

Algèbre Linéaire II

OLMA252

Université Paris-Saclay

Résumé

Ce sont des notes de cours pour le cours OLMA252, Algèbre linéaire 2, à l'Université Paris-Saclay. Elles sont basées sur des notes prises par Benjamin Schraen, Maria Paulo Gomez Aparicio et Johannes Anschütz de 2018 à 2025.

Les notes de cours traitent des espaces euclidiens, des déterminants et la réduction des endomorphismes. Dans l'appendice, les résultats nécessaires d'un cours d'algèbre linéaire 1 sont mis à disposition.

Chaque exposé est suivi d'une section de questions de compréhension et d'exercices typiques qui servent à la préparation des examens.

Les notes de cours sont corrigées par Alexis Vo.¹



1. Merci beaucoup!

Table des matières

1	Exposé 1 : Espaces euclidiens	4
1.1	Produit scalaire	4
1.2	Orthogonalité	7
1.3	Questions de compréhension et exercices typiques	11
2	Exposé 2 : Bases orthonormales	12
2.1	Bases orthonormales	12
2.2	Projections orthogonales	16
2.3	L'algorithme de Gram-Schmidt	17
2.4	Questions de compréhension et exercices typiques	18
3	Exposé 3 : Isométries et adjoints	19
3.1	Isométries	19
3.2	Adjoint d'un endomorphisme	21
3.3	Le groupe orthogonal	24
3.4	Questions de compréhension et exercices typiques	27
4	Exposé 4 : Déterminants, I	28
4.1	Déterminants d'ordre 2	28
4.2	Déterminants d'ordre supérieur	29
4.3	Opérations sur les colonnes d'un déterminant, I	30
4.4	Questions de compréhension et exercices typiques	31
5	Exposé 5 : Déterminants, II	32
5.1	Opérations sur les colonnes d'un déterminant, II	32
5.2	Déterminant d'une famille de vecteurs	35
5.3	Rappels sur les permutations	36
5.4	Une formule explicite pour le déterminant	37
5.5	Opérations sur les lignes	38
5.6	Questions de compréhension et exercices typiques	42
6	Exposé 6 : Déterminants, III	43
6.1	Volume et déterminants	43
6.2	Formes multilinéaires alternées	44
6.3	Formule de Cramer	45
6.4	Utilisation d'un déterminant pour évaluer le rang	46
6.5	Questions de compréhension et exercices typiques	47

7	Exposé 7 : Réduction des endomorphismes, I	48
7.1	Motivation pour la réduction des endomorphismes	48
7.2	Une application aux systèmes d'équations différentielles linéaires . .	49
7.3	Bases adaptées	52
7.4	Questions de compréhension et exercices typiques	54
8	Exposé 8 : Réduction des endomorphismes, II	55
8.1	Sous-espaces propres et valeurs propres	55
8.2	Le polynôme caractéristique	59
8.3	Passage de $\mathbb{R}$ à $\mathbb{C}$	61
8.4	Questions de compréhension et exercices typiques	62
9	Exposé 9 : Réduction des endomorphismes, III	63
9.1	Polynômes et division euclidienne	63
9.2	Trigonalisation	65
9.3	Polynômes d'endomorphismes	68
9.4	Le théorème de Cayley-Hamilton	69
10	Exposé 10 : Réduction des endomorphismes, IV	73
10.1	La trace et le déterminant	73
10.2	Réduction des endomorphismes auto-adjoint	73
10.3	Les groupes $O_2(\mathbb{R})$ et $O_3(\mathbb{R})$	75
10.4	Décomposition $d + n$	77
10.5	L'exponentielle pour des matrices	78
A	Appendix : Pré-requis sur des espaces vectoriels	81
A.1	Rappels sur les corps $\mathbb{R}$ et $\mathbb{C}$	81
A.2	Espaces vectoriels	82
A.3	Sous-espaces vectoriels	83
A.4	Familles génératrices, familles libres, bases	86
A.5	Dimension et sommes directes	90
A.6	Applications linéaires	90
A.7	Noyau, image et rang	91
A.8	Opérations sur les applications linéaires	93
A.9	Matrices	94
A.10	Représentation matricielle d'une application linéaire	95
A.11	Changements de base	96
B	Appendix : Solution d'exercices	98
B.1	Exercice 1.2	98
B.2	Exercice 2.1	98

B.3	Exercice 3.1		101
B.4	Exercice 4.1		103
B.5	Exercice 5.1		104
B.6	Exercice 6.1		105

1 Exposé 1 : Espaces euclidiens

Les espaces vectoriels considérés dans ce chapitre sont *réels*. Soit E un $\mathbb{R}$ -espace vectoriel.

But du exposé 1 : Sans de donné supplémentaire (un “produit scalaire”), il n’est pas raisonnable de parler de longueur d’un vecteur $v \in E$, ou de faire de la géométrie dans E (angle, triangle,...). On va résoudre ce problème en introduisant des produits scalaires.

1.1 Produit scalaire

Définition 1.1. Une forme bilinéaire sur E est une application

$$b : \begin{array}{ccc} E \times E & \rightarrow & \mathbb{R} \\ (x, y) & \mapsto & b(x, y) \end{array}$$

vérifiant les conditions suivantes :

$$\forall (x, y, z) \in E^3, \forall \lambda \in \mathbb{R}, \quad \begin{aligned} b(x + \lambda y, z) &= b(x, z) + \lambda b(y, z), \\ b(x, y + \lambda z) &= b(x, y) + \lambda b(x, z). \end{aligned}$$

La forme bilinéaire b est dite

- (i) symétrique si $\forall (x, y) \in E^2, b(y, x) = b(x, y)$;
- (ii) positive si $\forall x \in E, b(x, x) \geq 0$;
- (iii) définie si $b(x, x) = 0 \Leftrightarrow x = 0$.

Un produit scalaire est une forme bilinéaire symétrique définie positive sur E .

Remarque 1.2. On a :

$$b(0, x) = b(0 + 1 \cdot 0, x) = b(0, x) + 1 \cdot b(0, x) = b(0, x) + b(0, x).$$

En soustrayant $b(0, x)$ de chaque côté, on obtient $b(0, x) = 0$ pour tout $x \in E$. De manière similaire, on montre que $b(x, 0) = 0$ pour tout $x \in E$.

Pour un produit scalaire, on utilise souvent la notation :

$$\langle u, v \rangle$$

pour $u, v \in E$.

Exemple 1.3. (i) Si $E = \mathbb{R}^n$, le *produit scalaire canonique* sur E est de la forme

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = {}^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} (x_1, \dots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n x_i y_i.$$

(ii) Si $E = \mathbb{R}^2$, alors

$$\langle x, y \rangle := 2x_1y_1 + x_2y_2$$

définit un produit scalaire sur $\mathbb{R}^2$, différent du produit scalaire canonique.

▲ Un espace vectoriel non nul possède une infinité de produits scalaires différents !

(iii) Si $E = \mathcal{C}^0([-1, 1], \mathbb{R}) := \{f : [-1, 1] \rightarrow \mathbb{R} \mid f \text{ continue}\}$, alors

$$\langle f, g \rangle := \int_{-1}^1 f(t)g(t) dt$$

définit un produit scalaire sur E .

(iv) Si $E = \text{Mat}_n(\mathbb{R})$, alors

$$\langle A, B \rangle := \text{Tr}({}^tAB)$$

définit un produit scalaire sur E .

Définition 1.4. Un espace euclidien est un $\mathbb{R}$ -espace vectoriel de dimension finie muni d'un produit scalaire $\langle \cdot, \cdot \rangle$.

Définition 1.5. Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien. La norme euclidienne d'un élément $x \in E$ pour ce produit scalaire est définie par $\|x\| := \sqrt{\langle x, x \rangle}$.

$\|x\|$ est bien défini car $\langle x, x \rangle \geq 0$ (un produit scalaire est positif).

Lemme 1.6 (Inégalité de Cauchy-Schwarz). Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien. Pour tout $x, y \in E$, on a :

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

De plus, l'égalité a lieu si et seulement si x et y sont colinéaires, c'est-à-dire s'il existe $t \in \mathbb{R}$ tel que

$$x = ty \quad \text{ou} \quad y = tx.$$

Démonstration. ● Si $v = 0$, on obtient $0 = 0$.

-) Si $v \neq 0$: on considère $\forall t \in \mathbb{R}$,

$$0 \leq \|u + tv\|^2 = \langle u + tv, u + tv \rangle = \langle u, u \rangle + t\langle u, v \rangle + t\langle v, u \rangle + t^2\langle v, v \rangle,$$

(car $\langle \cdot, \cdot \rangle$ est bilinéaire)

$$= \|u\|^2 + 2t\langle u, v \rangle + t^2\|v\|^2 =: f(t)$$

(car $\langle \cdot, \cdot \rangle$ est symétrique). Car $\|v\| \neq 0$, $\mathbb{R} \rightarrow \mathbb{R}$, $t \mapsto f(t)$ est une fonction polynomiale tel que $f(t) \geq 0 \forall t \in \mathbb{R}$. Donc, $f(t)$ n'a pas deux zéros différents, et son discriminant $\Delta = 4\langle u, v \rangle^2 - 4\|u\|^2\|v\|^2$ est ≤ 0 . Donc, on obtient l'inégalité.

-) Cas de l'égalité : on a alors que $\Delta = 0$. Donc, $\exists t \in \mathbb{R}$, $f(t) = 0$. Donc, $\|u + tv\|^2 = \langle u + tv, u + tv \rangle = 0$, Donc $u + tv = 0 \Leftrightarrow u = -tv$ car $\langle \cdot, \cdot \rangle$ est défini. Donc, u, v sont colinéaires. L'égalité est facile à voir si u, v sont colinéaires.

□

La définition suivante sera étudiée dans le cours d'Analyse :

Définition 1.7. Soit E un espace vectoriel. On dit que $N: E \rightarrow \mathbb{R}_{\geq 0}$ est une norme si

- (i) $N(\lambda u) = |\lambda|N(u), \forall \lambda \in \mathbb{R}, \forall u \in E$,
- (ii) $N(u) = 0 \Rightarrow u = 0$
- (iii) $N(u + v) \leq N(u) + N(v), \forall u, v \in E$ (Inégalité triangulaire).

Lemme 1.8. Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien. La norme euclidienne $\|\cdot\|: E \rightarrow \mathbb{R}$, $u \mapsto \|u\| = \sqrt{\langle u, u \rangle}$ est bien une norme sur E .

Démonstration. •) (i) et (ii) de 1.7 sont évidents.

-) (iii) : On a $\|u + tv\|^2 = \|u\|^2 + \|v\|^2 + 2\langle u, v \rangle \leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| = (\|u\| + \|v\|)^2$, en utilisant l'inégalité de Cauchy-Schwarz (1.6). □

Proposition 1.9. On a les identités suivantes : $\forall u, v \in E$

- (i) Identité du parallélogramme : $\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$
- (ii) Identité de polarisation : $\langle u, v \rangle = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2)$

Démonstration.

$$\|u + v\|^2 = \langle u + v, u + v \rangle = \|u\|^2 + 2\langle u, v \rangle + \|v\|^2 \quad (1)$$

et

$$\|u - v\|^2 = \langle u - v, u - v \rangle = \|u\|^2 - 2\langle u, v \rangle + \|v\|^2 \quad (2)$$

Donc, (1)+(2) donne :

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2),$$

d'où 1). Pour 2), on considère (1)-(2) :

$$\|u + v\|^2 - \|u - v\|^2 = 2\langle u, v \rangle - (-2\langle u, v \rangle) = 4\langle u, v \rangle.$$

□

Exercice 1.1. Soient $x, y, z \in \mathbb{R}^3$ tel que $x^2 + y^2 + z^2 = 1$. On veut borner $x + 2y + 3z$.

Solution : Posons $u = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$. Alors,

$$1 = x^2 + y^2 + z^2 = \|u\|^2 \quad \text{et} \quad x + 2y + 3z = \langle u, v \rangle \quad \text{avec} \quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

D'après Cauchy-Schwarz 1.6 : $|\langle u, v \rangle| \leq \|u\| \|v\| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$
De plus, on a l'égalité si (u, v) est liée.

1.2 Orthogonalité

Soit E un $\mathbb{R}$ -espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$.

Définition 1.10.

- (i) Deux vecteurs x et y de E sont dits orthogonaux si $\langle x, y \rangle = 0$. On note $x \perp y$
- (ii) Deux sous-ensembles $A, B \subseteq E$ sont orthogonaux si

$$\forall u \in A, \forall v \in B, \quad \langle u, v \rangle = 0.$$

- (iii) Si $A \subseteq E$ est un sous-ensemble, on appelle orthogonal de A , noté $A^\perp$, l'ensemble

$$A^\perp := \{u \in E \mid \langle u, v \rangle = 0 \quad \forall v \in A\}.$$

Proposition 1.11.

- (i) Si $A \subseteq E$, alors $A^\perp$ est un sous-espace vectoriel de E .
- (ii) Si $A \subseteq B$, alors $B^\perp \subseteq A^\perp$.

(iii) $A^\perp = \text{Vect}(A)^\perp$ où $\text{Vect}(A)$ est le sous-espace vectoriel engendré par le sous-ensemble A .

(iv) $A \subseteq (A^\perp)^\perp$ (**▲** : Pas égal en général, par exemple si A n'est pas un sous-espace vectoriel)

Démonstration. Exercices.

□

Exemple 1.12.

- 1) $E = \mathcal{C}^0([-1, 1], \mathbb{R})$, $\langle f, g \rangle := \int_{-1}^1 f(t)g(t) dt$. Alors, pour $f(t) := \cos(t)$ et $g(t) := \sin(t)$, on a f et g orthogonaux :

$$\int_{-1}^1 \cos(t) \sin(t) dt = \frac{1}{2} \int_{-1}^1 \sin(2t) dt = 0,$$

car $\sin$ est une fonction impaire, c'est-à-dire $\sin(-t) = -\sin(t)$. Nous avons utilisé la formule trigonométrique : $2 \cos(t) \sin(t) = \sin(2t)$.

- 2) $E = \mathbb{R}^n$, avec le produit scalaire canonique $\langle \cdot, \cdot \rangle$. Soit $H \subseteq \mathbb{R}^n$ l'hyperplan défini par

$$H := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \mid x_1 + \cdots + x_n = 0 \right\}.$$

Alors, $H = \text{Vect}(u)^\perp$ où $u := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ et $H^\perp = \text{Vect}(u)$.

Définition 1.13. Soit E un espace euclidien. On appelle dual de E l'espace vectoriel

$$\mathcal{L}(E, \mathbb{R}) := \{f : E \rightarrow \mathbb{R} \mid f \text{ est linéaire}\}.$$

On le note généralement E^* . Un élément $f \in \mathcal{L}(E, \mathbb{R})$ est appelé une forme linéaire.

Rappels : Soient F, F' deux espaces vectoriels de dimension finie. Grâce à A.48, l'espace vectoriel $\mathcal{L}(F, F')$ des applications linéaires de F vers F' est isomorphe à l'espace vectoriel $\mathcal{M}_{m,n}(\mathbb{R})$ des matrices de taille $n := \dim F$ et $m := \dim F'$.

Ainsi, on a :

$$\dim \mathcal{L}(F, F') = n \times m.$$

En particulier, cela implique que $\dim E^* = \dim E$.

Si $f : F \rightarrow F'$ est une application linéaire, on a le résultat suivant par le théorème du rang A.41 :

$$\dim F = \dim \text{Ker}(f) + \dim \text{Im}(f).$$

Proposition 1.14. Soient F et F' deux espaces vectoriels de dimension finie tels que $\dim F = \dim F'$. Si $f \in \mathcal{L}(F, F')$, alors :

$$f \text{ est un isomorphisme d'espaces vectoriels} \quad \Leftrightarrow \quad \text{Ker}(f) = \{0\}.$$

Démonstration. On rappelle que si G, G' sont des sous-espaces vectoriels de dimension finie d'un même espace vectoriel, alors $G = G'$ si et seulement si $G \subseteq G'$ et $\dim G = \dim G'$. En effet, on peut prolonger une base de G à une base de G' , et la longueur d'une base d'un espace vectoriel est indépendante de la base.

-) “ $\Rightarrow$ ” Si f est injective, alors $\text{Ker}(f) = \{0\}$, d'après A.40.
-) “ $\Leftarrow$ ” Supposons que $\text{Ker}(f) = \{0\}$. Alors, $\dim \text{Ker}(f) = 0$ et, d'après le théorème du rang A.41, on a $\dim F = \dim \text{Im}(f)$. Par conséquent, $\text{Im}(f) = F'$. On applique la propriété précédente avec $G := \text{Im}(f) \subseteq G' := F'$. Donc, f est surjective.

□

Lemme 1.15 (Lemme de Riesz). *Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien et $f \in E^*$ une forme linéaire sur E . Alors, il existe un unique vecteur $u \in E$ tel que*

$$f(x) = \langle u, x \rangle, \quad \forall x \in E,$$

c'est-à-dire que la forme linéaire f est représentée par un produit scalaire.

Notation : pour tout $v \in E$, on note f_v l'application :

$$f_v : E \rightarrow \mathbb{R}, \quad x \mapsto \langle v, x \rangle.$$

Ainsi, $f_v \in E^*$ est une forme linéaire pour tout $v \in E$, car le produit scalaire $\langle \cdot, \cdot \rangle$ est linéaire dans le second argument.

Démonstration du lemme de Riesz. On considère l'application $\varphi : E \rightarrow E^*$ définie par

$$\varphi(u) = f_u \quad \text{pour tout } u \in E,$$

où $f_u : E \rightarrow \mathbb{R}$ est la forme linéaire donnée par $f_u(x) = \langle u, x \rangle$ pour tout $x \in E$.

Nous vérifions que φ est linéaire en utilisant que le produit scalaire $\langle \cdot, \cdot \rangle$ est linéaire dans le premier argument.

-) **Injectivité** : Montrons que φ est injective. Supposons que $u \in \text{Ker}(\varphi)$, c'est-à-dire que $f_u(x) = 0$ pour tout $x \in E$. En particulier, pour $x = u$, on a

$$f_u(u) = \langle u, u \rangle = 0.$$

Comme $\langle \cdot, \cdot \rangle$ est défini positif, cela implique que $u = 0$. Donc, $\text{Ker}(\varphi) = \{0\}$, ce qui montre que φ est injective.

-) **Bijektivité** : Puisque φ est linéaire et injective, et que $\dim(E) = \dim(E^*)$, il en découle que φ est bijective (par le théorème de la dimension).

Par conséquent, pour tout $f \in E^*$, il existe un unique $u \in E$ tel que $f = f_u$. Cela conclut la démonstration. $\square$

Exemple 1.16. Soit $E = \mathbb{R}^n$, et $\langle \cdot, \cdot \rangle$ le produit scalaire canonique. Le lemme de Riesz se comprend facilement dans ce cas. Soit $f: \mathbb{R}^n \rightarrow \mathbb{R}$ une forme linéaire. Notons $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. Tout vecteur $x \in \mathbb{R}^n$ peut s'écrire sous la forme

$$x = \sum_{i=1}^n \alpha_i e_i, \quad \alpha_i \in \mathbb{R}, \forall i = 1, \dots, n.$$

Ainsi, la forme linéaire f s'écrit

$$f(x) = \sum_{i=1}^n \alpha_i f(e_i).$$

En utilisant la définition du produit scalaire canonique, cela devient

$$f(x) = \left\langle \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \right\rangle,$$

où $u_i := f(e_i)$ pour chaque i . Ainsi, f est simplement le produit scalaire par le vecteur

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

$\square$

1.3 Questions de compréhension et exercices typiques

- (i) Quelle est la signification du produit scalaire canonique $\langle \cdot, \cdot \rangle$ sur $\mathbb{R}^n$?
- (ii) Donner un exemple d'un produit scalaire (ou d'un espace euclidien) qui n'est pas le produit scalaire canonique sur $\mathbb{R}^n$ (ou sur $\mathbb{R}^n$).
- (iii) Quelle est la norme d'un vecteur dans un espace euclidien ?
- (iv) Quelle est la signification de $v \perp u$ ou de $A^\perp$?
- (v) Est-il vrai que $E^\perp = \{0\}$ ou $E^\perp = E$ pour un espace euclidien E ?
- (vi) Que dit le *lemme de Riesz* ?

Exercice 1.2. (i) Prouver que $B: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y), (v, w) \mapsto 2xv + yw$ est un produit scalaire sur $\mathbb{R}^2$.

- (ii) Démontrer $\begin{pmatrix} 1 \\ 1 \end{pmatrix} \perp \begin{pmatrix} -1 \\ 2 \end{pmatrix}$ dans l'espace euclidien $(\mathbb{R}^2, B)$

2 Exposé 2 : Bases orthonormales

But de l'exposé 2 : Dans $\mathbb{R}^n$ muni de son produit scalaire canonique $\langle \cdot, \cdot \rangle$ la base canonique $(e_1, \dots, e_n)$ est spéciale : elle est orthonormale, i.e., $\langle e_i, e_j \rangle = 0$ si $i \neq j$, et $\langle e_i, e_i \rangle = 1$, $\forall i = 1, \dots, n$. Maintenant, on va généraliser ça à des espaces euclidiens généraux, et en particulier prouver l'existence des bases orthonormales.

2.1 Bases orthonormales

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien et $F \leq E$ un sous-espace vectoriel de E (automatiquement $\dim F < \infty$ car $\dim E < \infty$).

Rappel (1.10) : $F^\perp := \{x \in E \mid \langle x, z \rangle = 0, \forall z \in F\}$, l'orthogonal de F .

Théorème 2.1. *Dans un espace euclidien E , on a la décomposition orthogonale :*

$$E = F \oplus F^\perp$$

En particulier :

-) $\dim F^\perp = \dim E - \dim F$,
-) $F = (F^\perp)^\perp$.

Démonstration. Nous devons montrer :

- (i) $F \cap F^\perp = \{0\}$,
- (ii) $E = F + F^\perp$, c'est-à-dire que pour tout $x \in E$, il existe $x' \in F$ et $x'' \in F^\perp$ tels que $x = x' + x''$.

1. Orthogonalité : $F \cap F^\perp = \{0\}$ Soit $x \in F \cap F^\perp$. Par définition de l'orthogonalité, on a :

$$\langle x, z \rangle = 0, \quad \forall z \in F.$$

En particulier, en prenant $z = x$, il vient :

$$\langle x, x \rangle = 0 \Rightarrow x = 0.$$

Ainsi, $F \cap F^\perp = \{0\}$.

2. Décomposition de $x \in E$ en somme de deux composantes Soit $x \in E$. Considérons l'application linéaire $f_x \in E^*$ définie par :

$$f_x : E \rightarrow \mathbb{R}, \quad y \mapsto \langle x, y \rangle.$$

Notons $f = f_x|_F$ la restriction de f_x à F , qui appartient à F^* . Par le lemme de Riesz (1.15), il existe un unique $x' \in F$ tel que :

$$f = f_{x'}.$$

Cela signifie que pour tout $z \in F$,

$$f_x(z) = f_{x'}(z),$$

mais cette égalité ne s'étend pas nécessairement à tout $y \in E$ (**A**).

Définissons alors $x'' := x - x'$, ce qui permet d'écrire :

$$x = x' + x''.$$

Montrons que $x'' \in F^\perp$: Pour tout $z \in F$,

$$\langle x'', z \rangle = \langle x - x', z \rangle = \langle x, z \rangle - \langle x', z \rangle = f_x(z) - f_{x'}(z) = 0.$$

Ainsi, $x'' \in F^\perp$, et nous obtenons bien $E = F \oplus F^\perp$.

3. Preuve que $F = (F^\perp)^\perp$ On sait que $F \subseteq (F^\perp)^\perp$ d'après (1.11). De plus, on a la décomposition suivante :

$$E = F^\perp \oplus (F^\perp)^\perp.$$

En comparant les dimensions :

$$\dim F = \dim E - \dim F^\perp = \dim E - (\dim E - \dim(F^\perp)^\perp) = \dim(F^\perp)^\perp.$$

Ainsi, $F = (F^\perp)^\perp$. □

Définition 2.2. Soit E un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$.

-) Une famille $(v_i)_{i \geq 0}$ de vecteurs de E est dite **orthogonale** si, pour tout $i \neq j$, on a :

$$\langle v_i, v_j \rangle = 0, \quad \text{c'est-à-dire } v_i \perp v_j.$$

-) Une famille **orthonormale** de E est une famille orthogonale $(v_i)_{i \geq 0}$ qui vérifie en plus :

$$\|v_i\| = 1, \quad \forall i \geq 0.$$

Exemple 2.3. 1) $E = \mathbb{R}^n$ muni du produit scalaire canonique. La base canonique $(e_1, \dots, e_n)$ est orthonormale car

$$\langle e_i, e_j \rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

2) Dans $E = \mathcal{C}^0([-1, 1], \mathbb{R})$ muni de $\langle f, g \rangle := \int_{-1}^1 f(t)g(t)dt$, la famille $(\cos(t), \sin(t))$ est orthogonale (1.12). La famille $(1, t^2)$ n'est pas orthogonale : $\langle 1, t^2 \rangle = \int_{-1}^1 1 \cdot t^2 dt = \frac{2}{3} \neq 0$

Proposition 2.4. *Une famille orthogonale constituée de vecteurs non nuls est libre. En particulier, une famille orthonormale est libre.*

Démonstration. Supposons que $(v_1, \dots, v_n)$ soit une famille orthogonale avec $v_i \neq 0$ pour tout $i = 1, \dots, n$. Montrons qu'elle est libre.

Soit une combinaison linéaire nulle :

$$\sum_{j=1}^n \alpha_j v_j = 0, \quad \text{où } \alpha_j \in \mathbb{R}.$$

Prenons le produit scalaire avec un vecteur v_i de la famille :

$$\left\langle v_i, \sum_{j=1}^n \alpha_j v_j \right\rangle = \sum_{j=1}^n \alpha_j \langle v_i, v_j \rangle.$$

Or, par définition de l'orthogonalité, on a $\langle v_i, v_j \rangle = 0$ pour $i \neq j$, donc :

$$0 = \alpha_i \|v_i\|^2.$$

Comme $v_i \neq 0$, on a $\|v_i\|^2 \neq 0$, ce qui implique $\alpha_i = 0$ pour tout i .

Ainsi, $(v_1, \dots, v_n)$ est une famille libre.

Si de plus la famille est orthonormale, alors $\|v_i\| = 1$, ce qui garantit encore $v_i \neq 0$ pour tout i . La démonstration reste donc valide dans ce cas. $\square$

Définition 2.5. *Si $(E, \langle, \rangle)$ est un espace euclidien, une famille $\mathcal{B} = (e_1, \dots, e_n)$ est une base orthonormale (abrégée BON) si c'est une base et une famille orthonormale.*

Théorème 2.6. *Soit $(E, \langle, \rangle)$ un espace euclidien. Alors, il admet une BON.*

Démonstration. Soit $n := \dim E$. Soit $(e_1, \dots, e_p)$ une famille orthogonale maximale (du point de vue du cardinal p), telle que $e_i \neq 0$ pour tout $i = 1, \dots, p$. Supposons, par l'absurde, que $p < n$.

Posons $F = \text{Vect}(e_1, \dots, e_p)$. Alors, par la décomposition orthogonale (2.1), on a :

$$E = F \oplus F^\perp,$$

et on sait que $\dim F \leq p < n$, donc $\dim F^\perp \neq 0$, c'est-à-dire $F^\perp \neq \{0\}$.

Soit $x \in F^\perp$ et $x \neq 0$. Alors, $(e_1, \dots, e_p, x)$ est une famille orthogonale de cardinal strictement supérieur à p .

Cela contredit l'hypothèse $p < n$, donc il doit être vrai que $p = n$. Ainsi, $(e_1, \dots, e_n)$ est une base de E , car elle est libre et de cardinal $\dim E$.

Pour obtenir une famille orthonormale $(e'_1, \dots, e'_n)$, il suffit de définir :

$$e'_i := \frac{1}{\|e_i\|} e_i, \quad \forall i = 1, \dots, n.$$

Cela garantit que chaque vecteur e'_i est unitaire et la famille $(e'_1, \dots, e'_n)$ reste orthogonale. $\square$

Proposition 2.7. Soit E un espace euclidien et soit $(e_1, \dots, e_n)$ une base orthonormale de E . Si $x \in E$, on a

$$x = \sum_{i=1}^n \langle x, e_i \rangle e_i.$$

Autrement dit le réel $\langle x, e_i \rangle$ est la i -ème coordonnée de x dans la base $(e_1, \dots, e_n)$.

Démonstration. Posons $y = \sum_{i=1}^n \langle x, e_i \rangle e_i$. Alors, $\forall j = 1, \dots, n$

$$\langle x - y, e_j \rangle = \langle x, e_j \rangle - \sum_{i=1}^n \langle x, e_i \rangle \langle e_i, e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle = 0.$$

Donc, $x - y \in \text{Vect}(e_j, 1 \leq j \leq n)^\perp = E^\perp = \{0\}$. Donc, $x - y = 0 \Rightarrow x = y$. $\square$

Corollaire 2.8. $\forall x \in E$, $\|x\|^2 = \sum_{i=1}^n \langle x, e_i \rangle^2$.

Démonstration. Si $x = \sum_{i=1}^n \underbrace{x_i}_{\in \mathbb{R}} e_i$, $x = \langle x, e_i \rangle$ et

$$\|x\|^2 = \left\langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n x_j e_j \right\rangle = \sum_{i,j=1}^n x_i x_j \langle e_i, e_j \rangle = \sum_{i=1}^n x_i^2. \quad \square$$

Proposition 2.9. Soient $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien, et $\mathcal{E} := (e_1, \dots, e_n)$ une base orthonormale de E . Soit $f \in \mathcal{L}(E, E)$ et $A = (a_{ij})_{1 \leq i, j \leq n}$ la matrice représentative de f dans la base $\mathcal{E}$, c'est-à-dire $A = \text{Mat}_{\mathcal{E}}(f)$. Alors, pour tous $i, j \in \{1, \dots, n\}$, on a :

$$a_{ij} = \langle f(e_i), e_j \rangle.$$

Démonstration. La matrice A est définie comme étant la matrice dont les colonnes sont les vecteurs $f(e_j)$, exprimés dans la base $\mathcal{E}$. Autrement dit :

$$A = (c_1 | \dots | c_n),$$

où chaque colonne c_j est donnée par :

$$c_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}, \quad f(e_j) = a_{1j}e_1 + \dots + a_{nj}e_n.$$

Par la proposition 2.7 ci-dessus, on sait que :

$$a_{ij} = \langle f(e_j), e_i \rangle.$$

□

2.2 Projections orthogonales

Soit E un espace euclidien et $F \subset E$ un sous-espace vectoriel. D'après le théorème 2.1, on a $E = F \oplus F^\perp$. Ainsi, pour tout $x \in E$, il existe une décomposition unique de x sous la forme :

$$x = \underset{F}{\overset{\cap}{x_F}} + \underset{F^\perp}{\overset{\cap}{x_{F^\perp}}}$$

Définition 2.10. La projection orthogonale de E dans F est la projection p_F de E sur F , parallèlement à $F^\perp$. Autrement dit, $p_F: E = F \oplus F^\perp \rightarrow F$, où $x = x_F + x_{F^\perp} \mapsto x_F$.

Remarque 2.11. (i) p_F est linéaire, c'est-à-dire $p_F \in \mathcal{L}(E, F)$.

(ii) Pour tout $x \in E$, $p_F(x)$ est complètement caractérisé par la propriété suivante : soit $y \in E$. Alors,

$$y = p_F(x) \Leftrightarrow (y \in F \text{ et } x - y \in F^\perp),$$

et $p_F(x)$ est l'unique élément de F tel que $p_F(x) \in F$ et $x - p_F(x) \in F^\perp$. En particulier, on a $\langle p_F(x), x - p_F(x) \rangle = 0$ et $\|x\|^2 = \|p_F(x)\|^2 + \|x - p_F(x)\|^2$. On en déduit la conséquence suivante : si $(v_1, \dots, v_k)$ est une BON de F , alors

$$\forall x \in E, \quad p_F(x) = \sum_{i=1}^k \langle x, v_i \rangle v_i.$$

En effet, il suffit de vérifier que le vecteur $y = \sum_{i=1}^k \langle x, v_i \rangle v_i$ vérifie

$$y \in F \text{ et } x - y \in F^\perp.$$

Proposition 2.12. Soit $x \in E$. Alors,

$$\|x - p_F(x)\| = \inf \{\|x - y\| \mid y \in F\},$$

c'est-à-dire que $\|x - p_F(x)\|$ est la distance de x à F .

