Algèbre 3

Benjamin Schraen

1 Espaces vectoriels et applications linéaires

Dans ce cours, la lettre K désigne l'ensemble $\mathbb{Q}$ des nombres rationnels ou l'ensemble $\mathbb{R}$ des nombres réels ou l'ensemble $\mathbb{C}$ des nombres complexes.

1.1 Espaces vectoriels

Soit $n \in \mathbb{N}$ un entier. On note $\mathbb{R}^n$ l'ensemble des n-uplets $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ où $x_1, \dots, x_n$ sont des nombres réels. Cet ensemble est muni de deux opérations

— l'addition : si $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $w = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, on pose $v + w = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$;

— la multiplication par un scalaire $\lambda \in \mathbb{R}$: si $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, on pose $\lambda \cdot v = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$.

Ces opérations ont les propriétés suivantes. Notons, pour des raisons de référence dans la suite, $K = \mathbb{R}$, $E = \mathbb{R}^n$ et $0_E = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$. On vérifie alors les propriétés suivantes.

$$\begin{array}{ccc} E \times E & \longrightarrow & E \\ (v, w) & \longmapsto & v + w \end{array} \qquad \begin{array}{ccc} K \times E & \longrightarrow & E \\ (\lambda, v) & \longmapsto & \lambda \cdot v. \end{array}$$

Ces deux lois doivent vérifier les propriétés suivantes

a) La loi + est *associative* :

$$\forall u, v, w \in E, \quad (u + v) + w = u + (v + w).$$

b) La loi + est *commutative* :

$$\forall v, w \in E, \quad v + w = w + v.$$

c) Il existe un unique élément neutre $0_E \in E$ pour la loi $+$:

$$\forall v \in E, \quad v + 0_E = 0_E + v = v.$$

d) Tout élément $v \in E$ possède pour symétrique $-v = (-1) \cdot v \in E$:

$$v + (-v) = (-v) + v = 0_E.$$

e) La loi $\cdot$ est distributive par rapport aux lois $+$ de E et K :

$$\begin{gathered}\forall \lambda \in K,\ \forall v, w \in E, \quad \lambda \cdot (v + w) = (\lambda \cdot v) + (\lambda \cdot w)\\ \forall \lambda, \mu \in K,\ \forall v E, \quad (\lambda + \mu) \cdot v = (\lambda \cdot v) + (\mu \cdot v).\end{gathered}$$

f) La loi $\cdot$ est compatible à la multiplication dans K :

$$\forall \lambda, \mu \in K,\ \forall v \in E, \quad \lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v, \quad 1 \cdot v = v.$$

Voici un autre exemple. Rappelons que K désigne $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$.

Définition 1.1. *Soient $m \geqslant 1$ et $n \geqslant 1$ deux entiers. Une* matrice *de taille $m \times n$ à coefficients dans K est un tableau rectangulaire A ayant m lignes et n colonnes contenant des éléments de A. On note $a_{i,j}$ ses coefficients et on les indexe de la façon suivante*

$$A = (a_{i,j})_{\substack{1 \leqslant i \leqslant m\\ 1 \leqslant j \leqslant n}} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n}\\ a_{2,1} & a_{2,2} & \ddots & a_{2,n}\\ \vdots & \ddots & \ddots & \vdots\\ a_{m,1} & a_{m,2} & \cdots & a_{m,n}\end{pmatrix}.$$

On note $\mathcal{M}_{m,n}(K)$ l'ensemble des matrices à m lignes et n colonnes. On définit les opérations suivantes :

— $A = (a_{i,j})_{\substack{1 \leqslant i \leqslant m\\ 1 \leqslant j \leqslant n}}$, $B = (b_{i,j})_{\substack{1 \leqslant i \leqslant m\\ 1 \leqslant j \leqslant n}}$, alors $C = A + B$ où $C = (a_{i,j} + b_{i,j})_{\substack{1 \leqslant i \leqslant m\\ 1 \leqslant j \leqslant n}}$.

— $A = (a_{i,j})_{\substack{1 \leqslant i \leqslant m\\ 1 \leqslant j \leqslant n}}$, $\lambda \in K$, alors $\lambda \cdot A = (\lambda a_{i,j})_{\substack{1 \leqslant i \leqslant m\\ 1 \leqslant j \leqslant n}}$.

Alors en posant $E = \mathcal{M}_{m,n}(K)$, les propriétés a) à f) sont vérifiées.

Si $m = n$, une matrice de $\mathcal{M}_n(K) = \mathcal{M}_{m,n}(K)$ est appelée *matrice carrée* de taille n.

Définition 1.2. *On appelle K-*espace vectoriel*, ou simplement* espace vectoriel*, un ensemble E muni de deux opérations*

$$\begin{array}{ccc} E \times E & \longmapsto & E\\ (u, v) & \longmapsto & u + v \end{array} \qquad \begin{array}{ccc} K \times E & \longrightarrow & E\\ (\lambda, v) & \longmapsto & \lambda \cdot v \end{array}$$

vérifiant les propriétés a) à f).

Les éléments de l'ensemble E sont appelés *vecteurs* et les éléments du corps K sont appelés *scalaires*. Il faut bien prendre garde au fait que les lois $+$ et $\cdot$ sont définies sur des ensembles différents. La loi $+$ part de deux vecteurs et produit un vecteur alors que la loi $\cdot$ part d'un scalaire et d'un vecteur et produit un vecteur. Si λ est un scalaire et v un vecteur, il faut imaginer le vecteur $\lambda \cdot v$ comme étant le vecteur v dilaté au moyen du coefficient λ.

Remarque 1.3. Soit E un K-espace vectoriel. Si $v \in E$, on a $0 \cdot v + 0 \cdot v = (0+0) \cdot v = 0 \cdot v$. On en déduit $0 \cdot v = 0_E$ pour tout $v \in E$.

Si $(v_1, \dots, v_n)$ est une famille d'éléments de E, on appelle *combinaison linéaire* de $v_1, \dots, v_n$ un élément de la forme $\lambda_1 \cdot v_1 + \cdots + \lambda_n \cdot v_n$ pour $\lambda_1, \dots, \lambda_n \in K$.

1.2 Bases et coordonnées

Soit E un espace vectoriel.

Définition 1.4. *On appelle* famille génératrice *de E une famille finie de vecteurs $(v_1, \dots, v_n)$ telle que tout vecteur de E est combinaison linéaire de $v_1, \dots, v_n$. On dit qu'un espace vectoriel est de* dimension finie *s'il possède une famille génératrice finie.*

Exemple 1.5. Si $E = K^n$, pour $1 \leqslant i \leqslant n$, notons e_i l'élément $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ où 1 apparaît sur la i-ième ligne. Tout élément de E peut s'écrire

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \cdot e_1 + x_2 e_2 + \cdots + x_n \cdot e_n$$

donc $(e_1, \dots, e_n)$ est une famille génératrice de K^n.

Dans l'exemple précédent, on a envie de dire que K^n est en fait de dimension n. Pour définir correctement la notion de dimension, il est nécessaire d'introduire les notions de familles libres et de bases.

Définition 1.6. *Soit $(v_1, \dots, v_n)$ une famille finie d'éléments de E. On dit que la famille $(v_1, \dots, v_n)$ est* libre *si pour tout n-uplet de scalaires $(\lambda_1, \dots, \lambda_n) \in K^n$, on a*

$$\lambda_1 \cdot v_1 + \cdots + \lambda_n \cdot v_n = 0_E \Rightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0.$$

Autrement dit, la seule combinaison linéaire nulle des vecteurs $v_1, \dots, v_n$ est la combinaison dont les coefficients sont nuls.

Exemple 1.7. Si $E = K^n$, la famille $(e_1, \dots, e_n)$ est libre. En effet, supposons que $\lambda_1 \cdot e_1 + \cdots + \lambda_n \cdot e_n = 0_{\mathbb{R}^n}$, alors

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 \cdot e_1 + \cdots + \lambda_n \cdot e_n = 0_{\mathbb{R}^n} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

donc $\lambda_1 = \cdots = \lambda_n = 0$.

Définition 1.8. *Une* base *de E est une famille finie d'éléments de E qui est à la fois libre et génératrice.*

Exemple 1.9. Dans $\mathbb{R}^n$, la famille $(e_1, \dots, e_n)$ est une base appelée *base canonique.*

Soit $\mathcal{B} = (v_1, \dots, v_n)$ une base de E et soit $v \in E$. Comme la famille $\mathcal{B}$ est génératrice, il existe des scalaires $x_1, \dots, x_n$ tels que $v = x_1 \cdot v_1 + \cdots + x_n \cdot v_n$. Comme de plus la famille $\mathcal{B}$ est libre, les scalaires $x_1, \dots, x_n$ sont uniquement déterminés par v. On les appelle les *coordonnées* de v dans la base $\mathcal{B}$.

Il peut être commode, du point de vue du calcul matriciel, de noter les coordonnées d'un vecteur sous forme de vecteur colonne. Si E est un espace vectoriel, $\mathcal{B}$ une base de E et $v \in E$, on note

$$[v]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

le vecteur colonne des coordonnées de v dans la base $\mathcal{B}$.

Exemple 1.10. Notons $\mathcal{B}_{\text{can}} = (e_1, \dots, e_n)$ la base canonique de K^n. Si $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, on a

$$v = x_1 \cdot e_1 + \cdots x_n \cdot e_n$$

donc les scalaires $x_1, \dots, x_n$ sont les coordonnées du vecteur $(x_1, \dots, x_n)$ dans la base $\mathcal{B}_{\text{can}}$. Ainsi on a

$$\left[\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right]_{\mathcal{B}_{\text{can}}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Mais attention, il existe bien d'autres bases dans $\mathbb{R}^n$ et cette égalité ne vaut que pour la base canonique !

Exemple 1.11. Soit $E = K^2$. Posons $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $v = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. La famille (u, v) est libre. En effet, si $x \cdot u + y \cdot v = 0_{K^2}$, alors

$$\begin{pmatrix} x + y \\ x - y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} x + y & = 0 \\ x - y & = 0 \end{cases} \Leftrightarrow \begin{cases} x & = 0 \\ y & = 0. \end{cases}$$

De plus si $x, y \in K$, on a

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{x+y}{2} \cdot u + \frac{x-y}{2} \cdot v.$$

Ainsi la famille (u, v) est génératrice. C'est donc une base de K^2. De plus, les coordonnées du vecteur $\begin{pmatrix} x \\ y \end{pmatrix}$ dans la base (u, v) sont x, y. On a donc

$$\left[\begin{pmatrix} x \\ y \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} \frac{x+y}{2} \\ \frac{x-y}{2} \end{pmatrix}.$$

Exemple 1.12. Soit $m, n \geqslant 1$ deux entiers. On note $E_{i,j} \in \mathcal{M}_{m,n}(K)$ la matrice dont toutes les entrées sont nulles, sauf l'entrée sur la i-ème ligne et j-ème colonne qui vaut 1. Par exemple, si $m = 2$ et $n = 3$,

$$E_{1,1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_{2,1} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad E_{2,3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

La famille $\mathcal{B} = (E_{1,1}, E_{1,2}, E_{1,3}, E_{2,1}, E_{2,2}, E_{2,3})$ est une base de $\mathcal{M}_{2,3}(K)$ et on a

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix} = a_{1,1} \cdot E_{1,1} + a_{1,2} \cdot E_{1,2} + \cdots + a_{2,3} \cdot E_{2,3}.$$

Ainsi $\left[\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} a_{1,1} \\ a_{1,2} \\ a_{1,3} \\ a_{2,1} \\ a_{2,2} \\ a_{2,3} \end{pmatrix}.$

Exemple 1.13. Soit $K[X]$ l'ensemble des polynômes en une variable à coefficients dans K. Si $n \geqslant 0$, on note $K_n[X]$ l'ensemble des polynômes de degré inférieur ou égal à n. La famille $\mathcal{B} = (1, X, X^2 \ldots, X^n)$ est une base de $K_n[X]$. De plus on a

$$[a_0 + a_1 X + \cdots + a_n X^n]_{\mathcal{B}} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Théorème 1.14. *Soit E un K-espace vectoriel de dimension finie.*

1) *Il existe au moins une base de E.*
2) *Toutes les bases de E ont le même nombre d'éléments.*
3) *On peut toujours compléter une famille libre de E en une base.*
4) *On peut toujours extraire une base d'une famille génératrice de E.*

Si E est un K-espace vectoriel de dimension finie. La longueur d'une base de E est appelée *dimension* de E. On la note $\dim_K E$ (ou encore $\dim E$ lorsque le corps des scalaires est défini sans ambiguïté).

Exemple 1.15. L'exemple 1.9 montre que la famille des vecteurs élémentaires $(e_1, \ldots, e_n)$ est une base de K^n, on en conclut que $\dim_K K^n = n$.

Exemple 1.16. Dans l'exemple 1.12, on a vu que la famille $(E_{1,1}, \ldots, E_{2,3})$ est une base d e$\mathcal{M}_{2,3}(K)$. Ainsi $\dim_K \mathcal{M}_{2,3}(K) = 6$. Plus généralement, on a $\dim_K \mathcal{M}_{m,n}(K) = mn$.

Exemple 1.17. Dans l'exemple 1.13, on a vu que la base $(1, X, X^2, \ldots, X^n)$ est une base du K-espace vectoriel $K_n[X]$. On en conclut que $\dim_K K_n[X] = n + 1$.

1.3 Produit matriciel, changement de base

Rappelons que l'on peut multiplier les matrices.

Si $A \in \mathcal{M}_{m,n}(K)$ et $B \in \mathcal{M}_{n,p}(K)$, on définit le *produit* de A avec B par

$$AB = (c_{i,j})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant p}} \in \mathcal{M}_{m,p}(K)$$

où $c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$.

Proposition 1.18. *Si $A \in \mathcal{M}_{n,p}(K)$, $B \in \mathcal{M}_{p,q}(K)$ et $C \in \mathcal{M}_{q,r}(K)$, on a*

- *$A(BC) = (AB)C$;*
- *$A(B + C) = (AB) + (AC)$;*
- *$(A + B)C = (AC) + (BC)$;*
- *si $\lambda \in K$, on a $\lambda(AB) = (\lambda A)B = A(\lambda B)$.*

Si $n \geqslant 1$, on note I_n la matrice $(a_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ de $\mathcal{M}_n(K)$ définie par $a_{i,j} = 0$ si $i \neq j$ et $a_{i,j} = 1$ si $i = j$. On l'appelle la *matrice identité*. On vérifie que $AI_n = A = I_m A$ pour toute matrice $A \in \mathcal{M}_{m,n}(K)$.

Remarque 1.19. Le produit matriciel n'est pas commutatif. Voici un exemple dans $\mathcal{M}_2(K)$:

$$\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}.$$

Définition 1.20. *Une matrice $A \in \mathcal{M}_n(K)$ est dite* inversible *s'il existe une matrice $B \in \mathcal{M}_n(K)$ telle que $AB = BA = I_n$. Si elle existe, la matrice B est unique et est appelée* inverse de A. *On la note alors A^{-1}.*

Si A et B sont inversibles alors AB est aussi inversible et

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Soit E un K-espace vectoriel de dimension finie n. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E. La *matrice de passage* de $\mathcal{B}$ à $\mathcal{B}'$ est la matrice carrée de taille n dont la j-ème colonne est donnée par les coordonnées du j-ème vecteur de $\mathcal{B}'$ dans la base $\mathcal{B}$. Autrement dit, si $\mathcal{B} = (v_1, \dots, v_n)$ et $\mathcal{B}' = (v'_1, \dots, v'_n)$, et que

$$\forall 1 \leqslant j \leqslant n,\ v'_j = \sum_{i=1}^{n} a_{i,j} v_i,$$

alors

$$P = P_{\mathcal{B}}^{\mathcal{B}'} = P_{\mathcal{B} \leftarrow \mathcal{B}'} = (a_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} = \begin{pmatrix} [v'_1]_{\mathcal{B}} & \cdots & [v'_n]_{\mathcal{B}} \end{pmatrix}.$$

Proposition 1.21. *Si $v \in E$, on a alors*

$$[v]_{\mathcal{B}} = P_{\mathcal{B} \leftarrow \mathcal{B}'} [v]_{\mathcal{B}'}.$$

Démonstration. En effet, supposons que $[v]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $[v]_{\mathcal{B}'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$, c'est-à-dire $v = \sum_{i=1}^{n} x_i v_i = \sum_{j=1}^{n} x'_j v'_j$.

Comme $v'_j = \sum_{i=1}^{n} a_{i,j} v_i$, on en conclut

$$\sum_{i=1}^{n} x_i v_i = \sum_{j=1}^{n} x'_j \sum_{i=1}^{n} a_{i,j} v_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{i,j} x'_j \right) v_i.$$

Comme $(v_1, \dots, v_n)$ est une base de E, on a bien

$$\forall 1 \leqslant i \leqslant n, \qquad x_i = \sum_{j=1}^{n} a_{i,j} x'_j$$

c'est-à-dire $[v]_{\mathcal{B}} = P_{\mathcal{B} \leftarrow \mathcal{B}'} [v]_{\mathcal{B}'}$. □

Si $\mathcal{B}$, $\mathcal{B}'$ et $\mathcal{B}''$ sont trois bases de E, on a alors

$$P_{\mathcal{B} \leftarrow \mathcal{B}''} = P_{\mathcal{B} \leftarrow \mathcal{B}'} P_{\mathcal{B}' \leftarrow \mathcal{B}''}.$$

Remarque 1.22. Si $\mathcal{B} = \mathcal{B}'$, on a $P_{\mathcal{B} \leftarrow \mathcal{B}} = I_n$. On a donc, en général

$$P_{\mathcal{B} \leftarrow \mathcal{B}'} P_{\mathcal{B}' \leftarrow \mathcal{B}} = P_{\mathcal{B} \leftarrow \mathcal{B}} = I_n.$$

Ainsi la matrice $P_{\mathcal{B} \leftarrow \mathcal{B}'}$ est inversible et $P_{\mathcal{B} \leftarrow \mathcal{B}'}^{-1} = P_{\mathcal{B}' \leftarrow \mathcal{B}}$.

1.4 Applications linéaires

Soient E et F deux espaces vectoriels.

Définition 1.23. *Une* application linéaire *de E dans F est une application f de E dans F telle que*

- $\forall v, w,\ f(v+w) = f(v) + f(w)$;
- $\forall \lambda \in K,\ \forall v \in E,\ f(\lambda \cdot v) = \lambda \cdot f(v)$.

Si E et F sont deux espaces vectoriels, on note $\mathcal{L}(E,F)$ l'ensemble des applications linéaires de E dans F.

Lorsque $E = F$, on note $\mathcal{L}(E) = \mathcal{L}(E,E)$. Un élément de $\mathcal{L}(E)$ est appelé *endomorphisme* de E.

Exemple 1.24. L'application $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+y \\ y+2x \end{pmatrix}$ est linéaire.

