principales du système

$$x_1 v_1 + x_2 v_2 + \cdots + x_k v_k + y_1 w_1 + \cdots + y_\ell w_\ell = 0_E,$$

forment une base de $V+W$.

8.3. Somme directe.

Proposition 31 – notion de composante lorsque cela a un sens

Soient $(E, +, .)$ un espace vectoriel, et V, W deux sous-espaces vectoriels de E. Tout vecteur y de $V+W$ s'écrit de manière unique $y = v + w$ avec $v \in V$ et $w \in W$ si et seulement si $V \cap W = \{0_E\}$. Dans ce cas, on dit que v est la *composante de y dans V* et w la *composante de y dans W*.

Exercice de cours 39. Démontrer la proposition.

Définition 32 – somme directe

Soient $(E, +, .)$ un espace vectoriel de dimension finie, et V, W deux sous-espaces vectoriels de E.

1. On dit que V et W sont en *somme directe* lorsque $V \cap W = \{0_E\}$. Dans ce cas on note $\boxed{F = V \oplus W}$ le sous-espace vectoriel $F = V + W$.

2. Si de plus $V \oplus W = E$, on dit que V et W sont *supplémentaires dans E*.

Le critère suivant sera très utile lorsque nous parlerons de *projections* (voir par exemple la section 6.1).

Proposition 33 – caractérisation pour que deux sous-espaces vectoriels soient en somme directe

Soient $(E, +, .)$ un espace vectoriel de dimension finie, et V, W deux sous-espaces vectoriels de E. Soit $(v_1, v_2, \dots, v_p)$ une base de V et $(w_1, w_2, \dots, w_q)$ une base de W.

1. V et W sont en somme directe si et seulement si $(v_1, v_2, \dots, v_p, w_1, w_2, \dots, w_q)$ est une famille libre.

2. V et W sont supplémentaires dans E si et seulement si $(v_1, v_2, \dots, v_p, w_1, w_2, \dots, w_q)$ est une base de E.

Exercice de cours 40. Démontrer cette proposition.

8.4. Formule de la dimension. La proposition suivante est parfois connue sous le nom de *formule de Grassmann.*

Hermann Günther Grassmann *(1809-1877) est un mathématicien et indianiste allemand. Polymathe, il était connu de ses contemporains en tant que linguiste. Physicien, néo-humaniste, érudit mais aussi éditeur, Hermann Grassmann est avec Niels Abel, Évariste Galois et Georg Cantor l'un des grands mathématiciens «malheureux» du XIX^e siècle. Il est considéré aujourd'hui comme le fondateur du calcul tensoriel et de la théorie des espaces vectoriels.*

Proposition 34 – formule de la dimension ou formule de Grassmann

Soient $(E, +, .)$ un espace vectoriel de dimension finie, et V, W deux sous-espaces vectoriels de E. On a toujours
$$\dim(V + W) = \dim V + \dim W - \dim V \cap W.$$

Exercice de cours 41. Démontrer cette proposition.

8.5. Comment construire une base de l'intersection de deux sous-espaces vectoriels ? La preuve ci-dessus donne en fait un renseignement plus précis que la proposition : elle donne un procédé de construction d'une base de $V + W$. On peut même en extraire un procédé de construction d'une base de $V \cap W$ connaissant une base de V et une base de W, ce que nous faisons maintenant.

Soient $\mathscr{B}_V = (v_1, v_2, \ldots, v_k)$ une base de V, et $\mathscr{B}_W = (w_1, w_2, \ldots, w_\ell)$ une base de W. Un vecteur $x \in E$ appartient à $V \cap W$ si et seulement s'il appartient à V et à W, c'est-à-dire si et seulement s'il existe des réels $\lambda_1, \lambda_2, \ldots, \lambda_k, \tilde{\mu}_1, \tilde{\mu}_2, \ldots, \tilde{\mu}_\ell$ tels que
$$x = \lambda_1 v_1 + \ldots \lambda_k v_k \qquad \text{et} \qquad x = \tilde{\mu}_1 w_1 + \cdots + \tilde{\mu}_\ell w_\ell.$$
Autrement dit, $x \in V \cap W$ si et seulement s'il existe des réels $\lambda_1, \lambda_2, \ldots, \lambda_k, \mu_1, \mu_2, \ldots, \mu_\ell$ tels que (on notera que l'on a remplacé les $\tilde{\mu}_j$ par $\mu_j = -\tilde{\mu}_j$),
$$x = \lambda_1 v_1 + \ldots \lambda_k v_k \qquad \text{et} \qquad \lambda_1 v_1 + \ldots \lambda_k v_k + \mu_1 w_1 + \cdots + \mu_\ell w_\ell = 0.$$
Soit alors (S) le système linéaire homogène associé à cette deuxième équation, dont les inconnues sont les λ_j et les μ_j. On notera qu'il s'agit du même système que celui qui permet de trouver une base de $V + W$ (cf. la proposition 30).

On écrit (S) sous forme échelonnée réduite à l'aide de la méthode de Gauss, et l'on note $e_1, \ldots, e_r$ les vecteurs correpondant aux inconnues principales $\pi_1, \ldots, \pi_r$, et $e_{r+1}, \ldots, e_{r+q}$ les vecteurs correspondants aux inconnues secondaires $\sigma_{r+1}, \ldots, \sigma_{r+q}$ de (S). Il est important de remarquer que

— On a $r + q = k + \ell$, ou encore $q = k + \ell - r$.

— Pour $j \in \{1, \ldots, k\}$, $e_j = v_j$. En effet, si l'on s'arrête aux k premières colonnes, les vecteurs correspondant aux inconnues principales donnent une base de V extraite de $\mathscr{B}_V$, qui ne peut donc être que $\mathscr{B}_V$.

— Pour $j \in \{k + 1, \ldots, k + \ell\}$, les e_j sont donc des vecteurs de $\mathscr{B}_W$, rangés dans un ordre différent : ceux qui correspondent aux inconnues principales d'abord, pour $j \in \{k + 1, \ldots, r\}$, puis ceux qui correspondent aux inconnues secondaires pour $j \in \{r + 1, \ldots, k + \ell\}$.

En particulier, si $(\pi_1, \ldots, \pi_r, \sigma_{r+1}, \ldots, \sigma_{r+q})$ est une solution de (S), il lui correspond l'élément $x \in V \cap W$ donné par
$$x = \sum_{j=1}^{k} \pi_j e_j = -\sum_{j=k+1}^{r} \pi_j e_j - \sum_{j=r+1}^{r+q} \sigma_j e_j.$$

On sait que l'on obtient une solution de (S), donc un élément de $V \cap W$, à chaque fois que l'on fixe la valeur des inconnues secondaires $\sigma_{r+1}, \ldots, \sigma_{r+q}$. On note donc $x_1, x_2, \ldots, x_q$ les éléments de $V \cap W$ qu'on obtient en prenant successivement
$$\begin{array}{rcl} (\sigma_{r+1}, \ldots, \sigma_{r+q}) & = & (1, 0, 0, \ldots, 0), \\ (\sigma_{r+1}, \ldots, \sigma_{r+q}) & = & (0, 1, 0, \ldots, 0), \\ & \vdots & \\ (\sigma_{r+1}, \ldots, \sigma_{r+q}) & = & (0, 0, \ldots, 0, 1). \end{array}$$

***Johann Carl Friedrich Gauss**, né le 30 avril 1777 à Brunswick et mort le 23 février 1855 à Göttingen, est un mathématicien, astronome et physicien allemand. Doté d'un grand génie, il a apporté de très importantes contributions à ces trois sciences. Surnommé «le prince des mathématiciens», il est considéré comme l'un des plus grands mathématiciens de tous les temps.*

Proposition 35

La famille $\mathscr{F} = \{x_1, \dots, x_q\}$ ainsi construite est une base de $V \cap W$.

Exercice de cours 42. On considère les deux sous-espaces vectoriels V et W de $\mathbb{R}^4$ définis par

$$V = \mathrm{Vect}((1,0,1,1),(5,-1,6,1)) \text{ et } W = \mathrm{Vect}((1,1,2,0),(1,2,2,0),(3,1,6,0)).$$

Donner une base de V, de W, puis une base de $V + W$ et une base de $V \cap W$.

2

Matrices

Ce chapitre est pour l'essentiel un chapitre de révisions sur les matrices vues au premier semestre.

1. L'espace vectoriel des matrices

Définition 36 – matrice

Rappelons qu'une *matrice* à n lignes et p colonnes est un tableau de nombres de la forme

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix}.$$

Les nombres $a_{i,j}$ sont appelés *coefficients* de la matrice A, et on note souvent

$$A = (a_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p},$$

ou même $A = (a_{i,j})$ quand il n'y a pas de confusion possible. Les nombres n et p (dans cet ordre) sont appelés dimensions de la matrice.

On note $\mathscr{M}_{n,p}(\mathbb{R})$ l'ensemble des matrices à n lignes et p colonnes de nombres réels. Si les dimensions d'une matrice sont égales, i.e., s'il y a autant de lignes que de colonnes, on dit que la matrice est *carrée*, et on dit que la matrice est *d'ordre n*. On note simplement $\mathscr{M}_n(\mathbb{R})$ l'ensemble des matrices à n lignes et n colonnes.

EXEMPLE 19. **1.** On appelle matrice nulle la matrice dont tous les coefficients sont nuls :

$$0 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$$

2. La matrice d'un système linéaire à n équations et p inconnues

$$\left\{ \begin{array}{lcl} a_{1,1}X_1 + a_{1,2}X_2 + \cdots + a_{1,p}X_p & = & b_1, \\ a_{2,1}X_1 + a_{2,2}X_2 + \cdots + a_{2,p}X_p & = & b_2, \\ \vdots & \vdots & \vdots \\ a_{n,1}X_1 + a_{n,2}X_2 + \cdots + a_{n,p}X_p & = & b_n, \end{array} \right.$$

est le tableau à n lignes et p colonnes dont les coefficients sont les $a_{i,j}$:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix}.$$

La matrice complète du système est la matrice à n lignes et $p+1$ colonnes définie par

$$\left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \dots & a_{1,p} & b_1 \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n,1} & a_{n2} & \dots & a_{n,p} & b_p \end{array}\right).$$

On notera la présence de la barre verticale pour séparer les seconds membres du reste de la matrice.

3. On appelle *matrice élémentaire* toute matrice dont tous les coefficients sont nuls sauf l'un d'entre eux qui vaut 1. On note

$$E_{i,j} = \begin{pmatrix} & & \vdots & \\ & & \vdots & \\ \dots & \dots & 1 & \dots \\ & & \vdots & \end{pmatrix}$$

la matrice élémentaire dont l'élément non nul est sur la ligne i et la colonne j.

4. Lorsqu'une matrice n'a qu'une ligne, on parle de matrice-ligne. Lorsqu'elle n'a qu'une colonne, on parle de matrice-colonne. Par exemple on identifie souvent un vecteur v de $\mathbb{R}^n$ avec la matrice colonne de ses coordonnées dans la base canonique $(e_1, e_2, \dots, e_n)$ de $\mathbb{R}^n$: si $v = (x_1, x_2, \dots, x_n)$, on a $v = x_1e_1 + x_2e_2 + \dots + x_ne_n$ et on écrit (voir la proposition 15)

$$v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Définition 37 – somme de deux matrices

Soient $A = (a_{i,j})$ et $B = (b_{i,j})$ deux matrices de $\mathscr{M}_{n,p}(\mathbb{R})$. On note $C = A + B$ la matrice de $\mathscr{M}_{n,p}(\mathbb{R})$ définie par $C = (a_{i,j} + b_{i,j})$. Plus explicitement

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix} + \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,p} \\ b_{2,1} & b_{2,2} & \dots & b_{2,p} \\ \vdots & \vdots & & \vdots \\ b_{n,1} & b_{n,2} & \dots & b_{n,p} \end{pmatrix} =$$

$$\begin{pmatrix} a_{1,1}+b_{1,1} & a_{1,2}+b_{1,2} & \dots & a_{1,p}+b_{1,p} \\ a_{2,1}+b_{2,1} & a_{2,2}+b_{2,2} & \dots & a_{2,p}+b_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1}+b_{n,1} & a_{n,2}+b_{n,2} & \dots & a_{n,p}+b_{n,p} \end{pmatrix}.$$

On a ainsi défini une loi de composition interne «+» sur l'ensemble $\mathscr{M}_{n,p}(\mathbb{R})$. Comme il s'agit d'ajouter les coefficients terme à terme (donc de faire des additions de nombres réels), il est très facile de vérifier que cette loi «+» vérifie les quatre premières propriétés de la définition 1. En particulier l'élément neutre pour l'addition des matrices est la matrice nulle, et la matrice symétrique d'une matrice $A = (a_{i,j})$ est la matrice, notée $-A$, définie par

$$-A = \begin{pmatrix} -a_{1,1} & -a_{1,2} & \dots & -a_{1,p} \\ -a_{2,1} & -a_{2,2} & \dots & -a_{2,p} \\ \vdots & \vdots & & \vdots \\ -a_{n,1} & -a_{n,2} & \dots & -a_{n,p} \end{pmatrix}.$$

Pour résumer :

Proposition 38

$(\mathscr{M}_{n,p}(\mathbb{R}), +)$ est un groupe commutatif.

On définit maintenant une loi de composition externe «.» sur l'ensemble $\mathscr{M}_{n,p}(\mathbb{R})$.

Définition 39 – loi externe sur l'ensemble des matrices

Soit $A = (a_{i,j})$ une matrice de $\mathscr{M}_{n,p}(\mathbb{R})$ et λ un réel. On note $C = \lambda.A$ la matrice de $\mathscr{M}_{n,p}(\mathbb{R})$ définie par $C = (\lambda a_{i,j})$. Plus explicitement

$$\lambda \ . \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix} = \begin{pmatrix} \lambda a_{1,1} & \lambda a_{1,2} & \dots & \lambda a_{1,p} \\ \lambda a_{2,1} & \lambda a_{2,2} & \dots & \lambda a_{2,p} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n,1} & \lambda a_{n,2} & \dots & \lambda a_{n,p} \end{pmatrix}.$$

Comme il s'agit de multiplier chaque coefficient par λ, il est également facile de vérifier que cette loi «.» vérifie les 4 dernières propriétés de la définition 1.

Proposition 40 – l'ensemble des matrices est un espace vectoriel de dimension finie

$(\mathscr{M}_{n,p}(\mathbb{R}), +)$ est un espace vectoriel, de dimension np.

Exercice de cours 43. Démontrer la proposition.

Remarque 8. On notera que la famille $(E_{1,1},\, E_{1,2},\, \dots,\, E_{n,p-1},\, E_{n,p})$ n'est pas indexée par une partie de $\mathbb{N}$ mais par une partie de $\mathbb{N}\times\mathbb{N}$. Ce n'est qu'une façon de l'écrire : l'énumération $(E_{1,1}, E_{1,2}, \dots, E_{n,p-1}, E_{n,p}) = (\tilde{E}_1, \tilde{E}_2, \dots, \tilde{E}_{np})$ fournit une nouvelle numérotation de la famille pour des indices entre 1 et np. De manière générale n'importe quel ensemble fini peut servir à énumérer une base.

2. Produit de matrices

On rappelle maintenant comment est défini le produit de deux matrices. Il est naturel de penser multiplier terme à terme les coefficients de chacune, mais cette multiplication-là ne sert à rien (ou disons que c'est une autre histoire). Ce que l'on veut, c'est retrouver le membre de gauche des équations du système

$$\left\{ \begin{array}{rcl} a_{1,1}X_1 + a_{1,2}X_2 + \dots + a_{1,p}X_p & = & b_1, \\ a_{2,1}X_1 + a_{2,2}X_2 + \dots + a_{2,p}X_p & = & b_2, \\ \vdots & \vdots & \vdots \\ a_{n,1}X_1 + a_{n,2}X_2 + \dots + a_{n,p}X_p & = & b_n, \end{array} \right.$$

en multipliant sa matrice par la matrice colonne des symboles $X_1, X_2, \dots X_p$. Autrement dit, on veut que

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix} \times \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} a_{1,1}X_1 + a_{1,2}X_2 + \dots + a_{1,p}X_p \\ a_{2,1}X_1 + a_{2,2}X_2 + \dots + a_{2,p}X_p \\ \vdots \\ a_{n,1}X_1 + a_{n,2}X_2 + \dots + a_{n,p}X_p \end{pmatrix},$$

et c'est cela que l'on prend comme définition du produit d'une matrice par une matrice-colonne. Il est essentiel de noter que le nombre de lignes de la matrice colonne doit être égal au nombre de colonnes de la matrice.

