

Algèbre linéaire

Table des matières

1	Les matrices	3
1.1	Vocabulaire	3
1.2	Opérations sur les matrices	4
1.3	L'algèbre des matrices carrées	7
1.4	Transposée d'une matrice	10
1.5	Différentes interprétations du produit matriciel	12
1.6	Trace d'une matrice	14
1.7	Matrices décomposées en blocs	15
1.7.1	Matrices extraites	15
1.7.2	Matrices blocs	15
1.7.3	Opérations sur les matrices blocs	16
2	Familles de vecteurs	17
2.1	Familles libres et génératrices	17
2.2	Dimension d'un espace vectoriel	18
2.3	Base canonique	22
2.4	Exemples	23
2.5	Application linéaire associée à une famille de vecteurs	27
2.6	Image d'une famille par une application linéaire	27
2.7	Rang d'une famille de vecteurs	32
2.8	Matrice d'une application linéaire	32
3	Les systèmes linéaires	38
3.1	Trois interprétations d'un système linéaire	38
3.2	Les opérations élémentaires	39
3.3	Méthode du pivot de Gauss	42
3.4	Méthode du pivot total	45
3.5	Méthode de Gauss-Jordan	46

4	Somme de sous-espaces vectoriels	47
4.1	Sommes et sommes directes	47
4.2	Supplémentaires d'un sous-espace vectoriel	49
4.3	Rang d'une application linéaire	52
4.4	Propriétés des sommes directes	53
4.4.1	Un moyen de définir une application linéaire	53
4.4.2	Formules dimensionnelles	54
4.4.3	Associativité des sommes directes	56
4.4.4	Base adaptée à une décomposition en somme directe	58
4.5	Les projecteurs	59
4.6	Sous-espaces propres	62
4.6.1	Définitions	62
4.6.2	Exemples	65
4.6.3	Propriétés	65
5	Changement de base	66
5.1	Matrice de passage	66
5.2	Diagonalisation et trigonalisation	68
5.3	Trace d'un endomorphisme	71
5.4	Matrices équivalentes et matrices semblables	71
5.4.1	Matrices équivalentes	71
5.4.2	Propriétés du rang d'une matrice	73
5.4.3	Matrices semblables	74
6	Les hyperplans	75
6.1	En dimension quelconque	75
6.2	En dimension finie	76
6.3	Les hyperplans affines	77
6.4	Application aux systèmes linéaires	78
7	Déterminants	80
7.1	Applications multilinéaires	80
7.2	Les trois notions de déterminants	84
7.2.1	Volume	84
7.2.2	Déterminant d'un système de n vecteurs	87
7.2.3	Déterminant d'une matrice	89
7.2.4	Déterminant d'un endomorphisme	90
7.3	Propriétés du déterminant	90
7.4	Calcul des déterminants	92
7.5	Formules de Cramer	95
7.6	Exemples de déterminants.	96
7.6.1	Déterminant de Vandermonde	96
7.6.2	Déterminants tridiagonaux	99

7.6.3	Déterminants circulants	101
7.7	Le polynôme caractéristique	102
7.7.1	Définition	102
7.7.2	Propriétés du polynôme caractéristique	104
7.7.3	Caractérisation des endomorphismes diagonalisables	107

Notation. \mathbb{K} désigne un corps quelconque.

Selon le programme, “en pratique, \mathbb{K} est égal à \mathbb{R} ou \mathbb{C} ”.

Notation. Symbole de Kronecker : Si i et j sont deux objets mathématiques, on convient que $\delta_{i,j} = 0$ lorsque $i \neq j$ et $\delta_{i,i} = 1$ lorsque $i = j$.

1 Les matrices

1.1 Vocabulaire

Définition. Soit $(n, p) \in \mathbb{N}^{*2}$. On appelle **matrice** à n lignes et à p colonnes (à coefficients dans \mathbb{K}) toute famille de scalaires indexée par $\mathbb{N}_n \times \mathbb{N}_p$.

Si $M = (m_{i,j})_{(i,j) \in \mathbb{N}_n \times \mathbb{N}_p} = (m_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$, on représente M sous la forme suivante :

$$M = \begin{pmatrix} m_{1,1} & \cdots & m_{1,p} \\ \vdots & & \vdots \\ m_{n,1} & \cdots & m_{n,p} \end{pmatrix},$$

où le (i, j) ^{ème} coefficient est situé à l'intersection de la i ^{ème} ligne et de la j ^{ème} colonne. Une matrice est donc un tableau de scalaires.

Notation.

- L'ensemble des matrices à coefficients dans \mathbb{K} , à n lignes et p colonnes est noté $\mathcal{M}_{\mathbb{K}}(n, p)$ ou $\mathcal{M}_{n,p}(\mathbb{K})$.
- De plus, $\mathcal{M}_{\mathbb{K}}(n, n)$ est souvent noté $\mathcal{M}_{\mathbb{K}}(n)$ ou $\mathcal{M}_n(\mathbb{K})$.

Définitions :

- Une **matrice ligne** est une matrice ne possédant qu'une ligne.
- Une **matrice colonne** est une matrice ne possédant qu'une colonne.
- Une **matrice carrée** est une matrice possédant autant de lignes que de colonnes.
- Soit $M = (m_{i,j}) \in \mathcal{M}_{\mathbb{K}}(n, p)$.
 M est une **matrice triangulaire supérieure** si et seulement si
 $\forall (i, j) \in \mathbb{N}_n \times \mathbb{N}_p \ (i > j \implies m_{i,j} = 0)$.
- M est une **matrice triangulaire inférieure** si et seulement si
 $\forall (i, j) \in \mathbb{N}_n \times \mathbb{N}_p \ (i < j \implies m_{i,j} = 0)$.
- M est une **matrice diagonale** si et seulement si
 $\forall (i, j) \in \mathbb{N}_n \times \mathbb{N}_p \ (i \neq j \implies m_{i,j} = 0)$.
On note alors $M = \text{diag}(m_{1,1}, \dots, m_{n,n})$.
- Soit M une matrice diagonale et carrée. On dit que M est une **matrice scalaire** si et seulement si tous ses coefficients diagonaux sont égaux.
En particulier, lorsque tous ses coefficients diagonaux sont égaux à 1, on obtient la matrice identité, notée I_n .
Ainsi, M est une matrice scalaire si et seulement s'il existe $\lambda \in \mathbb{K}$ tel que $M = \lambda I_n$.

Définition. Soit $M = (\alpha_{i,j}) \in \mathcal{M}_{\mathbb{K}}(n, p)$.

Pour $j \in \mathbb{N}_p$, on appelle $j^{\text{ème}}$ vecteur colonne de M la quantité $(\alpha_{1,j}, \dots, \alpha_{n,j}) \in \mathbb{K}^n$.

Pour $i \in \mathbb{N}_n$, on appelle $i^{\text{ème}}$ vecteur ligne de M la quantité $(\alpha_{i,1}, \dots, \alpha_{i,p}) \in \mathbb{K}^p$.

Remarque. Lorsqu'aucune ambiguïté n'est possible, on identifie \mathbb{K}^n avec $\mathcal{M}_{\mathbb{K}}(n, 1)$ (ensemble des matrices colonnes).

Plus rarement, un vecteur de \mathbb{K}^n sera vu comme un élément de $\mathcal{M}_{\mathbb{K}}(1, n)$ (ensemble des matrices lignes).

1.2 Opérations sur les matrices

Définition. $\mathcal{M}_{\mathbb{K}}(n, p) = \mathbb{K}^{\mathbb{N}_n \times \mathbb{N}_p}$, or \mathbb{K} est un \mathbb{K} -espace vectoriel, donc $\mathcal{M}_{\mathbb{K}}(n, p)$ est un \mathbb{K} -espace vectoriel : On dispose ainsi des lois d'addition et de multiplication par un scalaire.

Exemple. $3 \begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 3 \end{pmatrix} - \begin{pmatrix} 0 & 4 & -1 \\ 2 & 3 & 1 \end{pmatrix} = \dots$

Définition du produit matriciel : Soit $(n, p, q) \in (\mathbb{N}^*)^3$.

Soient $A = (a_{i,j}) \in \mathcal{M}_{\mathbb{K}}(n, p)$ et $B = (b_{j,k}) \in \mathcal{M}_{\mathbb{K}}(p, q)$.

On appelle **produit des matrices** A et B la matrice $C = (c_{i,k}) \in \mathcal{M}_{\mathbb{K}}(n, q)$ définie par

$$\forall (i, k) \in \mathbb{N}_n \times \mathbb{N}_q \quad c_{i,k} = \sum_{j=1}^p a_{i,j} b_{j,k}.$$

Exemple. Soit $a \in \mathbb{K}$:

$$\begin{pmatrix} a & 1 \\ -a & a \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 1 \\ a+2 & 1-a \\ a & 2a \\ 0 & 0 \end{pmatrix}$$

Convention : sauf précision du contraire, lorsque A est une matrice, on notera $A_{i,j}$ son coefficient de position (i, j) .

Ainsi, lorsque A et B sont deux matrices telles que le nombre p de colonnes de A est égal au nombre de lignes de B , la définition du produit matriciel se résume par :

$$[AB]_{i,j} = \sum_{k=1}^p A_{i,k} B_{k,j}.$$

Formule pour le produit de trois matrices : Soit $(n, m, l, p) \in (\mathbb{N}^*)^4$.

Soient $A = (a_{i,j}) \in \mathcal{M}_{\mathbb{K}}(n, m)$, $B = (b_{j,k}) \in \mathcal{M}_{\mathbb{K}}(m, l)$ et $C = (c_{k,h}) \in \mathcal{M}_{\mathbb{K}}(l, p)$.

Pour tout $i, h \in \mathbb{N}_n \times \mathbb{N}_p$, $[(AB)C]_{i,h} = [A(BC)]_{i,h} = \sum_{\substack{1 \leq j \leq m \\ 1 \leq k \leq l}} A_{i,j} B_{j,k} C_{k,h}.$

Démonstration.

Soit $(i, h) \in \mathbb{N}_n \times \mathbb{N}_p$.

$$[(AB)C]_{i,h} = \sum_{k=1}^l [AB]_{i,k} C_{k,h} = \sum_{k=1}^l \left(\sum_{j=1}^m A_{i,j} B_{j,k} \right) C_{k,h} = \sum_{j=1}^m \sum_{k=1}^l A_{i,j} B_{j,k} C_{k,h}. \quad \square$$

Remarque. On pourrait généraliser en donnant l'expression des coefficients du produit de N matrices en fonction des coefficients de ces N matrices.

Exemple. On considère un ensemble $S = \{s_1, \dots, s_n\}$ fini de cardinal n , dont les éléments sont appelés des sommets, et une partie A de S^2 , dont les éléments sont appelés des arêtes. On a ainsi défini un graphe orienté, dont l'ensemble des sommets est S et l'ensemble des arêtes est A .

La matrice d'adjacence de ce graphe est, par définition, la matrice $M \in \mathcal{M}_n(\mathbb{R})$ suivante : pour tout $i, j \in \{1, \dots, n\}$, $M_{i,j} = 1$ si $(s_i, s_j) \in A$ et $M_{i,j} = 0$ sinon.

Alors, pour tout $p \in \mathbb{N}$, $[M^p]_{i,j}$ correspond au nombre de chemins de longueur p permettant de passer de s_i à s_j : en effet, $[M^p]_{i,j} = \sum_{1 \leq i_1, \dots, i_{p-1} \leq n} M_{i,i_1} M_{i_1,i_2} \cdots M_{i_{p-1},j}$, et

$M_{i,i_1} M_{i_1,i_2} \cdots M_{i_{p-1},j}$ est égal à 1 si et seulement si le chemin

$s_i \rightarrow s_{i_1} \rightarrow \cdots \rightarrow s_{i_{p-1}} \rightarrow s_j$ est bien une succession d'arêtes du graphe, et il est égal à 0 sinon.

Propriété. La multiplication matricielle est associative.

Propriété. La multiplication matricielle est distributive par rapport à l'addition.

Démonstration.

Soit $A \in \mathcal{M}_{n,p}$ et $B, C \in \mathcal{M}_{p,q}$. Alors $A(B + C) = AB + AC$

$$\text{car } [A(B + C)]_{i,k} = \sum_{j=1}^p A_{i,j} (B_{j,k} + C_{j,k}) = \sum_{j=1}^p A_{i,j} B_{j,k} + \sum_{j=1}^p A_{i,j} C_{j,k} = [AB + AC]_{i,k}.$$

De même, on montre que si $D \in \mathcal{M}_{n,p}(\mathbb{K})$, $(A + D)B = AB + DB$. \square

Propriété. Soit $A \in \mathcal{M}_{n,p}$, $B \in \mathcal{M}_{p,q}$ et $a \in \mathbb{K}$. Alors $a(AB) = (aA)B = A(aB)$.

Démonstration.

$$[a(AB)]_{i,k} = \sum_{j=1}^p a A_{i,j} B_{j,k} = [(aA)B]_{i,k} = [A(aB)]_{i,k}. \quad \square$$

Propriété. Pour tout $M \in \mathcal{M}_{\mathbb{K}}(n, p)$, $I_n M = M I_p = M$.

Démonstration.

$$[I_n M]_{i,k} = \sum_{1 \leq j \leq n} \delta_{i,j} M_{j,k} = M_{i,k}. \quad \square$$

Propriété. Soit $n, p \in \mathbb{N}^*$ et $M \in \mathcal{M}_{\mathbb{K}}(n, p)$.

Pour tout $X \in \mathbb{K}^p = \mathcal{M}_{\mathbb{K}}(p, 1)$, $MX \in \mathcal{M}_{\mathbb{K}}(n, 1) = \mathbb{K}^n$.

Si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$, alors $\forall i \in \{1, \dots, n\}$, $[MX]_i = \sum_{j=1}^p M_{i,j} x_j$.

$$\text{Si } X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, \text{ alors } MX = x_1 M_1 + \cdots x_p M_p,$$

en notant M_1, \dots, M_p les colonnes de M .

Démonstration.

$$[MX]_{i,1} = \sum_{j=1}^p M_{i,j} X_{j,1}, \text{ donc avec d'autres notations, } [MX]_i = \sum_{j=1}^p M_{i,j} x_j.$$

$$\text{Ainsi, } [MX]_i = \sum_{j=1}^p x_j [M_j]_i, \text{ où } [M_j]_i \text{ désigne la } i\text{-ème composante de } M_j. \square$$

Propriété. Soit $n, p \in \mathbb{N}^*$ et $M \in \mathcal{M}_{\mathbb{K}}(n, p)$.

Soit $j \in \{1, \dots, p\}$. La j -ème colonne de M est $M c_j$, où $c_j = (\delta_{i,j})_{1 \leq i \leq n} \in \mathbb{K}^p$.

Propriété. Soit $n, p \in \mathbb{N}^*$ et $M \in \mathcal{M}_{\mathbb{K}}(n, p)$.

Alors l'application $\begin{matrix} \tilde{M} : \mathbb{K}^p & \longrightarrow & \mathbb{K}^n \\ X & \longmapsto & MX \end{matrix}$ est une application linéaire que l'on appelle

l'application linéaire canoniquement associée à la matrice M .

Propriété. Soit $n, p \in \mathbb{N}^*$. Alors $\begin{matrix} \mathcal{M}_{\mathbb{K}}(n, p) & \longrightarrow & L(\mathbb{K}^p, \mathbb{K}^n) \\ M & \longmapsto & \tilde{M} \end{matrix}$ est un isomorphisme d'espaces vectoriels.

Démonstration.

◇ Soit $A, B \in \mathcal{M}_{\mathbb{K}}(n, p)$ et $\lambda \in \mathbb{K}$. Pour tout $X \in \mathbb{K}^p$,

$(\lambda A + B)(X) = (\lambda A + B)X = \lambda AX + BX = \lambda \tilde{A}(X) + \tilde{B}(X) = (\lambda \tilde{A} + \tilde{B})(X)$. Ceci prouve que $M \mapsto \tilde{M}$ est linéaire.

◇ Soit $u \in L(\mathbb{K}^p, \mathbb{K}^n)$. S'il existe $M \in \mathcal{M}_{\mathbb{K}}(n, p)$ telle que $\tilde{M} = u$, alors, pour tout $j \in \{1, \dots, p\}$, la j -ème colonne de M est égale à $M c_j = \tilde{M}(c_j) = u(c_j)$. Donc si M existe, elle est unique.

Pour la synthèse, notons M la matrice de $\mathcal{M}_{\mathbb{K}}(n, p)$ dont la j -ème colonne est égale à $u(c_j) = M_j$, pour tout $j \in \{1, \dots, p\}$. Ainsi, pour tout $j \in \{1, \dots, p\}$, $u(c_j) = \tilde{M}(c_j)$.

$$\text{Soit } X \in \mathbb{K}^p. \text{ En posant } X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, X = \sum_{j=1}^p x_j c_j,$$

$$\text{donc } u(X) = \sum_{j=1}^p x_j u(c_j) = \sum_{j=1}^p x_j M_j = MX = \tilde{M}(X), \text{ donc } u = \tilde{M}.$$

Ainsi, pour tout $u \in L(\mathbb{K}^p, \mathbb{K}^n)$, il existe une unique $M \in \mathcal{M}_{\mathbb{K}}(n, p)$ telle que $\tilde{M} = u$, ce qui prouve que $M \mapsto \tilde{M}$ est une bijection. \square

Remarque. Avec les notations précédentes, pour tout $X \in \mathbb{K}^p$, $\tilde{M}(X) = MX$.

Il est fréquent que l'on identifie M et \tilde{M} . Alors, pour tout $X \in \mathbb{K}^p$, $M(X) = MX$. Cette identification n'est pas systématique cependant.

Définition. Soit $M \in \mathcal{M}_{\mathbb{K}}(n, p)$.

$$\text{Ker}(M) \triangleq \text{Ker}(\tilde{M}) = \{X \in \mathbb{K}^p / MX = 0\}.$$

$$\text{Im}(M) \stackrel{\Delta}{=} \text{Im}(\tilde{M}) = \{MX \mid X \in \mathbb{K}^p\}.$$

Corollaire. Soit $(M, M') \in \mathcal{M}_{\mathbb{K}}(n, p)$. Alors,

$$(\forall X \in \mathbb{K}^p \quad MX = M'X) \iff M = M'.$$

Démonstration.

Si $\forall X \in \mathbb{K}^p \quad MX = M'X$, alors $\tilde{M} = \widetilde{M'}$, donc par injectivité, $M = M'$. \square

1.3 L'algèbre des matrices carrées

Notation. On fixe un entier n non nul.

Propriété. $(\mathcal{M}_n(\mathbb{K}), +, \cdot, \times)$ est une \mathbb{K} -algèbre, non commutative dès que $n \geq 2$.

ATTENTION : $(\mathcal{M}_n(\mathbb{K}), +, \cdot, \times)$ n'est pas intègre, dès que $n \geq 2$.

En effet, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = 0$.

Ainsi, lorsque $A, B, C \in \mathcal{M}_n(\mathbb{K})$, $AB = 0 \not\Rightarrow (A = 0) \vee (B = 0)$

et $(C \neq 0) \wedge (AC = BC) \not\Rightarrow A = B$.

Cependant, lorsque C est inversible, alors $AC = BC \implies A = B$.

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

On dit que A est nilpotente si et seulement si il existe $p \in \mathbb{N}^*$ tel que $A^p = 0$.

ATTENTION : L'anneau $(\mathcal{M}_n(\mathbb{K}), +, \times)$ n'est pas commutatif, ce qui nous prive d'un certain nombre de règles : si $A, B \in \mathcal{M}_n(\mathbb{K})$,

- $(AB)^p \neq A^p B^p$;
- $(A - B)(A + B) \neq A^2 - B^2$; plus généralement, la formule de Bernoulli n'est plus valable, lorsque A et B ne commutent pas.
- $(A + B)^2 \neq A^2 + B^2 + 2AB$. plus généralement, la formule de Newton n'est plus valable, lorsque A et B ne commutent pas.

Propriété. $\begin{array}{ccc} \mathcal{M}_{\mathbb{K}}(n) & \longrightarrow & L(\mathbb{K}^n) \\ M & \longmapsto & \tilde{M} \end{array}$ est un isomorphisme d'algèbres.

Démonstration.

$$\varphi(I_n) = Id_{\mathbb{K}^n}.$$

Pour tout $A, B \in \mathcal{M}_{\mathbb{K}}(n)$ et $X \in \mathbb{K}^n$,

$$\widetilde{(AB)}(X) = (AB)X = A(BX) = \tilde{A}(\tilde{B}(X)) = [\tilde{A} \circ \tilde{B}](X). \quad \square$$

Corollaire. Soit $A \in \mathcal{M}_{\mathbb{K}}(n)$. A est inversible dans $\mathcal{M}_{\mathbb{K}}(n)$ si et seulement si \tilde{A} est inversible dans $L(\mathbb{K}^n)$ et dans ce cas, $\widetilde{M^{-1}} = \tilde{M}^{-1}$.

Corollaire. Soit $A \in \mathcal{M}_{\mathbb{K}}(n)$. A est inversible dans $\mathcal{M}_{\mathbb{K}}(n)$ si et seulement si, pour tout $Y \in \mathbb{K}^n$, il existe un unique $X \in \mathbb{K}^n$ tel que $AX = Y$.

Formule : Dans $\mathcal{M}_2(\mathbb{K})$, $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible si et seulement si

$\det(M) \triangleq ad - cb \neq 0$, et dans ce cas

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\det(M)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Démonstration.

◇ Supposons que $\det(M) \neq 0$.

On calcule $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \times \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & -cb + ad \end{pmatrix} = \det(M)I_2$,

donc $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \times \frac{1}{\det(M)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = I_2$.

De même on vérifie que $\frac{1}{\det(M)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \times \begin{pmatrix} a & b \\ c & d \end{pmatrix} = I_2$,

donc M est inversible et $M^{-1} = \frac{1}{\det(M)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

◇ Supposons que $ad - bc = 0$.

Alors $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d \\ -c \end{pmatrix} = 0$ et $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} b \\ -a \end{pmatrix} = 0$, donc $\begin{pmatrix} d \\ -c \end{pmatrix}, \begin{pmatrix} b \\ -a \end{pmatrix} \in \text{Ker}(\tilde{M})$.

On en déduit que $\text{Ker}(\tilde{M}) \neq \{0\}$ lorsque $M \neq 0$, donc que \tilde{M} n'est pas inversible, puis que M n'est pas inversible. Lorsque $M = 0$, comme dans tout anneau non nul, M n'est pas inversible. □

Exemple. $\begin{pmatrix} 1 & 5 \\ -3 & 2 \end{pmatrix}^{-1} = \frac{1}{17} \begin{pmatrix} 2 & -5 \\ 3 & 1 \end{pmatrix}$.

On en déduit les formules de Cramer pour la résolution d'un système linéaire de deux équations à deux inconnues :

Formule de Cramer : Soit $a, b, c, d, e, f \in \mathbb{K}^4$. On considère le système linéaire (S) :

$$\begin{cases} ax + by = e \\ cx + dy = f \end{cases}, \text{ en les inconnues } (x, y) \in \mathbb{K}^2.$$

Lorsque $\det = ad - cb \triangleq \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$, $(S) \iff \begin{cases} x = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\det} \\ y = \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\det} \end{cases}.$

Démonstration.

Posons $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \end{pmatrix}$ et $Y = \begin{pmatrix} e \\ f \end{pmatrix}$. Alors $(S) \iff MX = Y$.

On suppose que $\det(M) \neq 0$, donc M est inversible.

Alors $(S) \iff X = M^{-1}Y = \frac{1}{\det(M)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} e \\ f \end{pmatrix}$. □

Exemple. $\begin{cases} x + 5y = 1 \\ -3x + 2y = 1 \end{cases} \iff \begin{cases} x = -\frac{3}{17} \\ y = \frac{4}{17} \end{cases}.$

Notation. On note $GL_n(\mathbb{K})$ le groupe des inversibles de $\mathcal{M}_n(\mathbb{K})$. On l'appelle le groupe linéaire de degré n .

Exemple. Un automorphisme intérieur de $\mathcal{M}_n(\mathbb{K})$ est un automorphisme sur $\mathcal{M}_n(\mathbb{K})$ de la forme $M \mapsto AMA^{-1}$ où $A \in GL_n(\mathbb{K})$.

Propriété. L'ensemble des matrices diagonales de $\mathcal{M}_n(\mathbb{K})$ est une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{K})$.

Propriété. Pour tout $i \in \mathbb{N}_n$, on pose $c_i = (\delta_{i,j})_{1 \leq j \leq n} \in \mathbb{K}^n$.

Pour tout $i \in \{0, \dots, n\}$, on note $F_i = \text{Vect}(c_k)_{1 \leq k \leq i}$. Ainsi, $F_0 = \{0\}$ et pour tout

$$i \in \{1, \dots, n\}, F_i = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} / x_1, \dots, x_i \in \mathbb{K} \right\}.$$

Si $M \in \mathcal{M}_n(\mathbb{K})$, alors M est triangulaire supérieure si et seulement si, pour tout $j \in \{1, \dots, n\}$, F_j est stable par \tilde{M} .

Démonstration.

M est triangulaire supérieure si et seulement si, pour tout $j \in \{1, \dots, n\}$, la j -ème colonne de M , égale à $\tilde{M}(c_j)$, est une combinaison linéaire de c_1, \dots, c_j , donc si et seulement si $(C) : \forall j \in \{1, \dots, n\}, \tilde{M}(c_j) \in F_j$.

Or si tout les F_j sont stables par \tilde{M} , alors pour tout j , sachant que $c_j \in F_j$, $\tilde{M}(c_j) \in F_j$, donc (C) est vérifiée.

Réciproquement, si l'on suppose (C) , pour j fixé dans $\{1, \dots, p\}$, pour tout $k \in \{1, \dots, j\}$, $\tilde{M}(c_k) \in F_k \subset F_j$, donc $\text{Vect}(\{\tilde{M}(c_k) / 1 \leq k \leq j\}) \subset F_j$, or

$$\begin{aligned} \text{Vect}(\{\tilde{M}(c_k) / 1 \leq k \leq j\}) &= \left\{ \sum_{k=1}^j \alpha_k \tilde{M}(c_k) / \alpha_1, \dots, \alpha_j \in \mathbb{K} \right\} \\ &= \left\{ \tilde{M} \left(\sum_{k=1}^j \alpha_k c_k \right) / \alpha_1, \dots, \alpha_j \in \mathbb{K} \right\} \\ &= \tilde{M}(\text{Vect}(\{c_k / 1 \leq k \leq j\})), \end{aligned}$$

donc $\text{Vect}(\{\tilde{M}(c_k) / 1 \leq k \leq j\}) = \tilde{M}(F_j)$, si bien que $\tilde{M}(F_j) \subset F_j$. \square

Propriété. On suppose que $n \geq 2$.

- L'ensemble des matrices triangulaires supérieures (respectivement : inférieures) de $\mathcal{M}_n(\mathbb{K})$ est une sous-algèbre non commutative de $\mathcal{M}_n(\mathbb{K})$.
- Le produit d'une matrice triangulaire supérieure dont la diagonale est (a_1, \dots, a_n) par une matrice triangulaire supérieure dont la diagonale est (b_1, \dots, b_n) est une matrice triangulaire supérieure dont la diagonale est $(a_1 b_1, \dots, a_n b_n)$.

Démonstration.

Première démonstration :

◇ Notons \mathcal{T} l'ensemble des matrices supérieures. $I_n \in \mathcal{T}$, \mathcal{T} est stable pour l'addition et la multiplication par un scalaire, donc pour montrer que \mathcal{T} est une sous-algèbre, il reste

à montrer qu'il est stable pour le produit. Or, si $A, B \in \mathcal{T}$, pour tout $i \in \{1, \dots, n\}$, $\widetilde{(AB)}(F_i) = \widetilde{A}(\widetilde{B}(F_i)) \subset \widetilde{A}(F_i)$, car $B \in \mathcal{T}$ donc $\widetilde{B}(F_i) \subset F_i$, donc $\widetilde{(AB)}(F_i) \subset F_i$, ce qui prouve que $AB \in \mathcal{T}$.

◇ Soit $j \in \{1, \dots, n\}$. Ac_j est la j -ème colonne de A et $A \in \mathcal{T}$, donc $Ac_j = A_{j,j}c_j + d$, où $d \in F_{j-1}$. De même, $Bc_j = B_{j,j}c_j + d'$, où $d' \in F_{j-1}$.

Ainsi, $(AB)c_j = A(B_{j,j}c_j + d') = B_{j,j}Ac_j + Ad' = A_{j,j}B_{j,j}c_j + B_{j,j}d + Ad'$.

Or $A \in \mathcal{T}$ et $d' \in F_{j-1}$, donc $Ad' \in F_{j-1}$ et $d \in F_{j-1}$. Ceci démontre que le coefficient de position (j, j) de AB est égal à $A_{j,j}B_{j,j}$.

Seconde démonstration : Par calcul matriciel direct.

◇ Soient $A = (a_{i,j})$ et $B = (b_{i,j})$ deux matrices triangulaires supérieures.

Soit $(i, j) \in \mathbb{N}_n^2$ avec $i > j$. Le coefficient de position (i, j) de AB vaut $\sum_{k=1}^n a_{i,k}b_{k,j}$. Mais,

pour tout $k \in \mathbb{N}_n$, $i > k$ ou $k > j$ (sinon, $i \leq k \leq j$, donc $i \leq j$, ce qui est faux). Or A et B sont triangulaires supérieures, donc, pour tout $k \in \mathbb{N}_n$, $a_{i,k} = 0$ ou $b_{k,j} = 0$, ce qui prouve que le (i, j) ^{ème} coefficient de AB est nul. Ainsi AB est une matrice triangulaire supérieure.

◇ De plus, le coefficient de position (i, i) de AB vaut $\sum_{k=1}^n a_{i,k}b_{k,i}$. Mais, pour tout

$k \in \mathbb{N}_n$, $a_{i,k}b_{k,i}$ est non nul si et seulement si $i \leq k$ et $k \leq i$, donc si et seulement si $k = i$. Ainsi, le (i, i) ^{ème} coefficient de AB vaut $a_{i,i}b_{i,i}$. □

Exercice. Soit $M \in \mathcal{M}_{\mathbb{K}}(n)$ une matrice triangulaire supérieure stricte, c'est-à-dire triangulaire supérieure et de diagonale nulle.

Montrer que pour tout $k \in \{1, \dots, n\}$, M^k est une matrice triangulaire supérieure dont les k diagonales supérieures (en partant de la diagonale principale) sont nulles. En déduire que M est nilpotente.

Solution : En adaptant ce qui précède, on montre que pour tout $i \in \{1, \dots, n\}$, $\widetilde{M}(F_i) \subset F_{i-1}$, puis par récurrence sur k que, pour tout $k \in \{1, \dots, n\}$, et $i \in \{1, \dots, n\}$, $\widetilde{M}^k(F_i) \subset F_{i-k}$ en convenant que pour tout $h \in \mathbb{Z} \setminus \mathbb{N}$, $F_h = \{0\}$.

1.4 Transposée d'une matrice

Définition. Soit $A = (\alpha_{i,j}) \in \mathcal{M}_{\mathbb{K}}(n, p)$. On appelle **transposée de la matrice** A et on note tA la matrice $(\beta_{i,j}) \in \mathcal{M}_{\mathbb{K}}(p, n)$ définie par

$$\boxed{\forall (i, j) \in \mathbb{N}_p \times \mathbb{N}_n \quad \beta_{i,j} = \alpha_{j,i}.}$$

En résumé, $[{}^tA]_{i,j} = A_{j,i}$.

ATTENTION : Si $A = (\alpha_{i,j}) \in \mathcal{M}_{\mathbb{K}}(n, p)$, alors $(\alpha_{j,i})_{\substack{1 \leq j \leq n \\ 1 \leq i \leq p}} = A$.

Exemples.

— ${}^tI_n = I_n$. Plus généralement, pour toute matrice diagonale $D \in \mathcal{M}_{\mathbb{K}}(n)$, ${}^tD = D$.

— La transposée de $\begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 4 & 5 \end{pmatrix}$ est $\begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & 5 \end{pmatrix}$.

— La transposée d'une matrice triangulaire supérieure est triangulaire inférieure.

Propriété. Pour tout $A \in \mathcal{M}_{\mathbb{K}}(n, p)$, ${}^t({}^tA) = A$.

Propriété. L'application $\begin{matrix} \mathcal{M}_{\mathbb{K}}(n, p) & \longrightarrow & \mathcal{M}_{\mathbb{K}}(p, n) \\ M & \longmapsto & {}^tM \end{matrix}$ est un isomorphisme d'espaces vectoriels.

Propriété. Soit $(A, B) \in \mathcal{M}_{\mathbb{K}}(n, p) \times \mathcal{M}_{\mathbb{K}}(p, q)$. Alors, ${}^t(AB) = {}^tB {}^tA$.

Démonstration.

Soit $i \in \{1, \dots, n\}$ et $k \in \{1, \dots, q\}$.

$$[{}^tB {}^tA]_{k,i} = \sum_{j=1}^p [{}^tB]_{k,j} [{}^tA]_{j,i} = \sum_{j=1}^p B_{j,k} A_{i,j} = [AB]_{i,k} = [{}^t(AB)]_{k,i}. \quad \square$$

Corollaire. Si $A \in GL_n(\mathbb{K})$, ${}^tA \in GL_n(\mathbb{K})$ et $({}^tA)^{-1} = {}^t(A^{-1})$.

Démonstration.

Soit $A \in GL_n(\mathbb{K})$. Alors $AA^{-1} = I_n$, donc $I_n = {}^tI_n = {}^t(AA^{-1}) = {}^t(A^{-1}){}^tA$.

De même, on montre que ${}^tA {}^t(A^{-1}) = I_n$. \square

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

M est une **matrice symétrique** si et seulement si ${}^tM = M$.

M est une **matrice antisymétrique** si et seulement si ${}^tM = -M$.

Exemples. $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 2 \\ 1 & 2 & 4 \end{pmatrix}$ est symétrique et $\begin{pmatrix} 0 & a & 1 \\ -a & 0 & 2 \\ -1 & -2 & 0 \end{pmatrix}$ est antisymétrique.

Remarque. Lorsque $\text{car}(\mathbb{K}) \neq 2$, si $M \in \mathcal{M}_n(\mathbb{K})$ est antisymétrique, sa diagonale est nulle.

Notation. $\mathcal{S}_n(\mathbb{K})$ désigne l'ensemble des matrices symétriques d'ordre n .

$\mathcal{A}_n(\mathbb{K})$ désigne l'ensemble des matrices antisymétriques d'ordre n .

Propriété. $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$, mais ce ne sont pas des sous-algèbres. Cependant, elles sont stables par passage à l'inverse :

si $A \in \mathcal{S}_n(\mathbb{K}) \cap GL_n(\mathbb{K})$ (resp : $A \in \mathcal{A}_n(\mathbb{K}) \cap GL_n(\mathbb{K})$) ,

alors $A^{-1} \in \mathcal{S}_n(\mathbb{K})$ (resp : $A \in \mathcal{A}_n(\mathbb{K})$).

Démonstration.

Notons T l'opérateur de transposition. Alors $\mathcal{S}_n(\mathbb{K}) = \text{Ker}(T - \text{Id}_{\mathcal{M}_n(\mathbb{K})})$

et $\mathcal{A}_n(\mathbb{K}) = \text{Ker}(T + \text{Id}_{\mathcal{M}_n(\mathbb{K})})$, donc ce sont des sous-espaces vectoriels.

$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, donc $\mathcal{S}_n(\mathbb{K})$ n'est pas stable pour le produit : ce n'est pas un sous-anneau.

$I_n \notin \mathcal{A}_n(\mathbb{K})$, donc ce n'est pas un sous-anneau.

Soit A une matrice inversible. Si ${}^t A = A$, alors ${}^t(A^{-1}) = ({}^t A)^{-1} = A^{-1}$ donc A^{-1} est symétrique.

C'est analogue si A est antisymétrique. \square

1.5 Différentes interprétations du produit matriciel

Au niveau des coefficients :

Si $A \in \mathcal{M}_{\mathbb{K}}(n, \mathbf{p})$ et $B \in \mathcal{M}_{\mathbb{K}}(\mathbf{p}, q)$, $[AB]_{i,k} = \sum_{j=1}^p A_{i,j} B_{j,k}$.

Remarque. D'après cette formule, les coefficients de la k -ème colonne de AB ne dépendent que de A et de la k -ème colonne de B . Plus précisément, on voit que la k -ème colonne de AB est égale à AB_k si B_k désigne la k -ème colonne de B .

D'où l'interprétation suivante :

Au niveau des colonnes de la matrice de droite :

Soit $A \in \mathcal{M}_{\mathbb{K}}(n, \mathbf{p})$ et $B \in \mathcal{M}_{\mathbb{K}}(\mathbf{p}, q)$.

Notons B_1, \dots, B_q les colonnes de B , ce qui permet d'écrire $B = [B_1 | B_2 | \dots | B_q]$.

Alors $AB = [AB_1 | AB_2 | \dots | AB_q]$. En résumé, si B_1, \dots, B_q sont des vecteurs colonnes de \mathbb{K}^p , $A \times [B_1 | B_2 | \dots | B_q] = [AB_1 | AB_2 | \dots | AB_q]$.

En décomposant la matrice de droite en blocs de colonnes :

Soit $A \in \mathcal{M}_{\mathbb{K}}(n, \mathbf{p})$ et $B \in \mathcal{M}_{\mathbb{K}}(\mathbf{p}, q)$. Soit $r \in \{1, \dots, q-1\}$.

Notons B' la matrice constituée des r premières colonnes de B et B'' celle qui est constituée des colonnes suivantes : $B' = (B_{j,k})_{\substack{1 \leq j \leq p \\ 1 \leq k \leq r}}$ et $B'' = (B_{j,k+r})_{\substack{1 \leq j \leq p \\ 1 \leq k \leq q-r}}$.

Ainsi, on peut décomposer B en blocs : $B = [B' | B'']$. Alors $AB = [AB' | AB'']$.

En résumé, $A \times [B' | B''] = [AB' | AB'']$.

Au niveau des colonnes de la matrice de gauche :

— Si $M \in \mathcal{M}_{\mathbb{K}}(n, p)$ et $X \in \mathbb{K}^p$, MX est une combinaison linéaire des colonnes de

M . Plus précisément, si l'on note M_1, \dots, M_p les colonnes de M et $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$,

$$MX = x_1 M_1 + \dots + x_p M_p.$$

— Soient $A \in \mathcal{M}_{\mathbb{K}}(n, p)$ et $B \in \mathcal{M}_{\mathbb{K}}(p, q)$. Les colonnes de AB sont des combinaisons linéaires des colonnes de A : en notant A_1, \dots, A_p les colonnes de A et $B = (b_{i,j})$, la j -ème colonne de AB est égale à $b_{1,j}A_1 + \dots + b_{p,j}A_p$.

Exemple. Si $M \in \mathcal{M}_{\mathbb{K}}(n, p)$, la j -ème colonne de M est égale à $M \times (\delta_{i,j})_{1 \leq i \leq p}$.

La première colonne privée de la dernière est égale à $M \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}$.

Propriété. Pour tout $M \in \mathcal{M}_{\mathbb{K}}(n, p)$,

$\text{Im}(M)$ est l'espace vectoriel engendré par les colonnes de M .

Démonstration.

$$\text{Im}(M) = \{MX \mid X \in \mathbb{K}^p\} = \left\{ \sum_{j=1}^p x_j M_j \mid x_1, \dots, x_p \in \mathbb{K} \right\}. \square$$

Remarque. En prenant la transposée de ces différentes relations, on obtient des interprétations au niveau des lignes des matrices :

Au niveau des lignes de la matrice de gauche :

Soit $A \in \mathcal{M}_{\mathbb{K}}(n, p)$ et $B \in \mathcal{M}_{\mathbb{K}}(p, q)$. Notons ${}_1A, \dots, {}_nA$ les lignes de A , ce qui permet

d'écrire $A = \begin{pmatrix} \boxed{{}_1A} \\ \vdots \\ \boxed{{}_nA} \end{pmatrix}$. Alors $AB = \begin{pmatrix} \boxed{{}_1AB} \\ \vdots \\ \boxed{{}_nAB} \end{pmatrix}$. En résumé, si ${}_1A, \dots, {}_nA$ sont des

vecteurs lignes de taille n ,

$$\begin{pmatrix} \boxed{{}_1A} \\ \vdots \\ \boxed{{}_nA} \end{pmatrix} \times B = \begin{pmatrix} \boxed{{}_1AB} \\ \vdots \\ \boxed{{}_nAB} \end{pmatrix}.$$

Démonstration.

En transposant l'égalité $A \times \begin{bmatrix} B_1 & B_2 & \dots & B_q \end{bmatrix} = \begin{bmatrix} AB_1 & AB_2 & \dots & AB_q \end{bmatrix}$,

on obtient $\begin{pmatrix} \boxed{{}^tB_1} \\ \vdots \\ \boxed{{}^tB_q} \end{pmatrix} \times {}^tA = \begin{pmatrix} \boxed{{}^tB_1 {}^tA} \\ \vdots \\ \boxed{{}^tB_q {}^tA} \end{pmatrix}. \square$

En décomposant la matrice de gauche en blocs de lignes :

Soit $A \in \mathcal{M}_{\mathbb{K}}(n, p)$ et $B \in \mathcal{M}_{\mathbb{K}}(p, q)$. Soit $r \in \{1, \dots, n-1\}$.

Notons A' la matrice constituée des r premières lignes de A et A'' celle qui est constituée des lignes suivantes : $A' = (A_{i,j})_{\substack{1 \leq i \leq r \\ 1 \leq j \leq p}}$ et $A'' = (A_{i+r,j})_{\substack{1 \leq i \leq n-r \\ 1 \leq j \leq p}}$.

Ainsi, on peut décomposer A en blocs : $A = \begin{pmatrix} \boxed{A'} \\ \boxed{A''} \end{pmatrix}$. Alors $AB = \begin{pmatrix} \boxed{A'B} \\ \boxed{A''B} \end{pmatrix}$.

En résumé, $\begin{pmatrix} \boxed{A'} \\ \boxed{A''} \end{pmatrix} \times B = \begin{pmatrix} \boxed{A'B} \\ \boxed{A''B} \end{pmatrix}.$

Au niveau des lignes de la matrice de droite :

- Si $M \in \mathcal{M}_{\mathbb{K}}(n, p)$ et $X \in \mathcal{M}_{1,n}$, XM est une combinaison linéaire des lignes de M . Plus précisément, si l'on note ${}_1M, \dots, {}_nM$ les lignes de M et $X = (x_1 \ \dots \ x_n)$, $\boxed{XM = x_1 \times {}_1M + \dots + x_n \times {}_nM}$.
- Soient $A \in \mathcal{M}_{\mathbb{K}}(n, p)$ et $B \in \mathcal{M}_{\mathbb{K}}(p, q)$. Les lignes de AB sont des combinaisons linéaires des lignes de B : en notant ${}_1B, \dots, {}_pB$ les lignes de B et $A = (a_{i,j})$, $\boxed{\text{la } i^{\text{ème}} \text{ ligne de } AB \text{ est égale à } a_{i,1} \times {}_1B + \dots + a_{i,p} \times {}_pB}$.

Exemple. Notons $U \in \mathbb{K}^n$ le vecteur Attila, dont toutes les composantes sont égales à 1. Pour tout $A \in \mathcal{M}_n(\mathbb{K})$, AU est un vecteur colonne, obtenu en sommant toutes les

colonnes de A , tUA est un vecteur ligne, obtenu en sommant toutes les lignes de A , et tUAU est un scalaire, égal à la somme de tous les coefficients de A .

1.6 Trace d'une matrice

Définition. Soit $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$.

La **trace de la matrice** M est $\text{Tr}(M) = \sum_{i=1}^n m_{i,i}$.

Exemple. $\text{Tr}(I_n) = n$.

Propriété. La trace est une forme linéaire de $\mathcal{M}_n(\mathbb{K})$.

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

Démonstration.

Notons $A = (a_{i,j})$ et $B = (b_{i,j})$.

$$\text{Tr}(AB) = \sum_{i=1}^n \left(\sum_{k=1}^p a_{i,k} b_{k,i} \right) = \sum_{\substack{1 \leq i \leq n \\ 1 \leq k \leq p}} b_{k,i} a_{i,k} = \sum_{k=1}^p \left(\sum_{i=1}^n b_{k,i} a_{i,k} \right) = \text{Tr}(BA). \quad \square$$

Exemple. Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$. Alors $\text{Tr}(X {}^tX) = {}^tXX = \sum_{i=1}^n x_i^2$.

Remarque. Si $A \in \mathcal{M}_{n,p}(\mathbb{R})$, $\text{Tr}({}^tAA) = \sum_{i=1}^p \sum_{j=1}^n A_{j,i}^2$, donc $A = 0 \iff \text{Tr}({}^tAA) = 0$.

ATTENTION : Si $(A, B, C) \in \mathcal{M}_n(\mathbb{K})^3$, on peut écrire

$\text{Tr}(ABC) = \text{Tr}((AB)C) = \text{Tr}(C(AB)) = \text{Tr}(CAB)$, ou $\text{Tr}(ABC) = \text{Tr}(BCA)$, mais en général $\text{Tr}(ABC) \neq \text{Tr}(ACB)$.