Si f et g sont deux applications linéaires de E dans F, on note $f+g$ l'application de E dans F définie par

$$\forall v \in E,\ (f+g)(v) = f(v) + g(v).$$

Il s'agit d'une application linéaire de E dans F. De même si $\lambda \in K$, on note $\lambda \cdot f$ l'application de E dans F définie par

$$\forall v \in E,\ (\lambda \cdot f)(v) = \lambda \cdot f(v).$$

Il s'agit encore d'une application linéaire de E dans F. Muni des opérations $+$ et $\cdot$ définies ci-dessus, l'ensemble $\mathcal{L}(E,F)$ est un espace vectoriel.

Soient E et F deux K-espaces vectoriels de dimension finie. On fixe $\mathcal{B}_E = (v_1, \ldots, v_n)$ une base de E et $\mathcal{B}_F = (w_1, \ldots, w_m)$ une base de F. Soit $f \in \mathcal{L}(E,F)$ une application linéaire de E dans F. Pour $1 \leqslant j \leqslant n$, on note $(a_{i,j})_{1\leqslant i\leqslant m}$ les coordonnées de $f(v_j)$ dans la base $\mathcal{B}_F$. Autrement dit

$$f(v_j) = a_{1,j}w_1 + \cdots + a_{m,j}w_m \quad \text{ou encore} \quad [f(v_j)] = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{pmatrix}.$$

La matrice de taille $m \times n$ de termes $(a_{i,j})_{\substack{1\leqslant i\leqslant m \\ 1\leqslant j\leqslant n}}$ est appelée *matrice de f dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$*. On la note $\mathrm{Mat}_{(\mathcal{B}_E,\mathcal{B}_F)}(f)$. Si $E = F$ et $\mathcal{B}_E = \mathcal{B}_F$, on note $\mathrm{Mat}_{\mathcal{B}_E}(f) = \mathrm{Mat}_{(\mathcal{B}_E,\mathcal{B}_E)}(f)$.

Remarque 1.25. D'autres notations sont parfois utilisées pour la matrice $\mathrm{Mat}_{(\mathcal{B}_E,\mathcal{B}_F)}(f)$. On peut la noter $\mathrm{Mat}_{\mathcal{B}_E\to\mathcal{B}_F}(f)$ ou encore $\mathrm{Mat}_{\mathcal{B}_F\leftarrow\mathcal{B}_E}(f)$. La dernière notation est particulièrement bien adaptée aux formules de changement de base et de composition.

Proposition 1.26. *Soit $f \in \mathcal{L}(E,F)$. Soit $\mathcal{B}_E$ une base de E et soit $\mathcal{B}_F$ une base de F. Alors, pour tout $v \in E$, on a*

$$[f(v)]_{\mathcal{B}_F} = \mathrm{Mat}_{(\mathcal{B}_E,\mathcal{B}_F)}(f)[v]_{\mathcal{B}_E}.$$

Remarque 1.27. On note Id_E l'endomorphisme *identité* de l'espace vectoriel E, c'est l'application linéaire de E dans E définie simplement par $\mathrm{Id}_E(v) = v$ pour tout $v \in E$. Si $\mathcal{B}$ est une base de E, on vérifie facilement que $\mathrm{Mat}_{\mathcal{B}}(\mathrm{Id}_E) = I_n$. Il faut prendre garde au fait que si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E (potentiellement différentes), alors

$$\mathrm{Mat}_{(\mathcal{B},\mathcal{B}')}(\mathrm{Id}_E) = P_{\mathcal{B}'\leftarrow\mathcal{B}}(= \mathrm{Mat}_{\mathcal{B}'\leftarrow\mathcal{B}}(\mathrm{Id}_E)).$$

1.5 Composition des applications linéaires

Soient E, F et G trois espaces vectoriels. Si $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$, l'application composée $g \circ f$ est une application linéaire de E dans G.

Proposition 1.28. *Soient E, F et G trois espaces vectoriels de dimension finie. Soient $\mathcal{B}_E$ une base de E, $\mathcal{B}_F$ une base de F et $\mathcal{B}_G$ une base de G. Alors si $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$, on a*

$$\mathrm{Mat}_{(\mathcal{B}_E,\mathcal{B}_G)}(g \circ f) = \mathrm{Mat}_{(\mathcal{B}_F,\mathcal{B}_G)}(g)\,\mathrm{Mat}_{(\mathcal{B}_E,\mathcal{B}_F)}(f).$$

Ou encore, en notations alternatives

$$\mathrm{Mat}_{\mathcal{B}_G\leftarrow\mathcal{B}_E}(g \circ f) = \mathrm{Mat}_{\mathcal{B}_G\leftarrow\mathcal{B}_F}(g)\,\mathrm{Mat}_{\mathcal{B}_F\leftarrow\mathcal{B}_E}(f).$$

Définition 1.29. *Une application linéaire f de E dans F est appelée* isomorphisme *si elle est bijective.*

Si f est un isomorphisme de E dans F son application réciproque f^{-1} de F dans E est linéaire et bijective et est donc également un isomorphisme. Rappelons que l'application réciproque f^{-1} d'une application bijective est l'unique application telle que $f \circ f^{-1} = \mathrm{Id}_F$. On a alors, de façon équivalente, $f^{-1} \circ f = \mathrm{Id}_E$.

Si f est un isomorphisme de E dans F, si $\mathcal{B}_E$ est une base de E et $\mathcal{B}_F$ une base de F, on a alors

$$\mathrm{Mat}_{(\mathcal{B}_E,\mathcal{B}_F)}(f)\,\mathrm{Mat}_{(\mathcal{B}_F,\mathcal{B}_E)}(f^{-1}) = \mathrm{Mat}_{\mathcal{B}_F}(\mathrm{Id}_E) = I_n.$$

Ainsi $\mathrm{Mat}_{(\mathcal{B}_{\mathcal{E}},\mathcal{B}_F)}(f)$ est inversible et son inverse est $\mathrm{Mat}_{(\mathcal{B}_{\mathcal{F}},\mathcal{B}_E)}(f^{-1})$. La réciproque est vraie également.

Proposition 1.30. *Soit $\mathcal{B}_E$ une base de E et soit $\mathcal{B}_F$ une base de F. Une application linéaire f de E dans F est inversible si et seulement si $\mathrm{Mat}_{(\mathcal{B}_E,\mathcal{B}_F)}(f)$ est inversible. Dans ce cas*

$$\mathrm{Mat}_{(\mathcal{B}_E,\mathcal{B}_F)}(f)^{-1} = \mathrm{Mat}_{(\mathcal{B}_F,\mathcal{B}_E)}(f^{-1}).$$

Proposition 1.31. *Soient $\mathcal{B}_E$ et $\mathcal{B}'_E$ deux bases de E et $\mathcal{B}_F$ et $\mathcal{B}'_F$ deux bases de F. Si $f \in \mathcal{L}(E,F)$, on a*

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

En particulier, si $E = F$, $\mathcal{B}_E = \mathcal{B}_{\mathcal{F}}$ et $\mathcal{B}'_E = \mathcal{B}'_F$, on a

$$\mathrm{Mat}_{\mathcal{B}'_E}(f) = P^{-1}_{\mathcal{B}_E \leftarrow \mathcal{B}'_E} \mathrm{Mat}_{\mathcal{B}_E}(f) P_{\mathcal{B}_E \leftarrow \mathcal{B}'_E}.$$

Démonstration. On peut en effet écrire

$$\begin{aligned}
\mathrm{Mat}_{(\mathcal{B}'_E,\mathcal{B}'_F)}(f) &= \mathrm{Mat}_{\mathcal{B}'_F \leftarrow \mathcal{B}'_E}(\mathrm{Id}_F \circ f \circ \mathrm{Id}_E) \\
&= \mathrm{Mat}_{\mathcal{B}'_F \leftarrow \mathcal{B}_F}(\mathrm{Id}_F) \, \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f) \, \mathrm{Mat}_{\mathcal{B}_E \leftarrow \mathcal{B}'_E}(\mathrm{Id}_E) \\
&= P_{\mathcal{B}'_F \leftarrow \mathcal{B}_F} \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(f) P_{\mathcal{B}_E \leftarrow \mathcal{B}'_E} \\
&= P^{-1}_{\mathcal{B}_F \leftarrow \mathcal{B}'_F} \mathrm{Mat}_{(\mathcal{B}_E,\mathcal{B}_F)}(f) P_{\mathcal{B}_E \leftarrow \mathcal{B}'_E}
\end{aligned}$$

où l'on a utilisé les résultats de la proposition 1.28 et des remarques 1.27 et 1.22. □

1.6 Sous-espaces vectoriels

Définition 1.32. *Soit E un K-espace vectoriel. Un sous-K-espace vectoriel de E (ou simplement sous-espace vectoriel lorsque K est sous-entendu) est une partie F de E vérifiant*

(i) $0_E \in F$;

(ii) pour tous v et w dans F, on a $v + w \in F$;

(iii) pour tout v dans F et tout λ dans K, on a $\lambda v \in F$.

Exemple 1.33. Dans $\mathbb{R}^3$, l'ensemble des éléments $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ vérifiant la relation $2x_1 + 3x_2 + 5x_3 = 0$ est un sous-espace vectoriel. Il s'agit d'un exemple de *sous-espace vectoriel défini par une équation.*

Exemple 1.34. Dans $\mathbb{R}^3$ posons $v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ et $w = \begin{pmatrix} 2 \\ 7 \\ 3 \end{pmatrix}$. Alors l'ensemble

$$\{\lambda v + \mu w \mid (\lambda, \mu) \in \mathbb{R}^2\} = \left\{ \begin{pmatrix} \lambda + 2\mu \\ 2\lambda + 7\mu \\ 3\lambda + 3\mu \end{pmatrix} \mid (\lambda, \mu) \in \mathbb{R}^2 \right\}$$

est un sous-espace vectoriel de $\mathbb{R}^3$. Il s'agit d'un exemple de sous-espace vectoriel engendré par une famille de vecteurs, ici les vecteurs v et w.

Plus généralement si $v_1, \ldots, v_n \in E$, on note $\mathrm{Vect}(v_1, \ldots, v_n)$ l'ensemble des combinaisons linéaires des vecteurs $v_1, \ldots, v_n$. C'est un sous-espace vectoriel de E appelé *sous-espace engendré* par $v_1, \ldots, v_n$.

Définition 1.35. *Si E et F sont deux espaces vectoriels et si f est une application linéaire de E dans F, on note* $\mathrm{Im}\, f$ *l'image $f(E)$ de f et* $\mathrm{Ker}\, f$ *le* noyau *de f, c'est-à-dire l'ensemble $f^{-1}(\{0_F\}) = \{x \in E \mid f(x) = 0_F\}$.*

On vérifiera à titre d'exercice que $\mathrm{Ker}\, f$ est un sous-espace vectoriel de E et que $\mathrm{Im}\, f$ est un sous-espace vectoriel de F.

Proposition 1.36. *Si f est une application linéaire entre deux espaces vectoriels, alors f est injective si et seulement si* $\mathrm{Ker}\, f = \{0_E\}$.

Démonstration. Si f est injective, alors $\mathrm{Ker}\, f$ est réduit à $\{0_E\}$. Réciproquement supposons que $\mathrm{Ker}\, f = \{0_E\}$. Si x et y sont deux éléments de E tels que $f(x) = f(y)$, alors $f(x) - f(y) = 0_F$. Par linéarité de f, on a alors $f(x - y) = 0_F$ et donc $x - y \in \mathrm{Ker}\, f$. Ainsi on doit avoir $x - y = 0_E$, c'est-à-dire $x = y$. On a prouvé que l'application f est injective. $\square$

La dimension de l'espace vectoriel $\mathrm{Im}\, f$ est appelé le *rang* de f et est notée $\mathrm{rg}\, f$. Ainsi on a, par définition, $\mathrm{rg}\, f = \dim \mathrm{Im}\, f$.

Théorème 1.37 (Théorème du rang)**.** *Soit E un espace vectoriel de dimension finie et soit F un espace vectoriel. Si f est une application linéaire de E dans F, alors* $\mathrm{Im}\, f$ *est un sous-espace vectoriel de dimension finie de F et on a*

$$\dim \mathrm{Im}\, f = \dim E - \dim \mathrm{Ker}\, f.$$

Exemple 1.38. Considérons l'application linéaire $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x + y \\ 2x + y \end{pmatrix}.$$

Déterminons son noyau. Un élément $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ est dans le noyau de f si et seulement si il est solution du système d'équations linéaires homogène

$$\begin{cases} x + y &= 0 \\ 2x + y &= 0 \end{cases} \Leftrightarrow \begin{cases} x + y &= 0 \\ x &= 0 \end{cases} \Leftrightarrow \begin{cases} x &= 0 \\ y &= 0. \end{cases}$$

Ainsi $\mathrm{Ker}\, f = 0$ et l'application f est injective. Le théorème du rang (théorème 1.37) nous donne alors $\mathrm{rg}\, f = 2$.

1.7 Sommes de sous-espaces vectoriels

On peut effectuer des opérations sur les sous-espaces vectoriels afin d'en produire de nouveaux.

Si E est un espace vectoriel et si F_1 et F_2 sont deux sous-espaces vectoriels, l'intersection $F_1 \cap F_2$ est un sous-espace vectoriel de E. L'ensemble

$$F_1 + F_2 = \{v + w \mid (v, w) \in F_1 \times F_2\}$$

est un sous-espace vectoriel de E appelé *somme* des sous-espaces F_1 et F_2. Plus généralement, si $F_1, \ldots, F_n$ sont des sous-espaces vectoriels, la partie

$$F_1 + \cdots + F_n = \{v_1 + \cdots + v_n \mid (v_1, \ldots, v_n) \in F_1 \times \cdots \times F_n\}$$

est un sous-espace vectoriel de E appelé *somme* de la famille de sous-espaces $(F_1, \ldots, F_n)$.

Soit E un K-espace vectoriel, ainsi que F_1 et F_2 deux sous-espaces vectoriels de E. Par définition, tout élément de $F_1 + F_2$ s'écrit sous la forme $v + w$ avec $v \in F_1$ et $w \in F_2$. Cette écriture n'est pas toujours unique. Considérons par exemple le cas de $E = \mathbb{R}^2$ avec $F_1 = \mathrm{Vect}\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right)$ et $F_2 = \mathrm{Vect}\left(\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right)$. Alors on peut écrire

$$\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}1\\1\end{pmatrix} = -\begin{pmatrix}0\\1\end{pmatrix} + 2\begin{pmatrix}1\\1\end{pmatrix}.$$

Cependant lorsque cette décomposition est unique, on dit que les sous-espaces F_1 et F_2 sont en *somme directe*. Plus généralement, on peut définir la notion de somme directe pour une famille finie de sous-espaces vectoriels.

Définition 1.39. *Soit E un espace vectoriel. Si $F_1, \ldots, F_n$ sont des sous-espaces vectoriels de E, on dit qu'ils sont en* somme directe *si tout élément $v \in F_1 + \cdots + F_n$ s'écrit de façon unique sous la forme $x = v_1 + \cdots + v_n$ avec $v_i \in F_i$ pour $1 \leqslant i \leqslant n$. Lorsque tel est le cas, on note également $F_1 \oplus \cdots \oplus F_n$ le sous-espace vectoriel $F_1 + \cdots + F_n$.*

Pour vérifier que des sous-espaces vectoriels sont en somme directe, on peut utiliser le critère suivant.

Proposition 1.40. *Les sous-espaces $F_1 + \cdots + F_n$ sont en somme directe si et seulement si pour tout $(v_1, \ldots, v_n) \in F_1 \times \cdots \times F_n$, l'égalité $v_1 + \cdots + v_n = 0_E$ implique $v_1 = v_2 = \cdots = v_n = 0_E$.*

Démonstration. Le sens $\Rightarrow$ est immédiat, il s'agit juste d'appliquer la définition d'une somme directe à la décomposition de l'élément 0. Montrons que si l'égalité $v_1 + \cdots + v_n = 0_E$ implique $v_1 = v_2 = \cdots = v_n = 0$ pour tout $(v_1, \ldots, v_n) \in F_1 \times \cdots \times F_n$, alors les sous-espaces $F_1, \ldots, F_n$ sont en somme directe. Il faut donc prouver que si $(v_1, \ldots, v_n) \in F_1 \times \cdots \times F_n$ et si $(w_1, \ldots, w_n) \in F_1 \times \cdots \times F_n$ vérifient

$$v_1 + \cdots + v_n = w_1 + \cdots + w_n, \tag{1}$$

alors $v_1 = w_1, \dots, v_n = w_n$. Pour cela, réécrivons l'égalité (1) sous la forme

$$(v_1 - w_1) + (v_2 - w_2) + \cdots + (v_n - w_n) = 0_E.$$

L'hypothèse implique alors $v_1 - w_1 = \cdots = v_n - w_n = 0_E$, c'est-à-dire $v_1 = w_1, \dots, v_n = w_n$. □

Le cas de deux sous-espaces vectoriels est particulier.