Démonstration. Comme $p_F(x) \in F$, il suffit de prouver que, pour tout $y \in F$, on a $\|x - p_F(x)\| \leq \|x - y\|$. Or, on a

$$\|x - y\|^2 = \|x - p_F(x)\|^2 + \|p_F(x) - y\|^2,$$

car $(x - p_F(x)) \perp (p_F(x) - y)$. En effet, $p_F(x) - y \in F$ et $x - p_F(x) \in F^\perp$. □

2.3 L'algorithme de Gram-Schmidt

Théorème 2.13. Soit E un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$. Soit $(v_1, \dots, v_n)$ une famille libre d'éléments de E . Alors, il existe une famille $(w_1, \dots, w_n)$ orthogonale telle que, pour tout $1 \leq i \leq n$,

$$\text{Vect}(v_1, \dots, v_i) = \text{Vect}(w_1, \dots, w_i).$$

Démonstration. Par récurrence sur i :

$i = 1$: On pose $w_1 := v_1$, ce qui suffit.

$i \geq 1$: Supposons que $(w_1, \dots, w_i)$ ont été construits. Posons $F_i := \text{Vect}(w_1, \dots, w_i)$. On définit alors

$$w_{i+1} := v_{i+1} - p_{F_i}(v_{i+1}),$$

où $p_{F_i}(v_{i+1})$ est la projection orthogonale de v_{i+1} sur F_i . Ainsi, $w_{i+1} \in F_i^\perp$ (par la caractérisation de la projection p_{F_i}), et la famille $(w_1, \dots, w_{i+1})$ est orthogonale.

On vérifie par récurrence que $F_i = \text{Vect}(w_1, \dots, w_i) = \text{Vect}(v_1, \dots, v_i)$. On note que $p_{F_i}(v_{i+1}) \in F_i$, donc

$$\begin{aligned} \text{Vect}(w_1, \dots, w_i, w_{i+1}) &= \text{Vect}(w_1, \dots, w_i, v_{i+1} - p_{F_i}(v_{i+1})) \\ &= \text{Vect}(w_1, \dots, w_i, v_{i+1}) \\ &= \text{Vect}(v_1, \dots, v_i, v_{i+1}). \end{aligned}$$

□

Remarque 2.14. On note que les preuves de 2.13 et 2.11 donnent une recette concrète pour construire une base orthogonale (BON) : Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien et $(v_1, \dots, v_n)$ une base de E . Le but est de construire une base $(w_1, \dots, w_n)$ orthogonale de E telle que $\text{Vect}(w_1, \dots, w_i) = \text{Vect}(v_1, \dots, v_i)$ pour tout i .

- (i) Posons $w'_1 := v_1$.
- (ii) Supposons que $(w_1, \dots, w_i)$ soit déjà construite. Posons

$$w_{i+1} := v_{i+1} - \sum_{j=1}^i \frac{\langle v_{i+1}, w_j \rangle}{\|w_j\|^2} w_j$$

En effet, $p_{F_i}(v_{i+1}) = \sum_{j=1}^i \frac{\langle v_{i+1}, w_j \rangle}{\|w_j\|^2} w_j$ par 2.11 (on applique 2.11 à la BON $(\frac{1}{\|w_1\|} w_1, \dots, \frac{1}{\|w_i\|} w_i)$ de F_i).

- (iii) Pour obtenir une BON $(w'_1, \dots, w'_n)$ de E , posons

$$w'_i := \frac{1}{\|w_i\|} w_i, \quad \forall i = 1, \dots, n.$$

2.4 Questions de compréhension et exercices typiques

- (i) Quelle est la dimension de $F^\perp$?
- (ii) Qu'est-ce qu'une BON ?
- (iii) Donner deux exemples différents de base orthonormale de $\mathbb{R}^2$.
- (iv) Que sont les coordonnées d'un vecteur dans une base orthonormale ? Donner une preuve de 2.7.
- (v) Qu'est-ce qu'une projection orthogonale ?
- (vi) Qu'est-ce que la distance d'un vecteur $x \in \mathbb{R}^3$ à un plan $H \subseteq \mathbb{R}^3$, i.e., $\dim H = 2$? Comment peut-on la calculer ?
- (vii) Comment appliquer l'algorithme de Gram-Schmidt ?

Exercice 2.1. Muni de $\mathbb{R}^3$ et du produit scalaire canonique, soient les vecteurs

$$v_1 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_3 := \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

- (i) Trouver une base orthogonale (e_1, e_2, e_3) de $\mathbb{R}^3$, avec $\text{Vect}(e_1, \dots, e_i) = \text{Vect}(v_1, \dots, v_i)$ pour $1 \leq i \leq 3$.
- (ii) Calculer la projection orthogonale de v_3 sur $\text{Vect}(v_1, v_2)$.

3 Exposé 3 : Isométries et adjoints

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien.

But de l'exposé 3 : Un isomorphisme d'espaces vectoriels permet de transporter chaque propriété d'un espace à l'autre (une base est envoyée vers une autre base, etc.). Pour les espaces euclidiens, on souhaite aussi transporter la propriété du produit scalaire, et cela est possible avec des isométries. De plus, nous allons généraliser la transposée d'une matrice.

3.1 Isométries

Définition 3.1. Une isométrie de E est un endomorphisme $f \in \mathcal{L}(E) := \mathcal{L}(E, E)$ préservant le produit scalaire de E , c'est-à-dire

$$\forall (x, y) \in E^2, \quad \langle f(x), f(y) \rangle = \langle x, y \rangle.$$

Définition 3.2. Soient x et y deux vecteurs non nuls de E . D'après Cauchy-Schwarz $\frac{\langle x, y \rangle}{\|x\| \|y\|} \in [-1, 1]$. On définit l'angle entre les vecteurs x et y comme l'unique élément $\theta \in [0, \pi]$ tel que

$$\langle x, y \rangle = \|x\| \|y\| \cos \theta.$$

On note θ aussi par $\theta(x, y)$.

Proposition 3.3. Soit $f: E \rightarrow E$ une isométrie.

- (i) f est bijective, i.e., $f \in \text{Aut}(E)$.
- (ii) f préserve la distance euclidienne et les angles des vecteurs.

Démonstration.

- (i) : Soit E de dimension finie, donc il suffit de montrer que f est injective. Supposons que $f(v) = 0$. Alors, $\|v\|^2 = \langle v, v \rangle = \langle f(v), f(v) \rangle = 0 \Rightarrow v = 0$.
- (ii) : On a $\|v\| = \sqrt{\langle v, v \rangle} = \sqrt{\langle f(v), f(v) \rangle} = \|f(v)\|$, $\forall v \in E$, donc $\|f(v) - f(u)\| = \|f(u - v)\| = \|u - v\|$. Ainsi, f préserve les distances. Pour les angles, soit $v, u \in E \setminus \{0\}$, donc $f(v), f(u) \in E \setminus \{0\}$ (car f est injective). On a alors :

$$\cos(\theta(f(v), f(u))) = \frac{\langle f(v), f(u) \rangle}{\|f(v)\| \|f(u)\|} = \frac{\langle v, u \rangle}{\|v\| \|u\|} = \cos(\theta(v, u)),$$

donc $\theta(f(v), f(u)) = \theta(v, u)$. □

Proposition 3.4. $f \in \mathcal{L}(E)$ isométrie $\Leftrightarrow \forall u \in E, \|f(u)\| = \|u\|$.

Démonstration.

“ $\Rightarrow$ ” : $\checkmark$ (d’après 3.3)

“ $\Leftarrow$ ” : Si $u, v \in E$, on a $\langle u, v \rangle = \frac{1}{2}(\|u+v\|^2 - \|u\|^2 - \|v\|^2)$ (1.9), donc $\langle f(v), f(u) \rangle = \frac{1}{2}(\|f(u) + f(v)\|^2 - \|f(u)\|^2 - \|f(v)\|^2) = \frac{1}{2}(\|u + v\|^2 - \|u\|^2 - \|v\|^2) = \langle u, v \rangle$. $\square$

Définition 3.5. Soit F un sous-espace vectoriel de E . Donc $E = F \oplus F^\perp$ 2.1, et tout vecteur $v \in E$ s’écrit uniquement $v = v_1 + v_2$ où $v_1 \in F$ et $v_2 \in F^\perp$. On pose :

$$s_F(v) := v_1 - v_2.$$

On appelle s_F la symétrie orthogonale d’axe F .

Proposition 3.6. Une symétrie orthogonale est une isométrie.

Démonstration.

-) Soit $v = v_1 + v_2$, où $v_1 \in F$ et $v_2 \in F^\perp$. On a alors $v_1 = p_F(v)$, où p_F est la projection orthogonale sur F , et $v_2 = v - p_F(v) = (\text{Id}_E - p_F)(v)$. Ainsi, l’application $s_F = p_F - (\text{Id}_E - p_F)$ est une application linéaire de E dans E .
-) Remarque : Pour $v, w \in E$ tels que $v \perp w$, on a

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\langle v, w \rangle = \|v\|^2 + \|w\|^2,$$

ce qui est une version du théorème de Pythagore.

-) s_F préserve la norme : Soit $v \in E$, et décomposons v en $v = v_1 + v_2$, où $v_1 \in F$ et $v_2 \in F^\perp$. On a alors

$$s_F(v) = v_1 - v_2.$$

Comme $v_1 \perp v_2$ (et donc $v_1 \perp -v_2$), on obtient :

$$\|s_F(v)\|^2 = \|v_1\|^2 + \|-v_2\|^2 = \|v_1\|^2 + \|v_2\|^2 = \|v\|^2$$

par le théorème de Pythagore. Ainsi, s_F est une isométrie, d’après la proposition 3.4.

$\square$

Exemple 3.7. (i) Si $v \in F$, alors $s_F(v) = v$. Si $v \in F^\perp$, alors $s_F(v) = -v$.

(ii) Si $\dim F = n - 1$, où $\dim E = n$ (i.e., F est un hyperplan), alors s_F s’appelle la réflexion orthogonale d’axe F .

3.2 Adjoint d'un endomorphisme

Soit E un espace euclidien.

Proposition 3.8. *Soit $f \in \mathcal{L}(E)$. Il existe une unique application linéaire $f^* \in \mathcal{L}(E)$ telle que, pour tous $x, y \in E$, on ait*

$$\langle f(x), y \rangle = \langle x, f^*(y) \rangle.$$

Démonstration. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormale de E (voir 2.6).

-) Supposons d'abord que f^* existe. Il vérifie alors

$$f^*(v) \underset{2.7}{=} \sum_{i=1}^n \langle e_i, f^*(v) \rangle e_i = \sum_{i=1}^n \langle f(e_i), v \rangle e_i,$$

et f^* doit être unique.

-) **Existence** : Posons, pour $v \in E$,

$$f^*(v) := \sum_{i=1}^n \langle f(e_i), v \rangle e_i.$$

Cette définition donne bien une application linéaire car f^* est la somme d'applications linéaires de la forme $E \rightarrow E$, $v \mapsto \langle f(e_i), v \rangle e_i$, pour chaque $i = 1, \dots, n$.

-) Vérification que f^* satisfait la condition demandée : Pour tous $u, v \in E$, on a

$$\begin{aligned} \langle u, f^*(v) \rangle &= \left\langle u, \sum_{i=1}^n \langle f(e_i), v \rangle e_i \right\rangle \\ &= \sum_{i=1}^n \langle f(e_i), v \rangle \langle u, e_i \rangle \\ &= \left\langle \sum_{i=1}^n \langle u, e_i \rangle f(e_i), v \right\rangle \\ &= \left\langle f \left(\sum_{i=1}^n \langle u, e_i \rangle e_i \right), v \right\rangle \\ &= \langle f(u), v \rangle. \end{aligned}$$

Ainsi, f^* est bien l'application linéaire unique vérifiant $\langle f(x), y \rangle = \langle x, f^*(y) \rangle$ pour tous $x, y \in E$.

□

Définition 3.9. f^* s'appelle l'adjoint de f .

Proposition 3.10. Soient $f, g \in \mathcal{L}(E)$.

- (i) $(f^*)^* = f$
- (ii) $(f \circ g)^* = g^* \circ f^*$
- (iii) $\text{Ker}(f^*) = (\text{Im } f)^\perp$ et $\text{Im } f^* = (\text{Ker } f)^\perp$

Démonstration.

-) (i), (ii) Faciles à vérifier.
-) (iii) Soit $u \in \text{Ker}(f^*)$. Par définition, cela signifie que pour tout $v \in E$, on a $\langle f^*(u), v \rangle = 0$. Or, par définition de f^* , on a $\langle f^*(u), v \rangle = \langle u, f(v) \rangle$, ce qui donne :

$$\langle u, f(v) \rangle = 0, \quad \forall v \in E.$$

Cela montre que $u \in (\text{Im } f)^\perp$, donc $\text{Ker}(f^*) \subseteq (\text{Im } f)^\perp$.

Réciproquement, si $u \in (\text{Im } f)^\perp$, alors pour tout $v \in E$, $\langle u, f(v) \rangle = 0$, ce qui implique que $\langle f^*(u), v \rangle = 0$, et donc $u \in \text{Ker}(f^*)$. Ainsi, $\text{Ker}(f^*) = (\text{Im } f)^\perp$.

-) (iv) Soit $y \in \text{Im}(f^*)$. Alors, il existe $z \in E$ tel que $y = f^*(z)$. Soit maintenant $x \in \text{Ker}(f)$. On a alors :

$$\langle y, x \rangle = \langle f^*(z), x \rangle = \langle z, f(x) \rangle = \langle z, 0 \rangle = 0.$$

Ainsi, $y \in (\text{Ker } f)^\perp$, ce qui prouve que $\text{Im}(f^*) \subseteq (\text{Ker } f)^\perp$.

Enfin, par le théorème du rang (voir A.41), on a :

$$\dim \text{Im}(f^*) = \dim E - \dim \text{Ker}(f^*) = \dim E - \dim(\text{Im } f)^\perp = \dim \text{Im } f = \dim(\text{Ker } f)^\perp.$$

Cela montre que $\text{Im}(f^*) = (\text{Ker } f)^\perp$.

□

Remarque 3.11. Soit $\mathcal{E} = (e_1, \dots, e_n)$ une BON de E , soit $A := \text{Mat}_{\mathcal{E}}(f) = (a_{ij})_{1 \leq i, j \leq n}$ la matrice représentative de f . D'après 2.9

$$a_{ji} = \langle f(e_i), e_j \rangle = \langle e_i, f^*(e_j) \rangle = \langle f^*(e_j), e_i \rangle, \quad \forall i, j \in \{1, \dots, n\}.$$

Donc, $\text{Mat}_{\mathcal{E}}(f^*) = {}^t A$ est la transposée de A .

Définition 3.12. Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien et $f \in \mathcal{L}(E)$.

-) f est symétrique (ou auto-adjointe) si $f = f^*$.
-) f est anti-symétrique si $f^* = -f$.

-) f est orthogonale si f est inversible et $f^* = f^{-1}$.

Lemme 3.13. $f \in \mathcal{L}(E)$ est une isométrie $\Leftrightarrow f$ est orthogonale.

Démonstration.

$\Rightarrow$: Supposons que f est une isométrie. Par 3.3, on sait que f est inversible. On a alors, pour tous $u, v \in E$:

$$\langle f(v), u \rangle = \langle f(v), f(f^{-1}(u)) \rangle \underset{f \text{ isomorphisme}}{=} \langle v, f^{-1}(u) \rangle.$$

Cela montre que $f^{-1} = f^*$.

$\Leftarrow$: Supposons maintenant que f est orthogonale. Pour tous $u, v \in E$:

$$\langle f(u), f(v) \rangle = \langle u, f^*(f(v)) \rangle = \langle u, f^{-1}(f(v)) \rangle = \langle u, v \rangle.$$

Cela montre que f préserve les produits scalaires, et donc f est une isométrie. $\square$

Proposition 3.14. Soit $f \in \mathcal{L}(E)$.

- (i) $f = f^* \Leftrightarrow \exists \text{ BON } \mathcal{E} \text{ de } E \text{ telle que } \text{Mat}_{\mathcal{E}}(f) \text{ symétrique} \Leftrightarrow \forall \text{ BON } \mathcal{E} \text{ de } E, \text{Mat}_{\mathcal{E}}(f) \text{ symétrique.}$
- (ii) $f = -f^* \Leftrightarrow \exists \text{ BON } \mathcal{E} \text{ de } E \text{ telle que } \text{Mat}_{\mathcal{E}}(f) \text{ antisymétrique} \Leftrightarrow \forall \text{ BON } \mathcal{E} \text{ de } E, \text{Mat}_{\mathcal{E}}(f) \text{ antisymétrique.}$

Démonstration. Faciles d'après 3.11. $\square$

Proposition 3.15. Soit $F \subseteq E$ un sous-espace stable par $f \in \mathcal{L}(E)$, c'est-à-dire $f(F) \subseteq F$.

- (i) $f^*(F^\perp) \subseteq F^\perp$
- (ii) Si f est symétrique, orthogonale ou antisymétrique, alors $f(F^\perp) \subseteq F^\perp$.

Démonstration.

- (i) Soit $y \in F^\perp$. Montrons que $f^*(y) \in F^\perp$. Soit $x \in F$, on a :

$$\langle f^*(y), x \rangle = \langle y, f(x) \rangle = 0,$$

car $y \in F^\perp$ et donc $\langle y, f(x) \rangle = 0$ pour tout $x \in F$. Ainsi, $f^*(y) \in F^\perp$.

- (ii) Si $f^* = f$, alors d'après (i), on a $f^*(F^\perp) \subseteq F^\perp$ et donc $f(F^\perp) \subseteq F^\perp$.
Si $f^* = -f$, alors par (i), on a $f^*(F^\perp) \subseteq F^\perp$, et donc $f(F^\perp) \subseteq F^\perp$ car $F^\perp$ est un sous-espace vectoriel stable par f .
Si $f^* = f^{-1}$, alors par (i), $f^{-1}(F^\perp) \subseteq F^\perp$. Comme f^{-1} est bijective, il préserve la dimension, et donc $f^{-1}(F^\perp) = F^\perp$, ce qui implique $f(F^\perp) = F^\perp$ par la théorie des ensembles. $\square$

Remarque 3.16. *Application fondamentale de 3.15 :* Supposons f symétrique (ou anti-symétrique ou orthogonale) et F un sous-espace de E tel que $f(F) \subseteq F$, i.e., F est stable par f . Soit $\mathcal{E}'$ une base de F et $\mathcal{E}''$ une base de $F^\perp$. Comme $E = F \oplus F^\perp$, $\mathcal{E} := \mathcal{E}' \cup \mathcal{E}''$ est une base de E (concaténation de bases).

On a :

$$\text{Mat}_{\mathcal{E}}(f) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

(diagonale par blocs) où $A = \text{Mat}_{\mathcal{E}'}(f|_F)$ et $B = \text{Mat}_{\mathcal{E}''}(f|_{F^\perp})$.

3.3 Le groupe orthogonal

Rappels : $I_n = \text{Diag}(1, \dots, 1) \in \mathcal{M}_n(\mathbb{R})$ est la matrice identité.

Définition 3.17.

- (i) Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ telle que ${}^tAA = I_n$ s'appelle une matrice orthogonale.
- (ii) Le groupe orthogonal de taille $n \times n$, noté $O_n(\mathbb{R})$, est l'ensemble des matrices orthogonales :

$$O_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) \mid A \text{ est orthogonale}\}.$$

Proposition 3.18. $O_n(\mathbb{R})$ est un sous-groupe de $\text{GL}_n(\mathbb{R}) := \{A \in \mathcal{M}_n(\mathbb{R}) \mid A \text{ inversible}\}$.

Démonstration.

-) L'élément neutre est $I_n \in O_n(\mathbb{R})$: c'est une matrice orthogonale.
-) Soient $U, V \in O_n(\mathbb{R})$. On montre que $UV \in O_n(\mathbb{R})$:

$${}^t(UV)(UV) = {}^tV {}^tU UV = {}^tV I_n V = {}^tV V = I_n.$$

Donc, $UV \in O_n(\mathbb{R})$.

-) Soit $U \in O_n(\mathbb{R})$. On montre que $U^{-1} \in O_n(\mathbb{R})$:

$${}^tUU = I_n \quad \Rightarrow \quad U^{-1} = {}^tU.$$

Ensuite,

$${}^t({}^tU){}^tU = U({}^tU) = I_n,$$

ce qui montre que $U^{-1} \in O_n(\mathbb{R})$.

□

Théorème 3.19. Soit $M \in \mathcal{M}_n(\mathbb{R})$. Alors, $M \in \text{O}_n(\mathbb{R})$ si et seulement si les colonnes de M forment une base orthonormale de $\mathbb{R}^n$, muni du produit scalaire canonique.

Démonstration. Soient $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. On a

$$\langle X, Y \rangle = \sum_{i=1}^n x_i y_i = {}^t X Y \in \mathcal{M}_{1,1}(\mathbb{R}),$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire standard sur $\mathbb{R}^n$.

Soit $M = (C_1 | \dots | C_n)$ une matrice dont les colonnes sont $C_i \in \mathbb{R}^n$. On a alors

$${}^t M M = ({}^t C_i C_j)_{1 \leq i, j \leq n} = (\langle C_i, C_j \rangle)_{1 \leq i, j \leq n}.$$

Par définition de la matrice orthogonale, ${}^t M M = I_n$ si et seulement si

$$\langle C_i, C_j \rangle = \begin{cases} 1, & \text{si } i = j, \\ 0, & \text{si } i \neq j. \end{cases}$$

Ce qui signifie que la famille $(C_1, \dots, C_n)$ est une base orthonormale de $\mathbb{R}^n$. $\square$

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien.

Proposition 3.20. $f \in \mathcal{L}(E)$ est une isométrie $\Leftrightarrow$ l'image d'une BON est une BON.

Démonstration.

“ $\Rightarrow$ ” : Supposons f une isométrie et soit $\mathcal{B}(e_1, \dots, e_n)$ une BON. Alors,

$$\langle f(e_i), f(e_j) \rangle = \langle e_i, e_j \rangle = \delta_{ij} = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j, \end{cases}$$

donc $(f(e_1), \dots, f(e_n))$ est une BON.

$$\text{“}\Leftarrow\text{”} : \|x\| = \|f(\sum_{i=1}^n x_i e_i)\|^2 = \|\sum_{i=1}^n x_i f(e_i)\|^2 \underset{(f(e_i))_i \text{ BON}}{=} \sum_{i=1}^n x_i^2 \underbrace{\|f(e_i)\|^2}_{=1} = \|x\|^2,$$

donc f est bien une isométrie (3.4). $\square$

Corollaire 3.21. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une BON de E et soit $\mathcal{V} = (v_1, \dots, v_n)$ une autre base de E . Soit $M := P_{\mathcal{B} \rightarrow \mathcal{V}}$ l'unique matrice telle que $M e_i = v_i$, $\forall i = 1, \dots, n$. Alors, $\mathcal{V}$ BON $\Leftrightarrow M \in \text{O}_n(\mathbb{R})$.

Démonstration. Soit $f \in \mathcal{L}(E)$ l'application linéaire telle que $f(e_i) = v_i$, $\forall i = 1, \dots, n$. On note $M = \text{Mat}_{\mathcal{B}}(f)$ (par définition). Alors, $\mathcal{V} \text{ BON} \xleftrightarrow[3.20]{3.13} f$ isométrie $\xleftrightarrow[3.13]{3.20} M$ orthogonale. $\square$

Exemple 3.22. Si $\theta \in \mathbb{R}$, les colonnes de la matrice

$$R(\theta) := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

(appelée *matrice de rotation* par θ) sont une BON de $\mathbb{R}^2$:

$$\left\langle \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \right\rangle = \cos^2(\theta) + \sin^2(\theta) = 1 = \left\langle \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}, \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix} \right\rangle$$

et

$$\left\langle \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix} \right\rangle = \cos(\theta)\cos(\theta) - \sin(\theta)\sin(\theta) = 0,$$

donc $R(\theta) \in O_2(\mathbb{R})$ (3.21).

Exercice : $R(\theta)^{-1} = R(-\theta)$, $\forall \theta \in \mathbb{R}$.

Lemme 3.23. Soit $v = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$. Alors, il existe $\theta \in \mathbb{R}$ et $\lambda \in \mathbb{R}_{\geq 0}$ tels que

$$v = \lambda R(\theta) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ où } R(\theta) \text{ est la matrice de rotation de l'angle } \theta.$$

Démonstration.

-) Posons $\lambda := \|v\| = \sqrt{a^2 + b^2}$.
-) Si $\lambda = 0$, alors $v = 0$ et la propriété est triviale.
-) Supposons maintenant que $\lambda \neq 0$. Posons $x := \frac{a}{\lambda}$ et $y := \frac{b}{\lambda}$. On a alors $x^2 + y^2 = 1$.
-) Par conséquent, il existe un $\theta \in \mathbb{R}$ tel que $x = \cos(\theta)$ et $y = \sin(\theta)$. Ainsi, on peut écrire

$$v = \lambda R(\theta) \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

où $R(\theta)$ est la matrice de rotation d'angle θ .

$\square$

3.4 Questions de compréhension et exercices typiques

- (i) Qu'est-ce qu'une isométrie? Donner deux exemples. Faire des dessins sur leur comportement.
- (ii) Soit $u \perp v$ dans un espace euclidien E , et $f: E \rightarrow E$ linéaire. Est-il vrai que $f(v) \perp f(u)$?
- (iii) Que dit le *lemme de Cauchy-Schwartz*? Pourquoi est-ce important pour la définition d'un angle entre deux vecteurs?
- (iv) Qu'est-ce que l'adjoint d'un endomorphisme?
- (v) Comment l'écrire en termes de matrices?
- (vi) Prouver que $f^*(F^\perp) \subseteq F^\perp$.
- (vii) Qu'est-ce qu'un groupe orthogonal? Donner deux éléments différents dans ce groupe.

Exercice 3.1. Soit $\mathbb{R}^n$ muni du produit scalaire canonique.

- (i) Soit $\mathcal{B} = (f_1, \dots, f_n)$ une base orthonormée de $\mathbb{R}^n$. Prouver qu'il existe une matrice $P \in O_n(\mathbb{R})$ telle que $(Pf_1, \dots, Pf_n)$ soit la base canonique de $\mathbb{R}^n$.
- (ii) Soit $(E, \langle, \rangle)$ un espace euclidien, et $f, g \in \mathcal{L}(E)$ deux isométries. Prouver que $f \circ g$ est une isométrie de E .

4 Exposé 4 : Déterminants, I

But de l'exposé 4 : Dans les précédents chapitres, nous avons abordé les notions de longueur, d'angle, et d'autres propriétés des vecteurs dans un espace euclidien. Maintenant, nous allons nous intéresser à la question du *volume*. Cette question sera traitée à l'aide des déterminants des matrices.

4.1 Déterminants d'ordre 2

Question : On cherche le volume $\text{Vol}(v, w)$ du parallélogramme engendré par

$$v := \begin{pmatrix} a \\ b \end{pmatrix}, \quad w := \begin{pmatrix} c \\ d \end{pmatrix} \in \mathbb{R}^2.$$

Solution : $\text{Vol}(v, w) = |\det(M)|$ est la valeur absolue du *déterminant* de $M = (v|w) \in \mathcal{M}_2(\mathbb{R})$ où

$$\det(M) := ad - bc \in \mathbb{R}.$$

Comment le vérifier ? On suppose que $\text{Vol}(v, w)$ satisfait les propriétés suivantes :

- (i) Si $\theta \in \mathbb{R}$, $\text{Vol}(R(\theta)v, R(\theta)w) = \text{Vol}(v, w)$. (Invariance par rotation)
- (ii) Si $a, b \in \mathbb{R}$, $\text{Vol}\left(\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} v, \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} w\right) = |a||b| \text{Vol}(v, w)$. (Prolongement par des scalaires)
- (iii) $\text{Vol}(v, w) = \text{Vol}(w, v)$. (Invariance par permutation)
- (iv) (e_1, e_2) la base canonique de $\mathbb{R}^2$, alors $\text{Vol}(e_1, e_2) = 1$. (Volume standard)
- (v) Si $\lambda \in \mathbb{R}$, alors $\text{Vol}(v, \lambda v + w) = \text{Vol}(v, w)$. (Invariance par déformation)

On va vérifier que la valeur absolue $|\det(M)|$ du déterminant satisfait les mêmes propriétés, et de plus $|\det(M)| = \text{Vol}(v, w)$.

Remarque 4.1. Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$. Alors,

$$\det(M) = ad - bc \neq 0 \Leftrightarrow M \text{ est inversible.}$$

En fait, M inversible $\Leftrightarrow \text{Ker}(M) = 0$, et nous avons :

-) Si

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0,$$

cela donne le système d'équations :

$$\begin{cases} L_1 : ax + by = 0 \\ L_2 : cx + dy = 0 \end{cases}$$

En multipliant L_1 par d et L_2 par $-b$, on obtient :

$$dL_1 - bL_2 \Rightarrow (ad - bc)x = 0,$$

et en multipliant L_2 par a et L_1 par $-c$, on obtient :

$$aL_2 - cL_1 \Rightarrow (ad - bc)y = 0.$$

Ainsi, si $\det(M) = ad - bc \neq 0$, cela implique que $x = y = 0$.

-) Si $\det(M) = 0$, on peut dériver la relation suivante :

$$d \begin{pmatrix} a \\ c \end{pmatrix} - c \begin{pmatrix} b \\ d \end{pmatrix} = 0,$$

ce qui montre que les colonnes de M ne sont pas linéairement indépendantes. Cela prouve que M n'est pas inversible car les colonnes ne forment pas une base (ou si $d = c = 0$).

4.2 Déterminants d'ordre supérieur

On suppose $K = \mathbb{R}$ ou $K = \mathbb{C}$.

Notations : Si $n \geq 1$ et $p \geq 1$ sont deux entiers. Si $C_1, \dots, C_n$ sont n vecteurs colonnes de K^p , on note $(C_1 | \dots | C_n)$ la matrice $\mathcal{M}_{p,n}(K)$ dont les colonnes successives sont $C_1, \dots, C_n$. Si $A \in \mathcal{M}_n(K)$ et si i et j sont deux entiers compris entre 1 et n , on note $A_{i,j} \in \mathcal{M}_{n-1}(K)$ la matrice carrée de taille $n - 1$ obtenue à partir de A en supprimant dans A la i -ème ligne et la j -ème colonne.

On définit par récurrence sur $n \geq 1$ une application

$$\det : \mathcal{M}_n(K) \longrightarrow K$$

de la façon suivante :

-) si $n = 1$, et $A = (a)$, on pose $\det A = a$;
-) si $n > 1$ et si $\det : \mathcal{M}_{n-1}(K) \rightarrow K$ est bien définie, on définit, pour $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(K)$,

$$\begin{aligned} \det A &:= a_{1,1} \det A_{1,1} - a_{1,2} \det A_{1,2} + \dots + (-1)^{n-1} a_{1,n} \det A_{1,n} \\ &= \sum_{i=1}^n (-1)^{i-1} a_{1,i} \det A_{1,i}. \end{aligned}$$

On a donc bien défini, pour toute matrice $A \in \mathcal{M}_n(K)$, un scalaire $\det A \in K$, appelé *le déterminant de A*. On le note aussi

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

Exemple 4.2. (i) Si $n = 2$, on a

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

comme défini dans 4.1.