On étend ensuite cette définition au cas du produit d'une matrice A par une matrice B quelconque (pourvu que le nombre de colonnes de A soit égal au nombre de lignes de B), en construisant la matrice dont la j-ième colonne est le produit de A par la j-ième colonne de B. Le produit de deux matrices se fait donc «ligne par colonne».

Exemple 20. Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 2 & 3 \end{pmatrix}$. Le produit $A \times B$ est bien défini puisque A a trois colonnes et B trois lignes. Comme A a deux lignes et B a 2 colonnes, le produit $A \times B$ sera une matrice à 2 lignes et 2 colonnes.

Le produit de A par la première colonne de B est

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1*1 & + & 2*0 & + & 3*2 \\ 0*1 & + & (-1)*0 & + & 1*2 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}.$$

Celui de A par la seconde colonne de B est

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 1 \end{pmatrix} \times \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1*(-1) & + & 2*1 & + & 3*3 \\ 0*(-1) & + & (-1)*1 & + & 1*3 \end{pmatrix} = \begin{pmatrix} 10 \\ 2 \end{pmatrix}.$$

On obtient donc

$$A \times B = \begin{pmatrix} 7 & 10 \\ 2 & 2 \end{pmatrix}.$$

On peut formaliser cette définition.

Définition 41 – produit de deux matrices

Soient $A = (a_{i,j})$ une matrice à n lignes et **p colonnes**, et $B = (b_{i,j})$ une matrice à **p lignes** et m colonnes. On note $C = A \times B$ la matrice à n lignes et m colonnes dont les coefficients $c_{i,j}$ sont donnés par

$$c_{i,j} = \sum_{k=1}^{p} a_{i,k} b_{k,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots + a_{i,p}b_{p,j}.$$

⚠ Le produit de deux matrices n'est pas commutatif. Autrement dit, en général $A \times B \neq B \times A$. Par exemple

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

D'ailleurs l'un des produits peut être bien défini alors que l'autre ne l'est pas en fonction des dimensions de A et de B.

Dans la pratique, il est commode pour calculer le produit $C = A \times B$ de deux matrices de disposer les calculs de la manière suivante

$$\begin{array}{c|c} & \begin{pmatrix} b_{1,1} & \dots & b_{1,j} & \dots & b_{1,m} \\ b_{2,1} & \dots & b_{2,j} & \dots & b_{2,m} \\ \vdots & & \vdots & & \vdots \\ b_{p,1} & \dots & b_{p,j} & \dots & b_{p,m} \end{pmatrix} \\ \hline \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ \vdots & \vdots & & \vdots \\ a_{i,1} & a_{i,2} & \dots & a_{i,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix} & \begin{pmatrix} & \vdots & \\ \dots & \displaystyle\sum_{k=1}^{p} a_{i,k} b_{k,j} & \dots \\ & \vdots & \end{pmatrix} \end{array}$$

Proposition 42 – associativité du produit de deux matrices

Le produit des matrices est associatif : si $A \in \mathscr{M}_{m,n}(\mathbb{R})$, $B \in \mathscr{M}_{n,p}(\mathbb{R})$ et $C \in \mathscr{M}_{p,q}(\mathbb{R})$, on a

$$A \times (B \times C) = (A \times B) \times C.$$

Proposition 43 – distributivité du produit matriciel par rapport à l'addition

Le produit des matrices est distributif par rapport à l'addition, à gauche et à droite : si $A \in \mathscr{M}_{n,p}(\mathbb{R})$, $B \in \mathscr{M}_{p,m}(\mathbb{R})$ et $C \in \mathscr{M}_{p,m}(\mathbb{R})$, on a

$$A \times (B + C) = A \times B + A \times C,$$

et, si $B \in \mathscr{M}_{n,p}(\mathbb{R})$, $C \in \mathscr{M}_{n,p}(\mathbb{R})$ et $A \in \mathscr{M}_{p,m}(\mathbb{R})$,

$$(B + C) \times A = B \times A + C \times A.$$

Exercice de cours 44. Montrer que pour $\lambda \in \mathbb{R}$, $A \in \mathscr{M}_{n,p}(\mathbb{R})$ et $B \in \mathscr{M}_{p,q}(\mathbb{R})$ on a

$$(\lambda A) \times B = \lambda(A \times B) = A \times (\lambda B).$$

Dans les espaces de matrices carrées $\mathscr{M}_n(\mathbb{R})$, la multiplication est une loi interne. Ce n'est bien sûr pas le cas dans $\mathscr{M}_{n,p}(\mathbb{R})$ quand $n \neq p$: la multiplication de deux éléments n'est alors même pas définie. Dans $\mathscr{M}_n(\mathbb{R})$, la multiplication a un élément neutre :

Proposition 44 – existence d'un élément neutre pour le produit matriciel

Soit $I \in \mathscr{M}_n(\mathbb{R})$ la matrice dont les éléments sont tous nuls, sauf ceux qui se trouvent sur la diagonale qui valent 1 :

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Pour toute matrice $A \in \mathscr{M}_n(\mathbb{R})$, on a

$$A \times I = I \times A = A.$$

Cette matrice I, qu'on note aussi I_n pour faire référence à sa taille, est appelée *matrice identité*. On se rappelle qu'il ne peut y avoir qu'un seul élément neutre pour une loi interne (voir l'exercice 1).

Définition 45 – puissance d'une matrice carrée

Soient $A \in \mathscr{M}_n(\mathbb{R})$ une matrice carrée, et $k \in \mathbb{N}^*$. On note A^k, et on lit A puissance k, la matrice de $\mathscr{M}_n(\mathbb{R})$ définie par

$$A^k = \underbrace{A \times A \times \cdots \times A}_{k \text{ facteurs}}.$$

On note aussi $A^0 = I$.

Exercice de cours 45. Soit $E_{k,\ell}$ une matrice élémentaire. Montrer que $(E_{k,\ell})^2 = 0$ si $k \neq \ell$, et que $(E_{k,k})^2 = E_{k,k}$.

3. Opérations sur les lignes d'une matrice

Pour résoudre les systèmes linéaires, on a été amené à utiliser trois opérations sur leurs lignes. Ces opérations n'affectent pas les inconnues, ni ne changent leur ordre, et on peut très bien effectuer ces opérations sur la matrice (complète si le système n'est pas homogène) associée au système. Il se trouve que l'on peut interpréter ces opérations sur les lignes d'une matrice comme un produit par une matrice particulière.

Proposition 46 – action de la multiplication à gauche par une matrice élémentaire

Soit $A = (a_{i,j})$ une matrice de $\mathscr{M}_{n,p}(\mathbb{R})$, et $E_{k,\ell} = (e_{i,j}) \in \mathscr{M}_n(\mathbb{R})$ la matrice élémentaire dont le seul élément non nul est $e_{k,\ell} = 1$. On a

$$E_{k,\ell} \times A = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \\ a_{\ell,1} & \dots & a_{\ell,p} \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} \leftarrow \text{ligne } k.$$

Ainsi, multiplier une matrice A par $E_{k,\ell}$ permet d'extraire la ℓ-ième ligne de A et de la placer sur la ligne k.

Proposition 47 – opérations élémentaires et multiplications par une matrice élémentaire

Soit $A \in \mathscr{M}_{n,p}(\mathbb{R})$.
- Effectuer l'opération $L_i \leftarrow L_i + \alpha L_j$ sur la matrice A, avec $i \neq j$, revient à multiplier A à gauche par $I_n + \alpha E_{i,j} \in \mathscr{M}_n(\mathbb{R})$.
- L'opération $L_i \leftarrow \alpha L_i$ correspond à multiplier A à gauche par $I_n + (\alpha - 1)E_{i,i} \in \mathscr{M}_n(\mathbb{R})$.

L'opération d'échange de deux lignes d'une matrice A peut aussi s'interpréter comme la multiplication par une matrice. Pour $i, j \in \{1, \dots, n\}$, on note $\widetilde{E}_{i,j} \in \mathscr{M}_n(\mathbb{R})$ la matrice définie par

$$\widetilde{E}_{i,j} = I_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}.$$

Il s'agit simplement de la matrice I_n dont on a déplacé les «1» qui se trouvaient à la i-ième et j-ième place sur la diagonale $(-E_{i,i} - E_{j,j})$, en les portant aux places (i, j) et (j, i) respectivement $(+E_{i,j} + E_{j,i})$.

Proposition 48 – échange entre deux lignes

Soit $A \in \mathscr{M}_{n,p}(\mathbb{R})$. La matrice $\tilde{A} = \widetilde{E}_{i,j} \times A$ est la matrice de $\mathscr{M}_{n,p}(\mathbb{R})$ obtenue en échangeant les lignes i et j de la matrice A.

Remarque 9. La proposition est en particulier valable pour $A = I_n$. Or dans ce cas

$$\tilde{A} = \widetilde{E}_{i,j} \times I_n = \widetilde{E}_{i,j}.$$

On en déduit que $\widetilde{E}_{i,j}$ est tout simplement la matrice obtenue en échangeant les lignes i et j de la matrice identité.

Remarque 10. Pour opérer sur les colonnes, on multiplie à droite par une matrice d'opération élémentaire. Autrement dit, pour opérer sur les colonnes, il suffit donc de remplacer dans la proposition 47, «ligne» par «colonne» et «gauche» par «droite». Par exemple,

$$\tilde{A} = A \times \widetilde{E}_{i,j}$$

est la matrice de $\mathscr{M}_{n,p}(\mathbb{R})$ obtenue en échangeant les colonnes i et j de la matrice A.

4. Inverse d'une matrice carrée

4.1. Définition. On se pose maintenant la question de l'existence pour une matrice $A \in \mathscr{M}_n(\mathbb{R})$ donnée, d'un symétrique pour la loi $\times$.

$$
\widetilde{E}_{i,j} = \begin{array}{c}
\begin{array}{ccccccccc} & & i & & & & j & & \\ & & \downarrow & & & & \downarrow & & \end{array} \\
\begin{pmatrix}
1 & & & & & & & & \\
& \ddots & & & & & & & \\
& & 0 & & & & 1 & & \\
& & & 1 & & & & & \\
& & & & \ddots & & & & \\
& & & & & 1 & & & \\
& & 1 & & & & 0 & & \\
& & & & & & & \ddots & \\
& & & & & & & & 1
\end{pmatrix}
\end{array}
\begin{array}{l} \\ \\ \\ \leftarrow \text{ligne } i \\ \\ \\ \\ \\ \leftarrow \text{ligne } j \\ \\ \\ \end{array}
$$

Figure 1. La matrice $\widetilde{E}_{i,j}$

Définition 49 – matrice inverse

Soit $A \in \mathscr{M}_n(\mathbb{R})$. On dit que A admet la matrice $Q \in \mathscr{M}_n(\mathbb{R})$ pour *inverse* lorsque

$$A \times Q = Q \times A = I.$$

On a déjà montré (voir l'exercice 1) qu'un élément ne peut avoir plus d'un symétrique pour une loi interne associative. Ayant en vue la Proposition 51 ci-dessous, on reprend la preuve dans le cadre actuel.

Proposition 50 – une matrice carrée admet au plus une matrice inverse

Une matrice carrée A admet au plus une matrice inverse.

Exercice de cours 46. Démontrer la proposition en utilisant seulement que Q_1 est inverse à droite de A, c'est-à-dire $A \times Q_1 = I$, et que Q_2 est inverse à gauche de A, c'est-à-dire $Q_2 \times A = I$.

L'exercice précédent a pour conséquence la proposition suivante.

Proposition 51 – une condition suffisante pour qu'une matrice admette un inverse

Si A admet une matrice inverse à droite Q_1 et une matrice inverse à gauche Q_2, alors $Q_1 = Q_2$ et A admet $Q_1 = Q_2$ comme inverse.

Lorsqu'une matrice A admet une matrice inverse, on dit que A est *inversible*, et on note A^{-1} la matrice inverse de A.

Exemple 21. **1.** La matrice I est inversible, d'inverse elle-même puisque $I \times I = I$.

2. D'après la proposition 48, la matrice $\widetilde{E}_{i,j}$ est inversible et son inverse est $\widetilde{E}_{j,i}$.

3. D'après la proposition 47, la matrice $I_n + \alpha E_{i,j}$ est inversible pour $i \neq j$, d'inverse $I_n - \alpha E_{i,j}$. D'ailleurs on peut vérifier directement que

$$(I_n + \alpha E_{i,j}) \times (I_n - \alpha E_{i,j}) = I_n - \alpha E_{i,j} + \alpha E_{i,j} - \alpha^2 E_{i,j} \times E_{i,j} = I_n,$$

puisque $E_{i,j} \times E_{i,j} = 0$ quand $i \neq j$.

Reste à trouver un moyen de savoir si une matrice donnée est inversible, et le cas échéant, de calculer son inverse.

4.2. Existence. Une manière de voir ce problème est de considérer le système linéaire $A \times X = B$ associé à la matrice A avec un second membre $B \in \mathscr{M}_{n,1}(\mathbb{R})$ quelconque, c'est-à-dire l'équation d'inconnue X qui s'écrit

$$A \times X = B.$$

S'il existe Q telle que $Q \times A = I$, cette équation entraine

$$X = (Q \times A) \times X = Q \times (A \times X) = Q \times B,$$

donc la seule solution possible est $X = Q \times B$. Si on a aussi $A \times Q = I$, alors

$$A \times (Q \times B) = (A \times Q) \times B = B,$$

ce qui prouve que $X = Q \times B$ est bien une solution du système.

Autrement dit : si A est inversible, pour tout second membre $B \in \mathscr{M}_{n,1}(\mathbb{R})$ le système de n équations à n inconnues $A \times X = B$ a une unique solution.

Examinons la réciproque : supposons que pour tout second membre $B \in \mathscr{M}_{n,1}(\mathbb{R})$ le système $A \times X = B$ a une unique solution. C'est en particulier le cas si B est (le vecteur colonne associé à) l'un des vecteurs e_j de la base canonique de $\mathbb{R}^n$, c'est à dire l'un des vecteurs

$$B_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; B_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \; \ldots, \; B_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

On note Q_1, Q_2,. . .,Q_n la solution correspondante du système, c'est-à dire l'unique vecteur tel que

$$A \times Q_j = B_j,$$

et Q la matrice dont les colonnes sont $Q_1, Q_2, \ldots, Q_n$. Par définition du produit de A par Q, on a

$$A \times Q = I,$$

donc Q est inverse à droite de A.

D'autre part, puisque le système $A \times X = B$ a une unique solution pour n'importe quel second membre B, toutes ses inconnues sont principales, donc sa forme échelonnée réduite est

$$I \times X = C.$$

On passe du système $A \times X = B$ à sa forme échelonnée réduite en effectuant des opérations du type $L_i \leftrightarrow L_j$, $L_i \leftarrow L_i + \alpha L_j$ ou $L_i \leftarrow \alpha L_i$, c'est-à-dire compte tenu des propositions 47 et 48, en multipliant A à gauche par des matrices de la forme $\widetilde{E}_{i,j}$ ou $(I + \beta E_{i,j})$ avec $\beta = \alpha$ ou $\beta = \alpha - 1$. Notant R le produit de toutes ces matrices, on a,

$$R \times A = I,$$

donc R est un inverse à gauche de A. On a démontré que si pour tout second membre $B \in \mathscr{M}_{n,1}(\mathbb{R})$ le système $A \times X = B$ a une unique solution, alors A est inversible. En résumé, on a obtenu la proposition suivante.