Proposition 1.41. *Soit E un K-espace vectoriels et soient F_1 et F_2 deux sous-espaces vectoriels de E. Alors F_1 et F_2 sont en somme directe si et seulement si $F_1 \cap F_2 = \{0_E\}$.*

Démonstration. Supposons tout d'abord que F_1 et F_2 sont en somme directe. Soit $v \in F_1 \cap F_2$. Alors $0_E = v - v = 0_E - 0_E$. Comme F_1 et F_2 sont en somme directe, une telle décomposition est unique, ainsi $v = 0_E$ et donc $F_1 \cap F_2 = \{0_E\}$. Réciproquement supposons $F_1 \cap F_2 = 0$. Soit $(v_1, v_2) \in F_1 \times F_2$ tel que $v_1 + v_2 = 0_E$. On a alors $v_1 = -v_2 \in F_2$, donc $v_1 \in F_1 \cap F_2$ et donc $v_1 = 0_E$. On en déduit $v_2 = 0_E$. Ainsi la proposition 1.40 implique que F_1 et F_2 sont en somme directe. □

Remarque 1.42. Attention, la proposition 1.41 ne se généralise pas verbatim au cas deux n sous-espaces vectoriels avec $n \geqslant 3$. Considérons par exemple le cas de $E = \mathbb{R}^2$, $F_1 = \operatorname{Vect}\left(\begin{pmatrix}1\\0\end{pmatrix}\right)$, $F_2 = \operatorname{Vect}\left(\begin{pmatrix}0\\1\end{pmatrix}\right)$ et $F_3 = \operatorname{Vect}\left(\begin{pmatrix}1\\1\end{pmatrix}\right)$. On vérifie que $F_1 \cap F_2 = F_2 \cap F_3 = F_1 \cap F_3 = \{0_{\mathbb{R}^2}\}$ (en particulier $F_1 \cap F_2 \cap F_3 = \{0_{\mathbb{R}^2}\}$). Pourtant ces trois sous-espaces ne sont pas en somme directe puisque

$$\begin{pmatrix}1\\1\end{pmatrix}) = 0_{\mathbb{R}^2} + 0_{\mathbb{R}^2} + \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix} + 0_{\mathbb{R}^2}.$$

Proposition 1.43. *Soit E un espace vectoriel et soient F et G deux sous-espaces vectoriels de E. Si F et G sont de dimension finie, alors $F + G$ et $F \cap G$ sont de dimension finie et on a*

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

Proposition 1.44. *Soit $(F_1, \dots, F_n)$ une famille de sous-espaces vectoriels de dimension finie de E. Si ces sous-espaces sont en somme directe, alors pour toute base $\mathcal{B}_1$ de F_1, $\mathcal{B}_2$ de F_2,...,$\mathcal{B}_n$ de F_n, la famille $\mathcal{B}$ obtenue en concaténant $\mathcal{B}_1$, $\mathcal{B}_2$,...,$\mathcal{B}_n$ est une famille libre de E. Si de plus $F_1 \oplus \cdots \oplus F_n = E$, alors $\mathcal{B}$ est une base de E.*

Proposition 1.45. *Soient $F_1, \dots, F_n$ des sous-espaces vectoriels de dimension finie de E. Ces sous-espaces sont en somme directe si et seulement si*

$$\dim(F_1 + \cdots + F_n) = \sum_{i=1}^{n} \dim F_i.$$

1.8 Outils pratiques

Soit $A = (a_{i,j}) \in \mathcal{M}_{m,n}(K)$ une matrice. Le *système linéaire homogène* associé est le système

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n & = 0 \\ \vdots & \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n & = 0. \end{cases}$$

La méthode du pivot permet de ramener ce système à un système de la forme

$$\begin{cases} x_{\sigma(1)} + b_{1,2}x_{\sigma(2)} + \cdots + b_{1,n}x_{\sigma(n)} & = 0 \\ 0 + x_{\sigma(2)} + b_{2,3}x_{\sigma(3)} + \cdots + b_{2,n}x_{\sigma(n)} & = 0 \\ 0 + 0 + x_{\sigma(3)} + \cdots & = 0 \\ \vdots & \vdots \\ 0 + \cdots + 0 + x_{\sigma(r)} + \cdots + b_{r,n}x_{\sigma(n)} & = 0 \end{cases} \tag{2}$$

où $x_{\sigma(1)}, \ldots, x_{\sigma(n)}$ est une permutation des variables $x_1, \ldots, x_n$ et r est un entier $0 \leqslant r \leqslant n$. On dit que le système a été mis sous forme *échelonnée*. Soit $0 \leqslant r \leqslant n$ l'entier tel que les variables $x_{\sigma(1)}, \ldots, x_{\sigma(r)}$ apparaissent sur la diagonale du système échelonné. Les variables $x_{\sigma(1)}, \ldots, x_{\sigma(r)}$ sont appelées les *variables principales* et $x_{\sigma(r+1)}, \ldots, x_{\sigma(n)}$ les *variables libres*. On peut alors extraire les informations suivantes du système mis sous forme échelonnée :

— le range de la matrice A est r, le nombre de variables principales ;

— on obtient une base du noyau de A en considérant $(X_1, \ldots, X_{n-r})$ où X_i est l'unique solution du système telle que $x_{\sigma(r+i)} = 1$ et $x_{\sigma(r+j)} = 0$ pour $1 \leqslant j \leqslant n-r$ et $j \neq i$;

— on obtient une base de l'image de A en prenant

$$(Ae_{\sigma(1)}, \ldots, Ae_{\sigma(r)})$$

(on rappelle que $(e_1, \ldots, e_n)$ désigne la base canonique de $\mathbb{R}^n$).

Ces techniques peuvent s'appliquer au calcul du rang, du noyau et de l'image d'une application linéaire entre espaces vectoriels de dimension finie. Soient E et F deux espaces vectoriels de dimension finie. Soit $f \in \mathcal{L}(R, F)$. On fixe $\mathcal{B}_E$ une base de E et $\mathcal{B}_F$ une base de F. Posons $A = \mathrm{Mat}_{\mathcal{B}_F \leftarrow \mathcal{B}_E}(K)$. On a alors

— $\mathrm{rg}(f) = \mathrm{rg}(A)$, où $\mathrm{rg}(A)$ peut se déterminer par la méthode du pivot comme ci-dessus ;

— soit $(X_1, \ldots, X_{n-r})$ une base de $\mathrm{Ker}(A)$, on obtient une base de $\mathrm{Ker}(f)$ en prenant $(v_1, \ldots, v_{n-r})$ telle que $[v_i]_{\mathcal{B}_E} = X_i$;

— soit $(Y_1, \ldots, Y_r)$ une base de $\mathrm{Im}(A)$, on obtient une base de $\mathrm{Ker}(f)$ en prenant $(v_1, \ldots, v_r)$ telle que $[v_i]_{\mathcal{B}_F} = Y_i$.

2 Le groupe symétrique

2.1 Définition

Soit X un ensemble. Une *permutation* de X est une application bijective $f : X \to X$. L'ensemble des permutations de X est noté $\mathfrak{S}(X)$. Comme la composition de deux applications bijectives est encore bijective, si f et g sont deux permutations de X, leur composée $f \circ g$ est encore une permutation de X. On a donc définit une opération sur l'ensemble $\mathfrak{S}(X)$: l'opération de *composition* qui prend deux éléments f et g de $\mathfrak{S}(X)$ et en fournit un troisième $f \circ g$.

Nous nous intéresserons désormais uniquement au cas où l'ensemble X est l'ensemble $\{1, 2, \ldots, n\}$ des entiers de 1 à n pour un entier $n \geqslant 1$.

Définition 2.1. *On appelle* groupe symétrique *et on note* $\mathfrak{S}_n$ *l'ensemble permutations de l'ensemble* $\{1, \ldots, n\}$.

Une façon standard de décrire une permutation σ est de l'écrire sous la forme d'un tableau à deux lignes, la première ligne étant la liste $1, 2, \ldots, n$ et la deuxième la liste $\sigma(1), \sigma(2), \ldots, \sigma(n)$. Voici un exemple si $n = 4$. La permutation

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

est la permutation

$$\begin{cases} 1 \mapsto 2 \\ 2 \mapsto 1 \\ 3 \mapsto 4 \\ 4 \mapsto 3 \end{cases}$$

Théorème 2.2. *Soit $n \geqslant 1$ un entier. L'ensemble $\mathfrak{S}_n$ est fini de cardinal $n!$.*

Démonstration. Il faut compter combien de permutations de l'ensemble fini $\{1, \ldots, n\}$ sont possibles. Se donner une permutation de $\{1, \ldots, n\}$ revient à se donner n entiers $\sigma(1), \ldots, \sigma(n)$ deux à deux distincts et compris entre 1 et n. Il y a donc n choix possibles pour $\sigma(1)$. Une fois $\sigma(1)$ choisi, il n'y a plus que $n - 1$ choix pour $\sigma(2)$, puis $n - 2$ choix pour $\sigma(3)$ etc. et une unique possibilité pour $\sigma(n)$. Au final, il y a donc $n(n-1)(n-2) \cdots 1$ choix possibles de permutations de $\{1, \ldots, n\}$. $\square$

2.2 Exemples d'éléments

Si $1 \leqslant i < j \leqslant n$, on note (i, j) l'unique permutation de $\{1, \ldots, n\}$ qui échange i et j et fixe tous les autres éléments. Une telle permutation s'appelle une *transposition.*

Si $2 \leqslant k \leqslant n$ et si $a_1, \ldots, a_k$ sont des éléments distincts de $\{1, \ldots, n\}$, on note $(a_1, \ldots, a_k)$ la permutation σ définie par

$$\begin{aligned}
\sigma(a_1) &= a_2 \\
\sigma(a_2) &= a_3 \\
&\vdots \\
\sigma(a_k) &= a_1 \\
\sigma(x) &= x \text{ si } x \notin \{a_1, \ldots, a_k\}.
\end{aligned}$$

Une telle permutation est appelée un *k-cycle*.

Remarque 2.3. Les 2-cycles sont exactement les transpositions.

Exemple 2.4.

$$(2,3) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad (1,3,2) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

2.3 Structure de groupe

La loi de composition $\circ$ de $\mathfrak{S}_n$ possède une propriété importante : elle est *associative*. Ce la signifie que, pour $\sigma_1, \sigma_2, \sigma_3$ dans $\mathfrak{S}_n$, on a

$$\sigma_1 \circ (\sigma_2 \circ \sigma_3) = (\sigma_1 \circ \sigma_2) \circ \sigma_3.$$

Démontrons-le.

Proposition 2.5. *La loi de composition de $\mathfrak{S}_n$ est associative.*

Démonstration. Rappelons que par définition, la permutation $\sigma_1 \circ \sigma_2$ est la permutation définie par $(\sigma_1 \circ \sigma_2)(x) = \sigma_1(\sigma_2(x))$. Le principe de la démonstration est donc de calculer, pour tout $x \in \{1, \ldots, n\}$, les éléments $(\sigma_1 \circ (\sigma_2 \circ \sigma_3))(x)$ et $((\sigma_1 \circ \sigma_2) \circ \sigma_3)(x)$ et de vérifier qu'ils sont égaux. Commençons par

$$(\sigma_1 \circ (\sigma_2 \circ \sigma_3))(x) = \sigma_1((\sigma_2 \circ \sigma_3)(x)) = \sigma_1(\sigma_2(\sigma_3(x))).$$

Et finissons par

$$((\sigma_1 \circ \sigma_2) \circ \sigma_3)(x) = (\sigma_1 \circ \sigma_2)(\sigma_3(x)) = \sigma_1(\sigma_2(\sigma_3(x))).$$

On a donc $(\sigma_1 \circ (\sigma_2 \circ \sigma_3))(x) = ((\sigma_1 \circ \sigma_2) \circ \sigma_3)(x)$ pour tout $x \in \{1, \ldots, n\}$, ce qui signifie que $\sigma_1 \circ (\sigma_2 \circ \sigma_3) = (\sigma_1 \circ \sigma_2) \circ \sigma_3$. $\square$

L'application identité $\mathrm{Id}_{\{1,\ldots,n\}}$ (que nous noterons simplement Id par la suite) est une permutation de X et vérifie $\sigma \circ \mathrm{Id} = \sigma = \mathrm{Id}_\circ \sigma$. On dit que c'est un *élément neutre* pour la loi $\circ$.

Si $\sigma \in \mathfrak{S}_n$, l'application réciproque σ^{-1} est un autre élément de $\mathfrak{S}_n$ qui vérifie $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \mathrm{Id}$.

Ces observations peuvent se résumer en disant que la paire $(\mathfrak{S}_n, \circ)$ est un *groupe*.

Définition 2.6. *Un ensemble G muni d'une loi de composition interne $*$ est appelé un* groupe *si*

(i) la loi $*$ est associative : $a * (b * c) = (a * b) * c$ pour tous a, b, c dans G ;

(ii) la loi x possède un élément neutre $e \in G$: $e * a = a * e = a$ pour tout $a \in G$;

(iii) tout élément a de G possède un symétrique a^{-1} pour $*$: $a * a^{-1} = a^{-1} * a = e$.

Exemple 2.7. La paire $(\mathfrak{S}_n, \circ)$ est un groupe. Si E est un espace vectoriel, la paire $(E, +)$ possède également une structure de groupe.

Ainsi tout élément possède un inverse et $(\mathfrak{S}(X), \circ)$ est un groupe.

Exemple 2.8. Considérons les éléments de $\mathfrak{S}_3$

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \sigma' = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

et calculons $\sigma \circ \sigma'$. On a $(\sigma \circ \sigma')(1) = \sigma(2) = 3$ et $(\sigma \circ \sigma')(2) = \sigma(1) = 2$. Comme $\sigma \circ \sigma'$ est une permutation, on a nécessairement $(\sigma \circ \sigma')(3) = 1$:

$$\begin{cases} 1 \overset{\sigma'}{\longmapsto} 2 \overset{\sigma}{\longmapsto} 3 \\ 2 \mapsto 1 \mapsto 2 \\ 3 \mapsto 3 \mapsto 1 \end{cases}$$

et donc $\sigma \circ \sigma' = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$. À titre d'exercice, vérifier que $\sigma'\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$. On remarque que $\sigma \circ \sigma' \neq \sigma' \circ \sigma$, le groupe S_3 n'est donc pas commutatif ! L'ordre de composition est donc très important.

Lorsqu'un groupe $(G, *)$ vérifie de plus la propriété $a * b = b * a$ pour tous a, b dans G, on dit qu'il est *commutatif*. On vient de voir que le groupe $\mathfrak{S}_3$ (et par extension le groupe $\mathfrak{S}_n$ pour $n \geqslant 3$) n'est pas commutatif.

Théorème 2.9. *Tout élément de $\mathfrak{S}_n$ s'écrit comme un produit de transpositions.*

Démonstration. Pour $n \geqslant 2$, soit H_n l'hypothèse de récurrence « toute permutation de $\{1, \dots, n\}$ est un produit de transpositions ». Alors H_2 est vrai car $\mathfrak{S}_2 = \{\mathrm{Id}, (12)\}$. Supposons H_n vrai et démontrons H_{n+1}. Soit $\sigma \in \mathfrak{S}_{n+1}$ et posons

$$\sigma' = \begin{cases} \sigma & \text{si } \sigma(n+1) = n+1 \\ (n+1, \sigma(n+1)) \circ \sigma & \text{si } \sigma(n+1) \neq n+1. \end{cases}$$

Alors $\sigma'(n+1) = n+1$. La restriction de σ' à $\{1, \dots, n\}$ est un élément de $\mathfrak{S}_n$ et s'écrit comme un produit de transpositions par H_n. Comme $\sigma = (n+1, \sigma(n+1)) \circ \sigma'$, on en conclut que σ est un produit de transpositions. $\square$

2.4 Signature

Si $\sigma \in \mathfrak{S}_n$, on pose

$$\ell(\sigma) = \mathrm{Card}\{(i,j) \in \{1,\ldots,n\}^2 \mid i < j \text{ et } \sigma(i) > \sigma(j)\}.$$

Il s'agit du nombre d'*inversions* de σ. On pose alors $\varepsilon(\sigma) = (-1)^{\ell(\sigma)}$. Le nombre $\varepsilon(\sigma)$ s'appelle la *signature* de σ.

Théorème 2.10. *Pour tous $\sigma, \tau \in \mathfrak{S}_n$, on a $\varepsilon(\sigma\circ\tau) = \varepsilon(\sigma)\varepsilon(\tau)$. De plus on a $\varepsilon(\tau) = -1$ si τ est une transposition.*

Démonstration. Soient σ et τ deux éléments de $\mathfrak{S}_n$. Il faut vérifier que $\varepsilon(\sigma\circ\tau) = \varepsilon(\sigma)\varepsilon(\tau)$. On utilise la formule suivante

$$\varepsilon(\sigma) = \prod_{1\leqslant i<j\leqslant n} \mathrm{sgn}(\sigma(j) - \sigma(i))$$

pour tout $\sigma \in \mathfrak{S}_n$. On a

$$\begin{aligned}
\varepsilon(\sigma\circ\tau)\varepsilon(\tau) &= \prod_{1\leqslant i<j\leqslant n} \mathrm{sgn}((\sigma\circ\tau)(j) - (\sigma\circ\tau)(i)) \prod_{1\leqslant i<j\leqslant n} \mathrm{sgn}(\tau(j) - \tau(i)) \\
&= \prod_{1\leqslant i<j\leqslant n} \underbrace{[\mathrm{sgn}(\sigma\circ\tau)(j)) - (\sigma\circ\tau)(i)))\mathrm{sgn}(\tau(j) - \tau(i))]}_{\text{symétrique en } i \text{ et } j} \\
&= \prod_{1\leqslant \tau(i)<\tau(j)\leqslant n} [\mathrm{sgn}(\sigma(\tau(j)) - \sigma(\tau(i)))\mathrm{sgn}(\tau(j) - \tau(i))] \\
&= \prod_{1\leqslant \tau(i)<\tau(j)\leqslant n} \mathrm{sgn}(\sigma(\tau(j)) - \sigma(\tau(i))) \\
&= \prod_{1\leqslant i<j\leqslant n} \mathrm{sgn}(\sigma(j) - \sigma(i)) = \varepsilon(\sigma).
\end{aligned}$$

On a donc $\varepsilon(\sigma\circ\tau)\varepsilon(\tau) = \varepsilon(\tau)$. Comme $\varepsilon(\tau) \in \{\pm 1\}$, on a $\varepsilon(\tau) = \varepsilon(\tau)^{-1}$ et donc $\varepsilon(\sigma\circ\tau) = \varepsilon(\sigma)\varepsilon(\tau)$.

Il reste à vérifier si τ est une transposition, on a $\varepsilon(\tau) = -1$. Supposons que $\tau = (i,j)$ avec $i < j$ et soient $k < \ell$.

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

Ainsi $\ell(\tau) = 2(j - i - 1) + 1$ et donc $\varepsilon(\tau) = -1$. $\square$

Si $(G, *)$ et $(H, *')$ sont deux groupes. On appelle *morphisme* d'un groupe G vers un groupe H une application $f : G \to H$ telle que $f(g * h) = f(g) *' f(h)$ pour tous g et h dans G.

Exemple 2.11. Considérons le cas où $G = (\mathfrak{S}_n, \circ)$ et $H = (\{\pm 1\}, \times)$. Alors la signature ε est un morphisme du groupe $\mathfrak{S}_n$ vers le groupe $\{\pm 1\}$.

Proposition 2.12. *Pour tout $\sigma \in \mathfrak{S}_n$, on a*

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

Démonstration. On a $\varepsilon(\mathrm{Id}) = 1$. Par ailleurs si $\sigma \in \mathfrak{S}_n$, on a

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

Ainsi $\varepsilon(\sigma)^{-1} = \varepsilon(\sigma^{-1})$. Comme $\varepsilon(\sigma) \in \{\pm 1\}$, on a $\varepsilon(\sigma)^2 = 1$ et donc $\varepsilon(\sigma) = \varepsilon(\sigma)^{-1}$. $\square$

Corollaire 2.13. *Soit $\sigma \in \mathfrak{S}_n$. La parité du nombre de transposition dans une décomposition de σ en produit de transpositions ne dépend que de σ.*

Démonstration. Soient $\sigma = \tau_1 \circ \cdots \circ \tau_r$ et $\sigma = \tau'_1 \circ \cdots \circ \tau'_s$ deux décomposition de σ en produit de transposition. Comme $\varepsilon(\tau_i) = \varepsilon(\tau'_j) = -1$ pour tous i et j, on a $\varepsilon(\sigma) = (-1)^r = (-1)^s$. On en déduit que r et s ont la même parité. $\square$

3 Déterminants

3.1 Définition

Définition 3.1. *Soit $A = (a_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_n(K)$. On appelle* déterminant *de la matrice A la quantité*

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}.$$

On utilise également la notation

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

Exemple 3.2. Si $n = 2$, on a $\mathfrak{S}_2 = \{\mathrm{Id}, (1,2)\}$ et $\varepsilon((1,2)) = -1$. Ainsi

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

On retrouve la notion de déterminant d'une matrice 2×2 vue en première année.