(ii) Si $n = 3$, on a

$$\begin{aligned} \det(A) &= \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} \\ &= a_{1,1} \det(A_{1,1}) - a_{1,2} \det(A_{1,2}) + a_{1,3} \det(A_{1,3}) \\ &= a_{1,1} \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} - a_{1,2} \begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} + a_{1,3} \begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} \\ &= a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{3,2}a_{2,3} - a_{2,1}a_{1,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} - a_{3,1}a_{2,2}a_{1,3}. \end{aligned}$$

4.3 Opérations sur les colonnes d'un déterminant, I

Théorème 4.3. *L'application $\det$ est linéaire en chaque colonne. Plus précisément, étant donné $n \geq 1$, ainsi que $1 \leq j \leq n$ et $n - 1$ vecteurs colonnes $C_1, \dots, C_{j-1}, C_{j+1}, \dots, C_n$ de K^n , alors l'application*

$$\begin{array}{ccc} K^n & \longrightarrow & K \\ C & \longmapsto & \det(C_1 | \dots | C_{j-1} | \underbrace{C}_{j} | C_{j+1} | \dots | C_n) \end{array}$$

est linéaire, c'est-à-dire

$$\begin{aligned} \forall C, C' \in K^n, \forall \lambda \in K, \quad & \det(C_1 | \dots | C_{j-1} | C + \lambda C' | C_{j+1} | \dots | C_n) \\ &= \det(C_1 | \dots | C_{j-1} | C | C_{j+1} | \dots | C_n) + \lambda \det(C_1 | \dots | C_{j-1} | C' | C_{j+1} | \dots | C_n). \end{aligned}$$

Démonstration. Par récurrence sur n :

-) $n = 1$ ✓

-) $n \geq 2$: Supposons vrai pour $n - 1$. Montrons que $\det(A)$ est linéaire par rapport à la k -ème colonne. Par définition, $\det(A)$ est une somme de termes de la forme $(-1)^{1+j} a_{1j} \det(A_{1j})$, avec $j = 1, \dots, n$. Il suffit de montrer que chacun de ces termes dépend linéairement de la k -ème colonne de A :
 -) Si $j \neq k$, on a que a_{1j} ne dépend pas de la k -ème colonne de A , et la matrice A_{1j} contient la colonne

$$\begin{pmatrix} a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix}$$

(comme k -ème colonne si $k < j$, comme $(k - 1)$ -ème colonne si $k > j$). Par hypothèse, $\det(A_{1j})$ est linéaire par rapport à cette colonne. Donc, $(-1)^{1+j} a_{1j} \det(A_{1j})$ dépend linéairement de la k -ème colonne de A pour $j = 1, \dots, n$.

-) Si $j = k$, $a_{1j} = a_{1k}$ dépend linéairement de la k -ème colonne de A et A_{1k} ne contient pas la k -ème colonne. Donc, $(-1)^{1+j} a_{1j} \det(A_{1j})$ dépend linéairement de la k -ème colonne de A pour $j = 1, \dots, n$. $\square$

4.4 Questions de compréhension et exercices typiques

- Quel est le déterminant de $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, de $\begin{pmatrix} d & c \\ a & b \end{pmatrix}$, de $\begin{pmatrix} 2 & 0 \\ 2 & 5 \end{pmatrix}$?
- Que dit la linéarité de l'application $\det$ par rapport à chaque colonne ?

Exercice 4.1. Vérifier que l'application $\mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$, $(v, w) \mapsto |\det(v|w)|$ satisfait les propriétés (i) – (v) dans 4.1.

5 Exposé 5 : Déterminants, II

But de l'exposé 5 : La définition d'un déterminant n'est pas la façon dont on veut le calculer (car la définition inductive donne une somme avec beaucoup de termes). On va maintenant développer des méthodes pour faciliter le calcul des déterminants.

5.1 Opérations sur les colonnes d'un déterminant, II

Lemme 5.1. *Soit $A \in \text{Mat}_n(K)$ une matrice carrée de taille $n \times n$. Si deux colonnes adjacentes de A sont égales, alors*

$$\det(A) = 0.$$

Démonstration. Par récurrence sur n .

-) **Cas de base :** $n = 2$. Si $A = \begin{pmatrix} a & a \\ c & c \end{pmatrix}$, alors

$$\det(A) = ac - ca = 0.$$

Donc, le résultat est vérifié pour $n = 2$. ✓

-) **Hypothèse de récurrence :** Supposons que la propriété est vraie pour une matrice de taille $n - 1$. Montrons qu'elle est vraie pour $n \geq 3$.

Soit $A = (C_1 | C_2 | \dots | C_n)$ une matrice de taille $n \times n$, avec $C_i \in K^n$ pour tous $i = 1, \dots, n$. Supposons que $C_k = C_{k+1}$ pour un certain k .

-) Si $j \notin \{k, k+1\}$, alors la matrice A_{1j} (qui résulte de la suppression de la première ligne et de la colonne j de A) contient deux colonnes égales adjacentes, ce qui implique par hypothèse de récurrence que $\det(A_{1j}) = 0$.
-) Donc, le déterminant de A peut s'écrire comme

$$\det(A) = (-1)^{1+k} a_{1k} \det(A_{1k}) + (-1)^{1+k+1} a_{1,k+1} \det(A_{1,k+1}).$$

Or, comme $C_k = C_{k+1}$, nous avons $A_{1k} = A_{1,k+1}$. Par conséquent, nous obtenons

$$\det(A) = (-1)^{1+k} a_{1k} \det(A_{1k}) + (-1)^{1+k+1} a_{1k} \det(A_{1k}) = 0.$$

Ainsi, $\det(A) = 0$ lorsque $C_k = C_{k+1}$, ce qui conclut la démonstration. $\square$

Lemme 5.2. *Si dans un déterminant on échange entre elles deux colonnes adjacentes, le déterminant change de signe.*

Démonstration. Soit $A = (C_1 | \dots | C_n) \in \text{Mat}_n(K)$. Remplaçons dans A les colonnes j et $j+1$ par leur somme, donc par 5.1 et 4.3

$$\begin{aligned} 0 &= \det((C_1 | \dots | C_j + C_{j+1} | C_j + C_{j+1} | \dots | C_n)) \\ &= \underbrace{\det((C_1 | \dots | C_j | C_j | \dots | C_n))}_{=0} + \det((C_1 | \dots | C_{j+1} | C_j | \dots | C_n)) \\ &\quad + \det((C_1 | \dots | C_{j+1} | C_j | \dots | C_n)) + \underbrace{\det((C_1 | \dots | C_{j+1} | C_{j+1} | \dots | C_n))}_{=0}, \end{aligned}$$

alors $\det((C_1 | \dots | C_{j+1} | C_j | \dots | C_n)) = -\det((C_1 | \dots | C_j | C_{j+1} | \dots | C_n))$. $\square$

Corollaire 5.3. *Soit $A = (C_1 | \dots | C_n) \in \text{Mat}_n(K)$.*

- (i) *Si $C_i = C_k$ pour $i \neq k$, alors, $\det(A) = 0$.*
- (ii) *Si on échange deux colonnes de A , le $\det(A)$ change de signe.*
- (iii) *On ne change pas un déterminant en ajoutant à une colonne une combinaison linéaire des autres colonnes.*

Démonstration. (i) Par échanges successifs entre colonnes adjacentes, on peut modifier la matrice A jusqu'à ce que les colonnes égales deviennent adjacentes. Chaque échange de colonnes change le signe du déterminant. Après les échanges, les colonnes adjacentes étant égales, on conclut que $\det(A) = 0$ grâce au Lemme 5.1.

(ii) La démonstration de 5.2 est applicable ici, grâce à (i) et à l'idée que l'échange de colonnes adjacentes n'affecte pas l'égalité des colonnes, ce qui implique que le déterminant est nul.

(iii) Soit $1 \leq k \leq n$ et $\lambda_1, \dots, \lambda_{k-1}, \lambda_{k+1}, \dots, \lambda_n \in K$. Considérons la matrice obtenue en remplaçant la k -ème colonne par la combinaison linéaire

$$C_k + \sum_{i \neq k} \lambda_i C_i.$$

Par la linéarité du déterminant en chaque colonne (cf. Théorème 4.3), on a :

$$\det((C_1 | \dots | C_k + \sum_{i \neq k} \lambda_i C_i | \dots | C_n)) = \det((C_1 | \dots | C_k | \dots | C_n)) + \sum_{i \neq k} \lambda_i \det((C_1 | \dots | C_i | \dots | C_n)).$$

Pour $i \neq k$, on a $\det((C_1 | \dots | C_i | \dots | C_n)) = 0$ car deux colonnes sont égales (la k -ème et l' i -ème). Ainsi, on obtient :

$$\det((C_1 | \dots | C_k + \sum_{i \neq k} \lambda_i C_i | \dots | C_n)) = 0.$$

$\square$

Exemple 5.4. 5.3 aide *énormement* pour calculer des déterminants :

$$\begin{vmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 2 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 2 & -3 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 2 & -3 & 5 \end{vmatrix} = 1 \begin{vmatrix} 1 & 1 \\ -3 & 5 \end{vmatrix} = 8.$$

Par contre, la définition du déterminant donnerait :

$$\begin{vmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 2 & -1 & 1 \end{vmatrix} = 1 \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} - 1 \begin{vmatrix} 0 & 1 \\ 2 & 1 \end{vmatrix} - 2 \begin{vmatrix} 0 & 1 \\ 2 & -1 \end{vmatrix} = 1 \cdot 2 - 1 \cdot (-2) - 2 \cdot (-2) = 8$$

Corollaire 5.5. Soit $n \geq 1$ et soit $A \in \mathcal{M}_n(K)$. La matrice A est inversible si et seulement si $\det(A) \neq 0$.

Démonstration. Récurrence sur n . Pour $n = 1$, on a déjà la validité du résultat (cas trivial).

Soit B une matrice obtenue à partir de A par les opérations suivantes :

- (i) permutation des colonnes,
- (ii) addition à une colonne une combinaison linéaire d'autres colonnes,
- (iii) multiplication d'une colonne par un scalaire $\lambda \neq 0$.

Les colonnes de A et de B engendrent le même sous-espace de $\mathbb{K}^n$, ce qui implique que A est inversible si et seulement si B l'est. Par les résultats des Lemmes 5.3 et 4.3, on a aussi :

$$\det(A) \neq 0 \Leftrightarrow \det(B) \neq 0.$$

Ainsi, nous pouvons remplacer A par une telle matrice B .

$\Rightarrow$: Si $A = (a_{i,j})_{1 \leq i,j \leq n}$ est inversible, alors il existe un $j \in \{1, \dots, n\}$ tel que $a_{1,j} \neq 0$. On peut obtenir une matrice B telle que la première ligne de B soit $(1, 0, \dots, 0)$. Ainsi, $\det(B) = \det(B_{1,1})$ où $B_{1,1}$ est une sous-matrice de B . Par hypothèse de récurrence, $B_{1,1}$ est inversible, donc $\det(B_{1,1}) \neq 0 \Rightarrow \det(B) \neq 0 \Rightarrow \det(A) \neq 0$.

$\Leftarrow$: Si A n'est pas inversible, les colonnes $C_i, \forall i = 1, \dots, n$, ne sont pas linéairement indépendantes. Ainsi, il existe des scalaires $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tels que :

$$\lambda_1 C_1 + \lambda_2 C_2 + \dots + \lambda_n C_n = 0$$

et pour un certain j , $\lambda_j \neq 0$. En ajoutant $\sum_{i \neq j} \frac{\lambda_i}{\lambda_j} C_i$ à la colonne C_j , on obtient une matrice $B = (C_1 | \dots | 0 | \dots | 0)$ où la j -ème colonne devient nulle. Ainsi, $\det(B) = 0$ et donc $\det(A) = 0$. $\square$

5.2 Déterminant d'une famille de vecteurs

Soit E un espace vectoriel de dimension finie $n \geq 1$ et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . On rappelle l'application linéaire

$$X_{\mathcal{B}}: E \rightarrow K^n, v \mapsto \begin{pmatrix} a_1 \\ \dots a_n \end{pmatrix}$$

où $v = a_1 e_1 + \dots + a_n e_n$ (A.10).

Définition 5.6. Soit $(v_1, \dots, v_n)$ une famille de n vecteurs de E . Le déterminant de la famille $(v_1, \dots, v_n)$ dans la base $\mathcal{B}$ est le scalaire

$$\det_{\mathcal{B}}(v_1, \dots, v_n) := \det((X_{\mathcal{B}}(v_1) | \dots | X_{\mathcal{B}}(v_n))) \in K.$$

Autrement dit $\det_{\mathcal{B}}(v_1, \dots, v_n)$ est le déterminant de la matrice carrée de taille $n \times n$ dont la i -ième colonne est le vecteur colonne des coordonnées de v_i dans la base $\mathcal{B}$.

Théorème 5.7. 1) L'application $\det_{\mathcal{B}} : E^n \rightarrow K$ est linéaire en chaque variable. Autrement dit si $1 \leq i \leq n$ et si $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ sont $n-1$ vecteurs de E , alors l'application

$$\begin{array}{ccc} E & \longrightarrow & K \\ v & \longmapsto & \det_{\mathcal{B}}(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_n) \end{array}$$

est linéaire.

2) L'application $\det_{\mathcal{B}}$ est alternée, autrement dit pour $(v_1, \dots, v_n) \in E^n$, et pour $i < j$, on a

$$\det_{\mathcal{B}}(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = - \det_{\mathcal{B}}(v_1, \dots, \underbrace{v_j}_{\text{indice } i}, \dots, \underbrace{v_i}_{\text{indice } j}, \dots, v_n).$$

Démonstration. Conséquence immédiate du 4.3 et 5.3 et de la linéarité de $X_{\mathcal{B}}(-)$. $\square$

Théorème 5.8. Soit E un K -espace vectoriel de dimension n . Soit $\mathcal{B}$ une base de E . Soit $(v_1, \dots, v_n) \in E^n$ une famille de n vecteurs de E . Alors $\det_{\mathcal{B}}(v_1, \dots, v_n) \neq 0$ si et seulement si la famille $(v_1, \dots, v_n)$ est une base de E .

Démonstration. La famille $(v_1, \dots, v_n)$ est une base de $E \Leftrightarrow (X_{\mathcal{B}}(v_1), \dots, X_{\mathcal{B}}(v_n))$ est une base de K^n (car $X_{\mathcal{B}}: E \rightarrow K^n$ est un isomorphisme) $\Leftrightarrow$ la matrice $(X_{\mathcal{B}}(v_1) | \dots | X_{\mathcal{B}}(v_n)) \in \mathcal{M}_n(K)$ est inversible $\stackrel{5.5}{\Leftrightarrow} \det_{\mathcal{B}}(v_1, \dots, v_n) = \det((X_{\mathcal{B}}(v_1) | \dots | X_{\mathcal{B}}(v_n))) \neq 0$. $\square$

5.3 Rappels sur les permutations

Soit $n \geq 1$.

Définition 5.9. Une permutation (de n éléments) est une application $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ bijective. On note par S_n l'ensemble des permutations σ de $\{1, \dots, n\}$, et par $\circ: S_n \times S_n \rightarrow S_n$, $(\sigma, \tau) \mapsto \sigma \circ \tau$ la composition des permutations.

Lemme 5.10. $(S_n, \circ)$ est un groupe de cardinal $n!$.

Démonstration. •) $\circ$ est bien défini : la composition des applications bijectives est bijective.

-) L'identité de $\{1, \dots, n\}$ est l'élément neutre.
-) Par bijectivité, chaque $\sigma \in S_n$ admet un inverse.
-) S_n a $n!$ élément : Pour $\sigma(1)$ on a n choix, pour $\sigma(2)$ on a $n - 1$ choix (car $\sigma(1) \neq \sigma(2)$), ..., pour $\sigma(n)$ on a un seul choix.

□

Remarque 5.11.

-) $\sigma \in S_n$ est un *cycle* (de longueur $1 \leq k \leq n$) si et seulement si il existe des éléments distincts $a_1, \dots, a_k \in \{1, \dots, n\}$, tels que :

$$\sigma(a_1) = a_2, \sigma(a_2) = a_3, \dots, \sigma(a_{k-1}) = a_k, \sigma(a_k) = a_1,$$

et $\sigma(b) = b$ pour $b \in \{1, \dots, n\} \setminus \{a_1, \dots, a_k\}$.

-) Toute permutation $\sigma \in S_n$ peut s'écrire comme un produit de cycles disjoints. Par exemple, la permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 3 & 4 \end{pmatrix}$$

qui agit comme suit :

$$\sigma(1) = 2, \sigma(2) = 1, \sigma(3) = 5, \sigma(4) = 3, \sigma(5) = 4,$$

s'écrit comme produit de cycles disjoints :

$$\sigma = (1 \ 2)(3 \ 5 \ 4),$$

où le 2-cycle $(1 \ 2)$ et le 3-cycle $(3 \ 5 \ 4)$ sont disjoints, c'est-à-dire que leurs supports sont disjoints. Le *support* d'une permutation $\tau \in S_n$ est défini par :

$$\text{supp}(\tau) = \{k \in \{1, \dots, n\} \mid \sigma(k) \neq k\}.$$

-) Une *transposition* est un cycle de longueur 2, c'est-à-dire un 2-cycle.
-) Si $\sigma_1, \sigma_2 \in S_n$ et $\text{supp}(\sigma_1) \cap \text{supp}(\sigma_2) = \emptyset$, alors les cycles σ_1 et σ_2 commutent, c'est-à-dire que $\sigma_1\sigma_2 = \sigma_2\sigma_1$.
-) La *signature* est un morphisme de groupes $\varepsilon: S_n \rightarrow \{\pm 1\}$ tel que, pour un k -cycle σ , on a :

$$\varepsilon(\sigma) = (-1)^{k+1}.$$

On a les propriétés suivantes :

$$\varepsilon(\sigma_1 \circ \sigma_2) = \varepsilon(\sigma_1)\varepsilon(\sigma_2),$$

et

$$\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)^{-1} = \varepsilon(\sigma),$$

car ε est un morphisme de groupes avec valeurs dans $\{\pm 1\}$.

-) Le noyau de ε est le groupe alterné, noté A_n . Ce groupe, pour $n \geq 2$, est un groupe d'ordre $n!/2$.

Lemme 5.12. *Toute permutation $\sigma \in S_n$ est un produit de transpositions.*

Démonstration. Si $\sigma \in S_n$ est un k -cycle $\sigma(a_1) = a_2, \dots, \sigma(a_k) = a_1$, alors $\sigma = (a_1a_2)(a_2a_3) \dots (a_{k-1}a_k)$. En général, $\sigma \in S_n$ est un produit de cycle. $\square$

5.4 Une formule explicite pour le déterminant

Proposition 5.13. *Soit $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(K)$. Alors,*

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}.$$

Cette formule n'est pas bien pour calculer $\det(A)$, par exemple pour $n = 5$ on a $120 = 5!$ termes. Mais elle aide pour la théorie des déterminants.

Exemple 5.14. Soit $A = (a_{i,j})_{1 \leq i,j \leq 3} \in \mathcal{M}_3(K)$ et $(e_1, \dots, e_3)$ la base canonique de $\mathbb{R}^3$, alors

$$\det(A) = \det\left(\sum_{i=1}^3 a_{i,1}e_i \mid \sum_{j=1}^3 a_{j,2}e_j \mid \sum_{k=1}^3 a_{k,3}e_k\right) = \sum_{i,j,k=1}^3 a_{i,1}a_{j,2}a_{k,3} \det(e_i|e_j|e_k).$$

On a $\det(e_i|e_j|e_k) = 0$ si (au moins) deux des indices i, j, k sont égaux. Il reste donc dans cette somme seulement les termes correspondants à (i, j, k) pour i, j et k différents et variant de 1 à 3 $\Rightarrow$ correspond à toutes les permutations de $\{1, 2, 3\}$.

On change alors la notation et on pose : $i = \sigma(1), j = \sigma(2), k = \sigma(3)$ pour $\sigma \in S_3$, et on a

$$\det(A) = \sum_{\sigma \in S_3} a_{\sigma(1),1} a_{\sigma(2),2} a_{\sigma(3),3} \det(e_{\sigma(1)} | e_{\sigma(2)} | e_{\sigma(3)}).$$

Par 5.12 et 5.3 on a

$$\det(e_{\sigma(1)} | e_{\sigma(2)} | e_{\sigma(3)}) = \varepsilon(\sigma) \underbrace{\det(e_1 | e_2 | e_3)}_{=1} = \varepsilon(\sigma).$$

Démonstration de 5.13. La même comme pour $n = 3$ (5.14) grâce à la formule :

$$\det\left(\sum_{i_1=1}^n a_{i_1,1} e_{i_1} \mid \dots \mid \sum_{i_n=1}^n a_{i_n,n} e_{i_n}\right) = \sum_{i_1, \dots, i_n=1}^n a_{i_1,1} a_{i_2,2} \cdots a_{i_n,n} \det(e_{i_1} \mid \dots \mid e_{i_n}).$$

En fait, si $i_k = i_j$ pour $i \neq j$ on a $\det(e_{i_1} \mid \dots \mid e_{i_n}) = 0$ (par 5.3). Donc, la dernière somme est

$$\sum_{\sigma \in S_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \underbrace{\det(e_{\sigma(1)} \mid \dots \mid e_{\sigma(n)})}_{=\varepsilon(\sigma)}. \quad \square$$

5.5 Opérations sur les lignes

Théorème 5.15. Soit $A \in \mathcal{M}_n(K)$. On a $\det({}^t A) = \det(A)$.

Démonstration. Soit $A = (a_{i,j})_{1 \leq i,j \leq n}$. Par la formule explicite du déterminant via les permutations (cf. 5.13), on a :

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n},$$

où S_n est l'ensemble des permutations de n éléments, et $\varepsilon(\sigma)$ est la signature de la permutation σ .

L'application $\sigma \mapsto \sigma^{-1}$ de S_n dans S_n est bijective, et on a $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$. Ainsi, en effectuant le changement de variable $\sigma \mapsto \sigma^{-1}$, on obtient :

$$\sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n}.$$

Maintenant, en utilisant le fait que pour toute permutation σ , on a

$$a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n} = \prod_{i=1}^n a_{\sigma^{-1}(i),i} = \prod_{j=1}^n a_{j,\sigma(j)},$$

où l'indice j est lié à l'indice i par $\sigma(j) = i$, on obtient :

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} = \det({}^t A).$$

Ainsi, on a bien $\det({}^t A) = \det(A)$. □

Théorème 5.16. Soit $n \geq 1$.

- 1) L'application $\det: \mathcal{M}_n(K) \rightarrow K$ est linéaire en chaque ligne.
- 2) Une matrice carrée $A \in \mathcal{M}_n(K)$ ayant deux lignes égales vérifie $\det A = 0$.
- 3) Si on échange deux lignes d'une matrice carrée $A \in \mathcal{M}_n(K)$, on multiplie son déterminant par -1 .
- 4) Soit $n \geq 1$ et soit $A \in \mathcal{M}_n(K)$. Soit B la matrice obtenue à partir de A en ajoutant à une ligne de A une combinaison linéaire des autres lignes de A . Alors $\det B = \det A$.
- 5) $\det(A) \neq 0 \Leftrightarrow$ les vecteurs lignes de A sont indépendants.

Démonstration. Grâce à 5.15 on déduit le cas des colonnes. □

Théorème 5.17. Soit $n \geq 1$ et soit $A \in \mathcal{M}_n(K)$.

- 1) Soit $1 \leq i \leq n$. On a alors

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{i,j} \det A_{i,j};$$

on dit qu'on a développé $\det A$ selon la i -ième ligne.

- 2) Soit $1 \leq j \leq n$. On a alors

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{i,j} \det A_{i,j};$$

on dit qu'on a développé $\det A$ selon la j -ième colonne.

Démonstration. Par définition, l'énoncé est vrai pour le développement selon la première ligne (4.2). Par permutation et transposition on déduit ce cas. □

Définition 5.18. Soit $A \in \mathcal{M}_n(K)$. On appelle cofacteurs de A les scalaires définis par

$$\text{cof}_{i,j}(A) := (-1)^{i+j} \det(A_{i,j}),$$

pour $1 \leq i, j \leq n$, où $A_{i,j}$ désigne la sous-matrice obtenue en supprimant la i -ième ligne et la j -ième colonne de A (cf. 4.2).

Exemple 5.19. Soit

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

Nous allons calculer $\det(A)$ de deux manières différentes :

•) Développement selon la troisième ligne :

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

Calculons les mineurs :

$$\det(A) = 1 \cdot ((-2)(0) - (3)(1)) + 1 \cdot ((1)(0) - (3)(2)) + 2 \cdot ((1)(1) - (-2)(2)).$$

Ce qui donne :

$$\det(A) = 1 \cdot (-3) + 1 \cdot (-6) + 2 \cdot (5) = -3 - 6 + 10 = 1.$$

•) Développement selon la troisième colonne :

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

Calculons les mineurs :

$$\det(A) = 3 \cdot ((2)(-1) - (1)(1)) + 2 \cdot ((1)(1) - (-2)(2)).$$

Ce qui donne :

$$\det(A) = 3 \cdot (-3) + 2 \cdot (5) = -9 + 10 = 1.$$

Ainsi, dans les deux cas, nous obtenons $\det(A) = 1$.

Remarque 5.20. Soit $A \in \mathcal{M}_n(K)$. But : calculer $\det(A)$. Méthode :

- (i) Appliquer des opérations sur les lignes ou les colonnes afin d'obtenir une ligne ou une colonne comportant beaucoup de zéros (par exemple, $n - 1$ zéros). Supposons qu'il s'agit de la i -ème colonne.
- (ii) Développer selon une ligne ou une colonne qui contient beaucoup de zéros (par exemple, la i -ème colonne obtenue à l'étape précédente).
- (iii) Ainsi, le calcul de $\det(A)$ se réduit au calcul de déterminants de matrices $B_1, \dots, B_k$ de taille $n - 1$.
- (iv) Répéter ce processus avec les matrices $B_1, \dots$.

Exemple 5.21.

$$\begin{vmatrix} 1 & -2 & 1 & 3 & 4 \\ 1 & -1 & 0 & 2 & 4 \\ 2 & 1 & 3 & 1 & 2 \\ -1 & 0 & 1 & 1 & 3 \\ 0 & 1 & -1 & 1 & 3 \end{vmatrix} \stackrel{\text{ajouté 1.ligne}}{=} \begin{vmatrix} 1 & -2 & 1 & 3 & 4 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 5 & 1 & -5 & -6 \\ 0 & -2 & 2 & 4 & 7 \\ 0 & 1 & -1 & 1 & 3 \end{vmatrix}$$

$$\begin{aligned}
& \text{développé selon 1.colonne} \stackrel{=}{=} 1 \cdot \begin{vmatrix} 1 & -1 & -1 & 0 \\ 5 & 1 & -5 & -6 \\ -2 & 2 & 4 & 7 \\ 1 & -1 & 1 & 3 \end{vmatrix} \stackrel{\text{ajouté 1. colonne}}{=} \begin{vmatrix} 1 & 0 & 0 & 0 \\ 5 & 6 & 0 & -6 \\ -2 & 0 & 2 & 7 \\ 1 & 0 & 2 & 3 \end{vmatrix} \\
& \stackrel{=}{=} \begin{vmatrix} 6 & 0 & -6 \\ 0 & 2 & 7 \\ 0 & 2 & 3 \end{vmatrix} \stackrel{\text{développé selon 1.colonne}}{=} 6 \cdot \begin{vmatrix} 2 & 7 \\ 2 & 3 \end{vmatrix} = 6(6 - 14) = -48.
\end{aligned}$$

Définition 5.22. Soit $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(K)$ une matrice carrée. On dit qu'elle est triangulaire supérieure si $a_{i,j} = 0$ pour $i > j$. Elle est triangulaire inférieure si $a_{i,j} = 0$ pour $i < j$.

Théorème 5.23. Soit $A \in \mathcal{M}_n(K)$ une matrice triangulaire supérieure ou inférieure. Alors

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

En particulier, une matrice triangulaire est inversible si et seulement si tous ses termes diagonaux sont non nuls.

Démonstration. On développe n -fois selon la première ligne ou selon la première colonne 5.17. $\square$

Théorème 5.24. Soit $A \in \mathcal{M}_n(K)$ une matrice triangulaire supérieure ou inférieure. Alors

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

En particulier, une matrice triangulaire est inversible si et seulement si tous ses termes diagonaux sont non nuls.

Démonstration. Soit $A = (a_{ij})_{i,j}, B = (b_{ij})_{i,j}$ et soit $C = AB, C = (c_{ik})_{i,k}$. Donc, $c_{ik} = \sum_{j=1}^n a_{ij}b_{jk}$. Si on note $C_k, k = 1, \dots, n$ les colonnes de C on a :

$$C_k = \sum_{j=1}^n b_{jk}a_j,$$

où $A = (a_1 | \dots | a_n)$. Donc, $\det(AB) = \det(C_1 | \dots | C_n) = \det\left(\sum_{j=1}^n b_{j1}a_j | \dots | \sum_{j=1}^n b_{jn}a_j\right) \stackrel{\text{linéarité}}{=} \sum_{i_1, \dots, i_n=1}^n b_{i_1,1}b_{i_2,2} \dots b_{i_n,n} \det(a_{i_1}, a_{i_2}, \dots, a_{i_n})$. Si $i_j = i_k$ pour $j \neq k$, le terme correspondant est nul, donc $\det(AB) = \sum_{\sigma \in S_n} b_{\sigma(1),1} \dots b_{\sigma(n),n} \det(a_{\sigma(1)} | \dots | a_{\sigma(n)}) = \sum_{\sigma \in S_n} \varepsilon(\sigma) b_{\sigma(1),1} \dots b_{\sigma(n),n} \det(a_1 | \dots | a_n) = \det(B) \det(A)$. $\square$

Corollaire 5.25. Si $A \in \mathcal{M}_n(K)$ est inversible, alors $\det(A^{-1}) = \frac{1}{\det(A)}$.

Démonstration. Puisque A est inversible, on a $AA^{-1} = I_n$, où I_n est la matrice identité de taille n .

En appliquant le déterminant des deux côtés de l'égalité, on obtient

$$\det(AA^{-1}) = \det(I_n).$$

D'une part, par propriété des déterminants, on a

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

D'autre part, $\det(I_n) = 1$ car la matrice identité a un déterminant égal à 1.

Ainsi, on a l'égalité

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

En isolant $\det(A^{-1})$, on obtient

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

□

Corollaire 5.26. Si $A, B \in \mathcal{M}_n(K)$ sont semblables, alors $\det(A) = \det(B)$.

Démonstration. $\exists P \in \mathcal{M}_n(K)$ inversible telle que $A = PBP^{-1}$. Donc, $\det(A) = \det(PBP^{-1}) \stackrel{??}{=} \det(P) \det(B) \det(P^{-1}) \stackrel{5.25}{=} \det(B)$. □

5.6 Questions de compréhension et exercices typiques

- (i) Faire le preuve de 5.3 (ii), et des quelques points de 5.6.
- (ii) Qu'est-ce que le signe $\varepsilon(\sigma)$ d'une permutation $\sigma \in S_n$? Comment peut-on le calculer?
- (iii) Vérifier 5.4 pour $n = 2$ en écrivant les deux côtés explicitement.

Exercice 5.1. (i) Calculer $\varepsilon(\sigma)$ pour $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 2 & 1 & 6 & 4 \end{pmatrix}$, et écrire σ comme un produit de cycle.

(ii) Trouver $\begin{vmatrix} 2 & 0 & 2 & 4 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 2 & 0 & 2 & 5 \end{vmatrix}$ et $\begin{vmatrix} 2 & 0 & 2 & 4 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 0 & 2 & 5 \end{vmatrix}$.

6 Exposé 6 : Déterminants, III

But de l'exposé 6 : On va finalement expliquer le lien entre le volume et le déterminant (énoncé dans l'exposé 4). On va prouver la formule de Cramer qui donne une façon pour trouver l'inverse d'une matrice inversible.

6.1 Volume et déterminants

Remarque 6.1. Soit $\theta \in \mathbb{R}$. On a d'après 3.22 :

$$\det(R(\theta)) = \begin{vmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{vmatrix} = \cos(\theta)^2 + \sin(\theta)^2 = 1.$$

On revient à la discussion du volume de l'exposé 4.

Corollaire 6.2. Soit $K = \mathbb{R}$, $n = 2$, et $v, w \in \mathbb{R}^2$. On définit le volume $\text{Vol}(v, w) := |\det(v|w)|$, qui est la seule fonction $\mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}_{\geq 0}$ vérifiant les propriétés suivantes :

- (i) Si $\theta \in \mathbb{R}$, alors $\text{Vol}(R(\theta)v, R(\theta)w) = \text{Vol}(v, w)$, où $R(\theta)$ désigne une rotation dans $\mathbb{R}^2$. (Invariance par rotation)
- (ii) Si $a, b \in \mathbb{R}$, alors

$$\text{Vol}\left(\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} v, \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} w\right) = |a||b| \text{Vol}(v, w).$$

(Prolongement par des scalaires)

- (iii) $\text{Vol}(v, w) = \text{Vol}(w, v)$. (Invariance par permutation)
- (iv) Soit (e_1, e_2) la base canonique de $\mathbb{R}^2$, alors $\text{Vol}(e_1, e_2) = 1$. (Volume standard)
- (v) Si $\lambda \in \mathbb{R}$, alors $\text{Vol}(v, \lambda v + w) = \text{Vol}(v, w)$. (Invariance par déformation)

On appelle $\text{Vol}(v, w)$ le volume du parallélogramme engendré par v et w .