Définition. Soit $A, B \in \mathcal{M}_n(\mathbb{K})$. On dit que A et B sont semblables si et seulement si il existe $P \in GL_n(\mathbb{K})$ telle que $B = PAP^{-1}$.

La relation de similitude ("être semblable à") est une relation d'équivalence sur $\mathcal{M}_n(\mathbb{K})$.

Remarque. Pour une matrice A donnée dans $\mathcal{M}_n(\mathbb{K})$, "réduire A ", c'est trouver une matrice semblable à A aussi simple que possible.

La théorie de la réduction des matrices est au centre du programme d'algèbre de seconde année.

Définition. Une matrice de $\mathcal{M}_n(\mathbb{K})$ est diagonalisable (resp : trigonalisable) si et seulement si elle est semblable à une matrice diagonale (resp : triangulaire supérieure).

Propriété. Deux matrices semblables ont la même trace, mais la réciproque est fausse.

Démonstration.

◇ Soient $(M, M') \in \mathcal{M}_n(\mathbb{K})$ un couple de matrices semblables. Il existe $P \in GL_n(\mathbb{K})$ tel que $M' = P^{-1}MP$. Ainsi $\text{Tr}(M') = \text{Tr}((P^{-1}M)P) = \text{Tr}(P(P^{-1}M)) = \text{Tr}(M)$.

◇ Prenons $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$: $\text{Tr}(A) = 0 = \text{Tr}(0)$, mais si A était semblable à la matrice nulle, il existerait $P \in GL_n(\mathbb{K})$ telle que $A = P0P^{-1} = 0$, ce qui est faux. \square

1.7 Matrices décomposées en blocs

1.7.1 Matrices extraites

Définition. Soit $n, p \in \mathbb{N}$ et soit I et J deux parties de \mathbb{N} telles que $|I| = n$ et $|J| = p$. Notons $0 \leq i_1 \leq i_2 \leq \dots \leq i_n$ les éléments de I et $0 \leq j_1 \leq j_2 \leq \dots \leq j_p$ les éléments de J .

Alors on convient d'identifier toute famille $(M_{i,j})_{(i,j) \in I \times J}$ de **scalaires** indexée par $I \times J$ avec la matrice $(M_{i_h, j_k})_{\substack{1 \leq h \leq n \\ 1 \leq k \leq p}} \in \mathcal{M}_{\mathbb{K}}(n, p)$.

Ainsi, on identifie globalement $\mathcal{M}_{\mathbb{K}}(n, p) = \mathbb{K}^{\mathbb{N}_n \times \mathbb{N}_p}$ avec $\mathbb{K}^{I \times J}$.

Exemple. Il est en particulier parfois pratique de faire débiter les indices de lignes et de colonnes à partir de 0.

Par exemple, $(\max(i, j))_{\substack{0 \leq i \leq 2 \\ 0 \leq j \leq 2}} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 2 & 2 \end{pmatrix}$.

Remarque. Lorsque I ou J est vide, $I \times J = \emptyset$ et $\mathbb{K}^{I \times J}$ possède un unique élément, que l'on appellera la matrice vide.

Définition. Soit $n, p \in \mathbb{N}^*$ et $M \in \mathcal{M}_{\mathbb{K}}(n, p)$. Une matrice extraite de M est une matrice de la forme $(M_{i,j})_{(i,j) \in I \times J}$, où $I \subset \mathbb{N}_n$ et $J \subset \mathbb{N}_p$.

Exemple. ...

1.7.2 Matrices blocs

Définition. Soient $(n_1, \dots, n_a) \in (\mathbb{N}^*)^a$ et $(p_1, \dots, p_b) \in (\mathbb{N}^*)^b$.

On pose $n = \sum_{i=1}^a n_i$ et $p = \sum_{j=1}^b p_j$.

Pour tout $(i, j) \in \mathbb{N}_a \times \mathbb{N}_b$, considérons une matrice $M_{i,j} \in \mathcal{M}_{\mathbb{K}}(n_i, p_j)$.

Alors la famille de ces matrices $M = (M_{i,j})_{\substack{1 \leq i \leq a \\ 1 \leq j \leq b}}$ peut être identifiée à une matrice possédant n lignes et p colonnes. On dit que M est une **matrice décomposée en blocs**, de dimensions (n_1, \dots, n_a) et (p_1, \dots, p_b) .

Exemple. Posons $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

$\begin{pmatrix} A & A & B \\ B & B & A \end{pmatrix}$ est une matrice décomposée en blocs.

Elle est égale à $\begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 1 \\ -1 & 0 & -1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & -1 & 0 \end{pmatrix}$.

Remarque. Choisissons $A = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ et $C = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$.

Au sens de la définition précédente, $\begin{pmatrix} A \\ B & C \end{pmatrix}$ n'est pas une matrice décomposée en blocs. Ainsi, le théorème ci-dessous, relatif au produit matriciel, ne s'applique pas à ce type de matrices. Cependant, cette décomposition peut être utilisée pour décrire une matrice.

Définition. La définition précédente peut être généralisée : Soit $n, p \in \mathbb{N}^*$. Soit $(I_i)_{1 \leq i \leq a}$ et $(J_j)_{1 \leq j \leq b}$ des partitions respectivement de \mathbb{N}_n et de \mathbb{N}_p . Alors on peut identifier toute matrice $M = (m_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ de $\mathcal{M}_{\mathbb{K}}(n, p)$ avec la famille des matrices extraites $(M_{i,j})_{\substack{1 \leq i \leq a \\ 1 \leq j \leq b}}$, où $M_{i,j} = (m_{h,k})_{(h,k) \in I_i \times J_j}$. On dit encore que $(M_{i,j})_{\substack{1 \leq i \leq a \\ 1 \leq j \leq b}}$ est une écriture par blocs de la matrice M , associée aux partitions $(I_i)_{1 \leq i \leq a}$ et $(J_j)_{1 \leq j \leq b}$. Avec ces notations, pour tout $\alpha, \beta \in \mathbb{N}_n \times \mathbb{N}_p$, $m_{\alpha,\beta} = [M_{i,j}]_{\alpha,\beta}$, où (i, j) est l'unique couple tel que $\alpha \in I_i$ et $\beta \in J_j$.

Définition. Reprenons les notations de la première définition.

La matrice $M = (M_{i,j})_{\substack{1 \leq i \leq a \\ 1 \leq j \leq b}}$ est une **matrice triangulaire supérieure par blocs** si et seulement si, pour tout $(i, j) \in \mathbb{N}_a \times \mathbb{N}_b$ tel que $i > j$, $M_{i,j} = 0$.

De même on définit la notion de matrice triangulaire inférieure par blocs.

La matrice $M = (M_{i,j})_{\substack{1 \leq i \leq a \\ 1 \leq j \leq b}}$ est une **matrice diagonale par blocs** si et seulement si, pour tout $(i, j) \in \mathbb{N}_a \times \mathbb{N}_b$ tel que $i \neq j$, $M_{i,j} = 0$.

1.7.3 Opérations sur les matrices blocs

Combinaison linéaire de matrices décomposées en blocs :

Soient $M = (M_{i,j})_{\substack{1 \leq i \leq a \\ 1 \leq j \leq b}}$ et $N = (N_{i,j})_{\substack{1 \leq i \leq a \\ 1 \leq j \leq b}}$ deux matrices décomposées en blocs selon les mêmes partitions $(I_i)_{1 \leq i \leq a}$ et $(J_j)_{1 \leq j \leq b}$ respectivement de \mathbb{N}_n et de \mathbb{N}_p .

Pour tout $(u, v) \in \mathbb{K}^2$, $uM + vN$ se décompose en blocs selon la formule suivante :

$$uM + vN = (uM_{i,j} + vN_{i,j})_{\substack{1 \leq i \leq a \\ 1 \leq j \leq b}}.$$

Démonstration.

Notons $M = (m_{\alpha,\beta})_{\substack{1 \leq \alpha \leq n \\ 1 \leq \beta \leq p}}$ et $N = (n_{\alpha,\beta})_{\substack{1 \leq \alpha \leq n \\ 1 \leq \beta \leq p}}$.

Soit $\alpha \in \mathbb{N}_n$ et $\beta \in \mathbb{N}_p$. Il existe un unique $(i, j) \in \mathbb{N}_a \times \mathbb{N}_b$ tel que $\alpha \in I_i$ et $\beta \in J_j$.

Alors $m_{\alpha,\beta} = [M_{i,j}]_{\alpha,\beta}$ et $n_{\alpha,\beta} = [N_{i,j}]_{\alpha,\beta}$, donc $u m_{\alpha,\beta} + v n_{\alpha,\beta} = [u M_{i,j} + v N_{i,j}]_{\alpha,\beta}$. \square

Produit matriciel de deux matrices décomposées en blocs : soit $n, p, q \in \mathbb{N}^*$.

Soit $M = (M_{i,j})_{\substack{1 \leq i \leq a \\ 1 \leq j \leq b}}$ une matrice décomposée en blocs selon les partitions $(I_i)_{1 \leq i \leq a}$ et $(J_j)_{1 \leq j \leq b}$ respectivement de \mathbb{N}_n et de \mathbb{N}_p .

Soit $N = (N_{j,k})_{\substack{1 \leq j \leq b \\ 1 \leq k \leq c}}$ une matrice décomposée en blocs selon **la même partition** $(J_j)_{1 \leq j \leq b}$ de \mathbb{N}_p et une partition $(K_k)_{1 \leq k \leq c}$ de \mathbb{N}_q .

Alors MN peut être vue comme une matrice décomposée en blocs selon les partitions $(I_i)_{1 \leq i \leq a}$ de \mathbb{N}_n et $(K_k)_{1 \leq k \leq c}$ de \mathbb{N}_q et :

$$MN = \left(\sum_{j=1}^b M_{i,j} N_{j,k} \right)_{\substack{1 \leq i \leq a \\ 1 \leq k \leq c}}.$$

En résumé, le produit de deux matrices par blocs se comporte comme le produit matriciel usuel.

Démonstration.

Notons $M = (m_{\alpha,\beta})_{\substack{1 \leq \alpha \leq n \\ 1 \leq \beta \leq p}}$ et $N = (n_{\beta,\gamma})_{\substack{1 \leq \beta \leq p \\ 1 \leq \gamma \leq q}}$. Soit $\alpha, \gamma \in \mathbb{N}_n \times \mathbb{N}_q$.

Il s'agit de montrer que $\sum_{\beta=1}^p m_{\alpha,\beta} n_{\beta,\gamma} = \left[\sum_{j=1}^b M_{i,j} N_{j,k} \right]_{\alpha,\gamma}$, où (i, k) est l'unique couple

tel que $\alpha \in I_i$ et $\gamma \in K_k$. Or,

$$\begin{aligned} \left[\sum_{j=1}^b M_{i,j} N_{j,k} \right]_{\alpha,\gamma} &= \sum_{j=1}^b [M_{i,j} N_{j,k}]_{\alpha,\gamma} = \sum_{j=1}^b \sum_{\beta \in J_j} [M_{i,j}]_{\alpha,\beta} [N_{j,k}]_{\beta,\gamma}, \text{ donc} \\ \left[\sum_{j=1}^b M_{i,j} N_{j,k} \right]_{\alpha,\gamma} &= \sum_{j=1}^b \sum_{\beta \in J_j} m_{\alpha,\beta} n_{\beta,\gamma} = \sum_{\beta=1}^p m_{\alpha,\beta} n_{\beta,\gamma}. \quad \square \end{aligned}$$

Application : Produit de matrices triangulaires (resp : diagonales) par blocs, puisances de telles matrices.

2 Familles de vecteurs

Notation. Pour ce chapitre, on fixe un \mathbb{K} -espace vectoriel E et un ensemble quelconque I (éventuellement infini).

2.1 Familles libres et génératrices

Définition. Soit $(x_i)_{i \in I}$ une famille de vecteurs de E .

- Elle est **libre** si et seulement si

$$\boxed{\forall (\alpha_i)_{i \in I} \in \mathbb{K}^{(I)} \quad \left(\sum_{i \in I} \alpha_i x_i = 0 \implies (\forall i \in I \quad \alpha_i = 0) \right)}.$$

- Elle est **liée** si et seulement si elle n'est pas libre, c'est-à-dire si et seulement si

$$\boxed{\exists (\alpha_i)_{i \in I} \in \mathbb{K}^{(I)} \setminus \{0\} \quad \sum_{i \in I} \alpha_i x_i = 0.}$$

- Elle est **génératrice** dans E si et seulement si

$$\forall x \in E \quad \exists (\alpha_i)_{i \in I} \in \mathbb{K}^{(I)} \quad \sum_{i \in I} \alpha_i x_i = x.$$

Ainsi $(x_i)_{i \in I}$ est toujours génératrice dans $\text{Vect}(x_i)_{i \in I}$.

- C'est une **base** de E si et seulement si elle est libre et génératrice dans E .

Remarque. Dans la définition d'une famille liée, la condition " $(\alpha_i)_{i \in I} \neq 0$ " est essentielle. En effet, si on l'oublie, comme $0 = \sum_{i \in I} 0 \cdot x_i$, toute famille serait liée.

Remarque. Lorsque $I = \emptyset$, on dit que $(x_i)_{i \in I}$ est la famille vide.

La famille vide est libre, car la propriété " $\forall i \in \emptyset \quad \alpha_i = 0$ " est vraie.

D'après les conventions du cours sur les groupes, l'unique combinaison linéaire de cette famille est 0, donc la famille vide n'est pas génératrice dans E , sauf si $E = \{0\}$.

En particulier, on remarquera que la famille vide est l'unique base de $\{0\}$.

Exemple. Dans $\mathbb{K}[X]$, la famille $(1, X + 1, X - 1)$ est liée.

Dans $\mathbb{R}^{\mathbb{R}}$, la famille (\cos, \sin) est libre, car si $\alpha \sin + \beta \cos = 0$, alors $0 = \alpha \sin(0) + \beta \cos(0) = \beta$ et $0 = \alpha \sin(\frac{\pi}{2}) + \beta \cos(\frac{\pi}{2}) = \alpha$.

Définition. Deux vecteurs x et y d'un \mathbb{K} -espace vectoriel E sont **colinéaires** si et seulement si la famille (x, y) est liée.

Propriété. Soit $e = (e_i)_{i \in I}$ une famille de vecteurs de E .

e est une base de E si et seulement si

$$\forall x \in E \quad \exists ! (\alpha_i)_{i \in I} \in \mathbb{K}^{(I)} \quad \sum_{i \in I} \alpha_i e_i = x.$$

Dans ce cas, pour $x \in E$, on appelle coordonnées de x dans la base $(e_i)_{i \in I}$ l'unique famille presque nulle de scalaire $(\alpha_i)_{i \in I}$ telle que $x = \sum_{i \in I} \alpha_i e_i$.

2.2 Dimension d'un espace vectoriel

Définition. E est de dimension finie si et seulement s'il admet une famille génératrice contenant un nombre fini d'éléments. Sinon, on dit qu'il est de dimension infinie.

Remarque. On dira qu'un espace vectoriel est de "dimension quelconque" lorsqu'il est de dimension finie ou de dimension infinie.

Exemple.

$\mathbb{R}^2 = \{x(1, 0) + y(0, 1) \mid (x, y) \in \mathbb{R}^2\} = \text{Vect}((1, 0), (0, 1))$ est de dimension finie.

$\mathbb{K}_n[X] = \text{Vect}(1, X, \dots, X^n)$ est de dimension finie.

Lemme : Soit $n \in \mathbb{N}$ et $e_1, \dots, e_n \in E$.

Toute famille (x_1, \dots, x_{n+1}) de $n + 1$ vecteurs de $\text{Vect}(e_1, \dots, e_n)$ est liée.

Démonstration.

Soit $n \in \mathbb{N}$. On note $R(n)$ l'assertion : pour tout $e_1, \dots, e_n \in E$, toute famille (x_1, \dots, x_{n+1}) de $n+1$ vecteurs de $\text{Vect}(e_1, \dots, e_n)$ est liée.

◇ Pour $n = 0$, si $x_1 \in \text{Vect}(\emptyset) = \{0\}$, x_1 est nul, donc (x_1) est une famille liée.

◇ Supposons que $n \geq 1$ et que $R(n-1)$ est vraie.

Soit $e_1, \dots, e_n \in E$ et (x_1, \dots, x_{n+1}) une famille de $n+1$ vecteurs de $\text{Vect}(e_1, \dots, e_n)$.

Pour tout $j \in \{1, \dots, n+1\}$, il existe $(\alpha_{i,j})_{1 \leq i \leq n} \in \mathbb{K}^n$ tel que $x_j = \sum_{i=1}^n \alpha_{i,j} e_i$.

Premier cas : si, pour tout $j \in \{1, \dots, n+1\}$, $\alpha_{n,j} = 0$, alors (x_1, \dots, x_n) est une famille de n vecteurs de $\text{Vect}(e_1, \dots, e_{n-1})$, donc d'après $R(n-1)$ elle est liée. Alors (x_1, \dots, x_{n+1}) est aussi liée.

Second cas : supposons maintenant qu'il existe $j_0 \in \{1, \dots, n+1\}$ tel que $\alpha_{n,j_0} \neq 0$: nous allons l'utiliser comme un pivot. Quitte à réordonner les vecteurs x_1, \dots, x_{n+1} , on peut supposer que $j_0 = n+1$. Ainsi, $\alpha_{n,n+1} \neq 0$.

Pour tout $j \in \{1, \dots, n\}$, posons $y_j = x_j - \frac{\alpha_{n,j}}{\alpha_{n,n+1}} x_{n+1} : y_j \in \text{Vect}(e_1, \dots, e_{n-1})$, donc d'après $R(n-1)$, la famille (y_1, \dots, y_n) est liée. Ainsi, il existe $(\beta_1, \dots, \beta_n) \in \mathbb{K}^n \setminus \{0\}$ tel que $0 = \sum_{j=1}^n \beta_j y_j = \sum_{j=1}^n \beta_j x_j + \lambda x_{n+1}$, où $\lambda \in \mathbb{K}$. La famille de scalaires $(\beta_1, \dots, \beta_n, \lambda)$ est non nulle, donc (x_1, \dots, x_{n+1}) est liée. □

Corollaire. Si (e_1, \dots, e_n) est une famille génératrice de E , alors toute famille libre de E est de cardinal inférieur ou égal à n .

Théorème de la base incomplète : Soient E un \mathbb{K} -espace vectoriel de dimension finie et $(e_i)_{i \in I}$ une famille génératrice de E (on ne suppose pas qu'elle contient un nombre fini d'éléments). Soit $J \subset I$ tel que $(e_i)_{i \in J}$ est une famille libre.

Alors il existe un ensemble L avec $J \subset L \subset I$ tel que $(e_i)_{i \in L}$ est une base de E .

Cela signifie que toute famille libre f de E peut être complétée en une base de E à l'aide de vecteurs d'une famille génératrice de E (Inutile : qui contient f).

Démonstration.

On note D l'ensemble des cardinaux des familles libres $(e_i)_{i \in K}$, où $J \subset K \subset I$.

D est une partie non vide de \mathbb{N} . De plus, E est de dimension finie, donc il possède une famille génératrice finie. Notons n son cardinal. Alors d'après le corollaire précédent, D est majorée par n . D possède donc un plus grand élément, noté m et il existe une famille libre $b = (e_i)_{i \in L}$ avec $J \subset L \subset I$ tel que $|L| = m$.

Supposons que, pour tout $i_0 \in I \setminus L$, $e_{i_0} \in \text{Vect}(b)$. Alors pour tout $i \in I$, $e_i \in \text{Vect}(b)$, donc $E = \text{Vect}(e_i)_{i \in I} \subset \text{Vect}(b)$, donc b est génératrice et libre, ce qui prouve que c'est une base de E . Il suffit donc de montrer que pour tout $i_0 \in I \setminus L$, $e_{i_0} \in \text{Vect}(b)$.

Raisonnons par l'absurde en supposant qu'il existe $i_0 \in I \setminus L$ tel que $e_{i_0} \notin \text{Vect}(b)$.

Posons $L' = L \cup \{i_0\}$. Montrons que $b' = (e_i)_{i \in L \cup \{i_0\}}$ est libre, ce qui contredira la définition de m . Soit $(\alpha_i)_{i \in L \cup \{i_0\}}$ une famille de scalaires telle que $\sum_{i \in L \cup \{i_0\}} \alpha_i e_i = 0$.

Si $\alpha_{i_0} \neq 0$, on peut écrire $e_{i_0} = -\frac{1}{\alpha_{i_0}} \sum_{i \in L} \alpha_i e_i$, ce qui est exclu car $e_{i_0} \notin \text{Vect}(b)$.

Ainsi, $\alpha_{i_0} = 0$, donc $\sum_{i \in L} \alpha_i e_i = 0$, or b est libre, donc pour tout $i \in L$, $\alpha_i = 0$. \square

Remarque. En adaptant la démonstration, on a également prouvé les propriétés suivantes, en dimension quelconque :

Propriété. Soit $(e_i)_{i \in I}$ une famille libre de vecteurs de E . Soit $e_j \in E$, où $j \notin I$. La famille $(e_i)_{i \in I \cup \{j\}}$ est libre si et seulement si $e_j \notin \text{Vect}(e_i)_{i \in I}$.

Propriété.

Soient E un \mathbb{K} -espace vectoriel et $g = (e_i)_{i \in I}$ une famille génératrice de E .

On dit qu'une sous-famille libre $(e_i)_{i \in J}$ de g est maximale dans g si et seulement si pour tout $i_0 \in I \setminus J$, la famille $(e_i)_{i \in J \cup \{i_0\}}$ est liée.

Si $(e_i)_{i \in J}$ est libre maximale dans g , alors c'est une base de E .

Corollaire. Une famille libre de vecteurs de E est maximale si et seulement si en lui ajoutant un vecteur elle devient liée.

Toute famille libre maximale de vecteurs de E est une base de E .

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

Toute famille libre de E peut être complétée en une base de E .

Démonstration.

On applique le théorème de la base incomplète en prenant comme famille génératrice de E la famille $(x)_{x \in E}$ de tous les vecteurs de E . \square

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

E admet au moins une base. Toutes les bases de E sont finies et ont même cardinal. Ce cardinal est appelé la **dimension** de E et est noté $\dim(E)$ ou $\dim_{\mathbb{K}}(E)$.

Démonstration.

La famille vide est libre et d'après le corollaire précédent, on peut la compléter en une base.

Soit b et b' deux bases de E , de cardinaux n et n' .

b est génératrice de E et b' est libre, donc $n' = |b'| \leq |b| = n$.

De même on montre que $n \leq n'$. \square

Propriété. Soit E un \mathbb{K} -espace vectoriel de dimension finie égale à n et soit e une famille de E . Les propriétés suivantes sont équivalentes.

1. e est une base de E .
2. e est libre et de cardinal n .
3. e est génératrice et de cardinal n .

Démonstration.

$1 \implies 2$ et $1 \implies 3$ sont claires.

$2 \implies 1$: supposons que e est libre et de cardinal n . On peut la compléter en une base b de E , mais b est alors de cardinal n , donc $e = b$ et e est bien une base de E .

$3 \implies 1$: supposons que e est génératrice et de cardinal n . On peut en extraire une base b de E , mais b est alors de cardinal n , donc $e = b$ et e est bien une base de E . \square

Propriété. Soit E un \mathbb{K} -espace vectoriel de dimension finie égale à n .

Toute famille libre de E a au plus n éléments et

toute famille génératrice de E a au moins n éléments.

Remarque. Tout espace de dimension finie possède donc au moins une base. C'est aussi le cas pour les espaces vectoriels de dimension infinie, si l'on accepte l'axiome du choix. Cependant, on peut construire des modèles de ZF dans lesquels \mathbb{R} n'admet aucune base en tant que \mathbb{Q} -espace vectoriel.

Exemples. Soit E un \mathbb{K} -espace vectoriel de dimension finie ou infinie.

◇ $\{0\}$ est un sous-espace vectoriel de E de dimension 0 car \emptyset est une base de $\{0\}$.

◇ Si $u \in E \setminus \{0\}$, alors (u) est une base de $\text{Vect}(u) = \{\lambda u / \lambda \in \mathbb{K}\}$, donc la droite vectorielle $\text{Vect}(u)$ est de dimension 1.

◇ Si (u, v) est une famille libre de vecteurs de E , $\text{Vect}(u, v) = \{\alpha u + \beta v / \alpha, \beta \in \mathbb{K}\}$ admet comme base (u, v) , donc il est de dimension 2. On dit que $\text{Vect}(u, v)$ est le *plan vectoriel* engendré par (u, v) .

Exemple.

◇ Soit I un intervalle d'intérieur non vide de \mathbb{R} et $b : I \longrightarrow \mathbb{C}$ une application continue. Alors l'ensemble des solutions de l'équation différentielle $(E) : y' = b(t)y$ est la droite vectorielle engendrée par l'application $t \longmapsto e^{B(t)}$, où B est une primitive de b . C'est un sous-espace vectoriel du \mathbb{C} -espace vectoriel $C^1(I, \mathbb{C})$.

◇ Soit $a, b \in \mathbb{R}$ tels que $a^2 - 4b > 0$. Alors l'ensemble des solutions de l'équation différentielle $(E) : y'' + ay' + by = 0$ est le plan vectoriel engendré par $(t \longmapsto e^{r_1 t}, t \longmapsto e^{r_2 t})$, où r_1 et r_2 sont les racines du polynôme caractéristique $X^2 + aX + b$. C'est un sous-espace vectoriel du \mathbb{R} -espace vectoriel $C^\infty(\mathbb{R}, \mathbb{R})$.

Théorème. Soit E un \mathbb{K} -espace vectoriel de dimension quelconque.

Soit F et G deux sous-espaces vectoriels de E avec G de dimension finie et $F \subset G$.

Alors F est de dimension finie avec $\dim(F) \leq \dim(G)$.

De plus $[F = G \iff \dim(F) = \dim(G)]$.

Démonstration.

Toute famille libre de F est une famille libre de G , donc elle est de cardinal inférieure à $\dim(G)$. Ainsi, l'ensemble D des cardinaux des familles libres de F est une partie non vide et majorée de \mathbb{N} , donc elle possède un maximum noté m . Il existe une famille libre de vecteurs de F de cardinal m . Par construction c'est une famille libre maximale de F , donc c'est une base de F . Ainsi F est de dimension finie et $\dim(F) = m \leq n$.

Supposons que $\dim(F) = \dim(G)$. Alors la famille maximale précédente est une famille libre de G de cardinal égal à la dimension de G , donc c'est aussi une base de G . Alors cette famille engendre à la fois F et G , donc $F = G$. \square

Remarque. Ainsi, en dimension finie, pour montrer que deux sous-espaces vectoriels sont égaux, il suffit de montrer une inclusion et l'égalité des dimensions.

Propriété. \mathbb{C} est un \mathbb{R} -espace vectoriel de dimension 2, dont $(1, i)$ est une base.

Démonstration.

Pour tout $z \in \mathbb{C}$, il existe $x, y \in \mathbb{R}$ tel que $z = x + iy$, donc la famille de complexes $(1, i)$ est une famille génératrice de \mathbb{C} , vu comme un \mathbb{R} -espace vectoriel.

De plus, si $x + iy = 0$ avec $x, y \in \mathbb{R}$, on sait alors que $x = y = 0$, donc $(1, i)$ est une famille libre.

Ainsi $(1, i)$ est une base du \mathbb{R} -espace vectoriel \mathbb{C} . Ceci prouve que \mathbb{C} est un \mathbb{R} -espace vectoriel de dimension finie et que $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.

Dans la base $(1, i)$, les coordonnées d'un complexe sont ses parties réelle et imaginaire. \square

Exemple. Prenons $E = \mathbb{R}^2$ et posons $c_1 = (1, 0)$ et $c_2 = (0, 1)$. Montrons que $c = (c_1, c_2)$ est une base de E .

Pour tout $(a, b) \in E$, $(a, b) = a(1, 0) + b(0, 1)$, donc c est une famille génératrice de E .

De plus, si $ac_1 + bc_2 = 0$, avec $a, b \in \mathbb{R}$, alors $0 = ac_1 + bc_2 = (a, b)$, donc $a = b = 0$. Ainsi c est une base de \mathbb{R}^2 , que l'on appelle la base canonique de \mathbb{R}^2 .

En particulier, on a montré que \mathbb{R}^2 est de dimension finie et que $\dim(\mathbb{R}^2) = 2$.

La forte ressemblance avec la démonstration précédente provient du fait que l'application
$$\begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{C} \\ (a, b) & \longmapsto & a + ib \end{array}$$
 est un isomorphisme de \mathbb{R} -espace vectoriel.

On peut généraliser sans difficulté à \mathbb{K}^n :

2.3 Base canonique

Propriété. Soit $n \in \mathbb{N}^*$. \mathbb{K}^n est un \mathbb{K} -espace vectoriel de dimension n dont une base est $c = (c_1, \dots, c_n)$, où pour tout $i \in \{1, \dots, n\}$, $c_i = (\delta_{i,j})_{1 \leq j \leq n}$. c est appelé la *base canonique* de \mathbb{K}^n .

Pour tout $x = (x_1, \dots, x_n) \in \mathbb{K}^n$, les coordonnées de x dans la base c sont exactement ses composantes x_1, \dots, x_n .

On peut encore généraliser :

Propriété. Soit I un ensemble quelconque.

Pour tout $i \in I$, on note $c_i = (\delta_{i,j})_{j \in I}$. Ainsi $c = (c_i)_{i \in I}$ est une famille de $\mathbb{K}^{(I)}$.

C'est une base de $\mathbb{K}^{(I)}$, appelée la *base canonique* de $\mathbb{K}^{(I)}$.

De plus, pour tout $x = (\alpha_i)_{i \in I} \in \mathbb{K}^{(I)}$, $x = \sum_{i \in I} \alpha_i c_i$. Ainsi, les *coordonnées* de x dans la base canonique sont exactement ses composantes.

Démonstration.

Soit $(\alpha_i) \in \mathbb{K}^{(I)}$. Pour tout $i \in I$, $\alpha_i = \sum_{j \in I} \delta_{i,j} \alpha_j$,

donc $(\alpha_i)_{i \in I} = (\sum_{j \in I} \delta_{i,j} \alpha_j)_{i \in I} = \sum_{j \in I} \alpha_j (\delta_{i,j})_{i \in I}$.

Ainsi $(\alpha_i)_{i \in I} = \sum_{j \in I} \alpha_j c_j$.

Ceci prouve que c est une famille génératrice de $\mathbb{K}^{(I)}$.

De plus, si $\sum_{j \in I} \alpha_j c_j = 0$, $(\alpha_i)_{i \in I} = 0$, donc, pour tout $i \in I$, $\alpha_i = 0$, ce qui prouve que c est aussi une famille libre. \square

Corollaire. La base canonique de $\mathbb{K}[X]$ est la famille $(X^n)_{n \in \mathbb{N}}$.

Ainsi, pour $P \in \mathbb{K}[X]$, l'écriture $P = \sum_{k \in \mathbb{N}} a_k X^k$ est la décomposition de P dans la base canonique de $\mathbb{K}[X]$.

Démonstration.

$\mathbb{K}[X] = \mathbb{K}^{(\mathbb{N})}$ et pour tout $n \in \mathbb{N}$, $X^n = (\delta_{i,n})_{i \in \mathbb{N}}$. \square

Corollaire. Soit $n \in \mathbb{N}$. $(1, X, \dots, X^n)$ est une base de $\mathbb{K}_n[X]$, encore appelée la base canonique de $\mathbb{K}_n[X]$. On en déduit que $\dim(\mathbb{K}_n[X]) = n + 1$.

Corollaire. Soit $n, p \in \mathbb{N}^*$. La base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$ est la famille de matrices $(E_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ définie par : Pour tout $i \in \{1, \dots, n\}$ et $j \in \{1, \dots, p\}$, $E_{i,j} = (\delta_{a,i} \delta_{b,j})_{\substack{1 \leq a \leq n \\ 1 \leq b \leq p}}$. $E_{i,j}$ est appelée la (i, j) -ième matrice élémentaire de $\mathcal{M}_{n,p}(\mathbb{K})$. Tous ses coefficients sont nuls, sauf celui de position (i, j) qui est égal à 1.

Ainsi, pour tout $M \in \mathcal{M}_{n,p}(\mathbb{K})$, $M = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} M_{i,j} E_{i,j}$.

On en déduit que $\dim(\mathcal{M}_{n,p}(\mathbb{K})) = np$.

Exercice. Lorsque $n = p$, calculer $E_{i,j} E_{h,k}$.

Exercice. Soit $M \in \mathcal{M}_n(\mathbb{K})$. Montrer qu'il existe $P \in \mathbb{K}[X] \setminus \{0\}$ tel que $P(M) = 0$. Un tel polynôme est dit annulateur de M . Que peut-on dire de l'ensemble des polynômes annulateurs de M ?

2.4 Exemples

Propriété. Dans \mathbb{K}^2 , deux vecteurs $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ et $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ forment une base de \mathbb{K}^2 si et seulement si $u_1 v_2 - u_2 v_1 \stackrel{\Delta}{=} \det_c(u, v) \neq 0$.

Démonstration.

$\dim(\mathbb{K}^2) = 2$, donc

$$\begin{aligned}
(u, v) \text{ libre} &\iff (u, v) \text{ est une base de } \mathbb{K}^2 \\
&\iff \forall w \in \mathbb{K}^2, \exists! (\alpha, \beta) \in \mathbb{K}^2, \alpha u + \beta v = w \\
&\iff \forall w \in \mathbb{K}^2, \exists! \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in \mathbb{K}^2, (u|v) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = w,
\end{aligned}$$

où $(u|v)$ désigne la matrice dont les colonnes sont u et v . Ainsi,

$$\begin{aligned}
(u, v) \text{ libre} &\iff \widetilde{(u|v)} \in GL(\mathbb{K}^2) \\
&\iff (u|v) \in GL_2(\mathbb{K}) \quad \square \\
&\iff \det(u|v) \neq 0 \\
&\iff u_1 v_2 - u_2 v_1 \neq 0.
\end{aligned}$$

Propriété. Soient $n \in \mathbb{N}$ et $(a_0, a_1, \dots, a_n) \in \mathbb{K}^{n+1}$ une famille de $n+1$ scalaires *deux à deux distincts*. Pour tout $i \in \{0, \dots, n\}$, notons $L_i = \prod_{\substack{0 \leq j \leq n \\ j \neq i}} \frac{X - a_j}{a_i - a_j}$ le i -ème polynôme

d'interpolation de Lagrange associé à la famille (a_0, a_1, \dots, a_n) .

Alors $L = (L_i)_{0 \leq i \leq n}$ est une base de $\mathbb{K}_n[X]$, appelée la base de Lagrange associé à la famille (a_0, a_1, \dots, a_n) .

De plus, pour tout $P \in \mathbb{K}_n[X]$, la décomposition de P dans cette base s'écrit

$$P = \sum_{i=0}^n \tilde{P}(a_i) L_i : \text{les } \tilde{P}(a_i) \text{ sont les coordonnées de } P \text{ dans la base } L.$$

Démonstration.

Soit $P \in \mathbb{K}_n[X]$. $P - \sum_{i=0}^n \tilde{P}(a_i) L_i$ admet au moins pour racines a_0, \dots, a_n , ce qui fait au moins $n+1$ racines, or c'est un polynôme de degré inférieur ou égal à n , donc il est identiquement nul. Ceci prouve que L est une famille génératrice de $\mathbb{K}_n[X]$. De plus, $|L| = \dim(\mathbb{K}_n[X])$, donc c'est une base de $\mathbb{K}_n[X]$. \square

Exercice. Soit $(P_n)_{n \in \mathbb{N}}$ une suite de polynômes de $\mathbb{K}[X]$. On suppose que cette suite de polynômes est étagée c'est-à-dire que, $\forall n \in \mathbb{N} \quad \deg(P_n) = n$.

Montrer que pour tout $N \in \mathbb{N}$, $(P_n)_{0 \leq n \leq N}$ est une base de $\mathbb{K}_N[X]$.

En déduire que $(P_n)_{n \in \mathbb{N}}$ est une base de $\mathbb{K}[X]$.

Solution :

◇ Soit $N \in \mathbb{N}$. Soit $(\alpha_n)_{0 \leq n \leq N}$ une famille de scalaires telle que $\sum_{n=0}^N \alpha_n P_n = 0$.

Supposons que cette famille est non nulle. Alors $\{n \in \{0, \dots, N\} / \alpha_n \neq 0\}$ est un ensemble non vide inclus dans \mathbb{N} et majoré par N , donc il admet un maximum que l'on note p .

Ainsi, $P_p = -\frac{1}{\alpha_p} \sum_{n=0}^{p-1} \alpha_n P_n$, donc $p = \deg(P_p) \leq \max_{0 \leq n < p} \deg(\alpha_n P_n) \leq p-1$. C'est

faux, donc la famille $(\alpha_n)_{0 \leq n \leq N}$ est nulle. Ainsi la famille $(P_n)_{0 \leq n \leq N}$ est une famille libre de $\mathbb{K}_N[X]$ de cardinal $N+1$. Or la famille $(X^n)_{0 \leq n \leq N}$ est la base canonique de $\mathbb{K}_N[X]$, donc $\mathbb{K}_N[X]$ est un espace vectoriel de dimension $N+1$. Ainsi $(P_n)_{0 \leq n \leq N}$ est une base de $\mathbb{K}_N[X]$.

- Soit $(\alpha_n)_{n \in \mathbb{N}} \in \mathbb{K}^{(\mathbb{N})}$ telle que $\sum_{n \in \mathbb{N}} \alpha_n P_n = 0$. Cette famille de scalaires étant presque nulle, il existe $N \in \mathbb{N}$ tel que pour tout $n > N$, $\alpha_n = 0$.

Ainsi $\sum_{n=0}^N \alpha_n P_n = 0$, or la famille $(P_n)_{0 \leq n \leq N}$ est libre, donc pour tout $n \leq N$, $\alpha_n = 0$. Ainsi la famille $(P_n)_{n \in \mathbb{N}}$ est libre.

- *Remarque.* Plus généralement, le même argument permet de démontrer qu'une famille $(x_i)_{i \in I}$ de vecteurs d'un espace vectoriel est libre si et seulement si toute sous-famille finie de $(x_i)_{i \in I}$ est libre.

- Il reste à montrer que $(P_n)_{n \in \mathbb{N}}$ est une famille génératrice de $\mathbb{K}[X]$.

Soit $P \in \mathbb{K}[X]$. Notons N son degré. $P \in \mathbb{K}_N[X]$, donc il existe $(\alpha_n)_{0 \leq n \leq N} \in \mathbb{K}^{N+1}$

telle que $P = \sum_{n=0}^N \alpha_n P_n$. On complète cette famille finie en la suite $(\alpha_n)_{n \in \mathbb{N}}$ en

posant $\alpha_n = 0$ pour tout $n > N$. Ainsi $P = \sum_{n \in \mathbb{N}} \alpha_n P_n$. Donc la famille $(P_n)_{n \in \mathbb{N}}$ engendre $\mathbb{K}[X]$.

Exercice. Soit f un endomorphisme de E . Montrer que f est une homothétie si et seulement si pour tout $u \in E$, $(u, f(u))$ est lié.

Propriété. Toute sur-famille d'une famille génératrice est génératrice.

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

Propriété. Une famille de vecteurs est libre si et seulement si toute sous-famille finie de cette famille est libre.

Exercice. Pour $n \in \mathbb{N}^*$, on pose
$$\begin{aligned} f_n : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{x^2 + n^2} \end{aligned}$$
 La famille $(f_n)_{n \in \mathbb{N}^*}$ est-elle libre ?

Solution : Soit $(\alpha_n)_{n \in \mathbb{N}^*} \in \mathbb{R}^{(\mathbb{N}^*)}$ telle que $\sum_{n \in \mathbb{N}^*} \alpha_n f_n = 0$.

On dispose ainsi de deux décompositions en éléments simples de la fraction nulle dans $\mathbb{R}(X)$. D'après l'unicité d'une telle décomposition, les α_n sont tous nuls. La famille $(f_n)_{n \in \mathbb{N}^*}$ est donc libre.

Exercice. Pour $n \in \mathbb{N}$, on pose
$$\begin{aligned} f_n : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \sin(x^n) \end{aligned}$$
 La famille $(f_n)_{n \in \mathbb{N}}$ est-elle libre ?

Solution : Supposons que $\sum_{n \in \mathbb{N}} \alpha_n f_n = 0$.

Ainsi, pour tout $x \in \mathbb{R}$, (1) : $\sum_{n \in \mathbb{N}} \alpha_n f_n(x) = 0$.

[Pour montrer qu'une famille de fonctions est libre, il existe d'autres méthodes que celle consistant à substituer x par des valeurs bien choisies pour se ramener à un système linéaire dont les inconnues sont les α_n . On peut en effet faire de l'analyse, c'est-à-dire dériver ou intégrer la

relation (1). On peut également faire un développement limité de (1) au voisinage d'un point bien choisi.]

Pour tout $n \in \mathbb{N}^*$, $x^n \xrightarrow{x \rightarrow 0} 0$, donc $f_n(x) = x^n + o(x^n)$.

Ainsi, $0 = \alpha_0 \sin(1) + \sum_{n \geq 1} (\alpha_n x^n + o(x^n))$.

On suppose qu'il existe un $\alpha_n \neq 0$ avec $n \geq 1$.

Appelons k le plus petit $n \geq 1$ tel que $\alpha_n \neq 0$.

Alors $0 = \alpha_0 \sin(1) + \alpha_k x^k + o(x^k)$. D'après l'unicité du développement limité (de l'application identiquement nulle), $\alpha_k = 0$, ce qui est faux.

Ainsi les α_n sont nuls pour $n \geq 1$. On en déduit que $\alpha_0 \sin(1) = 0$, donc que α_0 est aussi nul. La famille $(f_n)_{n \in \mathbb{N}}$ est donc libre.

Théorème. Si E_1, \dots, E_n sont n sous-espaces vectoriels de dimensions finies, alors $E_1 \times \dots \times E_n$ est de dimension finie et

$$\dim(E_1 \times \dots \times E_n) = \dim(E_1) + \dots + \dim(E_n).$$

Démonstration.

Le cas général se déduit du cas où $n = 2$ par récurrence car

$E_1 \times \dots \times E_n = (E_1 \times \dots \times E_{n-1}) \times E_n$ (Logique et vocabulaire ensembliste page 7).

Soit donc E et F deux espaces vectoriels de dimensions p et q . Notons $e = (e_1, \dots, e_p)$ une base de E et $f = (f_1, \dots, f_q)$ une base de F .

Pour tout $(x, y) \in E \times F$, avec $x = \sum_{i=1}^p x_i e_i$ et $y = \sum_{j=1}^q y_j f_j$,

$$(x, y) = (x, 0) + (0, y) = \left(\sum_{i=1}^p x_i e_i, 0 \right) + \left(0, \sum_{j=1}^q y_j f_j \right), \text{ donc}$$

$$(x, y) = \sum_{i=1}^p x_i (e_i, 0) + \sum_{j=1}^q y_j (0, f_j) : \text{ ceci démontre que la concaténation des familles}$$

$((e_1, 0), \dots, (e_p, 0))$ et $((0, f_1), \dots, (0, f_q))$ est une famille génératrice de $E \times F$.

Notons g cette famille.

Pour montrer qu'elle est libre, supposons que $\sum_{i=1}^p x_i (e_i, 0) + \sum_{j=1}^q y_j (0, f_j) = 0$, où

$(x_i)_{1 \leq i \leq p}$ et $(y_j)_{1 \leq j \leq q}$ sont deux familles de scalaires. D'après le calcul précédent,

$$(x, y) = 0, \text{ où } x = \sum_{i=1}^p x_i e_i \text{ et } y = \sum_{j=1}^q y_j f_j, \text{ or } e \text{ et } f \text{ sont libres, donc pour tout } i$$

et j , $x_i = y_j = 0$. Ainsi g est une base de $E \times F$.

Alors $\dim(E \times F) = |g| = p + q = \dim(E) + \dim(F)$. \square

2.5 Application linéaire associée à une famille de vecteurs

Propriété. Soit $x = (x_i) \in E^I$. Notons

$$\Psi_x : \begin{array}{ccc} \mathbb{K}^{(I)} & \longrightarrow & E \\ (\alpha_i)_{i \in I} & \longmapsto & \sum_{i \in I} \alpha_i x_i. \end{array}$$

Ψ_x est une application linéaire.

- x est une famille libre si et seulement si Ψ_x est injective.
- x est une famille génératrice si et seulement si Ψ_x est surjective.
- x est une base si et seulement si Ψ_x est un isomorphisme.

Ψ_x est appelée l'application linéaire associée à la famille de vecteurs x .

Démonstration.

- Soit $((\alpha_i), (\beta_i)) \in (\mathbb{K}^{(I)})^2$.

$$\Psi((\alpha_i) + (\beta_i)) = \Psi((\alpha_i + \beta_i)) = \sum_{i \in I} (\alpha_i + \beta_i) x_i = \Psi((\alpha_i)) + \Psi((\beta_i)).$$

$$\text{Soit } ((\alpha_i), \lambda) \in \mathbb{K}^{(I)} \times \mathbb{K}. \quad \Psi(\lambda(\alpha_i)) = \Psi((\lambda\alpha_i)) = \sum_{i \in I} (\lambda\alpha_i) x_i = \lambda \Psi((\alpha_i)).$$

Ainsi Ψ_x est une application linéaire.