Si $n = 3$, on a $\mathfrak{S}_3 = \{\mathrm{Id}, (12), (23), (13), (123), (132)\}$ et

$$\varepsilon(\mathrm{Id}) = \varepsilon((123)) = \varepsilon((132)) = 1, \quad \varepsilon((12)) = \varepsilon((23)) = \varepsilon((13)) = -1.$$

Ainsi

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = a_{1,1}a_{2,2}a_{3,3} - a_{2,1}a_{1,2}a_{3,3} - a_{1,1}a_{3,2}a_{2,3} - a_{3,1}a_{2,2}a_{1,3} \\ + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3}.$$

Voici un cas particulièrement simple de calcul du déterminant. On dit qu'une matrice $A \in \mathcal{M}_n(K)$ est *triangulaire supérieure* si $a_{i,j} = 0$ dès que $i > j$. Autrement dit il s'agit d'une matrice de la forme

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

Proposition 3.3. *Soit $A \in \mathcal{M}_n(K)$. Une matrice triangulaire supérieure. Alors*

$$\det(A) = \prod_{i=1}^{n} a_{i,i} = a_{1,1}a_{2,2}\cdots a_{n,n}.$$

Démonstration. Soit $\sigma \in \mathfrak{S}_n$. Si $\sigma \neq \mathrm{Id}$, il existe $1 \leqslant i \leqslant n$ tel que $\sigma(i) > i$. Comme A est triangulaire supérieure, on a $a_{\sigma(i),i} = 0$. On en déduit que $\prod_{i=1}^{n} a_{\sigma(i),i} = 0$ dès que $\sigma \neq \mathrm{Id}$. On déduit alors la formule de la définition du déterminant. □

On admet également la formule suivante pour le calcul des matrices triangulaires par blocs.

Proposition 3.4. *Si $A_1 \in \mathcal{M}_r(K)$, $A_2 \in \mathcal{M}_{n-r}(K)$ et $B \in \mathcal{M}_{r,n-r}(K)$, alors*

$$\begin{vmatrix} A_1 & B \\ 0_{n-r,r} & A_2 \end{vmatrix} = \det(A_1)\det(A_2).$$

Définition 3.5. *Si $A \in \mathcal{M}_{n,p}(K)$, on appelle* transposée de A *et on note ${}^tA$ la matrice de $\mathcal{M}_{p,n}(K)$ définie par*

$${}^tA = (a_{j,i})_{\substack{1\leqslant i\leqslant p \\ 1\leqslant j\leqslant n}}.$$

On a alors ${}^t(AB) = {}^tB\,{}^tA$.

Proposition 3.6. *Soit $A \in \mathcal{M}_n(K)$ une matrice carrée. On a alors* $\det({}^tA) = \det(A)$.

Démonstration. Soit $A = (a_{i,j})_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}} \in \mathcal{M}_n(K)$. Notons ${}^tA = (b_{i,j})_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}}$. Par définition de la transposée, on a $b_{i,j} = a_{j,i}$ pour $1 \leqslant i, j \leqslant n$. Ainsi, par définition du déterminant, on a donc

$$\det({}^tA) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma^{-1}(i),i} = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma^{-1}) \prod_{i=1}^{n} a_{\sigma(i),i}$$

où la dernière égalité provient du fait que l'application $\sigma \mapsto \sigma^{-1}$ est une bijection de $\mathfrak{S}_n$ sur $\mathfrak{S}_n$. On conclut en remarquant que $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$ pour tout $\sigma \in \mathfrak{S}_n$. $\square$

3.2 Opérations sur les lignes et les colonnes d'un déterminant

Théorème 3.7. *1) Si on échange deux colonnes d'indices distincts d'une matrice carrée, on multiplie son déterminant par -1.*

2) Si deux colonnes d'indice distincts d'une matrice carrée A sont identiques, alors $\det A = 0$.

3) On ne change pas le déterminant d'une matrice carrée A en ajoutant à une de ses colonnes une combinaison linéaire des autres colonnes. Autrement dit si $1 \leqslant i \leqslant n$ et si $(\lambda_j)_{j \neq i}$ est une famille de scalaires, on a

$$\det(C_1 \cdots C_n) = \det(C_1 \cdots C_{i-1}(C_i + \sum_{j \neq i} \lambda_j C_j) \cdots C_n).$$

Comme la transposition échange les lignes et colonnes d'une matrice et que le déterminant ne change pas par transposition, tous les résultats portant sur les colonnes d'un déterminant ont un analogue sur les lignes. On en déduit donc le résultat suivant.

Théorème 3.8. *Soit $n \geqslant 1$.*

1) Si on échange deux lignes d'indices distincts d'une matrice carrée $A \in \mathcal{M}_n(K)$, on multiplie son déterminant par -1.

2) Une matrice carrée $A \in \mathcal{M}_n(K)$ ayant deux lignes d'indices distincts qui sont identiques vérifie $\det A = 0$.

3) On ne change pas le déterminant d'une matrice carrée A en ajoutant à une de ses lignes une combinaison linéaire des autres lignes.

Pour calculer un déterminant, on peut donc commencer par le mettre sous forme triangulaire supérieure (ou inférieure) en effectuant des opérations élémentaires sur ses lignes ou ses colonnes et utiliser la formule permettant de calculer le déterminant d'une matrice triangulaire.

Nous allons à présent démontrer les énoncés ci-dessus.

Proposition 3.9. *L'application* det *est linéaire en chaque colonne. Plus précisément, étant donné $n \geqslant 1$, ainsi que $1 \leqslant j \leqslant n$ et $n-1$ vecteurs colonnes*

$$C_1, \ldots, C_{j-1}, C_{j+1}, \ldots, C_n \in K^n,$$

alors l'application

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

est linéaire, c'est-à-dire

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

Démonstration. Fixons $C_1, \dots, C_{j-1}, C_{j+1}, C_n$, $n-1$ vecteurs colonnes de K^n. Soient C et C' deux autres vecteurs colonnes et $\lambda \in K$ un scalaire. Il faut prouver que

$$\begin{aligned}&\det(C_1C_2 \cdots C_{j-1}(C + \lambda C')C_{j+1} \cdots C_n)\\ &\qquad = \det(C_1C_2 \cdots C_{j-1}C_jC_{j+1} \cdots C_n) + \lambda \det(C_1C_2 \cdots C_{j-1}C'_jC_{j+1} \cdots C_n).\end{aligned}$$

Notons donc $(a_{i,k})_{1\leqslant i\leqslant n}$ les coefficients de C_k pour $1 \leqslant k \leqslant n$, $(a_{i,j})_{1\leqslant i\leqslant n}$ les coefficients de C et $(a'_{i,j})_{1\leqslant i\leqslant n}$ les coefficients de C'. On a donc

$$\begin{aligned}\det(C_1C_2 \cdots C_{j-1}(C + \lambda C')C_{j+1} \cdots C_n) &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)(a_{\sigma(j),j} + \lambda a'_{\sigma(j),j} \prod_{k\neq j} a_{\sigma(k),k}\\ &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \left(a_{\sigma(j),j} \prod_{k\neq j} a_{\sigma(k),k} + \lambda a'_{\sigma(j),j} \prod_{k\neq j} a_{\sigma(k),k} \right)\\ &= \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^{n} a_{\sigma(k),k} + \lambda \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) a'_{\sigma(j),j} \prod_{k\neq j} a_{\sigma(k),k}\\ &= \det(C_1C_2 \cdots C_{j-1}CC_{j+1} \cdots C_n)\\ &\quad + \lambda \det(C_1C_2 \cdots C_{j-1}C'C_{j+1} \cdots C_n).\end{aligned}$$ □

Proposition 3.10. *1) Soit $A \in \mathcal{M}_n(K)$ une matrice de colonnes $C_1, \dots, C_n$. Si $\tau \in \mathfrak{S}_n$, on a*

$$\det(C_{\tau(1)} \dots C_{\tau(n)}) = \varepsilon(\tau) \det(C_1 \dots C_n).$$

2) Si on échange deux colonnes distinctes d'une matrice carrée, on multiplie son déterminant par -1.

3) Si deux colonnes distinctes d'une matrice carrée A sont égales, alors $\det A = 0$.

Démonstration. Prouvons le point 1). Soit $A = (a_{i,j})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}} \in \mathcal{M}_n(K)$ et soit $\tau \in \mathfrak{S}_n$ une permutation. Notons B la matrice $(C_{\tau(1)} \dots C_{\tau(n)})$. On a alors

$$\begin{aligned}
\det(B) &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{\sigma(j),\tau(j)} \\
&= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{(\sigma\circ\tau^{-1})(\tau(j)),\tau(j)} \\
&= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma \circ \tau^{-1}) \varepsilon(\tau) \prod_{j=1}^{n} a_{(\sigma\circ\tau^{-1})(j),j} \\
&= \varepsilon(\tau) \sum_{\sigma' \in \mathfrak{S}_n} \varepsilon(\sigma') \prod_{j=1}^{n} a_{\sigma'(j),j} \\
&= \varepsilon(\tau) \det(A)
\end{aligned}$$

L'avant dernière égalité provient du fait que l'application $\sigma \mapsto \sigma \circ \tau^{-1}$ est une permutation de $\mathfrak{S}_n$. En effet sa réciproque est donnée par $\sigma \mapsto \sigma \circ \tau$.

Prouvons 2). Si $i < j$ et si $\tau = (i, j)$, on a $\varepsilon(\tau) = -1$, on déduit donc de 1) que

$$\det(C_1 \cdots C_j \cdots C_j \cdots C_n) = -\det(C_1 \cdots C_i \cdots C_j \cdots C_n).$$

Prouvons à présent le point 3). Si $i < j$, et si $C_i = C_j$, on a

$$\begin{aligned}
\det(C_1 \cdots C_i \cdots C_j \cdots C_n) &= -\det(C_1 \cdots C_j \cdots C_i \cdots C_n) \\
&= \det(C_1 \cdots C_j \cdots C_i \cdots C_n)
\end{aligned}$$

où la première égalité provient de 1) et la seconde égalité de $C_i = C_j$. On a donc bien

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

□

Exemple 3.11.

$$\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}.$$

Corollaire 3.12. *On ne change pas le déterminant d'une matrice carrée A en ajoutant à une de ses colonnes une combinaison linéaire des autres colonnes. Autrement dit si $1 \leqslant i \leqslant n$ et si $(\lambda_j)_{j \neq i}$ est une famille de scalaires, on a*

$$\det(C_1 \cdots C_n) = \det(C_1 \cdots C_{i-1}(C_i + \sum_{j \neq i} \lambda_j C_j) \cdots C_n).$$

Démonstration. En effet on a

$$\begin{aligned}
\det(C_1 \cdots C_{i-1}(C_i + \sum_{j \neq i} \lambda_j C_j) \cdots C_n) &= \det(C_1 \cdots C_i \cdots C_n) \sum_{j \neq i} \lambda_j \det(C_1 \cdots C_{i-1} C_j \cdots C_n) \\
&= \det(C_1 \cdots C_{i-1} C_i \cdots C_n)
\end{aligned}$$

□

Comme la transposition échange les lignes et colonnes d'une matrice et que le déterminant ne change pas par transposition, tous les résultats portant sur les colonnes d'un déterminant ont un analogue sur les lignes. On en déduit donc le résultat suivant.

Proposition 3.13. *Soit $n \geqslant 1$.*

1) L'application $\det : \mathcal{M}_n(K)$ *est linéaire en chaque ligne.*

2) Soit $A \in \mathcal{M}_n(K)$ une matrice de lignes $L_1, \ldots, L_n$. Si $\tau \in \mathfrak{S}_n$, on a

$$\det \begin{pmatrix} L_{\tau(1)} \\ \vdots \\ L_{\tau(n)} \end{pmatrix} = \varepsilon(\tau) \det \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix}.$$

3) Si on échange deux lignes d'une matrice carrée $A \in \mathcal{M}_n(K)$, on multiplie son déterminant par -1.

4) Une matrice carrée $A \in \mathcal{M}_n(K)$ ayant deux lignes égales vérifie $\det A = 0$.

5) On ne change pas le déterminant d'une matrice carrée A en ajoutant à une de ses lignes une combinaison linéaire des autres lignes.

3.3 Développement d'un déterminant selon une ligne ou une colonne

Si $A \in \mathcal{M}_n(K)$ et si i et j sont deux entiers compris entre 1 et n, on note $A_{i,j} \in \mathcal{M}_{n-1}(K)$ la matrice carrée de taille $n-1$ obtenue à partir de A en supprimant dans A la i-ème ligne et la j-ème colonne.

Théorème 3.14. *Soit $n \geqslant 1$ et soit $A \in \mathcal{M}_n(K)$.*

1) Soit $1 \leqslant i \leqslant n$. On a alors

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

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

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

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

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

Démonstration. On prouve la formule de développement selon une colonee. La formule selon une ligne s'en déduit en utilisant le fait que $\det({}^t A) = \det(A)$.

Soit $1 \leqslant j \leqslant n$. La j-ème colonne C_j de la matrice A peut s'écrire $A = a_{1,j} e_1 + \cdots + a_{n,j} e_n$ où $e_1, \ldots, e_n$ désigne les vecteurs colonnes formant la base canonique de K^n. On déduit donc de la proposition 3.9 que

$$\det(A) = a_{1,j} \det(C_1 \ldots C_{j-1} e_1 C_{j+1} \ldots C_n) + \cdots + a_{n,j} \det(C_1 \ldots C_{j-1} e_n C_{j+1} \ldots C_n).$$

Or en appliquant 3.10 1) avec $\tau = (1, 2, \dots, j)$, on obtient, pour tout $1 \leqslant i \leqslant n$,

$$\begin{aligned}
\det(C_1 \dots C_{j-1} e_i C_{j+1} \dots C_n) &= \varepsilon(\tau) \det(e_i C_1 \dots C_{j-1} C_{j+1} \dots C_n) \\
&= \begin{vmatrix} 0 & a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_{i,1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}
\end{aligned}$$

La même opération de permutation sur les lignes, en utilisant le cycle $\tau' = (1, 2, \dots, i)$, donne

$$\begin{aligned}
\det(C_1 \dots C_{j-1} e_i C_{j+1} \dots C_n) &= \varepsilon(\tau)\varepsilon(\tau') \begin{vmatrix} 1 & a_{i,1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{i,n} \\ 0 & a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i,j+1} & \cdots & a_{i-1,n} \\ 0 & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix} \\
&= \varepsilon(\tau)\varepsilon(\tau') \begin{vmatrix} 1 & * \\ 0_{n-1,1} & A_{i,j} \end{vmatrix} = \varepsilon(\tau)\varepsilon(\tau') \det(A_{i,j}).
\end{aligned}$$

Comme τ est un cycle de longueur j et τ' un cycle de longueur i, on a $\varepsilon(\tau) = (-1)^{j-1}$ et $\varepsilon(\tau') = (-1)^{i-1}$ donc $\varepsilon(\tau)\varepsilon(\tau') = (-1)^{i+j-2} = (-1)^{i+j}$. On en déduit la formule recherchée. □

3.4 Multiplicativité du déterminant

Théorème 3.15. *Soit $n \geqslant 1$ et soient A et B deux matrices de $\mathcal{M}_n(K)$. On a alors*

$$\det(AB) = \det(A)\det(B).$$

Démonstration. Fixons $A = (a_{i,j})$ et $B = (b_{i,j})$ dans $\mathcal{M}_n(K)$. Notons $C_1, \dots, C_n$ les colonnes de B. Rappelons que l'on note $e_1, \dots, e_n$ les vecteurs de la base canonique de K^n. On a donc, pour tout $1 \leqslant j \leqslant n$,

$$C_j = \sum_{i=1}^{n} b_{i,j} e_i.$$

Par définition du produit matriciel, les colonnes de AB sont les vecteurs colonnes $AC_1, \dots, AC_n$. Pour $1 \leqslant j \leqslant n$, on a donc

$$AC_j = \sum_{i=1}^{n} b_{i,j} A e_i.$$

On peut donc calculer le déterminant $\det(AB)$ en utilisant la linéarité par rapport à chaque colonne (proposition 3.9). On a donc

$$\begin{aligned}\det(AB) = \det(AC_1, \dots, AC_n) &= \det\left(\sum_{i_1=1}^{n} b_{i_1,1} Ae_1, AC_2, \dots, AC_n\right) \\ &= \sum_{i_1=1}^{n} b_{i_1,1} \det(Ae_{i_1}, AC_2, \dots, AC_n) \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} b_{i_1,1} b_{i_2,2} \det(Ae_{i_1}, Ae_{i_2}, C_3, \dots, C_n) \\ &= \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} b_{i_1,1} \cdots b_{i_n,n} \det(Ae_{i_1}, \dots, Ae_{i_n}).\end{aligned}$$

Calculons alors $\det(Ae_{i_1}, \dots, Ae_{i_n})$ pour tout valeur de $(i_1, \dots, i_n) \in \{1, \dots, n\}^n$.

— Si il existe $k < \ell$ tels que $i_k = i_\ell$, alors $\det(1e_{i_1}, \dots, Ae_{i_n}) = 0$ d'après la proposition 3.10 3).

— Sinon il existe $\sigma \in \mathfrak{S}_n$ tel que $(i_1, \dots, i_n) = (\sigma(1), \dots, \sigma(n))$ et alors

$$\det(Ae_{i_1}, \dots, Ae_{i_n}) = \varepsilon(\sigma) \det(Ae_1, \dots, Ae_n)$$

d'après la proposition 3.9 1).

On a donc

$$\det(AB) = \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n),n}\right) \det(Ae_1, \dots, Ae_n) = \det(B) \det(Ae_1, \dots, Ae_n).$$

Comme $Ae_1, \dots, Ae_n$ sont les colonnes de la matrice A, on a $\det(Ae_1, \dots, Ae_n) = \det(A)$ ce qui fournit $\det(AB) = \det(B)\det(A)$. $\square$

Théorème 3.16. *Soit $A \in \mathcal{M}_n(K)$. Alors A est inversible si et seulement si* $\det(A) \neq 0$. *De plus, si A est inversible, alors*

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

Démonstration. Supposons A inversible. Alors il existe $A^{-1} \in \mathcal{M}_n(K)$ telle que $AA^{-1} = I_n$. On en déduit

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

En particulier $\det(A) \neq 0$ et $\det(A^{-1}) = \det(A)^{-1}$.

Réciproquement supposons que A n'est pas inversible. Alors $\mathrm{Ker}(A) \neq \{0_{K^n}\}$. En particulier il existe un vecteur $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n \smallsetminus \{0_{K^n}\}$ tel que $AX = 0_{K^n}$. En notant $C_1, \dots, C_n$ les colonnes de A, on en déduit $\sum_{i=1}^{n} x_i C_i = 0_{K^n}$. Ainsi les colonnes de A sont liées et $\det(A) = 0$. $\square$

Corollaire 3.17. *Considérons un système linéaire à n équations et n inconnues*

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n & = y_1 \\ \vdots & \vdots \\ a_{n,1}x_1 + \cdots + a_{n,n}x_n & = y_n \end{cases} \tag{3}$$

et considérons $A = (a_{i,j}) \in \mathcal{M}_n(K)$. Si $\det(A) \neq 0$, alors le système (3) possède une et une seule solution. En particulier si $y_1 = \cdots = y_n = 0$, la seule solution du système (3) est $x_1 = x_2 = \cdots = 0$.