Proposition 52 – conditions équivalentes pour qu'une matrice carrée soit inversible

Soit $A \in \mathscr{M}_n(\mathbb{R})$. Les conditions suivantes sont équivalentes :

1. La matrice A est inversible.
2. Le système $A \times X = B$ a une unique solution pour tout second membre $B \in \mathscr{M}_{n,1}(\mathbb{R})$.
3. La matrice de la forme échelonnée réduite du système est la matrice identité I.

4.3. Calcul pratique de l'inverse. La discussion qui précède donne deux procédés pour calculer l'inverse d'une matrice inversible $A \in \mathscr{M}_n(\mathbb{R})$:

— Par résolution de n systèmes linéaires $A \times X = B_j$, $j = 1, \ldots, n$, où les B_j sont les n vecteurs de la base canonique de $\mathbb{R}^n$.

— Par transformation du système en sa forme échelonnée réduite à l'aide de la méthode du pivot, qu'on a réinterprété comme la multiplication de A par des matrices inversibles de la forme $\widetilde{E}_{i,j}$ ou $I + \beta E_{i,j}$.

En fait ces deux procédés conduisent exactement aux mêmes calculs. Commençons par le second : si l'on effectue les mêmes substitutions de lignes sur la matrice identité que celle que l'on fait sur A pour l'amener à sa forme échelonnée réduite, on multiplie la matrice identité par A^{-1}, et donc on récupère l'inverse A^{-1} de A.

Voici un exemple d'un tel calcul. On place à droite de A la matrice identité, et on effectue sur les lignes de la grande matrice obtenue toutes les opérations nécessaires pour obtenir sa forme échelonnée réduite. Si cette forme est la matrice identité, c'est que A est inversible et dans ce cas, la matrice de droite est A^{-1}.

Exemple 22. Pour savoir si la matrice $A = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 2 & -1 \end{pmatrix}$ est inversible et calculer son inverse éventuel, on écrit

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array}\right),$$

et on effectue successivement les opérations $L_3 \leftarrow L_3 - \frac{1}{2} L_1$

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 3/2 & -1/2 & -1/2 & 0 & 1 \end{array}\right),$$

$L_3 \leftarrow L_3 - \frac{3}{2} L_2$

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & -2 & -1/2 & -3/2 & 1 \end{array}\right),$$

$L_3 \leftarrow -\frac{1}{2} L_3$

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1/4 & 3/4 & -1/2 \end{array}\right),$$

$L_2 \leftarrow L_2 - L_3$

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1/4 & 1/4 & 1/2 \\ 0 & 0 & 1 & 1/4 & 3/4 & -1/2 \end{array}\right),$$

$L_1 \leftarrow L_1 + L_3$

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 0 & 5/4 & 3/4 & -1/2 \\ 0 & 1 & 0 & -1/4 & 1/4 & 1/2 \\ 0 & 0 & 1 & 1/4 & 3/4 & -1/2 \end{array}\right),$$

$L_1 \leftarrow L_1 - L_2$

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 3/2 & 1/2 & -1 \\ 0 & 1 & 0 & -1/4 & 1/4 & 1/2 \\ 0 & 0 & 1 & 1/4 & 3/4 & -1/2 \end{array}\right),$$

et enfin $L_1 \leftarrow \frac{1}{2} L_1$, ce qui donne

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3/4 & 1/4 & -1/2 \\ 0 & 1 & 0 & -1/4 & 1/4 & 1/2 \\ 0 & 0 & 1 & 1/4 & 3/4 & -1/2 \end{array}\right).$$

On en conclut que A est inversible, d'inverse $A^{-1} = \begin{pmatrix} 3/4 & 1/4 & -1/2 \\ -1/4 & 1/4 & 1/2 \\ 1/4 & 3/4 & -1/2 \end{pmatrix}$.

Le lien avec la résolution des n systèmes linéaires est clair si l'on remarque que les colonnes de la matrice identité sont les B_j. Les substitutions de lignes ci-dessus sont exactement les opérations nécessaires à la résolution par la méthode du pivot d'un système $A \times X = B$ pour un second membre quelconque. Effectuer ces opérations sur chacune des colonnes de la matrice identité, c'est en fait résoudre simultanément les n systèmes $A \times X = B_j$.

3

Applications linéaires

1. Généralités

1.1. Définition.

Définition 53 – application linéaire

Soient $(E, +, .)$ et $(F, +, .)$ deux espaces vectoriels. On dit qu'une application $f\colon E \to F$ est *linéaire* lorsque pour tous a, b de E et tout λ dans $\mathbb{R}$,

$$f(a+b) = f(a) + f(b) \quad \text{et} \quad f(\lambda.a) = \lambda.f(a).$$

On note $\mathscr{L}(E, F)$ l'ensemble des applications linéaires de E dans F.

Les applications linéaires sont aussi appelées «morphismes d'espaces vectoriels». Comme pour les sous-espaces vectoriels, on peut vérifier les deux propriétés de la définition en une fois : une application $f\colon E \to F$ est linéaire si et seulement si pour tous réels λ, μ et tous x, y dans E, on a

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

Exercice de cours 47 (l'image du vecteur nul par une application linéaire est toujours le vecteur nul). Montrer que l'image du vecteur nul de E par une application linéaire $f : E \to F$ est toujours 0_F.

⚠ Dans la définition, on a noté «+» et «.» les lois de compositions sur E et sur F. Il faut prendre garde au fait que les ensembles E et F contiennent des objets a priori différents, et que les opérations sur ces objets différents n'ont a priori pas de rapport. Il est utile de se demander de quelle loi il s'agit quand on lit la définition ci-dessus.

1.2. Exemples.

1. On prend $E = F = \mathbb{R}$ munis de l'addition et de la multiplication externe (qui est aussi la multiplication interne dans ce cas) usuelles. Si $f : \mathbb{R} \to \mathbb{R}$ est une application linéaire, alors pour tout $x \in \mathbb{R}$, on a $f(x) = f(x.1) = x.f(1)$. Notant $a = f(1)$, on doit donc avoir $f : x \mapsto ax$. Réciproquement, on vérifie facilement qu'une telle application est bien linéaire.
2. Maintenant $E = F$ est l'espace vectoriel des vecteurs du plan. L'application $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(\vec{u}) = \lambda.\vec{u}$ est linéaire. Il s'agit de l'homothétie vectorielle de rapport λ. Si $\vec{v}$ est un vecteur fixé non nul, la translation de vecteur $\vec{v}$, définie par $f(\vec{u}) = \vec{u} + \vec{v}$ n'est pas linéaire : par exemple $f(\vec{u}_1 + \vec{u}_2) = \vec{u}_1 + \vec{u}_2 + \vec{v}$ alors que $f(\vec{u}_1) + f(\vec{u}_2) = \vec{u}_1 + \vec{u}_2 + 2\vec{v}$. On aurait aussi pu noter que $f(\vec{0}) = \vec{v} \neq \vec{0}$.
3. Soit $A \in \mathscr{M}_{n,p}(\mathbb{R})$ une matrice. L'application $f : \mathbb{R}^p \to \mathbb{R}^n$ définie par $f(X) = A \times X$ est une application linéaire. En effet d'après la proposition 43 et la remarque qui la précède, on a

   $$f(\lambda_1 X_1 + \lambda_2 X_2) = A \times (\lambda_1 X_1 + \lambda_2 X_2) = A \times (\lambda_1 X_1) + A \times (\lambda_2 X_2) = \lambda_1 f(X_1) + \lambda_2 f(X_2).$$

   Dans le paragraphe 4, on verra qu'une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$ est toujours de cette forme.

4. Voici un exemple d'application linéaire qui sort du cadre des espaces vectoriels de dimension finie. On note $\mathscr{C}^0([0,1])$ l'espace vectoriel des fonctions continues de $[0,1]$ dans $\mathbb{R}$. On sait qu'une telle fonction est intégrable sur $[0,1]$ et que pour tout réels λ, μ

$$\int_0^1 (\lambda f(x) + \mu g(x))dx = \lambda \int_0^1 f(x)dx + \mu \int_0^1 g(x)dx.$$

Autrement dit, l'application $\varphi : \mathscr{C}^0([0,1]) \to \mathbb{R}$ définie par

$$\varphi(f) = \int_0^1 f(x)dx$$

est une application linéaire.

1.3. Somme et composition d'applications linéaires.

Proposition 54 – l'ensemble des applications linéaires est un espace vectoriel

L'ensemble $\mathscr{L}(E,F)$ des applications linéaires de E dans F, muni de l'addition des fonctions et de la multiplication d'une fonction par un nombre, est un espace vectoriel.

Exercice de cours 48. Démontrer la proposition.

On peut aussi composer des applications linéaires :

Proposition 55 – composition des applications linéaires

Soient E, F, G trois espaces vectoriels, et $f\colon E \to F$, $g\colon F \to G$ deux applications linéaires. Alors $g \circ f\colon E \to G$ est une application linéaire.

Exercice de cours 49. Démontrer la proposition.

⚠ Attention à l'ordre dans la composition !

2. Noyau et Image

2.1. Injection, surjection, bijection. Les notions que l'on rappelle maintenant ne sont pas liées à la linéarité.

Définition 56 – application injective

On dit qu'une application $f\colon E \to F$ est *injective* lorsque pour tous $x, y \in E$,

$$f(x) = f(y) \quad \Rightarrow \quad x = y.$$

Exemple 23. L'application $f\colon \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$ n'est pas injective, puisque $f(-1) = f(1)$. En revanche l'application $f : \mathbb{R}^+ \to \mathbb{R}$ définie par $f(x) = x^2$ est injective.

Définition 57 – application surjective

On dit qu'une application $f\colon E \to F$ est *surjective* lorsque pour tout élément y de F, il existe au moins un élément x de E tel que $f(x) = y$.

Exemple 24. L'application $f\colon \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$ n'est pas surjective, puisqu'il n'existe pas de réel x tel que $f(x) = -1$. En revanche l'application $f\colon \mathbb{R} \to \mathbb{R}^+$ définie par $f(x) = x^2$ est surjective.

Définition 58 – application bijective

On dit qu'une application $f\colon E \to F$ est *bijective* lorsque pour tout élément y de F, il existe exactement un élément x de E tel que $f(x) = y$.

EXEMPLE 25. L'application $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ définie par $f(x) = x^2$ est bijective, puisque pour tout $y \in \mathbb{R}^+$, il existe un unique x tel que $f(x) = y$: il s'agit de $x = \sqrt{y}$.

Proposition 59 – une application est bijective si et seulement si elle est injective et surjective

Une application $f\colon E \to F$ est bijective si et seulement si elle est injective et surjective.

EXERCICE DE COURS 50. Démontrer la proposition.

Définition 60 – bijection réciproque

Soit $f\colon E \to F$ une application bijective. L'application $f^{-1}\colon F \to E$ qui à y dans F associe l'unique $x \in E$ tel que $f(x) = y$ est appelée *bijection réciproque* de f.

On est allé un peu vite : il faut justifier le fait que f^{-1} est bien une bijection :

Proposition 61 – la bijection réciproque est bien une bijection!

Soient $f\colon E \to F$ une application bijective, et $f^{-1}\colon F \to E$ sa bijection réciproque.

1. f^{-1} est bijective, et sa bijection réciproque est f.

2. On a $f^{-1} \circ f = I_E$ et $f \circ f^{-1} = I_F$.

EXERCICE DE COURS 51. Démontrer la proposition.

Dans le cas des applications linéaires, on a un peu de vocabulaire spécifique :

Définition 62 – endomorphisme, isomorphisme, automorphisme

Les applications linéaires de E dans E sont appelées *endomorphismes*. Les applications linéaires bijectives de E dans F sont appelées *isomorphismes*. Les endomorphismes bijectifs sont appelés *automorphismes*.

Plutôt que $\mathscr{L}(E, E)$, on note simplement $\mathscr{L}(E)$ l'espace vectoriel des endomorphismes de E.

2.2. Noyau d'une application linéaire.

Définition 63 – noyau d'une application linéaire

Soit $f\colon E \to F$ une application linéaire. Le sous-ensemble de E des éléments dont l'image est 0_F est appelé noyau de f. On le note

$$\operatorname{Ker} f = \{x \in E\colon f(x) = 0_F\}.$$

Proposition 64 – le noyau d'une application linéaire est un sous-espace vectoriel

Soit $f\colon E \to F$ une application linéaire. Le noyau de f est un sous-espace vectoriel de E.

Exercice de cours 52. Démontrer la proposition.

Proposition 65 – une application linéaire est injective si et seulement si son noyau est nul

Une application linéaire $f\colon E \to F$ est injective si et seulement si son noyau ne contient que le vecteur nul de E, c'est-à-dire $\operatorname{Ker} f = \{0_E\}$.

Exercice de cours 53. Démontrer la proposition.

Exercice de cours 54. Soit $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ l'application linéaire définie par $f(x,y,z) = (y+z, x+z, x-y)$. Déterminer son noyau. L'application f est-elle injective ?

2.3. Image d'une application linéaire.

Définition 66 – image d'une application linéaire

Le sous-ensemble de F constitué des images par f des éléments de E est appelé *image* de f. On le note

$$\operatorname{Im} f = \{f(x) \colon x \in E\}.$$

Autrement dit $\operatorname{Im} f$ est pour une application linéaire ce que l'on note $f(E)$ pour une application quelconque. Par définition donc, une application linéaire $f\colon E \to F$ est surjective si et seulement si $\operatorname{Im} f = F$.

Proposition 67 – l'image d'une application linéaire est un sous-espace vectoriel

Soit $f\colon E \to F$ une application linéaire. L'image de f est un sous-espace vectoriel de F.

Exercice de cours 55. Démontrer la proposition.

Exercice de cours 56. Soit $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ l'application linéaire définie par $f(x,y,z) = (y+z, x+z, x-y)$. Déterminer son image. L'application f est-elle surjective ?

2.4. Applications linéaires bijectives. Voici une particularité des applications bijectives qui sont linéaires :

Proposition 68 – la bijection réciproque d'une application linéaire bijective est linéaire

Soit $f\colon E \to F$ une application linéaire. Si f est bijective, alors son application réciproque $f^{-1}\colon F \to E$ est aussi linéaire.

Exercice de cours 57. Démontrer la proposition.

La proposition suivante, un peu technique, donne une caractérisation de l'injectivité et de la surjectivité d'une application linéaire f en termes de l'image d'une base par f. On retiendra en particulier qu'il ne peut pas y avoir d'application linéaire bijective entre deux espaces de dimensions finies différentes, et qu'une application linéaire est bijective si et seulement si elle transforme toute base de l'espace de départ en une base de l'espace d'arrivée :

Proposition 69 – caractérisation de l'injectivité et de la surjectivité d'une application linéaire

Soient E un espace vectoriel de dimension finie, et F un espace vectoriel quelconque. Soient $\mathscr{B}_E$ une base de E et $f \in \mathscr{L}(E,F)$ une application linéaire.

1. f est injective si et seulement si $f(\mathscr{B}_E)$ est une famille libre de F ; en particulier, si f est injective on a $\dim E \leqslant \dim F$.
2. f est surjective si et seulement si $f(\mathscr{B}_E)$ est une famille génératrice de F ; en particulier, si f est surjective on a $\dim F \leqslant \dim E$.
3. f est bijective si et seulement si $f(\mathscr{B}_E)$ est une base de F ; en particulier, si f est bijective on a $\dim E = \dim F$.

Exercice de cours 58. Démontrer la proposition.

2.5. Le théorème du rang.

Définition 70 – rang d'une application linéaire

Soit $f \in \mathscr{L}(E,F)$. On appelle *rang de f* la dimension de $\operatorname{Im} f$, et on note

$$\operatorname{rang}(f) = \dim \operatorname{Im} f.$$

Théorème 71 – théorème du rang

Soient E un espace vectoriel de dimension finie, et $f \in \mathscr{L}(E,F)$. On a

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

Exercice de cours 59. Le but de cet exercice est de démontrer le théorème 71.
On utilise le théorème de la base incomplète (voir le théorème 21). Soit $\mathcal{K} = (u_1, \dots, u_k)$ une base de $\operatorname{Ker} f$. Comme E est de dimension finie – appelons la n –, on peut compléter $\mathcal{K}$ en une famille $\mathscr{B} = (u_1, \dots, u_k, u_{k+1}, \dots, u_n)$ qui est une base de E.