- x est libre si et seulement si la seule famille (α_i) presque nulle de scalaires vérifiant $\Psi_x((\alpha_i)) = 0$ est la famille nulle, donc si et seulement si le noyau de Ψ_x est réduit à $\{0\}$.
- x est génératrice si et seulement si pour tout élément v de E , il existe $(\alpha_i) \in \mathbb{K}^{(I)}$ telle que $v = \Psi_x((\alpha_i))$, donc si et seulement si Ψ_x est surjective. \square

Remarque. Lorsque $e = (e_i)_{i \in I}$ est une base de E , pour tout $x \in E$,

$\Psi_e^{-1}(x)$ est la famille des coordonnées de x dans la base e .

Propriété. Soit $x = (x_i)_{i \in I}$ une famille de vecteurs de E .

x est libre si et seulement si, pour tout $y \in \text{Vect}(x)$, il existe une unique famille presque nulle de scalaires $(\alpha_i)_{i \in I}$ telle que $y = \sum_{i \in I} \alpha_i x_i$.

Démonstration.

x est libre si et seulement si Ψ_x est injective, donc si et seulement si, pour tout $y \in \text{Im}(\Psi_x)$, il existe un unique $(\alpha_i)_{i \in I} \in \mathbb{K}^{(I)}$ tel que $y = \Psi_x((\alpha_i))$.

Or $\text{Im}(\Psi_x) = \left\{ \sum_{i \in I} \alpha_i x_i / (\alpha_i)_{i \in I} \in \mathbb{K}^{(I)} \right\} = \text{Vect}(x)$. \square

Propriété. Si $e = (e_i)_{i \in I}$ est une base de E , alors E est isomorphe à $\mathbb{K}^{(I)}$.

Démonstration.

Ψ_e est en effet un isomorphisme de $\mathbb{K}^{(I)}$ dans E . \square

2.6 Image d'une famille par une application linéaire

Définition. Soient E et F deux \mathbb{K} -espaces vectoriels, $u \in L(E, F)$

et $x = (x_i)_{i \in I} \in E^I$. La famille $(u(x_i))_{i \in I}$ est appelée l'image de la famille x par l'application linéaire u .

On notera $(u(x_i))_{i \in I} = u(x)$. On remarquera que $x \in E^I$ et que $u(x) \in F^I$.

Propriété. Avec les notations précédentes, $\Psi_{u(x)} = u \circ \Psi_x$.

Démonstration.

Soit $(\alpha_i) \in \mathbb{K}^{(I)}$.

$$\Psi_{u(x)}((\alpha_i)) = \sum_{i \in I} \alpha_i u(x_i) = u \left(\sum_{i \in I} \alpha_i x_i \right) = (u \circ \Psi_x)((\alpha_i)). \quad \square$$

Théorème.

- L'image d'une famille libre par une injection linéaire est une famille libre.
- L'image d'une famille génératrice par une surjection linéaire est génératrice.
- L'image d'une base par un isomorphisme est une base.

Démonstration.

Soient E et F deux \mathbb{K} -espaces vectoriels, $u \in L(E, F)$ et $x = (x_i)_{i \in I} \in E^I$.

- Supposons que x est libre et que u est injective. Alors u et Ψ_x sont injectives, donc $\Psi_{u(x)}$ est injective en tant que composée de deux applications injectives. Ainsi $u(x)$ est une famille libre.
- Supposons que x est génératrice et que u est surjective. Alors u et Ψ_x sont surjectives, donc $\Psi_{u(x)}$ est surjective en tant que composée de deux applications surjectives. Ainsi $u(x)$ est une famille génératrice. \square

Théorème. Deux espaces de dimensions finies ont la même dimension si et seulement si ils sont isomorphes.

Démonstration.

Soit E et F deux \mathbb{K} -espaces vectoriels de dimensions finies respectives n et m .

Suppose qu'il existe un isomorphisme f de E dans F . Soit $e = (e_1, \dots, e_n)$ une base de E . Alors $(f(e_1), \dots, f(e_n))$ est une base de F , donc $\dim(F) = n = \dim(E)$.

Réciproquement, supposons que $n = m$. On sait que l'application Ψ_e canoniquement associée à e est un isomorphisme de \mathbb{K}^n dans E . De même on montre que F et \mathbb{K}^n sont isomorphes, donc E et F sont isomorphes. \square

Remarque. On dispose ainsi de deux techniques importantes pour calculer la dimension d'un espace vectoriel E : rechercher une base de E et calculer son cardinal ou bien chercher un isomorphisme entre E et un espace de dimension connue.

Exercice. Montrer que le cardinal d'un corps fini est de la forme p^n où $p \in \mathbb{P}$ et $n \in \mathbb{N}^*$

Solution : Soit \mathbb{L} un corps fini. Alors $\text{car}(\mathbb{L}) \neq 0$, donc en posant $p = \text{car}(\mathbb{L})$, $p \in \mathbb{P}$. Notons \mathbb{K} le sous-corps premier de \mathbb{L} . On sait que \mathbb{K} est isomorphe à \mathbb{F}_p , donc il est de cardinal p .

\mathbb{L} est un \mathbb{K} -espace vectoriel, et $(x)_{x \in \mathbb{L}}$ est une famille génératrice finie de \mathbb{L} , donc \mathbb{L} est de dimension finie. Posons $n = \dim_{\mathbb{K}}(\mathbb{L})$. Alors \mathbb{L} est un \mathbb{K} -espace vectoriel isomorphe à \mathbb{K}^n , donc $|\mathbb{L}| = p^n$.

Propriété. Soit E et F deux espaces de dimensions finies et soit $f \in L(E, F)$.

Si f est injective, alors $\dim(E) \leq \dim(F)$.

Si f est surjective, alors $\dim(E) \geq \dim(F)$.

Propriété. Soient E et F deux \mathbb{K} -espaces vectoriels de dimensions quelconques. Soient $u \in L(E, F)$ et G un sous-espace vectoriel de E de dimension finie. Alors $u(G)$ est de dimension finie et $\dim(u(G)) \leq \dim(G)$, avec égalité lorsque u est injective.

Démonstration.

$u|_G^{u(G)}$ est surjective. Elle est bijective lorsque u est injective. \square

Propriété. L'image d'une famille génératrice par une application linéaire u est une famille génératrice de $Im(u)$.

Démonstration.

Soient E et F deux \mathbb{K} -espaces vectoriels, $u \in L(E, F)$ et $x = (x_i)_{i \in I} \in E^I$ une famille génératrice de E .

L'application $\begin{array}{ccc} E & \longrightarrow & Im(u) \\ y & \longmapsto & u(y) \end{array}$ est surjective, donc $u(x)$, en tant qu'image de x par cette application, est une famille génératrice de $Im(u)$. \square

Propriété. L'image d'une famille liée par une application linéaire est liée.

Démonstration.

Soient E et F deux \mathbb{K} -espaces vectoriels, $u \in L(E, F)$ et $x = (x_i)_{i \in I} \in E^I$ une famille liée de E . Ainsi il existe $(\alpha_i)_{i \in I} \in \mathbb{K}^{(I)} \setminus \{0\}$ tel que $\sum_{i \in I} \alpha_i x_i = 0$.

$0 = u(\sum_{i \in I} \alpha_i x_i) = \sum_{i \in I} \alpha_i u(x_i)$, donc $u(x)$ est aussi une famille liée. \square

Théorème.

On suppose que E est un \mathbb{K} -espace vectoriel admettant une base $e = (e_i)_{i \in I}$.

Soit $f = (f_i)_{i \in I}$ une famille quelconque de vecteurs d'un second \mathbb{K} -espace vectoriel F .

Il existe une unique application linéaire $u \in L(E, F)$ telle que, $\forall i \in I$ $u(e_i) = f_i$.

De plus, $(f_i)_{i \in I}$ est $\begin{cases} \text{libre} \\ \text{génératrice} \\ \text{une base} \end{cases}$ si et seulement si u est $\begin{cases} \text{injective} \\ \text{surjective} \\ \text{bijective} \end{cases}$.

Démonstration.

Soit $u \in L(E, F)$.

$$\begin{aligned} (\forall i \in I \quad u(e_i) = f_i) &\iff \forall (\alpha_i) \in \mathbb{K}^{(I)} \quad \sum_{i \in I} \alpha_i u(e_i) = \sum_{i \in I} \alpha_i f_i \\ &\iff \Psi_{u(e)} = \Psi_f \iff u \circ \Psi_e = \Psi_f \\ &\iff u = \Psi_f \circ \Psi_e^{-1} \end{aligned}$$

car e étant une base Ψ_e est un isomorphisme.

Ainsi il existe une unique application linéaire $u \in L(E, F)$ telle que $\forall i \in I$ $u(e_i) = f_i$.

Il s'agit de $u = \Psi_f \circ \Psi_e^{-1}$.

f est libre si et seulement si Ψ_f est injective, donc si et seulement si $u = \Psi_f \circ \Psi_e^{-1}$ est injective.

f est génératrice si et seulement si Ψ_f est surjective, donc si et seulement si $u = \Psi_f \circ \Psi_e^{-1}$ est surjective. \square

Remarque. Ce théorème affirme notamment qu'une application linéaire $u \in L(E, F)$ est injective (resp : surjective, bijective) si et seulement si l'image par u d'une base de E est une famille libre (resp : une famille génératrice, une base) de F .

Corollaire.

Soit E et F deux espaces vectoriels de dimensions finies et soit $u \in L(E, F)$.

Si $\dim(E) = \dim(F)$, alors u injective $\iff u$ surjective $\iff u$ bijective.

Démonstration.

Notons $n = \dim(E) = \dim(F)$. Soit e une base de E . Alors u est injective si et seulement si $u(e)$ est libre dans F , or $|u(e)| = n = \dim(F)$, donc u est injective si et seulement si $u(e)$ est une base de F , c'est-à-dire si et seulement si u est bijective. On raisonne de même pour la surjectivité. \square

Exercice. Soit $u \in L(\mathbb{K}[X])$ tel que pour tout $P \in \mathbb{K}[X]$, $\deg(u(P)) = \deg(P)$. Montrer que u est un automorphisme sur $\mathbb{K}[X]$.

Solution : Soit $P \in \text{Ker}(u)$: $\deg(P) = \deg(u(P)) = -\infty$, donc $u(P) = 0$. Ainsi, $\text{Ker}(u) = \{0\}$ et u est injective.

Soit $n \in \mathbb{N}$. Pour tout $P \in \mathbb{K}_n[X]$, $\deg(u(P)) = \deg(P) \leq n$, donc $\mathbb{K}_n[X]$ est stable par u et l'endomorphisme u_n induit par u sur $\mathbb{K}_n[X]$ est bien défini. u_n est linéaire injective de $\mathbb{K}_n[X]$ dans $\mathbb{K}_n[X]$ qui est de dimension finie, donc c'est un automorphisme de $\mathbb{K}_n[X]$.

Soit $P \in \mathbb{K}[X]$. Il existe $n \in \mathbb{N}$ tel que $P \in \mathbb{K}_n[X]$, donc il existe $Q \in \mathbb{K}_n[X]$ tel que $u_n(Q) = P$. Alors $u(Q) = P$, ce qui prouve que u est surjective.

Par exemple, $P \mapsto P + P' + 2P''$ est un automorphisme de $\mathbb{K}[X]$.

Propriété. Soit E un \mathbb{K} -espace vectoriel de dimension finie et $u \in L(E)$. Alors

$$\begin{aligned} u \text{ inversible dans } L(E) &\iff u \text{ inversible à droite dans } L(E) \\ &\iff u \text{ inversible à gauche dans } L(E) \end{aligned}$$

Démonstration.

Si u est inversible à droite, il existe $v \in L(E)$ tel que $vu = Id_E$. Alors u est injectif, mais E est de dimension finie, donc u est un isomorphisme, c'est-à-dire que u est inversible dans $L(E)$.

On raisonne de même à gauche. \square

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

$$\begin{aligned} A \text{ inversible dans } \mathcal{M}_n(\mathbb{K}) &\iff A \text{ inversible à droite dans } \mathcal{M}_n(\mathbb{K}) \\ &\iff A \text{ inversible à gauche dans } \mathcal{M}_n(\mathbb{K}) \end{aligned}$$

Démonstration.

A est inversible à droite dans $\mathcal{M}_n(\mathbb{K})$ si et seulement si il existe $B \in \mathcal{M}_n(\mathbb{K})$ telle que $BA = I_n$, i.e telle que $\tilde{B}\tilde{A} = Id_{\mathbb{K}^n}$, donc si et seulement si \tilde{A} est inversible à droite dans $L(\mathbb{K}^n)$.

De même, A est inversible à gauche dans $\mathcal{M}_n(\mathbb{K})$ si et seulement si \tilde{A} est inversible à gauche dans $L(\mathbb{K}^n)$.

De plus, on sait que A est inversible dans $\mathcal{M}_n(\mathbb{K})$ si et seulement si \tilde{A} est inversible dans $L(\mathbb{K}^n)$. \square

Exercice. Soit A une \mathbb{K} -algèbre et B une sous-algèbre de A de dimension finie. Soit $b \in B$. Montrer que si b est inversible dans A , alors $b^{-1} \in B$.

Solution : Notons $u : \begin{matrix} B & \longrightarrow & B \\ x & \longmapsto & bx \end{matrix}$. u est linéaire. De plus, pour $x \in B$,
 $x \in \text{Ker}(u) \iff bx = 0 \iff b^{-1}bx = 0 \iff x = 0$, donc u est injective. Or B est de dimension finie, donc u est un isomorphisme. En particulier, 1_A possède un antécédent par u : il existe $c \in B$ tel que $bc = 1_A$. Alors $b^{-1} = c \in B$.

Propriété. Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice triangulaire supérieure, dont la diagonale est notée (a_1, \dots, a_n) . Alors A est inversible si et seulement si pour tout $i \in \{1, \dots, n\}$, $a_i \neq 0$, et dans ce cas, A^{-1} est encore triangulaire supérieure et sa diagonale est $\left(\frac{1}{a_1}, \dots, \frac{1}{a_n}\right)$.

Démonstration.

◇ Notons \mathcal{T} l'ensemble des matrices triangulaires supérieures. On a déjà vu que \mathcal{T} est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$ et nous sommes en dimension finie, donc d'après l'exercice précédent, si $A \in \mathcal{T}$ est inversible, alors $A^{-1} \in \mathcal{T}$. Les éléments diagonaux de $I_n = AA^{-1}$ sont égaux au produit des éléments diagonaux de A et de A^{-1} , donc les éléments diagonaux de A sont non nuls et les éléments diagonaux de A^{-1} sont les inverses des éléments diagonaux de A .

◇ Réciproquement, supposons que pour tout $i \in \mathbb{N}_n$, $A_{i,i} \neq 0$.

Soit $X \in \mathbb{K}^n$ tel que $AX = 0$. Pour tout $i \in \mathbb{N}_n$,

$$(E_i) : 0 = [AX]_i = \sum_{j=1}^n A_{i,j}X_j = \sum_{j=i}^n A_{i,j}X_j, \text{ car } A \text{ est triangulaire supérieure.}$$

Pour $i = n$, (E_n) se réduit à $A_{n,n}X_n = 0$, or $A_{n,n} \neq 0$, donc $X_n = 0$. Alors (E_{n-1}) se réduit à $A_{n-1,n-1}X_{n-1} = 0$, or $A_{n-1,n-1} \neq 0$, donc $X_{n-1} = 0$. Par récurrence descendante, on en déduit que $X = 0$.

Ainsi $\text{Ker}(\tilde{A}) = \{0\}$, donc \tilde{A} est une application linéaire injective de \mathbb{K}^n dans lui-même, or \mathbb{K}^n est de dimension finie, donc \tilde{A} est un isomorphisme, ce qui prouve que A est inversible. □

Propriété. Soit E un \mathbb{K} -espace vectoriel admettant une base $(e_i)_{i \in I}$.

Alors $L(E, F)$ est isomorphe à F^I .

Démonstration.

L'application $L(E, F) \longrightarrow F^I$
 $u \longmapsto (u(e_i))_{i \in I}$ est une application linéaire et le théorème précédent montre que toute famille $(f_i)_{i \in I}$ de F^I admet un unique antécédent pour cette application, donc qu'elle est bijective. Ainsi cette application est un isomorphisme. □

Théorème. Soit E et F deux espaces vectoriels de dimensions finies. Alors

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

Démonstration.

Posons $n = \dim(E)$ et notons $e = (e_i)_{1 \leq i \leq n}$ une base de E . D'après le précédent corollaire, $\dim(L(E, F)) = \dim(F^{\{1, \dots, n\}}) = \dim(F^n) = n \times \dim(F)$. □

2.7 Rang d'une famille de vecteurs

Définition. Soient E un espace vectoriel et x une famille de vecteurs de E .

Le rang de x est $\text{rg}(x) \triangleq \dim(\text{Vect}(x)) \in \mathbb{N} \cup \{+\infty\}$.

Propriété. Pour une famille x de vecteurs d'un \mathbb{K} -espace vectoriel E ,

- $\text{rg}(x) \leq \#(x)$. Lorsque $\text{rg}(x)$ est fini, il y a égalité si et seulement si x est libre.
- $\text{rg}(x) \leq \dim(E)$. Lorsque $\text{rg}(x)$ est fini, il y a égalité si et seulement si x est génératrice.

Démonstration.

◇ x est une famille génératrice de $\text{Vect}(x)$, donc son cardinal est plus grand que la dimension de $\text{Vect}(x)$, égale au rang de x .

De plus, il y a égalité si et seulement si x est une base de $\text{Vect}(x)$, donc si et seulement si x est une famille libre.

◇ $\text{Vect}(x)$ est un sous-espace vectoriel de E , donc sa dimension est inférieure à la dimension de E . De plus, il y a égalité si et seulement si $\text{Vect}(x) = E$, c'est-à-dire si et seulement si x est une famille génératrice de E . □

Propriété.

Soient E et F deux espaces vectoriels, x une famille de vecteurs de E et $u \in L(E, F)$. Alors $\text{rg}(u(x)) \leq \text{rg}(x)$. Lorsque $\text{rg}(x)$ est fini, il y a égalité lorsque u est injective.

Démonstration.

Notons $G = \text{Vect}(x)$ et $x = (x_i)_{i \in I}$. $\text{rg}(u(x)) = \dim(\text{Vect}(u(x)))$. Or

$\text{Vect}(u(x)) = \text{Vect}(u(x_i))_{i \in I} = u\left(\text{Vect}(x_i)_{i \in I}\right) = u(G)$, donc $\text{rg}(u(x)) = \dim(u(G))$.

La propriété résulte alors du fait que $\dim(u(G)) \leq \dim(G)$, avec égalité lorsque u est injective. □

Propriété. Soit $(x_i)_{i \in I}$ une famille de vecteurs d'un \mathbb{K} -espace vectoriel E . Alors $\text{rg}((x_i)_{i \in I})$ n'est pas modifié si l'on échange l'ordre de deux vecteurs, si l'on multiplie l'un des vecteurs x_i par un scalaire non nul, ou bien si l'on ajoute à l'un des x_i une combinaison linéaire des autres x_j .

2.8 Matrice d'une application linéaire

Remarque. On a vu que pour construire une application linéaire u de E dans F , si $(e_i)_{i \in I}$ est une base de E , il suffit de donner la famille $(u(e_i))_{i \in I}$ des images des e_i par u .

Par exemple, on peut définir un endomorphisme u sur $\mathbb{R}_3[X]$ par les conditions : $u(X^3) = X$, $u(X^2) = X^2 + 1$, $u(X) = X^3 - X$, et $u(1) = 1$.

u est nécessairement l'unique endomorphisme tel que : pour tout

$$P = a_3X^3 + a_2X^2 + a_1X + a_0, \quad u(P) = a_3X + a_2(X^2 + 1) + a_1(X^3 - X) + a_0.$$

Définition. Soient E et F deux \mathbb{K} -espaces vectoriels de dimensions respectives $p > 0$ et $n > 0$. Soient $e = (e_1, \dots, e_p)$ une base de E et $f = (f_1, \dots, f_n)$ une base de F .

Si $u \in L(E, F)$, on appelle **matrice de l'application linéaire** u dans les bases e et f la matrice notée $\text{mat}(u, e, f) = (\alpha_{i,j}) \in \mathcal{M}_{\mathbb{K}}(n, p)$ définie par : pour tout $i \in \{1, \dots, n\}$ et $j \in \{1, \dots, p\}$, $\alpha_{i,j}$ est la $i^{\text{ème}}$ coordonnée du vecteur $u(e_j)$ dans la base f .

C'est donc l'unique matrice $(\alpha_{i,j}) \in \mathcal{M}_{\mathbb{K}}(n, p)$ vérifiant : $\forall j \in \mathbb{N}_p \quad u(e_j) = \sum_{i=1}^n \alpha_{i,j} f_i$.

C'est l'unique matrice dont la j -ème colonne contient les coordonnées de $u(e_j)$ dans la base f , pour tout j : la j -ème colonne est égale à $\Psi_f^{-1}(u(e_j))$.

On peut également dire que la matrice de u dans les bases e et f est définie par $[\text{mat}(u, e, f)]_{i,j} = f_i^*(e_j)$, pour tout $i \in \mathbb{N}_n$ et $j \in \mathbb{N}_p$.

Interprétation tabulaire : Avec les notations précédentes,

$$\text{mat}(u, e, f) = \begin{pmatrix} u(e_1) & \cdots & u(e_p) \\ m_{1,1} & \cdots & m_{1,p} \\ \vdots & & \vdots \\ m_{n,1} & \cdots & m_{n,p} \end{pmatrix} \begin{matrix} f_1 \\ \vdots \\ f_n \end{matrix}.$$

Notation. Lorsque $E = F$ et que l'on choisit $e = f$, on note $\text{mat}(u, e)$ au lieu de $\text{mat}(u, e, e)$.

Exemple. L'endomorphisme u de $\mathbb{R}_3[X]$ défini ci-dessus a pour matrice dans la base

$$\text{canonique : } \text{mat}(u, (1, X, X^2, X^3)) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Exemple. La forme linéaire $u : \mathbb{K}^4 \longrightarrow \mathbb{K}$ définie par $u \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = x + y - 2z + 3t$ a

pour matrice dans les bases canoniques de \mathbb{K}^4 et \mathbb{K} la matrice ligne $(1 \ 1 \ -2 \ 3)$.
Plus généralement la matrice d'une forme linéaire est toujours une matrice ligne.

Exemple. Considérons $u : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ $\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 2x + y \\ y \\ 3x - y \end{pmatrix}$.

u est une application linéaire et, si l'on note $B = (b_1, b_2)$ et $C = (c_1, c_2, c_3)$ les bases canoniques de \mathbb{R}^2 et de \mathbb{R}^3 respectivement,

$$\text{mat}(u, B, C) = \begin{pmatrix} 2 & 1 \\ 0 & 1 \\ 3 & -1 \end{pmatrix}. \text{ En effet, } u(b_1) = u \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}$$

$$\text{et } u(b_2) = u \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

En fait, si l'on pose $M = \begin{pmatrix} 2 & 1 \\ 0 & 1 \\ 3 & -1 \end{pmatrix}$, on voit que pour tout $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$,

$u \begin{pmatrix} x \\ y \end{pmatrix} = M \begin{pmatrix} x \\ y \end{pmatrix}$, donc u n'est autre que l'application linéaire canoniquement associée à M .

Remarque. Plus généralement, si $M \in \mathcal{M}_{\mathbb{K}}(n, p)$, on a défini $\tilde{M} \in L(\mathbb{K}^p, \mathbb{K}^n)$ par : $\forall X \in \mathbb{K}^p$, $\tilde{M}(X) = MX$.

On a vu que, si l'on note $c = (c_1, \dots, c_p)$ la base canonique de \mathbb{K}^p , alors pour tout $j \in \mathbb{N}_p$, $\tilde{M}(c_j) = Mc_j$ est la j -ème colonne de M , donc :

Propriété. Pour tout $n, p \in \mathbb{N}^*$, pour tout $M \in \mathcal{M}_{\mathbb{K}}(n, p)$, $\boxed{\text{mat}(\tilde{M}, c, c') = M}$, en notant c et c' les bases canoniques de \mathbb{K}^p et de \mathbb{K}^n .

Remarque. Nous disposons maintenant de deux manières équivalentes de définir l'application linéaire canoniquement associée à une matrice $M \in \mathcal{M}_{\mathbb{K}}(n, p)$: c'est l'application $\begin{matrix} \tilde{M} : \mathbb{K}^p & \longrightarrow & \mathbb{K}^n \\ X & \longmapsto & \tilde{M}(X) = MX \end{matrix}$, ou bien c'est l'unique application $\tilde{M} \in L(\mathbb{K}^p, \mathbb{K}^n)$ telle que $\text{mat}(\tilde{M}, c, c') = M$.

Exemple. Reprenons l'exemple précédent, et déterminons la matrice de u pour un autre couple de bases : Posons $e_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $e_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.
 $\det(e_1, e_2) = -2 \neq 0$, donc $E = (e_1, e_2)$ est une base de \mathbb{R}^2 .

Posons $f_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $f_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ et $f_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

On vérifie que $F = (f_1, f_2, f_3)$ est une base de \mathbb{R}^3 .

En effet, si $a_1 f_1 + a_2 f_2 + a_3 f_3 = 0$, où $a_1, a_2, a_3 \in \mathbb{R}$, alors $0 = a_1 + a_2 = a_2 = a_1 + a_3$, donc $0 = a_2 = a_1 = a_3$. Ainsi F est libre, de cardinal 3, or $\dim(\mathbb{R}^3) = 3$, donc c'est bien une base de \mathbb{R}^3 .

Soit $(\alpha, \beta) \in \mathbb{R}^2$. $u(\alpha e_1 + \beta e_2) = \alpha u \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \beta u \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \alpha \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix}$,

donc $u(\alpha e_1 + \beta e_2) = x f_1 + y f_2 + z f_3$, où $\begin{cases} 3\alpha + \beta = x + y \\ \alpha - \beta = y \\ 2\alpha + 4\beta = x + z \end{cases}$, donc $y = \alpha - \beta$,

$x = 2\alpha + 2\beta$ et $z = 2\beta$. On en déduit que $\text{mat}(u, E, F) = \begin{pmatrix} 2 & 2 \\ 1 & -1 \\ 0 & 2 \end{pmatrix}$.

Nous verrons plus loin une formule de changement de bases qui permet d'aller plus vite.

Propriété. Soient E et F deux \mathbb{K} -espaces vectoriels de dimensions respectives $p > 0$ et $n > 0$. Soient $e = (e_1, \dots, e_p)$ une base de E et $f = (f_1, \dots, f_n)$ une base de F .

L'application
$$\begin{array}{ccc} L(E, F) & \longrightarrow & \mathcal{M}_{\mathbb{K}}(n, p) \\ u & \longmapsto & \text{mat}(u, e, f) \end{array}$$
 est un isomorphisme d'espaces vectoriels.

Démonstration.

◇ Soit $u, v \in L(E, F)$ et $\lambda \in \mathbb{K}$.

Posons $U = \text{mat}(u, e, f)$, $V = \text{mat}(v, e, f)$ et $M = \text{mat}(\lambda u + v, e, f)$.

Soit $i \in \mathbb{N}_n$ et $j \in \mathbb{N}_p$. $M_{i,j}$ est la i -ème coordonnée de $(\lambda u + v)(e_j)$ dans la base f . C'est donc $f_i^*(\lambda u(e_j) + v(e_j))$, or f_i^* est linéaire, donc $M_{i,j} = \lambda f_i^*(u(e_j)) + f_i^*(v(e_j)) = \lambda U_{i,j} + V_{i,j}$, donc $M = \lambda U + V$,

c'est-à-dire $\text{mat}(\lambda u + v, e, f) = \lambda \text{mat}(u, e, f) + \text{mat}(v, e, f)$. Ceci prouve la linéarité.

◇ Soit $M = (\alpha_{i,j}) \in \mathcal{M}_{\mathbb{K}}(n, p)$ et $u \in L(E, F)$. On a vu que $M = \text{mat}(u, e, f)$ si et seulement si pour tout $j \in \mathbb{N}_p$, $u(e_j) = g_j$, où $g_j = \sum_{i=1}^n \alpha_{i,j} f_i$, or on a déjà énoncé

qu'il existe une unique $u \in L(E, F)$ telle que, pour tout $j \in \mathbb{N}_p$, $u(e_j) = g_j$ (i.e : une application linéaire est uniquement déterminée par la donnée des images des vecteurs d'une base de l'espace de départ). Ainsi, M possède un unique antécédent : l'application est bijective. □

Remarque. En particulier, lorsque $E = \mathbb{K}^p$ et $F = \mathbb{K}^n$, où $n, p \in \mathbb{N}^*$, en notant c_n et c_p les bases canoniques de \mathbb{K}^n et de \mathbb{K}^p , on vient de montrer que

$$\begin{array}{ccc} \Psi : L(\mathbb{K}^p, \mathbb{K}^n) & \longrightarrow & \mathcal{M}_{\mathbb{K}}(n, p) \\ u & \longmapsto & \text{mat}(u, c_p, c_n) \end{array}$$
 est un isomorphisme. On le savait déjà car c'est l'isomorphisme réciproque de
$$\begin{array}{ccc} \mathcal{M}_{\mathbb{K}}(n, p) & \longrightarrow & L(\mathbb{K}^p, \mathbb{K}^n) \\ M & \longmapsto & \tilde{M} \end{array}.$$

Théorème. Soient E, F et G trois \mathbb{K} -espaces vectoriels de dimensions respectives q, p et n , munis de bases e, f et g . Soient $u \in L(E, F)$ et $v \in L(F, G)$.

Alors, $\text{mat}(v \circ u, e, g) = \text{mat}(v, f, g) \times \text{mat}(u, e, f)$.

Démonstration.

Posons $U = \text{mat}(u, e, f) \in \mathcal{M}_{\mathbb{K}}(p, q)$, $V = \text{mat}(v, f, g) \in \mathcal{M}_{\mathbb{K}}(n, p)$ et $M = \text{mat}(v \circ u, e, g) \in \mathcal{M}_{\mathbb{K}}(n, q)$. Soit $k \in \mathbb{N}_q$ et $i \in \mathbb{N}_n$.

$$M_{i,k} = g_i^*(vu(e_k)) = g_i^*\left[v\left(\sum_{j=1}^p U_{j,k} f_j\right)\right] = \sum_{j=1}^p U_{j,k} g_i^*[v(f_j)],$$

donc $M_{i,k} = \sum_{j=1}^p U_{j,k} V_{i,j} = [VU]_{i,k}$, ce qui prouve que $M = VU$. □

Exemple. On peut ainsi remplacer un calcul matriciel par un calcul sur des applications linéaires. Par exemple, on peut retrouver que, dans $\mathcal{M}_{\mathbb{K}}(n)$, $E_{i,j} E_{h,k} = \delta_{j,h} E_{i,k}$: Notons $e = (e_1, \dots, e_n)$ la base canonique de \mathbb{K}^n .

Pour $(i, j) \in \mathbb{N}_n^2$, notons $u_{i,j}$ l'endomorphisme de \mathbb{K}^n canoniquement associé à $E_{i,j}$.

Pour tout $k \in \mathbb{N}_n$, $u_{i,j}(e_k) = \delta_{k,j} e_i$.

Soit $l \in \mathbb{N}_n$. $u_{i,j} \circ u_{h,k}(e_l) = u_{i,j}(\delta_{k,l} e_h) = \delta_{k,l} \delta_{j,h} e_i$, donc $u_{i,j} \circ u_{h,k}(e_l) = \delta_{j,h} u_{i,k}(e_l)$.

Ainsi, $u_{i,j} \circ u_{h,k} = \delta_{j,h} u_{i,k}$, puis en prenant les matrices de ces endomorphismes,

$$E_{i,j} E_{h,k} = \delta_{j,h} E_{i,k}.$$

Propriété. Soient E et F deux \mathbb{K} -espaces vectoriels de dimensions respectives $p > 0$ et $n > 0$, munis des bases $e = (e_1, \dots, e_p)$ et $f = (f_1, \dots, f_n)$, et soit $u \in L(E, F)$.

On note M la matrice de u dans les bases e et f .

Soit $(x, y) \in E \times F$. On note X la matrice colonne des coordonnées de x dans la base e , et Y celle des coordonnées de y dans la base f . C'est-à-dire qu'en posant $x = \sum_{j=1}^p x_j e_j$,

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, \text{ et qu'en posant } y = \sum_{i=1}^n y_i f_i, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

C'est aussi dire que $X = \Psi_e^{-1}(x)$ et $Y = \Psi_f^{-1}(y)$.

On écrira également que $X = \text{mat}(x, e)$ et $Y = \text{mat}(y, f)$. Alors,

$$\boxed{u(x) = y \iff MX = Y.}$$

Démonstration.

On pourrait bien sûr passer aux coordonnées et mener un calcul analogue à celui de la démonstration de la propriété précédente. Essayons plutôt d'utiliser cette dernière propriété.

Si $x \in E$, notons $\hat{x} : \mathbb{K} \longrightarrow E$
 $\lambda \longmapsto \lambda x$: $\hat{x} \in L(\mathbb{K}, E)$.

L'application $\begin{matrix} E & \longrightarrow & L(\mathbb{K}, E) \\ x & \longmapsto & \hat{x} \end{matrix}$ est un isomorphisme, dont la bijection réciproque est $\begin{matrix} L(\mathbb{K}, E) & \longrightarrow & E \\ a & \longmapsto & a(1) \end{matrix}$. De même, à tout $y \in F$, on associe $\hat{y} = (\lambda \longmapsto \lambda y) \in L(\mathbb{K}, F)$.

On vérifie que $\widehat{u(\hat{x})} = u \circ \hat{x}$, car pour tout $\lambda \in \mathbb{K}$, $\widehat{u(\hat{x})}(\lambda) = \lambda u(x) = u(\lambda x) = u(\hat{x}(\lambda))$.

Alors $y \longmapsto \hat{y}$ étant injective,

$$\begin{aligned} u(x) = y &\iff \widehat{u(x)} = \hat{y} \\ &\iff \text{mat}(u \circ \hat{x}, 1, f) = \text{mat}(\hat{y}, 1, f) \\ &\iff M \times \text{mat}(\hat{x}, 1, e) = \text{mat}(\hat{y}, 1, f). \end{aligned}$$

Or $\text{mat}(\hat{x}, 1, e)$ est une matrice colonne dont les composantes sont les coordonnées de $\hat{x}(1) = x$ dans la base e , donc $\text{mat}(\hat{x}, 1, e) = X$, et de même, $\text{mat}(\hat{y}, 1, e) = Y$. \square

Propriété.

On reprend les notations de la propriété précédente et on suppose de plus que $n = p$.

Alors u est un isomorphisme si et seulement si M est une matrice inversible et dans ce cas, $\text{mat}(u, e, f)^{-1} = \text{mat}(u^{-1}, f, e)$.

Démonstration.

u est bijective si et seulement si pour tout $y \in F$, il existe un unique $x \in E$ tel que $u(x) = y$, donc si et seulement si pour tout $Y \in \mathbb{K}^n$, il existe un unique $X \in \mathbb{K}^n$ tel que $MX = Y$, c'est-à-dire si et seulement si M est inversible.

Dans ce cas, posons $N = \text{mat}(u^{-1}, f, e)$.

Alors $MN = \text{mat}(uu^{-1}, f, f) = \text{mat}(Id_F, f, f) = I_n$, donc $N = M^{-1}$. \square

Exercice. Pour tout $i, j \in \{0, \dots, n\}$, on pose $m_{i,j} = \binom{j}{i}$, en convenant que $\binom{j}{i} = 0$ lorsque $i > j$. Montrer que $M = (m_{i,j})_{\substack{0 \leq i \leq n \\ 0 \leq j \leq n}}$ est inversible et calculer son inverse.

Solution : Notons $u : \mathbb{K}_n[X] \longrightarrow \mathbb{K}_n[X]$ définie par $u(P) = P(X+1)$.
Notons c la base canonique de $\mathbb{K}_n[X]$. Pour tout $j \in \{0, \dots, n\}$,

$$(X+1)^j = \sum_{i=0}^j \binom{j}{i} X^i, \text{ donc } M = \text{mat}(u, c).$$

Or u est inversible et $u^{-1} : P \longmapsto P(X-1)$, donc M est une matrice inversible et $M^{-1} = \text{mat}(u^{-1}, c)$. Pour tout $j \in \{0, \dots, n\}$,

$$(X-1)^j = \sum_{i=0}^j \binom{j}{i} X^i (-1)^{j-i}, \text{ donc } M^{-1} = \left(\binom{j}{i} (-1)^{j-i} \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}.$$

Propriété. Soit E un \mathbb{K} -espace vectoriel de dimension finie égale à n , muni d'une base e . L'application
$$\begin{array}{ccc} L(E) & \longrightarrow & \mathcal{M}_n(\mathbb{K}) \\ u & \longmapsto & \text{mat}(u, e) \end{array}$$
 est un isomorphisme d'algèbres.

Remarque. Si e' est une seconde base de E , L'application
$$\begin{array}{ccc} L(E) & \longrightarrow & \mathcal{M}_n(\mathbb{K}) \\ u & \longmapsto & \text{mat}(u, e, e') \end{array}$$
 n'est pas un morphisme d'algèbres.

C'est pourquoi, le plus souvent, lorsque l'on considère la matrice d'un endomorphisme, on choisit la base d'arrivée égale à la base de départ.

Propriété. Si E est un \mathbb{K} -espace vectoriel de dimension finie égale à n , muni d'une base e , l'application
$$\begin{array}{ccc} GL(E) & \longrightarrow & GL_n(\mathbb{K}) \\ u & \longmapsto & \text{mat}(u, e) \end{array}$$
 est un isomorphisme de groupes.

3 Les systèmes linéaires

3.1 Trois interprétations d'un système linéaire

Définition. Une équation linéaire à p inconnues scalaires est une équation de la forme $(E) : \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_p x_p = b$, où $\alpha_1, \dots, \alpha_p, b \in \mathbb{K}$ sont des paramètres, et où $x_1, \dots, x_p \in \mathbb{K}$ sont les inconnues.

Exemples.

- $(E) : 2x - 5y + z = 5$ est une équation linéaire.
- $(E) : 2x = 1 + 2z$ est aussi linéaire.
- $(E) : x^2 + y + z - t = 2$ n'est pas linéaire.

Notation. Fixons $(n, p) \in \mathbb{N}^{*2}$ et considérons un système linéaire à n équations et p inconnues, c'est-à-dire un système d'équations de la forme suivante :

$$(S) : \begin{cases} \alpha_{1,1}x_1 + \cdots + \alpha_{1,p}x_p = b_1 \\ \vdots \\ \alpha_{i,1}x_1 + \cdots + \alpha_{i,p}x_p = b_i \\ \vdots \\ \alpha_{n,1}x_1 + \cdots + \alpha_{n,p}x_p = b_n \end{cases},$$

où, pour tout $(i, j) \in \{1, \dots, n\} \times \{1, \dots, p\}$, $\alpha_{i,j} \in \mathbb{K}$, pour tout $i \in \{1, \dots, n\}$, $b_i \in \mathbb{K}$, les p inconnues étant x_1, \dots, x_p , éléments de \mathbb{K} .

Le vecteur $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ est appelé le second membre du système, ou bien le membre constant.

Lorsqu'il est nul, on dit que le système est homogène.

Première interprétation. *Combinaison linéaire de vecteurs.*

Notons $C_1 = \begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{i,1} \\ \vdots \\ \alpha_{n,1} \end{pmatrix}$, $C_2 = \begin{pmatrix} \alpha_{1,2} \\ \vdots \\ \alpha_{i,2} \\ \vdots \\ \alpha_{n,2} \end{pmatrix}$, ..., $C_p = \begin{pmatrix} \alpha_{1,p} \\ \vdots \\ \alpha_{i,p} \\ \vdots \\ \alpha_{n,p} \end{pmatrix}$, et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{pmatrix}$. Il s'agit de $p+1$ vecteurs de \mathbb{K}^n . Alors

$$(S) \iff x_1 C_1 + x_2 C_2 + \cdots + x_p C_p = B.$$

Définition. On dit que (S) est **compatible** si et seulement s'il admet au moins une solution.

Propriété. (S) est compatible si et seulement si $B \in \text{Vect}(C_1, \dots, C_p)$.

Démonstration.

S est compatible si et seulement s'il existe $(x_1, \dots, x_p) \in \mathbb{K}^p$ tel que $B = x_1 C_1 + x_2 C_2 + \dots + x_p C_p$, c'est-à-dire si et seulement si $B \in \text{Vect}(C_1, \dots, C_p)$. \square

Deuxième interprétation. Matricielle.

Notons M la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ dont les colonnes sont C_1, \dots, C_p , et $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$.

Alors

$$(S) \iff MX = B.$$

Définition. On dit que (S) est un **système de Cramer** si et seulement si $n = p$ et si M est inversible. Dans ce cas, (S) admet une unique solution.

Démonstration.

Si M est inversible, $(S) \iff X = M^{-1}B$. \square

Troisième interprétation. A l'aide d'une application linéaire.

Soient E et F des \mathbb{K} -espaces vectoriels de dimensions p et n munis de bases $e = (e_1, \dots, e_p)$ et $f = (f_1, \dots, f_n)$. On note u l'unique application linéaire de $L(E, F)$ telle que $\text{mat}(u, e, f) = M$, x le vecteur de E dont les coordonnées dans e sont X et b le vecteur de F dont les coordonnées dans f sont B . Alors

$$(S) \iff u(x) = b.$$

Définition. On dit que (S) est un **système homogène** si et seulement si $b = 0$.

Définition. Le système homogène associé à (S) est $(S_H) : u(x) = 0$.

Propriété. L'ensemble des solutions de (S_H) est $\text{Ker}(u)$.

C'est un sous-espace vectoriel de dimension $p - r$, où r désigne le rang de u (ou de M).

3.2 Les opérations élémentaires

Définition. On appelle manipulations ou opérations élémentaires sur les lignes d'une matrice, les applications de $\mathcal{M}_{\mathbb{K}}(n, p)$ dans $\mathcal{M}_{\mathbb{K}}(n, p)$ suivantes :

- 1) Ajouter à une ligne le multiple d'une autre, opération notée :

$$L_i \longleftarrow L_i + \lambda L_j, \text{ où } i \neq j \text{ et } \lambda \in \mathbb{K}. \text{ C'est une transvection.}$$

- 2) Multiplier une ligne par un scalaire non nul, notée :

$$L_i \longleftarrow \alpha L_i, \text{ où } \alpha \in \mathbb{K}^*. \text{ C'est une affinité.}$$

- 3) Permuter deux lignes, notée :

$$L_i \longleftrightarrow L_j, \text{ où } i \neq j. \text{ C'est une transposition.}$$

Remarque. On définirait de même les opérations sur les colonnes.

Définition. Si $\sigma \in \mathcal{S}_n$, on note $P_\sigma = (\delta_{i,\sigma(j)}) \in \mathcal{M}_n(\mathbb{K})$.

Ainsi, pour tout $j \in \{1, \dots, n\}$, la $j^{\text{ème}}$ colonne de P_σ est constituée de 0, sauf pour le $\sigma(j)^{\text{ème}}$ coefficient qui vaut 1.

Propriété. Pour tout $(\sigma, \sigma') \in \mathcal{S}_n^2$, $P_{\sigma\sigma'} = P_\sigma P_{\sigma'}$.

Démonstration.

Notons $e = (e_1, \dots, e_n)$ la base canonique de \mathbb{K}^n . Si $s \in \mathcal{S}_n$, notons u_s l'endomorphisme canoniquement associé à la matrice P_s : Pour tout $j \in \{1, \dots, n\}$, $u_s(e_j) = e_{s(j)}$.

Soit $j \in \{1, \dots, n\}$: $(u_\sigma \circ u_{\sigma'})(e_j) = u_\sigma(e_{\sigma'(j)}) = e_{\sigma(\sigma'(j))} = u_{\sigma\sigma'}(e_j)$,

donc $u_\sigma \circ u_{\sigma'} = u_{\sigma\sigma'}$. En prenant les matrices de ces endomorphismes dans la base e , on en déduit que $P_{\sigma\sigma'} = P_\sigma P_{\sigma'}$. \square

Propriété.