Démonstration. Posons $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. Le système (3) est équivalent à l'équation $AX = Y$. Si $\det(A) \neq 0$, la matrice A est inversible d'après le théorème 3.16. On en déduit que l'équation $AX = Y$ est équivalente $X = A^{-1}Y$ qui a pour unique solution $A^{-1}Y$. De plus, si $Y = 0_{K^n}$, on voit que $X = 0_{K^n}$. □

3.5 Formule de Cramer

Définition 3.18. *Soit $A \in \mathcal{M}_n(K)$. La* comatrice *de A est la matrice* $\mathrm{Com}(A)$ *de $\mathcal{M}_n(K)$ définie par*

$$\mathrm{Com}(A) = ((-1)^{i+j} \det(A_{i,j}))_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}.$$

Théorème 3.19. *1) Si $A \in \mathcal{M}_n(K)$, on a*

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

2) Si $A \in \mathcal{M}_n(K)$ est inversible, on a

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

Démonstration. Prouvons le point 1). Soit $1 \leqslant i \leqslant n$ et soit $1 \leqslant j \leqslant n$. Calculons le coefficient $b_{i,j}$ de la ligne i et de la colonne j de la matrice $B = {}^t\,\mathrm{Com}(A)A$. Il s'agit de l'élément

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

Si $i = j$, on a, par développement du déterminant de A selon colonne $i = j$,

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

Supposons à présent $i \neq j$. Soit C la matrice obtenue en remplaçant la i-ième colonne de A par sa j-ième colonne. En développant le déterminant de C selon sa i-ième colonne, on a donc

$$b_{i,j} = \sum_{k=1}^{n} (-1)^{k+j} \det(A_{k,i}) a_{k,j} = \det(C).$$

Comme C a deux colonnes égales (sa i-ième et sa j-ième), on a $\det(C) = 0$ et donc $b_{i,j} = 0$ si $i \neq j$. Ainsi la matrice B est la matrice $\det(A)I_n$, ce qui prouve la formule.

Si A est inversible, on a alors $\det(A) \neq 0$, et on déduit facilement de la formule 1) la formule 2). □

Exemple 3.20. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(K)$. Alors

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

3.6 Déterminant d'une famille de vecteurs

Fixons un entier $n \geqslant 1$. Si $v_1, \ldots, v_n$ sont n vecteurs de K^n, on appelle *déterminant de la famille* $(v_1, \ldots, v_n)$ et on note note $\det(v_1, \ldots, v_n)$ le déterminant de la matrice dont les colonnes sont $v_1, v_2, \ldots, v_n$.

Théorème 3.21. *Soient $v_1, \ldots, v_n$ n vecteurs de K^n. Alors $\det(v_1, \ldots, v_n) \neq 0$ si et seulement si la famille $(v_1, \ldots, v_n)$ est une base de K^n.*

Démonstration. Comme K^n est de dimension n, la famille $(v_1, \ldots, v_n)$ est une base de K^n si et seulement si elle est libre, c'est-à-dire si et seulement si la matrice $(v_1, \cdots, v_n)$ a un noyau réduit à 0_{K^n}, c'est-à-dire si et seulement si la matrice $\det(v_1, \cdots, v_n) \neq 0$ d'après le théorème 3.16. □

Passons à présent au cas d'un espace vectoriel de dimension finie quelconque.

Définition 3.22. *Soit E un K-espace vectoriel. Si n est un entier, une* forme n-linéaire *sur E est une application*

$$f : E^n \longrightarrow K$$

telle que, pour tout $1 \leqslant i \leqslant n$ et tous vecteurs $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$ dans E, l'application de E dans K définie par $v \mapsto f(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$ est linéaire.

Une forme n-linéaire f sur E est dite alternée *si, pour tous $1 \leqslant i < j \leqslant n$ et tous vecteurs $v_1, \ldots, v_n$ dans E, on a $f(v_1, \ldots, v_n) = 0$ dès que $v_i = v_j$.*

Soit E un espace vectoriel de dimension finie $n \geqslant 1$ et soit $\mathcal{B}$ une base de E.

Définition 3.23. *Soit* $(v_1, \dots, v_n)$ *une famille de* n *vecteurs de* E*. Le* déterminant *de la famille* $(v_1, \dots, v_n)$ *dans la base* $\mathcal{B}$ *est le scalaire*

$$\det_{\mathcal{B}}(v_1, \dots, v_n) = \det([v_1]_{\mathcal{B}}, [v_2]_{\mathcal{B}}, \dots, [v_n]_{\mathcal{B}}).$$

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

Remarque 3.24. L'application $\det_{\mathcal{B}} : E^n \to K$ est n-linéaire alternée. C'est une conséquence des propositions 3.9 et 3.10 3).

Théorème 3.25. *Soit* E *un* K*-espace vectoriel de dimension* n*. Soit* $\mathcal{B}$ *une base de* E*. Soit* $(v_1, \dots, v_n) \in E^n$ *une famille de* n *vecteurs de* E*. Alors* $\det_{\mathcal{B}}(v_1, \dots, v_n) \neq 0$ *si et seulement si la famille* $(v_1, \dots, v_n)$ *est une base de* E*.*

Démonstration. Comme E est de dimension n, la famille $(v_1, \dots, v_n)$ est une base de E si et seulement si elle est libre, c'est-à-dire si et seulement si la matrice $([v_1]_{\mathcal{B}} \cdots [v_n]_{\mathcal{B}})$ a un noyau réduit à 0_{K^n}, c'est-à-dire si et seulement si la matrice $([v_1]_{\mathcal{B}} \cdots [v_n]_{\mathcal{B}})$, c'est-à-dire si et seulement si $\det([v_1]_{\mathcal{B}} \cdots [v_n]_{\mathcal{B}}) \neq 0$. □

Soit $\mathcal{B}_{\text{can}}$ la base canonique de K^n. Si $v_1, \dots, v_n$ sont n vecteurs de K^n, on a bien

$$\det(v_1, \dots, v_n) = \det_{\mathcal{B}_{\text{can}}}(v_1, \dots, v_n).$$

Ce paragraphe n'a pas été traité en cours mais figure dans ces notes comme complément.

Proposition 3.26. *Soit* f *une forme* n*-linéaire alternée sur un* K*-espace vectoriel* E*. Alors si* $(v_1, \dots, v_n) \in E^n$ *et si* $1 \leqslant i < j \leqslant n$*, on a*

$$f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n) = -f(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_n).$$

Plus généralement si $\sigma \in \mathfrak{S}_n$*, on a*

$$f(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \varepsilon(\sigma) f(v_1, \dots, v_n).$$

Démonstration. Il suffit de démontrer la première formule. On en déduit la seconde en décomposant σ en produit de transposition et en appliquant plusieurs fois la première formule.

Comme f est n-linéaire alternée, on a

$$\begin{aligned}
0 &= f(v_1, \dots, v_{i-1}, v_i + v_j, v_{i+1}, \dots, v_{j-1}, v_i + v_j, v_{j+1}, \dots, v_n) \\
&= f(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_{j-1}, v_i + v_j, v_{j+1}, \dots, v_n) \\
&+ f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i + v_j, v_{j+1}, \dots, v_n) \\
&= f(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n) \\
&+ f(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_n) \\
&+ f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n) \\
&+ f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_n) \\
&= 0 + f(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_n) \\
&+ f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n) + 0.
\end{aligned}$$

On en déduit le résultat. □

Exemple 3.27. Si E est un K-espace vectoriel de dimension n et si $\mathcal{B}$ est une base de E, l'application $\det_{\mathcal{B}}$ définit une forme n-linéaire alternée sur E.

Théorème 3.28. *Soit* E *un* K*-espace vectoriel de dimension* n*. Si* f *et* g *sont deux formes* n*-linéaires alternées sur* E *et si* $f \neq 0$*, il existe un unique scalaire* $\lambda \in K$ *tel que* $g = \lambda f$*.*

Démonstration. Fixons $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. Soient $v_1, \dots, v_n$ des vecteurs de E et décomposons-les dans la base $\mathcal{B}$. On a donc, pour tout $1 \leqslant j \leqslant n$,

$$v_j = \sum_{i=1}^{n} a_{i,j} e_i$$

avec $a_{i,j} \in K$. La multilinéarité de f implique alors

$$\begin{aligned} f(v_1, \dots, v_n) &= \sum_{i_1=1}^{n} a_{i_1,1} f(e_{i_1}, v_2, \dots, v_n) \\ &= \sum_{i=1}^{n} \sum_{i_2=1}^{n} a_{i_1,1} a_{i_2,2} f(e_{i_1}, e_{i_2}, v_3, \dots, v_n) \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} a_{i_1,1} a_{i_2,2} \cdots a_{i_n,n} f(e_{i_1}, \dots, e_{i_n}). \end{aligned}$$

Calculons alors $f(e_{i_1}, \dots, e_{i_n})$ pour tout valeur de $(i_1, \dots, i_n) \in \{1, \dots, n\}^n$.

- Si il existe $k < \ell$ tels que $i_k = i_\ell$, alors $f(e_{i_1}, \dots, e_{i_n}) = 0$ puisque f est alternée.
- Sinon il existe $\sigma \in \mathfrak{S}_n$ tel que $(i_1, \dots, i_n) = (\sigma(1), \dots, \sigma(n))$ et alors

$$f(e_{i_1}, \dots, e_{i_n}) = \varepsilon(\sigma) f(e_1, \dots, e_n).$$

On a donc

$$f(v_1, \dots, v_n) = \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \right) f(e_1, \dots, e_n).$$

Remarquons que le même raisonnement nous donne

$$g(v_1, \dots, v_n) = \left(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \right) g(e_1, \dots, e_n).$$

En particulier, comme $f \neq 0$, on a nécessairement $f(e_1, \dots, e_n) \neq 0$. Choisissons alors $\lambda = g(e_1, \dots, e_n) f(e_1, \dots, e_n)^{-1}$. Les formules précédentes montrent alors que $g(v_1, \dots, v_n) = \lambda f(v_1, \dots, v_n)$ et ceci pour toute valeur de $(v_1, \dots, v_n)$. □

Corollaire 3.29. *Soit E un K-espace vectoriel de dimension et soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de E. L'application $\det_{\mathcal{B}}$ est l'unique application $f : E^n \to K$ qui est n-linéaire alternée et telle que $f(b_1, \dots, b_n) = 1$.*

3.7 Déterminant et volume

Dans cette section on suppose que $K = \mathbb{R}$. Si $v_1, \dots, v_n$ sont n vecteurs de K^n, le *parallélépipède* défini par les vecteurs $v_1, \dots, v_n$ est l'ensemble

$$P = \{a_1 \cdot v_1 + a_2 \cdot v_2 + \cdots + a_n \cdot v_n \mid (a_1, \dots, a_n) \in [0,1]^n\}.$$

Théorème 3.30. *1) Si u et v sont deux vecteurs de $\mathbb{R}^2$, l'aire du parallélogramme défini par u et v vaut $|\det(u,v)|$.*

2) Si u, v, w sont trois vecteurs de $\mathbb{R}^3$, le volume du parallélépipède défini par u, v, w vaut $|\det(u,v,w)|$.

Démonstration. Démontrons la formule 1). Soient $u = \begin{pmatrix} a \\ c \end{pmatrix}$ et $v = \begin{pmatrix} b \\ d \end{pmatrix}$ deux vecteurs de $\mathbb{R}^2$ et soit P le parallélogramme défini par u et v. Soit $\theta \in [0, \pi[$ l'angle entre u et v, c'est-à-dire l'unique nombre $0 \leqslant \theta < \pi$ tel que $u \cdot v = ||u||||v|| \cos(\theta)$. L'aide de P est alors égale à $||u||||v|| \sin(\theta)$. On a donc

$$\begin{aligned}\mathcal{A}(P) &= ||u||||v||\sqrt{1-\cos^2(\theta)} = \sqrt{||u||^2||v||^2 - u \cdot v} \\ &= \sqrt{(a^2+c^2)(b^2+d^2)-(ab+cd)^2} = \sqrt{a^2d^2+b^2d^2-2abcd} = \sqrt{(ad-bc)^2} \\ &= |ad-bc| = |\det(u,v)|.\end{aligned}$$

Prouvons à présent 2). On note $u \wedge v$ le produit vectoriel de u et v. Alors $||u \wedge v||$ est égale à l'aire du parallélogramme défini par u et v. Comme de plus $u \wedge v$ est orthogonal au plan $\mathrm{Vect}(u, v)$, l'aire du parallélépipède P défini par u, v, w est égale à $|(u \wedge v) \cdot w|$. Posons alors, pour tous $v_1, v_2, v_3 \in \mathbb{R}^3$, $f(v_1, v_2, v_3) = (v_1 \wedge v_2) \cdot v_3$. On vérifie que f définit une forme 3-linéaire alternée sur $\mathbb{R}^3$. On déduit du corollaire 3.29 qu'il existe donc un réel c tel que $f(v_1, v_2, v_3) = c \det(v_1, v_2, v_3)$ pour tous $v_1, v_2, v_3 \in \mathbb{R}^3$. En évaluant f sur la base canonique (e_1, e_2, e_3) et en remarque que $e_1 \wedge e_2 = e_3$, on vérifie que $c = 1$. On peut donc bien conclure que le volume de P coïncide avec $|(u \wedge v) \cdot w| = |\det(u, v, w)|$. □

4 Réduction des endomorphismes, première partie

On fixe E un K-espace vectoriel.

4.1 Sous-espaces propres et valeurs propres

Définition 4.1. *Soit $f \in \mathcal{L}(E)$ un endomorphisme de E. Un* vecteur propre *de f est un vecteur $v \in E$ vérifiant les deux propriétés suivantes :*
- *v est non nul ;*
- *il existe $\lambda \in K$ tel que $f(v) = \lambda v$.*

Si v est un vecteur propre de f, le scalaire λ tel que $f(v) = \lambda v$ est appelé *valeur propre* de f correspondant à v. L'ensemble de toutes les valeurs propres de f est noté $\mathrm{Sp}(f)$, il s'agit du *spectre* de f.

Exemple 4.2. a) Considérons l'endomorphisme f de $\mathbb{R}^2$ défini par $f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ x \end{pmatrix}$. Le vecteur $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ est vecteur propre de f de valeur propre 1 car $u \neq 0_{\mathbb{R}^2}$ et $f(u) = u$. Le vecteur $v = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ est vecteur propre de f de valeur propre -1 car $f(v) = -v$. Le vecteur $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ n'est pas un vecteur propre de f car $f(w) \notin \mathrm{Vect}(w)$.

b) De façon générale, les vecteurs propres d'un endomorphisme f qui sont associés à la valeur propre 0 sont les éléments de $\mathrm{Ker}(f) \setminus \{0_E\}$. Ainsi $0 \in \mathrm{Sp}(f)$ si et seulement si $\mathrm{Ker}(f) \neq \{0_E\}$.

c) Soit E un K-espace vectoriel. Soit $\lambda \in K$. L'endomorphisme f de E défini par $f(v) = \lambda \cdot v$ pour tout $v \in E$ est appelé *homothétie de rapport* λ de E. Un vecteurs $v \in E$ est propre pour f si et seulement si $v \neq 0$. On a en effet $f(v) = \lambda \cdot v$ pour tout $v \in E$. Dans ce cas, f possède une unique valeur propre, λ, et $\mathrm{Sp}(f) = \{\lambda\}$.

d) Considérons $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ le $\mathbb{C}$-espace vectoriel des fonctions de classe $\mathcal{C}^\infty$ de $\mathbb{R}$ dans $\mathbb{C}$. Posons, pour $h \in E$, $f(h) = h'$. Alors f est un endomorphisme de E. Si $\lambda \in \mathbb{C}$, les vecteurs propres de f associés à la valeur propre λ sont les éléments $h \in E$ tels que $h' = \lambda h$ avec $h \neq 0$. Ce sont donc les fonctions de la forme $t \mapsto Ce^{\lambda t}$ avec $C \neq 0$. On a donc dans ce cas $\mathrm{Sp}(f) = \mathbb{C}$.

Définition 4.3. *Si $f \in \mathcal{L}(E)$ et si $\lambda \in K$, on note $E_\lambda(f) \coloneqq \mathrm{Ker}(f - \lambda \mathrm{Id}_E)$. Si λ est une valeur propre de f, on l'appelle le* sous-espace propre *de f associé à λ.*

4.2 Analogues matriciels

Définition 4.4. *Soit $A \in \mathcal{M}_n(K)$. On appelle* vecteur propre *de A est un vecteur colonne $X \in K^n$ vérifiant les deux propriétés suivantes :*

- *X est non nul;*
- *il existe $\lambda \in K$ tel que $AX = \lambda X$.*

Si X est un vecteur propre de A, le scalaire λ tel que $AX = \lambda X$ est appelé *valeur propre* de A correspondant à X. L'ensemble de toutes les valeurs propres de A est noté $\mathrm{Sp}(A)$, il s'agit du *spectre* de A. Si $\lambda \in K$, on note $E_\lambda(A) \coloneqq \mathrm{Ker}(A - \lambda I_n)$. Si λ est une valeur propre de A, on l'appelle le *sous-espace propre* de A associé à λ.

Remarque 4.5. Soit f_A l'unique endomorphisme de K^n dont la matrice dans la base canonique est A. Un vecteur colonne $X \in K^n$ est vecteur propre de A si et seulement si il est vecteur propre de f_A. De plus $\mathrm{Sp}(A) = \mathrm{Sp}(f_A)$ et, pour toute valeur propre λ de A, on a $E_\lambda(A) = E_\lambda(f_A)$.

Proposition 4.6. *Soit E un K-espace vectoriel de dimension finie. Soit $f \in \mathcal{L}(E)$ un endomorphisme de E. Soit $\mathcal{B}$ une base de E et soit $A = \mathrm{Mat}_\mathcal{B}(f)$. Soit $\lambda \in K$. Alors $X \in K^n$ est un vecteur propre de de A de valeur propre λ si et seulement si $X = [v]_\mathcal{B}$ pour $v \in E$ un vecteur propre de f de valeur propre λ. En particulier* $\mathrm{Sp}(f) = \mathrm{Sp}_K(A)$.