1. En utilisant la définition de $\operatorname{Im} f$, montrer que $(f(u_{k+1}), \dots, f(u_n))$ est une famille génératrice de $\operatorname{Im} f$.
2. Montrer que $(f(u_{k+1}), \dots, f(u_n))$ est une famille libre de $\operatorname{Im} f$.
3. En déduire que $\dim \operatorname{Im} f = n - k$. Conclure.

⚠ Le théorème du rang relie la dimension de E à celles de $\operatorname{Ker} f$ et $\operatorname{Im} f$ mais $\operatorname{Im} f$ n'est pas un sous-espace vectoriel de E en général !

Même lorsque $E = F$, le théorème du rang ne dit pas que les sous-espaces vectoriels $\operatorname{Ker} f$ et $\operatorname{Im} f$ sont supplémentaires dans E. Il se peut que le noyau et l'image d'un endomorphisme ne soient pas en somme directe.

Par exemple que si f est l'endomorphisme de dérivation (voir aussi l'exercice 62)

$$\begin{array}{rccc} f\colon & \mathbb{R}_n[X] & \longrightarrow & \mathbb{R}_n[X] \\ & P & \longmapsto & P' \end{array}$$

alors on a $\operatorname{Im} f \cap \operatorname{Ker} f = \operatorname{Vect}(1) = \{\text{polynômes constants}\} \neq \{0_{\mathbb{R}_n[X]}\}$.

Voici une des conséquences très utiles du théorème du rang :

Corollaire 72 – si $\dim E = \dim F$, l'injectivité équivaut à la surjectivité

Soit $f \in \mathscr{L}(E,F)$ avec E et F dimension finie. Si $\boxed{\dim E = \dim F}$, on a les équivalences

$$f \text{ bijective } \iff f \text{ surjective } \iff f \text{ injective.}$$

Voici une autre conséquence du théorème du rang :

Corollaire 73

Soit $f \in \mathscr{L}(E,F)$ avec E et F de dimension finie. Si $\dim E > \dim F$, f ne peut pas être injective. De même si $\dim E < \dim F$, f ne peut pas être surjective.

Exercice de cours 60. Démontrer les deux corollaires.

3. Dérivation et arithmétique dans $\mathbb{R}[X]$

L'espace vectoriel des polynômes est très important aussi bien en analyse qu'en algèbre. Nous approfondissons dans cette section son étude et donnons des exemples intéressants d'applications linéaires dans ce contexte.

3.1. Dérivation.

Définition 74 – polynôme dérivé

Pour tout $P = \sum_{i=0}^{n} a_n X^n$ de $\mathbb{R}[X]$, on appelle *polynôme dérivé*, et on note P', le polynôme défini par :

$$P' = \sum_{k=1}^{n} k a_k X^{k-1} = \sum_{k=0}^{n-1} (k+1) a_k X^k.$$

On note $P^{(0)} = P$, $P^{(1)} = P'$, $P^{(2)} = P'' = (P')'$ et, pour tout $k \in \mathbb{N}^*$, $P^{(k)} = (P^{(k-1)})'$.

Cette définition est très naturelle : le polynôme dérivée de P n'est rien d'autre que la fonction dérivée de la fonction polynomiale associée à P (qui est dérivable sur tout $\mathbb{R}$).

Exercice de cours 61. Soit $P \in \mathbb{R}[X]$.

1. Montrer que $\deg(P') = \deg(P) - 1$ si $\deg(P) \geqslant 1$ et $\deg(P') = -\infty$ si $\deg(P) \leqslant 0$, c'est-à-dire si P est un polynôme constant.
2. Montrer que l'application $\mathbb{R}[X] \to \mathbb{R}[X]$ qui à P associe sont polynôme dérivée est un endomorphisme de $\mathbb{R}[X] \to \mathbb{R}[X]$. Quel est son noyau ? Son image ? Mêmes questions pour sa restriction à $\mathbb{R}_n[X]$ où $n \in \mathbb{N}$.

Exercice de cours 62. Soient $n \in \mathbb{N}$ et $f\colon \mathbb{R}_n[X] \to \mathbb{R}_n[X]$ définie par $f(P) = \lambda P + P'$ avec $\lambda \neq 0$. Montrer que f est un isomorphisme sur $\mathbb{R}_n[X]$.

***Gottfried Wilhelm Leibniz**, né à Leipzig le 1er juillet 1646 et mort à Hanovre le 14 novembre 1716, est un philosophe, scientifique, mathématicien, logicien, diplomate, juriste, bibliothécaire et philologue allemand. Esprit polymathe, personnalité importante de la période Frühaufklärung, il occupe une place primordiale dans l'histoire de la philosophie et l'histoire des sciences (notamment des mathématiques) et est souvent considéré comme le dernier «génie universel».*

Proposition 75 – formule de Leibniz

Pour tous $P, Q \in \mathbb{R}[X]$, et tout $k \in \mathbb{N}$, on a
$$(PQ)^{(k)} = \sum_{i=0}^{k} \binom{k}{i} P^{(i)} Q^{(k-i)},$$
où $\binom{k}{i}$ est le coefficient binomial i parmi k.

Exercice de cours 63. Démontrer la proposition par récurrence sur k.

Le théorème suivant est crucial, et pas seulement pour ce cours !

Théorème 76 – théorème de Taylor pour les polynômes

Soient $n \in \mathbb{N}$ et $P \in \mathbb{R}[X]$ un polynôme de degré au plus n. Alors
$$P = \sum_{k=0}^{n} \frac{P^{(k)}(0)}{k!} X^k.$$

Autrement dit, les coordonnées du polynôme $P \in \mathbb{R}_n[X]$ dans la base canonique $1, X, X^2, \dots, X^n$ de $\mathbb{R}_n[X]$ sont $P(0), P'(0), P''(0)/2, \dots, P^{(n)}(0)/(n!)$.

***Brook Taylor**, est un homme de science anglais, né à Edmonton, aujourd'hui un quartier de Londres, le 18 août 1685, et mort à Londres le 29 décembre 1731. Principalement connu comme mathématicien, il s'intéressa aussi à la musique, à la peinture et à la religion.*

Exercice de cours 64.

1. Démontrer le théorème par récurrence sur n.
2. Soient $n \in \mathbb{N}$ et $a \in \mathbb{R}$.
   - **2.1** Dans les conditions du théorème, montrer que l'on a :
   $$P = \sum_{k=0}^{n} \frac{P^{(k)}(a)}{k!} (X-a)^k.$$
   C'est le théorème de Taylor pour les polynômes sous sa forme la plus générale.
   - **2.2** Montrer que $(1, X-a, (X-a)^2, \dots, (X-a)^n)$ forment une base de $\mathbb{R}_n[X]$. Quelles sont les coordonnées d'un polynôme $P \in \mathbb{R}_n[X]$ dans cette base ?

3.2. Arithmétique dans $\mathbb{R}[X]$ et racines d'un polynôme. Comme dans $\mathbb{Z}$, on peut dans $\mathbb{R}[X]$ additionner et multiplier des éléments entre eux, on a un *zéro* pour l'addition (le polynôme nul), et une *unité* (le polynôme constant égal à 1).

⚠ L'ensemble $\mathbb{Z}$ n'est pas un espace vectoriel, contrairement à $\mathbb{R}[X]$!

L'analogie entre $\mathbb{Z}$ et $\mathbb{R}[X]$ permet de faire de l'arithmétique dans $\mathbb{R}[X]$. Dans ce cours, on se contentera de quelques définitions.

Définition 77 – division entre polynômes

Soient $A, P \in \mathbb{R}[X]$. On dit que A *divise* P, et on note $A|P$, s'il existe $Q \in \mathbb{R}[X]$ tel que $P = AQ$.

Par exemple, $2X$ divise $3X^2$ car $3X^2 = (2X) \times (3/2X)$, $(X-2)$ divise clairement $(X-2)^2$ et $(X-2)(X-3)$, X ne divise pas $X^3 - 1$ (pourquoi ?). Tout polynôme divise le polynôme nul !

Exercice de cours 65. Vérifier que $X-2$ et $X-1$ divisent $X^2 - 3X + 2$. Quels sont tous les polynômes qui divisent $X^2 - 3X + 2$?

L'analogie entre $\mathbb{R}[X]$ et $\mathbb{Z}$ réside surtout dans l'énoncé suivant que l'on admet dans ce cours.

Théorème 78 – division euclidienne

Soient $A, B \in \mathbb{R}[X]$ avec $B \neq 0$. Il existe un unique couple $(Q, R) \in (\mathbb{R}[X])^2$ tel que :

$$\begin{cases} A = BQ + R \\ \deg(R) < \deg(B). \end{cases}$$

Le polynôme Q (respectivement R) s'appelle le *quotient* (respectivement *le reste*) de la *division euclidienne de A par B*.

Exercice de cours 66 (division euclidienne en pratique).

1. Effectuer la division euclidienne de $A = X^4 + 2X^3 - X + 6$ par $B = X^3 - 6X^2 + X + 4$.

2. Trouver les $a \in \mathbb{R}$ tels que $(X^2 - aX + 1)|X^4 - X + a$.

Soient $P \in \mathbb{R}[X]$ et $a \in \mathbb{R}$. Lorsque $X - a$ divise P, on dit que a est une *racine* de P. La proposition suivante assure que a est une racine de P si seulement si a est un zéro de la fonction polynomiale associée à P, c'est-à-dire que $P(a) = 0$.

Proposition 79 – les racines sont les zéros d'un polynôme

Soient $P \in \mathbb{R}[X]$ et $a \in \mathbb{R}$. On a l'équivalence : $(X - a)|P \iff P(a) = 0$.

Exercice de cours 67.

1. Démontrer l'implication «$\Rightarrow$» de la proposition.

2. Démontrer l'implication «$\Leftarrow$» de la proposition à l'aide de la division euclidienne.

3. Donner une autre démonstration de l'implication «$\Leftarrow$» à l'aide du théorème de Taylor généralisé pour les polynômes (voir l'exercice 64).

Exercice de cours 68. Soit $a \in \mathbb{R}$. Montrer que l'application de $\mathbb{R}[X]$ dans $\mathbb{R}[X]$ qui à P associe $P(a)$ est une application linéaire. Quel est son noyau ? Quelle est son image ?

La proposition 79 a de nombreuses conséquences intéressantes. Comme elles sont à la limite du programme de ce cours, nous les proposons sous forme d'exercice.

Exercice de cours 69 (racines/zéros d'un polynôme et conséquences). Soient $P \in \mathbb{R}[X]$ et $n \in \mathbb{N}^*$.

1. Montrer que si $x_1, \ldots, x_n$ sont des zéros deux à deux distincts de P, alors le produit $(X - x_1) \ldots (X - x_n)$ divise P.

2. Déduire de la question précédente que si $\deg P < n$ et si P admet au moins n zéros distincts, alors $P = 0$.

3. Montrer que si P s'annule en une infinité d'éléments de $\mathbb{R}$ alors $P = 0$.

Une jolie application de l'exercice précédent est l'étude *des polynômes d'interpolation de Lagrange* (voir la fiche d'exercice n°1).

Définition 80 – multiplicité d'un zéro

Soient $P \in \mathbb{R}[X]$, $a \in \mathbb{R}$ et $\alpha \in \mathbb{N}^*$. On dit que a est un *zéro d'ordre au moins α de P* si $(X-a)^\alpha | P$. On dit que a est un *zéro d'ordre exactement α de P* si $(X-a)^\alpha | P$ et $(X-a)^{\alpha+1} \nmid P$.

Exercice de cours 70. À l'aide du théorème de Taylor généralisé pour les polynômes (voir l'exercice 64), interpréter la multiplicité d'un zéro en termes des dérivées successives de P évaluées en a : $P(a), P'(a), \dots, P^{(\alpha-1)}(a), P^{(\alpha)}(a)$.

4. Matrices d'une application linéaire

4.1. Définition. Dans ce paragraphe, on considère deux espaces vectoriels E et F de dimension finie, respectivement p et n. On choisit une base $\mathscr{B}_E = (u_1, u_2, \dots, u_p)$ de E, et une base $\mathscr{B}_F = (v_1, v_2, \dots, v_n)$ de F.

Proposition 81 – la donnée de l'image des vecteurs d'une base détermine l'application linéaire

Soit $f \in \mathscr{L}(E, F)$ une application linéaire. La donnée de l'image par f des vecteurs de la base $\mathscr{B}_E$ détermine entièrement l'application f.

Exercice de cours 71. Démontrer la proposition.

D'autre part, les vecteurs $f(u_1), \dots, f(u_p)$ sont eux parfaitement déterminés par la donnée de leurs coordonnées dans la base $\mathscr{B}_F$. Autrement dit, la donnée du tableau à n lignes et p colonnes constituées des coordonnées des $f(u_j)$ dans la base $(v_1, v_2, \dots, v_n)$ de F détermine entièrement l'application linéaire f. Cela conduit à la définition suivante :

Définition 82

Soit $f \in \mathscr{L}(E, F)$ une application linéaire. On appelle *matrice de f dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$* la matrice $A \in \mathscr{M}_{n,p}(\mathbb{R})$ dont les p colonnes sont les coordonnées des images par f des vecteurs de $\mathscr{B}_E$ dans la base $\mathscr{B}_F$. On la note

$$A = \mathrm{Mat}_{\mathscr{B}_F \leftarrow \mathscr{B}_E}(f) = \begin{array}{cc} \begin{array}{cccc} f(u_1) & f(u_2) & \dots & f(u_p) \end{array} & \\ \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix} & \begin{array}{c} v_1 \\ v_2 \\ \vdots \\ v_n \end{array} \end{array}$$

avec, pour $j \in \{1, \dots, p\}$,

$$f(u_j) = a_{1,j} v_1 + a_{2,j} v_2 + \dots + a_{n,j} v_n = \sum_{i=1}^{n} a_{i,j} v_i.$$

⚠ La phrase «on considère la matrice de l'application linéaire $f \in \mathscr{L}(E, F)$» n'a pas de sens ; il faut préciser dans quelles bases de E et F on écrit la matrice !

Exemple 26. On considère l'espace vectoriel $\mathbb{R}^n$, et $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ une base quelconque de $\mathbb{R}^n$. La matrice de l'homothétie $h_\lambda \colon \mathbb{R}^n \to \mathbb{R}^n$ de rapport λ dans les bases $\mathscr{B}$ et $\mathscr{B}$ est

$$H_\lambda = \mathrm{Mat}_{\mathscr{B}\leftarrow\mathscr{B}}(h_\lambda) = \begin{array}{c c} \begin{array}{cccc} h_\lambda(e_1) & h_\lambda(e_2) & \ldots & h_\lambda(e_n) \end{array} & \\ \begin{pmatrix} \lambda & 0 & \ldots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda \end{pmatrix} & \begin{array}{c} e_1 \\ e_2 \\ \\ e_n \end{array} \end{array} = \lambda.I.$$

Exemple 27. On note $f_\theta \colon \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle θ dans $\mathbb{R}^2$, définie par

$$f_\theta(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta).$$

La matrice de f_θ dans la base canonique $\mathscr{B}$ de $\mathbb{R}^2$ est donc

$$A = \mathrm{Mat}_{\mathscr{B}\leftarrow\mathscr{B}}(f_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Exemple 28. On note $f_\theta \colon \mathbb{R}^3 \to \mathbb{R}^3$ la rotation d'axe $(0, 0, 1)$ et d'angle θ dans $\mathbb{R}^3$, définie par

$$f_\theta(x, y, z) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta, z).$$

La matrice de f_θ dans la base canonique $\mathscr{B}$ de $\mathbb{R}^3$ est donc

$$A = \mathrm{Mat}_{\mathscr{B}\leftarrow\mathscr{B}}(f_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Exemple 29. On considère l'espace vectoriel $\mathbb{R}_n[X]$ des polynômes de degré inférieur ou égal à n. La famille $\mathscr{B} = (1, X, X^2, \ldots, X^n)$ est une base de $\mathbb{R}_n[X]$. Soit alors $\varphi \colon \mathbb{R}_n[X] \to \mathbb{R}_n[X]$ l'application qui a un polynôme $P \in \mathbb{R}_n[X]$ associe son polynôme dérivé P'. L'application φ est linéaire (voir l'exercice 61), et $\varphi(1) = 0$, $\varphi(X) = 1$, …, $\varphi(X^n) = nX^{n-1}$. La matrice de φ dans les bases $\mathscr{B}$ et $\mathscr{B}$ est donc

$$A = \mathrm{Mat}_{\mathscr{B}\leftarrow\mathscr{B}}(\varphi) = \begin{array}{c c} \begin{array}{ccccc} \varphi(1) & \varphi(X) & \varphi(X^2) & \ldots & \varphi(X^n) \end{array} & \\ \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & n \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix} & \begin{array}{c} 1 \\ X \\ \vdots \\ X^{n-1} \\ X^n \end{array} \end{array}$$

Exercice de cours 72. Quelle est la matrice dans les bases $\mathscr{B}$ et $\mathscr{B}$ de l'application linéaire f de l'exercice 62 ?