En notant $(E_{i,j})_{(i,j) \in \{1, \dots, n\}^2}$ la base canonique de $\mathcal{M}_n(\mathbb{K})$, si $\lambda \in \mathbb{K}^*$ et $(i, j) \in \{1, \dots, n\}^2$ avec $i \neq j$, alors

$$\begin{array}{ccc} L_i \longleftarrow L_i + \lambda L_j : & \mathcal{M}_{\mathbb{K}}(n, p) & \longrightarrow \\ & M & \longmapsto \end{array} \quad \begin{array}{c} \mathcal{M}_{\mathbb{K}}(n, p) \\ (I_n + \lambda E_{i,j})M \end{array}$$

$$\begin{array}{ccc} L_i \longleftarrow \lambda L_i : & \mathcal{M}_{\mathbb{K}}(n, p) & \longrightarrow \\ & M & \longmapsto \end{array} \quad \begin{array}{c} \mathcal{M}_{\mathbb{K}}(n, p) \\ (I_n + (\lambda - 1)E_{i,i})M \end{array}$$

$$\begin{array}{ccc} L_i \longleftrightarrow L_j : & \mathcal{M}_{\mathbb{K}}(n, p) & \longrightarrow \\ & M & \longmapsto \end{array} \quad \begin{array}{c} \mathcal{M}_{\mathbb{K}}(n, p) \\ P_{(i,j)}M \end{array}$$

De même, en notant $(E_{i,j})_{(i,j) \in \{1, \dots, p\}^2}$ la base canonique de $\mathcal{M}_p(\mathbb{K})$, si $\lambda \in \mathbb{K}^*$ et $(i, j) \in \{1, \dots, p\}^2$ avec $i \neq j$, alors

$$\begin{array}{ccc} C_i \longleftarrow C_i + \lambda C_j : & \mathcal{M}_{\mathbb{K}}(n, p) & \longrightarrow \\ & M & \longmapsto \end{array} \quad \begin{array}{c} \mathcal{M}_{\mathbb{K}}(n, p) \\ M(I_p + \lambda E_{j,i}) \end{array}$$

$$\begin{array}{ccc} C_i \longleftarrow \lambda C_i : & \mathcal{M}_{\mathbb{K}}(n, p) & \longrightarrow \\ & M & \longmapsto \end{array} \quad \begin{array}{c} \mathcal{M}_{\mathbb{K}}(n, p) \\ M(I_p + (\lambda - 1)E_{i,i}) \end{array} \quad .$$

$$\begin{array}{ccc} C_i \longleftrightarrow C_j : & \mathcal{M}_{\mathbb{K}}(n, p) & \longrightarrow \\ & M & \longmapsto \end{array} \quad \begin{array}{c} \mathcal{M}_{\mathbb{K}}(n, p) \\ MP_{(i,j)} \end{array}$$

Propriété. Si l'on effectue une série d'opérations élémentaires sur les lignes d'une matrice M , alors on a multiplié M à gauche par une certaine matrice inversible.

Si l'on effectue une série d'opérations élémentaires sur les colonnes d'une matrice M , alors on a multiplié M à droite par une certaine matrice inversible.

Notation. Soit $(S) : MX = B$ un système linéaire de matrice $M \in \mathcal{M}_{n,p}(\mathbb{K})$ et de vecteur constant $B \in \mathbb{K}^n$.

On appellera matrice globale de (S) la matrice à n lignes et $p + 1$ colonnes dont les p premières colonnes sont celles de M et dont la dernière colonne est égale à B .

Propriété. Soient $(S) : MX = B$ et $(S') : M'X = B'$. On suppose que l'on peut passer de la matrice **globale** de (S) à celle de (S') à l'aide d'une série d'opérations élémentaires portant uniquement sur les lignes.

Alors ces deux systèmes sont équivalents.

Démonstration.

Une opération du type 1) revient à ajouter à l'équation i du système λ fois l'équation j . Le système après cette manipulation est équivalent au système avant cette manipulation.

Une opération du type 2) revient à multiplier une équation par un scalaire non nul, et une opération du type 3) revient à permuter l'ordre de deux équations du système.

À chaque étape on ne change pas l'espace des solutions du système. \square

En pratique : Pour résoudre un système linéaire, on tente de modifier la matrice globale du système par des manipulations élémentaires, afin de se ramener à une matrice qui, privée de sa dernière colonne, est diagonale ou bien triangulaire supérieure. Dans ce cas en effet, le système est simple à résoudre.

Remarque. Dans le système $(S) : MX = B$, permuter les colonnes de M revient à modifier l'ordre des inconnues. On peut donc autoriser ce type d'opération pour la résolution d'un système linéaire, mais il faudra les mémoriser pour connaître la position de chaque inconnue.

Propriété. Soit $M \in \mathcal{M}_n(\mathbb{K})$. On suppose que l'on peut transformer, par des opérations élémentaires portant uniquement sur les lignes, la matrice blocs $\begin{bmatrix} M & I_n \end{bmatrix} \in \mathcal{M}_{\mathbb{K}}(n, 2n)$ en une matrice de la forme $\begin{bmatrix} I_n & N \end{bmatrix} \in \mathcal{M}_{\mathbb{K}}(n, 2n)$.

Alors M est inversible et $M^{-1} = N$.

Démonstration.

Il existe une matrice $P \in GL_n(\mathbb{K})$ telle que $\begin{bmatrix} I_n & N \end{bmatrix} = P \times \begin{bmatrix} M & I_n \end{bmatrix}$,

or $P \times \begin{bmatrix} M & I_n \end{bmatrix} = \begin{bmatrix} PM & P \times I_n \end{bmatrix}$, donc $I_n = PM$ et $N = P$, ce qui montre bien que M est inversible et que son inverse est N . \square

Exemple. Inversion de la matrice de taille 4

$$M = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} :$$

Partons de la matrice $\left(\begin{array}{cccc|cccc} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{array} \right)$. Ajoutons à la première ligne

la somme des suivantes. On obtient comme nouvelle matrice :

$$\left(\begin{array}{cccc|cccc} 3 & 3 & 3 & 3 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{array} \right).$$
 On divise ensuite la première ligne par 3, ce qui donne

$$\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{array} \right),$$
 puis on enlève la première ligne aux suivantes. On obtient

$$\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & -1 & 0 & 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & -1 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & -1 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{array} \right).$$
 On multiplie ensuite les lignes autres que la première par -1 , ce qui fournit :

$$\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & 0 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{array} \right).$$
 Enfin, on enlève à la première ligne la somme des suivantes. On obtient :

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & 0 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{array} \right).$$
 Ainsi M est inversible et son inverse vaut

$$M^{-1} = \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}.$$

Remarque. Pour cette matrice, généralisée à une matrice de taille n , une autre méthode plus rapide consiste à remarquer que $(M + I_n)^2 = n(M + I_n)$, donc $M^2 + (2 - n)M + (1 - n)I_n = 0$, puis $M(M + (2 - n)I_n) = (n - 1)I_n$, ce qui montre que M est inversible et que $M^{-1} = \frac{1}{n - 1}(M + (2 - n)I_n)$.

Cette méthode se généralise à toute matrice pour laquelle on peut trouver simplement un polynôme annulateur.

Pour cette matrice, c'est facile car elle est combinaison linéaire de I_n et de la matrice U dont tous les coefficients sont égaux à 1. C'est plus généralement le cas de toute

matrice de la forme
$$\begin{pmatrix} a & b & \cdots & b \\ b & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & a \end{pmatrix}.$$

3.3 Méthode du pivot de Gauss

Notation. On souhaite résoudre le système $(S) : MX = B$ de n équations à p inconnues. La matrice globale du système sera notée $(a_{i,j}) \in \mathcal{M}_{\mathbb{K}}(n, p + 1)$.

Pour simplifier les notations, si on transforme $(a_{i,j})$ par des opérations élémentaires, le résultat sera encore noté $(a_{i,j})$: C'est la matrice globale d'un système équivalent à (S) .

But : On veut transformer la matrice globale en une matrice $(a_{i,j})$ de dimensions $(n, p+1)$ telle que

$$\forall (i, j) \in \{1, \dots, n\} \times \{1, \dots, p\} \quad i > j \implies a_{i,j} = 0.$$

Le système correspondant est alors triangulaire et facile à résoudre.

Pour cela on procède en $\min(p, n)$ étapes, en imposant qu'à l'étape r , la matrice globale commence comme une matrice triangulaire supérieure sur ses r premières colonnes, c'est-à-dire que

$$(E_r) : \quad \forall (i, j) \in \{1, \dots, n\} \times \mathbb{N}_r \quad i > j \implies a_{i,j} = 0.$$

Pour $r = 0$: (E_0) est toujours vérifiée.

Pour $0 < r \leq \min(n, p)$: On suppose que l'étape $r - 1$ est réalisée et on effectue l'étape r de la manière suivante :

Premier cas : $\forall i \in \{r, \dots, n\} \quad a_{i,r} = 0$.

Dans ce cas, il n'y a rien à faire car (E_r) est déjà vérifiée.

Second cas : $\exists i_0 \in \{r, \dots, n\} \quad a_{i_0,r} \neq 0$: On dit que $a_{i_0,r}$ est le pivot de l'étape r .

On permute d'abord les lignes L_{i_0} et L_r . Ainsi $a_{r,r} \neq 0$. Ensuite on effectue la série d'opérations élémentaires suivante :

$$\text{for } i \text{ from } r+1 \text{ to } n \text{ do } L_i \leftarrow L_i - \frac{a_{i,r}}{a_{r,r}} L_r \quad \text{od};$$

La nouvelle matrice vérifie (E_r) .

Remarque. Replaçons-nous dans la situation du début du second cas :

Il existe différentes stratégies pour choisir le pivot parmi les $a_{i,r} \neq 0$ où $i \in \{r, \dots, n\}$:

Stratégie du pivot partiel : On choisit comme pivot un coefficient $a_{i,r}$ dont le module est maximum. Cela présente l'avantage de minimiser les erreurs d'arrondis commises lorsque l'on divise par le pivot. C'est la stratégie la plus couramment utilisée lorsque cet algorithme est programmé en langage informatique.

Stratégie humaine : Dans les cas où on applique l'algorithme du pivot à la main, on souhaite éviter autant que possible l'apparition de fractions compliquées lors de la division par le pivot. La stratégie humaine consiste donc à choisir comme pivot 1 ou -1 quand c'est possible, sinon 2 ou -2 , etc. . .

Remarque. Comme on n'effectue que des opérations élémentaires sur les lignes, les lignes de la matrice finale du système engendrent le même espace vectoriel que les lignes de la matrice initiale. La méthode du pivot permet donc de déterminer une base de l'espace vectoriel engendré par les lignes (ou les colonnes en opérant sur les colonnes) d'une matrice.

La méthode du pivot permet aussi de déterminer une base de l'image d'une application linéaire : On considère sa matrice dans des bases données et on détermine une base de ses vecteurs colonnes en appliquant la méthode du pivot au niveau des colonnes.

Remarque. Le système final présente une matrice triangulaire supérieure, la dernière colonne exceptée. On dit que le système est échelonné.

Cependant, comme les pivots peuvent être nuls, il est assez difficile de programmer la résolution de ce système échelonné.

Exercice. Soit $\lambda \in \mathbb{R}$. Déterminez la compatibilité et les éventuelles solutions du système suivant :

$$\begin{cases} \lambda x & +y & +z & +t & = 1 \\ x & +\lambda y & +z & +t & = \lambda \\ x & +y & +\lambda z & +t & = \lambda^2 \\ x & +y & +z & +\lambda t & = \lambda^3 \end{cases}.$$

Résolution : La matrice globale du système est $\begin{pmatrix} \lambda & 1 & 1 & 1 & 1 \\ 1 & \lambda & 1 & 1 & \lambda \\ 1 & 1 & \lambda & 1 & \lambda^2 \\ 1 & 1 & 1 & \lambda & \lambda^3 \end{pmatrix}.$

Le pivot de la première étape est 1 (on adopte la stratégie 'humaine'). On obtient comme nouvelle matrice :

$$\begin{pmatrix} 1 & \lambda & 1 & 1 & \lambda \\ 0 & 1-\lambda^2 & 1-\lambda & 1-\lambda & 1-\lambda^2 \\ 0 & 1-\lambda & \lambda-1 & 0 & \lambda(\lambda-1) \\ 0 & 1-\lambda & 0 & \lambda-1 & \lambda(\lambda^2-1) \end{pmatrix}.$$

Premier cas : Si $\lambda \neq 1$, on simplifie les équations par $1-\lambda$. On obtient ainsi

$$\begin{pmatrix} 1 & \lambda & 1 & 1 & \lambda \\ 0 & 1+\lambda & 1 & 1 & 1+\lambda \\ 0 & 1 & -1 & 0 & -\lambda \\ 0 & 1 & 0 & -1 & -\lambda(1+\lambda) \end{pmatrix}.$$
 Pour la seconde étape, le pivot choisi est 1.

On aboutit à

$$\begin{pmatrix} 1 & \lambda & 1 & 1 & \lambda \\ 0 & 1 & -1 & 0 & -\lambda \\ 0 & 0 & 2+\lambda & 1 & (1+\lambda)^2 \\ 0 & 0 & 1 & -1 & -\lambda^2 \end{pmatrix}.$$
 Pour la troisième étape, le pivot choisi est 1.

On aboutit à

$$\begin{pmatrix} 1 & \lambda & 1 & 1 & \lambda \\ 0 & 1 & -1 & 0 & -\lambda \\ 0 & 0 & 1 & -1 & -\lambda^2 \\ 0 & 0 & 0 & 3+\lambda & (1+\lambda)^2 + \lambda^2(\lambda+2) \end{pmatrix}.$$

1.1 : Si $\lambda \neq -3$, le système est de Cramer et l'unique solution est donnée par :

$$t = \frac{\lambda^3 + 3\lambda^2 + 2\lambda + 1}{\lambda + 3}, \quad z = -\lambda^2 + \frac{\lambda^3 + 3\lambda^2 + 2\lambda + 1}{\lambda + 3} = \frac{2\lambda + 1}{\lambda + 3},$$

$$y = -\lambda + \frac{2\lambda + 1}{\lambda + 3} = \frac{-\lambda^2 - \lambda + 1}{\lambda + 3},$$

$$x = \lambda + \frac{-\lambda^3 - 3\lambda^2 - 2\lambda - 1 - 2\lambda - 1 + \lambda^3 + \lambda^2 - \lambda}{\lambda + 3} = \frac{-\lambda^2 - 2\lambda - 2}{\lambda + 3}.$$

1.2 : Si $\lambda = -3$, le système est incompatible.

Second cas : Si $\lambda = 1$, le système est compatible,

et $(S) \iff x = 1 - y - z - t$.

3.4 Méthode du pivot total

Notation. On reprend les notations du paragraphe précédent.

But : On veut transformer la matrice globale en une matrice $(a_{i,j})$ de dimensions $(n, p+1)$ telle qu'il existe $s \in \{0, \min(n, p)\}$ vérifiant

$$(F) : \begin{array}{ll} \forall (i, j) \in \mathbb{N}_s^2 & i > j \implies a_{i,j} = 0, \\ \forall r \in \mathbb{N}_s & a_{r,r} \neq 0, \\ \forall (i, j) \in \{s+1, \dots, n\} \times \{1, \dots, p\} & a_{i,j} = 0. \end{array} \text{ , et}$$

◇ Pour cela on procède en au plus $\min(p, n)$ étapes, en imposant qu'à l'étape r , la matrice globale commence comme une matrice triangulaire supérieure sur ses r premières colonnes, les coefficients diagonaux étant non nuls, c'est-à-dire que

$$(F_r) : (\forall (i, j) \in \{1, \dots, n\} \times \mathbb{N}_r \quad i > j \implies a_{i,j} = 0) \text{ et } (\forall i \in \mathbb{N}_r \quad a_{i,i} \neq 0).$$

◇ Pour $r = 0$: (F_0) est toujours vérifiée.

◇ Pour $0 < r \leq \min(n, p)$: On suppose que l'étape $r-1$ est réalisée et on effectue l'étape r de la manière suivante :

Premier cas : $\forall (i, j) \in \{r, \dots, n\} \times \{r, \dots, p\} \quad a_{i,j} = 0$.

Dans ce cas, avec $s = r-1$, la matrice vérifie (F) : on arrête l'algorithme.

Second cas : $\exists (i_0, j_0) \in \{r, \dots, n\} \times \{r, \dots, p\} \quad a_{i_0, j_0} \neq 0$: on dit que a_{i_0, j_0} est le pivot de l'étape r .

On permute les colonnes C_{j_0} et C_r , ce qui revient à modifier l'ordre des inconnues (il faudra mémoriser ce nouvel ordre). La suite de l'algorithme est identique à celui présenté au b).

◇ A la fin de l'algorithme, on obtient la matrice globale d'un système équivalent à (S) , vérifiant (F) .

◇ Le système est compatible si et seulement si $\forall i \in \{s+1, \dots, n\} \quad a_{i, p+1} = 0$.

Si le vecteur $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ du système est quelconque, ces conditions de compatibilité

s'expriment en fonction de b_1, \dots, b_n . Elles constituent un système d'équations de l'espace vectoriel engendré par les colonnes de (S) , car (S) est compatible si et seulement si B appartient à cet espace vectoriel.

Si la matrice de (S) est celle d'une application linéaire u dans des bases e et f , ces conditions de compatibilité constituent un système d'équations de $Im(u)$ dans la base f .

◇ En cas de compatibilité, le système triangulaire final peut être facilement résolu :

Définition. Résoudre un système $(S) : MX = B$ à n équations et p inconnues, c'est déterminer une partie I de $\{1, \dots, p\}$ et une famille $(b_{i,j})_{(i,j) \in (\{1, \dots, p\} \setminus I) \times I}$ telles que :

$$\forall i \in \{1, \dots, p\} \setminus I, \quad x_i = c_i + \sum_{j \in I} b_{i,j} x_j.$$

On dit que $(x_j)_{j \in I}$ est la famille des inconnues principales et que $(x_i)_{i \in \{1, \dots, p\} \setminus I}$ est la famille des inconnues secondaires (Attention : Le choix de la partie I n'est en général pas unique).

En résumé, résoudre un système, c'est exprimer les inconnues secondaires en fonction des inconnues principales.

Après déroulement de l'algorithme du pivot total, la résolution du système triangulaire final se fait en prenant naturellement comme inconnues principales x_{s+1}, \dots, x_n .

Attention : Ces inconnues ne sont pas nécessairement les $n - s$ dernières inconnues du système d'origine, car d'éventuelles permutations de colonnes ont peut-être modifiées l'ordre des inconnues.

3.5 Méthode de Gauss-Jordan

Notation. On reprend les notations du b), en supposant que (S) est un système de Cramer, c'est-à-dire que $M \in GL_n(\mathbb{K})$.

But : On veut transformer la matrice globale en une matrice de dimension $(n, n+1)$ dont les n premières colonnes correspondent à la matrice I_n , en utilisant uniquement des opérations élémentaires sur les lignes.

Pour cela on procède en n étapes, en imposant qu'à l'étape r , la matrice globale commence comme la matrice I_n sur ses r premières colonnes.

Pour $0 < r \leq n$: On suppose que l'étape $r - 1$ est réalisée et on effectue l'étape r de la manière suivante :

La matrice obtenue après l'étape $r - 1$, que l'on notera encore M , est égale au produit d'une matrice inversible avec la matrice initiale, donc elle est encore inversible. Alors, il existe $i_0 \in \{r, \dots, n\}$ $a_{i_0,r} \neq 0$.

En effet, si pour tout $i \in \{r, \dots, n\}$, $a_{i,r} = 0$, alors on vérifie que $M \begin{pmatrix} a_{1,r} \\ \vdots \\ a_{r-1,r} \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0$, ce

qui contredit l'inversibilité de la matrice M .

$a_{i_0,r}$ est le pivot de l'étape r .

On permute alors les lignes L_{i_0} et L_r . Ainsi $a_{r,r} \neq 0$. Ensuite on effectue la série d'opérations élémentaires suivante :

$$\forall i \in \{1, \dots, n\} \setminus \{r\}, \quad L_i \longleftarrow L_i - \frac{a_{i,r}}{a_{r,r}} L_r.$$

Enfin, on réalise l'opération

$$L_r \longleftarrow \frac{1}{a_{r,r}} L_r.$$

Les r premières colonnes de la nouvelle matrice sont bien celles de la matrice I_n .

A la fin de l'algorithme, le système est immédiatement résolu puisque

$$(S) \iff \forall i \in \{1, \dots, n\} \quad x_i = a_{i,n+1}.$$

Remarque. Le dernier théorème du paragraphe 3.2 montre comment modifier cet algorithme pour calculer l'inverse d'une matrice. Ce nouvel algorithme sera encore appelé algorithme de Gauss-Jordan.

Remarque. Si, lors de la recherche du pivot de l'étape r , pour tout $i \in \{r, \dots, n\}$, $a_{i,r} = 0$, c'est que la matrice initiale du système n'était pas inversible. L'algorithme de Gauss-Jordan peut donc constituer un test efficace d'inversibilité d'une matrice.

Remarque. Ceci montre que toute matrice inversible est un produit de matrices de permutation, d'affinités et de transvections.

4 Somme de sous-espaces vectoriels

Notation. E désigne un \mathbb{K} -espace vectoriel.

4.1 Sommes et sommes directes

Définition. Soient $k \in \mathbb{N}^*$ et $(E_i)_{1 \leq i \leq k}$ une famille de k sous-espaces vectoriels de E .

On note $E_1 + \dots + E_k = \text{Vect} \left(\bigcup_{i=1}^k E_i \right)$.

Propriété. Avec les notations précédentes,

$$E_1 + \dots + E_k = \left\{ \sum_{i=1}^k x_i \mid \forall i \in \{1, \dots, k\} \quad x_i \in E_i \right\}.$$

Démonstration.

Notons $F = \left\{ \sum_{i=1}^k x_i \mid \forall i \in \{1, \dots, k\} \quad x_i \in E_i \right\}$.

- Pour tout $i \in \{1, \dots, k\}$, $0 \in E_i$, donc $0 = \sum_{i=1}^k 0 \in F$. Ainsi, $F \neq \emptyset$.

Soit $(\alpha, \beta, x, y) \in \mathbb{K}^2 \times F^2$. Il existe $(x_i)_{1 \leq i \leq k} \in E_1 \times \dots \times E_k$ et

$(y_i)_{1 \leq i \leq k} \in E_1 \times \dots \times E_k$ tels que $x = \sum_{i=1}^k x_i$ et $y = \sum_{i=1}^k y_i$.

$\alpha x + \beta y = \sum_{i=1}^k (\alpha x_i + \beta y_i) \in F$, donc F est stable par combinaison linéaire.

Ainsi, F est un sous-espace vectoriel.

- Soit $x \in \bigcup_{i=1}^k E_i$. Il existe $j \in \{1, \dots, k\}$ tel que $x \in E_j$.

Pour $i = j$, posons $x_i = x$ et pour tout $i \in \{1, \dots, k\} \setminus \{j\}$, posons $x_i = 0$.

$(x_i)_{1 \leq i \leq k} \in E_1 \times \dots \times E_k$ et $x = \sum_{i=1}^k x_i$, donc $x \in F$.

Ainsi F contient $\bigcup_{i=1}^k E_i$.

- Soit G un sous-espace vectoriel qui contient $\bigcup_{i=1}^k E_i$.

Si $(x_1, \dots, x_k) \in E_1 \times \dots \times E_k$, $\sum_{i=1}^k x_i \in G$, donc $F \subset G$.

Ainsi F est bien le plus petit sous-espace vectoriel contenant $\bigcup_{i=1}^k E_i$. \square

Exemples. Si F est un \mathbb{K} -espace vectoriel, $F + \{0\} = F = F + F$.

Notation. On note également, $E_1 + \dots + E_k = \sum_{i=1}^k E_i$.

Définition. On dit que la somme précédente est **directe** si et seulement si

$$\forall (x_1, \dots, x_k) \in E_1 \times \dots \times E_k \quad \left(\sum_{i=1}^k x_i = 0 \implies (\forall i \in \{1, \dots, k\} \quad x_i = 0) \right).$$

Dans ce cas, la somme est notée $E_1 \oplus \dots \oplus E_k$ ou encore $\bigoplus_{1 \leq i \leq k} E_i$.

Remarque. k sous-espaces vectoriels non nuls de E notés E_1, \dots, E_k sont en somme directe si et seulement si pour tout $(x_1, \dots, x_k) \in E_1 \times \dots \times E_k$, avec pour tout i , $x_i \neq 0$, la famille (x_1, \dots, x_k) est libre.

Les notions de famille libre et de somme directe sont donc très proches.

Propriété.

Reprenons les notations ci-dessus. $\sum_{i=1}^k E_i$ est une somme directe si et seulement si

$$\forall x \in \sum_{i=1}^k E_i, \exists! (x_1, \dots, x_k) \in E_1 \times \dots \times E_k \quad x = \sum_{i=1}^k x_i.$$

Démonstration.

Notons $\varphi : E_1 \times \dots \times E_k \longrightarrow \sum_{i=1}^k E_i$ l'application définie par : $\varphi(x_1, \dots, x_k) = \sum_{i=1}^k x_i$.

On vérifie que φ est linéaire. Elle est surjective par définition de $\sum_{i=1}^k E_i$.

De plus, $\text{Ker}(\varphi) = \left\{ (x_1, \dots, x_k) \in E_1 \times \dots \times E_k \mid \sum_{i=1}^k x_i = 0 \right\}$.

Ainsi, la somme est directe si et seulement si $\text{Ker}(\varphi) = \{0\}$, donc si et seulement si φ est injective, c'est-à-dire bijective, ce qui est équivalent à l'existence, pour tout $x \in \sum_{i=1}^k E_i$,

d'un unique $(x_1, \dots, x_k) \in E_1 \times \dots \times E_k$ tel que $x = \sum_{i=1}^k x_i$. \square

Exemple. Soit $n \in \mathbb{N}^*$. Notons $c = (c_1, \dots, c_n)$ la base canonique de \mathbb{K}^n . Pour tout $i \in \{1, \dots, n\}$, notons $E_i = \text{Vect}(c_i)$. Alors $\mathbb{K}^n = \bigoplus_{i=1}^n E_i$.

4.2 Supplémentaires d'un sous-espace vectoriel

Propriété. Soient F et G deux sous-espaces vectoriels de E .

$$F \text{ et } G \text{ forment une somme directe si et seulement si } F \cap G = \{0\}.$$

Démonstration.

• Supposons que F et G forment une somme directe.

Soit $x \in F \cap G$. $0 = x + (-x)$ avec $x \in F$ et $-x \in G$, donc $x = 0$. Ainsi $F \cap G \subset \{0\}$.

• Réciproquement, supposons que $F \cap G = \{0\}$.

Soit $(x, y) \in F \times G$ tel que $x + y = 0$. Alors $x = -y \in F \cap G$, donc $x = y = 0$. \square

Propriété. Soit F un sous-espace vectoriel de E .

$$\text{Si } x \in E \setminus F, F \text{ et } \mathbb{K}x \text{ sont en somme directe.}$$

Démonstration.

Soit $y \in F \cap \mathbb{K}x$. Il existe $\lambda \in \mathbb{K}$ tel que $y = \lambda x$.

Si $\lambda \neq 0$, $x = \frac{1}{\lambda}y \in F$, ce qui est faux, donc $\lambda = 0$, ce qui prouve que $y = 0$.

Ainsi, $F \cap \mathbb{K}x = \{0\}$, donc la somme de F et de $\mathbb{K}x$ est directe. \square

Corollaire. Deux droites vectorielles distinctes sont en somme directe.

Démonstration.

Soient E un \mathbb{K} -espace vectoriel et D et D' deux droites vectorielles distinctes de E .

Il existe $(x, x') \in (E \setminus \{0\})^2$ tel que $D = \text{Vect}(x)$ et $D' = \text{Vect}(x')$.

Si $x' \in D$, on montre que $D' = D$, ce qui est faux, donc $x' \in D' \setminus D$. Ainsi D et $D' = \mathbb{K}x'$ sont en somme directe. \square

Définition. On dit que deux sous-espaces vectoriels F et G de E sont **supplémentaires** (dans E) si et seulement si ils vérifient l'une des conditions équivalentes suivantes :

- i) $E = F \oplus G$.
- ii) $E = F + G$ et $F \cap G = \{0\}$.
- iii) $\forall x \in E \quad \exists!(x_1, x_2) \in F \times G \quad x = x_1 + x_2$.

Propriété. Soient E un \mathbb{K} -espace vectoriel de dimension finie et F un sous-espace vectoriel de E . F admet au moins un supplémentaire, et pour tout supplémentaire G de F , $\dim(F) + \dim(G) = \dim(E)$.

Démonstration.

F étant de dimension finie, F possède au moins une base, notée (e_1, \dots, e_p) . D'après le théorème de la base incomplète, on peut la compléter en une base $e = (e_1, \dots, e_n)$ de E . Posons $G = \text{Vect}(e_{p+1}, \dots, e_n)$.

Soit $x \in E$, on peut écrire $x = \sum_{i=1}^n x_i e_i$ où $x_i \in \mathbb{K}$.

Alors $x = \sum_{i=1}^p x_i e_i + \sum_{i=p+1}^n x_i e_i \in F + G$, donc $E = F + G$.

Soit $x \in F \cap G$. On peut écrire $x = \sum_{i=1}^p x_i e_i = \sum_{i=p+1}^n x_i e_i$, où les x_i sont des scalaires.

Alors $\sum_{i=1}^p x_i e_i - \sum_{i=p+1}^n x_i e_i = 0$, or e est une base, donc pour tout i , $x_i = 0$. Ainsi, $x = 0$, puis $F \cap G = \{0\}$. On a ainsi construit un supplémentaire de F . \square

Remarque. Tout sous-espace vectoriel de E possède au moins un supplémentaire, si l'on accepte l'axiome du choix, ce qui formellement place ce résultat hors programme.

Exemple. Deux droites vectorielles D et D' dans \mathbb{K}^2 sont supplémentaires si et seulement si elles sont distinctes. Il est important de noter que $D \cap D' = \{0\} \neq \emptyset$

et que $D \cup D' \neq \mathbb{K}^2$, ainsi D' n'a aucun rapport avec le *complémentaire* de D . En particulier, le raisonnement “ $x \notin D$, donc $x \in D'$ ” est complètement faux.

Remarque. Plus généralement, les notions d'union d'ensembles et de somme de sous-espaces vectoriels sont très voisines, mais il ne faut pas les confondre.

Si A et B sont deux parties d'un ensemble E , $A \cup B$ est le plus petit *ensemble* contenant A et B . Si ce sont des sous-espaces vectoriels, $A + B$ est le plus petit *espace vectoriel* contenant A et B .

Ainsi, la somme de sous-espaces vectoriels est à l'algèbre linéaire ce qu'est l'union de parties à la théorie des ensembles.

Cependant, et nous insistons, la somme de deux sous-espaces vectoriels *n'est pas* l'union de ces deux sous-espaces vectoriels.

En particulier, la situation suivante est possible. $E = F \oplus G$, $x \in E$, $x \notin F$ et $x \notin G$.

Propriété. On suppose que la caractéristique de \mathbb{K} est différente de 2.

$\mathcal{M}_n(\mathbb{K}) = \mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(\mathbb{K})$. Pour tout $M \in \mathcal{M}_n(\mathbb{K})$, la décomposition de M selon cette somme directe est $M = \frac{1}{2}(M + {}^tM) + \frac{1}{2}(M - {}^tM)$.

De plus $\dim(\mathcal{S}_n(\mathbb{K})) = \frac{n(n+1)}{2}$ et $\dim(\mathcal{A}_n(\mathbb{K})) = \frac{n(n-1)}{2}$.

Démonstration.

◇ Pour tout $M \in \mathcal{M}_n(\mathbb{K})$, $M = \frac{1}{2}(M + {}^tM) + \frac{1}{2}(M - {}^tM) \in \mathcal{S}_n(\mathbb{K}) + \mathcal{A}_n(\mathbb{K})$, donc $\mathcal{M}_n(\mathbb{K}) = \mathcal{S}_n(\mathbb{K}) + \mathcal{A}_n(\mathbb{K})$. De plus, si $M \in \mathcal{S}_n(\mathbb{K}) \cap \mathcal{A}_n(\mathbb{K})$, alors $M = {}^tM = -M$, donc $M = 0$. Ainsi, $\mathcal{M}_n(\mathbb{K}) = \mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(\mathbb{K})$.

◇ Soit $M \in \mathcal{A}_n(\mathbb{K})$. On sait alors que pour tout i , $M_{i,i} = 0$,

donc (1) : $M = \sum_{1 \leq i, j \leq n} M_{i,j} E_{i,j} = \sum_{1 \leq i < j \leq n} M_{i,j} (E_{i,j} - E_{j,i})$, or pour tout i, j ,

$E_{i,j} - E_{j,i} \in \mathcal{A}_n(\mathbb{K})$, donc $(E_{i,j} - E_{j,i})_{1 \leq i < j \leq n}$ est une famille génératrice de $\mathcal{A}_n(\mathbb{K})$.

De plus si $\sum_{1 \leq i < j \leq n} M_{i,j} (E_{i,j} - E_{j,i}) = 0$, pour une famille de scalaires $(M_{i,j})_{1 \leq i < j \leq n}$,

en posant pour $i > j$, $M_{i,j} = -M_{j,i}$ et pour $i = j$, $M_{i,i} = 0$, la relation (1) affirme que $(M_{i,j}) = 0$, donc $(E_{i,j} - E_{j,i})_{1 \leq i < j \leq n}$ est libre. C'est une base de $\mathcal{A}_n(\mathbb{K})$, dont la

dimension vaut donc $|\{(i, j) \in \mathbb{N}/1 \leq i < j \leq n\}| = \binom{n}{2} = \frac{n(n-1)}{2}$. □

Remarque. Lorsque $\text{car}(\mathbb{K}) = 2$, $\mathcal{S}_n(\mathbb{K}) = \mathcal{A}_n(\mathbb{K})$.

Exercice. Soit $n \in \mathbb{N}$ et P un polynôme de $\mathbb{K}[X]$ de degré $n+1$.

Montrer que l'idéal $P\mathbb{K}[X]$ est un supplémentaire de $\mathbb{K}_n[X]$ dans $\mathbb{K}[X]$.

Solution : D'après le principe de la division euclidienne, pour tout $S \in \mathbb{K}[X]$, il existe un unique couple $(Q, R) \in \mathbb{K}[X]^2$ tel que $S = PQ + R$ et $\deg(R) \leq n$, c'est-à-dire qu'il existe un unique couple $(T, R) \in \mathbb{K}[X]^2$ tel que $S = T + R$, $T \in P\mathbb{K}[X]$ et $R \in \mathbb{K}_n[X]$.

Ainsi, $\mathbb{K}[X] = \mathbb{K}_n[X] \oplus P\mathbb{K}[X]$.

4.3 Rang d'une application linéaire

Théorème. Soit $u \in L(E, F)$.

Si H est un supplémentaire de $\text{Ker}(u)$ dans E , alors $u|_H^{\text{Im}(u)}$ est un isomorphisme. Ainsi $\text{Im}(u)$ est isomorphe à tout supplémentaire de $\text{Ker}(u)$.

Remarque. Soit $u \in L(E, F)$ et H un sous-espace vectoriel de E .

On a toujours $\text{Ker}(u|_H) = \{x \in H \mid u(x) = 0\} = H \cap \text{Ker}(u)$.

Démonstration.

Posons $v = u|_H^{\text{Im}(u)}$. $\text{Ker}(v) = H \cap \text{Ker}(u) = \{0\}$, donc v est injective.

Soit $y \in \text{Im}(u)$. Il existe $x \in E$ tel que $y = u(x)$, or $E = H \oplus \text{Ker}(u)$, donc il existe $(h, k) \in H \times \text{Ker}(u)$ tel que $x = h + k$. Ainsi $y = u(h) + u(k) = u(h) = v(h)$, car $u(k) = 0$ et $h \in H$. Ceci prouve que v est surjective. \square

Définition. Soient E et F deux espaces vectoriels et $u \in L(E, F)$. On note $\text{rg}(u) = \dim(\text{Im}(u)) \in \mathbb{N} \cup \{+\infty\}$: il s'agit du rang de l'application linéaire u .

Propriété. Si e est une base de E , alors $\text{rg}(u) = \text{rg}(u(e))$.

Démonstration.

$\text{rg}(u(e)) = \dim(\text{Vect}(u(e))) = \dim(u(\text{Vect}(e))) = \dim(\text{Im}(u)) = \text{rg}(u)$. \square

Formule du rang. Soient E un \mathbb{K} -espace vectoriel de dimension finie et F un second \mathbb{K} -espace vectoriel de dimension quelconque. Soit $u \in L(E, F)$. Alors $\text{Im}(u)$ est de dimension finie et

$$\dim(\text{Im}(u)) + \dim(\text{Ker}(u)) = \dim(E).$$

Démonstration.

Avec les notations du théorème précédent, $\text{Im}(u)$ est isomorphe à H qui est de dimension finie, donc $\text{Im}(u)$ est de dimension finie et $\text{rg}(u) = \dim(H) = \dim(E) - \dim(\text{Ker}(u))$. \square

Propriété. Soient E et F deux espaces vectoriels et $u \in L(E, F)$.

Alors $\text{rg}(u) \leq \min(\dim(E), \dim(F))$. De plus,

lorsque E est de dimension finie, $\text{rg}(u) = \dim(E)$ si et seulement si u est injective et lorsque F est de dimension finie, $\text{rg}(u) = \dim(F)$ si et seulement si u est surjective.

Démonstration.

Lorsque E est de dimension finie, $\text{rg}(u) = \dim(E) - \dim(\text{Ker}(u))$, donc $\text{rg}(u) = \dim(E)$ si et seulement si $\dim(\text{Ker}(u)) = 0$, c'est-à-dire si et seulement si u est injective. \square

Théorème. Soient E, F et G 3 espaces vectoriels. Soient $u \in L(E, F)$ et $v \in L(F, G)$ tels que $\text{Im}(u)$ et $\text{Im}(v)$ sont de dimensions finies. Alors $\boxed{\text{rg}(v \circ u) \leq \inf(\text{rg}(u), \text{rg}(v))}$. De plus, si u est bijective, alors $\text{rg}(v \circ u) = \text{rg}(v)$ et si v est bijective, $\text{rg}(v \circ u) = \text{rg}(u)$. Ainsi, on ne modifie par le rang d'une application linéaire en la composant avec un isomorphisme (à sa gauche ou à sa droite).

Démonstration.

$\text{rg}(vu) = \dim(v(\text{Im}(u))) \leq \dim(\text{Im}(u))$, avec égalité lorsque v est injective.

$\text{rg}(vu) = \dim(v(u(E))) \leq \dim(v(E))$, avec égalité lorsque u est surjective. \square

Définition. Si $M \in \mathcal{M}_{\mathbb{K}}(n, p)$, le rang de M est $\text{rg}(M) \stackrel{\Delta}{=} \text{rg}(\tilde{M}) = \dim(\text{Im}(M))$.

Le rang d'une matrice est aussi le rang de la famille de ses vecteurs colonnes.

Démonstration.

On a déjà vu que $\text{Im}(M)$ est l'espace vectoriel engendré par les colonnes de M . \square

Exemple. Le rang de $\begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 2 & 1 \\ 1 & 1 & 2 & 2 \end{pmatrix}$ est 2, car les deux premières colonnes sont

libres et les suivantes sont des combinaisons linéaires des deux premières.

Propriété. Soient E et F deux \mathbb{K} -espaces vectoriels de dimensions finies, munis de bases e et f et soit $u \in L(E, F)$. Alors $\text{rg}(\text{mat}(u, e, f)) = \text{rg}(u)$.

Démonstration.

Notons $e = (e_1, \dots, e_p)$ et $f = (f_1, \dots, f_n)$.

Alors les colonnes de $\text{mat}(u, e, f)$ sont les $\Psi_f^{-1}(u(e_j))$, donc

$\text{rg}(\text{mat}(u, e, f)) = \text{rg}(\Psi_f^{-1}(u(e_j))_{1 \leq j \leq p}) = \text{rg}(u(e))$ car Ψ_f^{-1} est injective.

De plus, e étant une base de E , on a déjà vu que $\text{rg}(u(e)) = \text{rg}(u)$. \square

Propriété. $M \in \mathcal{M}_n(\mathbb{K})$ est inversible si et seulement si $\text{rg}(M) = n$.

Propriété. Soit $(A, B) \in \mathcal{M}_{\mathbb{K}}(n, p) \times \mathcal{M}_{\mathbb{K}}(p, q)$. Alors, $\text{rg}(AB) \leq \min(\text{rg}(A), \text{rg}(B))$.
On ne modifie pas le rang d'une matrice en la multipliant à gauche ou à droite par une matrice inversible.

4.4 Propriétés des sommes directes

4.4.1 Un moyen de définir une application linéaire

Théorème. Soit $(E_i)_{1 \leq i \leq k}$ une famille de k sous-espaces vectoriels de E telle que $E = \bigoplus_{1 \leq i \leq k} E_i$. Soit F un second \mathbb{K} -espace vectoriel et, pour tout $i \in \{1, \dots, k\}$, soit u_i une application linéaire de E_i dans F .

Il existe une unique application linéaire u de E dans F telle que,

pour tout $i \in \{1, \dots, k\}$, la restriction de u à E_i est égale à u_i .

Ainsi, pour définir une application linéaire u de E dans F , on peut se contenter de préciser ses restrictions aux sous-espaces vectoriels E_i .

Démonstration.

◇ *Unicité.* Supposons qu'il existe une application linéaire u de E dans F telle que, pour tout $i \in \{1, \dots, k\}$, la restriction de u à E_i est égale à u_i .

Soit $x \in E$. Il existe un unique $(x_i)_{1 \leq i \leq k}$ tel que, pour tout i , $x_i \in E_i$ et $x = \sum_{i=1}^k x_i$.

Alors $u(x) = \sum_{i=1}^k u(x_i) = \sum_{i=1}^k u_i(x_i)$, ce qui prouve l'unicité : u est nécessairement

l'application qui à $x \in E$ associe $\sum_{i=1}^k u_i(x_i)$ où $\sum_{i=1}^k x_i$ est l'unique décomposition de x

dans $\bigoplus_{1 \leq i \leq k} E_i$.

◇ *Existence*. Notons u l'application ainsi définie. Il est clair que pour tout i , $u|_{E_i} = u_i$. Il reste à montrer que u est linéaire.

Soit $x, y \in E$ et $\alpha \in \mathbb{K}$. Ecrivons $x = \sum_{i=1}^k x_i$ et $y = \sum_{i=1}^k y_i$, avec pour tout i , $x_i, y_i \in E_i$.

Alors $\alpha x + y = \sum_{i=1}^k (\alpha x_i + y_i)$, donc par définition de u ,

$$u(\alpha x + y) = \sum_{i=1}^k u_i(\alpha x_i + y_i) = \sum_{i=1}^k [\alpha u_i(x_i) + u_i(y_i)],$$

$$\text{donc } u(\alpha x + y) = \alpha \sum_{i=1}^k u_i(x_i) + \sum_{i=1}^k u_i(y_i) = \alpha u(x) + u(y). \quad \square$$

Remarque. La propriété précédente signifie que
$$\begin{array}{ccc} L(E, F) & \longrightarrow & \prod_{i=1}^k L(E_i, F) \\ u & \longmapsto & (u|_{E_i})_{1 \leq i \leq k} \end{array}$$
 est un isomorphisme, la linéarité étant simple à démontrer.

4.4.2 Formules dimensionnelles

Propriété. Soit $k \in \mathbb{N}^*$ et soit E_1, \dots, E_k k sous-espaces vectoriels de dimensions finies d'un \mathbb{K} -espace vectoriel E . Alors, $\dim\left(\sum_{i=1}^k E_i\right) \leq \sum_{i=1}^k \dim(E_i)$,

avec égalité si et seulement si la somme est directe.

Démonstration.

L'application
$$\varphi : \prod_{i=1}^k E_i \longrightarrow \sum_{i=1}^k E_i$$
 est linéaire et surjective,

$$(x_1, \dots, x_k) \longmapsto \sum_{i=1}^k x_i$$

donc
$$\sum_{i=1}^k \dim(E_i) = \dim\left(\prod_{i=1}^k E_i\right) \geq \dim\left(\sum_{i=1}^k E_i\right).$$

De plus, il y a égalité des dimensions si et seulement si φ est injective, donc si et seulement si son noyau est réduit à $\{0\}$, ce qui signifie que $\sum_{i=1}^k E_i$ est une somme directe. \square

Remarque. Ainsi, lorsque E est de dimension finie, si F et G sont deux sous-espaces vectoriels de E , ils sont supplémentaires dans E si et seulement si $E = F + G$ et $\dim(E) = \dim(F) + \dim(G)$.

Exercice. A et B sont deux sous-espaces vectoriels de E .

On suppose que A' est un supplémentaire de $A \cap B$ dans A et que B' est un supplémentaire de $A \cap B$ dans B .

Montrer que $A + B = (A \cap B) \oplus A' \oplus B'$.

Solution :

• Montrons que $A + B = (A \cap B) + A' + B'$.

◇ Les trois sous-espaces vectoriels A' , B' et $A \cap B$ sont inclus dans $A + B$, donc $(A \cap B) + A' + B' = \text{Vect}((A \cap B) \cup A' \cup B')$ est un sous-espace vectoriel de $A + B$.

◇ Soit $x \in A + B$. Il existe $(a, b) \in A \times B$ tel que $x = a + b$.

$A = (A \cap B) \oplus A'$, donc il existe $(\alpha, a') \in (A \cap B) \times A'$ tel que $a = \alpha + a'$. De même, il existe $(\beta, b') \in (A \cap B) \times B'$ tel que $b = \beta + b'$.

Ainsi $x = (\alpha + \beta) + a' + b' \in (A \cap B) + A' + B'$, donc $A + B \subset (A \cap B) + A' + B'$.

• Soit $(\alpha, a', b') \in (A \cap B) \times A' \times B'$ tel que $\alpha + a' + b' = 0$.