Démonstration. Soit $v \in E$. On a $[f(v)]_\mathcal{B} = A[v]_\mathcal{B}$. Si $X = [v]_\mathcal{B}$, on a donc $AX = \lambda X$ si et seulement si $f(v) = \lambda v$. Comme de plus $X \neq 0$ si et seulement si $v \neq 0$, on obtient le résultat. □

Corollaire 4.7. *Soit $A \in \mathcal{M}_n(K)$ et soit $P \in \mathcal{M}_n(K)$ inversible. Alors, pour tout $\lambda \in K$, on a $E_\lambda(A) = PE_\lambda(P^{-1}AP)$. En particulier* $\dim E_\lambda(A) = \dim E_\lambda(P^{-1}AP)$ *et* $\mathrm{Sp}(A) = \mathrm{Sp}(P^{-1}AP)$.

Démonstration. Soit f_A l'endomorphisme de K^n défini par $f_A(X) = AX$ et soit $\mathcal{B}$ la base de K^n définie par les vecteurs colonnes de P. On a alors $P^{-1}AP = \mathrm{Mat}_{\mathcal{B}}(f_A)$. Si $X \in K^n$, on a $[X]_{\mathcal{B}} = P^{-1}X$, de sorte que, par la proposition 4.6, on a

$$X \in E_\lambda(f_A) \Leftrightarrow P^{-1}X \in E_\lambda(P^{-1}AP).$$

On a donc $E_\lambda(P^{-1}AP) = P^{-1}E_\lambda(f_A) = P^{-1}E_\lambda(A)$. Comme $X \mapsto P^{-1}X$ est un automorphisme de K^n, on a $\dim E_\lambda(A) = \dim E_\lambda(P^{-1}AP)$. On en déduit que $E_\lambda(A) \neq \{0_{K^n}\}$ si et seulement si $E_\lambda(P^{-1}AP) \neq \{0_{K^n}\}$, c'est-à-dire $\lambda \in \mathrm{Sp}(A)$ si et seulement si $\lambda \in \mathrm{Sp}(P^{-1}AP)$. Ainsi $\mathrm{Sp}(A) = \mathrm{Sp}(P^{-1}AP)$. □

4.3 Le polynôme caractéristique

Soit $A \in \mathcal{M}_n(K)$. On note χ_A la fonction de K dans K définie par

$$\chi_A(x) \coloneqq \det(xI_n - A).$$

Théorème 4.8. *Si $A \in \mathcal{M}_n(K)$, la fonction χ_A est un polynôme unitaire de degré n à coefficients dans K.*

Démonstration. Posons, pour $1 \leqslant i, j \leqslant n$,

$$b_{i,j}(x) = \begin{cases} x - a_{i,i} & \text{si } i = j \\ -a_{i,j} & \text{si } i \neq j \end{cases}.$$

Alors $b_{i,j}$ est un polynôme de degré 1 en x si $i = j$ et de degré 0 si $i \neq j$. Par conséquent, si $\sigma \in \mathfrak{S}_n$, le polynôme $\prod_{j=1}^n b_{\sigma(j),j}$ est de degré $\mathrm{Card}\{1 \leqslant j \leqslant n \mid \sigma(j) = j\}$. Il est donc de degré n quand $\sigma = \mathrm{Id}$ et de degré $\leqslant n-1$ dans les autres cas. Ainsi

$$\chi_A(x) = \prod_{i=1}^n (x - a_{i,i}) + Q(x)$$

où Q est un polynôme de degré $\leqslant n-1$. Comme $\prod_{i=1}^n (x - a_{i,i})$ est un polynôme unitaire de degré n, on en conclut que $\chi_A(x)$ est un polynôme unitaire de degré n. □

Le polynôme χ_A est appelé *polynôme caractéristique* de A.

Définition 4.9. *Soient A et B deux matrices de $\mathcal{M}_n(K)$. On dit que A et B sont* semblables *s'il existe une matrice $P \in \mathcal{M}_n(K)$ inversible telle que $B = PAP^{-1}$.*

Proposition 4.10. *Si A et B sont deux matrices semblables de $\mathcal{M}_n(K)$, alors $\chi_A(X) = \chi_B(X)$.*

Démonstration. Soit $x \in K$. On a alors

$$\chi_B(x) = \det(xI_n - PAP^{-1}) = \det(P(xI_n - A)P^{-1}) = \det(P)\chi_A(x)\det(P)^{-1} = \chi_A(x).$$

□

Proposition 4.11. *Soient A et B deux matrices de $\mathcal{M}_n(K)$. Soit E un K-espace vectoriel de dimension n et soit $\mathcal{B}$ une base de E. Les matrices A et B sont semblables si et seulement si il existe $f \in \mathcal{L}(E)$ et une base $\mathcal{B}'$ de E tels que $A = \mathrm{Mat}_{\mathcal{B}}(f)$ et $B = \mathrm{Mat}_{\mathcal{B}'}(f)$.*

Démonstration. Supposons que $A = \mathrm{Mat}_{\mathcal{B}}(f)$, $B = \mathrm{Mat}_{\mathcal{B}'}(f)$ avec $\mathcal{B}'$ base de E. Soit $P = P_{\mathcal{B}\leftarrow\mathcal{B}'}$. Alors on a $B = P^{-1}AP$ par la proposition 1.31, de sorte que A et B sont semblables. Réciproquement supposons que A et B sont semblables. Soit $P \in \mathcal{M}_n(K)$ inversible telle que $B = P^{-1}AP$. Soit f l'endomorphisme de E tel que $\mathrm{Mat}_{\mathcal{B}}(f) = A$. Soit $\mathcal{B}'$ la base de E telle que $P_{\mathcal{B}\leftarrow\mathcal{B}'} = P$. On a alors $B = \mathrm{Mat}_{\mathcal{B}'}(f)$. □

Supposons que E est de dimension finie n. Soient $\mathcal{B}_1$ et $\mathcal{B}_2$ deux bases de E. La proposition 4.11 implique que les matrices $\mathrm{Mat}_{\mathcal{B}_1}(f)$ et $\mathrm{Mat}_{\mathcal{B}_2}(f)$ sont semblables et donc, d'après la proposition 4.10,

$$\chi_{\mathrm{Mat}_{\mathcal{B}_1}(f)}(X) = \chi_{\mathrm{Mat}_{\mathcal{B}_2}(f)}(X).$$

Le polynôme $\chi_{\mathrm{Mat}_{\mathcal{B}}(f)}(X)$ ne dépend donc pas du choix de la base $\mathcal{B}$ mais uniquement de f. On le note donc $\chi_f(X)$ et on l'appelle le *polynôme caractéristique* de f.

Théorème 4.12. *1) Soit $n \geqslant 1$ un entier et soit $A \in \mathcal{M}_n(K)$. Alors les valeurs propres de A sont exactement les racines de $\chi_A(X)$ dans k.*

2) Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$. Alors les valeurs propres de f sont exactement les racines de $\chi_f(X)$ dans K.

Démonstration. Soit $A \in \mathcal{M}_n(K)$. Un élément $\lambda \in K$ est une valeur propre de A si et seulement si $E_\lambda(A) = \mathrm{Ker}(A - \lambda I_n)$ est non réduit à $\{0_{K^n}\}$, c'est-à-dire si et seulement si la matrice $A - \lambda I_n$ n'est pas inversible, c'est-à-dire si et seulement si $\det(\lambda I_n - A) = 0$. On en déduit le résultat.

Soit à présent E un K-espace vectoriel de dimension finie et soit $\lambda \in K$. Soit $\mathcal{B}$ une base de E. Alors λ est valeur propre de f si et seulement si λ est valeur propre de $\mathrm{Mat}_{\mathcal{B}}(f)$. Comme $\chi_f(X) = \chi_{\mathrm{Mat}_{\mathcal{B}}(f)}(X)$, on déduit le résultat de la première partie du théorème. □

Corollaire 4.13. *Soit $A \in \mathcal{M}_n(K)$. Alors A a au plus n valeurs propres.*

Démonstration. En effet le polynôme $\chi_A(X)$ est de degré n et possède donc au plus n racines dans K. □

Remarque 4.14. Il faut prendre garde au fait que deux matrices peuvent avoir le même polynôme caractéristique sans être semblables. Considérons par exemple

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Alors $\chi_A(X) = \chi_B(X) = (X-1)^2$. Par contre A et B ne sont pas semblables. Si c'était le cas, il existerait $P \in \mathcal{M}_2(\mathbb{R})$ telle que

$$A = PBP^{-1} = PI_2P^{-1} = PP^{-1} = I_2$$

ce qui est faux.

Si $P \in \mathrm{GL}_n(K)$, $A \in \mathcal{M}_n(K)$ et $x \in K$, on a

$$\begin{aligned} \det(xI_n - P^{-1}AP) &= \det(P^{-1}(xI_n - A)P) \\ &= \det(P)^{-1}\det(xI_n - A)\det(P) = \det(xI_n - A), \end{aligned}$$

donc

$$\chi_{P^{-1}AP} = \chi_A. \tag{4}$$

Le polynôme caractéristique de deux matrices semblables est identique. La réciproque est fausse !!! Deux matrices peuvent avoir le même polynôme caractéristique sans être semblables.

4.4 Influence du corps des scalaires sur les valeurs propres

Lors de l'étude des valeurs propres, le choix du corps des scalaires peut avoir de l'importance.

Exemple 4.15. Supposons que $K = \mathbb{R}$ et $E = \mathbb{R}^2$. Soit $f \in \mathcal{L}(E)$ défini par

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} -y \\ x \end{pmatrix}.$$

On a $\chi_f(X) = X^2 + 1$. Ce polynôme n'a pas de racine dans $\mathbb{R}$ de sorte que $\mathrm{Sp}(f) = \emptyset$.

Exemple 4.16. Supposons maintenant que $K = \mathbb{C}$ et $E = \mathbb{C}^2$. Soit $g \in \mathcal{L}(E)$ défini par

$$g\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} -y \\ x \end{pmatrix}.$$

On a $\chi_g(X) = X^2 + 1 = X - i)(X + i)$ de sorte que $\mathrm{Sp}(g) = \{i, -i\}$.

Proposition 4.17. *Si $K = \mathbb{C}$, si E est de dimension finie non nulle, alors pour tout $f \in \mathcal{L}(E)$, on a $\mathrm{Sp}(f) \neq \emptyset$. Si $n \geqslant 1$ et $A \in \mathcal{M}_n(\mathbb{C})$, on a $\mathrm{Sp}(A) \neq \emptyset$.*

Démonstration. D'après le théorème 4.8, les polynômes χ_f et χ_A sont des polynômes de degrés $\geqslant 1$ à coefficients dans $\mathbb{C}$. Le théorème fondamental de l'algèbre implique qu'ils possèdent au moins une racine dans $\mathbb{C}$. D'après le théorème 4.12, f et A possèdent au moins une valeur propre. $\square$

4.5 Endomorphismes diagonalisables

On suppose, dans cette partie, que E est un K-espace vectoriel de dimension finie.

Définition 4.18. *Soit $f \in \mathcal{L}(E)$ un endomorphisme de E. On apelle* base propre *de f une base de E constituée de vecteurs propres de f.*

On dit que l'endomorphisme f est diagonalisable *s'il possède une base propre.*

Soit $n \geqslant 1$ et soit $A \in \mathcal{M}_n(K)$. On dit que la matrice A est diagonalisable *dans K si l'endomorphisme $f_A : X \mapsto AX$ de l'espace K^n est diagonalisable.*

Proposition 4.19. *1) Soit $f \in \mathcal{L}(E)$ un endomorphisme de E. Une base $\mathcal{B}$ de E est une base propre de f si et seulement si $\mathrm{Mat}_{\mathcal{B}}(f)$ est une matrice diagonale.*

2) Soit $f \in \mathcal{L}(E)$ un endomorphisme de E. L'endomorphisme f est diagonalisable si et seulement si il existe une base $\mathcal{B}$ de E telle que la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ est diagonale.

3) Soit $n \geqslant 1$ et soit $A \in \mathcal{M}_n(K)$. La matrice A est diagonalisable si et seulement si elle est semblables à une matrice diagonale.

4) Soit $f \in \mathcal{L}(E)$ un endomorphisme. Soit $\mathcal{B}$ est une base de E. L'endomorphisme f est diagonalisable si et seulement si la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ est une matrice diagonalisable de $\mathcal{M}_{\dim E}(K)$.

Démonstration. Démontrons le point 1). Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et soit $A = \mathrm{Mat}_{\mathcal{B}}(f) = (a_{i,j})_{1 \leqslant i,j \leqslant n}$. Comme les vecteurs e_i sont non nuls, ce sont des vecteurs propres de f si et seulement si $a_{i,j} = 0$ pour $i \neq j$, c'est-à-dire si et seulement si la matrice A est diagonale.

Le point 2) est une conséquence immédiate du 1) et de la définition d'un endomorphisme diagonalisable.

Démontrons le point 3). Soit f_A l'endomorphisme de K^n défini par $f_A(X) = AX$ pour $X \in K^n$. Par définition la matrice A est diagonalisable si et seulement si l'endomorphisme f_A est diagonalisable. On déduit du point 2) que f_A est diagonalisable si et seulement si il existe une base $\mathcal{B}$ de K^n telle que $\mathrm{Mat}_{\mathcal{B}}(f_A)$ est diagonale. On déduit de la proposition 4.11 qu'il existe une telle base si et seulement si $A = \mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(f_A)$ est semblable à une matrice diagonale.

Démontrons le point 4). On déduit de 2) que l'endomorphisme f est diagonalisable si et seulement si il existe une base $\mathcal{B}'$ de E telle que $\mathrm{Mat}_{\mathcal{B}'}(f)$ est diagonale. On déduit alors de la proposition 4.11 que cette condition est équivalente à cé que $\mathrm{Mat}_{\mathcal{B}}(f)$ soit semblable à une matrice diagonale, c'est-à-dire à ce que $\mathrm{Mat}_{\mathcal{B}}(f)$ soit diagonalisable d'après 3). $\square$

4.6 Valeurs propres et sous-espaces propres d'une matrice diagonalisable

On suppose que E est un K-espace vectoriel de dimension finie.

Théorème 4.20. *Soit $A \in \mathcal{M}_n(K)$ une matrice diagonalisable. Il existe alors une matrice inversible $P \in \mathcal{M}_n(K)$ et une matrice diagonale $D \in \mathcal{M}_n(K)$ telles que $A = PDP^{-1}$. De plus* $\mathrm{Sp}(A)$ *est l'ensemble des entrées diagonales de D et, pour $\lambda \in \mathrm{Sp}(A)$,* $\dim E_\lambda(A)$ *est le nombre d'occurences de λ sur la diagonale de D.*

Démonstration. L'existence de P et D est une conséquence de la proposition 4.19 3). Comme $\mathrm{Sp}(A) = \mathrm{Sp}(PDP^{-1}) = \mathrm{Sp}(D)$ et $\dim E_\lambda(A) = \dim E_\lambda(D)$ d'après le corollaire 4.7, on est ramené à déterminer $\mathrm{Sp}(D)$ et $\dim E_\lambda(D)$. Soient $(d_1, \dots, d_n)$ les entrées diagonales de D (c'est-à-dire que $d_i = a_{i,i}$ si $D = (a_{i,j})$). Alors $\chi_D(X) = \prod_{i=1}^n (X - d_i)$ donc $\mathrm{Sp}(D) = \{d_1, \dots, d_n\}$. Soit $\lambda \in \{d_1, \dots, d_n\}$ et calculons la dimension de $E_\lambda(D)$. Un vecteurs $X = \begin{pmatrix} x_1 \\ \vdots \\ d_n \end{pmatrix}$ est dans $E_\lambda(D)$ si et seulement si $(x_1, \dots, x_n)$ est solution du système

$$\begin{cases} d_1 x_1 & = \lambda x_1 \\ \vdots & \vdots \\ d_n x_n & = \lambda x_n \end{cases} \Leftrightarrow \begin{cases} (d_1 - \lambda) x_1 & = 0 \\ \vdots & \vdots \\ (d_n - \lambda) x_n & = 0 \end{cases} \Leftrightarrow x_i = 0 \text{ si } d_i - \lambda \neq 0.$$

Ainsi $\dim E_\lambda(D)$ est égal au nombre d'indices $1 \leqslant i \leqslant n$ tels que $d_i = \lambda$, c'est-à-dire au nombre d'occurences de λ sur la diagonale de D. □

4.7 Indépendance linéaire des vecteurs propres

Théorème 4.21. *Soit $f \in \mathcal{L}(E)$ un endomorphisme de E. Soit $(v_1, \dots, v_n)$ une famille de vecteurs de E. On suppose que chaque v_i est un vecteur propre de f de valeur propre associée λ_i. Si l'on suppose que les valeurs $\lambda, \dots, \lambda_n$ sont deux à deux distinctes, ce qui signifie $\lambda_i \neq \lambda_j$ dès que $i \neq j$, alors la famille $(v_1, \dots, v_n)$ est une famille libre.*

Démonstration. On démontre le résultat par récurrence sur $n \geqslant 1$. Le cas $n = 1$ est une conséquence directe du fait qu'un vecteur propre est non nul et engendre donc une famille libre (à un élément). Supposons le résultat démontré au rang n et montrons-le au rang $n+1$. Soit $(v_1, \dots, v_{n+1})$ une famille de vecteurs propres telle que v_i est de valeur propre λ_i pour $1 \leqslant i \leqslant n+1$ et que $\lambda_i \neq \lambda_j$ dès que $i \neq j$. Supposons qu'il existe des scalaires $x_1, \dots, x_{n+1}$ tels que $x_1 v_1 + \cdots x_{n+1} v_{n+1} = 0_E$ et montrons que $x_1 = \cdots x_{n+1}$. En appliquant f à l'égalité $x_1 v_1 + \cdots x_{n+1} v_{n+1} = 0_E$ et en multipliant cette égalité par λ_{n+1}, on obtient

$$x_1 \lambda_1 v_1 + x_2 \lambda_2 v_2 + \cdots + x_{n+1} \lambda_{n+1} v_{n+1} = 0_E$$
$$x_1 \lambda_{n+1} v_1 + x_2 \lambda_{n+1} v_2 + \cdots + x_{n+1} \lambda_{n+1} v_{n+1} = 0_E.$$

En soustrayant ces deux égalité, on obtient

$$x_1(\lambda - \lambda_{n+1} v_1 + \cdots + x_n(\lambda_n - \lambda_{n+1}) v_n = 0_E.$$

L'hypothèse de récurrence implique alors que $x_i(\lambda_i - \lambda_{n+1}) = 0$ pour tout $1 \leqslant i \leqslant n$. Comme $\lambda_i \neq \lambda_{n+1}$ pour $1 \leqslant i \leqslant n$, on en déduit que $x_i = 0$ pour $1 \leqslant i \leqslant n$. On a alors $x_{n+1}v_{n+1} = 0_E$ et, puisque $v_{n+1} \neq 0_E$, on a également $x_{n+1} = 0$. □

Corollaire 4.22. *Soit E un K-espace vectoriel et soit $f \in \mathcal{L}(E)$ un endomorphisme de E. Soient $\lambda_1, \ldots, \lambda_n$ des valeurs propres de f deux à deux distinctes. Alors les sous-espaces $E_{\lambda_1}(f), \ldots, E_{\lambda_n}(f)$ sont en somme directe.*

Démonstration. Il faut prouver que si $i_1 < \cdots < i_r$ sont des éléments de $\{1, \ldots, n\}$ et $v_1, \ldots, v_r$ des vecteurs non nuls tels que $v_j \in E_{\lambda_{i_j}}(f)$, alors $v_1 + \cdots + v_r \neq 0_E$. Or les vecteurs $v_1, \ldots, v_r$ sont des vecteurs propres associés à des valeurs propres deux à deux distinctes. On déduit donc du théorème précédent que la famille $(v_1, \ldots, v_r)$ est libre et donc que $v_1 + \cdots + v_r \neq 0_E$. □

Corollaire 4.23. *1) Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$. Alors*

$$\sum_{\lambda \in \mathrm{Sp}(f)} \dim E_\lambda(f) \leqslant \dim E.$$

2) Soit $n \geqslant 1$ un entier et soit $A \in \mathcal{M}_n(K)$. Alors

$$\sum_{\lambda \in \mathrm{Sp}(A)} \dim E_\lambda(A) \leqslant n.$$

Démonstration. Il suffit de démontrer le premier point. Comme les sous-espaces $E_\lambda(f)$ sont en somme directe, on a

$$\sum_{\lambda \in \mathrm{Sp}(f)} E_\lambda(f) = \bigoplus_{\lambda \in \mathrm{Sp}(f)} E_\lambda(f)$$

et donc

$$\dim \left(\sum_{\lambda \in \mathrm{Sp}(f)} E_\lambda(f) \right) = \sum_{\lambda \in \mathrm{Sp}(f)} \dim E_\lambda(f).$$

Comme $\sum_{\lambda \in \mathrm{Sp}(f)} E_\lambda(f)$ est un sous-espace de E, on en déduit le résultat. □

4.8 Premier critère de diagonalisation

Théorème 4.24. *1) Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$. Alors les assertions suivantes sont équivalentes*

(i) l'endomorphisme f est diagonalisable ;
(ii) on a $\sum_{\lambda \in \mathrm{Sp}(f)} E_\lambda(f) = E$;
(iii) on a $\bigoplus_{\lambda \in \mathrm{Sp}(f)} E_\lambda(f) = E$;
(iv) on a $\sum_{\lambda \in \mathrm{Sp}(f)} \dim E_\lambda(f) = \dim E$;

(v) on a $\sum_{\lambda \in \mathrm{Sp}(f)} \dim E_\lambda(f) \geqslant \dim E$.