Remarque 11. La matrice d'une application linéaire dans des bases $\mathscr{B}_E$ et $\mathscr{B}_F$ dépend en général des bases $\mathscr{B}_E$ et $\mathscr{B}_F$. Comme on l'a vu plus haut, ce n'est cependant pas le cas pour les homothéties, et donc en particulier pour l'application identité (qu'on peut écrire h_λ avec $\lambda = 1$).

Remarque 12. On a parlé précédemment de l'application linéaire associée à une matrice A donnée de $\mathscr{M}_{n,p}(\mathbb{R})$. Il est maintenant clair qu'il s'agit de l'application linéaire f de $\mathbb{R}^p$ dans $\mathbb{R}^n$ dont la matrice dans les bases canoniques de $\mathbb{R}^p$ et $\mathbb{R}^n$ est A.

4.2. Liens entre une application linéaire et sa matrice dans des bases données. Soyons plus précis : on a vu que «la donnée du tableau à n lignes et p colonnes constituées des coordonnées des $f(u_j)$ dans la base $(v_1, v_2, \ldots, v_n)$ de F détermine entièrement l'application linéaire f», mais comment ?

Proposition 83 – écriture matricielle de l'image d'un vecteur par une application linéaire

Soit $f \in \mathscr{L}(E,F)$ une application linéaire. La matrice $A \in \mathscr{M}_{n,p}(\mathbb{R})$ est la matrice de f dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$, si et seulement si pour tout $x \in E$, de coordonnées $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ dans la base $\mathscr{B}_E$, les coordonnées $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ de son image $y = f(x)$ dans la base $\mathscr{B}_F$ vérifient $Y = A \times X$.

Exercice de cours 73. Démontrer la proposition.

Proposition 84 – lien entre produit matriciel et composition des applications linéaires

Soient E, F et G des espaces vectoriels de dimension finie, et deux applications linéaires $f\colon E \to F$ et $g\colon F \to G$. Soient aussi $\mathscr{B}_E$, $\mathscr{B}_F$ et $\mathscr{B}_G$ des bases respectives de E, F et G, et

$$A = \mathrm{Mat}_{\mathscr{B}_F \leftarrow \mathscr{B}_E}(f), \quad B = \mathrm{Mat}_{\mathscr{B}_G \leftarrow \mathscr{B}_F}(g).$$

Alors la matrice C de l'application linéaire $g \circ f\colon E \to G$ dans les bases $\mathscr{B}_E$ et $\mathscr{B}_G$ est donnée par le produit $B \times A$:

$$C = \mathrm{Mat}_{\mathscr{B}_G \leftarrow \mathscr{B}_E}(g \circ f) = \mathrm{Mat}_{\mathscr{B}_G \leftarrow \mathscr{B}_F}(g) \times \mathrm{Mat}_{\mathscr{B}_F \leftarrow \mathscr{B}_E}(f) = B \times A.$$

Exercice de cours 74. Démontrer la proposition.

Exercice de cours 75. Soient $f_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ et $f_{\theta'}\colon \mathbb{R}^2 \to \mathbb{R}^2$ les rotations d'angle θ et θ' respectivement. Quelle est la matrice de $f_\theta \circ f'_\theta$ dans la base canonique ? Pouvait-on prédire géométriquement ce résultat ?

4.3. Rang d'une application linéaire et rang de sa matrice.

Proposition 85 – rang d'une application linéaire en terme du rang d'un système linéaire

Soient $f\colon E \to F$ une application linéaire, et $A = \mathrm{Mat}_{\mathscr{B}_F \leftarrow \mathscr{B}_E}(f)$ sa matrice dans $\mathscr{B}_E$ et $\mathscr{B}_F$. Le rang de f est le rang du système linéaire associé à la matrice A.

Exercice de cours 76. À l'aide des résultats du paragraphe 6.3, démontrer la proposition.

On peut même extraire de cette preuve un procédé pour calculer le rang de f : c'est la dimension de l'espace vectoriel engendré par les vecteurs colonnes $(f(u_1), \dots f(u_p))$ de A. Cela justifie la définition suivante.

Définition 86 – rang d'une matrice

Soit $A \in \mathscr{M}_{n,p}(\mathbb{R})$ une matrice. On note $C_1, C_2, \dots, C_p$ les vecteurs colonnes de A. On appelle *rang* de A et on note $\mathrm{rang}(A)$ la dimension du sous-espace vectoriel $\mathrm{Vect}(C_1, \dots, C_p)$ de $\mathbb{R}^n$.

Remarque 13. Si $u_1, \dots, u_p$ sont des vecteurs d'un espace vectoriel E de dimension finie, on appelle *rang* de la famille $(u_1, \dots, u_p)$ la dimension du sous-espace vectoriel $\mathrm{Vect}(u_1, \dots, u_p)$.

Exemple 30. On veut déterminer le rang de l'application linéaire f dont la matrice dans la base canonique de $\mathbb{R}^3$ est $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 3 & 2 & -1 \end{pmatrix}$. Il s'agit de déterminer la dimension de $\mathrm{Vect}(C_1, C_2, C_3)$, où $C_1 = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$, $C_2 = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$ et $C_3 = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$. Il s'agit de compter les inconnues principales du système homogène correspondant à l'équation $\lambda_1 C_1 + \lambda_2 C_2 + \lambda_3 C_3 = 0$, dont la matrice associée est A. En échelonnant ce système, on trouve facilement $\mathrm{rang}(f) = \mathrm{rang}(A) = 2$.

Remarque 14. On vient de prouver un point important. Le rang d'un système linéaire a été défini comme le nombre de lignes non nulles du système échelonné obtenu en appliquant l'algorithme du pivot au sytème en question. Obtiendrait-on le même rang si l'on utilisait un autre procédé pour trouver un système échelonné équivalent au premier ? Pour les systèmes compatibles, en particulier les systèmes homogènes, la réponse est oui : ce que l'on calcule, c'est la dimension de l'image de l'application linéaire f associée au système, et ce nombre ne dépend pas de la méthode utilisée pour le calculer.

4.4. Matrice d'une bijection.

Proposition 87 – caractérisation matricielle de la bijectivité d'une application linéaire

Soient E et F de dimension finie, respectivement p et n. Soient $f \in \mathscr{L}(E, F)$ une application linéaire, et $A \in \mathscr{M}_{n,p}(\mathbb{R})$ la matrice de f dans les bases $\mathscr{B}_E$ de E et $\mathscr{B}_F$ de F. L'application f est bijective si et seulement si la matrice A est une matrice carrée, donc $n = p$, qui est inversible. Dans ce cas

$$\mathrm{Mat}_{\mathscr{B}_E \leftarrow \mathscr{B}_F}(f^{-1}) = A^{-1}.$$

Exercice de cours 77. Le but de l'exercice est de démonter la proposition.

1. Supposons f bijective.

1.1 À l'aide de la proposition 69, montrer : $\dim E = \dim F$.

On en déduit que $A = \mathrm{Mat}_{\mathscr{B}_F \leftarrow \mathscr{B}_E}(f)$ est une matrice carrée de $\mathscr{M}_n(\mathbb{R})$. Notons $B = \mathrm{Mat}_{\mathscr{B}_E \leftarrow \mathscr{B}_F}(f^{-1})$.

1.2 Montrer que A est inversible et que $A^{-1} = B = \mathrm{Mat}_{\mathscr{B}_E \leftarrow \mathscr{B}_F}(f^{-1})$.

2. Supposons la matrice carrée $A = \mathrm{Mat}_{\mathscr{B}_F \leftarrow \mathscr{B}_E}(f)$ inversible. On note $g \colon F \to E$ l'application linéaire dont la matrice dans les bases $\mathscr{B}_F$ et $\mathscr{B}_E$ est A^{-1}. Montrer que $(g \circ f)(x) = x$ pour tout x, ou encore que $g \circ f = I_E$. Montrer de la même manière que $f \circ g = I_F$, et en déduire que f est bijective.

Exercice de cours 78. Soient $a \in \mathbb{R}$ et $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ l'application linéaire définie par $f(x, y) = (x + 2y, 2x + ay)$. L'application f est-elle bijective ? Si oui donner f^{-1}.

5. Changement de base

La matrice d'une application linéaire f donnée dépend en général des bases dans lesquelles on la définit. On s'intéresse dans ce paragraphe aux relations entre les matrices de f dans des bases différentes.

5.1. Matrice de passage d'une base dans une autre. Soit E un espace vectoriel de dimension n. Soient $\mathscr{B} = (u_1, u_2, \ldots, u_n)$ une base de E, et $\mathscr{B}' = (u'_1, u'_2, \ldots, u'_n)$ une «nouvelle base» de E. En général, les vecteurs de $\mathscr{B}' = (u'_1, u'_2, \ldots, u'_n)$ sont donnés par leurs coordonnées dans «l'ancienne base» $\mathscr{B}$, et il est naturel et très simple de former la matrice P des coordonnées des vecteurs de la nouvelle base $\mathscr{B}'$ dans l'ancienne base $\mathscr{B}$:

$$P = \begin{array}{c} \\ \\ \\ \\ \\ \end{array} \begin{array}{cccc} u'_1 & u'_2 & \dots & u'_n \\ \end{array}$$

$$P = \left(\begin{array}{cccc} p_{1,1} & p_{1,2} & & p_{1,n} \\ p_{2,1} & p_{2,2} & & p_{2,n} \\ \vdots & \vdots & & \vdots \\ p_{n,1} & p_{n,2} & & p_{n,n} \end{array} \right) \begin{array}{c} u_1 \\ u_2 \\ \\ u_n \end{array}$$

Cette matrice a un nom, pour l'instant un peu troublant :

Définition 88 – matrice de passage

La matrice P des coordonnées des vecteurs de la nouvelle base $\mathscr{B}'$ dans l'ancienne base $\mathscr{B}$ est appelée *matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$*. On la note $P_{\mathscr{B},\mathscr{B}'}$.

⚠ Attention à l'ordre des bases dans cette définition : $P_{\mathscr{B},\mathscr{B}'}$ est la matrice dont les colonnes sont les coordonnées dans la base $\mathscr{B}$ des vecteurs de la base $\mathscr{B}'$.

Exemple 31. Soient $\mathscr{B} = (e_1, e_2)$ la base canonique de $\mathbb{R}^2$ et $e'_1 = (1, -1)$, $e'_2 = (0, 2)$. Les vecteurs e'_1 et e'_2 ne sont pas colinéaires, donc $\mathscr{B}' = (e'_1, e'_2)$ est une base de $\mathbb{R}^2$. La matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est la matrice

$$P_{\mathscr{B},\mathscr{B}'} = \begin{array}{cc} e'_1 & e'_2 \end{array}$$

$$P_{\mathscr{B},\mathscr{B}'} = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \begin{array}{c} e_1 \\ e_2 \end{array}$$

Proposition 89 – une matrice de passage est une matrice associée à l'application identité

Soient E un espace vectoriel de dimension $n \in \mathbb{N}^*$, $\mathscr{B} = (u_1, u_2, \dots, u_n)$ et $\mathscr{B}' = (u'_1, u'_2, \dots, u'_n)$ deux bases de E. La matrice de de passage de $\mathscr{B}$ dans $\mathscr{B}'$ est la matrice de l'application identité de E dans les bases $\mathscr{B}'$ et $\mathscr{B}$. Autrement dit

$$P_{\mathscr{B},\mathscr{B}'} = \mathrm{Mat}_{\mathscr{B} \leftarrow \mathscr{B}'}(I).$$

Pour démontrer la proposition, il suffit de remarquer que

$$\begin{array}{cccc} u'_1 = I(u'_1) & u'_2 = I(u'_2) & \dots & u'_n = I(u'_n) \end{array}$$

$$P_{\mathscr{B},\mathscr{B}'} = \left(\begin{array}{cccc} p_{1,1} & p_{1,2} & & p_{1,n} \\ p_{2,1} & p_{2,2} & & p_{2,n} \\ \vdots & \vdots & & \vdots \\ p_{n,1} & p_{n,2} & & p_{n,n} \end{array} \right) \begin{array}{c} u_1 \\ u_2 \\ \\ u_n \end{array} = \mathrm{Mat}_{\mathscr{B} \leftarrow \mathscr{B}'}(I).$$

Remarque 15. Une matrice de passage est toujours inversible, puisque l'application linéaire identité l'est (voir la proposition 87) !

De la définition ci-dessus découle l'importante proposition suivante.

Proposition 90 – matrice d'un vecteur dans une nouvelle base grâce à la matrice de passage

Une matrice $P \in \mathscr{M}_n(\mathbb{R})$ est la matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{B}'$ si et seulement si, pour tout vecteur $x \in E$, ses coordonnées $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans la base $\mathscr{B}$ et $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ dans la base $\mathscr{B}'$ vérifient $X = P \times X'$.

Exercice de cours 79. Démontrer la proposition.

Exemple 32. On reprend les vecteurs de l'exemple 31 : $\mathscr{B} = (e_1, e_2)$ la base canonique de $\mathbb{R}^2$ et $\mathscr{B}' = (e_1', e_2')$ avec $e_1' = (1, -1)$, $e_2' = (0, 2)$. Si $x = (x_1, x_2) \in \mathbb{R}^2$, on a $x = x_1 e_1 + x_2 e_2$ et $x = x_1' e_1' + x_2' e_2'$ avec

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}.$$

Exercice de cours 80. Vérifier l'assertion de l'exemple précédent.

Proposition 91 – inverse d'une matrice de passage

Si P est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$, alors P^{-1} est la matrice de passage de $\mathscr{B}'$ à $\mathscr{B}$.

Exercice de cours 81. Démontrer la proposition à l'aide de la proposition 90 et de la remarque 15.

5.2. Matrices semblables.

Proposition 92 – formule de changement de bases pour les applications linéaires

Soient $f \in \mathscr{L}(E, F)$, et A la matrice de f dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$. Soient aussi $\mathscr{B}_E'$ et $\mathscr{B}_F'$ des nouvelles bases de E et F respectivement. Alors la matrice A' de f dans les bases $\mathscr{B}_E'$ et $\mathscr{B}_F'$ est donnée par

$$A' = (P_{\mathscr{B}_F, \mathscr{B}_F'})^{-1} \times A \times P_{\mathscr{B}_E, \mathscr{B}_E'}$$

On notera que cette relation s'écrit

$$\mathrm{Mat}_{\mathscr{B}_F' \leftarrow \mathscr{B}_E'}(f) = (P_{\mathscr{B}_F, \mathscr{B}_F'})^{-1} \times \mathrm{Mat}_{\mathscr{B}_F \leftarrow \mathscr{B}_E}(f) \times P_{\mathscr{B}_E, \mathscr{B}_E'}$$

ou encore

$$\mathrm{Mat}_{\mathscr{B}_F' \leftarrow \mathscr{B}_E'}(f) = \mathrm{Mat}_{\mathscr{B}_F' \leftarrow \mathscr{B}_F}(I) \times \mathrm{Mat}_{\mathscr{B}_F \leftarrow \mathscr{B}_E}(f) \times \mathrm{Mat}_{\mathscr{B}_E \leftarrow \mathscr{B}_E'}(I).$$

Cette formule est donc une conséquence de la proposition 84.