$a' = -\alpha - b' \in B$ et $a' \in A'$, donc $a' \in A' \cap (A \cap B)$, or $A' \cap (A \cap B) = \{0\}$ car A' et $A \cap B$ sont supplémentaires dans A . Ainsi, $a' = 0$. De même, on montre que $b' = 0$. On en déduit que $\alpha = 0$, donc la somme est bien directe.

Formule de Grassmann : Soit E un \mathbb{K} -espace vectoriel et soit F et G deux sous-espaces vectoriels de E de dimensions finies. Alors $F + G$ est de dimension finie et $\boxed{\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G)}$.

Démonstration.

Reprenons l'exercice précédent et supposons que A et B sont de dimensions finies. Alors $A + B = (A \cap B) \oplus A' \oplus B'$ est de dimension finie et

$\dim(A + B) = \dim(A \cap B) + \dim(A') + \dim(B')$, or par définition de A' et de B' , $\dim(A') = \dim(A) - \dim(A \cap B)$ et $\dim(B') = \dim(B) - \dim(A \cap B)$.

Ainsi, $\dim(A + B) = \dim(A) + \dim(B) - \dim(A \cap B)$. \square

Remarque. Cette formule est analogue à la formule donnant le cardinal d'une réunion de deux ensembles finis A et B : $|A \cup B| = |A| + |B| - |A \cap B|$. Cette dernière formule se généralise à une union de p ensembles finis : c'est la formule du crible. On peut noter que la formule du crible ne s'adapte pas au cadre des dimensions de sous-espaces vectoriels : on peut trouver 3 sous-espaces vectoriels F , G et H d'un même \mathbb{K} -espace vectoriel tels que

$$\begin{aligned} \dim(F + G + H) &\neq \dim(F) + \dim(G) + \dim(H) \\ &\quad - \dim(F \cap G) - \dim(F \cap H) - \dim(G \cap H) \\ &\quad + \dim(F \cap G \cap H). \end{aligned}$$

Il suffit en effet de prendre 3 droites deux à deux distinctes d'un plan vectoriel.

4.4.3 Associativité des sommes directes

Propriété. Associativité d'une somme directe.

Soient $k \in \mathbb{N}^*$ et E_1, \dots, E_k k sous-espaces vectoriels d'un \mathbb{K} -espace vectoriel E .

Soit $(I_i)_{1 \leq i \leq p}$ une partition de $\{1, \dots, k\}$.

E_1, \dots, E_k forment une somme directe si et seulement si

$$\begin{aligned} i) \quad & \forall i \in \{1, \dots, p\} \quad (E_j)_{j \in I_i} \text{ forment une somme directe,} \\ \text{et } ii) \quad & \left(\bigoplus_{j \in I_i} E_j \right)_{i \in \{1, \dots, p\}} \text{ forment une somme directe.} \end{aligned}$$

Exemple. On peut écrire $E_1 \oplus E_2 \oplus E_3 \oplus E_4$ si et seulement si on peut écrire $E_1 \oplus E_2$, $E_3 \oplus E_4$ et $(E_1 \oplus E_2) \oplus (E_3 \oplus E_4)$, ou bien si et seulement si on peut écrire $E_1 \oplus E_2$, $(E_1 \oplus E_2) \oplus E_3$ et $((E_1 \oplus E_2) \oplus E_3) \oplus E_4$.

Démonstration.

• Supposons que E_1, \dots, E_k forment une somme directe.

◇ Soit $i \in \{1, \dots, p\}$. Soit $(x_j)_{j \in I_i}$ une famille de vecteurs telle que

$$\forall j \in I_i \quad x_j \in E_j \text{ et } \sum_{j \in I_i} x_j = 0.$$

Complétons à l'aide de vecteurs nuls cette famille en une famille $(x_j)_{1 \leq j \leq k}$. Ainsi

$$\sum_{j=1}^k x_j = 0, \text{ mais } E_1, \dots, E_k \text{ forment une somme directe, donc la famille } (x_j)_{1 \leq j \leq k} \text{ est}$$

nulle. Ainsi $\forall j \in I_i \quad x_j = 0$, ce qui prouve i).

◇ Soit $(y_i)_{1 \leq i \leq p}$ une famille de vecteurs telle que, pour tout $i \in \{1, \dots, p\}$, $y_i \in \bigoplus_{j \in I_i} E_j$

$$\text{et } \sum_{i=1}^p y_i = 0.$$

Soit $i \in \{1, \dots, p\}$. Il existe une famille $(x_j)_{j \in I_i}$ telle que $\forall j \in I_i \quad x_j \in E_j$ et $y_i = \sum_{j \in I_i} x_j$.

$$\text{Ainsi } 0 = \sum_{i=1}^p y_i = \sum_{i=1}^p \sum_{j \in I_i} x_j = \sum_{j=1}^k x_j, \text{ or } E_1, \dots, E_k \text{ forment une somme directe, donc}$$

la famille $(x_j)_{1 \leq j \leq k}$ est nulle. Ainsi, pour tout $i \in \{1, \dots, p\}$, $y_i = 0$, ce qui prouve ii).

• Réciproquement, supposons que i) et ii) sont vraies.

Soit $(x_1, \dots, x_k) \in E_1 \times \dots \times E_k$ tel que $\sum_{h=1}^k x_h = 0$.

Pour tout $i \in \{1, \dots, p\}$, posons $y_i = \sum_{j \in I_i} x_j$. Ainsi, $y_i \in \bigoplus_{j \in I_i} E_j$.

$0 = \sum_{h=1}^k x_h = \sum_{i=1}^p y_i$ et d'après ii), $\left(\bigoplus_{j \in I_i} E_j\right)_{1 \leq i \leq p}$ forment une somme directe, donc,

pour tout $i \in \{1, \dots, p\}$, $y_i = 0$.

Soit $i \in \{1, \dots, p\}$. $0 = y_i = \sum_{j \in I_i} x_j$ et $(E_j)_{j \in I_i}$ forment une somme directe, donc pour

tout $j \in I_i$, $x_j = 0$.

Ainsi, pour tout $h \in \{1, \dots, k\}$, $x_h = 0$, ce qui prouve que E_1, \dots, E_k forment une somme directe. \square

Théorème. Soient k un entier supérieur ou égal à 2, et $(E_i)_{1 \leq i \leq k}$ une famille de k sous-espaces vectoriels de E . E_1, \dots, E_k sont en somme directe si et seulement si

$$\forall i \in \{2, \dots, k\} \quad E_i \cap \sum_{j=1}^{i-1} E_j = \{0\}.$$

Démonstration.

Effectuons une démonstration par récurrence.

Soit $k \geq 2$. Notons $R(k)$ l'assertion suivante : pour tout k -uplet (E_1, \dots, E_k) de sous-espaces vectoriels de E , E_1, \dots, E_k forment une somme directe si et seulement si

$$\forall i \in \{2, \dots, k\} \quad E_i \cap \sum_{j=1}^{i-1} E_j = \{0\}.$$

Pour $k = 2$, on a déjà montré $R(2)$ page 49, au début du paragraphe 4.2.

Pour $k \geq 2$, supposons $R(k)$.

Soit (E_1, \dots, E_{k+1}) un $(k+1)$ -uplet de sous-espaces vectoriels de E .

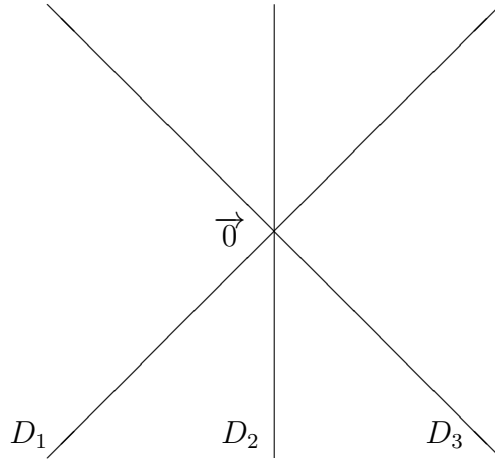
D'après l'associativité de la somme directe, (E_1, \dots, E_{k+1}) forment une somme directe

si et seulement si (E_1, \dots, E_k) forment une somme directe et si E_{k+1} et $\sum_{j=1}^k E_j$ forment une somme directe, c'est-à-dire si et seulement si pour tout $i \in \{2, \dots, k\}$,

$$E_i \cap \sum_{j=1}^{i-1} E_j = \{0\} \text{ (d'après } R(k)) \text{ et si } E_{k+1} \cap \sum_{j=1}^k E_j = \{0\} \text{ (d'après } R(2)).$$

Ceci prouve $R(k+1)$. \square

Figure.



Remarque. Une erreur fréquente est de croire que $(E_i)_{1 \leq i \leq k}$ constitue une somme directe si et seulement si, pour tout $(i, j) \in \{1, \dots, k\}$, avec $i \neq j$, $E_i \cap E_j = \{0\}$. C'est faux. En effet, la figure représente trois droites vectorielles d'un plan vectoriel P , notées D_1 , D_2 et D_3 . On sait qu'elles sont deux à deux en somme directe, donc $D_1 \cap D_2 = D_1 \cap D_3 = D_2 \cap D_3 = \{0\}$. Cependant, D_1 , D_2 et D_3 ne sont pas en somme directe, car il est facile de dessiner sur la figure ci-dessus 3 vecteurs non nuls sur D_1 , D_2 et D_3 dont la somme est nulle.

4.4.4 Base adaptée à une décomposition en somme directe

Théorème. Soit E un \mathbb{K} -espace vectoriel muni d'une base $(e_i)_{i \in I}$. Soit $(I_k)_{1 \leq k \leq n}$ une partition de I . Pour tout $k \in \{1, \dots, n\}$, on pose $E_k = \text{Vect}(e_i)_{i \in I_k}$. Alors $E = \bigoplus_{k=1}^n E_k$.

Démonstration.

- Soit $x \in E$. il existe $(\alpha_i)_{i \in I} \in \mathbb{K}^{(I)}$ telle que $x = \sum_{i \in I} \alpha_i e_i$.

$$x = \sum_{k=1}^n \left(\sum_{i \in I_k} \alpha_i e_i \right) \in \sum_{k=1}^n E_k, \text{ ainsi } E = \sum_{k=1}^n E_k.$$

- Soit $(x_k)_{1 \leq k \leq n} \in E_1 \times \dots \times E_n$ tel que $\sum_{k=1}^n x_k = 0$.

Pour tout $k \in \{1, \dots, n\}$, il existe $(\alpha_i)_{i \in I_k} \in \mathbb{K}^{(I_k)}$ telle que $x_k = \sum_{i \in I_k} \alpha_i e_i$.

Ainsi $0 = \sum_{k=1}^n x_k = \sum_{i \in I} \alpha_i e_i$, or (e_i) est une famille libre, donc, pour tout $i \in I$, $\alpha_i = 0$.

Ainsi, pour tout $k \in \{1, \dots, n\}$, $x_k = 0$, ce qui prouve que la somme est directe. \square

Théorème réciproque. Soit $(E_k)_{1 \leq k \leq n}$ une famille de sous-espaces vectoriels d'un

\mathbb{K} -espace vectoriel E tels que $E = \bigoplus_{k=1}^n E_k$. Pour tout $k \in \{1, \dots, n\}$, on suppose que E_k admet une base b_k . Alors la concaténation des bases $(b_k)_{1 \leq k \leq n}$, notée b , est une base de E . On dit que b est une **base adaptée à la décomposition en somme directe** $E = \bigoplus_{k=1}^n E_k$.

Démonstration.

- Soit $k \in \{1, \dots, n\}$. Notons $b_k = (e_i)_{i \in I_k}$. On peut supposer que les I_k sont deux à deux disjoints. Notons $I = \bigcup_{k=1}^n I_k$. Ainsi $b = (e_i)_{i \in I}$. C'est en ce sens que b est la concaténation des bases $(b_k)_{1 \leq k \leq n}$.

- Soit $(\alpha_i)_{i \in I} \in \mathbb{K}^{(I)}$ telle que $\sum_{i \in I} \alpha_i e_i = 0$.

$0 = \sum_{k=1}^n \left(\sum_{i \in I_k} \alpha_i e_i \right)$ et pour tout $k \in \{1, \dots, n\}$, $\sum_{i \in I_k} \alpha_i e_i \in E_k$. Or E_1, \dots, E_n forment une somme directe, donc, pour tout $k \in \{1, \dots, n\}$, $\sum_{i \in I_k} \alpha_i e_i = 0$. De plus, pour tout $k \in \{1, \dots, n\}$, $b_k = (e_i)_{i \in I_k}$ est libre, donc, pour tout $i \in I$, $\alpha_i = 0$, ce qui prouve que b est une famille libre.

- Soit $x \in E$. Il existe $(x_k)_{1 \leq k \leq n} \in E_1 \times \dots \times E_n$ tel que $x = \sum_{k=1}^n x_k$.

Pour tout $k \in \{1, \dots, n\}$, il existe $(\alpha_i)_{i \in I_k} \in \mathbb{K}^{(I_k)}$ telle que $x_k = \sum_{i \in I_k} \alpha_i e_i$.

Ainsi $x = \sum_{i \in I} \alpha_i e_i$, ce qui prouve que b est une famille génératrice. \square

Définition. Soient E un \mathbb{K} -espace vectoriel et F un sous-espace vectoriel de E . On appelle **base adaptée** à F toute base obtenue par “réunion” d’une base de F avec une base d’un supplémentaire de F , c’est-à-dire toute base de E obtenue en complétant une base de F .

4.5 Les projecteurs

Définition. $p \in L(E)$ est un **projecteur** si et seulement si $p^2 = p$.

Propriété. Soient F et G deux sous-espaces vectoriels supplémentaires de E .

Pour $x \in E$, on note $(p(x), q(x))$ l’unique couple de $F \times G$ tel que $x = p(x) + q(x)$.

p et q sont des projecteurs.

p est appelé le projecteur sur F parallèlement à G , et q le **projecteur associé** à p . On vérifie que $p + q = Id_E$ et $pq = qp = 0$.

Figure.

Démonstration.

◇ Soit $x, y \in E$ et $\alpha \in \mathbb{K}$.

Alors $x = p(x) + q(x)$ et $y = p(y) + q(y)$, donc $\alpha x + y = (\alpha p(x) + p(y)) + (\alpha q(x) + q(y))$ et $((\alpha p(x) + p(y)), (\alpha q(x) + q(y))) \in F \times G$. D'autre part, $\alpha x + y = p(\alpha x + y) + q(\alpha x + y)$ avec $(p(\alpha x + y), q(\alpha x + y)) \in F \times G$, donc d'après l'unicité de la décomposition d'un vecteur selon $F \oplus G$, $p(\alpha x + y) = \alpha p(x) + p(y)$ et $q(\alpha x + y) = \alpha q(x) + q(y)$.

On a montré que $p, q \in L(E)$.

◇ Soit $x \in E$. $p(x) \in F$, donc $p(p(x)) = p(x)$. De même, $q(q(x)) = q(x)$. Ainsi, p et q sont des projecteurs.

◇ Par définition, pour tout $x \in E$, $x = p(x) + q(x)$, donc $p + q = Id_E$.

◇ Soit $x \in E$. $p(x) \in F$, donc $q(p(x)) = 0$. Ainsi, $q \circ p = 0$. De même, $p \circ q = 0$. □

Exemples.

- Id_E est le projecteur sur E parallèlement à $\{0\}$.
- $0_{L(E)}$ est le projecteur sur $\{0\}$ parallèlement à E .
- L'application $p_1 : \mathbb{K}^2 \longrightarrow \mathbb{K}^2$ définie par $p_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$ est la projection sur la droite engendrée par $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ parallèlement à celle engendrée par $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.
- Soit $Q \in \mathbb{K}[X]$ un polynôme de degré $n \in \mathbb{N}^*$. L'application qui à $P \in \mathbb{K}[X]$ associe son reste pour la division euclidienne de P par Q est la projection sur $\mathbb{K}_{n-1}[X]$ parallèlement à l'idéal engendré par Q .

Propriété réciproque. Soit p un projecteur de E . Alors

p est le projecteur sur $\text{Im}(p)$ parallèlement à $\text{Ker}(p)$.

Pour tout $x \in E$, la décomposition de x selon la somme directe $E = \text{Im}(p) \oplus \text{Ker}(p)$ est $x = p(x) + (x - p(x))$, avec $p(x) \in F = \text{Im}(p)$ et $x - p(x) \in G = \text{Ker}(p)$.

Pour tout $x \in E$, $\boxed{x = p(x) \iff x \in F} : F = \text{Ker}(Id_E - p)$.

Démonstration.

Posons $F = \text{Im}(p)$ et $G = \text{Ker}(p)$.

◇ Soit $x \in E$ tel que $p(x) = x$. Alors $x \in \text{Im}(p) = F$.

Réciproquement, si $x \in F = \text{Im}(p)$, il existe $y \in E$ tel que $x = p(y)$,

donc $p(x) = p \circ p(y) = p(y) = x$ car p est un projecteur.

Ainsi $x \in F \iff p(x) = x$ et $\text{Im}(p) = F = \text{Ker}(Id_E - p)$.

◇ Soit $x \in E$. $p(x - p(x)) = p(x) - p^2(x) = 0$, car p est un projecteur, donc $x - p(x) \in \text{Ker}(p)$. De plus $p(x) \in \text{Im}(p)$, donc $x = \underbrace{p(x)}_{\in F} + \underbrace{(x - p(x))}_{\in G}$.

Ceci démontre que $E = F + G$.

◇ Soit $x \in F \cap G$. Alors $p(x) = x$ et $p(x) = 0$, donc $x = 0$. Ainsi $F \cap G = \{0\}$.

On a montré que $E = F \oplus G$.

◇ On peut donc considérer le projecteur u sur F parallèlement à G .

Soit $x \in E$. On a vu que $x = p(x) + (x - p(x))$ avec $p(x) \in F$ et $x - p(x) \in G$, donc $u(x) = p(x)$. Ainsi, $p = u$. □

Exercice. On note $p : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ l'application définie par

$p(x, y) = \left(\frac{x+y}{2}, \frac{x+y}{2}\right)$. Montrer que p est un projecteur et déterminer son noyau et son image.

Solution :

Première méthode : On vérifie que p est un endomorphisme et que pour tout $(x, y) \in \mathbb{R}^2$, $p^2(x, y) = p\left(\frac{x+y}{2}, \frac{x+y}{2}\right) = \left(\frac{x+y}{2}, \frac{x+y}{2}\right)$, donc p est bien un projecteur.

Soit $(x, y) \in \mathbb{R}^2$. $p(x, y) = 0 \iff \frac{x+y}{2} = 0 \iff (x, y) \in \text{Vect}(1, -1)$, donc $\text{Ker}(p)$ est la droite vectorielle engendrée par $(1, -1)$.

$p(x, y) = (x, y) \iff \frac{x-y}{2} = 0$, donc $\text{Im}(p)$ est la droite vectorielle engendrée par $(1, 1)$.

Ainsi, p est la projection sur la première diagonale parallèlement à la seconde diagonale.

Seconde méthode : Pour tout $(x, y) \in \mathbb{R}^2$,

$$(1) : (x, y) = \left(\frac{x+y}{2}, \frac{x+y}{2}\right) + \left(\frac{x-y}{2}, -\frac{x-y}{2}\right),$$

or $\left(\frac{x+y}{2}, \frac{x+y}{2}\right) \in D = \text{Vect}(1, 1)$ et $\left(\frac{x-y}{2}, -\frac{x-y}{2}\right) \in D' = \text{Vect}(1, -1)$,

donc (1) correspond à la décomposition de (x, y) selon la somme directe

$\mathbb{R}^2 = D \oplus D'$ (les deux droites D et D' sont distinctes, donc en somme directe, et $\dim(D \oplus D') = 2$, donc $\mathbb{R}^2 = D \oplus D'$). Ainsi p est la projection sur D parallèlement à D' .

ATTENTION : Pour un endomorphisme quelconque $u \in L(E)$, $\text{Im}(u)$ et $\text{Ker}(u)$ peuvent ne pas être supplémentaires. D'ailleurs on rencontre assez souvent des endomorphismes u tels que $u^2 = 0$, auquel cas $\text{Im}(u) \subset \text{Ker}(u)$.

Définition. $s \in L(E)$ est une **symétrie** si et seulement si $s^2 = \text{Id}_E$.

Propriété. Soient F et G deux sous-espaces vectoriels supplémentaires de E .

L'unique application s de E dans E telle que, pour tout $f, g \in F \times G$, $s(f+g) = f-g$ est une symétrie, appelée symétrie par rapport à F parallèlement à G .

Si l'on note p le projecteur sur F parallèlement à G , et q le projecteur associé à p , alors $s = p - q = 2p - \text{Id}_E$.

Figure.

Démonstration.

On sait que $p^2 = p$, $q^2 = q$ et $pq = qp = 0$, donc $s^2 = p^2 + q^2 - pq - qp = p + q = \text{Id}_E$.
□

Exemple. L'application $s : \mathbb{K}^2 \longrightarrow \mathbb{K}^2$ définie par $s \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$ est la symétrie

par rapport à la droite engendrée par $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ parallèlement à celle engendrée par $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

\mathbb{C} et \mathbb{R}^2 pouvant être identifiés par $\forall z \in \mathbb{C}, \text{Re}(z) + i\text{Im}(z) = (\text{Re}(z), \text{Im}(z))$, l'application précédente devient $z \longmapsto \bar{z}$: c'est la symétrie par rapport à la droite des réels

parallèlement à la droite des imaginaires purs, en regardant \mathbb{C} comme un \mathbb{R} -espace vectoriel.

Exemple. Lorsque $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$, l'application $s \in L(E)$ définie par $s(f)(x) = f(-x)$ est une symétrie car $s^2 = Id_E$.

Si l'on note \mathcal{P} (resp : \mathcal{I}) l'ensemble des fonctions paires (resp : impaires) de E dans E , pour tout $f \in E$ et $x \in \mathbb{R}$, $f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2} \in \mathcal{P} + \mathcal{I}$. On en déduit que $E = \mathcal{P} \oplus \mathcal{I}$ et que la projection sur \mathcal{P} parallèlement à \mathcal{I} est définie par $p(f)(x) = \frac{f(x) + f(-x)}{2}$. Son projecteur associé est défini par $q(f)(x) = \frac{f(x) - f(-x)}{2}$, donc $s = p - q$ est la symétrie par rapport à \mathcal{P} parallèlement à \mathcal{I} .

Propriété réciproque. On suppose que $\text{car}(\mathbb{K}) \neq 2$.

Pour toute symétrie s de E , il existe deux sous-espaces vectoriels supplémentaires F et G tels que s est la symétrie par rapport à F parallèlement à G . Il s'agit de $F = \text{Ker}(Id_E - s)$ et de $G = \text{Ker}(Id_E + s)$.

Démonstration.

Posons $p = \frac{1}{2}(s + Id_E)$. On sait que $s^2 = Id_E$, donc

$p^2 = \frac{1}{4}(s^2 + Id_E + 2s) = \frac{1}{2}(Id_E + s) = p$. Ainsi p est un projecteur.

Posons $F = \text{Im}(p) = \text{Ker}(Id_E - p) = \text{Ker}(Id_E - s)$ et $G = \text{Ker}(p) = \text{Ker}(Id_E + s)$.

On sait que $E = F \oplus G$ et que p est le projecteur sur F parallèlement à G , donc $s = p - q$ est la symétrie par rapport à F parallèlement à G . \square

Remarque. Cette réciproque est fautive lorsque $\text{car}(\mathbb{K}) = 2$. En effet dans ce cas, si s est la symétrie par rapport à F parallèlement à G (où $E = F \oplus G$), alors pour tout $(f, g) \in F \times G$, $s(f + g) = f - g = f + g$, car $1_{\mathbb{K}} = -1_{\mathbb{K}}$, donc $s = Id_E$.

Supposons que E de dimension finie $n \geq 2$ et fixons $e = (e_1, \dots, e_n)$ une base de E . Il existe un unique $s \in L(E)$ tel que $s(e_1) = e_2$, $s(e_2) = e_1$ et pour tout $i \in \{2, \dots, n\}$, $s(e_i) = e_i$. Alors $s^2 = s$, donc s est une symétrie, différente de l'identité.

4.6 Sous-espaces propres

Notation. On fixe un \mathbb{K} -espace vectoriel E et $u \in L(E)$.

4.6.1 Définitions

Introduction. Lorsque E est de dimension finie, la réduction de u consiste en la recherche d'une base e de E dans laquelle la matrice de u est aussi simple que possible. On considère que la matrice est d'autant plus simple qu'elle est voisine d'une matrice diagonale. Ainsi, il est intéressant de disposer d'un nombre important d'indices j tels que la $j^{\text{ème}}$ colonne de $\text{Mat}(u, e)$ a tous ses coefficients nuls, sauf éventuellement le $j^{\text{ème}}$, c'est-à-dire, en notant $e = (e_1, \dots, e_n)$, tels que $u(e_j)$ est colinéaire à e_j .

Ceci explique la présence des définitions suivantes.

Définition. $\lambda \in \mathbb{K}$ est une **valeur propre** de u si et seulement s'il existe un vecteur x non nul de E tel que $u(x) = \lambda x$.

Dans ce cas, tout vecteur y non nul tel que $u(y) = \lambda y$ est appelé un **vecteur propre** de u associé à la valeur propre λ .

De plus, toujours lorsque λ est une valeur propre de u , $\text{Ker}(\lambda \text{Id}_E - u)$ est appelé le **sous-espace propre** de u associé à la valeur propre λ . Il est noté E_λ , ou E_λ^u en cas d'ambiguïté.

Remarque. Dans la définition ci-dessus d'une valeur propre, la condition " x non nul" est essentielle. En effet, si on l'omettait, tout scalaire deviendrait une valeur propre, car, pour tout $\lambda \in \mathbb{K}$, $u(0) = \lambda \times 0$.

Remarque. Si λ est une valeur propre de u , l'ensemble des vecteurs propres de u pour la valeur propre λ est $E_\lambda \setminus \{0\}$.

Remarque. Même lorsque λ n'est pas une valeur propre de u , on note parfois $E_\lambda = \text{Ker}(\lambda \text{Id}_E - u)$, mais dans ce cas, $E_\lambda = \{0\}$.

Méthode : Pour rechercher les éléments propres de u , une méthode est d'étudier la condition $u(x) = \lambda x$. Si l'on regarde cette condition comme une équation en l'inconnue x , en considérant λ comme un paramètre, la résolution de cette équation donne les valeurs propres et les sous-espaces propres.

Exemple. Choisissons pour E l'ensemble des applications de classe C^∞ de \mathbb{R} dans \mathbb{R} et pour u ,

$$\begin{array}{ccc} E & \longrightarrow & E \\ f & \longmapsto & f' \end{array}$$

Soient $\lambda \in \mathbb{R}$ et $f \in E$. $u(f) = \lambda f \iff f' = \lambda f \iff [\exists C \in \mathbb{R} \ \forall x \in \mathbb{R} \ f(x) = Ce^{\lambda x}]$, donc tout réel est une valeur propre de u et, pour tout $\lambda \in \mathbb{R}$, E_λ est la droite vectorielle engendrée par l'application

$$\begin{array}{ccc} f_\lambda : \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & e^{\lambda x} \end{array}$$

Définition. Soient $n \in \mathbb{N}^*$ et $M \in \mathcal{M}_n(\mathbb{K})$: Les **éléments propres** de M (c'est-à-dire les valeurs propres, les vecteurs propres et les sous-espaces propres) sont les éléments propres de l'endomorphisme canoniquement associé à M .

En particulier, pour tout $\lambda \in \text{Sp}(M)$, $E_\lambda^M = \text{Ker}(\lambda I_n - M) = \{X \in \mathbb{K}^n / MX = \lambda X\}$.

Propriété.

$\lambda \in \mathbb{K}$ est une valeur propre de u si et seulement si $\lambda \text{Id}_E - u$ n'est pas injective.

En particulier, u est injectif si et seulement si 0 n'est pas une valeur propre de u .

Démonstration.

λ est une valeur propre de u si et seulement s'il existe un vecteur x non nul de E tel que $(\lambda \text{Id}_E - u)(x) = 0$, c'est-à-dire si et seulement si $\text{Ker}(\lambda \text{Id}_E - u) \neq \{0\}$. \square

Définition. On appelle **spectre** de u l'ensemble des valeurs propres de u .

Il est souvent noté $\text{Sp}(u)$.

Théorème.

La somme d'un nombre fini de sous-espaces propres de u est toujours directe.

Démonstration.

Effectuons une récurrence portant sur le nombre de sous-espaces propres.

Soit $h \in \mathbb{N}^*$. Notons $R(h)$ l'assertion suivante :

pour toute famille $(\lambda_1, \dots, \lambda_h)$ de h valeurs propres deux à deux distinctes de u , la

somme $\sum_{j=1}^h E_{\lambda_j}$ est directe.

- Supposons que $h = 1$. Toute famille $(E_i)_{1 \leq i \leq 1}$ constituée d'un seul sous-espace vectoriel de E forme une somme directe, car, pour tout $x_1 = (x_i)_{1 \leq i \leq 1} \in E_1$,

si $\sum_{i=1}^1 x_i = 0$, alors $x_1 = 0$.

Ainsi, la propriété $R(1)$ est vraie.

- Lorsque $h \geq 1$, supposons $R(h)$.

Soit $(\lambda_1, \dots, \lambda_{h+1})$ une famille de $h+1$ valeurs propres deux à deux distinctes de u .

Soit $(x_1, \dots, x_{h+1}) \in E_{\lambda_1} \times \dots \times E_{\lambda_{h+1}}$ tel que $\sum_{j=1}^{h+1} x_j = 0$.

Ainsi, (1) : $-x_{h+1} = \sum_{j=1}^h x_j$.

Prenons l'image de (1) par u . Ainsi, $-\lambda_{h+1}x_{h+1} = \sum_{j=1}^h \lambda_j x_j$.

Multiplions (1) par λ_{h+1} . Ainsi, $-\lambda_{h+1}x_{h+1} = \lambda_{h+1} \sum_{j=1}^h x_j$.

Effectuons la différence des deux égalités précédentes. $0 = \sum_{j=1}^h (\lambda_j - \lambda_{h+1})x_j$.

Or, pour tout $j \in \mathbb{N}_h$, $(\lambda_j - \lambda_{h+1})x_j \in E_{\lambda_j}$, et, d'après $R(h)$, la famille $(E_{\lambda_j})_{1 \leq j \leq h}$ constitue une somme directe. Ainsi, si $j \in \mathbb{N}_h$, $(\lambda_j - \lambda_{h+1})x_j = 0$, or $\lambda_j - \lambda_{h+1} \neq 0$, donc $x_j = 0$.

L'égalité (1) prouve alors que $x_{h+1} = 0$, ce qui montre $R(h+1)$. \square

Corollaire. Si $(x_i)_{i \in I}$ est une famille de vecteurs propres de u associés à des valeurs propres deux à deux distinctes, alors cette famille est libre.

Démonstration.

Pour tout $i \in I$, il existe $\lambda_i \in Sp(u)$ tel que $u(x_i) = \lambda_i x_i$.

De plus, par hypothèse, pour tout $(i, j) \in I^2$ tel que $i \neq j$, $\lambda_i \neq \lambda_j$.

Soit $(\alpha_i)_{i \in I}$ une famille presque nulle de scalaires telle que $\sum_{i \in I} \alpha_i x_i = 0$.

Supposons que (α_i) est une famille non nulle. Posons $J = \{i \in I / \alpha_i \neq 0\}$. J étant fini mais non vide, les sous-espaces propres E_{λ_i} , pour $i \in J$, constituent une somme directe, or $\sum_{i \in J} \alpha_i x_i = 0$ et, pour tout $i \in J$, $\alpha_i x_i \in E_{\lambda_i}$, donc, pour tout $i \in J$, $\alpha_i x_i = 0$.

Les x_i étant des vecteurs propres, ils sont non nuls, donc, pour tout $i \in J$, $\alpha_i = 0$, ce qui est faux. Ainsi la famille $(\alpha_i)_{i \in I}$ est nulle, ce qui prouve que la famille $(x_i)_{i \in I}$ est libre. \square

Exemple. Reprenons pour E l'ensemble des applications de classe C^∞ de \mathbb{R} dans \mathbb{R} et pour u , $\begin{matrix} E & \longrightarrow & E \\ f & \longmapsto & f' \end{matrix}$. En posant $\begin{matrix} f_\lambda : \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & e^{\lambda x} \end{matrix}$, l'étude précédente de u montre que $(f_\lambda)_{\lambda \in \mathbb{R}}$ est une famille libre de E .

4.6.2 Exemples

Lemme : Si F est un sous-espace vectoriel de E tel que $F \oplus E = E$, alors $F = \{0\}$.

Démonstration.

$$\{0\} = F \cap E = F. \quad \square$$

Propriété. Supposons que $E \neq \{0\}$.

Soient F et G deux sous-espaces vectoriels supplémentaires non nuls dans E .

- Si u est une homothétie de rapport λ , où $\lambda \in \mathbb{K}$, $Sp(u) = \{\lambda\}$ et $E_\lambda = E$.
- Si u est le projecteur sur F parallèlement à G , $Sp(u) = \{0, 1\}$, $E_1 = F$ et $E_0 = G$.
- Si u est la symétrie par rapport à F parallèlement à G , $Sp(u) = \{1, -1\}$, $E_1 = F$ et $E_{-1} = G$.

Démonstration.

- Supposons que $u = \lambda Id_E$, où $\lambda \in \mathbb{K}$.

Pour tout $x \in E \setminus \{0\}$, $u(x) = \lambda x$, donc x est un vecteur propre associé à λ .

Ainsi $Sp(u) = \{\lambda\}$ et $E_\lambda = E$.

- Supposons que u est le projecteur sur F parallèlement à G .

$x \in F \iff u(x) = x$, or $F \neq \{0\}$, donc $1 \in Sp(u)$ et $F = E_1$.

$x \in G \iff u(x) = 0$, or $G \neq \{0\}$, donc $0 \in Sp(u)$ et $G = E_0$.

Supposons qu'il existe une troisième valeur propre $\lambda \in \mathbb{K} \setminus \{0, 1\}$.

Alors $E_\lambda \oplus (E_0 \oplus E_1) = E$, donc d'après le lemme, $E_\lambda = \{0\}$, ce qui est impossible.

Ainsi, u admet 0 et 1 pour seules valeurs propres.

On a montré que $Sp(u) = \{0, 1\}$, avec $E_0 = G$ et $E_1 = F$.

- Supposons que u est la symétrie par rapport à F parallèlement à G .

$x \in F \iff u(x) = x$, or $F \neq \{0\}$, donc $1 \in Sp(u)$ et $F = E_1$.

$x \in G \iff u(x) = -x$, or $G \neq \{0\}$, donc $-1 \in Sp(u)$ et $G = E_{-1}$.

Comme dans le cas du projecteur, on montre que 1 et -1 sont les seules valeurs propres de u . \square

4.6.3 Propriétés

Propriété.

Si $v \in L(E)$ commute avec u , les sous-espaces propres de u sont stables par v .

Démonstration.

Supposons que v et u commutent. Soit $\lambda \in Sp(u)$.

Soit $x \in E_\lambda^u$. $u(v(x)) = v(u(x)) = v(\lambda x) = \lambda v(x)$, donc $v(x) \in E_\lambda^u$.

Ainsi, $v(E_\lambda^u) \subset E_\lambda^u$. \square

Propriété. Soit F un sous-espace vectoriel de E stable par u .

On note $u_{/F}$ l'endomorphisme induit par u sur F .

Alors $Sp(u_{/F}) \subset Sp(u)$ et pour tout $\lambda \in Sp(u_{/F})$, $E_\lambda^{u_{/F}} = E_\lambda^u \cap F$.

Démonstration.

Soit $\lambda \in Sp(u_{/F})$. Il existe $x \in F \setminus \{0\}$ tel que $u_{/F}(x) = \lambda x$,

or $u_{/F}(x) = u(x)$, donc $\lambda \in Sp(u)$.

De plus, si $x \in F$, $x \in E_\lambda^{u_{/F}} \iff u_{/F}(x) = \lambda x \iff u(x) = \lambda x \iff x \in E_\lambda^u$,

donc $E_\lambda^{u_{/F}} = E_\lambda^u \cap F$. \square

5 Changement de base

5.1 Matrice de passage

Notation. On fixe un \mathbb{K} -espace vectoriel E de dimension finie égale à $n \in \mathbb{N}^*$.

Propriété. Soit $e = (e_1, \dots, e_n)$ une base de E et $f = (f_j)_{1 \leq j \leq n} \in E^n$ une famille de n vecteurs de E . Pour tout $j \in \mathbb{N}_n$, on pose $p_{i,j} = e_i^*(f_j)$: c'est la $i^{\text{ème}}$ coordonnée dans

la base e du $j^{\text{ème}}$ vecteur de la famille f . Ainsi, pour tout $j \in \mathbb{N}_n$, $f_j = \sum_{i=1}^n p_{i,j} e_i$.

f est une base si et seulement si la matrice $P = (p_{i,j})$ est inversible. Dans ce cas, P est noté P_e^f (ou bien $P_{e \rightarrow f}$) et on dit que $P_e^f = (p_{i,j})$ est la **matrice de passage** de la base e vers la base f .

Démonstration.

$P = \text{mat}(u, e, f)$ où u est l'unique endomorphisme tel que $u(e) = f$. P est inversible si et seulement si u est inversible, ce qui est vrai si et seulement si f est une base de E .

\square

Interprétation tabulaire : Avec les notations précédentes,

$$P_e^f = \begin{pmatrix} f_1 & \cdots & f_n \\ p_{1,1} & \cdots & p_{1,n} \\ \vdots & & \vdots \\ p_{n,1} & \cdots & p_{n,n} \end{pmatrix} \begin{matrix} e_1 \\ \vdots \\ e_n \end{matrix}.$$

Remarque. La notation P_e^f reflète cette interprétation tabulaire.

Par analogie, on notera parfois $\text{mat}(u, e, f)$ sous la forme $\text{mat}(u)_f^e$.

Remarque. On pourrait également définir la matrice d'une famille de p vecteurs

$f = (f_j)_{1 \leq j \leq p}$ de E selon la base $e = (e_1, \dots, e_n)$ de E :

$$\text{mat}_e^f = (e_i^*(f_j))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in \mathcal{M}_{n,p}(\mathbb{K}).$$

Alors les colonnes de mat_e^f sont les $\Psi_e^{-1}(f_j)$, donc

$$\begin{aligned} \text{rg}(\text{mat}_e^f) &= \dim(\text{Vect}(\Psi_e^{-1}(f_j)_{1 \leq j \leq p})) \\ &= \dim(\Psi_e^{-1}(\text{Vect}(f_j)_{1 \leq j \leq p})) \\ &= \dim(\text{Vect}(f_j)_{1 \leq j \leq p}), \end{aligned}$$

car Ψ_e^{-1} est injective. Ainsi, $\text{rg}(\text{mat}_e^f) = \text{rg}(f)$.

On retrouve ainsi que f est une base si et seulement si mat_e^f est inversible.

Exemple. Notons $E = \text{Vect}(\cos, \sin)$, $f = (x \mapsto e^{ix})$ et $g = (x \mapsto e^{-ix})$. Alors $\text{mat}_{(\cos, \sin)}^{(f, g)} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$. Le déterminant de cette matrice est non nul, donc elle est

inversible. Ainsi (f, g) est une base de E et $P_{(\cos, \sin)}^{(f, g)} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$.

Propriété. Soit e une base de E : Pour toute matrice P inversible d'ordre n , il existe une unique base f de E telle que $P = P_e^f$.

Démonstration.

Soit $P \in GL_n(\mathbb{K})$. Si f est une base telle que $P = P_e^f$, alors pour tout $j \in \mathbb{N}_n$,

$$f_j = \sum_{i=1}^n P_{i,j} e_i, \text{ ce qui prouve l'unicité de } f.$$

Posons $f = u(e)$, où u est l'unique automorphisme de E tel que $P = \text{mat}(u, e, f)$. Alors $f = u(e)$ est une base et $P = P_e^f$. Ceci prouve l'existence. \square

Propriété. Soit e et e' deux bases de E . Alors $P_e^{e'} = \text{mat}(Id_E, e', e) = \text{mat}(Id_E)_e^{e'}$.

Remarque. ci-dessous, nous réutilisons la notation \hat{x} définie page 36.

Formule de changement de base pour les vecteurs :

Soit e et e' deux bases de E . Soit $x \in E$.

On pose $X = \Psi_e^{-1}(x)$ le vecteur colonne des coordonnées de x dans la base e .

On notera indifféremment $X = \Psi_e^{-1}(x) = \text{mat}(\hat{x}, 1, e) = \text{mat}(\hat{x})_e^1 = \text{mat}(x)_e$.

De même on pose $X' = \Psi_{e'}^{-1}(x) = \text{mat}(x)_{e'}$.

Alors, $\boxed{X = P_e^{e'} X'}$, ou encore $\text{mat}(x)_e = P_e^{e'} \text{mat}(x)_{e'}$.

Démonstration.

$$X = \text{mat}(\hat{x}, 1, e) = \text{mat}(Id_E \circ \hat{x}, 1, e) = \text{mat}(Id_E, e', e) \times \text{mat}(\hat{x}, 1, e'). \quad \square$$

Remarque. Cette formule n'est pas naturelle. En effet, P étant appelée la matrice de passage de la base (ancienne) e vers la (nouvelle) base e' , on est plutôt en droit d'attendre qu'elle permette d'exprimer simplement les nouvelles coordonnées X' en fonction des anciennes coordonnées X , or c'est l'inverse qui se produit.

Exemple. Soit E un \mathbb{R} -plan vectoriel muni d'une base (i, j) . Soit $\theta \in \mathbb{R}$. On pose $u_\theta = (\cos \theta)i + (\sin \theta)j$ et $v_\theta = -(\sin \theta)i + (\cos \theta)j$. Posons $P = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$: $\det(P) = 1 \neq 0$, donc P est inversible. Ainsi (u_θ, v_θ) est une base de E et $P = P_{(i, j)}^{(u_\theta, v_\theta)}$.

Soit M un point de E . Notons (x_0, y_0) les coordonnées de M dans (i, j) et (x_θ, y_θ) ses coordonnées dans (u_θ, v_θ) . Alors $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_\theta \\ y_\theta \end{pmatrix}$, donc en inversant cette relation, $\begin{pmatrix} x_\theta \\ y_\theta \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$.

Formule. Si e, e' et e'' sont trois bases de E , $P_e^{e''} = P_e^{e'} P_{e'}^{e''}$ et $(P_e^{e'})^{-1} = P_{e'}^e$.

Démonstration.

$Id_E = Id_E \circ Id_E$ et $P_e^{e'} \times P_{e'}^e = P_e^e = I_n$. \square

Formule de changement de bases pour les applications linéaires :

Soient E et F deux \mathbb{K} -espaces vectoriels de dimensions finies.

On suppose que e et e' sont deux bases de E et que f et f' sont deux bases de F .

Soit $u \in L(E, F)$. Notons $M = \text{mat}(u)_f^e$, $M' = \text{mat}(u)_{f'}^{e'}$, $P = P_e^{e'}$ et $Q = Q_f^{f'}$. Alors,

$$\boxed{M' = Q^{-1}MP} \text{ c'est-à-dire } \boxed{\text{mat}(u)_{f'}^{e'} = P_{f'}^f \times \text{mat}(u)_f^e \times P_e^{e'}}.$$

Démonstration.

$u = Id_F \circ u \circ Id_E$. \square

Formule de changement de bases pour les endomorphismes :

Soit E un \mathbb{K} -espace vectoriel de dimension finie et $u \in L(E)$. On suppose que e et e' sont deux bases de E . Notons $M = \text{mat}(u, e)$, $M' = \text{mat}(u, e')$ et $P = P_e^{e'}$. Alors,

$$\boxed{M' = P^{-1}MP}.$$

5.2 Diagonalisation et trigonalisation

Exemple. Considérons la matrice $M = \begin{pmatrix} 0 & -1 & 2 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix}$.

Commençons par en rechercher les éléments propres : soit $\lambda \in \mathbb{R}$ et $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$.

$$MX = \lambda X \iff \begin{cases} -y + 2z &= \lambda x \\ y &= \lambda y \\ x + y - z &= \lambda z \end{cases}.$$

Si $\lambda = 1$, $MX = X \iff x + y - 2z = 0$, donc $1 \in Sp(M)$ et

$E_1^M = \text{Ker}(M - I_3) = \text{Vect}\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}\right)$: il s'agit d'un plan vectoriel. Si $\lambda \neq 1$,

$MX = \lambda X \iff (y = 0, 2z = \lambda x, x = (\lambda + 1)z) \iff (y = 0, 2z = \lambda x, 2x = (\lambda + 1)\lambda x)$,
or $2 = (\lambda + 1)\lambda \iff (\lambda - 1)(\lambda + 2) = 0$. Si $\lambda = -2$, $MX = -2X \iff (y = 0, z = -x)$,

donc $-2 \in Sp(M)$ et $E_{-2}^M = \text{Vect}\left(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right)$.

Ainsi $Sp(M) = \{1, -2\}$ et si l'on pose $e_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ et $e_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $e = (e_1, e_2, e_3)$ constitue une base de vecteurs propres de M : en effet, (e_1, e_2) est une base de E_1^M donc e est une base de $E_1^M \oplus E_{-2}^M = \mathbb{R}^3$.