2) Soit $n \geqslant 1$ *un entier et soit* $A \in \mathcal{M}_n(K)$. *Alors les assertions suivantes sont équivalentes*

(i) la matrice A *est diagonalisable ;*

(ii) on a $\sum_{\lambda \in \mathrm{Sp}(A)} E_\lambda(A) = K^n$ *;*

(iii) on a $\bigoplus_{\lambda \in \mathrm{Sp}(A)} E_\lambda(A) = K^n$ *;*

(iv) on a $\sum_{\lambda \in \mathrm{Sp}(A)} \dim E_\lambda(A) = n$ *;*

(v) on a $\sum_{\lambda \in \mathrm{Sp}(A)} \dim E_\lambda(A) \geqslant n$.

Démonstration. On démontre uniquement le point 1), le point 2) s'en déduisant immédiatement. Comme les sous-espaces propres sont en somme directe, les points (ii) et (iii) sont équivalents. De même l'inégalité $\sum_{\lambda \in \mathrm{Sp}(f)} \dim E_\lambda(f) \leqslant \dim E$ montre que les points (iv) et (v) sont équivalents. De plus il est clair que les points (iii) et (iv) sont équivalents. Il reste donc à prouver que (i) et (ii) sont équivalents. Supposons donc que f est diagonalisable. Il existe donc une base propre $(v_1, \ldots, v_n)$ de f. Chaque v_i est un vecteur propre et appartient donc à un sous-espace propre $E_{\lambda_i}(f)$. On a donc

$$E = \mathrm{Vect}(v_1, \ldots, v_n) \subset \sum_{i=1}^{n} E_{\lambda_i}(f) \sum_{\lambda \in \mathrm{Sp}(f)} E_\lambda(f) \subset E.$$

On en déduit l'égalité (ii). Réciproquement supposons que $E = \sum_{\lambda \in \mathrm{Sp}(f)} E_\lambda(f)$. Il existe alors une famille génératrice de E constituée de vecteurs propres. On peut extraire une base de E de cette famille pour obtenir une base propre. Ainsi f est diagonalisable. □

Corollaire 4.25. *1) Soit* E *un* K*-espace vectoriel de dimension finie et soit* $f \in \mathcal{L}(E)$. *Alors si* $\mathrm{Card}(\mathrm{Sp}(f)) = \dim E$, *l'endomorphisme* f *est diagonalisable.*

2) Soit $n \geqslant 1$ *un entier et soit* $A \in \mathcal{M}_n(K)$. *Si* $\mathrm{Card}(\mathrm{Sp}(A)) = n$, *alors* A *est diagonalisable.*

Démonstration. Nous prouvons uniquement le premier point. Si $\lambda \in \mathrm{Sp}(f)$, alors $E_\lambda(f) \neq \{0_E\}$ et donc $\dim E_\lambda(f) \geqslant 1$. On a donc

$$\sum_{\lambda \in \mathrm{Sp}(f)} \dim E_\lambda(f) \geqslant \mathrm{Card}(\mathrm{Sp}(f)) = \dim E$$

de sorte que f est diagonalisable. □

Comme $\mathrm{Sp}(f)$ est l'ensemble des racines du polynôme caractéristique $\chi_f(X)$, on peut reformuler ce corollaire sous la forme suivante.

Corollaire 4.26. *1) Soit* E *un* K*-espace vectoriel de dimension finie et soit* $f \in \mathcal{L}(E)$. *Si le polynôme caractéristique de* f *est scindé à racines simples, alors l'endomorphisme* f *est diagonalisable.*

2) Soit $n \geqslant 1$ *un entier et soit* $A \in \mathcal{M}_n(K)$. *Si le polynôme caractéristique de* A *est scindé à racines simples, alors* A *est diagonalisable.*

5 Quelques applications des matrices diagonalisables

5.1 Puissances des matrices diagonalisables

Soit K un corps égal à $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$. Soit $n \geqslant 1$ un entier et soit $A \in \mathcal{M}_n(K)$.

Supposons à présent que A est diagonalisable. Il existe alors une matrice inversible $P \in \mathcal{M}_n(K)$ et une matrice diagonalisable $D \in \mathcal{M}_n(K)$ telles que $A = PDP^{-1}$. Notons $\lambda_1, \dots, \lambda_n$ les entrées diagonales de D, de sorte que

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Proposition 5.1. *Pour tout entier $k \geqslant 0$, on a*

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

De plus A est inversible si et seulement si $\lambda_i \neq 0$ pour tout $1 \leqslant i \leqslant n$, et dans ce cas

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

Démonstration. Prouvons par récurrence sur $k \geqslant 0$ que $A^k = PD^kP^{-1}$.

Si $k = 0$, on a $A^0 = I_n$ par convention et $\lambda_i^0 = 1$ pour tout $1 \leqslant i \leqslant n$ de sorte que l'on a bien $I_n = PI_nP^{-1}$.

Supposons le résultat prouvé pour un entier $k \geqslant 0$ et prouvons le pour $k+1$. On a alors

$$A^{k+1} = AA^k = PDP^{-1}PD^kP^{-1} = PDI_nD^kP^{-1} = PD^{k+1}P^{-1}.$$

On a alors

$$D^k = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^k \end{pmatrix}.$$

Comme $E_0(A) = \mathrm{Ker}(A)$, la matrice A est inversible si et seulement si 0 n'est pas valeur propre de A, c'est-à-dire si et seulement si tous les λ_i sont non nuls. Supposons A inversible. Alors l'inverse de D est la matrice

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

et on vérifie que

$$(PD^{-1}P^{-1})A = (PD^{-1}P^{-1})PDP^{-1} = PD^{-1}I_nDP^{-1} = PP^{-1} = I_n. \qquad \square$$

5.2 Systèmes de suite récurrentes

On fixe K un corps égal à $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$. Soit $A = (a_{i,j})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(K)$ une matrice carrée. On s'intéresse aux suites $(u_k^{(1)})_{k\geqslant 0}$, $(u_k^{(2)})_{k\geqslant 0}, \ldots, (u_k^{(n)})_{k\geqslant 0}$ à valeurs dans K satisfaisant au système récurrent d'ordre 1 couplé :

$$\forall k \geqslant 0, \qquad \begin{cases} u_{k+1}^{(1)} & = a_{1,1}u_k^{(1)} + \cdots + a_{1,n}u_k^{(n)} \\ \vdots & \vdots \\ u_{k+1}^{(n)} & = a_{n,1}u_k^{(1)} + \cdots + a_{n,n}u_k^{(n)} \end{cases}. \tag{5}$$

Posons, pour tout $k \geqslant 0$, $X_k = \begin{pmatrix} u_k^{(1)} \\ \vdots \\ u_k^{(n)} \end{pmatrix}$. Le système (5) est alors équivalent à

$$\forall k \geqslant 0, \qquad X_{k+1} = AX_k.$$

On suppose à présent que A est diagonalisable. On peut alors écrire $A = PDP^{-1}$ avec P inversible et D diagonale de la forme

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Posons, pour $k \geqslant 0$, $Y_k = P^{-1}X_k$. On a alors, pour tout $k \geqslant 0$,

$$\begin{aligned} X_{k+1} = AX_k \Leftrightarrow PY_{k+1} = APY_k &\Leftrightarrow Y_{k+1} = P^{-1}APY_k \\ &\Leftrightarrow Y_{k+1} = DY_k \end{aligned}$$

Posons, pour tout $k \geqslant 0$, $Y_k \begin{pmatrix} v_k^{(1)} \\ \vdots \\ v_k^{(n)} \end{pmatrix}$. Le système (5) est donc équivalent au système

$$\forall k \geqslant 0, \qquad \begin{cases} v_{k+1}^{(1)} & = \lambda_1 v_k^{(1)} \\ \vdots & \vdots \\ v_{k+1}^{(n)} & = \lambda_n v_k^{(n)} \end{cases} \tag{6}$$

qui est à présent un système découplé de n suites récurrentes linéaires d'ordre 1. On en déduit que

$$\forall k \geqslant 0, \quad \begin{cases} v_k^{(1)} & = \lambda_1^k v_0^{(1)} \\ \vdots & \vdots \\ v_k^{(n)} & = \lambda_n^k v_0^{(n)} \end{cases}$$

On peut utiliser P^{-1} pour exprimer $v_0^{(1)}, \dots, v_0^{(n)}$ en fonction de $u_0^{(n)}, \dots, u_0^{(n)}$ et utiliser P pour retrouver les valeurs de $u_k^{(1)}, \dots, u_k^{(n)}$.

Théorème 5.2. *Soient $a_0, \dots, a_{n-1} \in K$. On suppose que le polynôme $P(X) = X^n - a_{n-1}X^{n-1} - \cdots - a_1 X - a_0$ est scindé à racines simples dans K, c'est-à-dire de la forme $P(X) = \prod_{i=1}^n (X - \lambda_i)$ avec $\lambda_1, \dots, \lambda_n$ deux à deux distincts dans K. Alors l'ensemble $\mathcal{S}$ des suites $(u_k)_{k \geqslant 0}$ à valeurs dans K et satisfaisant la relation*

$$\forall k \geqslant 0, \qquad u_{k+n} = a_{n-1}u_{k+n-1} + \cdots + a_1 u_{k+1} + a_0 u_k \tag{7}$$

est l'ensembles des combinaisons linéaires, à coefficients dans K, des suites $(\lambda_i^k)_{k \geqslant 0}$ pour $i = 1, \dots, n$.

Démonstration. Comme λ_i est racine de P, on a $\lambda_i^n = a_{n-1}\lambda_i^{n-1} + \cdots + a_1\lambda_i + a_0$. On en déduit donc que la suite $(\lambda_i^k)_{k \geqslant 0}$ est dans $\mathcal{S}$. Toute combinaison linéaire de ces suites est donc également dans $\mathcal{S}$.

Réciproquement soit $(u_k)_{k \geqslant 0} \in \mathcal{S}$. On pose $X_k = \begin{pmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+n-1} \end{pmatrix}$. La relation (7) est alors équivalente à $X_{k+1} = AX_k$ où A désigne la matrice

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 \\ a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \end{pmatrix}.$$

Le polynôme caractéristique de la matrice A est égal au polynôme $P = X^n - a_{n-1}X^{n-1} - \cdots - a_1X - a_0$. Ainsi $\mathrm{Sp}(A) = \{\lambda_1, \ldots, \lambda_n\}$. Comme les λ_i sont deux à deux distincts, on a $\mathrm{Card}(\mathrm{Sp}(A)) = n$ et la matrice A est diagonalisable d'après le corollaire 4.26. Il existe donc une matrice inversible $P \in \mathcal{M}_n(K)$ tel que $A = PDP^{-1}$ avec

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

En posant $Y_k = P^{-1}X_k$ pour tout $k \geqslant 0$, on a donc, comme précédemment $Y_k = \begin{pmatrix} \lambda_1^k c_1 \\ \vdots \\ \lambda_n^k c_n \end{pmatrix}$ pour certaines conditions initiales $c_1, \ldots, c_n$. La relation $X_k = PY_k$ implique donc que la suite $(u_k)_{k\geqslant 0}$ est combinaison linéaire des suites $(\lambda_1^k)_{k\geqslant 0}, \ldots, (\lambda_n^k)_{k\geqslant 0}$. □

5.3 Systèmes d'équations différentielles linéaires

On fixe K un corps égal à $\mathbb{R}$ ou $\mathbb{C}$. Soit $A = (a_{i,j})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(K)$ une matrice carrée. On se propose de résoudre le problème suivant. On cherche des fonctions $x_1, \ldots, x_n$ de $\mathbb{R}$ dans K, de classe $\mathcal{C}^1$ telles que

$$\forall t \in \mathbb{R}, \quad \begin{cases} x_1'(t) & = a_{1,1}x_1(t) + \cdots + a_{1,}x_n(t) \\ \vdots & \vdots \\ x_n'(t) & = a_{n,1}x_1(t) + \cdots + a_{n,n}x_n(t) \end{cases}. \tag{8}$$

Posons, pour tout $t \in \mathbb{R}$,

$$X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \qquad X'(t) = \begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix}.$$

Résoudre le système (8) revient à rechercher les fonctions $t \mapsto X(t)$ de $\mathbb{R}$ dans K dont les coordonnées sont de classe $\mathcal{C}^1$ et telles que

$$\forall t \in \mathbb{R}, \qquad X'(t) = AX(t).$$

On suppose à présent que A est diagonalisable. On peut alors écrire $A = PDP^{-1}$ avec P inversible et D diagonale de la forme

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Posons, pour $t \in \mathbb{R}$, $Y(t) = P^{-1}X(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$. Alors les fonctions y_i sont des fonctions de classe $\mathcal{C}^1$ de $\mathbb{R}$ dans K et on a alors, pour tout $t \in \mathbb{R}$,

$$\begin{aligned} X'(t) = AX(t) &\Leftrightarrow PY(t) = APY(t) \Leftrightarrow Y(t) = P^{-1}APY(t) \\ &\Leftrightarrow Y'(t) = DY(t) \end{aligned}$$

Le système (8) est donc équivalent au système

$$\forall t \in \mathbb{R}, \qquad \begin{cases} y_1'(t) & = \lambda_1 y_1(t) \\ \vdots & \vdots \\ y_n'(t) & = \lambda_n y_n(t) \end{cases} \tag{9}$$

qui est à présent un système découplé de n équations différentielles liénaires d'ordre 1. On en déduit que

$$\forall t \in \mathbb{R}, \quad \begin{cases} y_1(t) & = e^{\lambda_1 t} y_1(0) \\ \vdots & \vdots \\ y_n(t) & = e^{\lambda_n t} y_n(0) \end{cases}$$

On peut utiliser P^{-1} pour exprimer $y_1(0), \dots, y_n(0)$ en fonction de $x_1(0), \dots, x_n(0)$ et utiliser P pour expliciter $x_1, \dots, x_n$.

Théorème 5.3. *Soient $a_0, \dots, a_{n-1} \in K$. On suppose que le polynôme $P(X) = X^n - a_{n-1}X^{n-1} - \cdots - a_1 X - a_0$ est scindé à racines simples dans K, c'est-à-dire de la forme $P(X) = \prod_{i=1}^n (X - \lambda_i)$ avec $\lambda_1, \dots, \lambda_n$ deux à deux distincts dans K. Alors l'ensemble $\mathcal{S}$ des fonctions x de classes $\mathcal{C}^n$ de $\mathbb{R}$ dans K satisfaisant l'équation différentielle*

$$\forall t \in \mathbb{R}, \qquad x^{(n)}(t) = a_{n-1}x^{(n-1)}(t) + \cdots + a_1 x'(t) + a_0 x(t) \tag{10}$$

est l'ensembles des combinaisons linéaires, à coefficients dans K, des fonctions $t \mapsto e^{\lambda_i t}$ pour $i = 1, \dots, n$.

Démonstration. Comme λ_i est racine de P, on a $\lambda_i^n = a_{n-1}\lambda_i^{n-1} + \cdots + a_1\lambda_i + a_0$. On en déduit facilement que la fonction $(t \mapsto e^{\lambda_i t})$ est dans $\mathcal{S}$. Toute combinaison linéaire de ces fonctions est donc également dans $\mathcal{S}$.

Réciproquement soit $x \in \mathcal{S}$. On pose, pour $t \in \mathbb{R}$, $X(t) = \begin{pmatrix} x(t) \\ x'(t) \\ \vdots \\ x^{(n-1)}(t) \end{pmatrix}$. La relation

(10) est alors équivalente à $X' = AX$ où A désigne la matrice

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 \\ a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \end{pmatrix}.$$

Comme dans la démonstration du théorème 5.2, le polynôme caractéristique de la matrice A est égal au polynôme $P = X^n - a_{n-1}X^{n-1} - \cdots - a_1X - a_0$. Ainsi $\mathrm{Sp}(A) = \{\lambda_1, \ldots, \lambda_n\}$. Comme les λ_i sont deux à deux distincts, on a $\mathrm{Card}(\mathrm{Sp}(A)) = n$ et la matrice A est diagonalisable d'après le corollaire 4.26. Il existe donc une matrice inversible $P \in \mathcal{M}_n(K)$ tel que $A = PDP^{-1}$ avec

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

En posant $Y(t) = P^{-1}X(t)$ pour tout $t \in \mathbb{R}$, on a donc, comme précédemment $Y(t) = \begin{pmatrix} e^{\lambda_1 t}c_1 \\ \vdots \\ e^{\lambda_n t}c_n \end{pmatrix}$ pour certaines conditions initiales $c_1, \ldots, c_n \in K$. La relation $X = PY$ implique donc que la focntion x est combinaison linéaire des fonctions $(t \mapsto e^{\lambda_i t})_{1 \leqslant i \leqslant n}$. $\square$

6 Sous-espace stables et polynômes d'endomorphismes

6.1 Notion de sous-espace stable

Définition 6.1. *Soit E un K-espace vectoriel et soit $f \in \mathcal{L}(E)$ un endomorphisme de E. Un sous-espace vectoriel $F \subset E$ est dit* stable *par f si $f(F) \subset F$.*

Proposition 6.2. *Si F_1 et F_2 sont deux sous-espaces vectoriels de E stables par f, alors $F_1 \cap F_2$ et $F_1 + F_2$ le sont aussi.*

Exemple 6.3. — Les sous-espaces vectoriels $\{0_E\}$ et E sont toujours stables.
- — Soit $v \in E$ avec $v \neq 0_E$. Alors $\mathrm{Vect}(v)$ est stable par f si et seulement si v est un vecteur propre de f.
- — Pour tout scalaire $\lambda \in K$, le sous-espace propre $E_\lambda(f)$ est stable par f.