Exercice de cours 82. Démontrer la proposition à l'aide de la proposition 90.

Corollaire 93 – formule de changement de bases pour les endomorphismes

Soient $f \in \mathscr{L}(E)$ un endomorphisme, et A la matrice de f dans la base $\mathscr{B}$. Soit aussi $\mathscr{B}'$ une nouvelle base de E. Alors la matrice A' de f dans la base $\mathscr{B}'$ vérifie

$$A' = P^{-1} \times A \times P,$$

où $P = P_{\mathscr{B}, \mathscr{B}'}$ est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$.

Définition 94 – matrice semblable

On dit que deux matrices A et B sont *semblables* lorsqu'il existe une matrice inversible $P \in \mathscr{M}_n(\mathbb{R})$ telle que $A = P \times B \times P^{-1}$.

De manière équivalente, deux matrices sont semblables si et seulement si elles représentent la même application linéaire dans des bases différentes, puisqu'une matrice inversible peut être vue comme une matrice de passage. On peut se demander s'il existe une base dans laquelle la matrice d'une application linéaire donnée est la plus simple possible. C'est l'objet central de la *réduction des endomorphismes* (voir le cours de deuxième année).

6. Projections et symétries vectorielles

Pour finir, on examine en détail deux types d'applications linéaires qui jouent un rôle important. Comme leur nom l'indique, il est utile pour bien comprendre ce qui suit d'avoir en tête l'exemple des projections et symétries du plan ou de l'espace.

6.1. Projections.

Définition 95 – définition algébrique d'une projection

Soit $(E, +, .)$ un espace vectoriel. On dit qu'un endomorphisme $p \in \mathscr{L}(E)$ est une *projection* lorsque $p \circ p = p$.

Exercice de cours 83.

1. Soit $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'application linéaire dont A est la matrice dans la base canonique de $\mathbb{R}^2$. Montrer que f est une projection.

2. Soit $p \in \mathscr{L}(E, E)$ une projection. Montrer que $I - p$ est aussi une projection.

Proposition 96 – image et noyau d'une projection sont supplémentaires dans E

Si $p \in \mathscr{L}(E)$ est une projection, alors $\operatorname{Ker} p$ et $\operatorname{Im} p$ sont des sous-espaces vectoriels supplémentaires de E :

$$E = \operatorname{Ker} p \oplus \operatorname{Im} p.$$

De plus la restriction de p à $\operatorname{Ker} p$ est l'application nulle, et celle de p à $\operatorname{Im} p$ est l'application identité.

Exercice de cours 84. Démontrer la proposition.

Soit $p \in \mathscr{L}(E)$ une projection. La proposition 96 dit que $E = \operatorname{Ker} p \oplus \operatorname{Im} p$, c'est-à dire que tout vecteur $x \in E$ s'écrit de manière unique $x = u + v$ avec $u \in \operatorname{Ker} p$ et $v \in \operatorname{Im} p$. Conformément à la proposition 31, on dit que u est la *composante* de x sur $\operatorname{Ker} p$ et que v est la *composante* de x sur $\operatorname{Im} p$.

Avec ces notations, on a $p(x) = p(u) + p(v) = p(v)$, et la proposition 96 nous dit que

$$p(x) = p(u) + p(v) = 0 + v = v$$

Finalement, on a donc

$$x = u + v, \quad u \in \operatorname{Ker} p,\ v \in \operatorname{Im} p \quad \Rightarrow \quad p(x) = v.$$

Cette remarque va nous permettre d'interpréter géométriquement la notion de projection.

Proposition 97 – interprétation géométrique d'une projection

Soient E_1 et E_2 deux sous-espaces vectoriels supplémentaires de l'espace vectoriel E. Soit $p : E \to E$ l'application qui à $x \in E$ associe sa composante sur E_2. Alors p est une projection ; son image est E_2 et son noyau est E_1. On dit que p est la *projection sur E_2 dans la direction de E_1* (ou encore : *parallèlement à E_1*).

Exercice de cours 85. Démontrer la proposition.

Dans le cas où E est l'ensemble des vecteurs du plan, et E_1, E_2 deux droites vectorielles distinctes, comme dans le cas où E est l'ensemble des vecteurs de l'espace, E_1 une droite et E_2 un plan ne contenant pas E_1, la projection sur E_2 dans la direction de E_1 est bien ce que l'on pense, comme l'illustre la figure 1.

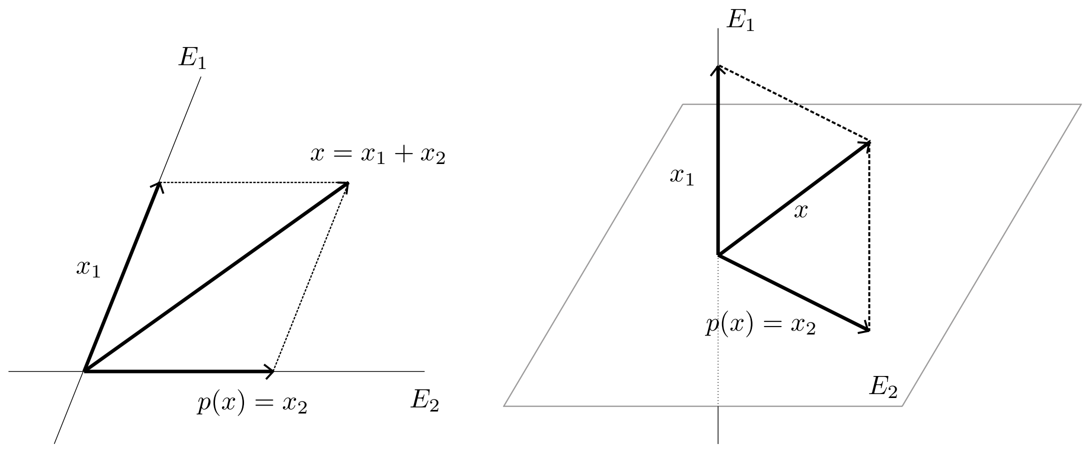


FIGURE 1. Projections vectorielles dans le plan et dans l'espace

Si A est la matrice d'une projection p dans des bases données, on a $A^2 = A \times A = A$, puisque que $A \times A$ est la matrice dans les mêmes bases de $p \circ p = p$. On peut choisir les bases de manière à ce que la matrice de p dans celles-ci soit très simple comme l'illustre l'exercice suivant.

EXERCICE DE COURS 86 (matrice d'une projection). Soient $p\colon E \to E$ une projection, $\mathscr{B}_1 = (u_1, \dots, u_{k_1})$ une base de $\operatorname{Im} p$ et $\mathscr{B}_2 = (v_1, \dots, v_{k_2})$ une base de $\operatorname{Ker} p$. Notant A la matrice de p dans les bases $\mathscr{B}$ et $\mathscr{B}$, montrer que l'on a :

$$A = \mathrm{Mat}_{\mathscr{B}\leftarrow\mathscr{B}}(p) = \begin{array}{c} \begin{array}{ccc|ccc} p(u_1) & \dots & p(u_{k_1}) & p(v_1) & \dots & p(v_{k_2}) \end{array} \\ \left(\begin{array}{ccc|ccc} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right) \end{array} \begin{array}{c} \\ u_1 \\ \\ u_{k_1} \\ v_1 \\ \\ v_{k_2} \end{array}$$

On pourra rapprocher ce résultat de l'exexercice 83 (1).

6.2. Symétries.

Définition 98 – définition algébrique d'une symétrie

Soit $(E, +, .)$ un espace vectoriel. On dit qu'un endomorphisme $s \in \mathscr{L}(E)$ est une *symétrie* lorsque $s \circ s = I$.

On notera qu'une symétrie est une application linéaire bijective, qui est sa propre inverse.

EXERCICE DE COURS 87. Soit $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ et $s\colon \mathbb{R}^2 \to \mathbb{R}^2$ l'application linéaire dont A est la matrice dans la base canonique de $\mathbb{R}^2$. Montrer que s est une symétrie.

EXERCICE DE COURS 88. Soit $(\mathscr{F}(\mathbb{R}, \mathbb{R}), +, .)$ l'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. À une fonction $f \in \mathscr{F}$ on associe la fonction $s(f)$ définie par

$$s(f)\colon x \mapsto f(-x).$$

Montrer que s est une symétrie.

Proposition 99 – lien entre symétries et projections

Soit $p \in \mathscr{L}(E)$. L'application p est une projection si et seulement si $s = I - 2p$ est une symétrie.

Exercice de cours 89. Démontrer la proposition.

Puisque s est une bijection, on a $\operatorname{Ker} s = \{0_E\}$ et $\operatorname{Im} s = E$, donc en particulier on a encore $E = \operatorname{Ker} s \oplus \operatorname{Im} s$, mais cette égalité est sans intérêt pour une symétrie. En revanche, on a aussi la proposition suivante.

Proposition 100 – deux sous-espaces vectoriels supplémentaires associées à une symétrie

Si $s \in \mathscr{L}(E)$ est une symétrie, alors $\operatorname{Ker}(s+I)$ et $\operatorname{Ker}(s-I)$ sont des sous-espaces vectoriels supplémentaires de E :

$$E = \operatorname{Ker}(s+I) \oplus \operatorname{Ker}(s-I).$$

Exercice de cours 90. Démontrer la proposition.

Exemple 33. On reprend l'exercice 88. On a vu que l'application $s\colon \mathscr{F} \to \mathscr{F}$ définie par $s(f)\colon x \to f(-x)$ est une symétrie. Une fonction $f \in \mathscr{F}$ appartient à $\operatorname{Ker}(s+I)$ si et seulement si pour tout $x \in \mathbb{R}$, on a $(s(f)+f)(x) = f(-x) + f(x) = 0$. Autrement dit $\operatorname{Ker}(s+I)$ est le sous-espace des fonction impaires. De la même manière, $f \in \operatorname{Ker}(s-I)$ si et seulement si $(s(f)-f)(x) = f(-x) - f(x) = 0$, donc $\operatorname{Ker}(s-I)$ est le sous-espace des fonctions paires. Dans ce cas, la proposition précédente dit donc que toute fonction de $\mathbb{R}$ dans $\mathbb{R}$ s'écrit de manière unique comme somme d'une fonction paire et d'une fonction impaire.

Soient $s\colon E \to E$ une symétrie, et $x \in E$. On peut écrire $x = x_1 + x_2$ avec $x_1 \in \operatorname{Ker}(s-I)$ et $x_2 \in \operatorname{Ker}(s+I)$, de sorte que $s(x) = s(x_1+x_2) = s(x_1) + s(x_2) = x_1 - x_2$.

Proposition 101 – interprétation géométrique d'une symétrie

Soient E_1 et E_2 deux sous-espaces vectoriels supplémentaires de l'espace vectoriel E. Soit $s\colon E \to E$ l'application qui à $x \in E$ associe $x_1 - x_2$ où x_1 et x_2 sont définis par $x = x_1 + x_2$ avec $x_1 \in E_1$ et $x_2 \in E_2$. Alors s est une symétrie. On dit que s est la *symétrie par rapport à E_1 dans la direction de E_2* (ou encore : *parallèlement à E_2*).

Exercice de cours 91. Démontrer la proposition.

Dans le cas où E est l'ensemble des vecteurs du plan, et E_1, E_2 deux droites vectorielles distinctes, comme dans le cas où E est l'ensemble des vecteurs de l'espace, E_1 un plan et E_2 une droite qui n'est pas contenue dans E_1, la symétrie par rapport à E_1 dans la direction de E_2 est bien ce que l'on pense, comme l'illustre la figure 2.

Si A est la matrice d'une symétrie s dans des bases données, on a $A^2 = A \times A = I$, puisque que $A \times A$ est la matrice dans les mêmes bases de $s \circ s = I$. Pour une symétrie aussi, on peut trouver des bases dans lesquelles la matrice de A est très simple comme l'illustre l'exercice suivant.

Exercice de cours 92. Soient E_1 et E_2 deux sous-espaces supplémentaires de E, et $s : E \to E$ la symétrie par rapport à E_1 dans la direction de E_2. Soient aussi $\mathscr{B}_1 = (u_1, \dots, u_{k_1})$ une base de E_1 et $\mathscr{B}_2 = (v_1, \dots, v_{k_2})$ une base de E_2. Notant A la matrice de s dans les bases $\mathscr{B}$ et $\mathscr{B}$, montrer que l'on a :

$$A = \mathrm{Mat}_{\mathscr{B}\leftarrow\mathscr{B}}(s) = \begin{array}{c} \begin{array}{ccc|ccc} s(u_1) & \dots & s(u_{k_1}) & s(v_1) & \dots & s(v_{k_2}) \end{array} \\ \left(\begin{array}{ccc|ccc} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & -1 & \dots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & -1 \end{array}\right) \end{array} \begin{array}{c} \\ u_1 \\ \\ u_{k_1} \\ v_1 \\ \\ v_{k_2} \end{array}$$

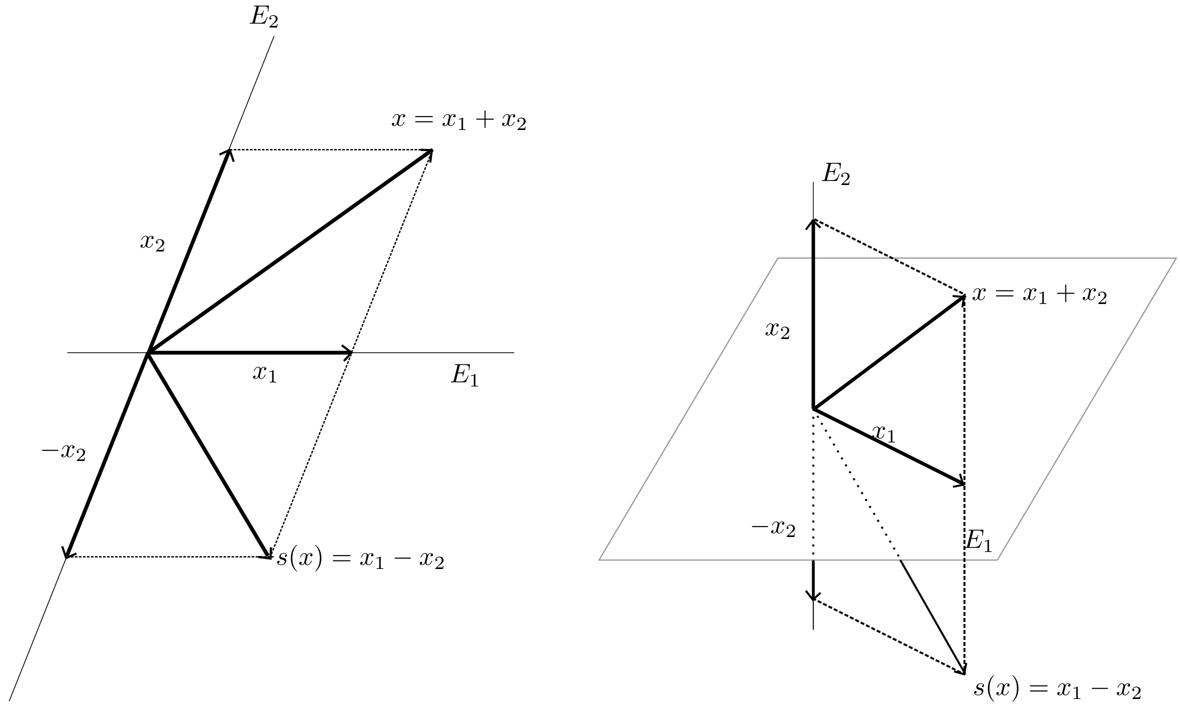


FIGURE 2. Symétries vectorielles dans le plan et dans l'espace

Là encore, on pourra rapprocher ce résultat de l'exercice 87.