Ainsi, M est diagonalisable. En effet, $M = \text{mat}(\tilde{M}, c)$, où c désigne la base canonique de \mathbb{R}^3 et $\text{mat}(\tilde{M}, e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$, que l'on notera D .

D'après la formule de changement de bases pour les endomorphismes, $M = PDP^{-1}$, où $P = P_c^e = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}$.

Définition. Soit E un \mathbb{K} -espace vectoriel de dimension $n \in \mathbb{N}^*$ et $u \in L(E)$.

Les propriétés suivantes sont équivalentes.

- i) Il existe une base e de E telle que $\text{mat}(u, e)$ est diagonale.
- ii) Il existe une base de E constituée de vecteurs propres de u .
- iii) $E = \bigoplus_{\lambda \in Sp_{\mathbb{K}}(u)} E_{\lambda}^u$.
- iv) $n = \sum_{\lambda \in Sp_{\mathbb{K}}(u)} \dim(E_{\lambda}^u)$.

Lorsqu'elles sont vraies, on dit que u est **diagonalisable**.

Démonstration.

- i) \implies ii). Supposons qu'il existe une base $e = (e_1, \dots, e_n)$ de E telle que $\text{mat}(u, e)$ est diagonale. Soit $j \in \mathbb{N}_n$. La $j^{\text{ème}}$ colonne de $\text{mat}(u, e)$ a tous ses coefficients nuls, sauf éventuellement le $j^{\text{ème}}$, donc $u(e_j)$ est colinéaire à e_j . Ainsi, pour tout $j \in \mathbb{N}_n$, e_j est un vecteur propre de u .
- ii) \implies iii). Supposons qu'il existe une base $e = (e_1, \dots, e_n)$ de E constituée de vecteurs propres de u .

Soit $j \in \mathbb{N}_n$. Il existe $\mu \in Sp(u)$ telle que $e_j \in E_{\mu}^u$, donc $e_j \in \bigoplus_{\lambda \in Sp_{\mathbb{K}}(u)} E_{\lambda}^u$.

Ainsi, $\{e_1, \dots, e_n\} \subset \bigoplus_{\lambda \in Sp_{\mathbb{K}}(u)} E_{\lambda}^u$, donc $E = \text{Vect}(\{e_1, \dots, e_n\}) \subset \bigoplus_{\lambda \in Sp_{\mathbb{K}}(u)} E_{\lambda}^u$.

- iii) \implies iv). Supposons que $E = \bigoplus_{\lambda \in Sp_{\mathbb{K}}(u)} E_{\lambda}^u$.

Alors, $n = \dim(E) = \dim \left(\bigoplus_{\lambda \in Sp_{\mathbb{K}}(u)} E_{\lambda}^u \right) = \sum_{\lambda \in Sp_{\mathbb{K}}(u)} \dim(E_{\lambda}^u)$.

- iv) \implies i). Supposons que $n = \sum_{\lambda \in Sp_{\mathbb{K}}(u)} \dim(E_{\lambda}^u)$.

Pour tout $\lambda \in Sp_{\mathbb{K}}(u)$, choisissons une base de E_{λ}^u , notée e_{λ} . D'après le théorème de la base adaptée à une décomposition en somme directe, la "réunion" des e_{λ} pour

$\lambda \in Sp_{\mathbb{K}}(u)$ est une famille libre de E . De plus elle est de cardinal

$\sum_{\lambda \in Sp_{\mathbb{K}}(u)} \dim(E_{\lambda}^u) = n = \dim(E)$, donc c'est une base de E , constituée de vecteurs propres de u . Ainsi, la matrice de u dans cette base est diagonale. \square

Propriété. les homothéties, les projecteurs et les symétries sont diagonalisables.

Démonstration.

On a déjà vu que, pour chacun de ces endomorphismes, la somme des sous-espaces propres est égale à l'espace E en entier. \square

Définition. Soit $M \in \mathcal{M}_n(\mathbb{K})$. On dit que M est diagonalisable si et seulement si son endomorphisme canoniquement associé est diagonalisable.

Propriété. $M \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable si et seulement si il existe $P \in GL_n(\mathbb{K})$ telle que $P^{-1}MP$ est une matrice diagonale.

Démonstration.

Notons u l'endomorphisme canoniquement associé à M et c la base canonique de \mathbb{K}^n .

- Supposons que M est diagonalisable. Alors u est diagonalisable, donc il existe une base e de \mathbb{K}^n telle que $\text{mat}(u, e)$ est diagonale.

Or $\text{mat}(u, e) = P^{-1}MP$, où $P = P_c^e \in GL_n(\mathbb{K})$.

- Réciproquement, supposons qu'il existe $P \in GL_n(\mathbb{K})$ et une matrice D diagonale telles que $M = PDP^{-1}$.

Il existe une base e de \mathbb{K}^n telle que P est la matrice de passage de c vers e . Alors, $\text{mat}(u, e) = P^{-1}\text{mat}(u, c)P = P^{-1}PDP^{-1}P = D$, donc $\text{mat}(u, e)$ est diagonale, ce qui prouve que u est diagonalisable. \square

Définition. Soit $M \in \mathcal{M}_n(\mathbb{K})$ une matrice diagonalisable.

"diagonaliser" M , c'est déterminer une matrice diagonale D et une matrice inversible P telles que $M = PDP^{-1}$.

Définition. Un endomorphisme u est **trigonalisable** si et seulement s'il existe une base dans laquelle la matrice de u est triangulaire supérieure.

Définition. $M \in \mathcal{M}_n(\mathbb{K})$ est trigonalisable si et seulement si l'endomorphisme canoniquement associé à M est trigonalisable, c'est-à-dire si et seulement si il existe $P \in GL_n(\mathbb{K})$ telle que $P^{-1}MP$ est triangulaire supérieure.

Démonstration.

Notons u l'endomorphisme canoniquement associé à M et $c = (c_1, \dots, c_n)$ la base canonique de \mathbb{K}^n .

- Supposons que M est trigonalisable. Ainsi, il existe une base $e = (e_1, \dots, e_n)$ de \mathbb{K}^n telle que $T = \text{mat}(u, e)$ est triangulaire supérieure. Donc, en notant $P = P_c^e$, $P^{-1}MP = T$ est triangulaire supérieure.

- Réciproquement, supposons qu'il existe $P \in GL_n(\mathbb{K})$ telle que $M = PTP^{-1}$, où T est triangulaire supérieure. Il existe une base e de \mathbb{K}^n telle que $P = P_c^e$. Comme $M = PTP^{-1}$, $T = \text{mat}(u, e)$, donc M est trigonalisable. \square

Définition. Soit $M \in \mathcal{M}_n(\mathbb{K})$. “**Trigonaliser**” M , c’est déterminer si M est trigonalisable, et dans ce cas, c’est calculer $P \in GL_n(\mathbb{K})$ et T triangulaire supérieure telles que $M = PTP^{-1}$.

5.3 Trace d’un endomorphisme

Définition.

Soit E un \mathbb{K} -espace vectoriel de dimension finie. La quantité $\text{Tr}(\text{mat}(u, e))$ ne dépend pas du choix de la base e de E . On la note $\text{Tr}(u)$. C’est la trace de l’endomorphisme u .

Démonstration.

Soient e et f deux bases de E .

$\text{mat}(u, e)$ et $\text{mat}(u, f)$ sont deux matrices semblables, donc elles ont la même trace. \square

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

Pour tout $u, v \in L(E)$, $\text{Tr}(uv) = \text{Tr}(vu)$.

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

Si p est un projecteur de E , alors $\text{Tr}(p) = \text{rg}(p)$.

Démonstration.

Il existe F et G avec $E = F \oplus G$ tels que p est le projecteur sur F parallèlement à G . Notons $\text{rg}(p) = r = \dim(F)$ et $n = \dim(E)$. Considérons une base (e_1, \dots, e_r) de F et une base (e_{r+1}, \dots, e_n) de G . La famille de vecteurs $e = (e_1, \dots, e_n)$ est une base de E , adaptée à la décomposition en somme directe $E = F \oplus G$.

La matrice de p dans e est égale à la matrice blocs suivante : $\begin{pmatrix} I_r & 0_{r, n-r} \\ 0_{n-r, r} & 0_{n-r, n-r} \end{pmatrix}$, dont la trace vaut r . Ainsi $\text{Tr}(p) = r = \text{rg}(p)$. \square

5.4 Matrices équivalentes et matrices semblables

5.4.1 Matrices équivalentes

Définition. Deux matrices M et M' de $\mathcal{M}_{\mathbb{K}}(n, p)$ sont **équivalentes** si et seulement s’il existe $P \in GL_p(\mathbb{K})$ et $Q \in GL_n(\mathbb{K})$ telles que $M' = QMP^{-1}$.

On définit ainsi une relation d’équivalence sur $\mathcal{M}_{\mathbb{K}}(n, p)$.

Propriété. Soient E et F deux \mathbb{K} -espaces vectoriels de dimensions respectives $p > 0$ et $n > 0$, munis des bases e et f , et soit $u \in L(E, F)$. On note $M = \text{mat}(u, e, f)$.

Soit $M' \in \mathcal{M}_{\mathbb{K}}(n, p)$: M' est équivalente à M si et seulement s’il existe des bases e' et f' telles que $M' = \text{mat}(u, e', f')$. En résumé, deux matrices sont équivalentes si et seulement si elles représentent une même application linéaire dans des bases différentes, autant pour la base de départ que pour la base d’arrivée.

Démonstration.

Supposons qu’il existe des bases e' et f' de E et F telles que $M' = \text{mat}(u, e', f')$.

En notant $P = P_e^{e'}$ et $Q = P_f^{f'}$, $M' = Q^{-1}MP$, donc M' est équivalente à M .

Réciproquement, supposons que M' est équivalente à M . Il existe $P \in GL_p(\mathbb{K})$ et $Q \in GL_n(\mathbb{K})$ telles que $M' = Q^{-1}MP$. Il existe une base e' de E et une base f' de F telles que $P = P_e^{e'}$ et $Q = P_f^{f'}$. Ainsi $M' = Q^{-1}MP = \text{mat}(u, e', f')$. \square

Propriété. Deux matrices sont équivalentes si et seulement si il est possible de transformer l'une en l'autre par une succession d'opérations élémentaires portant sur les lignes ou sur les colonnes.

Démonstration.

Notons M et M' ces deux matrices.

$\diamond \Leftarrow$: par hypothèse, il existe deux matrices inversibles P (qui provient des opérations élémentaires portant sur les colonnes) et Q (qui provient des opérations élémentaires portant sur les lignes) telles que $M' = QMP$, donc M et M' sont équivalentes.

$\diamond \Rightarrow$: par hypothèse, il existe deux matrices inversibles P et Q telles que $M' = QMP$. D'après l'algorithme de Gauss-Jordan d'inversion d'une matrice, la matrice inversible Q peut être transformée en la matrice identité par une succession d'opérations élémentaires portant sur les lignes. Ces opérations élémentaires correspondent globalement au fait de multiplier Q à sa gauche par une matrice Q' , donc ces mêmes opérations transforment M' en $Q'M' = (Q'Q)MP = MP$. On raisonne de même à droite pour la matrice P en utilisant des opérations élémentaires portant sur les colonnes. \square

Théorème. Soient E et F deux \mathbb{K} -espaces vectoriels de dimensions respectives $p > 0$ et $n > 0$, et soit $u \in L(E, F)$. Notons r le rang de u .

Il existe une base e de E et une base f de F telles que $\text{mat}(u, e, f)$ admet la décomposition en blocs suivante :

$$\text{mat}(u, e, f) = \begin{pmatrix} I_r & 0_{r, p-r} \\ 0_{n-r, r} & 0_{n-r, p-r} \end{pmatrix}.$$

Pour la suite, cette matrice sera notée $J_{n,p,r}$.

Démonstration.

D'après la formule du rang, $\dim(\text{Ker}(u)) = p - r$: notons (e_{r+1}, \dots, e_p) une base de $\text{Ker}(u)$, que l'on complète en une base $e = (e_1, \dots, e_p)$ de E .

Posons $H = \text{Vect}(e_1, \dots, e_r)$: d'après le théorème de la base adaptée à une décomposition en somme directe, $E = H \oplus \text{Ker}(u)$. On sait alors que $u|_H^{\text{Im}(u)}$ est un isomorphisme, donc $(u(e_1), \dots, u(e_r))$ est une base de $\text{Im}(u)$, que l'on complète en une base $f = (u(e_1), \dots, u(e_r), f_{r+1}, \dots, f_n)$ de F . Alors $\text{mat}(u, e, f) = J_{n,p,r}$. \square

Propriété. Si $M \in \mathcal{M}_{\mathbb{K}}(n, p)$, M est équivalente à $J_{n,p,r}$, où r désigne le rang de M .

Démonstration.

Notons u l'application linéaire canoniquement associée à M .

u a pour rang r , donc il existe une base e de \mathbb{K}^p et une base f de \mathbb{K}^n telles que

$\text{mat}(u, e, f) = \begin{pmatrix} I_r & 0_{r, p-r} \\ 0_{n-r, r} & 0_{n-r, p-r} \end{pmatrix} = J_{n,p,r}$. Or, dans les bases canoniques de \mathbb{K}^p et de \mathbb{K}^n , la matrice de u est M , donc M et $J_{n,p,r}$ sont équivalentes. \square

Corollaire. Deux matrices sont équivalentes si et seulement si elles ont le même rang.

Démonstration.

Soient M et M' deux matrices de $\mathcal{M}_{\mathbb{K}}(n, p)$. Si M et M' ont le même rang noté r , elles sont toutes deux équivalentes à $J_{n,p,r}$, donc M est équivalente à M' .

Réciproquement, supposons que M est équivalente à M' . Il existe $P \in GL_p(\mathbb{K})$ et $Q \in GL_n(\mathbb{K})$ telles que $M' = Q^{-1}MP$, donc, P étant inversible, $\text{rg}(M') = \text{rg}(Q^{-1}M)$, puis $\text{rg}(M') = \text{rg}(M)$. \square

5.4.2 Propriétés du rang d'une matrice

Propriété. Pour toute matrice $M \in \mathcal{M}_{\mathbb{K}}(n, p)$, $\text{rg}(M) = \text{rg}(^tM)$.

On en déduit que le rang de M est aussi le rang de la famille de ses vecteurs lignes.

Démonstration.

Si l'on pose $r = \text{rg}(M)$, on a vu qu'il existe des matrices Q et P inversibles telles que $M = QJ_{n,p,r}P$, donc $^tM = ^tPJ_{p,n,r}^tQ$. Mais tP et tQ sont inversibles, donc $\text{rg}(^tM) = \text{rg}(J_{p,n,r}) = r$. \square

Propriété. Si l'on effectue une série de manipulations élémentaires sur une matrice, on ne modifie pas le rang de cette matrice.

Remarque. Pour déterminer le rang d'une matrice, une méthode consiste donc à transformer cette matrice en une matrice dont on connaît le rang par une succession d'opérations élémentaires portant sur les lignes ou sur les colonnes. On peut en particulier utiliser l'algorithme du pivot.

En particulier :

Propriété. Le rang d'une matrice est égal au nombre d'étapes dans la méthode du pivot global.

Démonstration.

Soit $M \in \mathcal{M}_{\mathbb{K}}(n, p)$. Par la méthode du pivot total, M est équivalente à une matrice de la forme $\begin{pmatrix} T & 0_{k,p-k} \\ 0_{n-k,k} & 0_{n-k,p-k} \end{pmatrix}$, où T est une matrice triangulaire supérieure à coefficients diagonaux non nuls (ce sont les pivots successifs) de taille k : k correspond bien au nombre d'étapes lorsqu'on applique l'algorithme du pivot total à M .

On sait que T est une matrice inversible, donc $\text{rg}(T) = k$.

Le rang d'une matrice étant égal au rang de la famille de ses colonnes, on a encore $k = \text{rg}((T \ 0_{k,p-k}))$. Enfin, Le rang d'une matrice étant égal au rang de la famille de ses lignes, on a encore $k = \text{rg}(\begin{pmatrix} T & 0_{k,p-k} \\ 0_{n-k,k} & 0_{n-k,p-k} \end{pmatrix}) = \text{rg}(M)$. \square

Propriété. Soit $M \in \mathcal{M}_{\mathbb{K}}(n, p)$.

Si P est une matrice extraite de M , alors $\text{rg}(P) \leq \text{rg}(M)$.

Démonstration.

Il existe $I \subset \mathbb{N}_n$ et $J \subset \mathbb{N}_p$ tels que $P = (M_{i,j})_{\substack{i \in I \\ j \in J}}$. Posons $Q = (M_{i,j})_{\substack{i \in \mathbb{N}_n \\ j \in J}}$.

Alors, $\text{rg}(Q) = \dim(\text{Vect}(M_j)_{j \in J}) \leq \dim(\text{Vect}(M_j)_{j \in \mathbb{N}_p}) = \text{rg}(M)$.

De même, en raisonnant sur les lignes, on montre que $\text{rg}(P) \leq \text{rg}(Q)$. \square

Exemple. Posons $M = \begin{pmatrix} 1 & -2 & 3 & 2 \\ 0 & 2 & 1 & 3 \\ 0 & 0 & -1 & -1 \\ 4 & 5 & 6 & 15 \end{pmatrix}$. M n'est pas de rang 4 car sa dernière colonne est la somme des précédentes, mais M est de rang supérieur ou égal à 3, car la matrice extraite de M en ôtant la dernière ligne et la dernière colonne est inversible (car triangulaire supérieure à coefficients diagonaux non nuls). Ainsi $\text{rg}(M) = 3$.

Propriété. Soit $A \in \mathcal{M}_{\mathbb{K}}(n, p)$ une matrice non nulle.

$\text{rg}(A)$ est égal à la taille maximale des matrices inversibles extraites de A .

Démonstration.

Posons \mathcal{A} l'ensemble des entiers k tels qu'il existe une matrice inversible extraite de A de taille k .

◇ Soit $k \in \mathcal{A}$. Il existe une matrice inversible B extraite de A de taille k . Alors d'après la propriété précédente, $k = \text{rg}(B) \leq \text{rg}(A)$. Ainsi \mathcal{A} est majorée par $\text{rg}(A)$. De plus $1 \in \mathcal{A}$ car A est non nulle. Ainsi \mathcal{A} est non vide et majoré, donc il possède un maximum noté s et $s \leq \text{rg}(A)$.

◇ Posons $r = \text{rg}(A)$. $\text{Vect}(A_1, \dots, A_p)$ est de dimension r , donc d'après le théorème de la base extraite, il existe $J \subset \mathbb{N}_p$ de cardinal r tel que $(A_j)_{1 \leq j \leq r}$ est une base de $\text{Vect}(A_1, \dots, A_p) = \text{Im}(A)$. Posons $Q = (A_{i,j})_{\substack{i \in \mathbb{N}_n \\ j \in J}}$. Q est une matrice extraite de A et $\text{rg}(Q) = \text{rg}((A_j)_{j \in J}) = r$.

Ainsi r est la dimension de l'espace F engendré par les lignes de Q . On peut donc à nouveau extraire des lignes de Q une base de F : il existe $I \subset \mathbb{N}_n$ de cardinal r tel que $P = (A_{i,j})_{\substack{i \in I \\ j \in J}}$ est de rang r .

P est une matrice carrée de taille r et de rang r , donc elle est inversible, ce qui conclut. \square

5.4.3 Matrices semblables

Définition. Deux matrices carrées M et M' dans $\mathcal{M}_n(\mathbb{K})$ sont **semblables** si et seulement s'il existe $P \in GL_n(\mathbb{K})$ tel que $M' = PMP^{-1}$. On définit ainsi une seconde relation d'équivalence sur $\mathcal{M}_n(\mathbb{K})$, appelée relation de similitude.

Remarque. Si deux matrices sont semblables, elles sont équivalentes, mais la réciproque est fausse. En particulier, l'interprétation de la relation de similitude en termes d'opérations élémentaires est délicate et non productive (jusqu'à preuve du contraire).

Remarque. Les classes d'équivalence de la relation d'équivalence sont entièrement déterminées par un entier, égal au rang des représentants de cette classe. Au contraire, les classes d'équivalence de la relation de similitude sont nettement plus complexes. Leur étude constitue la théorie de la réduction que vous étudierez en détail en seconde année.

Exemple. Une matrice M de $\mathcal{M}_n(\mathbb{K})$ est diagonalisable (resp : trigonalisable) si et seulement si elle est semblable à une matrice diagonale (resp : triangulaire supérieure).

Propriété. Soit E un \mathbb{K} -espace vectoriel de dimension $n > 0$, muni d'une base e , et soit $u \in L(E)$. On note $M = \text{mat}(u, e)$. Soit $M' \in \mathcal{M}_n(\mathbb{K})$.

M' est semblable à M si et seulement s'il existe une base e' telle que $M' = \text{mat}(u, e')$. En résumé, deux matrices sont semblables si et seulement si elles représentent un même endomorphisme dans des bases différentes, en imposant de prendre une même base au départ et à l'arrivée.

Démonstration.

Il suffit d'adapter la démonstration de la propriété analogue portant sur la relation "être équivalente à". \square

Propriété. Soient $(M, M') \in \mathcal{M}_n(\mathbb{K})^2$ et $P \in GL_n(\mathbb{K})$ tels que $M' = PMP^{-1}$.

Alors, pour tout $n \in \mathbb{N}$, $M'^n = PM^nP^{-1}$ et pour tout $Q \in \mathbb{K}[X]$, $Q(M') = PQ(M)P^{-1}$. Si M' et M sont inversibles, pour tout $n \in \mathbb{Z}$, $M'^n = PM^nP^{-1}$.

Démonstration.

• Pour $n = 0$, $PM^0P^{-1} = PP^{-1} = I_n = M'^0$.

Pour $n \geq 0$, supposons que $M'^n = PM^nP^{-1}$.

Alors $M'^{n+1} = M'^n M' = PM^nP^{-1} PMP^{-1} = PM^{n+1}P^{-1}$.

On a donc montré par récurrence que, pour tout $n \in \mathbb{N}$, $M'^n = PM^nP^{-1}$.

• Soit $Q \in \mathbb{K}[X]$. Il existe $(a_n) \in \mathbb{K}^{(\mathbb{N})}$ tel que $Q = \sum_{n \in \mathbb{N}} a_n X^n$.

$$PQ(M)P^{-1} = P\left(\sum_{n \in \mathbb{N}} a_n M^n\right)P^{-1} = \sum_{n \in \mathbb{N}} a_n PM^nP^{-1} = \sum_{n \in \mathbb{N}} a_n M'^n = Q(M').$$

• Supposons que M et M' sont inversibles.

Soit $n \in \mathbb{N}$. $M'^{-n} = (M'^n)^{-1} = (PM^nP^{-1})^{-1} = (P^{-1})^{-1}M^{-n}P^{-1} = PM^{-n}P^{-1}$.

Ainsi, pour tout $n \in \mathbb{Z}$, $M'^n = PM^nP^{-1}$. \square

6 Les hyperplans

Dans tout ce chapitre, on fixe un \mathbb{K} -espace vectoriel E , où \mathbb{K} est un corps.

6.1 En dimension quelconque

Définition. Soit H un sous-espace vectoriel de E . On dit que H est un hyperplan si et seulement si il existe une droite vectorielle D telle que $H \oplus D = E$.

Ainsi, les hyperplans sont les supplémentaires des droites vectorielles.

Propriété. Soit H un hyperplan et D une droite non incluse dans H . Alors $H \oplus D = E$.

Démonstration.

Il existe $x_1 \in E \setminus H$ tel que $D = \mathbb{K}x_1$.

◊ Si $x \in H \cap D$, il existe $\lambda \in \mathbb{K}$ tel que $x = \lambda x_1$. Si $\lambda \neq 0$, alors $x_1 = \frac{1}{\lambda}x \in H$ ce qui est faux, donc $\lambda = 0$ puis $x = 0$, ce qui prouve que la somme $H + D$ est directe.

◇ Soit $x \in E$. H étant un hyperplan, il existe une droite vectorielle D_0 telle que $E = H \oplus D_0$. Il existe $x_0 \in E \setminus H$ tel que $D_0 = \mathbb{K}x_0$.

Ainsi, il existe $h \in H$ et $\lambda \in \mathbb{K}$ tel que $x = h + \lambda x_0$.

De plus $x_1 \in E = H \oplus \mathbb{K}x_0$, donc il existe également $\lambda_1 \in \mathbb{K}$ et $h_1 \in H$ tels que $x_1 = h_1 + \lambda_1 x_0$. $\lambda_1 \neq 0$ car $x_1 \notin H$, donc $x_0 = \frac{1}{\lambda_1}(x_1 - h_1)$.

Alors $x = (h - \frac{\lambda}{\lambda_1} h_1) + \frac{\lambda}{\lambda_1} x_1 \in H \oplus \mathbb{K}x_1$, ce qui prouve que $E = H \oplus D$. □

Propriété. Soit H une partie de E . H est un hyperplan de E si et seulement si il est le noyau d'une forme linéaire non nulle.

De plus, si $H = \text{Ker}(\varphi) = \text{Ker}(\psi)$, alors φ et ψ sont colinéaires. Ainsi H est le noyau d'une forme linéaire non nulle, unique à un coefficient multiplicatif près.

Démonstration.

• Supposons que $H = \text{Ker}(\varphi)$, où $\varphi \in L(E, \mathbb{K}) \setminus \{0\}$.

Il existe $x_0 \in E$ tel que $\varphi(x_0) \neq 0$. $x_0 \notin H$ donc $H \cap \mathbb{K}x_0 = \{0\}$ (raisonnement identique à celui de la démonstration précédente).

De plus, si $x \in E$, en posant $\lambda = \frac{\varphi(x)}{\varphi(x_0)}$, on a $\varphi(x) = \varphi(\lambda x_0)$, donc $x = (x - \lambda x_0) + \lambda x_0$ et $x - \lambda x_0 \in \text{Ker}(\varphi) = H$, ce qui prouve que $E = H \oplus \mathbb{K}x_0$, donc que H est un hyperplan.

• Réciproquement, supposons que H est un hyperplan de E .

Il existe donc $x \in E \setminus \{0\}$ tel que $E = H \oplus \mathbb{K}x$.

Notons φ l'unique forme linéaire telle que $\varphi|_H = 0$ et $\varphi(x) = 1$. φ est non nulle.

De plus, si $y = h + \lambda x \in E$, où $h \in H$ et $\lambda \in \mathbb{K}$,

$y \in \text{Ker}(\varphi) \iff \varphi(h) + \lambda\varphi(x) = 0 \iff \lambda = 0 \iff y = h \iff y \in H$, donc $\text{Ker}(\varphi) = H$.

Soit $\psi \in L(E, \mathbb{K})$ telle que $H = \text{Ker}(\psi)$. Posons $\lambda = \psi(x)$.

Soit $y = h + \alpha x \in E$, où $h \in H$ et $\alpha \in \mathbb{K}$. $\psi(y) = \alpha\lambda = \lambda\varphi(y)$, donc $\psi = \lambda\varphi$. □

Définition. Soient H un hyperplan de E et $\varphi \in L(E, \mathbb{K}) \setminus \{0\}$ tel que $H = \text{Ker}(\varphi)$. Alors $x \in H \iff [(E) : \varphi(x) = 0]$. On dit que (E) est **équation de H** .

Exemples. $\left\{ f \in C([0, 1], \mathbb{R}) / \int_0^1 f(t) dt = 0 \right\}$ est un hyperplan de $C([0, 1], \mathbb{R})$.

L'ensemble des matrices carrées de taille n et de trace nulle est un hyperplan de $\mathcal{M}_n(\mathbb{K})$.

6.2 En dimension finie

Notation. On suppose que E est un espace de dimension finie notée n , avec $n > 0$.

Si $e = (e_1, \dots, e_n)$ est une base de E , pour tout $i \in \{1, \dots, n\}$, on note e_i^* l'application qui associe à tout vecteur x de E sa $i^{\text{ème}}$ coordonnée dans la base e .

ATTENTION : e_i^* dépend non seulement de e_i mais également des autres e_j : si l'on change la valeur de e_1 dans la base e , on change a priori la valeur de e_2^* .

Propriété. Avec les notations précédentes, la famille $e^* = (e_i^*)_{1 \leq i \leq n}$ est une base de $L(E, \mathbb{K}) = E^*$, que l'on appelle la base duale de e .

Démonstration.

Soit $\Psi \in E^*$: Pour tout $x \in E$, $x = \sum_{i=1}^n e_i^*(x)e_i$, donc $\Psi(x) = \sum_{i=1}^n e_i^*(x)\Psi(e_i)$.

Ainsi, $\Psi = \sum_{i=1}^n \Psi(e_i)e_i^* \in \text{Vect}(e_1^*, \dots, e_n^*)$.

Ceci prouve que e^* est une famille génératrice de E^* . De plus

$\dim(L(E, \mathbb{K})) = \dim(E) \times \dim(\mathbb{K}) = n$, donc e^* est une base de E^* . \square

Remarque.

Les hyperplans de E sont les sous-espaces vectoriels de E de dimension $n - 1$.

Définition. Soit $e = (e_1, \dots, e_n)$ une base de E et H un hyperplan de E .

Si $H = \text{Ker}(\psi)$, où $\psi \in E^*$, en notant $\psi = \sum_{i=1}^n \alpha_i e_i^*$, l'équation de l'hyperplan H

devient $x = \sum_{i=1}^n x_i e_i \in H \iff \sum_{i=1}^n \alpha_i x_i = 0 : (E)$. On dit que (E) est une équation cartésienne de l'hyperplan H , c'est-à-dire une condition nécessaire et suffisante portant sur les coordonnées de x dans la base e pour que x appartienne à H .

Démonstration.

$$x \in H \iff \psi(x) = 0, \text{ or } \psi(x) = \left[\sum_{i=1}^n \alpha_i e_i^* \right] \left(\sum_{j=1}^n x_j e_j \right) = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \alpha_i x_j e_i^*(e_j), \text{ or } e_i^*(e_j)$$

représente la i -ème coordonnée de e_j , donc $e_i^*(e_j) = \delta_{i,j}$, puis $\psi(x) = \sum_{i=1}^n \alpha_i x_i$, ce qu'il

fallait démontrer. \square

Exemple. Dans un plan vectoriel rapporté à une base (\vec{i}, \vec{j}) , une droite vectorielle D a une équation cartésienne de la forme :

$$\vec{v} = x\vec{i} + y\vec{j} \in D \iff ax + by = 0, \text{ où } (a, b) \in \mathbb{R}^2 \setminus \{0\}.$$

Exemple. Dans un espace vectoriel de dimension 3 rapporté à une base $(\vec{i}, \vec{j}, \vec{k})$, un plan vectoriel P a une équation cartésienne de la forme :

$$\vec{v} = x\vec{i} + y\vec{j} + z\vec{k} \in P \iff ax + by + cz = 0, \text{ où } (a, b, c) \in \mathbb{R}^3 \setminus \{0\}.$$

6.3 Les hyperplans affines

Notation. Soit \mathcal{E} un espace affine de direction E . On fixe un point $O \in \mathcal{E}$.

Définition. On appelle hyperplan affine de \mathcal{E} tout sous-espace affine de \mathcal{E} dont la direction est un hyperplan de E .

Propriété. Soit \mathcal{H} une partie de \mathcal{E} .

\mathcal{H} est un hyperplan affine de \mathcal{E} si et seulement si il existe $\varphi \in L(E, \mathbb{K}) \setminus \{0\}$ et $a \in \mathbb{K}$ tel que, pour tout $M \in \mathcal{E}$, $[M \in \mathcal{H} \iff \varphi(\vec{OM}) = a]$.

Dans ce cas, la condition $\varphi(\overrightarrow{OM}) = a$ est appelée une équation de \mathcal{H} .

De plus, la direction de \mathcal{H} est l'hyperplan $\text{Ker}(\varphi)$, d'équation $\varphi(x) = 0$ en l'inconnue $x \in E$.

Démonstration.

◇ Supposons que \mathcal{H} est un hyperplan affine de direction H .

On sait qu'il existe $\varphi \in L(E, \mathbb{K}) \setminus \{0\}$ tel que $H = \text{Ker}(\varphi)$.

Il existe $A \in \mathcal{H}$. Ainsi, pour tout $M \in \mathcal{E}$,

$$M \in \mathcal{H} \iff \overrightarrow{AM} \in H \iff \varphi(\overrightarrow{AM}) = 0 \iff \varphi(\overrightarrow{OM}) = \varphi(\overrightarrow{OA}).$$

◇ Réciproquement, supposons qu'il existe $\varphi \in L(E, \mathbb{K}) \setminus \{0\}$ et $a \in \mathbb{K}$ tel que, pour tout $M \in \mathcal{E}$, $[M \in \mathcal{H} \iff \varphi(\overrightarrow{OM}) = a]$.

$\varphi \neq 0$, donc $\text{rg}(\varphi) \geq 1$, mais $\text{Im}(\varphi) \subset \mathbb{K}$, donc $\text{Im}(\varphi) = \mathbb{K}$. Ceci prouve que φ est surjective. En particulier, il existe $\overrightarrow{x_0} \in E$ tel que $\varphi(\overrightarrow{x_0}) = a$. De plus il existe $A \in \mathcal{E}$ tel que $\overrightarrow{x_0} = \overrightarrow{OA}$.

Ainsi, $M \in \mathcal{H} \iff \varphi(\overrightarrow{OM}) = \varphi(\overrightarrow{OA}) \iff \varphi(\overrightarrow{AM}) = 0 \iff \overrightarrow{AM} \in \text{Ker}(\varphi)$, donc $\mathcal{H} = A + \text{Ker}(\varphi)$, ce qui prouve que \mathcal{H} est un hyperplan affine de direction $\text{Ker}(\varphi)$. □

Remarque. Dans le cas particulier où $\mathcal{E} = E$ et où $O = \overrightarrow{0}$, l'équation devient $\varphi(M) = a$, donc les hyperplans affines de E sont exactement les $\varphi^{-1}(\{a\})$, avec $\varphi \in L(E, \mathbb{K}) \setminus \{0\}$ et $a \in \mathbb{K}$.

Exemple. $\{M \in \mathcal{M}_n(\mathbb{K}) / \text{Tr}(M) = 2\}$ est un hyperplan affine de $\mathcal{M}_n(\mathbb{K})$.

Propriété. Supposons que E est de dimension finie égale à $n \in \mathbb{N}^*$ et que E est muni d'une base $e = (e_1, \dots, e_n)$, dont la base duale est notée $e^* = (e_1^*, \dots, e_n^*)$.

Soit \mathcal{H} un hyperplan affine de \mathcal{E} , dont une équation est $\Psi(\overrightarrow{OM}) = a$.

Notons $(\alpha_1, \dots, \alpha_n)$ les coordonnées de Ψ dans e^* . Si M a pour coordonnées (x_1, \dots, x_n)

dans le **repère affine** (O, e) , alors $M \in \mathcal{H} \iff \sum_{i=1}^n \alpha_i x_i = a$.

Cette dernière condition est donc la forme générale d'une équation cartésienne d'hyperplan affine en dimension n .

Pour $n = 2$, on obtient la forme générale d'une équation de droite, et pour $n = 3$, il s'agit de la forme générale d'une équation de plan.

6.4 Application aux systèmes linéaires

Notation. On fixe $(n, p) \in \mathbb{N}^{*2}$ et on considère un système linéaire de n équations à p inconnues de la forme :

$$(S) : \begin{cases} \alpha_{1,1}x_1 + \cdots + \alpha_{1,p}x_p = b_1 \\ \vdots \\ \alpha_{i,1}x_1 + \cdots + \alpha_{i,p}x_p = b_i \\ \vdots \\ \alpha_{n,1}x_1 + \cdots + \alpha_{n,p}x_p = b_n \end{cases},$$

où, pour tout $(i, j) \in \{1, \dots, n\} \times \{1, \dots, p\}$, $\alpha_{i,j} \in \mathbb{K}$, pour tout $i \in \{1, \dots, n\}$, $b_i \in \mathbb{K}$, les p inconnues étant x_1, \dots, x_p , éléments de \mathbb{K} .

Propriété. Notons $M = (\alpha_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in \mathcal{M}_{\mathbb{K}}(n, p)$ la matrice de (S) .

Ainsi $(S) \iff MX = B$, où $B = (b_i)_{1 \leq i \leq n} \in \mathbb{K}^n$. Si (S) est compatible, l'ensemble des solutions de (S) est un sous-espace affine de \mathbb{K}^p de dimension $p - r$, où r désigne le rang de M et dont la direction est $\text{Ker}(M)$.

Démonstration.

Supposons que (S) est compatible. Il existe $X_0 \in \mathbb{K}^p$ tel que $MX_0 = B$.

Ainsi, $(S) \iff MX = MX_0 \iff X - X_0 \in \text{Ker}(M) \iff X \in (X_0 + \text{Ker}(M))$. \square

Notation. Soient E et F des \mathbb{K} -espaces vectoriels de dimensions p et n munis de bases $e = (e_1, \dots, e_p)$ et $f = (f_1, \dots, f_n)$. On note u l'unique application linéaire de $L(E, F)$ telle que $\text{mat}(u, e, f) = M$, x le vecteur de E dont les coordonnées dans e sont X et b le vecteur de F dont les coordonnées dans f sont B . Alors $(S) \iff u(x) = b$.

Propriété. L'ensemble des solutions de (S) est soit vide, soit un sous-espace affine de E de direction $\text{Ker}(u)$.

Quatrième interprétation d'un système linéaire : *A l'aide de formes linéaires.*

Notons $e^* = (e_1^*, \dots, e_p^*)$ la base duale de e . Pour tout $i \in \{1, \dots, n\}$, posons

$l_i = \sum_{j=1}^p \alpha_{i,j} e_j^*$. Ainsi, (l_1, \dots, l_n) est une famille de formes linéaires telles que

$(S) \iff [\forall i \in \{1, \dots, n\} \ l_i(x) = b_i]$ et l'ensemble des solutions de (S) est $\bigcap_{i=1}^n l_i^{-1}(\{b_i\})$.

C'est donc une intersection d'hyperplans affines, si pour tout $i \in \mathbb{N}_n$, $l_i \neq 0$.

Propriété. Si E est un \mathbb{K} -espace vectoriel de dimension p , l'intersection de r hyperplans vectoriels de E est un sous-espace vectoriel de dimension supérieure à $p - r$.

Réciproquement tout sous-espace vectoriel de E de dimension $p - r$ où $r \geq 1$ est une intersection de r hyperplans de E , donc est caractérisé par un système de r équations linéaires.

Démonstration.

◇ Considérons r (avec $r \geq 1$) hyperplans de E , notés H_1, \dots, H_r .

Pour chaque $i \in \{1, \dots, r\}$, il existe une forme linéaire non nulle ℓ_i telle que

$H_i = \text{Ker}(\ell_i)$. Notons $F = \bigcap_{i=1}^r H_i$. Fixons une base $e = (e_1, \dots, e_p)$ de E et notons à nouveau $e^* = (e_1^*, \dots, e_p^*)$ la base duale de e .

Pour tout $i \in \{1, \dots, r\}$, il existe $(\alpha_{i,j})_{1 \leq j \leq p} \in \mathbb{K}^p$ telle que $\ell_i = \sum_{j=1}^p \alpha_{i,j} e_j^*$.

Ainsi, pour tout $x \in E$,

$$x = \sum_{j=1}^p x_j e_j \in F \iff \forall i \in \{1, \dots, r\} \ l_i(x) = 0 \iff \forall i \in \{1, \dots, r\} \ \sum_{j=1}^p \alpha_{i,j} x_j = 0.$$

Alors avec les notations précédentes, $F = \text{Ker}(u)$ et $\dim(F) = p - \text{rg}(u) \geq p - r$, car la matrice de u possède r lignes.

◇ Pour la réciproque, soit F un sous-espace vectoriel de E de dimension $p - r$ avec $r \geq 1$. Notons (f_{r+1}, \dots, f_p) une base de F , que l'on complète en une base $f = (f_1, \dots, f_p)$ de E . Notons $f^* = (f_1^*, \dots, f_p^*)$ la base duale de f .

Si $x = \sum_{j=1}^p x_j f_j \in E$, $x \in F \iff x_1 = \dots = x_r = 0 \iff \forall i \in \{1, \dots, r\}, f_i^*(x) = 0$.

Ainsi, F est caractérisé par un système de r équations linéaires.

De plus, en notant $H_i = \text{Ker}(f_i^*)$, $[x \in F \iff \forall i \in \{1, \dots, r\}, x \in H_i]$, donc F est l'intersection des r hyperplans H_1, \dots, H_r . □

Exemple. Dans un espace vectoriel de dimension 3 rapporté à une base $(\vec{i}, \vec{j}, \vec{k})$, une droite vectorielle D est caractérisée par un système de 2 équations cartésiennes de la forme :

$$\vec{v} = x\vec{i} + y\vec{j} + z\vec{k} \in D \iff \begin{cases} ax + by + cz = 0 \\ a'x + b'y + c'z = 0 \end{cases}, \text{ où } \left(\begin{pmatrix} a \\ b \\ c \end{pmatrix}, \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} \right) \text{ est une}$$

famille libre dans \mathbb{K}^3 .

Propriété. Soit \mathcal{E} un espace affine de direction E et \mathcal{F} un sous-espace affine de \mathcal{E} dirigé par $F \neq E$. On sait que F est l'intersection de r hyperplans vectoriels (où $r = \dim(E) - \dim(F)$) H_1, \dots, H_r , qui sont chacun le noyau d'une forme linéaire non nulle, notée $\varphi_1, \dots, \varphi_r$.

Il existe $A \in \mathcal{E}$ tel que $\mathcal{F} = A + F$, donc

$$M \in \mathcal{F} \iff \overrightarrow{AM} \in F \iff (\forall i \in \{1, \dots, r\} \varphi_i(\overrightarrow{OM}) = \varphi_i(\overrightarrow{OA})).$$

Cette dernière condition constitue un système d'équations de \mathcal{F} .

Ainsi, tout sous-espace affine de \mathcal{E} peut être caractérisé par un système d'équations linéaires.

Corollaire. Tout sous-espace affine strictement inclus dans \mathcal{E} est une intersection d'un nombre fini d'hyperplans affines.

7 Déterminants

7.1 Applications multilinéaires

Définition. Soient $p \in \mathbb{N}^*$ et (E_1, \dots, E_p) une famille de p \mathbb{K} -espaces vectoriels.

Soient F un \mathbb{K} -espace vectoriel et f une application de $E_1 \times \dots \times E_p$ dans F .

f est une **application p -linéaire** si et seulement si, pour tout $j \in \mathbb{N}_p$

et pour tout $(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_p) \in E_1 \times \dots \times E_{j-1} \times E_{j+1} \times \dots \times E_p$,

l'application $\begin{matrix} E_j & \longrightarrow & F \\ x_j & \longmapsto & f(a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_p) \end{matrix}$ est linéaire.

Définition. Une **application bilinéaire** est une application 2-linéaire.

Notation.

- $L_p(E_1, \dots, E_p; F)$ désigne l'ensemble des applications p -linéaires de $E_1 \times \dots \times E_p$ dans F . C'est un sous-espace vectoriel de $\mathcal{F}(E_1 \times \dots \times E_p, F)$.
- On note $L_p(E, F) = L_p(\underbrace{E, \dots, E}_{p \text{ fois}}; F)$.
- Enfin, on note $L_p(E) = L_p(E, \mathbb{K})$.
Les éléments de $L_p(E)$ sont appelés des **formes p -linéaires** sur E .

Exemple. L'application $(\mathbb{R}^2)^3 \rightarrow \mathbb{R}$
 $((x, x'), (y, y'), (z, z')) \mapsto xy'z' - 2x'yz' + 3x'yz$
est une forme trilinéaire sur \mathbb{R}^2 .

Notation. Pour la suite de ce paragraphe, on fixe $p \in \mathbb{N}^*$
et deux \mathbb{K} -espaces vectoriels E et F .

Propriété. Soit u_1, \dots, u_p p applications linéaires de E dans \mathbb{K} .

Alors l'application $u : E^p \rightarrow \mathbb{K}$
 $(x_1, \dots, x_p) \mapsto \prod_{i=1}^p u_i(x_i)$ est une forme p -linéaire.

Démonstration.

Soient $j \in \mathbb{N}_p$ et $(a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_p) \in E^{p-1}$.

Notons $v : E \rightarrow \mathbb{K}$
 $x \mapsto u(x_1, \dots, x_{j-1}, x, x_{j+1}, \dots, x_p)$.

Pour tout $x \in E$, $v(x) = \left(\prod_{\substack{1 \leq i \leq p \\ i \neq j}} u_i(x_i) \right) u_j(x)$, or u_j est linéaire, donc v est linéaire.

Ainsi, u est une forme p -linéaire. \square

Propriété. On suppose que E est de dimension finie $n \in \mathbb{N}^*$ et on note
 $e = (e_1, \dots, e_n)$ une base de E . Une application f de E^p dans F est p -linéaire si
et seulement s'il existe une famille $(\alpha_u)_{u \in \mathbb{N}_n^p}$ de vecteurs de F telle que, pour tout

$$(x_1, \dots, x_p) = \left(\sum_{i=1}^n a_{i,j} e_i \right)_{1 \leq j \leq p} \in E^p, \quad f(x_1, \dots, x_p) = \sum_{u=(i_1, \dots, i_p) \in \mathbb{N}_n^p} a_{i_1,1} \dots a_{i_p,p} \alpha_u.$$

Démonstration.