Démonstration.

- La fonction $\text{Vol}(v, w)$ vérifie les propriétés (i)-(v) :

$$(i) \quad \text{Vol}(R(\theta)v, R(\theta)w) = |\det(R(\theta) \cdot (v|w))| \stackrel{??}{=} |\det(R(\theta)) \cdot \det(v|w)| \stackrel{6.1}{=} \text{Vol}(v, w).$$

- (ii) Comme (i), mais avec $R(\theta)$ remplacé par la matrice

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix},$$

qui a déterminant ab .

(iii) Par 5.3 :

$$\text{Vol}(v, w) = |\det(v|w)| = |-\det(w|v)| = \text{Vol}(w, v).$$

(iv) Si $I_2 \in \mathcal{M}_2(\mathbb{R})$ est la matrice identité, alors $\det(I_2) = 1$.

(v) Par 4.3.

-) Soit $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}_{\geq 0}$ la fonction qui vérifie les propriétés (i), (ii), (iii), (iv), et (v). Montrons que $\Phi(v, w) = \text{Vol}(v, w)$ pour tous $v, w \in \mathbb{R}^2$. Cela est clair si $v = w = 0$. Supposons donc que $v \neq 0$ ou $w \neq 0$. Par permutation, c'est-à-dire (iii), supposons $v \neq 0$. Par 3.23, il existe $\lambda \in \mathbb{R}$ et $\theta \in \mathbb{R}$ tels que $v = \lambda R(\theta)e_1$. Donc,

$$\Phi(v, w) \stackrel{(i),(ii)}{=} |\lambda| \Phi(e_1, R(\theta)^{-1}w),$$

et de même, $\text{Vol}(v, w) = |\lambda| \text{Vol}(e_1, R(\theta)^{-1}w)$. Ainsi, on peut supposer que $v = e_1$. Par (v), on peut réduire au cas où $w = \mu e_2$ pour un $\mu \in \mathbb{R}$. Donc,

$$\Phi(e_1, w) \stackrel{(ii)}{=} |\mu| \Phi(e_1, e_2) \stackrel{(iv)}{=} |\mu| = |\mu| \text{Vol}(e_1, w). \quad \square$$

Remarque 6.3 (sans démonstration). Plus généralement :

$$\text{Vol} : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}, (x_1, \dots, x_n) \mapsto |\det(x_1 | \dots | x_n)|$$

calcule le volume du cube engendré par $x_1, \dots, x_n \in \mathbb{R}^n$.

6.2 Formes multilinéaires alternées

Soit E un espace vectoriel de dimension finie n .

Définition 6.4. (i) Une forme n -linéaire sur E est une application $f : E^n \rightarrow K$ telle que $\forall 1 \leq i \leq n$ et pour tout $(n-1)$ -tuplet $(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n) \in E^{n-1}$, l'application

$$E \rightarrow K, v \mapsto f(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_n)$$

est linéaire.

(ii) Une forme n -linéaire sur E est dite alternée si $\forall (v_1, \dots, v_n) \in E^n$ et $i < j$ on a

$$f(v_1, \dots, \underset{i\text{-ème}}{v_i}, \dots, \underset{j\text{-ème}}{v_j}, \dots, v_n) = -f(v_1, \dots, \underset{i\text{-ème}}{v_j}, \dots, \underset{j\text{-ème}}{v_i}, \dots, v_n).$$

Exemple 6.5. Si $\mathcal{B}$ est une base de E , l'application $\det_{\mathcal{B}}$ est une forme n -linéaire alternée sur E . 5.2

Théorème 6.6. Soit f une forme n -linéaire alternée sur E et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Alors il existe $\alpha \in K$ tel que $f = \alpha \det_{\mathcal{B}}$.

Démonstration. Posons $\alpha = f(e_1, \dots, e_n)$ et $g = f - \alpha \det_{\mathcal{B}}$. Alors, g est une forme n -linéaire alternée (facile à vérifier, car une combinaison linéaire d'applications linéaires est linéaire). Montrons que g est nulle. Soit $(v_1, \dots, v_n) \in E$. Écrivons $v_j = \sum_{i=1}^n a_{ij} e_i$, $a_{ij} \in K$. Alors,

$$g(v_1, \dots, v_n) = \sum_{i_1=1}^n \dots \sum_{i_n=1}^n \prod_{j=1}^n a_{i_j, j} g(e_{i_1}, \dots, e_{i_n}).$$

Deux cas :

-) Si $\exists j < k$ tel que $i_k = i_j$, alors $g(e_{i_1}, \dots, e_{i_n}) = 0$, car g est alternée.
-) Sinon, on a $i_j = \sigma(j)$ pour une permutation σ , et comme g est alternée,

$$g(e_{i_1}, \dots, e_{i_n}) = \varepsilon(\sigma) g(e_1, \dots, e_n) = \varepsilon(\sigma) (f(e_1, \dots, e_n) - \alpha \det_{\mathcal{B}}(e_1, \dots, e_n)) = 0$$

(car $\det_{\mathcal{B}}(e_1, \dots, e_n) = 1$ et $\alpha = f(e_1, \dots, e_n)$). Donc, $g(v_1, \dots, v_n) = 0$ et $g = 0$. $\square$

Corollaire 6.7 (Formule du changement de base). Soient $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (f_1, \dots, f_n)$ deux bases de E . Alors on a, pour tout $(v_1, \dots, v_n) \in E^n$:

$$\det_{\mathcal{B}}(v_1, \dots, v_n) = \det_{\mathcal{B}}(f_1, \dots, f_n) = \det(P_{\mathcal{B} \rightarrow \mathcal{B}'}) \det_{\mathcal{B}'}(v_1, \dots, v_n).$$

Démonstration. On considère la forme n -linéaire alternée $\det_{\mathcal{B}}(-)$. Le théorème 6.6 affirme qu'il existe un $\alpha \in K$ tel que $\det_{\mathcal{B}}(-) = \alpha \det_{\mathcal{B}'}(-)$. En particulier, on a $\alpha = \det_{\mathcal{B}}(f_1, \dots, f_n)$ (car $\det_{\mathcal{B}'}(f_1, \dots, f_n) = 1$), et la matrice $\text{Mat}_{\mathcal{B}}(f_1 | \dots | f_n)$ est la matrice de changement de base $P_{\mathcal{B} \rightarrow \mathcal{B}'}$. $\square$

6.3 Formule de Cramer

Définition 6.8. Soit $A \in \mathcal{M}_n(K)$. La comatrice, ou matrice des cofacteurs, de A est la matrice $\text{Com}(A) \in \mathcal{M}_n(K)$ définie par

$$\text{Com}(A) = \left((-1)^{i+j} \det(A_{i,j}) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}.$$

Théorème 6.9. Soit $A \in \mathcal{M}_n(K)$. On a

$${}^t \text{Com}(A)A = A^t \text{Com}(A) = \det(A)I_n.$$

En particulier, si $A \in \mathcal{M}_n(K)$ est inversible, on a

$$A^{-1} = (\det A)^{-1} {}^t \text{Com}(A).$$

Démonstration. Posons $B = {}^t \text{Com}(A)A = (b_{ij})_{i,j}$. Alors :

$$b_{ij} = \sum_{k=1}^n \text{cof}_{k,i}(A)a_{k,j} = \sum_{k=1}^n (-1)^{k+i} \det(A_{k,i})a_{k,j}.$$

On a $b_{ii} = \det(A)$ (développement selon la i -ème colonne). Si $i \neq j$ on a $b_{ij} = 0$, car :

$$b_{i,j} = \text{cof}_{1,i}(A)a_{1,j} + \dots + \text{cof}_{n,i}(A)a_{n,j} = \begin{vmatrix} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,j} & \dots & a_{1,n} \\ \dots & & & & & & \dots \\ a_{n,1} & \dots & a_{n,j} & \dots & a_{n,j} & \dots & a_{n,n} \end{vmatrix} = 0$$

par développement selon la i -ème colonne. Donc, $b_{i,j} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$ et ${}^t \text{Com}(A)A = \det(A)I_n$. De même, $A^t \text{Com}(A) = \det(A)I_n$ par un calcul analogue. $\square$

Exemple 6.10. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(K)$. Alors

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

6.4 Utilisation d'un déterminant pour évaluer le rang

Soit $p, n \geq 1$ et $A \in \mathcal{M}_{p,n}(K)$. Soit $r \leq \min(p, n)$. On appelle *mineur d'ordre r* de A un déterminant d'une sous-matrice carrée de A obtenue en sélectionnant r lignes et r colonnes de A .

Exemple 6.11. La matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ possède trois mineurs d'ordre 2. Il s'agit de

$$\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = -3, \quad \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = -6, \quad \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = -3.$$

Théorème 6.12. Soit $A \in \mathcal{M}_{p,n}(K)$ et soit $r = \min(p, n)$. Alors A est de rang r si et seulement si A possède au moins un mineur d'ordre r non nul.

Démonstration.

-) Cas où $p \leq n$. Écrivons $A = (C_1 | \dots | C_n)$, où $C_i \in K^p$, pour tout $i = 1, \dots, n$.
On a

$\text{rg} A = p \Leftrightarrow$ la famille $(C_1, \dots, C_n)$ est une famille génératrice de K^p ,
 $\Leftrightarrow$ on peut choisir une base parmi les $C_1, \dots, C_n$ qui est une base de K^p ,
 $\Leftrightarrow \exists i_1 < \dots < i_p \in \{1, \dots, n\}$ tel que $(C_{i_1}, \dots, C_{i_p})$ est une base de K^p ,
 $\Leftrightarrow \exists$ un mineur d'ordre p non nul de A .

-) Si $p \geq n$, on utilise l'égalité $\text{rg} A = \text{rg}^t A$.

□

6.5 Questions de compréhension et exercices typiques

- Expliquer pourquoi la formule de Cramer n'est pas très utile pour trouver l'inverse d'une matrice inversible A . Quelle méthode plus adaptée ?
- Écrire les formules pour le développement d'une matrice par rapport à une ligne et par rapport à une colonne.

Exercice 6.1. Calculer le volume du parallélogramme engendré par $\begin{pmatrix} 5 \\ 7 \end{pmatrix}$ et $\begin{pmatrix} 4 & 3 \end{pmatrix}$.

7 Exposé 7 : Réduction des endomorphismes, I

But de l'exposé 7 : Nous commençons avec la théorie de réduction des endomorphismes. Le but de la réduction est de transformer une matrice dans une forme particulièrement simple, par exemple une matrice diagonale. Nous donnons une application aux systèmes d'équations différentielles.

Soit $K = \mathbb{R}$ ou $K = \mathbb{C}$.

7.1 Motivation pour la réduction des endomorphismes

Soit $n \geq 1$, $A \in \mathcal{M}_n(K)$ et $y = \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix} \in K^n$. En général, on s'intéresse aux solutions d'équation

$$\{x \in K^n \mid y = Ax\},$$

et dans le cours d'algèbre linéaire I on a appris des méthodes puissantes (algorithme de Gauss,...) pour trouver toutes solutions (et en particulier de décider si il y a une...).

Un cas est particulièrement simple.

Définition 7.1. $A \in \mathcal{M}_n(K)$ est diagonalisable si il existe $P \in \text{GL}_n(K)$ tq $P^{-1}AP = D = \text{Diag}(\lambda_1, \dots, \lambda_n)$ est une matrice diagonale des termes $\lambda_1, \dots, \lambda_n \in K$.

Dans ce cas, $y = Ax$ ssi $y' = Dx'$ où $y' := Py$, $x' := Px$. L'ensemble

$$\{x' \in K^n \mid y' = Dx'\}$$

est facile de trouver : si $\lambda_i \neq 0$ pour $i = 1, \dots, n$, alors on a

$$x' = \begin{pmatrix} x'_1 \\ \dots \\ x'_n \end{pmatrix} = \mu \begin{pmatrix} \frac{y'_1}{\lambda_1} \\ \dots \\ \frac{y'_n}{\lambda_n} \end{pmatrix}$$

pour un $\mu \in K^n$, et donc on trouve $x = P^{-1}x' = P^{-1}D^{-1}y' = P^{-1}D^{-1}Py$ aussi.

Un autre cas est aussi accessible.

Définition 7.2. $A \in \mathcal{M}_n(K)$ est triangulable si il existe $P \in \text{GL}_n(K)$ tq $P^{-1}AP = T$ est une matrice triangulaire supérieure.

En effet, si $A = (a_{i,j})_{i,j}$ est triangulaire supérieure et $y = Ax$, alors $x_1 = a_{1,1}^{-1}y_1$ (si $a_{1,1}$ est inversible), $x_2 = \frac{1}{a_{2,2}}(y_2 - a_{2,1}x_1)$ (si $a_{2,2}$ est inversible), $x_3 = \dots$. Donc, on peut par recursion trouver x à partir de y (si A est inversible).

Réduction des endomorphismes : Etant donné $A \in \mathcal{M}_n(K)$, on cherche $P \in \text{GL}_n(K)$ tq la matrice $B = P^{-1}AP$, qui est semblable à A (par définition), a une forme particulièrement simple.

Nous allons prouver grâce à la réduction des endomorphismes.

Théorème 7.3. Soit $K = \mathbb{R}$ et $A \in \mathcal{M}_n(\mathbb{R})$ symétrique. Alors, A est diagonalisable.

Démonstration. 10.2. □

Théorème 7.4. Soit $K = \mathbb{C}$ et $A \in \mathcal{M}_n(\mathbb{C})$. Alors, A est trigonalisable.

Démonstration. 9.17. □

Remarque 7.5. Si une matrice A est diagonalisable, on peut calculer A^i , $i \geq 0$ de la manière suivante : déterminer $P \in \text{GL}_n(K)$ tq $P^{-1}AP = D = \text{Diag}(\lambda_1, \dots, \lambda_n)$ soit diagonale. Alors,

$$A^i = (PDP^{-1})^i = \underbrace{PDP^{-1}PDP^{-1} \dots PDP^{-1}}_{i\text{-facteurs}} = PD^iP^{-1}$$

et $D^i = \text{Diag}(\lambda_1^i, \dots, \lambda_n^i)$ est facile à déterminer.

C'est montrer que la réduction des endomorphismes soit utile pour des autres questions aussi. Dans 7.2 nous donnons une autre application très importante.

7.2 Une application aux systèmes d'équations différentielles linéaires

Soit $I \subseteq \mathbb{R}$ une intervalle, $0 \in I$. Si $x: I \rightarrow \mathbb{C}$ est une fonction $\mathcal{C}^1$, i.e., dérivable, on pose $x' := \frac{dx}{dt}$.

Remarque 7.6. Soit $\lambda \in \mathbb{C}$. On a $x' = \lambda x$ si et seulement si $\exists c \in \mathbb{C}$ tq $x = ce^{\lambda t}$. En effet, $(ce^{\lambda t})' = c\lambda e^{\lambda t}$. D'autre côté, si $x' = \lambda x$, considérons $y = x/e^{\lambda t}$. Alors,

$$y' = \frac{x'e^{\lambda t} - \lambda x e^{\lambda t}}{e^{2\lambda t}} = 0$$

car $x' = \lambda x$. Donc, $y: I \rightarrow \mathbb{C}$ est une fonction constante, d'une valeur $c \in \mathbb{C}$. Alors, $x = ce^{\lambda t}$.

Considérons un système d'équations différentielles de la forme :

$$\begin{cases} x_1' = a_{1,1}x_1 + \dots + a_{1,n}x_n, \\ \dots \\ x_n' = a_{n,1}x_1 + \dots + a_{n,n}x_n, \end{cases}$$

où $x_i: I \rightarrow \mathbb{C}$ sont des fonctions $\mathcal{C}^1$, et $a_{i,j} \in \mathbb{C}$. Posons $A = (a_{i,j})_{i,j}$, et

$$X(t) := \begin{pmatrix} x_1(t) \\ \dots \\ x_n(t) \end{pmatrix}, \quad \text{et } X'(t) = \frac{dX(t)}{dt} = \begin{pmatrix} x'_1(t) \\ \dots \\ x'_n(t) \end{pmatrix}$$

(alors, $X: I \rightarrow \mathbb{C}^n$, $t \mapsto X(t)$ est une fonction $\mathcal{C}^1$).

On a $X'(t) = AX(t)$ par construction. Supposons A diagonalisable. Alors, $\exists P \in \text{GL}_n(K)$ tq $A = PDP^{-1}$ où $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$ une matrice diagonale. Posons $Y(t) := P^{-1}X(t)$. Alors, $Y'(t) = P^{-1}X'(t)$ car les coefficients $a_{i,j}$ sont constantes. On a

$$Y'(t) = P^{-1}X'(t) = P^{-1}AX(t) = \underbrace{P^{-1}AP}_{=D} \underbrace{P^{-1}X(t)}_{=Y(t)} = DY(t).$$

Ecrivons $Y(t) = \begin{pmatrix} y_1(t) \\ \dots \\ y_n(t) \end{pmatrix}$. Alors, $Y'(t) = DY(t)$ donne le système d'équations différentielles :

$$\begin{cases} y'_1(t) = \lambda_1 y_1(t), \\ \dots \\ y'_n(t) = \lambda_n y_n(t). \end{cases}$$

Grâce à 7.6 c'est facile à résoudre : il existe $c_1, \dots, c_n \in \mathbb{C}$, en effet $c_i = y_i(0)$, tq $y_i(t) = c_i e^{\lambda_i t}$ pour $i = 1, \dots, n$. Car $X(t) = PY(t)$, on trouve

$$x_i(t) = \sum_{j=1}^n p_{ij} c_j e^{\lambda_j t}$$

où $P = (p_{ij})_{i,j}$.

Résumé :

Théorème 7.7. Soit $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(\mathbb{C})$. On suppose que A est diagonalisable. Soit $\mathcal{B}$ une base de diagonalisation de A et soit P la matrice de passage de la base canonique à la base $\mathcal{B}$. Alors une famille $(x_1, \dots, x_n)$ de fonctions dérivables de I dans $\mathbb{C}$ vérifie le système d'équation différentielles

$$\begin{cases} x'_1 &= a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n \\ \vdots & \vdots \\ x'_n &= a_{n,1}x_1 + a_{n,2}x_2 + \dots + a_{n,n}x_n \end{cases}$$

si et seulement si il existe des constantes $c_1, \dots, c_n$ de $\mathbb{C}$ telles que

$$\forall t \in \mathbb{R}, \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = P \begin{pmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix},$$

où

$$A = P \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} P^{-1}.$$

Démonstration. On suit la discussion ci-dessus. $\square$

Remarque 7.8. Grâce à la décomposition $d + n$ (10.4) on peut aussi résoudre un système $X'(t) = AX(t)$ si A n'est pas diagonalisable (10.5).

Remarque 7.9. Soit $d \geq 1$ et soient $a_i \in \mathbb{C}$, $i = 0, \dots, d-1$. Alors, un fonction $f: I \rightarrow \mathbb{C}$, qui est d fois dérivable (avec derivations $f^0 := f$, $f^{(i)} := (f^{(i-1)})'$, $i = 1, \dots, d$), satisfait

$$f^{(d)} + \sum_{i=0}^{d-1} a_i f^{(i)} = 0$$

ssi $X(t) := \begin{pmatrix} f^{(0)} \\ \vdots \\ f^{(d-1)} \end{pmatrix}$ satisfait

$$X'(t) = AX(t)$$

pour

$$A := \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & & \cdots & 0 \\ \cdots & & \cdots & & \cdots \\ 0 & 0 & & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{d-2} & -a_{d-1} \end{pmatrix}$$

En effet, si $X(t) = {}^t(x_1(t), \dots, x_n(t))$ cet système d'équations différentielles est donné par :

$$\begin{cases} x'_1 = x_2 \\ x'_2 = x_3 \\ \cdots \\ x'_{d-1} = x_d \\ x'_d = -a_0 x_1 - a_1 x_2 - \cdots - a_{d-1} x_d \end{cases}$$

et donc $x_i = f^{(i-1)}$ pour $f = x_1$.

Exemple 7.10. Cherchons les solutions d'équation différentielles $f'' - f = 0$ (donc, $d = 2$ et $a_0 = -1, a_1 = 0$). Par 7.9, c'est équivalent à trouver les solutions de

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{A:=} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}.$$

Posons $P := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in \text{GL}_2(\mathbb{R})$ (car $\det(P) = -2 \neq 0$). Alors, $P^{-1} = \frac{1}{-2} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix}$ (par 6.10), et

$$P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Par 7.7 chaque solution $\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$ est une combinaison linéaire de $P \begin{pmatrix} e^t \\ 0 \end{pmatrix} = \begin{pmatrix} e^t \\ e^t \end{pmatrix}$ et $P \begin{pmatrix} 0 \\ e^{-t} \end{pmatrix} = \begin{pmatrix} e^{-t} \\ -e^{-t} \end{pmatrix}$. Les premières lignes donne alors que $f(t) = c_1 e^t + c_2 e^{-t}$ pour $c_1, c_2 \in \mathbb{C}$.

7.3 Bases adaptées

Soit E un K -espace vectoriel de dimension finie n .

Définition 7.11. Soient $F_1, \dots, F_r \subseteq E$ des sous-espaces vectoriels tq $E = F_1 \oplus \dots \oplus F_r$ est une somme direct. Une base $\mathcal{B}$ de E est dite adaptée à la décomposition si elle est de la forme $\mathcal{B} = \mathcal{B}_1 \amalg \dots \amalg \mathcal{B}_r$ pour des bases $\mathcal{B}_i$ de F_i , $i = 1, \dots, r$.

Lemme 7.12. Soit $f \in \mathcal{L}(E)$ et $\mathcal{B} = \mathcal{B}_1 \amalg \dots \amalg \mathcal{B}_r$ une base adaptée à une décomposition $E = F_1 \oplus \dots \oplus F_r$ de E .

(i) Équivalent :

- a) Chaque F_i est stable par f , i.e., $f(F_i) \subseteq F_i$, $i = 1, \dots, r$.
- b) La matrice $A := \text{Mat}_{\mathcal{B}}(f)$ est une matrice diagonale par blocs $\text{Diag}(A_1, \dots, A_r)$ où $A_i \in \text{Mat}_{\dim F_i}(K)$.

En effet, $A_i = \text{Mat}_{\mathcal{B}_i}(f|_{F_i})$ dans ce cas.

(ii) Équivalent :

- a) Pour $i = 1, \dots, r$ le sous-espace $F_1 \oplus \dots \oplus F_i$ est stable par f .
- b) La matrice $A := \text{Mat}_{\mathcal{B}}(f)$ est une matrice triangulaire supérieur par blocs avec blocs $A_1, \dots, A_r$ sur la diagonale avec $A_i \in \text{Mat}_{\dim F_i}(K)$.

Démonstration. Pour simplicité supposons $r = 2$ (la preuve en général est similaire). Soit $d := \dim F_1$. Donc,

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$$

avec $A_{1,1} \in \text{Mat}_{d,d}(K)$, $A_{1,2} \in \text{Mat}_{d,n-d}(K)$, $A_{2,1} \in \text{Mat}_{n-d,d}(K)$, $A_{2,2} \in \text{Mat}_{n-d,n-d}(K)$. Maintenant, $A_{i,j}$ est la matrice représentative dans les bases $\mathcal{B}_j$ et $\mathcal{B}_i$ de l'application obtenue comme composition

$$f_{i,j}: F_j \rightarrow E \xrightarrow{f} E \xrightarrow{\pi_i} F_i$$

avec $\pi_i: E = F_1 \oplus F_2 \rightarrow F_i$ la projection sur F_i . Donc, $f(F_i) \subseteq F_i$ ssi $f_{2,1} = 0$ ssi $A_{2,1} = 0$, et similaire $f(F_2) \subseteq F_2$ ssi $A_{1,2} = 0$. $\square$

Définition 7.13. Soit $A \in \mathcal{M}_n(K)$. On note

$$u_A: K^n \rightarrow K^n, v \mapsto Av.$$

On a $u_A \in \mathcal{L}(K^n)$ (et chaque $f \in \mathcal{L}(K^n)$ est $f = u_A$ pour $A = \text{Mat}_{\mathcal{B}}(f)$ si $\mathcal{B}$ est la base canonique de K^n).

Corollaire 7.14. Une matrice $A \in \text{Mat}_n(K)$ est diagonalisable ssi il existe une décomposition $K^n = F_1 \oplus \dots \oplus F_n$ de K^n par de sous-espaces $F_i \subseteq K^n$ de dimension 1 qui sont stable par u_A .

Démonstration. Soit $P = (C_1 | \dots | C_n) \in \text{GL}_n(K)$ et $(e_1, \dots, e_n) \in K^n$ la base canonique de K^n . Si $P^{-1}AP = \text{Diag}(\lambda_1, \dots, \lambda_n)$ est diagonal, alors $P^{-1}APe_i = \lambda_i e_i$, i.e., $APe_i = \lambda_i Pe_i$. Donc, $F_i := \text{Vect}(\underbrace{Pe_i}_{=C_i})$, $i = 1, \dots, n$, est stable par u_A .

Car $C_1, \dots, C_n$ est une famille libre, $E = F_1 \oplus \dots \oplus F_n$. Si $E = F_1 \oplus \dots \oplus F_n$ est une décomposition dans de sous-espaces $F_i = \text{Vect}(C_i)$, $C_i \in K^n$, $i = 1, \dots, n$, de dimension 1, qui sont stable par u_A . Donc $P := (C_1 | \dots | C_n)$ est inversible, $\mathcal{B} = (C_1, \dots, C_n)$ est une base adaptée à la décomposition $E = F_1 \oplus \dots \oplus F_n$, et $P^{-1}AP = \text{Mat}_{\mathcal{B}}(u_A)$ est diagonalisable par 7.12. $\square$

Corollaire 7.15. Une matrice $A \in \text{Mat}_n(K)$ est trigonalisable ssi il existe une décomposition $K^n = F_1 \oplus \dots \oplus F_n$ de K^n par de sous-espaces $F_i \subseteq K^n$ de dimension 1, tq pour $i = 1, \dots, n$ on a $u_A(F_1 \oplus \dots \oplus F_i) \subseteq F_1 \oplus \dots \oplus F_i$.

Démonstration. Ce se déduit comme 7.14 de 7.12. $\square$

7.4 Questions de compréhension et exercices typiques

- (i) Soit $A \in \mathcal{M}_n(K)$ une matrice triangulaire supérieure. Comment on trouve les solutions d'équation $y = Ax$ pour un $y \in K^n$?
- (ii) Qu'est-ce qu'une équation différentielle ?

Exercice 7.1. (i) Prouver que la matrice $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$ est diagonalisable avec valeurs propres $\{i, -i\}$.

- (ii) Trouver les solutions d'équations différentielles $f'' = -f$.

8 Exposé 8 : Réduction des endomorphismes, II

But de l'exposé 8 : Soit $f \in \mathcal{L}(E)$. Après 7 on s'intéresse à des sous-espaces $F \subseteq E$ qui sont stables par f . Dans cet exposé on se concentre sur des sous-espaces stables de dimension 1, qui sont engendrés par un vecteur propre. De plus, on va introduire le polynôme caractéristique d'une matrice dont les zéros sont exactement les valeurs propres de la matrice.

8.1 Sous-espaces propres et valeurs propres

On fixe E un K -espace vectoriel ($K = \mathbb{R}$ ou $K = \mathbb{C}$). Soit $f \in \mathcal{L}(E)$.

Définition 8.1. Un vecteur propre de f est un vecteur $v \in E$ vérifiant les deux propriétés suivantes :

-) v est non nul;
-) il existe $\lambda \in K$ tel que $f(v) = \lambda v$.

Remarque 8.2. Si v est un vecteur propre de f , le scalaire λ tel que $f(v) = \lambda v$ est appelé *valeur propre* de f correspondant à v . Elle est uniquement déterminée par v , car $v \neq 0$. L'ensemble de toutes les valeurs propres de f est noté $\text{Sp } f$, il s'agit du *spectre* de f .

Remarque 8.3. Soit $v \in E$. Alors, v est un vecteur propre de f ssi $\text{Vect}(v)$ est un sous-espace stable par f , qui est de dimension 1.

Exemple 8.4. $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y \\ x \end{pmatrix}$ a le vecteur propre $v := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ avec valeur propre 1 car $f(v) = v = 1 \cdot v$. Aussi, $w := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ est un *eigenvector* car $f(w) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -1 \cdot w$. Par contre, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ n'est pas un vecteur propre car $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et son image $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ par f sont indépendants.

Lemme 8.5. Soit $\lambda \in K$. On a $\lambda \in \text{Sp}(f) \Leftrightarrow \text{Ker}(f - \lambda \text{Id}_E) \neq \{0\}$. En particulier, $0 \in \text{Sp}(f) \Leftrightarrow f$ n'est pas injective.

Démonstration. On a $\text{Ker}(f - \lambda \text{Id}_E) = \{v \in E \mid f(v) = \lambda v\}$. □

Exemple 8.6. (i) $n \geq 1$, $f: \mathbb{R}[X]_n := \{P \in \mathbb{R}[X] \mid \deg(P) \leq n\} \rightarrow \mathbb{R}[X]_n$, $P \mapsto P'$. Si $P \neq 0$ et $P' = f(P) = \lambda P$ pour $\lambda \in \mathbb{R}$, alors $\lambda = 0$ ou $\deg(P') = \deg(P)$. Donc, $\lambda = 0$ seule possibilité, et nécessairement $P' = 0$. Donc, les vecteurs propres de f sont les polynômes constants *non nuls*.

- (ii) $K = \mathbb{C}$, $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$. Si $\lambda \in \mathbb{C}$, posons, $g_\lambda(t) = e^{\lambda t}$. Alors, $g_\lambda \in E$. Soit $f \in \mathcal{L}(E)$ définie par $f(g) = g'$, $g \in E$. Alors on a $f(g_\lambda) = (e^{\lambda t})' = \lambda e^{\lambda t} = \lambda g_\lambda(t)$ est $g_\lambda(t)$ est un vecteur propre de f de valeur propre λ .

Définition 8.7. Si $f \in \mathcal{L}(E)$ et si $\lambda \in K$, on note

$$E_\lambda(f) := \text{Ker}(f - \lambda \text{Id}_E) = \{v \in E \mid f(v) = \lambda v\}$$

est appelé sous-espace propre de f associé à λ .

On a alors que λ est une valeur propre de f ssi $E_\lambda(f) \neq 0$. L'ensemble $E_\lambda(f)$ est un sous-espace vectoriel de E (elle est le noyau d'une application linéaire).

Remarque 8.8. Pour $\lambda \in K$ le sous-espace $E_\lambda(f) \subseteq E$ est stable par f et si $\dim E < \infty$, $\mathcal{C}$ est une base de $E_\lambda(f)$ et $g := f|_{E_\lambda(f)} \in \mathcal{L}(E_\lambda(f))$, alors $\text{Mat}_{\mathcal{C}}(g) = \lambda I_d$ où $d = \dim(E_\lambda(f))$ ($d = 0$ est possible!).

Lemme 8.9. Si $\lambda, \mu \in K$, $\lambda \neq \mu$, alors $E_\lambda(f) \cap E_\mu(f) = \{0\}$.

Démonstration. Si $v \in E_\lambda(f) \cap E_\mu(f)$, on a

$$0 = f(v) - f(v) = \lambda v - \mu v = (\lambda - \mu)v,$$

donc $v = 0$ car $\lambda - \mu \neq 0$. □

Exemple 8.10. Si $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y \\ x \end{pmatrix}$, alors $\mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \subseteq E_1(f)$ et $\mathbb{R} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \subseteq E_{-1}(f)$. Note que $\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2$, donc $v := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, w := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ sont libre. Donc, $\mathbb{R}^2 = \mathbb{R}v \oplus \mathbb{R}w$ et donc, $E_1(f) = \mathbb{R}v$, $E_{-1}(f) = \mathbb{R}w$, et f a le spectre $\{1, -1\} \subseteq \mathbb{R}$.

Définition 8.11. Soit $A \in \mathcal{M}_n(K)$.