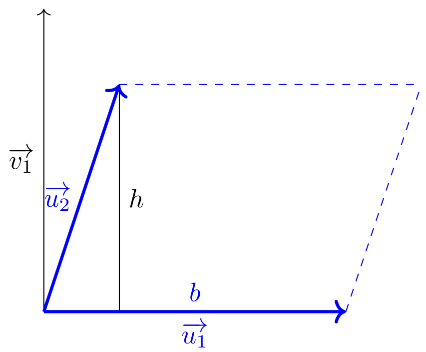
4 Déterminant d'une matrice

L'objet mathématique appelé déterminant est plutôt difficile à définir. Il ne s'agit de rien de moins que de mesurer comment changent les volumes et l'orientation sous l'effet d'une application linéaire. En petite dimension ($n = 1, 2$ ou 3), notre intuition et un peu de travail peuvent permettre de répondre à cette question, mais ce n'est pas immédiat sauf si $n = 1$. Par contre, le calcul d'un déterminant est relativement simple, et dans ce cours, où l'on veut être capable de dire si le déterminant d'une matrice donnée est nul ou pas, c'est l'essentiel.

1. Introduction

Le *déterminant* d'une matrice carrée peut être vu comme une généralisation (ou une formalisation) des notions d'aire et de volume, qui tient compte de l'*orientation*. Nous tentons ici de motiver sa définition dans le cas des matrices carrées d'ordre 2.

On se place dans $\mathbb{R}^2$ muni de sa base canonique $(\overrightarrow{i}, \overrightarrow{j})$ orthonormée, où $\overrightarrow{i} = (1, 0)$ et $\overrightarrow{j} = (0, 1)$, que l'on convient d'appeler *directe*. Les termes *orthonormée* et *directe* seront définis de manière rigoureuse en deuxième année. Ici, *orthonormée* est à comprendre au sens de la géométrie du lycée. Considérons deux vecteurs $\overrightarrow{u_1} = (x_1, y_1)$ et $\overrightarrow{u_2} = (x_2, y_2)$ de $\mathbb{R}^2$ écrit dans la base canonique $(\overrightarrow{i}, \overrightarrow{j})$. On appelle *parallélogramme engendré* par $\overrightarrow{u_1}$ et $\overrightarrow{u_2}$ l'ensemble $\{\alpha_1\overrightarrow{u_1} + \alpha_2\overrightarrow{u_2} : (\alpha_1, \alpha_2) \in [0, 1]^2\}$.



Calculons l'aire $\mathscr{A}$ de ce parallélogramme. On sait qu'elle est égale à la longueur b d'un côté (par exemple $\|\overrightarrow{u_1}\|$), multipliée par la hauteur h correspondante : $\mathscr{A} = b \times h$. Pour obtenir $\mathscr{A}$ en fonction des coordonnées de $\overrightarrow{u_1}$ et $\overrightarrow{u_2}$, il suffit d'introduire le vecteur $\overrightarrow{v_1} = (-y_1, x_1)$. Ce vecteur est en effet orthogonal à $\overrightarrow{u_1}$ et de même norme, de sorte que

$$|\overrightarrow{v_1} \,.\, \overrightarrow{u_2}| = h \times \|\overrightarrow{u_1}\| = h \times b = \mathscr{A},$$

où $\cdot$ désigne le produit scalaire canonique de $\mathbb{R}^2$. On a donc

$$\mathscr{A} = |-y_1x_2 + x_1y_2|.$$

Comme nous le verrons au cours de ce chapitre, pour $(x_1, y_1, x_2, y_2) \in \mathbb{R}^4$, le *déterminant* de la matrice carrée $A = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \in \mathscr{M}_2(\mathbb{R})$, noté $\det(A)$ ou $\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}$, est défini par :

$$\det(A) = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1y_2 - y_1x_2.$$

En particulier, $|\det(A)|$ représente l'aire du parallélogramme engendré par les vecteurs colonnes $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ et $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ de A. Le déterminant possède de plus un signe : il est strictement positif si les vecteurs $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ et $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ forment aussi une base *directe*, strictement négatif s'ils forment une base *indirecte*, et nul si ces vecteurs sont liés.

On peut de même introduire le déterminant d'une matrice carrée réelle d'ordre 3 de sorte que sa valeur absolue soit le volume du *parallélépipède* engendré par les vecteurs colonnes d'une telle matrice.

Dans ce chapitre, nous allons définir le déterminant d'une matrice carré d'ordre quelconque. Il sera bon de garder à l'esprit cet exemple introductif pour mieux comprendre les propriétés du déterminant et la façon dont il est construit.

2. Definition et calcul pratique du déterminant

2.1. Définition. On commence avec une proposition dont l'énoncé est simple, mais dont la preuve requiert un assez gros travail que l'on reporte à plus tard dans ce chapitre. Elle affirme l'existence et l'unicité d'une fonction vérifiant trois propriétés élémentaires.

Proposition 102 – existence et unicité de l'application «déterminant»

Pour tout $n \in \mathbb{N}^*$, il existe une unique application Δ_n de $\mathscr{M}_n(\mathbb{R})$ dans $\mathbb{R}$ qui vérifie les trois propriétés suivantes :

1. $A \mapsto \Delta_n(A)$ est linéaire par rapport à chacune des lignes de A.
2. Si $\tilde{A}$ se déduit de A par l'échange de deux lignes distinctes, alors $\Delta_n(\tilde{A}) = -\Delta_n(A)$.
3. $\Delta_n(I_n) = 1$.

Nous verrons plus loin que les propriétés analogues portant sur les colonnes (qui correspondent à l'intuition vue en introduction) se déduisent de celles sur les lignes (voir la proposition 114). Comme nous allons utiliser le pivot de Gauss, il est plus aisé de travailler sur les lignes dans un premier temps.

Définition 103 – déterminant

L'application donnée par la proposition précédente est appelée application *déterminant*. On la note $\det\colon \mathscr{M}_n(\mathbb{R}) \to \mathbb{R}$.

En dimension 1 et 2, on peut obtenir assez rapidement une formule explicite pour le déterminant à partir des propriétés (1), (2) et (3). C'est l'objet de l'exercice suivant.

Exercice de cours 93.

1. Supposons $n = 1$. Une matrice $A \in \mathscr{M}_1(\mathbb{R})$ s'écrit $A = (a)$ pour un $a \in \mathbb{R}$. Montrer que $\det(A) = a$.
2. Supposons $n = 2$. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathscr{M}_2(\mathbb{R})$. Montrer que

$$\det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = ad - bc.$$

On donnera plus loin une expression explicite du déterminant d'une matrice de taille $\geqslant 3$, mais cette expression est en général inutilisable pour le calculer. On va voir que la méthode du pivot donne un procédé effectif de calcul.

2.2. Calcul d'un déterminant par la méthode du pivot. On commence en étudiant l'effet des opérations sur les lignes d'une matrice sur son déterminant.

Échange de lignes ($L_i \leftrightarrow L_j$). Soit $A \in \mathscr{M}_n(\mathbb{R})$, et $A_{i\leftrightarrow j}$ la matrice obtenue en échangeant deux lignes de A ($L_i \leftrightarrow L_j$). On a

(5) $$\det A_{i\leftrightarrow j} = -\det A.$$

Exemple 34. D'après la remarque 9, $\widetilde{E}_{i,j}$ s'obtient en échangeant les lignes i et j de la matrice identité. On a donc $\det(\widetilde{E}_{i,j}) = -\det(I_n) = -1$.

On en déduit :

Corollaire 104 – une matrice avec deux lignes identiques est de déterminant nul

Si A a deux lignes identiques, alors $\det(A) = 0$.

Exercice de cours 94. Démontrer le corollaire.

Multiplication d'une ligne par un scalaire non nul ($L_i \leftarrow \alpha L_i$). Soit $A \in \mathscr{M}_n(\mathbb{R})$. On note $L_1, \dots, L_n$ ses lignes. Soit $i \in \{1, \dots, n\}$ fixé et $A_{i,\alpha}$ la matrice obtenue en multipliant la ligne L_i par $\alpha \in \mathbb{R}^*$. On a

(6) $$\det(A_{i,\alpha}) = \alpha \det(A).$$

En effet, puisque det est linéaire par rapport à la ligne L_i,

$$\det(L_1, \dots, L_{i-1}, \alpha.L_i, L_{i+1} \dots L_n) = \alpha \det(L_1, \dots, L_{i-1}, L_i, L_{i+1} \dots L_n) = \alpha \det(A).$$

Corollaire 105 – une matrice avec une ligne nulle est de déterminant nul

Si A a une ligne nulle, alors $\det(A) = 0$.

Exercice de cours 95. Démontrer le corollaire.

Ajout à une ligne d'un multiple d'une autre ligne ($L_i \leftarrow L_i + \alpha L_j$). Soient $A \in \mathscr{M}_n(\mathbb{R})$ et $\alpha \in \mathbb{R}^*$. On note $A_{i,\alpha,j}$ la matrice obtenue à partir de A en effectuant l'opération $L_i \leftarrow L_i + \alpha L_j$, où $i \neq j$. On a

(7) $$\det(A_{i,\alpha,j}) = \det A.$$

En effet, supposons par exemple que $i < j$. Alors, puisque det est linéaire par rapport à la i-ème ligne,

$$\begin{aligned}\det(A_{i,\alpha,j}) &= \det(L_1, \dots, L_i + \alpha L_j, \dots, L_j, \dots, L_n)\\ &= \det(L_1, \dots, L_i, \dots, L_j, \dots, L_n) + \alpha \det(L_1, \dots, L_j, \dots, L_j, \dots, L_n)\\ &= \det(A) + 0 = \det(A).\end{aligned}$$

Calcul effectif. Soient $A \in \mathscr{M}_n(\mathbb{R})$ une matrice donnée et $\tilde{A}$ la matrice obtenue à partir de A en appliquant l'algorithme du pivot. On se souvient qu'il s'agit d'effectuer sur A des opérations du type précédent jusqu'à obtenir une matrice échelonnée. Les formules (5), (6) et (7) donnent donc

$$\det(\tilde{A}) = (-1)^e \det A,$$

où $e \in \mathbb{N}$ est le nombre d'échanges de lignes qu'il a été nécessaire de faire. On distingue alors deux cas :

— $\tilde{A}$ a (au moins) une ligne nulle. D'après le corollaire 105, on a $\det(\tilde{A}) = 0$, donc finalement
$$\det A = 0.$$

— $\tilde{A}$ n'a pas de ligne nulle. Dans ce cas, on peut réduire la matrice $\tilde{A}$ jusqu'à obtenir la matrice I_n. Là encore il s'agit de faire subir à $\tilde{A}$ des opérations du type de celles de la section précédente, et plus précisément de multiplier chaque ligne L_i de $\tilde{A}$ par l'inverse de son pivot $\tilde{a}_{i,i}$, puis d'effectuer des opérations du type $L_i \leftarrow L_i + \alpha L_j$. Avec (6) et (7) on obtient
$$\det(I_n) = \frac{1}{\tilde{a}_{1,1}} \frac{1}{\tilde{a}_{2,2}} \dots \frac{1}{\tilde{a}_{n,n}} \det(\tilde{A}).$$
Dans ce cas, puisque $\det(I_n) = 1$, on a donc
$$\det(A) = (-1)^e \tilde{a}_{1,1}\tilde{a}_{2,2} \dots \tilde{a}_{n,n},$$
où les $\tilde{a}_{i,i}$ sont les coefficients diagonaux de la matrice échelonnée $\tilde{A}$.

Notant que A n'a pas de ligne nulle si et seulement si elle est de rang n, ou encore si et seulement si A est inversible, on a démontré la proposition suivante.

Proposition 106 – caractérisation de l'inversibilité d'une matrice

Soit $A \in \mathscr{M}_n(\mathbb{R})$.
— $\boxed{A \text{ n'est pas inversible si et seulement si } \det A = 0}$.
— Si A est inversible, le déterminant de A est le produit des éléments diagonaux de la matrice obtenue en échelonnant A, multiplié par -1 si le nombre d'échanges de lignes effectués est impair.

De plus, si A est inversible, le déterminant de A est le produit des éléments diagonaux de la matrice obtenue en échelonnant A, multiplié par -1 si le nombre d'échanges de lignes effectués est impair.

Remarque 16. Ce résultat constituerait une preuve de l'unicité dans la proposition 102 si l'on savait qu'il y a unicité de la matrice échelonnée associée à A, ce que l'on n'a jamais prouvé. En effet, dans ce cas, si $\Delta\colon \mathscr{M}_n(\mathbb{R}) \to \mathbb{R}$ est une application qui vérifie (i), (ii) et (iii), alors $\Delta(A)$ serait parfaitement déterminé pour n'importe quelle matrice A. Pour ce qui concerne l'existence, il faudrait prouver que l'application définie par la proposition 106 vérifie bien (i), (ii) et (iii), ce qui n'est pas immédiat.

3. Déterminant d'un produit de matrices

On s'intéresse au déterminant d'un produit $A \times B$ de matrices. On commence par le cas où A est une matrice d'opération sur les lignes :

Supposons que $A = \widetilde{E}_{i,j}$. D'après la proposition 48, $A \times B$ est la matrice qu'on obtient à partir de B en échangeant deux de ses lignes. Donc, d'après (5), on a $\det(A \times B) = -\det B$. D'un autre côté, on vient de voir que $\det(\widetilde{E}_{i,j}) = -1$. On a donc $\det(A \times B) = \det(A)\det(B)$ dans ce cas.

On considère maintenant le cas où $A = I + \alpha E_{i,j}$ avec $i \neq j$. D'après (7), on a $\det(A \times B) = \det(B)$. En particulier pour $B = I_n$, on obtient $\det(A) = 1$, et par conséquent, $\det(A \times B) = \det(A)\det(B)$ dans ce cas aussi.

Enfin si $A = I + (\alpha - 1)E_{i,i}$, on sait que $\det(A) = \alpha$ et que $A \times B$ est la matrice obtenue en multipliant la i-ième ligne L_i de B par α. Puisque le déterminant est linéaire par rapport à cette ligne, on obtient encore

$$\det(A \times B) = \det(L_1, \ldots, \alpha L_i, \ldots, L_n) = \alpha \det(L_1, \ldots, L_n) = \alpha \det(B) = \det(A)\det(B).$$

On est alors en mesure de démontrer le résultat très important suivant.

Proposition 107

Soient $A, B \in \mathscr{M}_n(\mathbb{R})$. On a $\det(AB) = \det(A)\det(B)$.

Exercice de cours 96. La démonstration de la proposition est un peu délicate. C'est l'objet de cet exercice.

1. Supposons que A ne soit pas inversible. Montrer que $A \times B$ n'est pas inversible non plus. *Indication : utiliser les les applications linéaires f et g associées à A et B respectivement.* Conclure que $\det(A \times B) = \det(A)\det(B)$ dans ce cas.

2. Supposons maintenant que A soit inversible. Remarquer que A et A^{-1} peuvent s'écrire comme produit de matrices de la forme $\widetilde{E}_{i,j}$ et $(I + \alpha E_{i,j})$. Utiliser alors ce qui précède la proposition pour montrer que $\det(A \times B) = \det(A)\det(B)$.

On en déduit immédiatement le résultat suivant.

Corollaire 108 – le déterminant de l'inverse d'une matrice inversible est l'inverse du déterminant

Soit $A \in \mathscr{M}_n(\mathbb{R})$ une matrice inversible. On a
$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Exercice de cours 97.

1. Vérifier l'assertion du corollaire
2. Montrer, toujours à l'aide de la proposition 107, que deux matrices semblables ont le même déterminant.

L'exercice précédent permet de définir le déterminant d'un endomorphisme.