- Supposons que $f \in L_p(E, F)$. Soit $(x_1, \dots, x_p) = \left(\sum_{i=1}^n a_{i,j} e_i \right)_{1 \leq j \leq p} \in E^p$.

$$f(x_1, \dots, x_p) = f\left(\sum_{i=1}^n a_{i,1} e_i, x_2, \dots, x_p\right),$$

donc en utilisant la linéarité selon la première variable,

$$f(x_1, \dots, x_p) = \sum_{i=1}^n a_{i,1} f(e_i, x_2, \dots, x_p), \text{ puis}$$

$$f(x_1, \dots, x_p) = \sum_{i=1}^n a_{i,1} \sum_{j=1}^n a_{j,2} f(e_i, e_j, x_3, \dots, x_p).$$

Le développement de x_1 a nécessité l'utilisation d'une variable notée i , puis le développement de x_2 a nécessité l'utilisation d'une seconde variable notée j . Pour développer x_1, x_2, \dots, x_p , nous avons besoin de p variables. Nous allons les noter i_1, i_2, \dots, i_p , ce qui impose de renommer i et j en i_1 et i_2 . Ainsi,

$$f(x_1, \dots, x_p) = \sum_{i_1=1}^n \sum_{i_2=1}^n a_{i_1,1} a_{i_2,2} f(e_{i_1}, e_{i_2}, x_3, \dots, x_p). \text{ En poursuivant ce calcul, on obtient } f(x_1, \dots, x_p) = \sum_{u=(i_1, \dots, i_p) \in \mathbb{N}_n^p} a_{i_1,1} \dots a_{i_p,p} f(e_{i_1}, \dots, e_{i_p}).$$

• Réciproquement, $L_p(E, F)$ étant un \mathbb{K} -espace vectoriel, il suffit de montrer que pour

$$\text{tout } u = (i_1, \dots, i_p) \in \mathbb{N}_n^p, \text{ l'application } \begin{matrix} E^p & \longrightarrow & \mathbb{K} \\ (\sum_{i=1}^n a_{i,j} e_i)_{1 \leq j \leq p} & \longmapsto & a_{i_1,1} \dots a_{i_p,p} \end{matrix} \text{ est } p\text{-}$$

linéaire, mais il s'agit de l'application $(x_1, \dots, x_p) \longmapsto \prod_{j=1}^p e_{i_j}^*(x_j)$, or les applications

coordonnées e_j^* sont des formes linéaires, donc d'après une propriété précédente, cette application est bien une forme p -linéaire. \square

Définition. Soient $\sigma \in \mathcal{S}_p$ et $f \in L_p(E, F)$.

$$\text{On note } \sigma(f) : \begin{matrix} E^p & \longrightarrow & F \\ (x_1, \dots, x_p) & \longmapsto & f(x_{\sigma(1)}, \dots, x_{\sigma(p)}) \end{matrix}.$$

On vérifie que $\sigma(f)$ est une application p -linéaire de E dans F .

Propriété. L'application $\begin{matrix} \mathcal{S}_p \times L_p(E, F) & \longrightarrow & L_p(E, F) \\ (\sigma, f) & \longmapsto & \sigma(f) \end{matrix}$, est une opération du groupe (\mathcal{S}_p, \circ) sur l'ensemble $L_p(E, F)$.

Démonstration.

Soient $(\sigma, \sigma') \in \mathcal{S}_p^2$, $f \in L_p(E, F)$ et $(x_1, \dots, x_p) \in E^p$.

$$\sigma'(\sigma(f))(x_1, \dots, x_p) = \sigma(f)(x_{\sigma'(1)}, \dots, x_{\sigma'(p)}).$$

Notons $(y_1, \dots, y_p) = (x_{\sigma'(1)}, \dots, x_{\sigma'(p)})$. Ainsi, pour tout $i \in \mathbb{N}_p$, $y_i = x_{\sigma'(i)}$, donc

$$\sigma'(\sigma(f))(x_1, \dots, x_p) = f(y_{\sigma(1)}, \dots, y_{\sigma(p)}) = f(x_{\sigma'(\sigma(1))}, \dots, x_{\sigma'(\sigma(p))}), \text{ ce qui prouve que}$$

$$\forall (\sigma, \sigma') \in \mathcal{S}_p^2 \quad \forall f \in L_p(E, F) \quad \sigma'(\sigma(f)) = (\sigma'\sigma)(f).$$

De plus, $Id_{\mathbb{N}_p}(f) = f$, donc $(\sigma, f) \longmapsto \sigma(f)$ est bien une opération de groupe. \square

Définition. Soit $f \in L_p(E, F)$. f est une **application p -linéaire symétrique** si et seulement si pour tout $\sigma \in \mathcal{S}_p$, $\sigma(f) = f$.

f est une **application p -linéaire antisymétrique** si et seulement si pour tout $\sigma \in \mathcal{S}_p$, $\sigma(f) = \varepsilon(\sigma)f$, où $\varepsilon(\sigma)$ désigne la signature de la permutation σ .

Propriété. Soit $f \in L_p(E, F)$.

f est symétrique si et seulement si pour toute transposition τ de \mathcal{S}_p , $\tau(f) = f$.

f est antisymétrique si et seulement si pour toute transposition τ de \mathcal{S}_p , $\tau(f) = -f$.

Démonstration.

Supposons que pour toute transposition τ de \mathcal{S}_p , $\tau(f) = -f$.

Soit $n \in \mathbb{N}^*$. Notons $R(n)$ l'assertion suivante :

pour toute famille de n transpositions (τ_1, \dots, τ_n) , $(\tau_1 \circ \dots \circ \tau_n)(f) = (-1)^n f$.

Par récurrence sur n , il est simple de montrer que $R(n)$ est vraie pour tout $n \in \mathbb{N}^*$.

Soit $\sigma \in \mathcal{S}_p$. L'ensemble des transpositions engendre \mathcal{S}_p , donc il existe un nombre fini de transpositions, notées τ_1, \dots, τ_n , telles que $\sigma = \tau_1 \circ \dots \circ \tau_n$.

D'après $R(n)$, $\sigma(f) = (-1)^n f = \varepsilon(\sigma)f$.

Ainsi, f est une application p -linéaire antisymétrique.

La réciproque est simple à prouver.

Pour démontrer que f est symétrique si et seulement si pour toute transposition τ de \mathcal{S}_p , $\tau(f) = f$, il suffit d'adapter la démonstration précédente. \square

Exemples.

- L'application $\begin{array}{ccc} \mathbb{K}^p & \longrightarrow & \mathbb{K} \\ (x_1, \dots, x_p) & \longmapsto & x_1 \times \dots \times x_p \end{array}$ est p -linéaire symétrique.
- L'application $\begin{array}{ccc} (\mathbb{K}^2)^2 & \longrightarrow & \mathbb{K} \\ \left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right) & \longmapsto & ad - bc = \det\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right) \end{array}$ est une forme bilinéaire antisymétrique.

Définition. Soit $f \in L_p(E, F)$. f est **une application p -linéaire alternée** si et seulement si elle annule tout p -uplet de vecteurs de E contenant au moins deux vecteurs égaux.

Propriété. Soit $f \in L_p(E, F)$.

Si f est alternée, alors elle est antisymétrique.

Lorsque $\text{car}(\mathbb{K}) \neq 2$, alternée \iff antisymétrique.

Démonstration.

Pour simplifier, on se limite au cas où $p = 2$, mais le principe de la démonstration est valable dans le cas général : il suffit d'adapter au prix de notations plus lourdes.

◇ Supposons que f est alternée. Soit $x, y \in E$.

$$0 = f(x + y, x + y) = f(x, x) + f(y, y) + f(x, y) + f(y, x) = f(x, y) + f(y, x),$$

donc $f(x, y) = -f(y, x)$.

◇ On suppose que f est antisymétrique et que $\text{car}(\mathbb{K}) \neq 2$. Soit $x \in E$.

$$f(x, x) = -f(x, x), \text{ donc } (2.1_{\mathbb{K}})f(x, x) = 0, \text{ or } 2.1_{\mathbb{K}} \neq 0, \text{ donc } f(x, x) = 0. \square$$

Remarque. Lorsque \mathbb{K} est de caractéristique 2, l'équivalence n'est plus vraie. Par exemple l'application $f : (\mathbb{Z}/2\mathbb{Z})^2 \longrightarrow \mathbb{Z}/2\mathbb{Z}$ définie par $f(x, y) = xy$ est dans $L_2(\mathbb{Z}/2\mathbb{Z})$, mais elle n'est pas alternée car $f(1, 1) = 1 \neq 0$. Pourtant elle est symétrique, donc antisymétrique, car dans $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$, $1_{\mathbb{K}} = -1_{\mathbb{K}}$.

Propriété. $f \in L_p(E, F)$ est alternée si et seulement si pour tout $(x_1, \dots, x_p) \in E^p$, $f(x_1, \dots, x_p)$ ne varie pas lorsque l'on ajoute à l'un des x_i une combinaison linéaire des autres x_j .

Plus formellement, f est alternée si et seulement si

$$\begin{aligned} \forall (x_1, \dots, x_p) \in E_1 \times \dots \times E_p \quad \forall i \in \mathbb{N}_p \quad \forall (\alpha_j)_{j \in \mathbb{N}_p \setminus \{i\}} \in \mathbb{K}^{p-1} \\ f(x_1, \dots, x_p) = f(x_1, \dots, x_{i-1}, x_i + \sum_{j \neq i} \alpha_j x_j, x_{i+1}, \dots, x_p) \end{aligned} .$$

Propriété. $f \in L_p(E, F)$ est alternée si et seulement si l'image par f de toute famille liée de vecteurs est nulle.

Démonstration.

◇ Supposons que f est alternée.

Si (x_1, \dots, x_p) est lié, il existe $(\alpha_1, \dots, \alpha_p) \in \mathbb{K}^p$ tel que $\sum_{i=1}^p \alpha_i x_i = 0$, avec

$(\alpha_1, \dots, \alpha_p) \neq 0$. Il existe $j \in \{1, \dots, p\}$ tel que $\alpha_j \neq 0$. Ainsi, $x_j = \frac{-1}{\alpha_j} \sum_{\substack{1 \leq i \leq p \\ j \neq i}} \alpha_i x_i$, donc

$$f(x_1, \dots, x_p) = f(x_1, \dots, x_{j-1}, 0, x_{j+1}, \dots, x_p) = 0.$$

◇ Réciproquement, supposons que l'image par f de toute famille liée de vecteurs est nulle. Tout p -uplet de vecteurs de E contenant au moins deux vecteurs égaux est lié, donc son image par f est nulle. Ceci prouve que f est alternée. □

Corollaire. Si E est de dimension $n \in \mathbb{N}^*$ et si $p > n$, toute forme p -linéaire alternée sur E est nulle.

7.2 Les trois notions de déterminants

Au sein de ce paragraphe, E désignera un \mathbb{K} -espace vectoriel de dimension finie égale à n , avec $n > 0$.

7.2.1 Volume

Supposons temporairement que $\mathbb{K} = \mathbb{R}$.

Enoncé du problème :

Pour tout $x = (x_1, \dots, x_n) \in E^n$, on note H_x l'hyperparallélépipède

$H_x = \left\{ \sum_{i=1}^n t_i x_i \mid t_1, \dots, t_n \in [0, 1] \right\}$: c'est l'unique hyperparallélépipède de E dont les côtés issus de l'origine sont x_1, \dots, x_n .

On souhaite définir une fonction $\text{vol} : E^n \rightarrow \mathbb{R}$ telle que, pour tout

$x = (x_1, \dots, x_n) \in E^n$, $|\text{vol}(x)|$ soit égal au volume de H_x , en s'appuyant sur l'idée intuitive que l'on a de la notion de volume d'une partie de E (vu comme un espace affine). On souhaite de plus que le signe de $\text{vol}(x)$ corresponde à l'orientation du n -uplet x , en s'appuyant également sur une idée intuitive de la notion d'orientation. On dira que $\text{vol}(x)$ est le volume algébrique de H_x et par opposition, que $|\text{vol}(x)|$ est son volume absolu.

Cas d'un hyperparallélépipède plat :

Si x est lié, H_x est "plat", donc on impose $\text{vol}(x) = 0$ dès que x est une famille liée de n vecteurs.

Homogénéité de la fonction vol :

Supposons maintenant que x est libre. Ainsi, x est une base de E .

Si l'on remplace l'un des x_i par λx_i , où $\lambda \in \mathbb{R}$, le volume de H_x doit être multiplié par $\pm \lambda$, car le volume absolu de H_x est intuitivement proportionnel à la longueur de chacun de ses côtés. Plutôt que de parler d'intuition, on peut dire que cette propriété est un axiome que doit vérifier toute notion de volume. De même, nous conviendrons que l'orientation d'un n -uplet de vecteurs de E est positive ou négative (il n'y a que deux valeurs possibles pour l'orientation) et que le fait de remplacer l'un des vecteurs x_i par λx_i avec $\lambda < 0$ change l'orientation du n -uplet. Alors la fonction vol doit vérifier la propriété suivante : pour tout $x = (x_1, \dots, x_n) \in E^n$, pour tout $i \in \mathbb{N}_n$, pour tout $\lambda \in \mathbb{R}$, $\text{vol}(x_1, \dots, x_{i-1}, \lambda x_i, x_{i+1}, \dots, x_n) = \lambda \text{vol}(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)$.

Lorsque x est lié, cette propriété est vraie car elle se réduit à $0 = 0$.

 n -linéarité de vol :

Fixons $i, j \in \mathbb{N}_n$ ainsi que $(x_k)_{k \in \mathbb{N}_n \setminus \{i, j\}}$ une famille de $n - 2$ vecteurs (si $n = 1$, E est de dimension 1 et vol est clairement linéaire d'après le point précédent, donc on peut supposer que $n \geq 2$). Pour tout $x_i, x_j \in E^2$, posons $f(x_i, x_j) = \text{vol}(x_k)_{k \in \mathbb{N}_n}$.

1°) Soit $a, b \in E^2$. Commençons par établir que $f(a + b, b) = f(a, b)$.

Notons également $a = x_i$ et $b = x_j$.

Posons $G = \left\{ \sum_{\substack{1 \leq k \leq n \\ k \neq \{i, j\}}} t_k x_k \mid \forall k, t_k \in [0, 1] \right\}$.

Alors $H_x = \{xa + yb \mid (x, y) \in [0, 1]^2\} + G$. Il s'agit de montrer que H_x a le même volume absolu que $K = \{x(a + b) + yb \mid (x, y) \in [0, 1]^2\} + G$ et que x a la même orientation que le n -uplet x' obtenu à partir de x en remplaçant x_i par $x_i + x_j$.

Figure

Conformément à la figure, posons $A = \{xa + yb \mid x, y \in [0, 1] \wedge y \leq x\} + G$, $B = \{xa + yb \mid x, y \in [0, 1] \wedge y \geq x\} + G$ et notons C l'image de A par la translation de vecteur b , c'est-à-dire $C = b + A$.

Selon la figure, $H_x = A \cup B$ et $\text{vol}(A \cap B) = 0$, $K = B \cup C$ et $\text{vol}(B \cap C) = 0$.

On peut le démontrer formellement :

1.1] $A \cap B = \{x(a + b) \mid x \in [0, 1]\} + G$: c'est l'hyperparallélépipède associé à x' déduit de x en remplaçant x_i et x_j par $\frac{x_i + x_j}{2}$. x' est lié, donc $\text{vol}(A \cap B) = 0$.

1.2] $C = \{xa + (1 + y)b \mid x, y \in [0, 1] \wedge y \leq x\} + G$,
donc $B \cap C = \{xa + b \mid x \in [0, 1]\} = b + \{xa \mid x \in [0, 1]\} + G$.

Il est raisonnable d'imposer à la notion de volume d'être invariante par translation. Alors $|\text{vol}(B \cap C)| = |\text{vol}(H_{x''})|$ où x'' se déduit de x en remplaçant x_i et x_j par $\frac{x_i}{2}$. Ainsi, on a bien $\text{vol}(B \cap C) = 0$.

1.3] Si $x, y \in [0, 1]$, on a bien sûr, $x \leq y$ ou bien $y \leq x$, donc $H_x = A \cup B$.

$$\begin{aligned} K &= \{xa + (x+y)b \mid (x, y) \in [0, 1]^2 \wedge x+y \leq 1\} + G \\ &\cup \{xa + (x+y)b \mid (x, y) \in [0, 1]^2 \wedge x+y \geq 1\} + G \\ &= \{xa + zb \mid x, z \in [0, 1] \wedge x \leq z\} + G \\ &\cup [b + \{xa + (x-z)b \mid x, z \in [0, 1] \wedge x-z \geq 0\} + G] \\ &= B \cup (b+A) = B \cup C. \end{aligned}$$

Ainsi, on a montré qu'on peut passer de l'hyperparallélépipède H_x à l'hyperparallélépipède K en découpant H_x en deux morceaux disjoints (au sens que l'intersection de ces deux morceaux est de volume nul), en translatant l'un des deux morceaux puis en effectuant à nouveau la réunion disjointe des deux morceaux. Il est raisonnable d'imposer aux notions de volume absolu et d'orientation d'être invariantes par cette opération. Alors H_x et K ont le même volume algébrique, ce qui montre que $f(a+b, b) = f(a, b)$, pour tout $a, b \in E^2$.

2°) Soit maintenant $\lambda \in \mathbb{R}^*$. Alors $f(a + \lambda b, b) = \frac{1}{\lambda} f(a + \lambda b, \lambda b) = \frac{1}{\lambda} f(a, \lambda b) = f(a, b)$. Ainsi $\text{vol}(x)$ n'est pas modifié si l'on ajoute à l'un des x_i le vecteur λx_j , où $\lambda \in \mathbb{R}$ et $j \in \mathbb{N}_n \setminus \{i\}$. Donc, pour tout $x \in E^n$, $\text{vol}(x)$ n'est pas modifié si l'on ajoute à l'un des x_i une combinaison linéaire des autres x_j .

3°) On reprend les notations du 1°).

a) pour tout $\lambda \in \mathbb{R}$, $f(a + \lambda a, b) = (1 + \lambda)f(a, b) = f(a, b) + f(\lambda a, b)$, donc lorsque c est colinéaire à a , $f(a + c, b) = f(a, b) + f(c, b)$ (c'est évident lorsque a est nul).

b) Si maintenant c est quelconque dans E , lorsque x est une base de E , on peut écrire $c = \lambda a + d$, où d est une combinaison linéaire des autres vecteurs x_j .

Alors d'après 2°), $f(a + c, b) = f(a + \lambda a, b) = f(a, b) + f(\lambda a, b) = f(a, b) + f(c, b)$.

c) Lorsque $\{b\} \cup \{x_k \mid k \in \mathbb{N}_n \setminus \{i, j\}\}$ est liée, alors $f(a + c, b) = 0 = f(a, b) + f(c, b)$.

d) Il reste à étudier le cas où x n'est pas une base mais où les $n-1$ vecteurs $\{b\} \cup \{x_k \mid k \in \mathbb{N}_n \setminus \{i, j\}\}$ sont libres. Alors $a = x_i$ est une combinaison linéaire des autres x_j , donc $f(a + c, b) = f(c, b) = f(a, b) + f(c, b)$.

4°) Ainsi, pour tout $a, b, c \in E$, $f(a + c, b) = f(a, b) + f(c, b)$, puis d'après la propriété d'homogénéité de vol , pour tout $a, b, c \in E$ et $\lambda \in \mathbb{R}$, $f(a + \lambda c, b) = f(a, b) + \lambda f(c, b)$. Ceci prouve que vol est bien n -linéaire. De plus elle est alternée, car elle s'annule sur tout n -uplet lié de vecteurs de E .

Conclusion :

Si vol est une application de E^n dans \mathbb{R} telle que, pour tout $x \in E^n$, $|\text{vol}(x)|$ représente le volume de H_x et le signe de $\text{vol}(x)$ représente l'orientation du n -uplet x , alors en imposant des contraintes raisonnables aux notions de volume absolu et d'orientation, l'application vol est nécessairement une forme n -linéaire alternée.

En particulier, il suffit d'appeler **orientation** toute application \mathcal{O} définie sur l'ensemble \mathcal{B} des bases de E à valeurs dans $\{1, -1\}$ telle que, pour tout $x \in E^n$:

- si l'on change l'un des vecteurs de x par son opposé, alors $\mathcal{O}(x)$ est remplacé par son opposé.

- On ne change pas $\mathcal{O}(x)$ si l'on multiplie l'un des vecteurs de x par un réel strictement positif, ou bien si l'on ajoute à l'un des vecteurs de x un autre vecteur de x .

7.2.2 Déterminant d'un système de n vecteurs

Notation. Notons $A_n(E)$ l'ensemble des formes n -linéaires alternées.

On choisit une base e de E . Soit $f : E^n \rightarrow \mathbb{K}$ une forme n -linéaire alternée. Alors, d'après un calcul présenté page 82,

pour tout $x = (x_1, \dots, x_n) = (\sum_{i=1}^n p_{i,j} e_i)_{1 \leq j \leq n} \in E^n$,

$$f(x_1, \dots, x_n) = \sum_{u=(i_1, \dots, i_n) \in \mathbb{N}_n^n} p_{i_1,1} \dots p_{i_n,n} f(e_{i_1}, \dots, e_{i_n}), \text{ ou, avec d'autres notations,}$$

$$f(x_1, \dots, x_n) = \sum_{u \in \mathcal{F}(\mathbb{N}_n, \mathbb{N}_n)} p_{u(1),1} \dots p_{u(n),n} f(e_{u(1)}, \dots, e_{u(n)}).$$

Si u est une injection de \mathbb{N}_n dans \mathbb{N}_n , alors on sait que u est une bijection, donc lorsque $u \in \mathcal{F}(\mathbb{N}_n, \mathbb{N}_n) \setminus \mathcal{S}_n$, u n'est pas injective : il existe $(i, j) \in \mathbb{N}_n^2$ avec $i \neq j$ tel que $u(i) = u(j)$. Alors, $e_{u(i)} = e_{u(j)}$, or f est alternée, donc $f(e_{u(1)}, \dots, e_{u(n)}) = 0$.

Ainsi, $f(x_1, \dots, x_n) = \sum_{u \in \mathcal{S}_n} p_{u(1),1} \dots p_{u(n),n} f(e_{u(1)}, \dots, e_{u(n)})$. De plus, f étant alternée,

elle est antisymétrique, donc $f(x_1, \dots, x_n) = \sum_{u \in \mathcal{S}_n} p_{u(1),1} \dots p_{u(n),n} \varepsilon(u) f(e_1, \dots, e_n)$.

$$\det_e : E^n \longrightarrow \mathbb{K}$$

Posons, $(x_1, \dots, x_n) = (\sum_{i=1}^n p_{i,j} e_i)_{1 \leq j \leq n} \mapsto \sum_{\sigma \in \mathcal{S}_n} p_{\sigma(1),1} \dots p_{\sigma(n),n} \varepsilon(\sigma)$.

On a montré que, pour tout $f \in A_n(E)$, (1) : $f = f(e_1, \dots, e_n) \det_e$.

Définition. Soit $e = (e_1, \dots, e_n)$ une base de E et $x = (x_1, \dots, x_n)$ un système de n vecteurs de E . On pose $P = \text{mat}_e(x)$, de sorte que $P_{i,j}$ désigne la i -ème coordonnée dans la base e de x_j . On appelle **déterminant du système de vecteurs** x dans la base e la quantité suivante :

$$\det_e(x_1, \dots, x_n) \triangleq \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n P_{\sigma(j),j}.$$

Exemple. Prenons $E = \mathbb{K}^2$ et $x = (\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix})$. Alors $\det_e(x) = ad - bc$. Ainsi, cette définition est cohérente avec nos précédentes définitions de déterminants.

Théorème. Soit e une base de E .

$$\boxed{\text{Si } f \text{ est une forme } n\text{-linéaire alternée sur } E, \text{ alors } f = f(e) \det_e.}$$

Démonstration.

Il s'agit de la formule (1). \square

Propriété. Avec les notations précédentes, on a aussi

$$\det_e(x_1, \dots, x_n) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n P_{j, \sigma(j)} = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n e_j^*(x_{\sigma(j)}).$$

Démonstration.

Posons $x = (x_1, \dots, x_n)$.

$$\det_e(x) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n P_{\sigma(j), \sigma^{-1}(\sigma(j))}, \text{ donc en posant } k = \sigma(j) \text{ dans le produit,}$$

$$\det_e(x) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{k=1}^n P_{k, \sigma^{-1}(k)}. \text{ De plus, l'application } \sigma \mapsto \sigma^{-1} \text{ étant une bijection}$$

de \mathcal{S}_n dans lui-même, on peut poser $s = \sigma^{-1}$. Ainsi, $\det_e(x) = \sum_{s \in \mathcal{S}_n} \varepsilon(s^{-1}) \prod_{j=1}^n P_{j, s(j)}$.

Mais ε est un morphisme de groupes à valeurs dans $\{1, -1\}$, donc pour tout $s \in \mathcal{S}_n$, $\varepsilon(s^{-1}) = \varepsilon(s)^{-1} = \varepsilon(s)$. Ainsi, $\det_e(x) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n P_{j, \sigma(j)}$. \square

Formule de Sarrus : En notant c la base canonique de \mathbb{K}^3 ,

$$\det_c \left(\begin{pmatrix} p_{1,1} \\ p_{2,1} \\ p_{3,1} \end{pmatrix}, \begin{pmatrix} p_{1,2} \\ p_{2,2} \\ p_{3,2} \end{pmatrix}, \begin{pmatrix} p_{1,3} \\ p_{2,3} \\ p_{3,3} \end{pmatrix} \right) = p_{1,1}p_{2,2}p_{3,3} + p_{2,1}p_{3,2}p_{1,3} + p_{3,1}p_{1,2}p_{2,3} \\ - p_{1,3}p_{2,2}p_{3,1} - p_{2,3}p_{3,2}p_{1,1} - p_{3,3}p_{1,2}p_{2,1}.$$

Démonstration.

$\mathcal{S}_3 = \{Id_{\mathbb{N}_3}, (1 \ 3 \ 2), (1 \ 2 \ 3), (1 \ 3), (2 \ 3), (1 \ 2)\}$. \square

Propriété. \det_e est une forme n -linéaire alternée.

Démonstration.

\diamond Soit $\sigma \in \mathcal{S}_n$. Posons f_σ l'application de E^n dans \mathbb{K} définie par

$$f_\sigma(x_1, \dots, x_n) = \prod_{j=1}^n e_j^*(x_{\sigma(j)}) = \prod_{j=1}^n e_{\sigma^{-1}(j)}^*(x_j). \text{ D'après une propriété du paragraphe}$$

précédent, f_σ est une forme n -linéaire, donc $\det_e = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) f_\sigma$ est aussi une forme

n -linéaire.

\diamond Soit $x = (x_1, \dots, x_n) \in E^n$. On suppose qu'il existe $h, k \in \{1, \dots, n\}$ tels que $h < k$ et $x_h = x_k$. Il s'agit de montrer que $\det_e(x) = 0$.

Notons τ la transposition $(h \ k)$ et \mathcal{A}_n le groupe alterné de degré n , c'est-à-dire le sous-groupe des permutations paires de \mathcal{S}_n . On sait que $\mathcal{S}_n = \mathcal{A}_n \sqcup \tau \mathcal{A}_n$.

$$\text{Ainsi, } \det_e(x) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n e_j^*(x_{\sigma(j)}) = \sum_{\sigma \in \mathcal{A}_n} \prod_{j=1}^n e_j^*(x_{\sigma(j)}) - \sum_{\sigma \in \mathcal{A}_n} \prod_{j=1}^n e_j^*(x_{\tau \sigma(j)}),$$

ou encore $\det_e(x) = \sum_{\sigma \in \mathcal{A}_n} \prod_{j=1}^n e_{\sigma^{-1}(j)}^*(x_j) - \sum_{\sigma \in \mathcal{A}_n} \prod_{j=1}^n e_{\sigma^{-1}(j)}^*(x_{\tau(j)})$.

Or, pour tout $j \in \mathbb{N}_n$, $x_j = x_{\tau(j)}$, donc $\det_e(x) = 0$. \square

Propriété. $\det_e(e) = 1$.

Démonstration.

$$\det_e(e) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n \delta_{j, \sigma(j)} = \varepsilon(\text{Id}_{\mathbb{N}_n}) = 1. \quad \square$$

Propriété. $A_n(E)$ est une droite vectorielle dirigée par \det_e .

Démonstration.

D'après le théorème précédent, tout élément f de $A_n(E)$ est colinéaire à \det_e , donc $A_n(E) \subset \text{Vect}(\det_e)$.

De plus, $\det_e \in A_n(E)$, donc $\text{Vect}(\det_e) \subset A_n(E)$. Ainsi, $A_n(E) = \text{Vect}(\det_e)$.

Or $\det_e \neq 0$, car $\det_e(e) = 1 \neq 0$, donc (\det_e) est un système libre qui engendre $A_n(E)$. \square

Remarque. Ainsi, à un coefficient multiplicatif non nul près, il n'y a qu'une forme n -linéaire alternée sur un \mathbb{K} -espace vectoriel de dimension n .

Dans un \mathbb{R} -espace vectoriel de dimension n , la seule façon raisonnable de définir le volume algébrique de l'hyperparallélépipède H_x associé à un n -uplet x de n vecteurs est donc de choisir une base e et de convenir que ce volume est égal à $\det_e(x)$. L'unité de volume est alors le volume de H_e . Changer le choix de la base e se limite à multiplier cette notion de volume par un réel non nul, ce qui change l'orientation si et seulement si ce réel est négatif.

En résumé, \det_e est la seule définition raisonnable du volume algébrique de H_x .

7.2.3 Déterminant d'une matrice

Définition. Soit $M \in \mathcal{M}_n(\mathbb{K})$. Le déterminant de M , noté $\det(M)$ est le déterminant des vecteurs colonnes de M dans la base canonique de \mathbb{K}^n .

Représentation tabulaire.

Si $M = (\alpha_{i,j}) \in \mathcal{M}_n(\mathbb{K})$. On note

$$\det(M) = \begin{vmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{vmatrix}.$$

Propriété. Soit $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$.

$$\det(M) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n m_{j, \sigma(j)} = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n m_{\sigma(j), j} = \det({}^t M).$$

Ainsi $\det(M)$ est aussi le déterminant des vecteurs lignes de M dans la base canonique de \mathbb{K}^n .

Formule de Sarrus :

$$\begin{vmatrix} p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,1} & p_{3,2} & p_{3,3} \end{vmatrix} = p_{1,1}p_{2,2}p_{3,3} + p_{2,1}p_{3,2}p_{1,3} + p_{3,1}p_{1,2}p_{2,3} \\ - p_{1,3}p_{2,2}p_{3,1} - p_{2,3}p_{3,2}p_{1,1} - p_{3,3}p_{1,2}p_{2,1}.$$

7.2.4 Déterminant d'un endomorphisme

Définition. Soit $u \in L(E)$. Le **déterminant de l'endomorphisme** u est l'unique scalaire, noté $\det(u)$, vérifiant $\boxed{\forall f \in A_n(E) \quad \forall x \in E^n \quad f(u(x)) = (\det(u))f(x)}$.

Démonstration.

Si $f \in A_n(E)$, notons f_u l'application de E^n dans \mathbb{K} définie par $f(x_1, \dots, x_n) = f(u(x_1), \dots, u(x_n))$. On vérifie que f_u est n -linéaire et alternée. Ainsi $\varphi_u : f \mapsto f_u$ est une application de $A_n(E)$ dans lui-même. On vérifie que φ_u est linéaire. Mais $A_n(E)$ est une droite vectorielle, donc $\dim(L(A_n(E))) = 1$ et $A_n(E) = \text{Vect}\{Id_{A_n(E)}\}$: il existe un unique scalaire, noté $\det(u)$ tel que $\varphi_u = \det(u).Id_{A_n(E)}$, c'est-à-dire tel que, pour tout $f \in A_n(E)$, $f_u = \det(u).f$. \square

Propriété. Soient e une base de E et $u \in L(E)$.

Pour tout $(x_1, \dots, x_n) \in E^n$, $\det_e(u(x_1), \dots, u(x_n)) = \det(u)\det_e(x_1, \dots, x_n)$.

En particulier, $\boxed{\det(u) = \det_e(u(e_1), \dots, u(e_n))}$.

Cependant, $\det(u)$ ne dépend pas du choix de la base e .

Propriété. Pour toute base e de E et pour tout $u \in L(E)$, $\det(u) = \det(Mat(u, e))$.

Démonstration.

$\det(u) = \det_e(u(e_1), \dots, u(e_n))$, mais si l'on note $Mat(u, e) = M = (m_{i,j})$, par définition

de \det_e , $\det(u) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n m_{\sigma(j),j}$, donc $\det(u) = \det(M)$. \square

Exemple. Si $\sigma \in \mathcal{S}_n$, on note $P_\sigma = (\delta_{i,\sigma(j)}) \in \mathcal{M}_n(\mathbb{K})$. Alors $\det(P_\sigma) = \varepsilon(\sigma)$.

Démonstration.

Notons $e = (e_1, \dots, e_n)$ la base canonique de \mathbb{K}^n . Si $s \in \mathcal{S}_n$, notons u_s l'endomorphisme canoniquement associé à la matrice P_s : Pour tout $j \in \{1, \dots, n\}$, $u_s(e_j) = e_{s(j)}$.

Alors $\det(P_\sigma) = \det_e(u_\sigma(e_1), \dots, u_\sigma(e_n)) = \varepsilon(\sigma)\det_e(e) = \varepsilon(\sigma)$. \square

7.3 Propriétés du déterminant

Notation.

On fixe $n \in \mathbb{N}^*$, E un \mathbb{K} -espace vectoriel de dimension n et e une base de E .

Propriété. \det_e est n -linéaire alternée, donc antisymétrique. $\det_e(e) = 1$.

$\det_e(x_1, \dots, x_n)$ n'est pas modifié si l'on ajoute à l'un des x_i une combinaison linéaire des autres x_j .

Propriété. Le déterminant d'une matrice M de $\mathcal{M}_n(\mathbb{K})$ est modifié en :

- $\det(M)$ pour une opération élémentaire du type $L_i \leftarrow L_i + \lambda L_j$
ou $C_i \leftarrow C_i + \lambda C_j$;
- $\alpha \det(M)$ pour une opération élémentaire du type $L_i \leftarrow \alpha L_i$ ou $C_i \leftarrow \alpha C_i$;
- $-\det(M)$ pour un échange entre deux lignes ou deux colonnes.

ATTENTION : En général, $\det(\alpha M + N) \neq \alpha \det(M) + \det(N)$.

Méthode : Pour calculer le déterminant d'une matrice, on tente de modifier la matrice par des manipulations élémentaires, afin de se ramener à une matrice dont on connaît le rang ou le déterminant.

Propriété. $\det(Id_E) = 1$, $\det(I_n) = 1$.

Pour tout $\lambda \in \mathbb{K}$ et $u \in L(E)$, $\det(\lambda u) = \lambda^n \det(u)$.

Pour tout $\lambda \in \mathbb{K}$ et $A \in \mathcal{M}_n(\mathbb{K})$, $\det(\lambda A) = \lambda^n \det(A)$.

Théorème. Si $f, g \in L(E)$, alors $\boxed{\det(fg) = \det(f) \times \det(g)}$.

Pour tout $A, B \in \mathcal{M}_n(\mathbb{K})$, $\det(AB) = \det(A)\det(B)$.

Démonstration.

$$\begin{aligned} \det(f \circ g) &= \det_e(f \circ g(e_1), \dots, f \circ g(e_n)) \\ &= \det(f) \det_e(g(e_1), \dots, g(e_n)) \quad \square \\ &= \det(f) \det(g). \end{aligned}$$

Formule de changement de base : Soient e et e' deux bases de E , et soit x une famille de n vecteurs de E . Alors, $\boxed{\det_{e'}(x) = \det_{e'}(e) \det_e(x)}$.

Démonstration.

C'est la formule (1) appliquée avec $f = \det_{e'}$. \square

Théorème. $\boxed{x \text{ est une base si et seulement si } \det_e(x) \neq 0.}$

Démonstration.

Supposons que x est une base, alors $1 = \det_x(x) = \det_x(e) \times \det_e(x)$, donc $\det_e(x) \neq 0$. Réciproquement, si x n'est pas une base, alors x est une famille liée de n vecteurs, or \det_e est alternée, donc $\det_e(x) = 0$. \square

Corollaire. Soit $u \in L(E)$ et $A \in \mathcal{M}_n(\mathbb{K})$.

$u \in GL(E)$ si et seulement si $\det(u) \neq 0$ et dans ce cas, $\det(u^{-1}) = \frac{1}{\det(u)}$.

$A \in GL_n(\mathbb{K})$ si et seulement si $\det(A) \neq 0$ et dans ce cas, $\det(A^{-1}) = \frac{1}{\det(A)}$.

Démonstration.

$u \in GL(E)$ si et seulement si $u(e)$ est une base de E , donc si et seulement si $\det(u) = \det_e(u(e)) \neq 0$.

Lorsque u est inversible, $\det(u) \times \det(u^{-1}) = \det(u \circ u^{-1}) = \det(Id_E) = 1$. \square

Remarque. \det est donc un morphisme du groupe $GL(E)$ vers (\mathbb{K}^*, \times) . Son noyau est un sous-groupe (distingué) de $GL(E)$, noté $SL(E)$. C'est le groupe spécial linéaire de E : $SL(E) = \{u \in L(E) \mid \det(u) = 1\}$.

On dispose en particulier de $SL_n(\mathbb{K}) = \{M \in \mathcal{M}_n(\mathbb{K}) \mid \det(M) = 1\}$: c'est le groupe spécial linéaire de degré n .

Propriété. Deux matrices carrées semblables ont le même déterminant.

Ainsi le déterminant, comme la trace et le rang, est un invariant de similitude.

7.4 Calcul des déterminants

Notation. $c = (c_1, \dots, c_n)$ désigne la base canonique de \mathbb{K}^n .

Lemme. On suppose que $n \geq 2$. Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$ dont la dernière colonne est c_n . Alors $\det(A)$ est égal au déterminant de la matrice extraite de A en ôtant la dernière colonne et la dernière ligne.

Démonstration.

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n A_{\sigma(j),j}, \text{ or } A_{\sigma(n),n} = \delta_{\sigma(n),n}, \text{ donc } \det(A) = \sum_{\substack{\sigma \in \mathcal{S}_n \\ \sigma(n)=n}} \varepsilon(\sigma) \prod_{j=1}^{n-1} A_{\sigma(j),j}.$$

Or l'application $\varphi : \{\sigma \in \mathcal{S}_n \mid \sigma(n) = n\} \longrightarrow \mathcal{S}_{n-1}$ définie par $\varphi(\sigma) = \sigma|_{\mathbb{N}_{n-1}}$ est une bijection et pour tout $\sigma \in \mathcal{S}_n$ telle que $\sigma(n) = n$, $\varepsilon(\varphi(\sigma)) = \varepsilon(\sigma)$: en effet, la décomposition de $\varphi(\sigma)$ en produit de transpositions de \mathcal{S}_{n-1} donne immédiatement une décomposition de σ en produit de transpositions de \mathcal{S}_n . Ainsi, en posant $s = \varphi(\sigma)$,

$$\text{on obtient } \det(A) = \sum_{s \in \mathcal{S}_{n-1}} \varepsilon(s) \prod_{j=1}^{n-1} A_{s(j),j}. \text{ Il s'agit bien du déterminant de la matrice}$$

extraite de A en ôtant la dernière colonne et la dernière ligne. \square

Remarque. Le lemme est encore vrai lorsque $n = 1$ si l'on convient que le déterminant de la matrice vide est égal à 1.

Définition. Soit $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$.

Pour tout $(i, j) \in \mathbb{N}_n^2$, notons ${}_{i,j}M$ la matrice extraite de M en ôtant la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne. La quantité $\det({}_{i,j}M)$ s'appelle le $(i, j)^{\text{ème}}$ **mineur** de M .

La quantité $C_{i,j} = (-1)^{i+j} \det({}_{i,j}M)$ s'appelle le $(i, j)^{\text{ème}}$ **cofacteur** de M .

Exemple. Pour $A = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix}$, les cofacteurs de la première colonne sont

$$C_{1,1} = 22 \times 33 - 32 \times 23 = -10, C_{2,1} = -(12 \times 33 - 32 \times 13) = 20$$

$$\text{et } C_{3,1} = 12 \times 23 - 22 \times 13 = -10.$$

Théorème. Avec ces notations,

— Pour tout $j \in \mathbb{N}_n$, $\det(M) = \sum_{i=1}^n m_{i,j} C_{i,j}$: c'est le **développement de $\det(M)$ selon sa $j^{\text{ème}}$ colonne.**

— Pour tout $i \in \mathbb{N}_n$, $\boxed{\det(M) = \sum_{j=1}^n m_{i,j} C_{i,j}}$: c'est le **développement de $\det(M)$ selon sa $i^{\text{ème}}$ ligne.**

Démonstration.

La seconde partie du théorème s'obtient en appliquant la première partie à tM .

Notons M_1, \dots, M_n les colonnes de M et fixons $j \in \mathbb{N}_n$. $M_j = \sum_{i=1}^n M_{i,j} c_i$, donc

$$\det(M) = \det_c(M_1, \dots, M_n) = \det_c(M_1, \dots, M_{j-1}, \sum_{i=1}^n M_{i,j} c_i, M_{j+1}, \dots, M_n), \text{ puis}$$

$$\det(M) = \sum_{i=1}^n M_{i,j} \det_c(M_1, \dots, M_{j-1}, c_i, M_{j+1}, \dots, M_n). \text{ Il reste donc à montrer que,}$$

pour tout $i \in \mathbb{N}_n$, $\det_c(M_1, \dots, M_{j-1}, c_i, M_{j+1}, \dots, M_n) = C_{i,j}$.

Soit $i \in \mathbb{N}_n$. En effectuant $n - j$ échanges de colonnes, on obtient

$$\det_c(M_1, \dots, M_{j-1}, c_i, M_{j+1}, \dots, M_n) = (-1)^{n-j} \det_c(M_1, \dots, M_{j-1}, M_{j+1}, \dots, M_n, c_i),$$

puis en effectuant $n - i$ échanges de lignes, on obtient

$$\det_c(M_1, \dots, M_{j-1}, c_i, M_{j+1}, \dots, M_n) = (-1)^{i+j} \begin{vmatrix} 0 & & \\ i,j M & & \\ & 0 & \\ * \dots * & & 1 \end{vmatrix} \text{ et le lemme permet de}$$

conclure. \square

Définition. On appelle **comatrice** de M la matrice $(C_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ des cofacteurs de M .

On la notera $Com(M)$ ou bien $Cof(M)$.

La transposée de la comatrice s'appelle la **matrice complémentaire** de M .

Théorème. $\forall M \in \mathcal{M}_n(\mathbb{K}) \quad \boxed{M^t Cof(M) = {}^t Cof(M) M = \det(M) I_n}$.

Démonstration.

Soit $i \in \mathbb{N}_n$. $[M^t Cof(M)]_{i,i} = \sum_{j=1}^n M_{i,j} C_{i,j} = \det(M)$ d'après la formule de développement de $\det(M)$ selon sa i -ème ligne.

De même $[{}^t Cof(M) M]_{i,i} = \sum_{j=1}^n C_{j,i} M_{j,i} = \det(M)$ d'après la formule de développement de $\det(M)$ selon sa i -ème colonne.

$$\text{Soit } i, j \in \mathbb{N}_n \text{ avec } i \neq j : [M^t Cof(M)]_{i,j} = \sum_{k=1}^n M_{i,k} C_{j,k}.$$

Notons L_1, \dots, L_n les lignes de M et notons A la matrice dont les lignes sont

$$L_1, \dots, L_{j-1}, L_i, L_{j+1}, \dots, L_n. \text{ Alors } [M^t Cof(M)]_{i,j} = \sum_{k=1}^n A_{j,k} C_{j,k} \text{ est le développement}$$

de $\det(A)$ selon sa j -ème ligne, car $C_{j,k}$ est bien le cofacteur de A de position (j, k) .

Ainsi, $[M^t Cof(M)]_{i,j} = \det(A) = 0$, car A possède deux lignes égales.

En raisonnant sur les colonnes, on montre de même que $[{}^t Cof(M) M]_{i,j} = 0$. \square

Corollaire. Lorsque M est inversible, $M^{-1} = \frac{1}{\det(M)} {}^t \text{Cof}(M)$.

Exemple. Avec $n = 2$, on retrouve que, lorsque $ad - bc \neq 0$,

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

Théorème. Soit $M = (M_{i,j})_{\substack{1 \leq i \leq a \\ 1 \leq j \leq a}}$ une matrice décomposée en blocs, où, pour tout $i, j \in \mathbb{N}_a$, $M_{i,j} \in \mathcal{M}_{n_i, n_j}(\mathbb{K})$.

Si M est triangulaire supérieure (ou inférieure) par blocs, alors, $\det(M) = \prod_{i=1}^a \det(M_{i,i})$

Démonstration.