- (i) Un vecteur propre de A est un élément de K^n qui est vecteur propre de u_A .
- (ii) De même une valeur propre de A est une valeur propre de u_A .
- (iii) Si λ est une valeur propre de A , on note $E_\lambda(A)$ le sous-espace $E_\lambda(u_A)$ de K^n (donc $E_\lambda(A) = \text{Ker}(A - \lambda \cdot I_n)$).
- (iv) $\text{Sp}(A) := \{\lambda \in K \mid E_\lambda(A) \neq 0\}$.

Lemme 8.12. Soit $A \in \mathcal{M}_n(K)$ et $\lambda \in K$. Alors,

$$\lambda \in \text{Sp}(A) \Leftrightarrow \text{Ker}(A - \lambda I_n) \neq 0 \Leftrightarrow \det(A - \lambda I_n) = 0.$$

De plus, si $P \in \text{GL}_n(K)$, on a $\text{Sp}(P^{-1}AP) = \text{Sp}(A)$ et pour $\lambda \in K^n$ on a une égalité de sous-espaces vectoriels de K^n

$$E_\lambda(PAP^{-1}) = PE_\lambda(A).$$

En particulier $\dim E_\lambda(PAP^{-1}) = \dim E_\lambda(A)$.

Démonstration. On a $\text{Ker}(A - \lambda I_n) \neq 0 \Leftrightarrow (A - \lambda I_n)$ n'est pas inversible $\Leftrightarrow \det(A - \lambda I_n) = 0$ (par 5.5).

Si $P \in \text{GL}_n(K)$ et $\lambda \in K$, on a $x \in E_\lambda(A) \Leftrightarrow AX = \lambda X \Leftrightarrow P^{-1}AX = \lambda P^{-1}X \Leftrightarrow P^{-1}APP^{-1}X = \lambda P^{-1}X \Leftrightarrow P^{-1}X \in E_\lambda(P^{-1}AP)$. Donc, $E_\lambda(A) = P.E_\lambda(P^{-1}AP)$ et car P est inversible, ce deux sous-espace ont la même dimension ($\Rightarrow \text{Sp}(A) = \text{Sp}(P^{-1}AP)$). $\square$

Exemple 8.13. On cherche les valeurs propres $\lambda \in K$ de la matrice

$$A := \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}.$$

On a λ valeur propre de $A \Leftrightarrow 0 = \det(A - \lambda I_n) = \begin{vmatrix} 1 - \lambda & 2 \\ -1 & 4 - \lambda \end{vmatrix} = \begin{vmatrix} 0 & -2 + \lambda \\ 3 - \lambda & 4 - \lambda \end{vmatrix} = (3 - \lambda)(-2 + \lambda) \Leftrightarrow \lambda \in \{2, 3\}$. Donc, $\text{Sp}(A) = \{2, 3\}$. Cherchons alors un vecteur propre pour la valeur propre 2, i.e., nous devons trouver $E_2(A) = \text{Ker}(A - 2I_n) = \text{Ker}\left(\begin{pmatrix} -1 & 2 \\ -1 & 2 \end{pmatrix}\right) = \text{Vect}\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right)$. De même, $E_3(A) = \text{Ker}(A - 3I_n) = \text{Ker}\left(\begin{pmatrix} -2 & 2 \\ -1 & 1 \end{pmatrix}\right) = \text{Vect}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$.

Exemple 8.14. Soit $D = \text{Diag}(d_1, \dots, d_n) \in \text{Mat}_n(K)$ diagonale. On a pour $\lambda \in K$:

$$\lambda \in \text{Sp}(D) \Leftrightarrow \text{Ker}(D - \lambda I_n) \neq 0 \Leftrightarrow 0 = \det(D - \lambda I_n) = \prod_{i=1}^n (d_i - \lambda) \Leftrightarrow \lambda \in \{d_1, \dots, d_n\}.$$

Donc, $\text{Sp}(D) = \{d_1, \dots, d_n\}$. Si A est diagonalisable, et $P \in \text{GL}_n(K)$ tq $D := P^{-1}AP$ diagonale, on trouve $\text{Sp}(A) = \text{Sp}(D)$.

Exemple 8.15. Soit $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, i.e., A est la matrice de $f = u_A$ de 8.10 dans la base canonique. Soient $v := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $w := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ les vecteur propres de A (avec valeur propre 1 et -1). Donc, dans la base $\mathcal{C} := (v, w)$ de $\mathbb{R}^2$ on a $\text{Mat}_{\mathcal{C}}(u_A) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Exemple 8.16. Soit $n \geq 1$ et soit $A \in \mathcal{M}_n(K)$. Si A est une matrice triangulaire supérieure

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

alors

$$\operatorname{Sp} A = \{a_{i,i}, 1 \leq i \leq n\} = \{a_{1,1}, a_{2,2}, \dots, a_{n,n}\}.$$

En effet, l'argument de 8.14 marche ici aussi.

Exemple 8.17. Une matrice triangulaire supérieure peut être diagonalisable ou non.

- (i) La matrice $A = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$ est diagonalisable. En effet, par 8.16 $\operatorname{Sp}(A) = \{1, 2\} = \dim K^2$ avec vecteurs propres $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ (on voit la diagonalisabilité aussi facilement grâce à 8.20).
- (ii) Étudions $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Donc, $\operatorname{Sp}(A) = \{1\}$. Mais, $\operatorname{Ker}(A - I_n) = \operatorname{Ker}\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right) = \operatorname{Vect}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$ et $\dim E_1(A) = 1 < 2$ et A n'est pas diagonalisable (7.14) car le seul sous-espace de K^n stable par u_A et de dimension 1 est $E_1(A)$.

Théorème 8.18. Soit $f \in \mathcal{L}(E)$. Soit $(\lambda_1, \dots, \lambda_n)$ une famille de valeurs propres de f . On suppose que $\lambda_i \neq \lambda_j$ pour $i \neq j$. Alors la famille de sous-espaces vectoriels $(E_{\lambda_i}(f))_{1 \leq i \leq n}$ est en somme directe. Autrement dit

$$\sum_{i=1}^n E_{\lambda_i}(f) = \bigoplus_{i=1}^n E_{\lambda_i}(f), \text{ i.e.,}$$

$$E_{\lambda_1}(f) + \dots + E_{\lambda_n}(f) = E_{\lambda_1}(f) \oplus \dots \oplus E_{\lambda_n}(f).$$

Démonstration. Posons $E_{\lambda_i} := E_{\lambda_i}(f)$, $\forall i = 1, \dots, n$. Montrons $\forall i = 2, \dots, n$, $(E_{\lambda_1} + \dots + E_{\lambda_{i-1}}) \cap E_i = \{0\}$ (A.13). Récurrence sur n .

-) $n = 1$: ✓
-) $n \geq 1$: Supposons $0 \neq v \in (E_{\lambda_1} + \dots + E_{\lambda_i}) \cap E_{\lambda_i}$. Donc, $v = v_1 + \dots + v_{i-1}$ tq $v_j \in E_{\lambda_j}$ pour $j = 1, \dots, i-1$. Alors, $E' = \operatorname{Vect}(v_1, \dots, v_{i-1})$ est un sous-espace de E de dimension finie. Car $f(v_j) = \lambda_j v_j$ le sous-espace E' est stable par f . Soit $g := f|_{E'}$. Par hypothèse de récurrence, $E' = \operatorname{Vect}(v_1) \oplus \dots \oplus \operatorname{Vect}(v_{i-1})$. Par 7.14 et 8.14 g est diagonalisable sur E' avec $\operatorname{Sp}(g) \subseteq \{\lambda_1, \dots, \lambda_{i-1}\}$ (notons que $v_j = 0$ est possible, donc pas forcément $\operatorname{Sp}(g) = \{\lambda_1, \dots, \lambda_{i-1}\}$!). Donc, $\lambda_i \notin \operatorname{Sp}(g)$ et alors

$$0 = \operatorname{Ker}(g - \lambda_i \operatorname{Id}_{E'}) = E_i \cap E'.$$

Mais, $0 \neq v \in E_i \cap E'$ par construction. Contradiction ! □

Exemple 8.19. La famille $g_n(t)$, $n \in \mathbb{N}$, de point (4) de 8.4 est libre dans $E := \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ car chaque sous-ensemble fini de $\{g_n(t) = e^{nt} \mid n \in \mathbb{N}\}$ est donné par des vecteurs propre pour des valeurs distinctes pour l'application linéaire $E \rightarrow E$, $g \mapsto g'$.

Corollaire 8.20. Soit $n := \dim E < \infty$, $f \in \mathcal{L}(E)$.

(i) $\text{Card}(\text{Sp}(f)) \leq n$.

(ii) Si $\text{Card}(\text{Sp}(f)) = n$, alors f est diagonalisable.

Démonstration. Par 8.18

$$\text{Card}(\text{Sp}(f)) \leq \sum_{\lambda \in \text{Sp}(f)} \dim E_{\lambda}(f) \leq n$$

et l'égalité $n = \text{Card}(\text{Sp}(f))$ force $E = \bigoplus_{\lambda \in \text{Sp}(f)} E_{\lambda}(f)$, et donc f serait diagonalisable par 7.14. $\square$

8.2 Le polynôme caractéristique

Soit $A \in \mathcal{M}_n(K)$.

Définition 8.21. La polynôme caractéristique de A est la fonction $\chi_A: K \rightarrow K$, $x \mapsto \det(xI_n - A)$.

Exemple 8.22. Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, on a $\chi_A(x) = \begin{vmatrix} x-a & -b \\ -c & x-d \end{vmatrix} = (x-a)(x-d) - bc = x^2 - (a+d)x + ad - bc$.

Remarque 8.23. Rappel :

- (i) Un polynôme à coefficients dans K est une fonction $P: K \rightarrow K$ tq exists $a_0, \dots, a_n \in K$ tq $P(x) = a_n x^n + \dots + a_1 x + a_0$ pour $x \in K$.
- (ii) On note $K[X]$ l'ensemble des polynômes à coefficients dans K .
- (iii) $(K[X], +, \cdot)$ est un anneau commutative, c'est aussi un K -espaces vectoriel. Pour $P(X) := a_0 + a_1 X + \dots + a_n X^n + \dots$, $Q(X) = b_0 + b_1 X + \dots + b_m X^m + \dots$ (alors, $b_i = 0$, $a_i = 0$ pour $i \gg 0$), on a

$$(P + Q)(X) := (a_0 + b_0) + (a_1 + b_1)X + \dots + (a_i + b_i)X^i$$

et

$$(PQ)(X) := a_0 b_0 + (a_0 b_1 + a_1 b_0)X + \dots + \left(\sum_{i=0}^n a_i b_{n-i} \right) X^n + \dots$$

(iv) On note $X^0 := 1$.

(v) Le degré de P est $\max\{k \in \mathbb{N} \mid a_k \neq 0\}$ (par convention : $\deg(0) = -\infty$).

(vi) les polynômes de degré zéro sont les polynômes constants non-nuls.

- (vii) Si $P, Q \in K[X]$, on a $\deg(P + Q) \leq \max\{\deg P, \deg Q\}$ et $\deg(PQ) = \deg(P) + \deg(Q)$.
- (viii) Si $d := \deg(P)$ et $a_d = 1$, le polynôme P est appelé unitaire.

Théorème 8.24. Soit $A \in \mathcal{M}_n(K)$.

- (i) La fonction χ_A est un polynôme unitaire de degré n à coefficients dans K .
- (ii) Soit $\lambda \in K$. Alors, $\lambda \in \text{Sp}(A)$ si et seulement si $\chi_A(\lambda) = 0$.

Démonstration. •) (ii) : ✓ par 8.12.

-) Soit $x \in K$. Écrivons $A = (a_{ij})_{ij}$ et $xI_n - A = (b_{ij})_{ij}$. Alors,

$$\det(xI_n - A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) b_{\sigma(1),1} \dots b_{\sigma(n),n}.$$

Car $b_{\sigma(i),i} = \begin{cases} x - a_{i,i}, & \sigma(i) = i \\ -a_{i,j}, & \sigma(i) \neq i \end{cases}$, on voit que $Q_\sigma := \varepsilon(\sigma) b_{\sigma(1),1} \dots b_{\sigma(n),n}$ est un polynôme de degré $\text{Card}\{i \in \{1, \dots, n\} \mid \sigma(i) = i\}$. En particulier, entre ces polynômes le seul de degré $\geq n$ est $Q_{\text{Id}} = (x - a_{1,1}) \dots (x - a_{n,n})$ que est unitaire. $\square$

Lemme 8.25. Si $P \in \text{GL}_n(K)$, $A \in \mathcal{M}_n(K)$, on a

$$\chi_A = \chi_{P^{-1}AP}.$$

En particulier, $\text{Sp}(A) = \text{Sp}(P^{-1}AP)$ et les matrices semblables ont le même polynôme caractéristique.

Démonstration. $\det(xI_n - P^{-1}AP) = \det(xP^{-1}P - P^{-1}AP) = \det(P^{-1}(xI_n - A)P) = \det(P^{-1}) \det(xI_n - A) \det(P) = \det(xI_n - A)$ en utilisant ?? $\square$

Remarque 8.26. Deux matrices $A, B \in \mathcal{M}_n(K)$ peuvent avoir le même polynôme caractéristique $\chi_A = \chi_B$ sans être semblables ! Par exemple : $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ont le polynôme caractéristique $x^2 - 1$, mais $\dim E_1(A) = 2 \neq \dim E_1(B) = 1$ (8.17, (ii)).

Lemme 8.27. Soit $f \in \mathcal{L}(E)$ et $n := \dim E$. On définit

$$\chi_f(x) := \chi_A(x) \in K[X],$$

où $A = \text{Mat}_{\mathcal{B}}(f) \in \mathcal{M}_n(K)$ pour une base $\mathcal{B}$ de E . Alors, $\chi_f(x)$ est bien défini et ne dépend pas du choix de la base $\mathcal{B}$.

Démonstration. Si $\mathcal{C}$ est une autre base, et $P := P_{\mathcal{C} \rightarrow \mathcal{B}}$ la matrice de changement de base, alors $\text{Mat}_{\mathcal{C}} = P^{-1} \text{Mat}_{\mathcal{B}} P$, et donc $\chi_{\text{Mat}_{\mathcal{C}}}(x) = \chi_A$ par 8.25. $\square$

Exemple 8.28. Soit $f \in \mathcal{L}(E)$ et $n := \dim E$. Si f est nilpotente, i.e., existe $k \geq 1$ tq $f^k = 0$, alors $\chi_f(x) = x^n$. En effet, après le choix d'une base, on peut supposer $E = K^n$ et $f = u_A$ pour une $A \in \mathcal{M}_n(K)$. De plus, on peut supposer $K = \mathbb{C}$ (ça ne change pas le polynôme caractéristique). Soit $\lambda \in \text{Sp}(A)$ avec un vecteur propre $0 \neq v \in K^n$. Donc,

$$0 = A^k v = A^{k-1}(\lambda v) = \lambda A^{k-1} v = \dots = \lambda^k v,$$

et $\lambda^k = 0$, i.e., $\lambda = 0$. Donc, $\chi_A(x) = x^n$ car $\chi_A(x) \in \mathbb{C}[x]$ est unitaire de degré n avec seule racine 0.

Remarque 8.29. Supposons $A \in \mathcal{M}_n(K)$ est diagonalisable. But : trouver $P \in \text{GL}_n(K)$ tq $D := P^{-1}AP$ soit une matrice diagonale. On suit le procédé suivant :

- (i) Calculer χ_A .
- (ii) Trouver les zéros $\lambda_1, \dots, \lambda_r \in K$ de χ_A (possible $r \leq n$!)
- (iii) Pour $i = 1, \dots, r$, calculer $E_{\lambda_i}(A) = \text{Ker}(A - \lambda_i I_n)$. Posons $d_i := \dim(E_{\lambda_i}(A))$.
- (iv) Pour $i = 1, \dots, r$, trouver une base $\mathcal{B} = (f_1^i, \dots, f_{d_i}^i)$ de vecteurs $f_k^i \in K^n$.
- (v) Posons $P = (f_1^1, \dots, f_{d_1}^1, f_1^2, \dots, f_{d_r}^r)$.
- (vi) Alors, $D := P^{-1}AP$ est la matrice diagonale $\text{Diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{d_1}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{d_r})$.

En effet, soit $e_1, \dots, e_n \in K^n$ la base canonique. Fix $i = 1, \dots, r$, alors $P e_i = f_k^j \in E_{\lambda_j}(A)$ pour un $1 \leq j \leq r$ et $1 \leq k \leq d_j$ alors $D e_i = P^{-1} A f_k^j = P^{-1} \lambda_j f_k^j = \lambda_j P^{-1} f_k^j = \lambda_j e_i$.

8.3 Passage de $\mathbb{R}$ à $\mathbb{C}$

Soit $n \geq 1$ et soit $A \in \mathcal{M}_n(\mathbb{R})$. Comme $\mathbb{R} \subset \mathbb{C}$ la matrice A peut également être vue comme une matrice de $\mathcal{M}_n(\mathbb{C})$.

Définition 8.30. Posons

- (i) $\text{Sp}_{\mathbb{R}}(A) := \{x \in \mathbb{R} \mid \text{Ker}(xI_n - A) \neq 0\}$ et
- (ii) $\text{Sp}_{\mathbb{C}}(A) := \{x \in \mathbb{C} \mid \text{Ker}(xI_n - A) \neq 0\}$.

Remarque 8.31. Il faut prendre garde au fait que le caractère diagonalisable de A peut dépendre du fait que l'on considère A comme une matrice de $\mathcal{M}_n(\mathbb{R})$ ou une matrice de $\mathcal{M}_n(\mathbb{C})$. Pour bien marquer la différence, on dira que A est diagonalisable dans $\mathcal{M}_n(\mathbb{R})$ (ou parfois juste *sur* $\mathbb{R}$) pour dire que A est diagonalisable comme matrice de $\mathcal{M}_n(\mathbb{R})$. On dira alors que A est diagonalisable dans $\mathcal{M}_n(\mathbb{C})$ (ou parfois just *sur* $\mathbb{C}$) lorsque l'on voit A comme matrice de $\mathcal{M}_n(\mathbb{C})$.

Exemple 8.32. Soit $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, d'où $\chi_A(x) = x^2 + 1$. Alors, $\text{Sp}_{\mathbb{R}}(A) = \{x \in \mathbb{R} \mid x^2 + 1 = 0\} = \emptyset$, par contre $\text{Sp}_{\mathbb{C}}(A) = \{x \in \mathbb{C} \mid x^2 + 1 = 0\} = \{i, -i\}$. On trouve $A \begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} -i \\ 1 \end{pmatrix} = -i \begin{pmatrix} 1 \\ i \end{pmatrix}$ car $-ii = 1$, et $A \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} i \\ 1 \end{pmatrix} = i \begin{pmatrix} 1 \\ -i \end{pmatrix}$. Donc, $A = P \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} P^{-1}$ pour $P := \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$.

8.4 Questions de compréhension et exercices typiques

- (i) Qu'est-ce que sont les valeurs et vecteurs propres d'une matrice diagonale ?
- (ii) Comment on peut déterminer le valeur propre d'une endomorphisme $f \in \mathcal{L}(E)$ si $\dim E < \infty$? Comment un vecteur propre associe ?

Exercice 8.1. Calculer les valeur propres et vecteurs propres d'une matrice donné.

9 Exposé 9 : Réduction des endomorphismes, III

9.1 Polynômes et division euclidienne

Soit $K = \mathbb{R}$ ou $K = \mathbb{C}$.

Théorème 9.1. $\forall P, Q \in K[X], Q \neq 0, \exists S, R \in K[X]$ tq $P = SQ + R$, avec $\deg(R) < \deg Q$. De plus, S, R sont unique.

Démonstration. Par récurrence sur $\deg(P)$. Si $\deg P < \deg Q$, posons $R = P$, $S = 0$. Si $\deg P \geq \deg Q$, soit $m := \deg P - \deg Q$. Il existe alors $s_m \in K$ tq $P_1 := P - aX^m Q$ a $\deg(P_1) < \deg(P)$. Par récurrence, il existe S_1, R_1 avec $P_1 = S_1 Q + R$ et $\deg(R) < \deg(Q)$. Donc,

$$P = aX^m Q + P_1 = (aX^m + S_1)Q + R,$$

et nous posons $S := aX^m + S_1$.

Supposons $SQ + R = P = S'Q + R'$ avec $\deg(R), \deg(R') < \deg(Q)$. Si $S \neq S'$, on a $\deg(S - S') \geq 0$. Donc, $(S - S')Q = R' - R$ et

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

contradiction. □

Exemple 9.2. Soit $P = X^2 + X + 1$, $Q = X - 1$. On veut réduire le degré de P . Donc, considérons $P_1 := P - XQ = X^2 + X + 1 - X^2 + X = 2X + 1$, et $P_1 - 2Q = 2X + 1 - 2X + 2 = 3$. Donc, $P = XQ + P_1 = XQ + 2Q + 3 = (X + 2)Q + 3$ et $R = 3$.

Définition 9.3. (i) On dit que Q divise P , où Q est un diviseur de P , si $R = 0$.

Notation : $Q \mid P$.

(ii) On dit que P, Q sont premiers entre eux si leurs seuls diviseurs communs sont les polynômes constants. Dans ce cas, on note $P \wedge Q = 1$.

On a $Q \mid 0$ pour $Q \in K[X] \setminus \{0\}$, car $0Q = 0$.

Théorème 9.4. $P \wedge Q = 1 \Leftrightarrow \exists U, V \in K[X]$ tq $UP + VQ = 1$.

Démonstration. $\Leftarrow$: Si $S \mid P$ et $S \mid Q$ alors, $S \mid 1 \Rightarrow S$ constant et $P \wedge Q = 1$.

$\Rightarrow$: Par récurrence sur le degré de Q .

-) Si $\deg(Q) = 0$, alors $Q = \lambda \in K^\times$, On pose $U = 0, V = \lambda^{-1}$, donc $UP + VQ = 1$.

-) Supposons la propriété vraie pour tous P tq $\deg(P) \leq d$. Soient $P, Q \in K[X]$, $Q \neq 0$, tq $P \wedge Q = 1$. Ecrivons $P = QS + R$ avec $\deg(R) < \deg(P)$ (par 9.1). On applique l'hypothèse à (P, R) : en effet, si $B \in K[X] \setminus \{0\}$ est tel que $B|P$ et $B|R$ alors, $B|Q \Rightarrow B = 1$ (car $P \wedge Q = 1$). Donc, $P \wedge R = 1$ et par l'hypothèse de récurrence $\exists U_1, V_1 \in K[X]$ tq $U_1P + V_1R = 1$. On a $V_1R = V_1(P - QS)$ (par définition de R), donc

$$1 = U_1P + V_1R = U_1P + V_1P - V_1QS = \underbrace{(U_1 + V_1)}_{U:=}P + \underbrace{(-V_1S)}_{V:=}Q. \quad \square$$

Exemple 9.5. Soit $a, b \in K$, $a \neq b$. Alors, $(X - a) \wedge (X - b) = 1$ (car le seul diviseur unitaire de $\deg > 0$ de $X - a$ est $X - a$).

Définition 9.6. Soit $P \in K[X]$ et soit $x \in K$. On dit que x est une racine de P si $P(x) = 0$. On note $\text{Rac}(P) \subseteq K$ l'ensemble des racines de P .

Remarque 9.7. Si $P, Q \in K[X]$, alors $\text{Rac}(PQ) = \text{Rac}(P) \cup \text{Rac}(Q)$ (en effet, pour $x \in K$ on a $P(x)Q(x) = 0 \Leftrightarrow P(x) = 0$ où $Q(x) = 0$).

On va admettre le théorème suivant (qui sera prouver dans un cours d'analyse complexe).

Théorème 9.8 (d'Alembert-Gauß). Soit $P \in \mathbb{C}[X]$. Si $\deg P > 0$, alors il existe au moins une racine de P dans $\mathbb{C}$.

Remarque 9.9. Si $P(X) = X^2 + 1 \in \mathbb{R}[X]$, le polynôme P n'a pas une racine dans $\mathbb{R}$, seulement sur $\mathbb{C}$. Donc, on ne peut pas remplacer $\mathbb{C}$ par $\mathbb{R}$ dans 9.6 ! Mais il y a une version plus faible : tout polynôme $P \in \mathbb{R}[X]$ dont le degré est *impair* possède une racine réelle (parce que $P(-x), P(x)$ ont un signe différent pour $x \in \mathbb{R}$ très grand).

Remarque 9.10. Si Q divise P , le polynôme R tel que $P = QR$ est unique. En effet, si $QR = QR'$, alors $Q(R - R') = 0$, et ça force $(R - R') = 0$.

Proposition 9.11. Soit $P \in K[X]$ et soit $a \in K$. Alors $P(a) = 0$ si et seulement si $X - a$ divise P .

Démonstration. $\exists S, R \in K[X]$ tq $P = S(X - a) + R$ et $\deg(R) < \deg(X - a) = 1$. (par 9.1). Donc, $R = c \in K$ est constant. On a $P(a) = 0 \Leftrightarrow S(a - a) + R(a) = 0 \Leftrightarrow R(a) = 0 \Leftrightarrow c = 0 \Leftrightarrow (X - a) \mid P$. $\square$

Théorème 9.12. Soit $P \in \mathbb{C}[X]$ un polynôme unitaire.

(i) $\text{Rac}(P)$ est fini.

(ii) Pour $a \in \text{Rac}(P)$ existe unique $m_a \in \mathbb{N}_{\geq 1}$ tq

$$P(X) = \prod_{a \in \text{Rac}(P)} (X - a)^{m_a}.$$

En particulier, $\deg(P) = \sum_{a \in \text{Rac}(P)} m_a$.

Démonstration. Récurrence sur $n := \deg(P)$.

-) $n = 1$: $\exists a \in K$ tq $P(X) = X - a$ (car P unitaire). Alors, $\text{Rac}(P) = \{a\}$ et $m_a = 1$.
-) $n \geq 2$: Car $n = \deg(P) > 0$, $\exists a \in \mathbb{C}$ tq $P(a) = 0$ (par 9.8). Par 9.11 $(X - a) \mid P$, et il existe Q tq $(X - a)Q = P$. Par hypothèse, $Q = \prod_{b \in \text{Rac}(Q)} (X - b)^{m'_b}$ pour $m'_b \geq 1$. On a $\text{Rac}(P) = \text{Rac}(Q) \cup \{a\}$. Si $a \notin \text{Rac}(Q)$, posons $m_a := 1$. Si $a \in \text{Rac}(Q)$, posons $m_a := m'_a + 1$. $\square$

Définition 9.13. (i) Un polynôme de $K[X]$ de la forme $\prod_{a \in M} (X - a)^{m_a}$ avec $M \subseteq K$ fini, $m_a \in \mathbb{N}_{\geq 1}$ si $a \in M$, est dit *scindé* sur K .

(ii) Soit $P(X) \in K[X]$, $a \in \text{Rac}(P)$. L'entier m_a de 9.12 s'appelle *multiplcité* de a dans $P(X)$.

(iii) Un polynôme $P \in K[X]$ est dit *simplement scindé* si pour $a \in \text{Rac}(P)$ on a $m_a = 1$.

Remarque 9.14. On a $m_a = \max\{m \geq 1 \mid (X - a)^m \text{ divise } P(X)\}$.

9.2 Trigonalisation

Proposition 9.15. Soient $1 \leq r \neq n$ des entiers et soit $M \in \mathcal{M}_n(K)$ une matrice de la forme

$$M = \begin{pmatrix} M_1 & M_2 \\ 0 & M_3 \end{pmatrix}$$

avec $M_1 \in \mathcal{M}_r(K)$, $M_2 \in \mathcal{M}_{r, n-r}(K)$ et $M_3 \in \mathcal{M}_{n-r}(K)$. Alors

$$\det(M) = \det(M_1) \det(M_3) \text{ et } \chi_M(X) = \chi_{M_1}(X) \chi_{M_3}(X).$$

Démonstration. Récurrence sur n :

-) $n = 1$: ✓
-) $n \geq 2$: Notons, que l'affirmation " $\det(M) = \det(M_1) \det(M_3)$ " est invariant si on fait des opérations/permutations sur les premier r -lignes de M (par 5.16), et aussi par des opérations/permutations sur les premières r -colonnes

(5.3). Donc, pouvons supposer que la premier colonne de M_1 est $\begin{pmatrix} \lambda \\ 0 \\ \dots \\ 0 \end{pmatrix}$ pour un $\lambda \in K$. Donc,

$$\det(M) = \lambda \det(M_{1,1}), \quad \det(M_1) = \lambda \det((M_1)_{1,1})$$

(en développement de M et M_1 par rapport à la premier colonne). Mais, $M_{1,1}$ est triangulaire supérieure de blocs $(M_1)_{1,1}, M_3$ sur la diagonale. Donc, par HR : $\det(M_{1,1}) = \det((M_1)_{1,1}) \det(M_3)$. Le même argument montre $\chi_M(x) = \chi_{M_1}(x) \chi_{M_3}(x)$ pour $x \in K$, donc $\chi_M = \chi_{M_1} \chi_{M_3}$. $\square$

Théorème 9.16. *Une matrice $A \in \mathcal{M}_n(K)$ est trigonalisable dans $\mathcal{M}_n(K)$ si et seulement si son polynôme caractéristique χ_A est scindé dans $K[X]$.*

Démonstration. $\Rightarrow$: Supposons A trigonalisable, et soit $P \in \text{GL}_n(K)$ tq $T := P^{-1}AP$ triangulaire supérieure. Écrivons $T = (t_{ij})_{ij}$. Alors, $\chi_T(x) = \prod_{i=1}^n (x - t_{ii})$, et χ_T est scindé.

$\Leftarrow$: Récurrence sur n .

$n = 1$: $\checkmark$

$n \geq 2$: Car χ_A est scindé, exists $\lambda \in K$ avec $\chi_A(\lambda) = 0$. Par 8.24 $\exists v \in K^n$ non-nul tq $Av = \lambda v$. Donc, $\exists$ une base $\mathcal{B} := (e_1, \dots, e_n)$ de K^n tq $e_1 = v$. On a

$$\text{Mat}_{\mathcal{B}}(u_A) = \begin{pmatrix} \lambda & \dots \\ 0 & B \end{pmatrix}$$

où $B \in \mathcal{M}_{n-1}(K)$ (7.12). On a $\chi_B \mid \chi_A$ par 9.15. Donc χ_B est scindé et, par HR, $\exists Q \in \text{GL}_{n-1}(K)$ tq $Q^{-1}BQ$ triangulaire supérieure. Posons $P = \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} \in \mathcal{M}_n(K)$. On a $\det(P) = \det(Q)$ par 9.15, donc $P \in \text{GL}_n(K)$ par 5.5. On voit que $P^{-1}AP = \begin{pmatrix} \lambda & \dots \\ 0 & Q^{-1}BQ \end{pmatrix}$ est triangulaire supérieure. $\square$

Corollaire 9.17 (7.4). *Toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ est trigonalisable dans $\mathcal{M}_n(\mathbb{C})$.*

Démonstration. On combine 9.16 et 9.12. $\square$

Exemple 9.18. Pour $\theta \in \mathbb{R}$ la matrice

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$$

n'est pas trigonalisable sur $\mathbb{R}$ car son polynôme caractéristique $X^2 + 1 \in \mathbb{R}[X]$ n'est pas scindé. Sur $\mathbb{C}$, $X^2 + 1 = (X - i)(X + i) \in \mathbb{C}[X]$ est $R(\theta)$ est diagonalisable sur $\mathbb{C}$ (par 8.20).