Définition 109 – déterminant d'un endomorphisme

Soient $f : E \to E$ un endomorphisme de l'espace vectoriel E, supposé de dimension finie. Soient $\mathscr{B}$ une base de E, et
$$A = \mathrm{Mat}_{\mathscr{B}\leftarrow\mathscr{B}}(f).$$
On appelle *déterminant de* f, et l'on note $\det(f)$ le nombre $\det(A)$. Ce nombre ne dépend pas du choix de $\mathscr{B}$.

4. Déterminant et matrice transposée

Définition 110 – transposée d'une matrice carrée

Soit $A = (a_{i,j})$ une matrice de $\mathscr{M}_{n,p}(\mathbb{R})$. On appelle *transposée* de A, et on note A^T la matrice de $\mathscr{M}_{p,n}(\mathbb{R})$ définie par
$$A^T = (t_{i,j}) \text{ avec } t_{i,j} = a_{j,i} \text{ pour tous } i \in \{1, \dots p\}, j \in \{1, \dots, n\}.$$

De manière un peu imprécise, on peut dire que A^T s'obtient à partir de A en échangeant les coefficients de A par symétrie par rapport à sa diagonale. De la définition il vient immédiatement que la transposée de la transposée d'une matrice est la matrice elle-même :

$$(A^T)^T = A. \tag{8}$$

Exemple 35. Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. On a $A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

Exercice de cours 98.

1. Soit $E_{k,\ell}$ une matrice élémentaire de $\mathscr{M}_{n,n}$. Vérifier que l'on a :
$$(E_{k,\ell})^T = E_{\ell,k}.$$
2. Vérifier que la transposée de la somme de deux matrices et la somme des transposées, et que pour $\lambda \in \mathbb{R}$ et $A \in \mathscr{M}_{n,p}(\mathbb{R})$,
$$(\lambda.A)^T = \lambda.A^T.$$
Autrement dit, l'application de $\mathscr{M}_{n,p}(\mathbb{R})$ dans $\mathscr{M}_{n,p}(\mathbb{R})$ qui à une matrice A associe sa transposée A^T est un endomorphisme de $\mathscr{M}_{n,p}(\mathbb{R})$.

La transposée d'un produit de matrices est le produit des transposées, dans l'ordre inverse. Plus précisément :

Proposition 111 – transposée d'un produit de matrices

Soient $A \in \mathscr{M}_{n,p}(\mathbb{R})$ et $B \in \mathscr{M}_{p,m}(\mathbb{R})$, de sorte que $A \times B \in \mathscr{M}_{n,m}(\mathbb{R})$ est bien définie. Alors $B^T \times A^T \in \mathscr{M}_{m,n}(\mathbb{R})$ est bien définie, et on a

$$(A \times B)^T = B^T \times A^T.$$

Exercice de cours 99. Démontrer la proposition.

Corollaire 112 – transposée d'une matrice inversible

Soit $A \in \mathscr{M}_n(\mathbb{R})$. La matrice A est inversible si et seulement si A^T l'est. Dans ce cas

$$(A^T)^{-1} = (A^{-1})^T.$$

Exercice de cours 100. Démontrer le corollaire.

Outre son intérêt propre, le résultat suivant ouvre la porte à l'étude des propriétés du déterminant d'une matrice comme fonction de ses colonnes (ce qui est d'une certaine façon plus naturel si on repense à l'introduction).

Proposition 113 – transposer une matrice ne change pas son déterminant

Soit $A \in \mathscr{M}_n(\mathbb{R})$. On a

$$\det(A^T) = \det(A).$$

Exercice de cours 101. Démontrer la proposition.
Indication : traiter à part comme dans l'exercice 96 le cas où A n'est pas inversible. Pour le cas où A est inversible, reprendre les idées de l'exercice 96 (2).

On en déduit que le déterminant possède les mêmes propriétés par rapport aux colonnes que par rapport aux lignes (voir la proposition 102).

Proposition 114 – propriétés du déterminant relatives aux colonnes

La fonction $\det\colon \mathscr{M}_n(\mathbb{R}) \to \mathbb{R}$ vérifie les propriétés suivantes :

1. $A \mapsto \det(A)$ est linéaire par rapport aux colonnes de A.

2. Si $\widetilde{A}$ se déduit de A par un échange de deux colonnes, alors $\det(\widetilde{A}) = -\det(A)$.

Exercice de cours 102. Démontrer la proposition.

5. Développement par rapport à une ligne ou une colonne

On démontre ici que le déterminant d'une matrice $n \times n$ peut être calculé à partir de n déterminants de taille $n-1$. La proposition ci-dessous à surtout un intérêt théorique, mais peut être utilisée pour calculer des déterminants pour des matrices de petite taille (typiquement $n = 3, 4$ ou 5). Cette formule est appelée «développement du déterminant suivant la j-ème colonne».

$$\widehat{A}_{i,j} = \begin{pmatrix} a_{1,1} & \dots & a_{1,j} & \dots & \dots & a_{1,n} \\ \vdots & & & & & \\ \vdots & & & & & \\ a_{i,1} & \dots & a_{i,j} & \dots & \dots & a_{i,n} \\ \vdots & & & & & \\ a_{n,1} & \dots & a_{n,j} & \dots & \dots & a_{n,n} \end{pmatrix}$$

FIGURE 1. La matrice extraite $\widehat{A}_{i,j}$

Proposition 115 – développement du déterminant suivant la j-ème colonne

Soit $A \in \mathscr{M}_n(\mathbb{R})$. Pour $i, j \in \{1, \dots, n\}$ on note $\widehat{A}_{i,j} \in \mathscr{M}_{n-1}(\mathbb{R})$ la matrice obtenue à partir de A en rayant la i-ième ligne et la j-ième colonne (cf. la figure 1). Pour tout $j \in \{1, \dots, n\}$, on a

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(\widehat{A}_{i,j}).$$

EXERCICE DE COURS 103. Démontrer la proposition. On commencera par le cas $j = 1$, et on remarquera que la première colonne C_1 s'écrit sous la forme

$$C_1 = a_{1,1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_{2,1} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + a_{n,1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \sum_{i=1}^{n} a_{i,1} C_i^0.$$

En utilisant la transposition des matrices, qui ne change pas le déterminant, on peut en déduire le résultat suivant (développement du déterminant par rapport à sa i-ième ligne).

Corollaire 116 – développement du déterminant par rapport à sa i-ième ligne

Pour tout $i \in \{1, \dots, n\}$, on a

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det(\widehat{A}_{i,j}).$$

6. Existence de la fonction déterminant

On démontre maintenant l'existence, pour tout $n \geqslant 2$, d'une fonction $\Delta_n \colon \mathscr{M}_n(\mathbb{R}) \to \mathbb{R}$ qui vérifie les propriétés (1), (2) et (3) de la proposition 102. La formule du développement d'un déterminant par rapport à sa première ligne

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(\widehat{A}_{i,j}) \tag{9}$$

va permettre de faire cette démonstration par récurrence. On note $\mathscr{P}(n)$ la propriété : «il existe une application $\Delta_n \colon \mathscr{M}_n(\mathbb{R}) \to \mathbb{R}$ qui vérifie (1), (2) et (3)».

– $\mathscr{P}(2)$ est vraie : on a vu que l'application $\Delta_2 \colon \mathscr{M}_2(\mathbb{R}) \to \mathbb{R}$ définie par

$$\Delta_2(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = ad - bc,$$

est le seul candidat possible. Or $\Delta_2(I_2) = 1$, donc Δ_2 vérifie (3). De plus

$$\Delta_2(\begin{pmatrix} c & d \\ a & b \end{pmatrix}) = cb - ad = -\Delta_2(\begin{pmatrix} a & b \\ c & d \end{pmatrix}),$$

$$\widehat{A}_{i_2,1} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & \dots & a_{1,n} \\ \vdots & & & & \\ a_{i_1,1} & a_{i_1,1} & \dots & \dots & a_{i_1,n} \\ \vdots & & & & \\ a_{i_2,1} & a_{i_2,1} & \dots & \dots & a_{i_2,n} \\ \vdots & & & & \\ a_{n,1} & a_{n,2} & \dots & \dots & a_{n,n} \end{pmatrix} \qquad \widehat{B}_{i_1,1} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & \dots & a_{1,n} \\ \vdots & & & & \\ a_{i_2,1} & a_{i_2,1} & \dots & \dots & a_{i_2,n} \\ \vdots & & & & \\ a_{i_1,1} & a_{i_1,1} & \dots & \dots & a_{i_1,n} \\ \vdots & & & & \\ a_{n,1} & a_{n,2} & \dots & \dots & a_{n,n} \end{pmatrix}$$

FIGURE 2. Les matrices extraites $\widehat{A}_{i_1,1}$ et $\widehat{B}_{i_2,1}$

donc Δ_2 vérifie (2). Il reste à vérifier la linéarité par rapport à chacune des lignes. On calcule

$$\begin{aligned} \Delta_2(\lambda_1(a_1\ b_1) + \lambda_2(a_2\ b_2), (c\ d)) &= \Delta_2((\lambda_1 a_1 + \lambda_2 a_2\ \lambda_1 b_1 + \lambda_2 b_2), (c\ d)) \\ &= (\lambda_1 a_1 + \lambda_2 a_2)d - (\lambda_1 b_1 + \lambda_2 b_2)c \\ &= \lambda_1(a_1 d - b_1 c) + \lambda_2(a_2 d - b_2 c) \\ &= \lambda_1 \Delta_2((a_1\ b_1), (c\ d)) + \lambda_2 \Delta_2((a_2\ b_2), (c\ d)), \end{aligned}$$

ce qui prouve la linéarité par rapport à la première ligne. La linéarité par rapport à la seconde ligne s'obtient de la même manière, ou bien en utilisant le comportement du déterminant par échange de ligne.

– Supposons $\mathscr{P}(n-1)$ vraie pour un $n \geqslant 3$ donné. On note, pour $A \in \mathscr{M}_n(\mathbb{R})$,

$$\Delta_n(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \Delta_{n-1}(\widehat{A}_{i,1}), \tag{10}$$

où Δ_{n-1} est la fonction dont l'existence est assurée par $\mathscr{P}(n-1)$. On veut montrer que Δ_n vérifie (1), (2) et (3). D'abord puisque $a_{i,1} = 1$ si $i = 1$ et $a_{i,1} = 0$ sinon, on a

$$\Delta_n(I_n) = \Delta_{n-1}(\widehat{A}_{1,1}) = \Delta_{n-1}(I_{n-1}) = 1,$$

ce qui prouve (3). La linéarité par rapport aux $n-1$ dernières colonnes vient de l'hypothèse de récurrence, et la linéarité par rapport à la première colonne est claire à la vue de (10). Cela prouve (1).

Soient $A = (a_{i,j}) \in \mathscr{M}_n(\mathbb{R})$ et $B = (b_{i,j})$ la matrice obtenue à partir de A en échangeant les lignes L_{i_1} et L_{i_2}, avec $i_1 < i_2$. On a

$$\begin{aligned} \Delta_n(B) &= \sum_{i=1}^{n} (-1)^{i+1} b_{i,1} \Delta_{n-1}(\widehat{B}_{i,1}) \\ &= \sum_{\substack{i \neq i_1 \\ i \neq i_2}} (-1)^{i+1} a_{i,1} \Delta_{n-1}(\widehat{B}_{i,1}) + (-1)^{i_1+1} b_{i_1,1} \Delta_{n-1}(\widehat{B}_{i_1,1}) + (-1)^{i_2+1} b_{i_2,1} \Delta_{n-1}(\widehat{B}_{i_2,1}). \end{aligned}$$

On traite chacun des termes de cette somme séparément. Pour $i \neq i_1, i \neq i_2$, la matrice $\widehat{B}_{i,1}$ s'obtient à partir de $A_{i,1}$ en échangeant les lignes L_{i_1} et L_{i_2}. Donc

$$\sum_{\substack{i \neq i_1 \\ i \neq i_2}} (-1)^{i+1} a_{i,1} \Delta_{n-1}(\widehat{B}_{i,1}) = - \sum_{\substack{i \neq i_1 \\ i \neq i_2}} (-1)^{i+1} a_{i,1} \Delta_{n-1}(\widehat{A}_{i,1}).$$

Ensuite $\widehat{B}_{i_1,1}$ devient $\widehat{A}_{i_2,1}$ en effectuant successivement les opérations $L_k \leftrightarrow L_{k-1}$ pour $k = i_2, i_2 - 1, \dots, i_1 + 2$ (cf. la figure 2). Donc

$$\begin{aligned} (-1)^{i_1+1} b_{i_1,1} \Delta_{n-1}(\widehat{B}_{i_1,1}) &= (-1)^{i_1+1} a_{i_2,1} (-1)^{i_2-i_1-1} \Delta_{n-1}(\widehat{A}_{i_2,1}) \\ &= -(-1)^{i_2+1} a_{i_2,1} \Delta_{n-1}(\widehat{A}_{i_2,1}) \end{aligned}$$

Le même raisonnement donne enfin

$$(-1)^{i_2+1} b_{i_2,1} \Delta_{n-1}(\widehat{B}_{i_2,1}) = -(-1)^{i_1+1} a_{i_1,1} \Delta_{n-1}(\widehat{A}_{i_1,1}),$$

ce qui prouve que $\Delta_n(B) = -\Delta_n(A)$, et donc que la fonction Δ_n définie par (10) vérifie $\mathscr{P}(n)$. Cela termine la partie "existence" de la proposition 102.

7. Une formule explicite pour le déterminant

Définition 117 – permutation

Soit $n \in \mathbb{N}$. Les bijections de l'ensemble $\{1, \ldots, n\}$ dans lui-même s'appellent *permutations* de $\{1, \ldots, n\}$. On note $\mathfrak{S}_n$ l'ensemble de ces permutations.

La composée $\sigma \circ \sigma'$ de deux permutations est aussi une permutation : la composition des applications est une loi interne sur $\mathfrak{S}_n$. On sait qu'elle est associative, et l'application identité sur $\{1, \ldots, n\}$, qui est une permutation, est l'élément neutre pour la composition. Par définition toute permutations est bijective, et sa bijection réciproque est une permutation. En résumé :

$$\boxed{(\mathfrak{S}_n, \circ) \text{ est un groupe.}}$$

Définition 118 – transposition

On appelle *transposition* toute application $\tau \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ telle qu'il existe $i, j \in \{1, \ldots, n\}$, avec $i \neq j$, pour lesquels

$$\tau(k) = \begin{cases} k & \text{si } k \notin \{i, j\}, \\ j & \text{si } k = i, \\ i & \text{si } k = j. \end{cases}$$

Une transposition est une permutation : pout tout $k \in \{1, \ldots, n\}$ l'équation $\tau(x) = k$ admet une unique solution. Voici la clé de l'étude de $\mathfrak{S}_n$.

Proposition 119 – Toute permutation s'écrit comme composée de transpositions

Toute permutation $\sigma \in \mathfrak{S}_n$ s'écrit comme composée d'un nombre fini de transpositions. Cette décomposition n'est pas unique, mais la parité du nombre des transpositions est toujours la même pour une permutation donnée.

Exercice de cours 104. Démontrer la proposition.

Définition 120 – signature

On appelle *signature* d'une permutation σ, et on note $\varepsilon(\sigma) \in \{-1, 1\}$ le nombre $(-1)^m$, où m est le nombre de transpositions d'une décomposition quelconque de σ.

On est finalement en mesure de donner une formule explicite pour le déterminant d'une matrice. On prendra garde que cette formule ne permet pas en général un calcul efficace du déterminant : la somme ci-dessous comprend $n!$ termes puisque c'est le nombre d'éléments dans $\mathfrak{S}_n$.

Proposition 121 – formule explicite pour le déterminant d'une matrice

Soient $n \in \mathbb{N}^*$, et $A = (a_{i,j}) \in \mathscr{M}_n(\mathbb{R})$. Si $\Delta \colon \mathscr{M}_n(\mathbb{R}) \to \mathbb{R}$ vérifie les propriétés (1), (2) et (3) de la proposition 102, alors on a

$$\Delta(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \ldots a_{n,\sigma(n)}. \tag{11}$$

Exercice de cours 105. Démontrer la proposition.