Au prix d'une récurrence, il suffit de montrer que, pour tout $p, q \in \mathbb{N}^*$, pour tout $A \in \mathcal{M}_{\mathbb{K}}(p, p)$, $B \in \mathcal{M}_{\mathbb{K}}(p, q)$ et $C \in \mathcal{M}_{\mathbb{K}}(q, q)$, $\begin{vmatrix} A & B \\ 0_{q,p} & C \end{vmatrix} = \det(A)\det(C)$.

Si A n'est pas inversible, les colonnes de A sont liées, donc les colonnes de $\begin{pmatrix} A & B \\ 0_{q,p} & C \end{pmatrix}$ sont également liées. Ainsi, $\begin{vmatrix} A & B \\ 0_{q,p} & C \end{vmatrix} = 0 = \det(A)\det(C)$.

Si A est inversible, alors $\begin{pmatrix} A & B \\ 0_{q,p} & C \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & I_q \end{pmatrix} \times \begin{pmatrix} I_p & A^{-1}B \\ 0 & C \end{pmatrix}$, donc

$$\begin{vmatrix} A & B \\ 0_{q,p} & C \end{vmatrix} = \begin{vmatrix} A & 0 \\ 0 & I_q \end{vmatrix} \times \begin{vmatrix} I_p & A^{-1}B \\ 0 & C \end{vmatrix}, \text{ ce qui permet de conclure, car en développant}$$

plusieurs fois selon la première colonne, on montre que $\begin{vmatrix} I_p & A^{-1}B \\ 0 & C \end{vmatrix} = \det(C)$ et, en

développant plusieurs fois selon la dernière colonne, que $\begin{vmatrix} A & 0 \\ 0 & I_q \end{vmatrix} = \det(A)$. \square

Exemple.
$$\begin{vmatrix} -3 & -7 & 54 & 1 & 0 \\ 2 & 4 & 5 & 7 & 0 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 5 & -3 & 4 \\ 0 & 0 & 0 & 1 & 5 \end{vmatrix} = \begin{vmatrix} -3 & -7 \\ 2 & 4 \end{vmatrix} \times \begin{vmatrix} 1 & 0 & 2 \\ 5 & -3 & 4 \\ 0 & 1 & 5 \end{vmatrix} = 2 \times (-9) = -18.$$

Exemple. Posons $\Delta_n = \det((\min(i, j))_{1 \leq i, j \leq n})$. Si l'on effectue les opérations élémentaires

$$L_i \longleftarrow L_i - L_1 \text{ pour tout } i \geq 2, \text{ on obtient } \Delta_n = \begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & & & \\ \vdots & [\min(i, j)]_{1 \leq i, j \leq n-1} & & \\ 0 & & & \end{vmatrix}, \text{ donc}$$

$$\Delta_n = \Delta_{n-1} = D_1 = 1.$$

Corollaire. Le déterminant d'une matrice triangulaire supérieure ou inférieure est égal au produit de ses éléments diagonaux.

Remarque. On retrouve ainsi qu'une matrice triangulaire supérieure est inversible si et seulement si ses coefficients diagonaux sont non nuls.

Exemple. Supposons que $E = F \oplus G$. Notons s la symétrie par rapport à F parallèlement à G . Alors $\det(s) = (-1)^{\dim(G)}$.

Démonstration.

Soit (e_1, \dots, e_r) une base de F et (e_{r+1}, \dots, e_n) une base de G . Notons $e = (e_1, \dots, e_n)$. C'est une base de E et $\text{mat}(u, e)$ est diagonale, les r premiers coefficients diagonaux étant égaux à 1 et les suivants à -1 . On conclut facilement. \square

Conclusion : Pour calculer un déterminant, le plus souvent, on ajoute à une ligne (ou à une colonne) une combinaison linéaire des autres lignes (ou des autres colonnes), ou bien on effectue un développement selon une ligne (ou selon une colonne), dans le but de se ramener à un déterminant que l'on sait calculer, ou dont la valeur est connue : déterminants d'ordre 2 ou 3, déterminants triangulaires, ou bien des déterminants classiques étudiés plus loin.

7.5 Formules de Cramer

Propriété. Considérons un système linéaire de Cramer $(S) : MX = B$,

où $M \in GL_n(\mathbb{K})$, $B \in \mathbb{K}^n$, dont l'unique solution est notée $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$.

Alors, pour tout $j \in \{1, \dots, n\}$, $x_j = \frac{\det({}_jM)}{\det(M)}$, où ${}_jM$ est la matrice dont les colonnes sont celles de M , sauf la $j^{\text{ème}}$ qui est égale à B .

Démonstration.

Soit $j \in \{1, \dots, n\}$. Notons M_j la j -ème colonne de M .

$$\begin{aligned} \det({}_jM) &= \det_c(M_1, \dots, M_{j-1}, B, M_{j+1}, \dots, M_n) \\ &= \det_c(M_1, \dots, M_{j-1}, \sum_{i=1}^n x_i M_i, M_{j+1}, \dots, M_n) \\ &= \sum_{i=1}^n x_i \det_c(M_1, \dots, M_{j-1}, M_i, M_{j+1}, \dots, M_n) \\ &= x_j \det(M), \end{aligned}$$

car \det_c est une application n -linéaire alternée. \square

Remarque. Ces formules de Cramer sont utiles sur le plan théorique. Pour résoudre un système de Cramer, ces formules sont idéales lorsque $n = 2$, mais elles sont inadaptées lorsque $n \geq 3$. En effet, l'utilisation de ces formules nécessite n divisions et le calcul de $n + 1$ déterminants d'ordre n .

Mais, pour $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, $\det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n a_{j,\sigma(j)}$, donc le calcul d'un déterminant d'ordre n , en procédant de manière naïve demande $n! - 1$ additions ou soustractions et $n!(n - 1)$ multiplications.

Ainsi, résoudre (S) par application directe des formules de Cramer demande à peu près $(n+2)!$ multiplications, $(n+1)!$ additions et n divisions.

Supposons que nous utilisons un ordinateur d'une puissance de 1GHz ($= 10^9$ cycles par seconde). Le nombre de multiplications qu'il peut effectuer pendant une durée égale à l'âge de l'univers est de l'ordre de $10 \times 10^9 \times 365 \times 24 \times 60^2 \times 10^9 \approx 10^{27}$. La résolution d'un système d'ordre 30 nécessite environ 10^{35} multiplications. Ainsi, en supposant que notre ordinateur travaille à cette tâche depuis la création de l'univers, il n'aura à notre époque réalisé que 10 milliardièmes des calculs nécessaires !

Il faut donc se tourner vers des algorithmes plus efficaces : on peut montrer que l'algorithme du pivot de Gauss nécessite de l'ordre de $\frac{n^3}{3}$ multiplications et additions et de l'ordre de $\frac{n^2}{2}$ divisions. La résolution d'un système d'ordre 30 ne demandera plus qu'environ 18000 opérations, ce qui sera effectué en 2×10^{-5} secondes...

7.6 Exemples de déterminants.

7.6.1 Déterminant de Vandermonde

Définition. Soient $n \in \mathbb{N}$ et $(a_0, \dots, a_n) \in \mathbb{K}^{n+1}$.

La **matrice de Vandermonde** est $\mathcal{V}(a_0, \dots, a_n) = (a_{i-1}^{j-1}) \in \mathcal{M}_{n+1}(\mathbb{K})$, et le **déterminant de Vandermonde** est $V(a_0, \dots, a_n) = \det(\mathcal{V}(a_0, \dots, a_n))$. Ainsi,

$$V(a_0, \dots, a_n) = \begin{vmatrix} 1 & a_0 & a_0^2 & \cdots & a_0^n \\ 1 & a_1 & a_1^2 & \cdots & a_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^n \end{vmatrix}.$$

Propriété. Soient $n \in \mathbb{N}^*$ et $(a_0, \dots, a_n) \in \mathbb{K}^{n+1}$.

$$V(a_0, \dots, a_n) = V(a_0, \dots, a_{n-1}) \prod_{i=0}^{n-1} (a_n - a_i).$$

Démonstration.

Afin d'illustrer les différentes techniques relatives au calcul des déterminants, nous allons présenter 4 démonstrations de ce résultat.

Première démonstration. par combinaison linéaire de colonnes.

Effectuons sur le déterminant de Vandermonde de (a_0, \dots, a_n) les opérations élémentaires suivantes : $C_j \leftarrow C_j - a_n C_{j-1}$ dans l'ordre suivant : j varie de $n+1$ à 2 (en effet, au rang j , on a besoin de C_{j-1} , donc, au rang j , la colonne d'indice $j-1$ ne doit pas avoir été modifiée).

On obtient

$$V(a_0, \dots, a_n) = \begin{vmatrix} 1 & a_0 - a_n & \cdots & a_0^{n-2}(a_0 - a_n) & a_0^{n-1}(a_0 - a_n) \\ 1 & a_1 - a_n & \cdots & a_1^{n-2}(a_1 - a_n) & a_1^{n-1}(a_1 - a_n) \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & a_{n-1} - a_n & \cdots & a_{n-1}^{n-2}(a_{n-1} - a_n) & a_{n-1}^{n-1}(a_{n-1} - a_n) \\ 1 & 0 & \cdots & 0 & 0 \end{vmatrix}.$$

Développons selon sa dernière ligne ce déterminant de taille $n+1$. On obtient

$$V(a_0, \dots, a_n) = (-1)^{(n+1)+1} \left(\prod_{i=0}^{n-1} (a_i - a_n) \right) V(a_0, \dots, a_{n-1}).$$

Deuxième démonstration, utilisant des combinaisons linéaires de colonnes et des polynômes.

Soit P un polynôme unitaire de degré n à coefficients dans \mathbb{K} , noté $P = X^n + \sum_{j=1}^n b_j X^{j-1}$.

Effectuons sur le déterminant de Vandermonde de (a_0, \dots, a_n) l'opération élémentaire suivante : $C_{n+1} \leftarrow C_{n+1} + \sum_{j=1}^n b_j C_j$.

Le $i^{\text{ème}}$ coefficient de la dernière colonne devient alors $a_i^n + \sum_{j=1}^n b_j a_i^{j-1} = P(a_i)$.

En particulier, si l'on choisit $P = \prod_{i=0}^{n-1} (X - a_i)$ (qui est bien un polynôme unitaire de degré n), les coefficients de la dernière colonne sont tous nuls, sauf le dernier, qui vaut $\prod_{i=0}^{n-1} (a_n - a_i)$. Ainsi, en développant par rapport à la dernière colonne,

$$V(a_0, \dots, a_n) = V(a_0, \dots, a_{n-1}) \prod_{i=0}^{n-1} (a_n - a_i).$$

Pour les deux démonstrations suivantes, on supposera que a_0, \dots, a_n sont deux à deux distincts. Ce n'est pas restrictif car lorsque, parmi a_0, \dots, a_n , deux scalaires au moins sont égaux, le déterminant $V(a_0, \dots, a_n)$ contient au moins deux lignes égales, donc il

est nul, ainsi que la quantité $V(a_0, \dots, a_{n-1}) \prod_{i=0}^{n-1} (a_i - a_n)$.

Troisième démonstration, utilisant des polynômes.

$$\text{Soit } x \in \mathbb{K}. V(a_0, \dots, a_{n-1}, x) = \begin{vmatrix} 1 & a_0 & a_0^2 & \cdots & a_0^n \\ 1 & a_1 & a_1^2 & \cdots & a_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \cdots & a_{n-1}^n \\ 1 & x & x^2 & \cdots & x^n \end{vmatrix}.$$

Si l'on développe ce déterminant selon sa dernière ligne, on obtient un polynôme en x de degré inférieur ou égal à n , dont le coefficient de degré n vaut $V(a_0, \dots, a_{n-1})$.

Soit $i \in \{0, \dots, n-1\}$. La matrice $\mathcal{V}(a_0, \dots, a_{n-1}, a_i)$ possède deux lignes identiques, donc $V(a_0, \dots, a_{n-1}, a_i)$ est nul. Ainsi, le polynôme $V(a_0, \dots, a_{n-1}, x)$ admet au moins n racines deux à deux distinctes, qui sont a_0, \dots, a_{n-1} . C'est donc un multiple de $\prod_{i=0}^{n-1} (x - a_i)$.

Il existe $Q \in \mathbb{K}[X]$ tel que $V(a_0, \dots, a_{n-1}, x) = Q(x) \prod_{i=0}^{n-1} (x - a_i)$.

Nécessairement, $\deg(Q) \leq 0$, donc Q est une constante, et, en égalant les coefficients de degré n , on obtient que cette constante vaut $V(a_0, \dots, a_{n-1})$.

Ainsi, $V(a_0, \dots, a_{n-1}, x) = V(a_0, \dots, a_{n-1}) \left(\prod_{i=0}^{n-1} (x - a_i) \right)$.

Quatrième démonstration, utilisant les polynômes d'interpolation de Lagrange.

Reprenons les notations de la page 24 et notons $c = (1, X, \dots, X^n)$ la base canonique de $\mathbb{K}_n[X]$. Soit $j \in \{0, \dots, n\}$.

Dans la base $L = (L_0, \dots, L_n)$ de $\mathbb{K}_n[X]$, les coordonnées du polynôme X^j sont $a_0^j, a_1^j, \dots, a_n^j$, donc la matrice $\mathcal{V}(a_0, \dots, a_n)$ est la matrice de passage de la base L vers la base c , notée P_L^c .

Remarquons, même si ce n'est pas exactement le but de la démonstration, que ce qui précède montre sans calcul que la matrice $\mathcal{V}(a_0, \dots, a_n)$ est inversible si et seulement si a_0, \dots, a_n sont deux à deux distincts.

De plus, ce qui précède permet d'inverser rapidement $\mathcal{V}(a_0, \dots, a_n)$. En effet, $\mathcal{V}^{-1}(a_0, \dots, a_n) = P_c^L$, donc le (i, j) ^{ème} coefficient de $\mathcal{V}^{-1}(a_0, \dots, a_n)$ est le coefficient de degré $i-1$ du polynôme L_{j-1} , que l'on pourrait exprimer en fonction de a_0, \dots, a_n en utilisant les relations entre coefficients et racines d'un polynôme.

Cependant, pour le calcul du déterminant de Vandermonde, on peut se contenter de calculer le coefficient de position $(n+1, n+1)$ de $\mathcal{V}^{-1}(a_0, \dots, a_n)$. Il s'agit du coefficient

dominant de $L_{n+1} = \prod_{i=0}^{n-1} \frac{X - a_i}{a_n - a_i}$, donc il est égal à $\prod_{i=0}^{n-1} \frac{1}{a_n - a_i}$.

D'autre part, $\mathcal{V}^{-1}(a_0, \dots, a_n) = \frac{1}{V(a_0, \dots, a_n)} {}^t \text{Cof}(\mathcal{V}(a_0, \dots, a_n))$, donc le coefficient

de position de $(n+1, n+1)$ de $\mathcal{V}^{-1}(a_0, \dots, a_n)$ est aussi égal à $\frac{C_{n+1, n+1}}{V(a_0, \dots, a_n)}$, où $C_{n+1, n+1}$ désigne le cofacteur de $\mathcal{V}(a_0, \dots, a_n)$ de position de $(n+1, n+1)$.

On en déduit que $\prod_{i=0}^{n-1} \frac{1}{a_n - a_i} = \frac{V(a_0, \dots, a_{n-1})}{V(a_0, \dots, a_n)}$. \square

Propriété. Soient $n \in \mathbb{N}$ et $(a_0, \dots, a_n) \in \mathbb{K}^{n+1}$: $V(a_0, \dots, a_n) = \prod_{0 \leq i < j \leq n} (a_j - a_i)$.

Ainsi $\mathcal{V}(a_0, \dots, a_n)$ est inversible si et seulement si a_0, \dots, a_n sont deux à deux distincts.

Démonstration.

Soit $n \in \mathbb{N}$. Notons $R(n)$ l'assertion suivante :

$$\forall (a_0, \dots, a_n) \in \mathbb{K}^{n+1} \quad V(a_0, \dots, a_n) = \prod_{0 \leq i < j \leq n} (a_j - a_i).$$

Démontrons par récurrence sur n que $R(n)$ est vraie pour tout n .

Pour $n = 0$, soit $a_0 \in \mathbb{K}$. $V(a_0) = \det(1) = 1 = \prod_{0 \leq i < j \leq 0} (a_j - a_i)$, car l'ensemble des

indices de ce produit est l'ensemble vide.

Pour $n \geq 1$, supposons $R(n-1)$.

Soit $(a_0, \dots, a_n) \in \mathbb{K}^{n+1}$. On a établi que

$$V(a_0, \dots, a_n) = V(a_0, \dots, a_{n-1}) \prod_{i=0}^{n-1} (a_n - a_i), \text{ donc, d'après l'hypothèse de récurrence,}$$

$$V(a_0, \dots, a_n) = \left(\prod_{0 \leq i < j \leq n-1} (a_j - a_i) \right) \left(\prod_{i=0}^{n-1} (a_n - a_i) \right) = \prod_{0 \leq i < j \leq n} (a_j - a_i), \text{ ce qui démontre}$$

$R(n)$. \square

7.6.2 Déterminants tridiagonaux

Définition. Soient n un entier supérieur ou égal à 2 et $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$.

M est une **matrice tridiagonale** si et seulement si, pour tout $(i, j) \in \mathbb{N}_n^2$,

$|i - j| \geq 2 \implies m_{i,j} = 0$. Un **déterminant tridiagonal** est le déterminant d'une matrice tridiagonale.

Notation. Fixons un entier n supérieur ou égal à 3 et $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ une matrice tridiagonale.

Pour tout $k \in \mathbb{N}_n$, notons $M_k = (m_{i,j})_{\substack{1 \leq i \leq k \\ 1 \leq j \leq k}}$. Ainsi, pour tout $k \in \mathbb{N}_n$, M_k est une matrice de taille k , extraite de M en ne retenant que ses k premières colonnes et ses k premières lignes. En particulier, $M = M_n$.

Pour tout $k \in \mathbb{N}_n$, notons Δ_k le déterminant de M_k .

Propriété. Pour tout $k \geq 3$, $\Delta_k = m_{k,k} \Delta_{k-1} - m_{k-1,k} m_{k,k-1} \Delta_{k-2}$

Démonstration.

Soit $k \geq 3$.

$$\Delta_k = \begin{vmatrix} m_{1,1} & m_{1,2} & 0 & \cdots & \cdots & \cdots & 0 \\ m_{2,1} & \ddots & m_{2,3} & \ddots & & & \vdots \\ 0 & m_{3,2} & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & m_{k-1,k} \\ 0 & \cdots & \cdots & \cdots & 0 & m_{k,k-1} & m_{k,k} \end{vmatrix},$$

donc, en développant selon la dernière colonne, on obtient

$$\Delta_k = m_{k,k} \begin{vmatrix} m_{1,1} & m_{1,2} & 0 & \cdots & \cdots & \cdots & 0 \\ m_{2,1} & \ddots & m_{2,3} & \ddots & & & \vdots \\ 0 & m_{3,2} & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & m_{k-2,k-1} \\ 0 & \cdots & \cdots & \cdots & 0 & m_{k-1,k-2} & m_{k-1,k-1} \end{vmatrix} - m_{k-1,k} \begin{vmatrix} m_{1,1} & m_{1,2} & 0 & \cdots & \cdots & \cdots & 0 \\ m_{2,1} & \ddots & m_{2,3} & \ddots & & & \vdots \\ 0 & m_{3,2} & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & m_{k-2,k-3} & m_{k-2,k-2} & m_{k-2,k-1} \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & m_{k,k-1} \end{vmatrix}.$$

Dans le membre de droite, le premier déterminant est Δ_{k-1} et, en développant le second déterminant selon la dernière ligne, on montre que ce dernier est égal à $m_{k,k-1}\Delta_{k-2}$. On obtient ainsi la relation annoncée. \square

Remarque. (Δ_k) est ainsi une suite récurrente linéaire d'ordre 2.

En particulier lorsque les suites $(m_{k,k})$ et $(m_{k-1,k}m_{k,k-1})$ sont constantes, on sait en déduire une expression de Δ_k en fonction de k .

Exemple. Pour $n \geq 2$, notons $M_n = (m_{i,j})$ la matrice de taille n dont les coefficients sont définis par les relations suivantes : pour tout $i \in \mathbb{N}_n$, $m_{i,i} = 2$, pour tout $i \in \mathbb{N}_{n-1}$, $m_{i,i+1} = 1$ et $m_{i+1,i} = 3$, et, pour tout $(i,j) \in \mathbb{N}_n^2$ tel que $|i-j| \geq 2$, $m_{i,j} = 0$.

D'après la propriété précédente, si l'on note Δ_n le déterminant de M_n , pour tout $n \geq 4$, $\Delta_n = 2\Delta_{n-1} - 3\Delta_{n-2}$.

Ainsi, $(\Delta_n)_{n \geq 2}$ est une suite récurrente linéaire d'ordre 2 à coefficients constants. Le polynôme caractéristique associé à cette suite est $P(X) = X^2 - 2X + 3$, dont les racines sont $1 + i\sqrt{2}$ et $1 - i\sqrt{2}$, donc il existe $(\lambda, \mu) \in \mathbb{C}^2$ tel que,

pour tout $n \geq 2$, (1) : $\Delta_n = \lambda(1 + i\sqrt{2})^n + \mu(1 - i\sqrt{2})^n$.

Déterminons λ et μ . On pourrait dans ce but calculer Δ_2 et Δ_3 et substituer leurs valeurs dans la relation (1), mais il y a plus simple. Nous allons prolonger la suite $(\Delta_n)_{n \geq 2}$ en l'unique suite $(\Delta_n)_{n \geq 0}$ qui vérifie la même relation de récurrence linéaire d'ordre 2.

Or, en posant $\Delta_1 = 2$, $\Delta_3 = 2\Delta_2 - 3\Delta_1$, et $\Delta_2 = \begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} = 2 \times 2 - 3$, donc il suffit de poser $\Delta_0 = 1$.

Substituons successivement n par 0 et par 1 dans la relation (1). On obtient : $\lambda + \mu = 1$

et $\lambda + \mu + i\sqrt{2}(\lambda - \mu) = 2$. Ainsi, $\lambda - \mu = -i\frac{\sqrt{2}}{2}$, donc $\lambda = \frac{1}{2} - i\frac{\sqrt{2}}{4}$ et $\mu = \bar{\lambda}$.

On en déduit que, pour tout $n \in \mathbb{N}$, $\Delta_n = 2\operatorname{Re} \left(\left(\frac{1}{2} - i\frac{\sqrt{2}}{4} \right) (1 + i\sqrt{2})^n \right)$.

En développant à l'aide de la formule du binôme de Newton, on obtient :

$$\Delta_n = \operatorname{Re} \left(\left(1 - i\frac{\sqrt{2}}{2} \right) \sum_{k=0}^n \binom{n}{2k} i^k 2^{\frac{k}{2}} \right) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2k} (-1)^k 2^k + \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n}{2k+1} (-1)^k 2^k.$$

7.6.3 Déterminants circulants

Notation. Fixons un entier n strictement positif.

Si $\sigma \in \mathcal{S}_n$ et $x = (a_1, \dots, a_n) \in \mathbb{K}^n$, on note $\sigma(x) = (a_{\sigma(1)}, \dots, a_{\sigma(n)})$.

Enfin, notons c le cycle de longueur n suivant : $c = (n, n-1, \dots, 2, 1)$.

Définition. Soit $M \in \mathcal{M}_n(\mathbb{K})$ une matrice dont les lignes sont notées L_1, \dots, L_n .

On dit que M est circulante si et seulement si, pour tout $i \in \mathbb{N}_n$, $L_i = c^{i-1}(L_1)$.

Ainsi, on passe d'une ligne à la suivante selon la permutation circulaire c .

Remarque. On peut calculer le déterminant d'une matrice circulante quelconque au moyen de la théorie de la réduction des matrices. Cependant, dans des cas simples, il n'est pas utile de faire appel au calcul général. Il est souvent suffisant de commencer par remplacer la première ligne par la somme de toutes les lignes. En effet, la somme de toutes les lignes est un vecteur de \mathbb{K}^n colinéaire à $(1, 1, \dots, 1)$. Ainsi, après mise en facteur, la première ligne ne contient que des "1". On peut alors, pour j variant de n à 2, effectuer les opérations $C_j \leftarrow C_j - C_{j-1}$.

Exemple. Soit $n \geq 2$. Calculez le déterminant Δ de la matrice d'ordre n dont le coefficient de position (i, j) vaut $j - i + 1$ si $i \leq j$ et $n + j - i + 1$ si $i \geq j$.

$$\Delta = \begin{vmatrix} 1 & 2 & 3 & \cdots & \cdots & \cdots & n \\ n & 1 & 2 & \ddots & & & n-1 \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & 3 \\ 3 & & & & n & \ddots & 2 \\ 2 & \cdots & \cdots & \cdots & & n & 1 \end{vmatrix}.$$

On vérifie qu'il s'agit bien du déterminant d'une matrice circulante.

$$\text{L'opération } L_1 \leftarrow \sum_{i=1}^n L_i \text{ donne } \Delta = \frac{n(n+1)}{2} \begin{vmatrix} 1 & 1 & 1 & \cdots & \cdots & \cdots & 1 \\ n & 1 & 2 & 3 & & & n-1 \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & 3 \\ 3 & & & & n & \ddots & 2 \\ 2 & 3 & \cdots & \cdots & & n & 1 \end{vmatrix}.$$

Pour j variant de n à 2, effectuons les opérations $C_j \leftarrow C_j - C_{j-1}$. Ainsi,

$$\Delta = \frac{n(n+1)}{2} \begin{vmatrix} 1 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ n & 1-n & 1 & \cdots & \cdots & \cdots & 1 \\ \vdots & 1 & \ddots & \ddots & & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \ddots & \vdots \\ 3 & 1 & & & 1 & 1-n & 1 \\ 2 & 1 & \cdots & \cdots & 1 & 1 & 1-n \end{vmatrix}.$$

Développons selon la première ligne et, sur le déterminant de taille $n-1$ ainsi obtenu,

effectuons l'opération $L_1 \leftarrow \sum_{i=1}^{n-1} L_i$. On obtient :

$$\Delta = \frac{n(n+1)}{2} \begin{vmatrix} -1 & -1 & \cdots & \cdots & \cdots & -1 \\ 1 & 1-n & 1 & \cdots & \cdots & 1 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 1 & & & 1 & 1-n & 1 \\ 1 & \cdots & \cdots & 1 & 1 & 1-n \end{vmatrix}.$$

Pour i variant de 2 à $n-1$, effectuons $L_i \leftarrow L_i + L_1$. Ainsi,

$$\Delta = \frac{n(n+1)}{2} \begin{vmatrix} -1 & -1 & \cdots & \cdots & \cdots & -1 \\ 0 & -n & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & & & 0 & -n & 0 \\ 0 & \cdots & \cdots & 0 & 0 & -n \end{vmatrix} = -\frac{n(n+1)}{2} (-n)^{n-2}.$$

7.7 Le polynôme caractéristique

Notation. On fixe un \mathbb{K} -espace vectoriel E de dimension $n \in \mathbb{N}^*$ et $u \in L(E)$.

7.7.1 Définition

Définition. Soit $M \in \mathcal{M}_n(\mathbb{K})$. On appelle polynôme caractéristique la quantité $\chi_M = \det(XI_n - M)$. C'est le déterminant d'une matrice dont les coefficients sont dans le corps $\mathbb{K}(X)$. χ_M est un élément de $\mathbb{K}[X]$.

De plus, pour tout $\lambda \in \mathbb{K}$, $\chi_M(\lambda) = \det(\lambda I_n - M)$.

Démonstration.

En posant $M = (m_{i,j})$, $\chi_M = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n (X\delta_{j,\sigma(j)} - m_{j,\sigma(j)})$.

On en déduit que, si $\lambda \in \mathbb{K}$, $\chi_M(\lambda) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n (\lambda\delta_{j,\sigma(j)} - m_{j,\sigma(j)}) = \det(\lambda I_n - M)$.

□

Remarque.

Vous rencontrerez parfois une définition légèrement différente du polynôme caractéristique. Il s'agit de $\det(M - XI_n)$. On passe de cette dernière convention à celle que nous avons adoptée en multipliant par $(-1)^n$. Dans un problème ou au sein d'un exercice, il est bon de se demander quelle est la convention (parfois implicitement) utilisée.

Représentation tabulaire. Si $M = \begin{pmatrix} m_{1,1} & \cdots & m_{1,n} \\ \vdots & & \vdots \\ m_{n,1} & \cdots & m_{n,n} \end{pmatrix}$,

$$\chi_M(\lambda) = \begin{vmatrix} \lambda - m_{1,1} & -m_{1,2} & \cdots & -m_{1,n} \\ -m_{2,1} & \lambda - m_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -m_{n-1,n} \\ -m_{n,1} & \cdots & -m_{n,n-1} & \lambda - m_{n,n} \end{vmatrix}.$$

Remarque. Souvent, le corps est de cardinal infini, ce qui permet d'identifier le polynôme χ_M avec l'application polynomiale $\lambda \mapsto \chi_M(\lambda)$ de \mathbb{K} dans \mathbb{K} .

Propriété. Pour tout $M \in \mathcal{M}_n(\mathbb{K})$, $\chi_{tM} = \chi_M$.

Démonstration.

$$\chi_{tM} = \det(XI_n - {}^tM) = \det({}^t(XI_n - M)) = \det(XI_n - M) = \chi_M. \quad \square$$

Propriété. Si M est une matrice triangulaire supérieure (ou inférieure) dont les coefficients diagonaux sont $\lambda_1, \dots, \lambda_n$, alors $\chi_M(X) = \prod_{i=1}^n (X - \lambda_i)$.

Propriété. Deux matrices semblables ont le même polynôme caractéristique.

Démonstration.

Soit $M \in \mathcal{M}_n(\mathbb{K})$ et $P \in GL_n(\mathbb{K})$.

$$\chi_{PMP^{-1}} = \det(XI_n - PMP^{-1}) = \det(P(XI_n - M)P^{-1}) = \det(XI_n - M) = \chi_M. \quad \square$$

Remarque. La réciproque est fausse.

Démonstration.

Posons $M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. $\chi_M(X) = (X - 1)^2 = \chi_{I_2}(X)$.

Supposons que M et I_2 sont semblables. Ainsi, il existe $P \in GL_2(\mathbb{R})$ tel que $M = PI_2P^{-1} = PP^{-1} = I_2$. C'est faux, donc M et I_2 ne sont pas semblables alors qu'elles ont le même polynôme caractéristique. \square

Définition. On déduit de la propriété précédente que la quantité $\chi_{mat(u,e)}$ ne dépend pas du choix de la base e de E . Cette quantité s'appelle le polynôme caractéristique de u .

Propriété. $(\lambda \in Sp(u)) \iff (\lambda \in \mathbb{K} \text{ et } \chi_u(\lambda) = 0)$.

Démonstration.

Soit $\lambda \in \mathbb{K}$. $\lambda \in Sp(u)$ si et seulement si $\lambda Id_E - u$ n'est pas injectif, donc si et seulement si $\det(\lambda Id_E - u) = 0$. \square

Corollaire. Pour toute $M \in \mathcal{M}_n(\mathbb{K})$, $Sp({}^tM) = Sp(M)$.

Corollaire. Le spectre d'une matrice triangulaire supérieure est égal l'ensemble de ses coefficients diagonaux.

Exemple. Choisissons $M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$.

Soit $\lambda \in \mathbb{R}$. $\chi_M(\lambda) = \begin{vmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -1 & 1 & \lambda - 1 \end{vmatrix}$. On effectue $L_1 \leftarrow L_1 + L_3$. On obtient

$\chi_M(\lambda) = \begin{vmatrix} \lambda - 1 & 0 & \lambda - 1 \\ 0 & \lambda & -1 \\ -1 & 1 & \lambda - 1 \end{vmatrix}$. On effectue $C_3 \leftarrow C_3 - C_1$. On obtient

$\chi_M(\lambda) = (\lambda - 1) \begin{vmatrix} 1 & 0 & 0 \\ 0 & \lambda & -1 \\ -1 & 1 & \lambda \end{vmatrix}$, donc

$\chi_M(\lambda) = (\lambda - 1)(\lambda^2 + 1)$. Ainsi $Sp(M) = \{1\}$.

Cependant, on peut aussi considérer M comme une matrice à coefficients complexes, auquel cas, $Sp(M) = \{1, i, -i\}$.

Remarque. Si $M \in \mathcal{M}_n(\mathbb{R})$, le spectre de M considérée comme matrice à coefficients dans \mathbb{R} (noté $Sp_{\mathbb{R}}(M)$) n'est pas en général égal au spectre de M considérée comme matrice à coefficients complexes (noté $Sp_{\mathbb{C}}(M)$). On dispose seulement de la relation $Sp_{\mathbb{R}}(M) = Sp_{\mathbb{C}}(M) \cap \mathbb{R}$. Il est donc important de préciser de quel spectre de M on parle.

De même, si $u \in E$ (où E est un \mathbb{R} -espace vectoriel), les valeurs propres de u sont dans \mathbb{R} , mais parfois, on appelle valeurs propres de u toutes les racines de χ_u , même celles appartenant à $\mathbb{C} \setminus \mathbb{R}$. Ici aussi, il est important de préciser quelle définition d'une valeur propre on utilise.

Par exemple, au sein d'un problème ou d'un exercice, la question "toutes les valeurs propres de u sont-elles dans \mathbb{R} ?" n'a d'intérêt que si l'on considère que les valeurs propres de u sont toutes les racines complexes de χ_u .

Définition. Soit $\lambda \in Sp(u)$.

On appelle **multiplicité** de λ sa multiplicité en tant que racine de χ_u .

Elle est notée $m(\lambda)$.

Si $m(\lambda) = 1$, on dit que λ est une **valeur propre simple** de u .

Si $m(\lambda) = 2$, on dit que λ est une **valeur propre double** de u .

Si $m(\lambda) = 3$, on dit que λ est une **valeur propre triple** de u .

Remarque. Lorsque $\lambda \notin Sp(u)$, on convient que $m(\lambda) = 0$.

7.7.2 Propriétés du polynôme caractéristique

Propriété. χ_u est un polynôme unitaire de degré n tel que

$$\chi_u(X) = X^n - Tr(u)X^{n-1} + \cdots + (-1)^n \det(u).$$

Remarque. Dans la formule ci-dessus, les “ $+\dots+$ ” indiquent qu’il y a des termes intermédiaires, mais il n’y a pas de formule simple donnant ces termes.

Démonstration.

Fixons une base e de E et notons $M = (m_{i,j}) = \text{mat}(u, e)$.

$$\chi_u = \chi_M = \det(XI_n - M) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n (X\delta_{j,\sigma(j)} - m_{j,\sigma(j)}).$$

- Soit $\sigma \in \mathcal{S}_n$. $\prod_{j=1}^n (X\delta_{j,\sigma(j)} - m_{j,\sigma(j)})$ est un polynôme en X de degré inférieur à n .

De plus, son degré est égal à n si et seulement si, pour tout $j \in \mathbb{N}_n$, $\sigma(j) = j$, c’est-à-dire si et seulement si $\sigma = \text{Id}_{\mathbb{N}_n}$.

Ainsi χ_u est un polynôme de degré n dont le coefficient dominant vaut $\varepsilon(\text{Id}_{\mathbb{N}_n}) = 1$.

- Soit $\sigma \in \mathcal{S}_n \setminus \{\text{Id}_{\mathbb{N}_n}\}$. Il existe $i \in \mathbb{N}_n$ tel que $\sigma(i) \neq i$. Posons $j = \sigma(i)$. Si $\sigma(j) = j$, σ étant injective, $j = i$, ce qui est faux. Ainsi $\{k \in \mathbb{N}_n / \sigma(k) \neq k\}$ est de cardinal supérieur ou égal à 2. Donc $\prod_{j=1}^n (X\delta_{j,\sigma(j)} - m_{j,\sigma(j)})$ est un polynôme en X de degré inférieur à $n - 2$.

Ainsi $\chi_u(X) = \prod_{j=1}^n (X - m_{j,j}) + Q(X)$ où Q est un polynôme de degré inférieur ou

égal à $n - 2$. On en déduit que $\chi_u(X) = X^n - \left(\sum_{j=1}^n m_{j,j} \right) X^{n-1} + R(X)$ où R est un polynôme de degré inférieur ou égal à $n - 2$.

D’autre part, le terme constant de χ_u est $\chi_u(0) = \det(-u) = (-1)^n \det(u)$. \square

Corollaire. Si $\mathbb{K} = \mathbb{C}$, u admet au moins un vecteur propre.

Démonstration.

χ_u est un polynôme de $\mathbb{C}[X]$ de degré supérieur à 1, or \mathbb{C} est algébriquement clos, donc χ_u admet au moins une racine dans \mathbb{C} . \square

Exercice. Soit E un \mathbb{C} -espace vectoriel de dimension $n \geq 1$ et $u, v \in L(E)$ tels que $uv = vu$. Montrer que u et v possèdent au moins un vecteur propre commun.

Résolution. u possède au moins une valeur propre $\lambda \in \mathbb{C}$. E_λ^u est un \mathbb{C} -espace vectoriel non nul et stable par v , donc $v|_{E_\lambda^u}$ possède un vecteur propre $x \in E_\lambda^u$. Alors x est un vecteur propre commun à u et v .

Contrexemple en dimension quelconque.

Choisissons $E = \mathbb{C}[X]$ et $u : \begin{matrix} E & \longrightarrow & E \\ P & \longmapsto & XP(X) \end{matrix}$.

Soit $P \in E \setminus \{0\}$. $\deg(u(P)) = 1 + \deg(P)$, donc il n’existe aucun $\lambda \in \mathbb{K}$ tel que $u(P) = \lambda P$.

Ainsi le spectre de u est égal à l’ensemble vide.

Corollaire. Si χ_u est scindé sur \mathbb{K} (c'est toujours le cas lorsque $\mathbb{K} = \mathbb{C}$),

$$\boxed{Tr(u) = \sum_{\lambda \in Sp_{\mathbb{K}}(u)} m(\lambda)\lambda, \quad \text{et } \det(u) = \prod_{\lambda \in Sp_{\mathbb{K}}(u)} \lambda^{m(\lambda)}.$$

Démonstration.

D'après les relations entre coefficients et racines d'un polynôme, $Tr(u)$ est égal à la somme des racines de χ_u , comptées avec multiplicité, or χ_u est scindé dans \mathbb{K} , donc l'ensemble des racines de χ_u est $Sp_{\mathbb{K}}(u)$.

Ainsi, $Tr(u) = \sum_{\lambda \in Sp_{\mathbb{K}}(u)} m(\lambda)\lambda$.

Le raisonnement est similaire pour la seconde formule. \square

Remarque. En pratique, pour déterminer les éléments propres d'une matrice M , on peut commencer par calculer χ_M . On détermine les racines de χ_M et, pour chacune d'entre elles, notée λ , on recherche une base du sous-espace propre en résolvant le

$$\text{système linéaire } (\lambda I_n - M) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0.$$

Parfois, on n'a pas besoin de déterminer précisément les sous-espaces propres, mais seulement de calculer leurs dimensions. Dans ce cas, il est commode d'utiliser la formule suivante :

$$\forall \lambda \in Sp(M) \quad \boxed{\dim(E_\lambda) = n - rg(\lambda I_n - M)}.$$

Démonstration.

D'après la formule du rang,

$$\dim(E_\lambda) = \dim(Ker(\lambda I_n - M)) = n - \dim(Im(\lambda I_n - M)) = n - rg(\lambda I_n - M). \quad \square$$

Propriété. Soit F un sous-espace vectoriel de E stable par u .

$$\boxed{\text{Si } u|_F \text{ est l'endomorphisme induit par } u \text{ sur } F, \text{ alors } \chi_{u|_F} | \chi_u.}$$

Démonstration.

Choisissons $e' = (e_1, \dots, e_p)$ une base de F que l'on complète en une base $e = (e_1, \dots, e_n)$ de E . Si l'on note $M = Mat(u, e)$ et $M' = Mat(u|_F, e')$, il existe deux matrices $A \in \mathcal{M}_{p, n-p}(\mathbb{K})$ et $B \in \mathcal{M}_{n-p, n-p}(\mathbb{K})$ telles que $M = \begin{pmatrix} M' & A \\ 0_{n-p, p} & B \end{pmatrix}$.

$$\text{Soit } \lambda \in \mathbb{K}. \chi_u(X) = \det(XI_n - M) = \det \begin{pmatrix} XI_p - M' & -A \\ 0_{n-p, p} & XI_{n-p} - B \end{pmatrix} = \chi_{u|_F}(X) \chi_B(X).$$

\square

Propriété. Soit (E_1, \dots, E_p) une famille de sous-espaces vectoriels de E telle que $E = \bigoplus_{i=1}^p E_i$. On suppose que u stabilise la famille (E_1, \dots, E_p) . Pour tout $i \in \mathbb{N}_p$, on

note u_i l'endomorphisme induit par u sur E_i . Alors $\chi_u = \prod_{i=1}^p \chi_{u_i}$.

Démonstration.

Pour tout $i \in \mathbb{N}_p$, on choisit une base de E_i notée e_i .

Notons e la “réunion” des e_i , pour i variant de 1 à p . e est une base de E .

On sait que $Mat(u, e)$ est diagonale par blocs, la “diagonale” étant constituée des p blocs suivants :

$$M_1 = mat(u_1, e_1), \dots, M_p = mat(u_p, e_p).$$

Le déterminant d’une matrice diagonale par blocs étant égal au produit des déterminants des blocs diagonaux, pour tout $\lambda \in \mathbb{K}$,

$$\chi_u(X) = \det(XI_n - M) = \prod_{i=1}^p \det(XI_{n_i} - M_i), \text{ où pour tout } i \in \mathbb{N}_p, n_i = \dim(E_i).$$

$$\text{Ainsi, } \chi_u(X) = \prod_{i=1}^p \chi_{u_i}(X). \quad \square$$

Notation. Pour tout $\lambda \in E_\lambda$, on note $q(\lambda) = \dim(E_\lambda)$.

Propriété. $\forall \lambda \in Sp(u) \quad 1 \leq q(\lambda) \leq m(\lambda).$

Démonstration.

- Soit $\lambda \in Sp(u)$. $E_\lambda \neq \{0\}$, donc $1 \leq q(\lambda)$.
- u commute avec lui-même, donc E_λ est stable par u . Notons u' l’endomorphisme induit par u sur E_λ .

Pour tout $x \in E_\lambda$, $u(x) = \lambda x$, donc $u' = \lambda Id_{E_\lambda}$.

Soit $\mu \in \mathbb{K}$. $\chi_{u'}(\mu) = \det(\mu Id_{E_\lambda} - u') = \det(\mu I_{q(\lambda)} - \lambda I_{q(\lambda)}) = (\mu - \lambda)^{q(\lambda)}$.

Ainsi, $\chi_{u'}(X) = (X - \lambda)^{q(\lambda)}$. Or $\chi_{u'} | \chi_u$, donc $m(\lambda) \geq q(\lambda)$. \square

Cas particulier. Si λ est une valeur propre simple de u , $1 = q(\lambda) = m(\lambda)$.

7.7.3 Caractérisation des endomorphismes diagonalisables

Théorème. u est diagonalisable si et seulement si χ_u est scindé sur \mathbb{K} et, pour tout $\lambda \in Sp(u)$, $m(\lambda) = q(\lambda)$.

Démonstration.

- Supposons que u est diagonalisable. Ainsi, il existe une base e de E dans laquelle la matrice de u est diagonale. Notons $\lambda_1, \dots, \lambda_n$ les coefficients diagonaux de cette matrice.

$$\diamond \quad \chi_u = \chi_{Mat(u, e)} = \prod_{i=1}^n (X - \lambda_i), \text{ donc } \chi_u \text{ est scindé sur } \mathbb{K}.$$

\diamond Soit $\lambda \in Sp(u)$. L’égalité précédente montre que $m(\lambda) = \text{Card}(\{i \in \mathbb{N}_n / \lambda_i = \lambda\})$.

Or, pour tout $i \in \mathbb{N}_n$, $e_i \in E_{\lambda_i}$, donc $Vect(\{e_i / \lambda_i = \lambda\}) \subset E_\lambda$.

Ainsi, $m(\lambda) = \dim(Vect(\{e_i / \lambda_i = \lambda\})) \leq \dim(E_\lambda) = q(\lambda)$.

L’inégalité contraire étant vraie pour tout endomorphisme, on a montré que, pour tout $\lambda \in Sp(u)$, $m(\lambda) = q(\lambda)$.

- Réciproquement, supposons que χ_u est scindé sur \mathbb{K} et que, pour tout $\lambda \in Sp(u)$, $m(\lambda) = q(\lambda)$.

Alors, $\sum_{\lambda \in Sp_{\mathbb{K}}(u)} q(\lambda) = \sum_{\lambda \in Sp_{\mathbb{K}}(u)} m(\lambda) = \deg(\chi_u) = n$, car χ_u est scindé sur \mathbb{K} , ce qui prouve que u est diagonalisable. \square

Cas particulier.

Si χ_u est scindé sur \mathbb{K} et si toutes ses racines sont simples, alors u est diagonalisable.

Démonstration.

Lorsqu'une valeur propre λ est simple, on a déjà établi que $m(\lambda) = q(\lambda)$. \square