Exemple 9.19. Soit $A = \begin{pmatrix} -4 & 0 & -2 \\ 0 & 1 & 0 \\ 5 & 1 & 3 \end{pmatrix}$. On a $\chi_A(x) = (X - 1)^2(X + 2)$ (vérifier !), qui est un polynôme scindé. Donc, A est trigonalisable (9.16). Soit $f := u_A$ l'endomorphisme $\mathbb{R}^3 \rightarrow \mathbb{R}^3$, $v \mapsto Av$. Nous cherchons une base $\mathcal{B} = (v_1, v_2, v_3)$ de

$\mathbb{R}^3$ tq $\text{Mat}_{\mathcal{B}}(u_A) = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & -2 \end{pmatrix}$ est triangulaire (les termes sur la diagonale sont donné par les racine de $\chi_A(x)$ compté avec leur multiplicité). En effet, on peut supposer $b = 0, c = 0$ si on demand

$$\text{Vect}(v_1) = E_\lambda(A), \text{ Vect}(v_3) = E_{-2}(A).$$

-) Calcul de v_1 : v_1 est un vecteur propre correspondant à $\lambda = 1$. Donc, on cherche

$$\text{Ker}(A - I_3) = \text{Ker} \begin{pmatrix} -5 & 0 & -2 \\ 0 & 0 & 0 \\ 5 & 1 & 2 \end{pmatrix} = \text{Ker} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 5 & 1 & 2 \end{pmatrix} = \text{Vect} \left(\underbrace{\begin{pmatrix} 2 \\ 0 \\ -5 \end{pmatrix}}_{=: v_1} \right)$$

Calcul de v_2 : On cherche $a \in K$ et $v_2 \in K^3$ tq $(f - I_3)v_2 = av_1$, i.e., $v_2 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, et

$$(A - I_3)v = \begin{pmatrix} -5 & 0 & -2 \\ 0 & 0 & 0 \\ 5 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = a \begin{pmatrix} 2 \\ 0 \\ 5 \end{pmatrix},$$

d'où, par exemple, $a = 1$ et $v_2 := \begin{pmatrix} -2 \\ -3 \\ 4 \end{pmatrix}$.

Calcul de v_3 : On prend v_3 comme vecteur propre associé à la valeur propre -2 .
Donc, on calcule

$$\text{Ker}(A + 2I_3) = \text{Ker} \begin{pmatrix} -2 & 0 & -2 \\ 0 & 3 & 0 \\ 5 & 1 & 5 \end{pmatrix} = \text{Ker} \begin{pmatrix} -2 & 0 & -2 \\ 0 & 3 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \text{Vect} \left(\underbrace{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}}_{=:v_3} \right)$$

Conclusion : Si on pose $P = (v_1|v_2|v_3) = \begin{pmatrix} 2 & -2 & 1 \\ 0 & -3 & 0 \\ -5 & 4 & -1 \end{pmatrix}$, on a $A = P \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} P^{-1}$.

9.3 Polynômes d'endomorphismes

Soit $M \in \mathcal{M}_n(K)$. Soit $P \in K[X]$. Alors, on peut écrire $P(X) = \sum_{i=0}^d a_i X^i$ et on pose

$$P(A) := \sum_{i=0}^d a_i A^i = a_0 I_n + a_1 A + \dots + a_d A^d \in \mathcal{M}_n(K).$$

Exemple 9.20. $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ et $P(X) = X^2 + X + 1$. Donc,

$$P(A) = \underbrace{\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}}_{=A^2} + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}$$

Analog pour les endomorphismes : Soit E un K -espace vectoriel. Soit $f \in \mathcal{L}(E)$.
Si $n \geq 1$, on note

$$f^n = \underbrace{f \circ \dots \circ f}_n$$

et on pose $f^0 = \text{Id}_E$.

Soit $P = \sum_{i=0}^n a_i X^i \in K[X]$ un polynôme. On définit alors

$$P(f) := \sum_{i=0}^n a_i f^i \in \mathcal{L}(E).$$

Si $\dim_K E = n$ est fini, et $\mathcal{B}$ est une base de E , alors

$$\text{Mat}_{\mathcal{B}}(P(f)) = P(\text{Mat}_{\mathcal{B}}(f))$$

car $\text{Mat}_{\mathcal{B}}(f^i) = \text{Mat}_{\mathcal{B}}(f)^i$.

Remarque 9.21. Soient $P, Q \in K[X]$, $f \in \mathcal{L}(E)$.

(i) Si $\lambda \in K$, on a $(P + \lambda Q)(f) = P(f) + \lambda Q(f)$.

(ii) $(PQ)(f) = P(f) \circ Q(f)$ (en particulier, $P(f) \circ Q(f) = Q(f) \circ P(f)$)

De même pour des matrices. De plus, $B^{-1}P(A)B = P(B^{-1}AB)$ si $B \in \text{GL}_n(K)$ (car $B^{-1}A^iB = \underbrace{B^{-1}AB^{-1}B \dots ABB^{-1}AB}_{i\text{-facteurs de } B^{-1}AB} = (B^{-1}AB)^i$).

9.4 Le théorème de Cayley-Hamilton

Soit E un K -espace vectoriel.

Définition 9.22. Soit $f \in \mathcal{L}(E)$. Un polynôme $Q \in K[X]$ est dit *annulation* de f si $Q(f) = 0$.

Théorème 9.23 (Cayley-Hamilton). Soit $A \in \mathcal{M}_n(K)$. Alors, $\chi_A(A) = 0$, i.e., le polynôme caractéristique de A est un polynôme d'annulation pour A (et u_A).

De même, on a $\chi_f(f)$ pour $f \in \mathcal{L}(E)$ et E un K -espace vectoriel de dimension finie.

Démonstration. • Si $K = \mathbb{R}$, l'affirmation $\chi_A(A) = 0$ ne dépend pas de fait que on voit A comme matrice dans $\mathcal{M}_n(\mathbb{R})$ ou dans $\mathcal{M}_n(\mathbb{C})$. Donc, supposons $K = \mathbb{C}$, et posons $f = u_A \in \mathcal{L}(K^n)$.

• Par 9.17 il existe une base $\mathcal{B} = (e_1, \dots, e_n)$ tq $\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & * \\ & \dots & \\ 0 & & \lambda_n \end{pmatrix}$ est triangulaire supérieure. Donc, $\chi_A(x) = \prod_{i=1}^n (x - \lambda_i)$ (8.25). Écrivons

$$g_i := (f - \lambda_1 \text{Id}) \circ \dots \circ (f - \lambda_i)$$

pour $i = 1, \dots, n$. Montrons par récurrence sur i que g_i annule les vecteurs $e_1, e_2, \dots, e_i$. Donc, $g_n = \chi_f(f) = 0$.

• $i = 1$: $g_1(e_1) = (f - \lambda_1 \text{Id})(e_1) = f(e_1) - \lambda_1 e_1 = \lambda_1 e_1 - \lambda_1 e_1 = 0$.

• $i \geq 1$: On a $g_i = (f - \lambda_i \text{Id}) \circ g_{i-1} = g_{i-1} \circ (f - \lambda_i \text{Id})$. Donc,

$$g_i(e_j) = (f - \lambda_i \text{Id}) \circ g_{i-1}(e_j) = 0$$

pour $1 \leq j \leq i-1$. Il reste à montrer que $g_i(e_i) = 0$. On a

$$(f - \lambda_i \text{Id})(e_i) = \lambda_i e_i + w - \lambda_i e_i = w$$

pour un $w \in \text{Vect}(e_1, \dots, e_{i-1})$ (car $\text{Mat}_{\mathcal{B}}(f)$ est triangulaire supérieure). Donc, $g_i(e_i) = g_{i-1}(f - \lambda_i \text{Id})(e_i) = g_{i-1}(w) = 0$ par hypothèse. $\square$

9.23 est très utile pour comprendre le comportement d'une endomorphisme $f \in \mathcal{L}(E)$, puisque on peut l'utiliser pour construire beaucoup des sous-espaces de E qui sont stable par f .

Définition 9.24. Soit $f \in \mathcal{L}(E)$. Supposons $\chi_f(x) = \prod_{\lambda \in \text{Sp}(f)} (x - \lambda)^{m_\lambda}$ est scindé, $m_\lambda \in \mathbb{N}_{\geq 1}$.

(i) m_λ est la multiplicité de valeur propre λ .

(ii) Le sous-espace

$$F_\lambda(f) := \text{Ker}((f - \lambda \text{Id}_E)^{m_\lambda}) \subseteq E$$

est le sous-espace caractéristique pour la valeur propre λ .

Si $A \in \mathcal{M}_n(K)$, on pose $F_\lambda(A) := F_\lambda(u_A)$ pour $\lambda \in \text{Sp}(A) = \text{Sp}(u_A)$.

Remarque 9.25. On a

$$E_\lambda(f) = \text{Ker}((f - \lambda \text{Id})) \subseteq \text{Ker}((f - \lambda)^2) \subseteq \dots \subseteq F_\lambda(f),$$

mais $E_\lambda(f) \neq F_\lambda(f)$ en général : si $A = \begin{pmatrix} 0 & 10 \\ 0 & 01 \end{pmatrix}$ et $(e_1, e_2, e_3) \in K^3$ est la base canonique, alors pour $\lambda = 0$

$$E_0(A) = \text{Vect}(e_1) \subseteq \text{Ker}(A^2) = \text{Vect}(e_1, e_2) \subseteq F_0(A) = \text{Vect}(e_1, e_2, e_3)$$

car $\chi_A(x) = X^3$.

Lemme 9.26. Soit E un K -espace vectoriel. Soit $f \in \mathcal{L}(E)$, et $P \in K[X]$. Alors,

$$E_P(f) := \text{Ker}(P(f)) = \{v \in E \mid P(f)v = 0\}$$

est un sous-espace vectoriel de E , qui est stable par f , i.e., $v \in E_P(f) \Rightarrow f(v) \in E_P(f)$.

Démonstration. $P(f) \in \mathcal{L}(E)$ est $E_P(f) = \text{Ker}(P(f))$, donc $E_P(f)$ est un sous-espace. Soit $v \in E_P(f)$, écrivons $P(X) = a_0 + a_1X + \dots + a_nX^n$. Alors, $P(f) \circ f = f \circ P(f)$ par 9.21. Donc,

$$P(f)(f(v)) = (P(f) \circ f)(v) = (f \circ P(f))(v) = f(P(f)(v)) = 0$$

et donc $f(v) \in E_P(f)$. □

Exemple 9.27. Soit $a \in K$, et $P(X) = X - a$. Donc,

$$E_P(f) = \text{Ker}(f - a) = E_a(f)$$

(notation : 9.26, 8.7).

Lemme 9.28 (Lemme des noyaux). *Soit $A \in \mathcal{M}_n(K)$. Soit $Q, R \in K[X]$. Soit $R \wedge Q = 1$, i.e., R, Q premiers entre eux. On a $E_{QR}(A) = E_Q(A) \oplus E_R(A)$.*

Démonstration. Car $Q \wedge R = 1$, $\exists U, V \in K[X]$ tq $UQ + VR = 1$ (9.4).

-) Par 9.26 appliquer à $f = u_A: K^n \rightarrow K^n$, $x \mapsto Ax$, $E_Q(A)$, $E_P(A)$ sont des sous-espaces vectoriels, stable par A .

-) Soit $v \in E_Q(A) \cap E_R(A)$. Donc, $Q(A)v = 0$ et $R(A)v = 0$. Alors,

$$v = A^0 v = (UQ + VR)(A)v = U(A)Q(A)v + V(A)R(A)v = 0.$$

-) Soit $v \in E_{QR}(A)$. Donc, $0 = Q(A)R(A)v = R(A)Q(A)v$, d'où $R(A)v \in E_Q(A)$ et $Q(A)v \in E_R(A)$. On a

$$v = \underbrace{U(A)Q(A)v}_{\in E_R(A)} + \underbrace{V(A)R(A)v}_{\in E_Q(A)}.$$

En effet, $E_Q(A)$ est stable par A , et donc aussi stable par $V(A)$. De même, $U(A)Q(A)v \in E_R(A)$.

□

Corollaire 9.29. *Soient $P_1, \dots, P_k \in K[X]$, deux à deux premiers entre eux, $k \geq 2$, alors $E_{P_1 \dots P_k}(f) = E_{P_1}(f) \oplus \dots \oplus E_{P_k}(f)$.*

Démonstration. Par récurrence : on pose $R = P_1 \dots P_{k-1}$, $Q = P_k$. Donc, $R \wedge Q = 1$, et

$$E_{RQ}(f) = E_R(f) \oplus E_Q(f)$$

par 9.28. On concluit en appliquant la HR à $R = P_1 \dots P_{k-1}$.

□

Corollaire 9.30. *Soit $f \in \mathcal{L}(E)$. Supposons $\chi_f(x) = \prod_{i=1}^r (x - \lambda_i)^{m_{\lambda_i}}$ où $\lambda_i \neq \lambda_j$ pour $i \neq j$. Alors,*

$$E = F_{\lambda_1}(f) \oplus \dots \oplus F_{\lambda_r}(f).$$

De plus, il existe une base adaptée $\mathcal{B}$ à cette décomposition tq la matrice $\text{Mat}_{\mathcal{B}}(f)$ est une matrice $\text{Diag}(A_1, \dots, A_r)$ par bloc, où A_i est une matrice triangulaire supérieure avec seule term λ_i sur la diagonale.

Démonstration. Par 9.23 on a $\chi_f(f) = 0$, donc $E_{\chi_f(x)}(f) = E$ et on concluit par 9.29. Car chaque $F_i := F_{\lambda_i}(f) = E_{(x-\lambda_i)^{m_{\lambda_i}}}$ est stable par f , on peut triagonaliser $f_i := f|_{F_i}$. En effet, $g_i := f_i - \lambda_i \text{Id}_{F_i}$ est un endomorphisme nilpotent par définition de F_i : $g_i^{m_{\lambda_i}} = (f_i - \lambda_i \text{Id}_{F_i})^{m_{\lambda_i}} = 0$. Donc, $x^{m_{\lambda_i}} \stackrel{8.28}{=} \chi_{g_i}(x) = \det(x \text{Id}_{F_i} - g_i) = \det(x \text{Id}_{F_i} - f_i + \lambda_i \text{Id}_{F_i}) = \chi_{f_i}(\lambda_i)$ et $\chi_{f_i}(x) = (x - \lambda_i)^{m_{\lambda_i}}$. □

Remarque 9.31. Grâce à la théorie de bloc de Jordan on peut prouver que il existe une base tq $A_i = (b_{jk})_{jk}$ est une matrice triangulaire superieure tq

$$\begin{cases} b_{jk} = \lambda_i & , j = k \\ b_{jk} \in \{0, 1\} & , j = k - 1 \\ b_{jk} = 0 & , \text{sinon.} \end{cases}$$

Donc, A_i est une matrice diagonale par bloc tq les blocs ont la forme

$$J_m(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 \\ & \dots & & \dots & \\ 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix} \in \mathcal{M}_m(K)$$

pour un $m \geq 1$. Si $(e_1, \dots, e_m)$ est la base canonique de K^k , on voit que $(f - \lambda_i)(e_m) = e_{m-1}$, $(f - \lambda_i)(e_{m-1}) = (f - \lambda_i)^2(e_m) = e_{m-2}, \dots (f - \lambda_i)^{m-1}(e_m) = e_1$, $(f - \lambda_i)^m(e_m) = 0$.

10 Exposé 10 : Réduction des endomorphismes, IV

10.1 La trace et le déterminant

Définition 10.1. Soit $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(K)$. La trace de A est le scalaire

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

Remarque 10.2. Si $A, B \in \mathcal{M}_n(K)$, on a $\text{Tr}(AB) = \text{Tr}(BA)$ (exercice!). En particulier si $P \in \text{GL}_n(K)$, alors

$$\text{Tr}(PAP^{-1}) = \text{Tr}(P^{-1}PA) = \text{Tr}(A).$$

Donc, les matrices semblable ont la même trace.

Corollaire 10.3. Soit $A \in \mathcal{M}_n(\mathbb{C})$. On a alors

$$\text{Tr}(A) = \sum_{\lambda \in \text{Sp } A} m_\lambda(A) \lambda, \quad \det(A) = \prod_{\lambda \in \text{Sp } A} \lambda^{m_\lambda(A)}.$$

Démonstration. Par 9.17 il existe $P \in \text{GL}_n(\mathbb{C})$ tq $B := PAP^{-1}$ est une matrice triangulaire supérieure. Car $\text{Tr}(A) = \text{Tr}(B)$, $\det(A) = \det(B)$, on peut supposer alors que A soit triangulaire supérieure. Par 9.15 on a $\chi_A(x) = \prod_{i=1}^n (x - a_{i,i}) = \prod_{\lambda \in \text{Sp}(A)} (x - \lambda)^{m_\lambda(A)}$. On peut conclure. $\square$

10.2 Réduction des endomorphismes auto-adjoint

Soit E un $(\mathbb{R})$ -espace euclidien. Soit $f \in \mathcal{L}(E)$. Rappel :

-) f est auto-adjoint (ou symétrique) si $f^* = f$ ($\Leftrightarrow$ si $\mathcal{B}$ est une BON de E la matrice $\text{Mat}_{\mathcal{B}}(f)$ est symétrique, 3.11).
-) Si $F \subseteq E$ est un sous-espace de E stable par f , alors $F^\perp$ est stable par f^* (3.15).

Proposition 10.4. Soit $f \in \mathcal{L}(E)$ auto-adjoint. Soit $\lambda, \mu \in \mathbb{R}$, $\lambda \neq \mu$. Alors, $E_\lambda(f) \perp E_\mu(f)$.

Démonstration. Soit $v \in E_\lambda(f)$, $w \in E_\mu(f)$. Alors, $\lambda \langle v, w \rangle = \langle \lambda v, w \rangle = \langle f(v), w \rangle \stackrel{f \text{ auto-adjoint}}{=} \langle v, f(w) \rangle = \mu \langle v, w \rangle$. Donc, $(\lambda - \mu) \langle v, w \rangle = 0$, donc $\langle v, w \rangle = 0$. $\square$

Lemme 10.5. Soit $M \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique réel. Alors, $\text{Sp}_{\mathbb{C}}(M) \subseteq \mathbb{R}$, donc toutes les valeurs propres dans $\mathbb{C}$ sont déjà dans $\mathbb{R}$.

Démonstration. Soit $\lambda \in \text{Sp}_{\mathbb{C}}(M)$ et $v = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{C}^n$ un vecteur propre pour λ .

Posons $\bar{v} := \begin{pmatrix} \overline{a_1} \\ \vdots \\ \overline{a_n} \end{pmatrix}$ où $\bar{\cdot} : \mathbb{C} \rightarrow \mathbb{C}$ est la conjugaison complexe.

$$\bullet) {}^t\bar{v}Mv = \lambda {}^t\bar{v}v$$

$$\bullet) {}^t\bar{v}Mv = {}^t(M\bar{v})v = {}^t(\overline{Mv})v \text{ (car } M \text{ a ses termes dans } \mathbb{R}), \text{ et de plus } = {}^t(\overline{\lambda vv})v = \bar{\lambda} {}^t\bar{v}v.$$

Donc, $\lambda {}^t\bar{v}v = \bar{\lambda} {}^t\bar{v}v$. Car ${}^t\bar{v}v = \sum_{i=1}^n \overline{a_i}a_i = \sum_{i=1}^n |a_i|^2 \neq 0$, donc $\lambda = \bar{\lambda}$ et $\lambda \in \mathbb{R}$. $\square$

Théorème 10.6 (7.3). Soit $f \in \mathcal{L}(E)$ auto-adjoint. Alors, f est diagonalisable et possède une base propre orthonormale, i.e., une base orthonormale de vecteurs propres.

Démonstration. Récurrence sur $n := \dim(E)$. Après 10.5 il existe $\lambda \in \text{Sp}_{\mathbb{R}}(f)$. Soit $0 \neq v \in E_{\lambda}(f)$, et $D := \text{Vect}(v)^{\perp}$. Donc, $E = \text{Vect}(v) \oplus D$, et $\text{Vect}(v), D$ sont stable par f . Posons $g := f|_D : D \rightarrow D$. Soit $\mathcal{C} = (e_2, \dots, e_n)$ une BON de D , et $\mathcal{B} = \text{Diag}(\frac{v}{|v|}, e_2, \dots, e_n)$. Alors, $\mathcal{B}$ est une BON de E . Donc, $\text{Mat}_{\mathcal{B}}(f) = \text{Diag}(\lambda, \text{Mat}_{\mathcal{C}}(g))$, et donc g est auto-adjoint (3.11). Par récurrence, posons supposer que $\mathcal{C}$ est une base propre orthonormale de g , d'où le thm. $\square$

Corollaire 10.7. Soit $A \in \mathcal{M}_n(\mathbb{R})$ symétrique. Alors, A est diagonalisable dans une base orthonormé, i.e., $\exists P \in \text{O}_n(\mathbb{R})$ tq $P^{-1}AP$ soit diagonale.

Démonstration. C'est la version matricielle de 10.6. $\square$

Exemple 10.8. Soit $A = \begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}$. On calcule $A^2 = 9I_3$, donc $P(A) = 0$

ou $P(X) = X^2 - 9 = (X - 3)(X + 3)$. Par 9.28, $\text{Sp}(A) = \{3, -3\}$. But : Trouver $P \in \text{O}_3(\mathbb{R})$ tq tPAP diagonale. On calcule

$$E_3(A) = \text{Vect}\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right), \quad E_{-3}(A) = E_3(A)^{\perp} = \text{Vect}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right).$$

Grâce à Gram-Schmidt on trouve une BON de $E_3(A)$ comme $v_1 := \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$, $v_2 := \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}$ et $v_3 := \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$ est une BON de $E_{-3}(A)$. Donc, posons $P := (v_1|v_2|v_3)$ et ${}^tPAP = \text{Diag}(3, 3, -3)$.

10.3 Les groupes $O_2(\mathbb{R})$ et $O_3(\mathbb{R})$

Rappel :

-) $O_n(\mathbb{R}) = \{M \in \mathcal{M}_n(\mathbb{R}) \mid {}^tM \cdot M = I_n\}$ est le groupe orthogonal, 3.17.
-) Si $M \in O_n(\mathbb{R})$, alors $1 = \det(I_n) = \det({}^tMM) = \det(M)^2$.
-) Si $M \in O_n(\mathbb{R})$, et $0 \neq v \in E_\lambda(M)$, $\lambda \in \mathbb{R}$, alors $\|v\| = \|Mv\| = \|\lambda v\| = |\lambda| \|v\|$, donc $|\lambda| = 1$, i.e., $\lambda \in \{1, -1\}$.

Proposition 10.9. Soit $R \in O_2(\mathbb{R})$. Alors, $\exists a, b \in \mathbb{R}$ tq $a^2 + b^2 = 1$, et

- (i) si $\det(R) = 1$, $R = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.
- (ii) si $\det(R) = -1$, $R = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$.

Démonstration. Ecrivons

$$R = \begin{pmatrix} a & x \\ b & y \end{pmatrix}.$$

Car R est orthogonale, les colonnes de R sont une BON (3.19). Donc, $ax + by = 0$, et $a^2 + b^2 = 1$. Il existe $t \in \mathbb{R}$ tq $x = tb$, $y = -ta$ (car $\text{Ker} \begin{pmatrix} a & b \end{pmatrix} = \text{Vect} \left(\begin{pmatrix} -b \\ a \end{pmatrix} \right)$). Car $1 = x^2 + y^2 = t^2(a^2 + b^2) = t^2$. Donc, $t \in \{1, -1\}$. De plus, $\det(R) = ay - xb = -ta^2 - tb^2 = -t$. $\square$

Remarque 10.10. Si $a^2 + b^2 = 1$, il existe $\theta \in \mathbb{R}$ tq $a = \cos(\theta)$, $b = \sin(\theta)$ (3.23), donc $R = R(\theta)$, ou $R = R(\theta) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Deux cas :

-) Si $\det(R) = 1$, donc $R = R(\theta)$, et R est la matrice de la rotation de centre 0 et d'angle θ .

-) Si $\det(R) = -1$, alors $R = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$ est symétrique, et donc diagonalisable par 10.6. En fait, R a le polynôme caractéristique $(X - \cos(\theta))(X + \cos(\theta)) - \sin(\theta)^2 = X^2 - (\cos(\theta)^2 + \sin(\theta)^2) = X^2 - 1 = (X - 1)(X + 1)$.
et $\mathbb{R}^2 = E_1(R) \oplus E_{-1}(R)$. On sait $E_1(R) \perp E_{-1}(R)$ (10.4). Donc, R est la symétrie orthogonale d'axe $E_1(R)$.

Proposition 10.11. *Soit $M \in O_3(\mathbb{R})$. Alors, M est semblable à*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix} \text{ ou } \begin{pmatrix} -1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}$$

avec $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O_2(\mathbb{R})$.

Démonstration. $\chi_M(X) \in \mathbb{R}[X]$ est un polynôme unitaire de degré 3 $\Rightarrow \exists \lambda \in \mathbb{R}$ tq $\chi_A(\lambda) = 0$. Donc, $\lambda \in \text{Sp}(M) \subseteq \{1, -1\}$. Soit $0 \neq v \in E_\lambda(M)$, et $D := \text{Vect}(v)^\perp$. Par 3.16 D est stable par M . De plus, $D \rightarrow D$, $v \mapsto Mv$ est une isométrie pour la restriction du produit scalaire canonique le long $D \subseteq \mathbb{R}^3$. Donc, en prenant une BON $\mathcal{B} = (v_1, v_2, v_3)$ adapté à la décomposition $\mathbb{R}^3 = \text{Vect}(v) \oplus D$, on voit que

$$\text{Mat}_{\mathcal{B}}(u_M) = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}$$

avec $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O_2(\mathbb{R})$. □

Définition 10.12. $SO_n(\mathbb{R}) := \{M \in O_n(\mathbb{R}) \mid \det(M) = 1\}$ est le groupe des matrices spéciales orthogonales.

Corollaire 10.13. *Toute matrice de $SO_3(\mathbb{R})$ est semblable à une matrice de la forme*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$$

pour un $\theta \in \mathbb{R}$.

Démonstration. Par 10.11 M est semblable à une matrice $\text{Diag}(\lambda, R)$ avec $R \in O_2(\mathbb{R})$ (diagonale par bloc). Donc, $1 = \det(M) = \lambda \det(R)$.

-) Si $\lambda = 1$, $\det(R) = 1$, et donc $R = R(\theta)$ pour un $\theta \in \mathbb{R}$.
-) Si $\lambda = -1$, alors R a 1 comme valeur propre 10.10 et la preuve 10.11 montre que on peut réduire au cas $\lambda = 1$.

□

10.4 Décomposition $d + n$

Soit E un K -espace vectoriel de dimension finie.

Théorème 10.14. *Soit $f \in \mathcal{L}(E)$ tq $\chi_f(x)$ est scindé. Alors, $\exists! d, n \in \mathcal{L}(E)$ tq d diagonalisable, n nilpotente et $f = d + n$ et $d \circ n = n \circ d$.*

Démonstration. •) Existence : Car $\chi_f(x) = \prod_{i=1}^r (x - \lambda_i)^{m_i}$ est scindé ($\lambda_i \neq \lambda_j$ si $i \neq j$),

$$E = F_1 \oplus \dots \oplus F_r$$

avec $F_i = \text{Ker}((f - \lambda_i \text{Id})^{m_i})$. Posons $d \in \mathcal{L}(E)$ tq $d|_{F_i} = \lambda_i \text{Id}_{F_i}$ (donc, dans une base $\mathcal{B} = \mathcal{B}_1 \amalg \dots \amalg \mathcal{B}_r$ adapté à la décomposition précédente, la matrice de d est $\text{Diag}(\lambda_1 I_{m_1}, \dots, \lambda_r I_{m_r})$). Donc, $d(v) = \sum_{i=1}^r \lambda_i v_i$ si $v = \sum_{i=1}^r v_i$ tq $v_i \in F_i$, et chaque vecteur de F_i , $i = 1, \dots, r$, est un vecteur propre de d et d est diagonalisable. De plus, $d \circ f = f \circ d$ car F_i , $i = 1, \dots, r$, est stable par f . Posons $s := \max\{m_1, \dots, m_r\}$, alors $s \geq m_i$, $\forall 1 \leq i \leq r$. Mq $(f - d)^s = 0$, donc $n := f - d$ est nilpotente ($\Rightarrow$ existence dans le thm car $n \circ d = d \circ n$). Soit $1 \leq i \leq r$, et $v \in F_i$. Donc, $d(v) = \lambda_i v$ est

$$(f - d)^s(v) = (f - \lambda_i \text{Id}_E)^s(v) = 0$$

car $s \geq m_i$ et $F_i = \text{Ker}((f - \lambda_i \text{Id}_E)^{m_i})$.

•) Unicité : Supposons $f = d + n$. Car d est diagonalisable,

$$E = E_{\lambda_1}(d) \oplus \dots \oplus E_{\lambda_r}(d)$$

si $\text{Sp}(d) = \{\lambda_1, \dots, \lambda_r\}$. Car $d \circ n = n \circ d$, $E_{\lambda_i}(d)$ est stable par n pour $1 \leq i \leq r$. Car n est nilpotente sur E (et donc aussi sur $E_{\lambda_i}(d)$), il existe vecteurs $0 \neq v_i \in E_{\lambda_i}(f)$ tq $n(v_i) = 0$ pour $1 \leq i \leq r$. Donc, $f(v_i) = (d+n)(v_i) = \lambda_i v_i$, et $\lambda_i \in \text{Sp}(f)$. Car $f - \lambda_i \text{Id}_E$ preserve $E_{\lambda_i}(d)$ et est nilpotente sur $E_{\lambda_i}(d)$ (car $f - d = n$ est nilpotente), on conclut que $E_{\lambda_i}(d)$ est forcément le sous-espace caractéristique F_i de f avec valeur propre λ_i . Par la preuve d'existence, on conclut que la décomposition $f = d + n$ comme dans l'annonce est unique. $\square$

Exemple 10.15. Soit $A := \begin{pmatrix} 2 & 1 & -1 \\ 3 & 3 & -4 \\ 3 & 1 & -2 \end{pmatrix}$. On a

$$\chi_A(x) = (x + 1)(x - 2)^2.$$

Les sous-espaces caractéristiques de A sont :

$$F := \text{Ker}(A + I_3), \quad G := \text{Ker}((A - 2I_3)^2).$$

On a $\dim G = m_2 = 2$. On calcule

$$\text{Ker}(A - 2I_3) = \text{Ker}\left(\begin{pmatrix} 0 & 1 & -1 \\ 3 & 1 & -4 \\ 3 & 1 & -4 \end{pmatrix}\right) = \text{Ker}\left(\begin{pmatrix} 0 & 1 & -1 \\ 3 & 1 & -4 \\ 0 & 0 & 0 \end{pmatrix}\right) = \text{Vect}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right),$$

$$(A - 2I_3)(G) = E_2(A)$$
$$P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$
$$G = \text{Ker}((A-2I_3)^2) = \text{Ker}\left(\begin{pmatrix} 0 & 1 & -1 \\ 3 & 1 & -4 \\ 3 & 1 & -4 \end{pmatrix}\right)^2 = \text{Ker}\left(\begin{pmatrix} 0 & 0 & 0 \\ -9 & 0 & 9 \\ -9 & 0 & 9 \end{pmatrix}\right) = \text{Vect}\left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right).$$

Donc, on peut poser $u_3 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, et donc $u_2 = (A - 2I_3)u_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. (mais $u_3 := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$)

et $u_2 := (A - 2I_3)u_3 = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$ serait un choix valable aussi).

On revient à la discussion dans 7.2. Soit $I \subseteq \mathbb{R}$ une intervalle, $0 \in I$. Soit $A \in \mathcal{M}_n(\mathbb{C})$. Considerons le systeme d'equations differentielles

$$x' = Ax.$$

$$e^B := I_n + B + \frac{B^2}{2!} + \frac{B^3}{3!} + \dots = \sum_{i=0}^{\infty} \frac{B^i}{i!}$$

78

Démonstration. •) Si $P \in \text{GL}_n(\mathbb{C})$, alors e^B converge si et seulement si $P^{-1}e^BP = P^{-1}(I_n + B + \frac{B^2}{2!} + \frac{B^3}{3!} + \dots)P = I_n + P^{-1}BP + \frac{(P^{-1}BP)^2}{2!} + \frac{(P^{-1}BP)^3}{3!} + \dots = e^{P^{-1}BP}$ est convergent.

-) Si $C \in \mathcal{M}_n(\mathbb{C})$, e^B, e^C sont convergent et $BC = CB$, on vérifie que e^{BC} est convergent et

$$e^B e^C = e^{BC}.$$

-) Grâce à 10.14 existe $B = D + N$, $D \in \mathcal{M}_n(\mathbb{C})$ diagonalisable, $N \in \mathcal{M}_n(\mathbb{C})$ nilpotent et $DN = ND$. Donc, il rest deux cas :

-) $B = D$ diagonalisable. Ici, on a

$$e^{\text{Diag}(\lambda_1, \dots, \lambda_n)} = \text{Diag}(e^{\lambda_1}, \dots, e^{\lambda_n}).$$

-) $B = N$ nilpotent. Soit $k \geq 1$, tq $N^k = 0$. On a

$$e^N = I_n + N + \frac{N^2}{2!} + \dots + \frac{N^k}{k!} + 0 + \dots$$

qui est une somme convergent.

□

Théorème 10.17. L'équation $x'(t) = Ax(t)$ a comme solutions les fonctions

$$x(t) = e^{tA}x_0$$

pour $x_0 \in \mathbb{C}^n$.

Démonstration. On calcule que $t \mapsto x(t) := e^{tA}x_0$ a le dérivatif

$$t \mapsto Ae^{tA}x_0.$$

Soit $x: I \rightarrow \mathbb{C}^n$ une solution. Posons $y(t) := e^{-tA}x(t)$. Alors,

$$y'(t) = Ae^{-tA}x(t) - e^{tA}x'(t) = Ae^{-tA}x(t) - e^{-tA}Ax(t) = 0.$$

Donc, $y: I \rightarrow \mathbb{C}^n$ est une fonction constante avec valeur $x_0 \in \mathbb{C}^n$.

□

Exemple 10.18. Soit $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Alors,

$$A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A^3 = 0.$$

Donc,

$$e^{tA} = I_3 + tA + \frac{t^2 A^2}{2!} = \begin{pmatrix} 1 & t & \frac{t}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

Notons que $x'(t) = Ax(t)$ correspond (comme dans 7.9) à l'équation différentielle $f^{(2)} = 0$, qui a des solutions $t \mapsto a + bt$ pour $a, b \in \mathbb{C}$.

Remarque 10.19. 10.17 est bien pour trouver une approximation d'une solution car on peut calculer juste le premier k -term d'exponentielle.

A Appendix : Pré-requis sur des espaces vectoriels

Ce chapitre contient des rappels du programme vu en L1 dans le cours Algèbre linéaire 1, qui constitue un pré-requis pour le cours d'Algèbre linéaire 2. Il ne sera pas étudié en cours.

A.1 Rappels sur les corps $\mathbb{R}$ et $\mathbb{C}$

Dans tout ce texte, la lettre K désigne l'ensemble $\mathbb{R}$ des nombres réels ou l'ensemble $\mathbb{C}$ des nombres complexes. L'ensemble K est donc muni d'une loi d'addition, notée $+$ et d'une loi de multiplication, notée parfois $\times$, mais le plus souvent par le symbole $\cdot$. Ces deux lois vérifient les propriétés suivantes :

-) elles sont *associatives* : si $(a, b, c) \in K^3$, alors $a + (b + c) = (a + b) + c$ et $a \cdot (b \cdot c) = (a \cdot b) \cdot c$. Cela signifie que pour additionner trois nombres a , b et c , on peut commencer par additionner a et b , puis additionner $a+b$ et c ou alors commencer par additionner b et c puis additionner a et $b + c$, le résultat sera le même. On note donc ce résultat $a + b + c$ car, quelque soit la position des parenthèses, il ne change pas. De même on note $a \cdot b \cdot c$ pour $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
-) elles sont *commutatives* : si $(a, b) \in K^2$, alors $a + b = b + a$ et $a \cdot b = b \cdot a$.
-) elles possèdent un *élément neutre*. L'élément neutre pour l'addition est l'élément 0 : pour tout $a \in K$, $a + 0 = 0 + a = a$. L'élément neutre pour la multiplication est l'élément 1 : pour tout $a \in K$, $a \cdot 1 = 1 \cdot a = a$. Attention l'élément neutre de l'addition diffère de l'élément neutre de la multiplication.
-) la loi de multiplication est *distributive* par rapport à la loi d'addition, cela signifie que si $(a, b, c) \in K^3$, alors $a \cdot (b+c) = (a \cdot b) + (a \cdot c)$. Par commutativité, on a aussi $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$.
-) tout élément de K , possède un *inverse* pour la loi d'addition. En effet si $a \in K$, alors $a + (-a) = (-a) + a = 0$. Dans le cadre de la loi d'addition, l'inverse d'un élément est souvent appelé *opposé*.
-) tout élément non nul de K possède un inverse pour la loi de multiplication. En effet si $a \in K^\times := K \setminus \{0\}$, alors $a \cdot (a^{-1}) = (a^{-1}) \cdot a = 1$. Il est nécessaire de supposer a non nul car l'élément 0 n'a pas d'inverse pour la loi de multiplication : pour tout $a \in K$, on a $a \cdot 0 = 0$, on ne peut donc pas trouver d'élément $b \in K$ tel que $a \cdot b = 1$.

Exemple A.1. Lorsque $K = \mathbb{C}$, on a $i \cdot i = i^2 = -1$. Ainsi $(-i) \cdot i = (-1) \cdot i \cdot i = (-1)^2 = 1$, donc $-i$ est l'inverse de i pour la loi de multiplication dans $\mathbb{C}$.

A.2 Espaces vectoriels

Un K -espace vectoriel (ou tout simplement espace vectoriel quand le corps K est sous-entendu) est un ensemble E muni de deux lois de composition

$$\begin{array}{ccc} E \times E & \longrightarrow & E \\ (x, y) & \longmapsto & x + y \end{array} \quad \begin{array}{ccc} K \times E & \longrightarrow & E \\ (\lambda, x) & \longmapsto & \lambda \cdot x. \end{array}$$

Ces deux lois doivent vérifier les propriétés suivantes

a) (i) La loi $+$ est *associative* :

$$\forall (x, y, z) \in E \times E \times E, \quad (x + y) + z = x + (y + z).$$

(ii) Il existe un unique élément neutre $0 \in E$ pour la loi $+$:

$$\forall x \in E, \quad x + 0 = 0 + x = x.$$

(iii) Tout élément $x \in E$ possède un unique inverse $-x \in E$ pour la loi $+$:

$$x + (-x) = (-x) + x = 0.$$

(iv) La loi $+$ est *commutative* :

$$\forall (x, y) \in E \times E, \quad x + y = y + x.$$

b) La loi $\cdot$ est distributive par rapport aux lois $+$ de E et K :

$$\begin{aligned} \forall (\lambda, x, y) \in K \times E \times E, \quad \lambda \cdot (x + y) &= (\lambda \cdot x) + (\lambda \cdot y) \\ \forall (\lambda, \mu, x) \in K \times K \times E, \quad (\lambda + \mu) \cdot x &= (\lambda \cdot x) + (\mu \cdot x). \end{aligned}$$

c) La loi $\cdot$ est compatible à la multiplication dans K :

$$\forall (\lambda, \mu, x) \in K \times K \times E, \quad \lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x, \quad 1 \cdot x = x.$$

Dans la pratique, on écrit très souvent λx pour $\lambda \cdot x$.

Les éléments de l'ensemble E sont appelés *vecteurs* et les éléments du corps K sont appelés *scalaires*. Il faut bien prendre garde au fait que les lois $+$ et $\cdot$ sont définies sur des ensembles différents. La loi $+$ part de deux vecteurs et produit un vecteur alors que la loi $\cdot$ part d'un scalaire et d'un vecteur et produit un vecteur. Si λ est un scalaire et v un vecteur, il faut imaginer le vecteur $\lambda \cdot v$ comme étant le vecteur v dilaté au moyen du coefficient λ .

Exemple A.2. (i) Soit E un K -espace vectoriel. Si $x \in E$, on a $0 \cdot x + 0 \cdot x = (0 + 0) \cdot x = 0 \cdot x$. On en déduit $0 \cdot x = 0$ pour tout $x \in E$.

(ii) D'après la loi de distributivité, pour $x \in E$, on a $x + (-1)x = (1 - 1)x = 0x = 0$. Ainsi $(-1)x$ est l'inverse de x pour la loi d'addition et on a donc $(-1)x = -x$.

Exemple A.3. Si $n \geq 1$ est un entier, l'ensemble K^n des n -uplets d'éléments de K est un espace vectoriel pour les lois suivantes. Si $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ sont deux éléments de K^n et $\lambda \in K$, on définit

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n), \quad \lambda \cdot (x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n).$$

Si $n = 1$, l'espace vectoriel K^n est juste le corps K et la loi $\cdot$ s'identifie alors à la loi de multiplication dans K .

Exemple A.4. Notons $K^{\mathbb{N}}$ l'ensemble des suites $(x_n)_{n \geq 0}$ d'éléments de K . Il s'agit d'un K -espace vectoriel pour les opérations suivantes

$$(x_n)_{n \geq 0} + (y_n)_{n \geq 0} := (x_n + y_n)_{n \geq 0}, \quad \lambda \cdot (x_n)_{n \geq 0} := (\lambda x_n)_{n \geq 0}.$$

Exemple A.5. Notons $\mathcal{F}(\mathbb{R}, K)$ l'ensemble des fonctions de $\mathbb{R}$ dans K . Il s'agit d'un espace vectoriel, l'addition de deux fonctions est la loi d'addition usuelle sur les fonctions : si f et g sont deux éléments de $\mathcal{F}(\mathbb{R}, K)$, la fonction $f + g$ est la fonction prenant la valeur $f(x) + g(x)$ en x . De même on définit la multiplication d'une fonction f par un scalaire λ en définissant $\lambda \cdot f$ comme étant la fonction prenant la valeur $\lambda f(x)$ en x . On peut encore écrire ceci $(\lambda \cdot f)(x) := \lambda f(x)$.

A.3 Sous-espaces vectoriels

Définition A.6. Soit E un K -espace vectoriel. Un sous-espace vectoriel K de E (ou simplement sous-espace vectoriel lorsque K est sous-entendu) est une partie F de E vérifiant

(i) $0 \in F$;

(ii) pour tous x et y dans F et λ dans K , on a $x + \lambda y \in F$.

Exemple A.7. Dans $\mathbb{R}^3$, l'ensemble des éléments (x_1, x_2, x_3) vérifiant la relation $2x_1 + 3x_2 + 5x_3 = 0$ est un sous-espace vectoriel. Il s'agit d'un exemple de *sous-espace vectoriel défini par une équation*.

Exemple A.8. Dans $\mathbb{R}^3$ posons $v = (1, 2, 3)$ et $w = (2, 7, 3)$. Alors l'ensemble

$$\{\lambda v + \mu w \mid (\lambda, \mu) \in \mathbb{R}^2\} = \{\lambda + 2\mu, 2\lambda + 7\mu, 3\lambda + 3\mu \mid (\lambda, \mu) \in \mathbb{R}^2\}$$

est un sous-espace vectoriel de $\mathbb{R}^3$. Il s'agit d'un exemple de sous-espace vectoriel engendré par une famille de vecteurs, ici les vecteurs v et w .

On peut effectuer des opérations sur les sous-espaces vectoriels afin d'en produire de nouveaux.

Si E est un espace vectoriel et si F_1 et F_2 sont deux sous-espaces vectoriels, l'intersection $F_1 \cap F_2$ est un sous-espace vectoriel de E . L'ensemble

$$F_1 + F_2 = \{v + w \mid (v, w) \in F_1 \times F_2\}$$

est un sous-espace vectoriel de E appelé *somme* des sous-espaces F_1 et F_2 . Plus généralement, si $F_1, \dots, F_n$ sont des sous-espaces vectoriels, la partie

$$F_1 + \dots + F_n = \{x_1 + \dots + x_n \mid (x_1, \dots, x_n) \in F_1 \times \dots \times F_n\}$$

est un sous-espace vectoriel de E appelé *somme* de la famille de sous-espaces $(F_1, \dots, F_n)$.

Si $v \in E$, on note Kv l'ensemble $\{\lambda v \mid \lambda \in K\}$. C'est un sous-espace vectoriel de E appelé *droite vectorielle* engendrée par v . Si $(v_1, \dots, v_n)$ est une famille finie d'éléments de E , le *sous-espace vectoriel engendré par* $(v_1, \dots, v_n)$ est noté $\text{Vect}(v_1, \dots, v_n)$ et défini par

$$\text{Vect}(x_1, \dots, x_n) := Kx_1 + \dots + Kx_n = \{\lambda_1 v_1 + \dots + \lambda_n v_n \mid (\lambda_1, \dots, \lambda_n) \in K^n\}.$$

Remarquons que dans l'exemple A.8, le sous-espace vectoriel peut être défini comme $\text{Vect}(v, w)$.

Soit E un K -espace vectoriel, ainsi que F_1 et F_2 deux sous-espaces vectoriels de E . Par définition, tout élément de $F_1 + F_2$ s'écrit sous la forme $v + w$ avec $v \in F_1$ et $w \in F_2$. Cette écriture n'est pas toujours unique. Considérons par exemple le cas de $E = \mathbb{R}^3$ avec $F_1 = \text{Vect}((1, 2, 3), (2, 1, 3))$ et $F_2 = \text{Vect}((2, 7, 3), (1, 8, 3))$. Alors on peut écrire

$$(3, 9, 6) = (1, 2, 3) + (2, 7, 3) = (2, 1, 3) + (1, 8, 3).$$

Cependant lorsque cette décomposition est unique, on dit que les sous-espaces F_1 et F_2 sont en *somme directe*. Plus généralement, on peut définir la notion de somme directe pour une famille finie de sous-espaces vectoriels.

Définition A.9. Soit E un espace vectoriel. Si $F_1, \dots, F_n$ sont des sous-espaces vectoriels de E , on dit qu'ils sont en somme directe si tout élément $x \in F_1 + \dots + F_n$ s'écrit de façon unique sous la forme $x = v_1 + \dots + v_n$ avec $v_i \in F_i$ pour $1 \leq i \leq n$. Lorsque tel est le cas, on note également $F_1 \oplus \dots \oplus F_n$ le sous-espace vectoriel $F_1 + \dots + F_n$.

Pour vérifier que des sous-espaces vectoriels sont en somme directe, on peut utiliser le critère suivant.

Proposition A.10. *Les sous-espaces $F_1 + \dots + F_n$ sont en somme directe si et seulement si pour tout $(v_1, \dots, v_n) \in F_1 \times \dots \times F_n$, l'égalité $v_1 + \dots + v_n = 0$ implique $v_1 = v_2 = \dots = v_n = 0$.*

Démonstration. Le sens $\Rightarrow$ est immédiat, il s'agit juste d'appliquer la définition d'une somme directe à la décomposition de l'élément 0. Montrons que si l'égalité $v_1 + \dots + v_n = 0$ implique $v_1 = v_2 = \dots = v_n = 0$ pour tout $(v_1, \dots, v_n) \in F_1 \times \dots \times F_n$, alors les sous-espaces $F_1, \dots, F_n$ sont en somme directe. Il faut donc prouver que si $(v_1, \dots, v_n) \in F_1 \times \dots \times F_n$ et si $(w_1, \dots, w_n) \in F_1 \times \dots \times F_n$ vérifient

$$v_1 + \dots + v_n = w_1 + \dots + w_n, \quad (3)$$

alors $v_1 = w_1, \dots, v_n = w_n$. Pour cela, réécrivons l'égalité (3) sous la forme

$$(v_1 - w_1) + (v_2 - w_2) + \dots + (v_n - w_n) = 0.$$

L'hypothèse implique alors $v_1 - w_1 = \dots = v_n - w_n = 0$, c'est-à-dire $v_1 = w_1, \dots, v_n = w_n$. $\square$

Le cas de deux sous-espaces vectoriels est particulier.

Proposition A.11. *Soit E un K -espace vectoriels et soient F_1 et F_2 deux sous-espaces vectoriels de E . Alors F_1 et F_2 sont en somme directe si et seulement si $F_1 \cap F_2 = 0$.*

Démonstration. Supposons tout d'abord que F_1 et F_2 sont en somme directe. Soit $v \in F_1 \cap F_2$. Alors $0 = v - v = 0 - 0$. Comme F_1 et F_2 sont en somme directe, une telle décomposition est unique, ainsi $v = 0$ et donc $F_1 \cap F_2 = 0$. Réciproquement supposons $F_1 \cap F_2 = 0$. Soit $(v_1, v_2) \in F_1 \times F_2$ tel que $v_1 + v_2 = 0$. On a alors $v_1 = -v_2 \in F_2$, donc $v_1 \in F_1 \cap F_2$ et donc $v_1 = 0$. On en déduit $v_2 = 0$. Ainsi la proposition A.10 implique que F_1 et F_2 sont en somme directe. $\square$

Remarque A.12. Attention, la proposition A.11 ne se généralise pas verbatim au cas de n sous-espaces vectoriels avec $n \geq 3$. Considérons par exemple le cas de $E = \mathbb{R}^2$, $F_1 = \mathbb{R}(1, 0)$, $F_2 = \mathbb{R}(0, 1)$ et $F_3 = \mathbb{R}(1, 1)$. On vérifie que $F_1 \cap F_2 = F_2 \cap F_3 = F_1 \cap F_3 = 0$ (en particulier $F_1 \cap F_2 \cap F_3 = 0$). Pourtant ces trois sous-espaces ne sont pas en somme directe puisque

$$(1, 1) = 0 + 0 + (1, 1) = (1, 0) + (0, 1) + 0.$$

Lemme A.13. *Soit E un K -espace vectoriel et soient $F_1, \dots, F_n \subseteq E$ des espaces vectoriels de E . Alors, $\sum_{i=1}^n F_i = \bigoplus_{i=1}^n F_i$ ssi $\forall i = 2, \dots, n$ on a $(F_1 + \dots + F_{i-1}) \cap F_i = \{0\}$.*

Démonstration. $\bullet \Rightarrow (F_1 + \dots + F_{i-1}) \cap F_i = (F_1 \oplus \dots \oplus F_{i-1}) \cap F_i = \{0\}$.

-) $\Leftrightarrow$ Réccurence sur n : Cas $n = 1$ ✓. Donc supposons vrai pour $n \geq 1$. Soient $v_i, w_i \in F_i$, $i = 1, \dots, n$. Supposons $v_1 + \dots + v_n = w_1 + \dots + w_n$. Then $(v_1 - w_1) + \dots + (v_{n-1} - w_{n-1}) = (v_n - w_n) \in (F_1 + \dots + F_{n-1}) \cap F_n = \{0\}$. Donc, $v_n = w_n$ et $v_1 + \dots + v_{n-1} = w_1 + \dots + w_{n-1}$. Par récurrence, $v_i = w_i$ pour $i = 1, \dots, n - 1$.

□

A.4 Familles génératrices, familles libres, bases

Soit E un espace vectoriel.

Définition A.14. On dit que E est de dimension finie s'il existe une famille finie $(x_1, \dots, x_n)$ d'éléments de E telle que $E = \text{Vect}(x_1, \dots, x_n)$. Une telle famille est appelée famille génératrice de E .

Exemple A.15. Si $E = K^n$, pour $1 \leq i \leq n$, notons e_i l'élément $(0, \dots, 0, \underbrace{1}_i, 0, \dots, 0)$.

Tout élément $(x_1, \dots, x_n)$ de E peut s'écrire

$$(x_1, \dots, x_n) = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$$

donc $K^n = \text{Vect}(e_1, \dots, e_n)$ ce qui montre que K^n est un espace vectoriel de dimension finie.

Dans l'exemple précédent, on a envie de dire que K^n est en fait de dimension n . Pour définir correctement la notion de dimension, il est nécessaire d'introduire les notions de familles libres et de bases.

Définition A.16. Soit $(v_1, \dots, v_n)$ une famille finie d'éléments de E . On dit que la famille $(v_1, \dots, v_n)$ est libre si pour tout n -uplet de scalaires $(\lambda_1, \dots, \lambda_n) \in K^n$, on a

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0 \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

Autrement dit, la seule combinaison linéaire nulle des vecteurs $v_1, \dots, v_n$ est la combinaison dont les coefficients sont nuls.

Exemple A.17. Si $E = K^n$, la famille $(e_1, \dots, e_n)$ est libre. En effet, supposons que $\lambda_1 e_1 + \dots + \lambda_n e_n = 0$, alors

$$(\lambda_1, \dots, \lambda_n) = \lambda_1 e_1 + \dots + \lambda_n e_n = 0 = (0, \dots, 0)$$

donc $\lambda_1 = \dots = \lambda_n = 0$.

Définition A.18. Une base de E est une famille finie d'éléments de E qui est à la fois libre et génératrice.

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et soit $v \in E$. Comme la famille $\mathcal{B}$ est génératrice, il existe des scalaires $x_1, \dots, x_n$ tels que $v = x_1 e_1 + \dots + x_n e_n$. Comme de plus la famille $\mathcal{B}$ est libre, les scalaires $x_1, \dots, x_n$ sont uniquement déterminés par v . On les appelle les *coordonnées* de v dans la base $\mathcal{B}$. Nous avons donc démontré le résultat suivant.

Proposition A.19. *Soit E un K -espace vectoriel de dimension finie et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Pour tout $v \in E$, il existe un unique n -uplet $(x_1, \dots, x_n) \in K^n$ tel que*

$$v = x_1 e_1 + x_2 e_2 + \dots + x_n e_n.$$

Exemple A.20. Si $E = K^n$, notons $\mathcal{B} = (e_1, e_2, \dots, e_n)$ la famille des vecteurs élémentaires (voir exemple A.15). Il s'agit d'une famille génératrice et libre, donc il s'agit d'une base de K^n . Si $v = (x_1, \dots, x_n) \in K^n$, on a

$$(x_1, x_2, \dots, x_n) = x_1 e_1 + \dots + x_n e_n$$

donc les scalaires $x_1, \dots, x_n$ sont les coordonnées du vecteur $(x_1, \dots, x_n)$ dans la base $\mathcal{B}$. Pour cette raison, on lui donne le nom de *base canonique* de K^n . Mais attention, il existe bien d'autres bases dans K^n ! (voir les feuilles d'exercices)

Exemple A.21. Soit $n \geq 0$ et soit $K[X]$ l'ensemble des applications polynômes de K dans K , c'est-à-dire des applications de K dans K de la forme $x \mapsto \sum_{i=0}^n a_i x_i$, pour $a_0, \dots, a_n \in K$ fixé. Il s'agit d'un sous-espace vectoriel du K -espace vectoriel $\mathcal{F}(K, K)$. Si $n \geq 0$, on note $K_n[X]$ le sous-espace des applications polynômes de degré inférieur ou égal à n . Si on note X^i l'application polynôme $x \mapsto x^i$, la famille $(1, X, X^2, \dots, X^n)$ est une base de $K_n[X]$.

Le corollaire suivant est une reformulation de la proposition A.19

Corollaire A.22. *Soit E un K -espace vectoriel de dimension finie et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Alors l'application*

$$\begin{array}{ccc} K^n & \longrightarrow & E \\ (x_1, \dots, x_n) & \longmapsto & x_1 e_1 + \dots + x_n e_n \end{array}$$

est une bijection (en fait une isomorphisme d'espaces vectoriels).

Le théorème qui va suivre est très utile pour construire des bases.

Théorème A.23. *Soit E un K -espace vectoriel de dimension finie. Soit $\mathcal{B}_1 = (e_1, \dots, e_r)$ une famille libre de E et soit $\mathcal{B}_2 = (e_1, \dots, e_r, e_{r+1}, \dots, e_s)$ une famille génératrice de E telle que $s \geq r$. Alors il existe un entier $0 \leq k \leq s-r$ et des entiers $i_1 < \dots < i_k$ compris entre $r+1$ et s tels que la famille $(e_1, \dots, e_r, e_{i_1}, \dots, e_{i_k})$ soit une base de E . En d'autres termes, d'une famille génératrice, on peut toujours extraire une base contenant une sous-famille libre donnée.*

Démonstration. Voir le cours de L1. □

Corollaire A.24. *Dans un espace vectoriel de dimension finie, toute famille génératrice contient une sous-famille qui est une base.*

Démonstration. On peut supposer $E \neq 0$. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une famille génératrice de E . Comme cette famille est génératrice, elle contient au moins un élément non nul. Quitte à permuter les éléments de cette famille on peut supposer que $e_1 \neq 0$. La famille réduite à e_1 est alors libre. Le théorème A.23 implique alors qu'il existe une sous-famille de $\mathcal{B}$ qui est une base de E . □

Corollaire A.25. *Dans un espace vectoriel de dimension finie, on peut toujours compléter une famille libre en une base.*

Démonstration. Soit $\mathcal{B} = (e_1, \dots, e_r)$ une famille libre de E . Comme E est de dimension finie, il possède une famille génératrice $(f_1, \dots, f_s)$. La famille $\mathcal{B}' = (e_1, \dots, e_r, f_1, \dots, f_s)$ est alors aussi une famille génératrice et on peut appliquer le théorème A.23 pour prouver qu'il existe une sous-famille de $\mathcal{B}'$ qui est une base et qui complète $\mathcal{B}$. □

Théorème A.26. *Tout K -espace vectoriel E de dimension finie possède une base. De plus toutes les bases de E ont le même nombre d'éléments.*

Démonstration. L'existence d'une base est une conséquence immédiate du corollaire A.24 puisqu'un espace vectoriel de dimension finie contient au moins une famille génératrice. La démonstration du fait que toutes les bases ont le même nombre d'éléments est plus délicate. Nous allons la décomposer en plusieurs étapes.

Lemme A.27. *Soit E un K -espace vectoriel de dimension finie et soit $\mathcal{B} = (e_1, \dots, e_n)$ une famille génératrice de E . Soit $v \in E$ et supposons que $v \notin \text{Vect}(e_1, \dots, e_r)$. Alors il existe $r + 1 \leq i \leq n$ tel que $(e_1, \dots, e_{i-1}, v, e_{i+1}, \dots, e_n)$ est une famille génératrice de E .*

Démonstration. Comme la famille $\mathcal{B}$ est génératrice, il existe $(\lambda_1, \dots, \lambda_n) \in K^n$ tel que $v = \lambda_1 e_1 + \dots + \lambda_n e_n$. Comme $v \notin \text{Vect}(e_1, \dots, e_r)$, il existe $r + 1 \leq i \leq n$ tel que $\lambda_i \neq 0$. On peut alors écrire

$$e_i = \lambda_i^{-1}(v - \lambda_1 e_1 - \dots - \lambda_{i-1} e_{i-1} - \lambda_{i+1} e_{i+1} - \dots - \lambda_n e_n).$$

Ainsi $e_i \in \text{Vect}(e_1, \dots, e_{i-1}, v, e_{i+1}, \dots, e_n)$. En conséquence $E = \text{Vect}(e_1, \dots, e_n) \subset \text{Vect}(e_1, \dots, e_{i-1}, v, e_{i+1}, \dots, e_n)$ et donc la famille $(e_1, \dots, e_{i-1}, v, e_{i+1}, \dots, e_n)$ est une famille génératrice de E . □

Soit $\mathcal{B} = (e_1, \dots, e_n)$ et soit $\mathcal{B}' = (f_1, \dots, f_m)$ deux bases de E . Supposons par l'absurde que $m > n$. En appliquant le lemme A.27, on construit une famille génératrice de E à n termes dont le premier terme est f_1 et les vecteurs suivants proviennent de $\mathcal{B}$. Comme $\mathcal{B}'$ est une base, la famille (f_1, f_2) est libre donc $f_2 \notin \text{Vect}(f_1)$. Le lemme A.27 permet alors de construire une famille génératrice à n termes dont les deux premiers termes sont f_1, f_2 et les suivants des vecteurs de $\mathcal{B}$. En itérant ce procédé, on obtient une famille génératrice à n termes constituée des vecteurs $(f_1, \dots, f_n)$. Mais alors $f_{n+1} \in \text{Vect}(f_1, \dots, f_n)$, ce qui contredit la liberté de la famille $\mathcal{B}'$. $\square$

Si E est un K -espace vectoriel de dimension finie. La longueur d'une base de E est appelée *dimension* de E . On la note $\dim_K E$.

Exemple A.28. L'exemple A.20 montre que la famille des vecteurs élémentaires $e_1, \dots, e_n$ est une base de K^n , on en conclut que $\dim_K K^n = n$.

Exemple A.29. Dans l'exemple A.21, on a vu que la base $(1, X, X^2, \dots, X^n)$ est une base du K -espace vectoriel $K_n[X]$. On en conclut que $\dim_K K_n[X] = n + 1$.

Corollaire A.30. Soit E un K -espace vectoriel de dimension finie. Si $(e_1, \dots, e_n)$ est une famille libre de E , alors $n \leq \dim_K E$.

Démonstration. D'après le corollaire A.25, on peut compléter la famille $(e_1, \dots, e_n)$ en une base de E . Le théorème A.26 implique alors que $\dim_K E \geq n$. $\square$

Corollaire A.31. Dans un espace vectoriel de dimension n , toute famille constituée de m vecteurs avec $m > n$ est liée.

Démonstration. Il s'agit d'une reformulation du corollaire A.30. $\square$

Corollaire A.32. Soit E un K -espace vectoriel de dimension finie et soit F un sous-espace vectoriel de E . Alors F est de dimension finie et $\dim_K F \leq \dim_K E$. De plus on a $\dim_K F = \dim_K E$ si et seulement si $F = E$.

Corollaire A.33. Soit E un K -espace vectoriel de dimension finie et soit $\mathcal{B} = (e_1, \dots, e_n)$ une famille de vecteurs de E .

- (i) Supposons $\mathcal{B}$ libre. Alors $\mathcal{B}$ est une base de E si et seulement si $n = \dim_K E$.
- (ii) Supposons $\mathcal{B}$ génératrice. Alors $\mathcal{B}$ est une base de E si et seulement si $n = \dim_K E$.

A.5 Dimension et sommes directes

Proposition A.34. Soit E un espace vectoriel et soient F et G deux sous-espaces vectoriels de E . Si F et G sont de dimension finie, alors $F + G$ et $F \cap G$ sont de dimension finie et on a

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

Démonstration. Voir le cours de L1. □

Proposition A.35. Soit $(F_1, \dots, F_n)$ une famille de sous-espaces vectoriels de dimension finie de E . Si ces sous-espaces sont en somme directe, alors pour toute base $\mathcal{B}_1$ de F_1 , $\mathcal{B}_2$ de $F_2, \dots, \mathcal{B}_n$ de F_n , la famille $\mathcal{B}$ obtenue en concaténant $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_n$ est une famille libre de E . Si de plus $F_1 \oplus \dots \oplus F_n = E$, alors $\mathcal{B}$ est une base de E .

Démonstration. Voir le cours de L1. □

Proposition A.36. Soient $F_1, \dots, F_n$ des sous-espaces vectoriels de dimension finie de E . Ces sous-espaces sont en somme directe si et seulement si

$$\dim(F_1 + \dots + F_n) = \sum_{i=1}^n \dim F_i.$$

Démonstration. Voir le cours de L1. □

A.6 Applications linéaires

Définition A.37. Soient E et F deux espaces vectoriels. Une application linéaire de E dans F est une application f de E dans F telle que

$$\forall (\lambda, x, y) \in K \times E \times E, \quad f(x + \lambda y) = f(x) + \lambda f(y).$$

Si E et F sont deux espaces vectoriels, on note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F .

Lorsque $E = F$, on note $\mathcal{L}(E) = \mathcal{L}(E, E)$. Un élément de $\mathcal{L}(E)$ est appelé *endomorphisme* de E .

Exemple A.38. (i) L'application $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ définie par $f(x, y) = (x + y, y + 2x)$ est linéaire.

(ii) Plus généralement, étant donnés des nombres réels $a_1, \dots, a_n$, l'application $f : \mathbb{R}^n \rightarrow \mathbb{R}$ définie par $f(x_1, \dots, x_n) = a_1 x_1 + \dots + a_n x_n$ est linéaire.

(iii) Encore plus généralement, une application $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ s'écrit sous la forme

$$f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), f_2(x_1, \dots, x_n), \dots, f_p(x_1, \dots, x_n)).$$

Chaque f_i est une application de $\mathbb{R}^n \rightarrow \mathbb{R}$ et s'appelle la *i-ème application coordonnée* de f . On vérifie facilement que f est linéaire si et seulement si toutes ses applications coordonnées f_i sont linéaires.

A.7 Noyau, image et rang

Définition A.39. Si E et F sont deux espaces vectoriels et si f est une application linéaire de E dans F , on note $\text{Im } f$ l'image $f(E)$ de f et $\text{Ker } f$ le noyau de f , c'est-à-dire l'ensemble $f^{-1}(\{0\}) = \{x \in E, f(x) = 0\}$.

On vérifiera que $\text{Ker } f$ est un sous-espace vectoriel de E et que $\text{Im } f$ est un sous-espace vectoriel de F .

Proposition A.40. Si f est une application linéaire entre deux espaces vectoriels, alors f est injective si et seulement si $\text{Ker } f = 0$.

Démonstration. Si f est injective, alors $\text{Ker } f$ est réduit à 0. Réciproquement supposons que $\text{Ker } f = \{0\}$. Si x et y sont deux éléments de E tels que $f(x) = f(y)$, alors $f(x) - f(y) = 0$. Par linéarité de f , on a alors $f(x - y) = 0$ et donc $x - y \in \text{Ker } f$. Ainsi on doit avoir $x - y = 0$, c'est-à-dire $x = y$. On a prouvé que l'application f est injective. $\square$

La dimension de l'espace vectoriel $\text{Im } f$ est appelé le *rang* de f et est notée $\text{rg } f$. Ainsi on a, par définition, $\text{rg } f = \dim \text{Im } f$.

Théorème A.41 (Théorème du rang). Soit E un espace vectoriel de dimension finie et soit F un espace vectoriel. Si f est une application linéaire de E dans F , alors $\text{Im } f$ est un sous-espace vectoriel de dimension finie de F et on a

$$\dim \text{Im } f = \dim E - \dim \text{Ker } f.$$

Démonstration. L'espace vectoriel E est de dimension finie. D'après le théorème A.26, il possède au moins une base. Soit donc $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Nous allons montrer que la famille $(f(e_1), \dots, f(e_n))$ est une famille génératrice de $\text{Im } f$. Tout d'abord, on a bien $f(e_i) \in \text{Im } f$ pour tout $1 \leq i \leq n$. Soit $w \in \text{Im } f$. Par définition de $\text{Im } f$, il existe $v \in E$ tel que $f(v) = w$. Comme $\mathcal{B}$ est une base de E , il existe des scalaires $x_1, \dots, x_n \in K$ tels que $v = x_1 e_1 + \dots + x_n e_n$. L'application f étant linéaire, on a

$$w = f(v) = x_1 f(e_1) + \dots + x_n f(e_n).$$

Ainsi w est combinaison linéaire des vecteurs $f(e_1), \dots, f(e_n)$. Ceci étant vrai pour tout choix de $w \in \text{Im } f$, la famille $(f(e_1), \dots, f(e_n))$ est une famille génératrice de $\text{Im } f$. On en conclut donc que $\text{Im } f$ est un espace vectoriel de dimension finie.

D'après le corollaire A.24, on peut extraire de la famille $(f(e_1), \dots, f(e_n))$ une base de $\text{Im } f$. Soit $r := \dim_K \text{Im } f$. Il existe donc des indices $1 \leq i_1 < \dots < i_r \leq n$ tels que $(f(e_{i_1}), \dots, f(e_{i_r}))$ est une base de $\text{Im } f$. Comme r est le rang de f , il nous reste donc à démontrer que

$$n = \dim_K E = r + \dim_K \text{Ker } f.$$

Soit $W := \text{Vect}(e_{i_1}, \dots, e_{i_r})$ le sous-espace vectoriel de E engendré par les vecteurs $e_{i_1}, \dots, e_{i_r}$. Montrons que la famille $(e_{i_1}, \dots, e_{i_r})$ est libre. Supposons qu'il existe $(\lambda_1, \dots, \lambda_r) \in K^r$ tel que

$$\lambda_1 e_{i_1} + \dots + \lambda_r e_{i_r} = 0.$$

Alors, par linéarité de f , on a

$$\lambda_1 f(e_{i_1}) + \dots + \lambda_r f(e_{i_r}) = f(\lambda_1 e_{i_1} + \dots + \lambda_r e_{i_r}) = f(0) = 0.$$

En utilisant le fait que $(f(e_{i_1}), \dots, f(e_{i_r}))$ est une famille libre de F , on en conclut que $(\lambda_1, \dots, \lambda_r) = (0, \dots, 0)$ et donc que $(e_{i_1}, \dots, e_{i_r})$ est une famille libre de E . Comme par ailleurs, $(e_{i_1}, \dots, e_{i_r})$ est, par définition de W , une famille génératrice de W , il s'agit en fait d'une base de W et donc $\dim_K W = r$.

Nous allons à présent montrer que $E = \text{Ker } f \oplus W$. Commençons par vérifier que $\text{Ker } f$ et W sont en somme directe. Soit $v \in \text{Ker } f \cap W$. D'une part $v \in W$, donc il existe $(\lambda_1, \dots, \lambda_r) \in K^r$ tel que $v = \lambda_1 e_{i_1} + \dots + \lambda_r e_{i_r}$. D'autre part, $v \in \text{Ker } f$, donc $f(v) = 0$. Ceci nous donne

$$\lambda_1 f(e_{i_1}) + \dots + \lambda_r f(e_{i_r}) = 0.$$

Comme $(f(e_{i_1}), \dots, f(e_{i_r}))$ est une famille libre, on en conclut que $(\lambda_1, \dots, \lambda_r) = (0, \dots, 0)$ et donc que $v = 0$. Ainsi, d'après la proposition A.11, les sous-espaces $\text{Ker } f$ et W sont en somme directe. Il reste donc à prouver que $E = \text{Ker } f + W$. Soit $v \in E$ un vecteur de E . Il existe $(\lambda_1, \dots, \lambda_r) \in K^r$ tel que

$$f(v) = \lambda_1 f(e_{i_1}) + \dots + \lambda_r f(e_{i_r}).$$

Posons alors $w := \lambda_1 e_{i_1} + \dots + \lambda_r e_{i_r}$. On a $w \in W$. De plus $f(v - w) = 0$, donc $v - w \in \text{Ker } f$. On a donc bien décomposé $v = (v - w) + w$ prouvant que $E = \text{Ker } f + W$.

On conclut donc de la proposition A.36 que $\dim_K E = \dim_K \text{Ker } f + \dim_K W$. $\square$

Exemple A.42. (i) Considérons l'application linéaire $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ définie par

$$f(x, y) = (x + y, 2x + y).$$

Déterminons son noyau. Un élément $(x, y) \in \mathbb{R}^2$ est dans le noyau de f si et seulement si il est solution du système d'équations linéaires homogène

$$\begin{cases} x + y &= 0 \\ 2x + y &= 0 \end{cases} \Leftrightarrow \begin{cases} x + y &= 0 \\ x &= 0 \end{cases} \Leftrightarrow \begin{cases} x &= 0 \\ y &= 0. \end{cases}$$

Ainsi $\text{Ker } f = 0$ et l'application f est injective. Le théorème du rang (théorème A.41) nous donne alors $\text{rg } f = 2$. Ainsi $\dim \text{Im } f = \dim \mathbb{R}^2$, ce qui implique $\text{Im } f = \mathbb{R}^2$ d'après le corollaire A.32. L'application f est donc surjective. Étant injective et surjective, elle est bijective.

(ii) Considérons à présent l'application $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie par

$$f(x, y, z) = (x + y + z, 2x + y + z, y + z).$$

Déterminons son noyau. Il faut résoudre le système

$$\begin{cases} x + y + z = 0 \\ 2x + y + z = 0 \\ y + z = 0 \end{cases} \Leftrightarrow \begin{cases} x + y + z = 0 \\ x = 0 \\ y = -z \end{cases}$$

Le noyau de f est donc le sous-espace vectoriel constitué des vecteurs de la forme $(0, y, -y)$, c'est-à-dire la droite vectoriel $\mathbb{R}(0, 1, -1)$. Ainsi $\dim \text{Ker } f = 1$. Le théorème du rang implique alors que $\dim \text{Im } f = 2$. Cette fois-ci, on a $\text{Im } f \subsetneq \mathbb{R}^3$, l'application f n'est ni injective ni surjective.

A.8 Opérations sur les applications linéaires

Soient E et F deux espaces vectoriels. Si f et g sont deux applications linéaires de E dans F , on note $f + g$ l'application de E dans F définie par

$$\forall x \in E, (f + g)(x) = f(x) + g(x).$$

Il s'agit d'une application linéaire de E dans F . De même si $\lambda \in K$, on note $\lambda \cdot f$ l'application de E dans F définie par

$$\forall x \in E, (\lambda \cdot f)(x) = \lambda \cdot f(x).$$

Il s'agit encore d'une application linéaire de E dans F . Muni des opérations $+$ et $\cdot$ définies ci-dessus, l'ensemble $\mathcal{L}(E, F)$ est un espace vectoriel.

Si E , F et G sont trois espaces vectoriels, et si $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$, l'application composée $g \circ f$ est une application linéaire de E dans G .

Théorème A.43. *Soient E et F deux espaces vectoriels. Supposons que E est de dimension finie et soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Alors pour toute famille $(f_1, \dots, f_n)$ de $n = \dim_K E$ éléments de F , il existe une unique application linéaire u de E dans F telle que, pour tout $1 \leq i \leq n$, on a $u(e_i) = f_i$. Si de plus la famille $(f_1, \dots, f_n)$ est libre, l'application u est injective. Si de plus la famille $(f_1, \dots, f_n)$ est une base de F , l'application u est bijective.*

Démonstration. Soit $v \in E$ et soit $(x_1, \dots, x_n)$ le système de coordonnées de v dans la base $\mathcal{B}$. On a donc $v = x_1 e_1 + \dots + x_n e_n$. On pose alors

$$u(v) := x_1 f_1 + \dots + x_n f_n.$$

On obtient ainsi une application u de E dans F . On vérifie facilement qu'elle est linéaire. Elle vérifie également la propriété $u(e_i) = f_i$ pour $1 \leq i \leq n$. Vérifions que c'est la seule application linéaire ayant cette propriété. Supposons en effet que u' soit une application linéaire de E dans F tel que $u'(e_i) = f_i$ pour $1 \leq i \leq n$. Si $v \in E$ et si $(x_1, \dots, x_n)$ est le système de coordonnées de v dans la base $\mathcal{B}$, on a

$$u'(v) = u'(x_1 e_1 + \dots + x_n e_n) = x_1 u'(e_1) + \dots + x_n u'(e_n) = x_1 f_1 + \dots + x_n f_n = u(v).$$

Comme $u'(v) = u(v)$ pour tout $v \in E$, on a bien $u' = u$. Ceci prouve l'unicité de u .

Le reste est laissé en exercice. $\square$

A.9 Matrices

Définition A.44. Soient $n \geq 1$ et $p \geq 1$ deux entiers. Une matrice de taille $n \times p$ à coefficients dans K est un tableau rectangulaire A ayant n lignes et p colonnes contenant des éléments de A . On note $a_{i,j}$ ses coefficients et on les indexe de la façon suivante

$$A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,p} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,p} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,p} \end{pmatrix}.$$

On note $\mathcal{M}_{n,p}(K)$ l'ensemble des matrices à n lignes et p colonnes. C'est un K -espace vectoriel pour les opérations suivantes

-) $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}, B = (b_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$, alors $C = A + B$ où $C = (a_{i,j} + b_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$.
-) $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}, \lambda \in K$, alors $\lambda \cdot A = (\lambda a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$.

Le K -espace vectoriel $\mathcal{M}_{n,p}(K)$ est de dimension finie. Une base de $\mathcal{M}_{n,p}(K)$ est donnée par la famille de matrices $E_{i,j}$ pour $1 \leq i \leq n$ et $1 \leq j \leq p$. La matrice $E_{i,j}$ est la matrice dont toutes les entrées sont nulles sauf l'entrée à l'intersection de la i -ème ligne et de la j -ème colonne qui vaut 1. On en déduit que

$$\dim_K \mathcal{M}_{n,p}(K) = np.$$

Si $n = p$, une matrice de $\mathcal{M}_n(K) = \mathcal{M}_{n,p}(K)$ est appelée *matrice carrée* de taille n .

Si $A \in \mathcal{M}_{n,p}(K)$ et $B \in \mathcal{M}_{p,q}(K)$, on définit le *produit* de A avec B par

$$AB = (c_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq q}} \in \mathcal{M}_{n,q}(K)$$

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

Proposition A.45. Si $A \in \mathcal{M}_{n,p}(K)$, $B \in \mathcal{M}_{p,q}(K)$ et $C \in \mathcal{M}_{q,r}(K)$, on a

-) $A(BC) = (AB)C$;
-) $A(B + C) = (AB) + (AC)$;
-) $(A + B)C = (AC) + (BC)$;
-) si $\lambda \in K$, on a $\lambda(AB) = (\lambda A)B = A(\lambda B)$.

Pour $n \geq 1$, on identifie l'espace vectoriel K^n à l'espace des matrices $\mathcal{M}_{n,1}(K)$ via l'opération $(x_1, \dots, x_n) \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Si $n \geq 1$, on note I_n la matrice $(a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ de $\mathcal{M}_n(K)$ définie par $a_{i,j} = 0$ si $i \neq j$ et $a_{i,j} = 1$ si $i = j$. On l'appelle la *matrice identité*.

Définition A.46. Si $A \in \mathcal{M}_{n,p}(K)$, on appelle transposée de A et on note tA la matrice de $\mathcal{M}_{p,n}(K)$ définie par

$${}^tA = (a_{j,i})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}.$$

On a alors ${}^t(AB) = {}^tB {}^tA$.

Définition A.47. Une matrice $A \in \mathcal{M}_n(K)$ est dite inversible s'il existe une matrice $B \in \mathcal{M}_n(K)$ telle que $AB = BA = I_n$. Si elle existe, la matrice B est unique et est appelée inverse de A . On la note alors A^{-1} .

On note $\text{GL}_n(K)$ l'ensemble des matrices inversibles de $\mathcal{M}_n(K)$. Si A et B sont dans $\text{GL}_n(K)$, alors $AB \in \text{GL}_n(K)$ et

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

A.10 Représentation matricielle d'une application linéaire

Soient E et F deux K -espaces vectoriels de dimension finie. On fixe $\mathcal{B}_E = (e_1, \dots, e_p)$ une base de E et $\mathcal{B}_F = (f_1, \dots, f_n)$ une base de F . Soit $u \in \mathcal{L}(E, F)$

une application linéaire de E dans F . Pour $1 \leq j \leq p$, on note $(a_{i,j})_{1 \leq i \leq n}$ les coordonnées de $u(e_j)$ dans la base $\mathcal{B}_F$. Autrement dit

$$u(e_j) = a_{1,j}f_1 + \cdots + a_{n,j}f_n.$$

La matrice de taille $n \times p$ de termes $(a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ est appelée *matrice de u dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$* . On la note $\text{Mat}_{(\mathcal{B}_E, \mathcal{B}_F)}(u)$ ou encore $\text{Mat}(u; \mathcal{B}_E, \mathcal{B}_F)$. Si $E = F$ et $\mathcal{B}_E = \mathcal{B}_F$, on note $\text{Mat}_{\mathcal{B}_E}(u) = \text{Mat}_{(\mathcal{B}_E, \mathcal{B}_E)}(u)$.

Soit $v \in E$. Si $(x_1, \dots, x_n)$ est le système de coordonnées de v dans la base $\mathcal{B}_E$, on note $X_{\mathcal{B}_E}(v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ le vecteur des coordonnées de v dans la base $\mathcal{B}_E$.

La matrice de u dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$ a alors la propriété suivante :

$$X_{\mathcal{B}_F}(u(x)) = \text{Mat}_{(\mathcal{B}_E, \mathcal{B}_F)}(u)X_{\mathcal{B}_E}(v).$$

Proposition A.48. *Soient E et F deux espaces vectoriels de dimension finie tels que $\dim E = p$ et $\dim F = n$. Soit $\mathcal{B}_E$ une base de E et soit $\mathcal{B}_F$ une base de F . Alors l'application de $\mathcal{L}(E, F)$ dans $\mathcal{M}_{n,p}(K)$ définie par $u \mapsto \text{Mat}_{(\mathcal{B}_E, \mathcal{B}_F)}(u)$ est linéaire et bijective.*

Démonstration. Voir le cours de L1. □

Proposition A.49. *Soient E , F et G trois espaces vectoriels de dimension finie tels que $\dim E = q$, $\dim F = p$ et $\dim G = n$. Soient $\mathcal{B}_E$ une base de E , $\mathcal{B}_F$ une base de F et $\mathcal{B}_G$ une base de G . Alors si $u \in \mathcal{L}(F, G)$ et $v \in \mathcal{L}(E, F)$, on a*

$$\text{Mat}_{(\mathcal{B}_E, \mathcal{B}_G)}(u \circ v) = \text{Mat}_{(\mathcal{B}_F, \mathcal{B}_G)}(u) \text{Mat}_{(\mathcal{B}_E, \mathcal{B}_F)}(v).$$

Démonstration. Voir le cours de L1. □

A.11 Changements de base

Soit E un K -espace vectoriel de dimension finie n . Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E . La *matrice de passage* de $\mathcal{B}$ à $\mathcal{B}'$ est la matrice carrée de taille n dont la j -ème colonne est donnée par les coordonnées du j -ème vecteur de $\mathcal{B}'$ dans la base $\mathcal{B}$. Autrement dit, si $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (e'_1, \dots, e'_n)$, et que

$$\forall 1 \leq j \leq n, e'_j = \sum_{i=1}^n a_{i,j}e_i,$$

alors

$$P_{\mathcal{B} \rightarrow \mathcal{B}'} = P_{\mathcal{B}}^{\mathcal{B}'} = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}.$$

Si $v \in E$, on a alors

$$X_{\mathcal{B}}(v) = P_{\mathcal{B} \rightarrow \mathcal{B}'} X_{\mathcal{B}'}(v).$$

En effet, soit $(x_1, \dots, x_n)$ le système de coordonnées de v dans la base $\mathcal{B}$ et $(x'_1, \dots, x'_n)$ le système de coordonnées de v dans la base $\mathcal{B}'$, c'est-à-dire $v = \sum_{i=1}^n x_i e_i = \sum_{j=1}^n x'_j e'_j$.

Comme $e'_j = \sum_{i=1}^n a_{i,j} e_i$, on en conclut

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

Comme $(e_1, \dots, e_n)$ est une base de E , on a bien

$$\forall 1 \leq i \leq n, \quad x_i = \sum_{j=1}^n a_{i,j} x'_j$$

c'est-à-dire $X_{\mathcal{B}}(v) = P_{\mathcal{B} \rightarrow \mathcal{B}'} X_{\mathcal{B}'}(v)$.

Si $\mathcal{B}$, $\mathcal{B}'$ et $\mathcal{B}''$ sont trois bases de E , on a alors

$$P_{\mathcal{B} \rightarrow \mathcal{B}''} = P_{\mathcal{B} \rightarrow \mathcal{B}'} P_{\mathcal{B}' \rightarrow \mathcal{B}''}.$$

De plus, $P_{\mathcal{B} \rightarrow \mathcal{B}} = I_n$ de sorte que la matrice $P_{\mathcal{B} \rightarrow \mathcal{B}'}$ est toujours inversible et que

$$P_{\mathcal{B} \rightarrow \mathcal{B}'}^{-1} = P_{\mathcal{B}' \rightarrow \mathcal{B}}.$$

Proposition A.50. Soient E et F deux espaces vectoriels de dimension finie. Soient $\mathcal{B}_E$ et $\mathcal{B}'_E$ deux bases de E et $\mathcal{B}_F$ et $\mathcal{B}'_F$ deux bases de F . Si $f \in \mathcal{L}(E, F)$, on a

$$\text{Mat}_{(\mathcal{B}'_E, \mathcal{B}'_F)}(f) = P_{\mathcal{B}_F \rightarrow \mathcal{B}'_F}^{-1} \text{Mat}_{(\mathcal{B}_E, \mathcal{B}_F)}(f) P_{\mathcal{B}_E \rightarrow \mathcal{B}'_E}.$$

En particulier, si $E = F$, $\mathcal{B}_E = \mathcal{B}_F$ et $\mathcal{B}'_E = \mathcal{B}'_F$, on a

$$\text{Mat}_{\mathcal{B}'_E}(f) = P_{\mathcal{B}_E \rightarrow \mathcal{B}'_E}^{-1} \text{Mat}_{\mathcal{B}_E}(f) P_{\mathcal{B}_E \rightarrow \mathcal{B}'_E}.$$

Deux matrices A et B de $\mathcal{M}_n(K)$ sont dites *semblables* s'il existe une matrice inversible $P \in \text{GL}_n(K)$ telle que

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

Deux matrices sont semblables si et seulement si ce sont les matrices du même endomorphisme de K^n dans des bases différentes.

B Appendix : Solution d'exercices

B.1 Exercice 1.2

Vérification de la bilinéarité de B

On définit $B : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ par

$$B((x, y), (v, w)) = 2xv + yw.$$

Vérifions si B est un produit scalaire :

-) **Bilinearité** : B est clairement linéaire par rapport à chaque variable.
-) **Symétrie** : Calculons $B((x, y), (v, w))$ et $B((v, w), (x, y))$:

$$B((x, y), (v, w)) = 2xv + yw,$$

$$B((v, w), (x, y)) = 2vx + yw.$$

-) **Définie** : $B((x, y), (x, y)) = 2x^2 + y^2 = 0 \Leftrightarrow (x, y) = 0$

Conclusion : B est un produit scalaire.

Vérification de l'orthogonalité de $(1, 1)$ et $(-1, 2)$

On vérifie si $B((1, 1), (-1, 2)) = 0$:

$$B((1, 1), (-1, 2)) = 2(1 \cdot 1) + (1 \cdot (-2)) = 2 - 2 = 0.$$

Donc $(1, 1)$ et $(-1, 2)$ sont orthogonaux.

B.2 Exercice 2.1

Trouver une base orthonormale de $\mathbb{R}^3$

Muni de $\mathbb{R}^3$ avec le produit scalaire canonique. Soient

$$v_1 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_3 := \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Nous appliquons l'algorithme de Gram-Schmidt pour construire une base orthonormale.

On applique l'algorithme de Gram-Schmidt aux vecteurs :

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Étape 1 : Calcul de e_1

$$\|v_1\| = \sqrt{1^2 + 1^2 + 0^2} = \sqrt{2}$$

$$e_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\boxed{e_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}$$

Étape 2 : Calcul de e_2

$$\text{proj}_{e_1}(v_2) = \frac{\langle v_2, e_1 \rangle}{\langle e_1, e_1 \rangle} e_1$$

$$\langle v_2, e_1 \rangle = \frac{1}{\sqrt{2}}$$

$$\text{proj}_{e_1}(v_2) = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$u_2 = v_2 - \text{proj}_{e_1}(v_2) = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

$$\|u_2\| = \sqrt{\frac{3}{2}}, \quad e_2 = \frac{u_2}{\|u_2\|} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

$$\boxed{e_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}}$$

Étape 3 : Calcul de e_3

$$\text{proj}_{e_1}(v_3) = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \text{proj}_{e_2}(v_3) = \frac{1}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

$$u_3 = v_3 - \text{proj}_{e_1}(v_3) - \text{proj}_{e_2}(v_3) = \begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix}$$

$$\|u_3\| = \sqrt{\frac{4}{3}}, \quad e_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix}$$

$$e_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix}$$

Conclusion

La base orthonormale obtenue est :

$$\left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \quad \frac{1}{\sqrt{3}} \begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix} \right)$$

Calcul de la projection orthogonale

Soit la projection orthogonale de v_3 sur $\text{Vect}(v_1, v_2)$:

$$p_{\text{Vect}(v_1, v_2)}(v_3) = \langle v_3, e_1 \rangle e_1 + \langle v_3, e_2 \rangle e_2$$

Calculons les produits scalaires :

$$\langle v_3, e_1 \rangle = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}$$

$$\langle v_3, e_2 \rangle = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \frac{1}{\sqrt{6}}$$

On en déduit :

$$\begin{aligned}
p_{\text{Vect}(v_1, v_2)}(v_3) &= \frac{1}{\sqrt{2}}e_1 + \frac{1}{\sqrt{6}}e_2 \\
&= \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{6}} \cdot \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2} + \frac{1}{6} \\ \frac{1}{2} - \frac{1}{6} \\ \frac{2}{6} \end{pmatrix} = \begin{pmatrix} \frac{3}{6} + \frac{1}{6} \\ \frac{3}{6} - \frac{1}{6} \\ \frac{2}{6} \end{pmatrix} \\
&= \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}
\end{aligned}$$

Ainsi, la projection orthogonale est :

$$p_{\text{Vect}(v_1, v_2)}(v_3) = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$$

B.3 Exercice 3.1

(i) **Existence d'une matrice orthogonale P telle que $(Pf_1, \dots, Pf_n)$ soit la base canonique de $\mathbb{R}^n$**

Soit $B = (f_1, f_2, \dots, f_n)$ une base orthonormée de $\mathbb{R}^n$. Nous devons prouver qu'il existe une matrice $P \in O(n)$ telle que $(Pf_1, \dots, Pf_n)$ soit la base canonique de $\mathbb{R}^n$, c'est-à-dire $(e_1, e_2, \dots, e_n)$, où e_i est le vecteur canonique avec 1 dans la i -ème position et 0 dans toutes les autres positions.

La matrice P est définie comme la matrice dont les colonnes sont les vecteurs $f_1, f_2, \dots, f_n$. Autrement dit,

$$P = \begin{pmatrix} | & | & \dots & | \\ f_1 & f_2 & \dots & f_n \\ | & | & \dots & | \end{pmatrix}.$$

Cela signifie que P est une matrice $n \times n$ dont les colonnes sont les vecteurs $f_1, f_2, \dots, f_n$, qui forment une base orthonormée de $\mathbb{R}^n$.

Puisque $B = (f_1, \dots, f_n)$ est une base orthonormée, les vecteurs $f_1, f_2, \dots, f_n$ sont orthogonaux et de norme 1. Cela implique que la matrice P est une matrice orthogonale. Ainsi, $P \in O(n)$, le groupe des matrices orthogonales.

La matrice P agit sur $f_1, f_2, \dots, f_n$ de la manière suivante :

$$Pf_1 = f_1, \quad Pf_2 = f_2, \quad \dots, \quad Pf_n = f_n.$$

Mais comme $f_1, f_2, \dots, f_n$ sont une base orthonormée et P est une matrice orthogonale, cela signifie que P^{-1} existe et que P^{-1} appliqué à $f_1, f_2, \dots, f_n$ donne les vecteurs de la base canonique $e_1, e_2, \dots, e_n$.

Donc il existe donc une matrice $P \in O(n)$ orthogonale telle que $(Pf_1, Pf_2, \dots, Pf_n) = (e_1, e_2, \dots, e_n)$, la base canonique de $\mathbb{R}^n$.

(ii) Prouver que $f \circ g$ est une isométrie si f et g sont des isométries

Soient $(E, \langle, \rangle)$ un espace euclidien, et $f, g : E \rightarrow E$ deux isométries, c'est-à-dire que pour tous $x, y \in E$,

$$\langle f(x), f(y) \rangle = \langle x, y \rangle \quad \text{et} \quad \langle g(x), g(y) \rangle = \langle x, y \rangle.$$

But : $f \circ g$, l'application composée de f et g , est également une isométrie, c'est-à-dire que pour tous $x, y \in E$,

$$\langle (f \circ g)(x), (f \circ g)(y) \rangle = \langle x, y \rangle.$$

Par définition de la composition des applications, on a :

$$(f \circ g)(x) = f(g(x)) \quad \text{et} \quad (f \circ g)(y) = f(g(y)).$$

Calculons :

$$\langle (f \circ g)(x), (f \circ g)(y) \rangle = \langle f(g(x)), f(g(y)) \rangle.$$

Puisque f est une isométrie, elle préserve le produit scalaire. Ainsi, nous avons :

$$\langle f(g(x)), f(g(y)) \rangle = \langle g(x), g(y) \rangle.$$

De même pour g :

$$\langle g(x), g(y) \rangle = \langle x, y \rangle.$$

Donc en combinant ces résultats, on a :

$$\boxed{\langle (f \circ g)(x), (f \circ g)(y) \rangle = \langle x, y \rangle.}$$

B.4 Exercice 4.1

Soit $v = \begin{pmatrix} a \\ b \end{pmatrix}$ et $w = \begin{pmatrix} c \\ d \end{pmatrix} \in \mathbb{R}^2$, et considérons la matrice $M = \begin{pmatrix} v & w \end{pmatrix}$. Le déterminant de cette matrice est donné par :

$$\det(M) = ad - bc.$$

Nous devons vérifier les cinq propriétés suivantes :

- (i) **Invariance par rotation** : On vérifié par un calcul direct que

$$\det(A \cdot B) = \det(A) \det(B)$$

pour $A, B \in \text{Mat}_2(\mathbb{R})$. De plus,

$$\det(R(\theta)) = \cos(\theta)^2 + \sin(\theta)^2 = 1.$$

Donc

$$\text{Vol}(R(\theta)v, R(\theta)w) = |\det(R(\theta)v | R(\theta)w)| = |\det(R(\theta)) \det(v | w)| = \text{Vol}(v, w).$$

- (ii) **Prolongement par des scalaires** : Soit $\lambda, \mu \in \mathbb{R}$, alors

$$\text{Vol}(\lambda v, \mu w) = |\det(\lambda v | \mu w)| = |\lambda \mu \det(v | w)| = |\lambda| |\mu| |\det(v | w)|.$$

Cette propriété est aussi vérifiée.

- (iii) **Invariance par permutation** : On a

$$\text{Vol}(v, w) = |\det(v | w)| = |\det(w | v)|,$$

car le déterminant change de signe lorsqu'on permute les colonnes, mais le module est invariant. Donc, cette propriété est vérifiée.

- (iv) **Volume standard** : Pour la base canonique $(e_1, e_2) = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$,

$$\text{Vol}(e_1, e_2) = |\det(e_1 | e_2)| = |1 \times 1 - 0 \times 0| = 1.$$

Cette propriété est également vérifiée.

- (v) **Invariance par déformation** : Soit $\lambda \in \mathbb{R}$, on a

$$\text{Vol}(v, \lambda v + w) = |\det(v | \lambda v + w)| = |\det(v | w)|,$$

car le volume ne change pas lorsqu'on effectue une déformation linéaire sur un vecteur tout en maintenant la relation entre les deux vecteurs. Ainsi, cette propriété est vérifiée.

En conclusion, l'application $(v, w) \mapsto |\det(v | w)|$ satisfait bien les propriétés (i) – (v).

B.5 Exercice 5.1

Calculons $\epsilon(\sigma)$ et écrivons σ comme un produit de cycles

$$\sigma(1) = 3, \sigma(2) = 5, \sigma(3) = 2, \sigma(4) = 1, \sigma(5) = 6, \sigma(6) = 4.$$

$$\sigma = (1\ 3\ 2\ 5\ 6\ 4).$$

Pour calculer $\epsilon(\sigma)$, il faut déterminer le nombre de transpositions nécessaires pour exprimer ce produit de cycles. Le cycle $(1\ 3\ 2\ 5\ 6\ 4)$ a une longueur de 6, donc il peut être écrit comme un produit de 5 transpositions :

$$(1\ 3\ 2\ 5\ 6\ 4) = (1\ 4)(1\ 6)(1\ 5)(1\ 2)(1\ 3).$$

On en déduit,

$$\boxed{\epsilon(\sigma) = (-1)^5 = -1}$$

Calcul de déterminants

Soit la matrice :

$$A = \begin{bmatrix} 2 & 0 & 2 & 4 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 2 & 0 & 2 & 5 \end{bmatrix}$$

$$L_4 = L_4 - L_1 :$$

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

Développement selon la première colonne :

$$\det(A) = 2 \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 2 \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 2 \times \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}$$

$$\det(A) = 1 \times (1 \times 1 - 1 \times 1) = 1 \times 0 = 0.$$

$$\boxed{\det(A) = 0}$$

Soit la matrice :

$$B = \begin{bmatrix} 2 & 0 & 2 & 4 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 0 & 2 & 5 \end{bmatrix}$$

De la même manière, on trouve pour $\det(B)$

$$\boxed{\det(B) = 2 \times 1 = 2}$$

B.6 Exercice 6.1

Le volume du parallélogramme engendré par les vecteurs

$$\mathbf{u} = \begin{pmatrix} 5 \\ 7 \end{pmatrix} \text{ et } \mathbf{v} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

est donné par la valeur absolue du déterminant de la matrice formée par ces deux vecteurs, c'est-à-dire :

$$\text{Volume} = \left| \det \begin{pmatrix} 5 & 4 \\ 7 & 3 \end{pmatrix} \right|.$$

$$\det \begin{pmatrix} 5 & 4 \\ 7 & 3 \end{pmatrix} = 5 \cdot 3 - 7 \cdot 4 = 15 - 28 = -13.$$

Donc le volume du parallélogramme est :

$$\boxed{\text{Volume} = |-13| = 13}$$