Proposition 6.4. *Soient f et g deux endomorphismes de E tels que $g \circ f = f \circ g$. Alors les sous-espaces* $\mathrm{Ker}(g)$ *et* $\mathrm{Im}(g)$ *sont stables par f.*

Démonstration. Soit $v \in \mathrm{Ker}(g)$. Alors $g(f(v)) = f(g(v)) = f(0_E) = 0_E$ donc $f(v) \in \mathrm{Ker}(g)$ et $\mathrm{Ker}(g)$ est stable par f.

Soit $v \in \mathrm{Im}(g)$. Alors il existe $w \in E$ tel que $v = g(w)$. On a donc $f(v) = f(g(w)) = g(f(w))$ de sorte que $f(v) \in \mathrm{Im}(g)$. Ainsi $\mathrm{Im}(g)$ est stable par f. □

6.2 Endomorphisme induit

Soit $F \subset E$ un sous-espace vectoriel stable par f. On note f_F l'endomorphisme de F défini par $f_F(v) = f(v)$ pour tout $v \in F$. Notons que f_F est bien défini puisque F est stable par f. On l'appelle l'endomorphisme de F *induit* par f.

Remarquons que si $f, g \in \mathcal{L}(E)$ et si F est stable par f et g, alors pour tout $\lambda \in K$ le sous-espace F est stable par $f + \lambda g$ et par $g \circ f$, et on a

$$(f + \lambda g)_F = f_F + \lambda g_F, \qquad (g \circ f)_F = g_F \circ f_F.$$

Exemple 6.5. Soit E un espace vectoriel de dimension 4 et soit $\mathcal{B} = (v_1, v_2, v_3, v_4)$ une base de E. Soit $f \in \mathcal{L}(E)$ tel que

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 2 & 3 & 5 \\ 0 & 7 & 3 & 0 \\ 0 & -1 & -2 & 0 \\ 4 & -3 & 6 & -2 \end{pmatrix}.$$

Soit $F = \mathrm{Vect}(v_1, v_4)$ et soit $\mathcal{B}_F = (v_1, v_4)$ une base de F. Alors F est stable par f et

$$\mathrm{Mat}_{\mathcal{B}_F}(f_F) = \begin{pmatrix} 1 & 5 \\ 4 & -2 \end{pmatrix}.$$

Remarque 6.6. Remarquons de plus que si v est un vecteur propre de f_F, alors v est un vecteur propre de f et que $\mathrm{Sp}(f_F) \subset \mathrm{Sp}(f)$.

6.3 Matrices triangulaires par blocs

Proposition 6.7. *Soit E un K-espace vectoriel de dimension finie. Soit $f \in \mathcal{L}(E)$ un endomorphisme de E. Soit $\mathcal{B} = (v_1, \ldots, v_n)$ une base de E. Soit $1 \leqslant r \leqslant n$ et posons $F = \mathrm{Vect}(v_1, \ldots, v_r)$ et $\mathcal{B}_F = (v_1, \ldots, v_r)$. Alors le sous-espace F est stable par f si et seulement si la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$ est de la forme*

$$\begin{pmatrix} A & B \\ 0_{n-r,r} & D \end{pmatrix}$$

avec $A \in \mathcal{M}_r(K)$, $B \in \mathcal{M}_{r,n-r}(K)$ et $D \in \mathcal{M}_{n-r}(K)$. Dans ce cas on a de plus $A = \mathrm{Mat}_{\mathcal{B}_F}(f_F)$.

Démonstration. Notons $a_{i,j}$ les coefficients de la matrice $\mathrm{Mat}_{\mathcal{B}}(f)$. Alors, pour tout $1 \leqslant j \leqslant n$, on a

$$f(v_j) = \sum_{i=1}^{n} a_{i,j} v_i.$$

Ainsi F est stable par f si et seulement si, pour tout $1 \leqslant j \leqslant r$, $f(v_j) \in F$, c'est-à-dire $a_{i,j} = 0$ pour $r+1 \leqslant i \leqslant n$. $\square$

Corollaire 6.8. *Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$. Soit F u sous-espace vectoriel de E. Si F est stable par f, le polynôme caractéristique de f_F divise le polynôme caractéristique de f :*

$$\chi_{f_F}(X) \mid \chi_f(X).$$

Démonstration. Soit $(v_1, \ldots, v_r)$ une base de F. On peut la compléter en une base $\mathcal{B} = (v_1, \ldots, v_n)$ de E. On a alors

$$M = \mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

avec $A = \mathrm{Mat}_{(v_1,\ldots,v_r)}(f_F)$. On a alors

$$\chi_f(X) = \chi_M(X) = \det(XI_n - M) = \det(XI_r - A)\det(XI_{n-r} - D) = \chi_{f_F}(X)\det(XI_{n-r} - D).$$

Ainsi $\chi_{f_F}(X)$ divise $\chi_f(X)$ dans $K[X]$. $\square$

6.4 Rappels sur les polynômes

Voir cours de première année.

6.5 Polynômes d'endomorphismes

Soit $A \in \mathcal{M}_n(K)$. Si $P(X) = \sum_{i=0}^{n} a_i X^i \in K[X]$, on pose

$$P(A) = \sum_{i=0}^{n} a_i A^i \in \mathcal{M}_n(K)$$

avec la convention $A^0 = I_n$ si $A \neq 0_n$ et $0_n^0 = 0_n$.

Exemple 6.9. Prenons $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ et $P(X) = X^2 + X + 1$. On a alors

$$P(A) = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}.$$

On peut remplacer A par un endomorphisme f d'un espace vectoriel E en remplaçant A^n par

$$f^n = \underbrace{f \circ \cdots \circ f}_{n}.$$

Par convention, on pose $f^0 = \mathrm{Id}_E$.

Soit $P = \sum_{i=0}^n a_i X^i \in K[X]$ un polynôme. On définit alors

$$P(f) \coloneqq \sum_{i=0}^n a_i f^i \in \mathcal{L}(E).$$

L'application de $K[X]$ vers $\mathcal{L}(E)$ définie par $P \mapsto P(f)$ est linéaire, ce qui signifie que, pour P et Q dans $K[X]$ et $\lambda \in K$, on a $(P + \lambda Q)(f) = P(f) + \lambda Q(f)$. On a également, pour P et Q dans $K[X]$,

$$(PQ)(f) = P(f) \circ Q(f) = Q(f) \circ P(f).$$

Remarque 6.10. Si E est un espace vectoriel de dimension finie et si $\mathcal{B}$ est une base de E, on a les relations

$$P(\mathrm{Mat}_{\mathcal{B}}(f)) = \mathrm{Mat}_{\mathcal{B}}\, P(f).$$

On peut déduire de cette formule que pour $P(X) \in K[X]$, $A, B \in \mathcal{M}_n(K)$ avec B inversible,

$$P(BAB^{-1}) = BP(A)B^{-1}.$$

Corollaire 6.11. *Soit $f \in \mathcal{L}(E)$ et soit $P(X) \in K[X]$. Alors* $\mathrm{Ker}(P(f))$ *et* $\mathrm{Im}(P(f))$ *sont des sous-espaces vectoriels de E stables par f.*

Démonstration. En effet, f commute avec $P(f)$. □

6.6 Polynômes annulateurs

Soit E un K-espace vectoriel et soit $f \in \mathcal{L}(E)$ un endomorphisme de E. Un polynôme $P \in K[X]$ est dit *annulateur* de f si $P(f) = 0_{\mathcal{L}(E)}$.

Remarque 6.12. Si P est annulateur de f, alors PQ aussi pour tout $Q \in K[X]$. En effet, on a alors

$$(PQ)(f) = P(f) \circ Q(f) = 0_{\mathcal{L}(E)} \circ Q(f) = 0_{\mathcal{L}(E)}.$$

Exemple 6.13. 1) Si $\lambda \in K$, alors $(X - \lambda)$ et $X(X - \lambda)$ sont des polynômes annulateurs de l'endomorphisme $\lambda \mathrm{Id}_E$.

2) Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(K)$, alors $X^2 - (a+d)X + ad - bc$ est annulateur de A.

Proposition 6.14. *Soit E un K-espace vectoriel de dimension finie. Alors tout endomorphsime f de E possède un polynôme annulateur non nul.*

Démonstration. Le K-espace vectoriel $\mathcal{L}(E)$ est de dimension n^2 où $n = \dim E$. La famille $(\mathrm{Id}_E, f, f^2, \ldots, f^{n^2})$ est donc liée. Il existe donc des scalaires $a_0, a_1, \ldots, a_{n^2} \in K$ non tous nuls tels que

$$a_0 \mathrm{Id}_E + a_1 f + \cdots + a_{n^2} f^{n^2} = 0_{\mathcal{L}(E)}.$$

Ainsi f est annulé par le polynôme non nul $a_0 + a_1 X + \cdots + a_{n^2} X^{n^2}$. □

Corollaire 6.15. *Soit $A \in \mathcal{M}_n(K)$. Alors il existe un polynôme non nul $P \in K[X]$ tel que $P(A) = 0$.*

Théorème 6.16. *Soit E un K-espace vectoriel et soit $f \in \mathcal{L}(E)$. Soit $P \in K[X]$ un polynôme annulateur de f, non nul et de degré minimal parmi les polynômes annulateurs non nuls de f. Alors $Q \in K[X]$ annule f si et seulement si P divise Q dans $K[X]$.*

Démonstration. On a déjà vu que si P divise Q, alors $Q(f) = 0_{\mathcal{L}(E)}$. Montrons l'implication réciproque. Supposons que $Q(f) = 0_{\mathcal{L}(E)}$ et effectuons la division euclidienne de Q par P. On a donc $Q = PR + S$ avec $R, S \in K[X]$ et $\deg S < \deg P$. Comme $Q(f) = P(f) = 0_{\mathcal{L}(E)}$, on a aussi $S(f) = 0_{\mathcal{L}(E)}$. Comme $\deg S < \deg P$ et comme P est de degré minimal parmi les polynômes annulateurs non nuls de f, on a nécessairement $S = 0$, c'est-à-dire que P divise Q. □

Corollaire 6.17. *Soit E un K-espace vectoriel de dimension finie. Alors il existe un unique polynôme unitaire annulant f et de degré minimal parmi les polynômes annulateurs non nuls de f.*

Démonstration. Supposons que P_1 et P_2 sont deux tels polynômes. Par minimalité de leur degré, on a $\deg P_1 = \deg P_2$. De plus le théorème implique que P_1 divise P_2. On peut écrire $P_2 = P_1 Q$ avec $\deg Q = 0$. Ainsi Q est un polynôme constant de valeur λ. Comme P_1 et P_2 sont unitaires, $\lambda = 1$ et $P_1 = P_2$. □

Corollaire 6.18. *Soit $n \geqslant 1$ un entier et soit $A \in \mathcal{M}_n(K)$. Alors il existe un unique polynôme unitaire annulant A et de degré minimal parmi les polynômes annulateurs non nuls de A.*

Définition 6.19. *Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$. On appelle* polynôme minimal *de f l'unique polynôme unitaire annulant f et de degré minimal parmi les polynômes annulateurs non nuls de f. On le note $\pi_f(X)$.*

Si $n \geqslant 1$ et si $A \in \mathcal{M}_n(K)$, on appelle polynôme minimal *de A l'unique polynôme unitaire annulant A et de degré minimal parmi les polynômes annulateurs non nuls de A. On le note $\pi_A(X)$.*

On a alors, pour tout $Q \in K[X]$,

$$Q(A) = 0 \Leftrightarrow \pi_A | Q.$$

Exemple 6.20. 1) Soit $\lambda \in K$ non nul et considérons $f = \lambda \mathrm{Id}$. Comme f est non nul, f n'est pas annulé par un polynôme de degré 0. Par contre le polynôme unitaire $X - \lambda$, de degré 1 annule f, on a donc $\pi_f(X) = X - \lambda$.

2) Pour un endomorphisme f, on a $\deg(\pi_f) = 0$ si et seulement si $f = 0_{\mathcal{L}(E)}$. En effet dans ce cas $\pi_f = 1$.

3) On a $\deg(\pi_f) = 1$ si et seulement si f est une homothétie non nulle. Dans ce cas on a en effet $\pi_f(X) = X - \lambda$. Comme $\pi_f(f) = 0_{\mathcal{L}(E)}$, on a bien $f = \lambda \mathrm{Id}_E$.

4) Considérons $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Comme A n'est pas une matrice d'homothétie, on a $\deg(\pi_1) \geqslant 2$. Par ailleurs $A^2 = 3A$. Ainsi A est annulée par $X^2 - 3X$ et $\pi_A(X) = X^2 - 3X$.

6.7 Le lemme des noyaux

Soit E un K-espace vectoriel et soit $f \in \mathcal{L}(E)$.

Théorème 6.21. *Soient P et Q deux polynômes de $K[X]$ premiers entre eux. On a alors*

$$\mathrm{Ker}((PQ)(f)) = \mathrm{Ker}(P(f)) \oplus \mathrm{Ker}(Q(f)).$$

Démonstration. Montrons dans un premier temps que $\mathrm{Ker}(P(f)) \cap \mathrm{Ker}(Q(f)) = \{0_E\}$. Soit $v \in \mathrm{Ker}(P(f)) \cap \mathrm{Ker}(Q(f))$. Comme P et Q sont premiers entre eux, le théorème de Bezout implique qu'il existe deux polynômes $A, B \in K[X]$ tels que $AP + BQ = 1$. En appliquant cette égalité à f, on obtient

$$\mathrm{Id}_E = A(f) \circ P(f) + B(f) \circ Q(f).$$

On en déduit que

$$v = A(f)(P(f)(v)) + B(f)(Q(f)(v)) = A(f)(0_E) + B(f)(0_E) = 0_E.$$

Ainsi $\mathrm{Ker}(P(f)) \cap \mathrm{Ker}(Q(f)) = \{0_E\}$.

Montrons à présent que $\mathrm{Ker}(P(f)) \subset \mathrm{Ker}((PQ(f)))$. Soit $v \in \mathrm{Ker}(P(f))$. On a alors

$$(PQ)(f)(v) = (Q(f) \circ P(f))(v) = Q(f)(P(f)(v)) = Q(f)(0_E) = 0_E.$$

Ainsi $v \in \mathrm{Ker}((PQ)(f))$. On en déduit que $\mathrm{Ker}(P(f)) \subset \mathrm{Ker}((PQ(f)))$.

On montre de même que $\mathrm{Ker}(Q(f)) \subset \mathrm{Ker}((PQ(f)))$ de sorte que

$$\mathrm{Ker}(P(f)) + \mathrm{Ker}(Q(f)) \subset \mathrm{Ker}((PQ(f))).$$

Montrons enfin que $\mathrm{Ker}((PQ)(f)) \subset \mathrm{Ker}(P(f)) + \mathrm{Ker}(Q(f))$. Soit $v \in \mathrm{Ker}((PQ)(f))$. Rappelons que l'on a

$$v = (AP)(f)(v) + (BQ)(f)(v).$$

Posons $v_1 = (BQ)(f)(v)$ et $v_2 = (AP)(f)(v)$. Il suffit de montrer que $v_1 \in \mathrm{Ker}(P(f))$ et $v_2 \in \mathrm{Ker}(Q(f))$ pour conclure que $v \in \mathrm{Ker}(P(f)) + \mathrm{Ker}(Q(f))$. Or on a

$$P(f)(v_1) = (P(f) \circ (BQ)(f))(v) = (PBQ)(f)(v) = B(f)((PQ)(f)(v)) = B(f)(0_E) = 0_E.$$

Ainsi $v_1 \in \mathrm{Ker}(P(f))$. On montre de façon analogue que $v_2 \in \mathrm{Ker}(Q(f))$. □

Corollaire 6.22. *Soient $P_1, \ldots, P_r \in K[X]$ des polynômes premiers entre eux deux à deux. On a alors*

$$\mathrm{Ker}((P_1 P_2 \cdots P_r)(f)) = \mathrm{Ker}(P_1(f)) \oplus \cdots \oplus \mathrm{Ker}(P_r(f)).$$

Démonstration. La démonstration se fait par récurruence sur r en utilisant le théorème. □

6.8 Un critère de diagonalisabilité

Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$.

Théorème 6.23. *Les assertions suivantes sont équivalentes :*

(i) l'endomorphisme f est diagonalisable ;

(ii) l'endomorphisme f est annulé par un polynôme scindé à racines simples ;

(iii) le polynôme minimal de f est scindé à racines simples.

Démonstration. Montrons que (i) implique (ii). Supposons f diagonalisable et posons $\mathrm{Sp}(f) = \{\lambda_1, \ldots, \lambda_r\}$ avec $\lambda_1, \ldots, \lambda_r$ distinctes. Posons $P = \prod_{i=1}^{r}(X - \lambda_i)$. Montrons que $P(f) = 0_{\mathcal{L}(E)}$. Comme f est diagonlisable, on a

$$E = E_{\lambda_1}(f) \oplus \cdots \oplus E_{\lambda_r}(f).$$

Soit $1 \leqslant i \leqslant r$. On peut écrire $P = Q_i(X)(X - \lambda_i)$ avec $Q_i(X) = \prod_{j \neq i}(X - \lambda_j)$. Si $v \in E_{\lambda_i}(f)$, on a

$$P(f)(v) = Q_i(f)((f - \lambda_i \mathrm{Id})(v)) = Q_i(f)(0_E) = 0_E.$$

Plus généralement, pour tout $v \in E$, on peut écrire

$$v = v_1 + \cdots + v_r$$

avec $v_i \in E_{\lambda_i}(f)$ pour tout $1 \leqslant i \leqslant r$ et on a

$$P(f)(v) = P(f)(v_1) + \cdots + P(f)(v_r) = 0_E.$$

Comme $P(f)(v) = 0_E$ pour tout $v \in E$, on a bien $P(f) = 0_{\mathcal{L}(E)}$.

Montrons que (ii) implique (iii). Supposons donc que f est annulé par un polynôme scindé à racines simples $P(X)$. Alors d'après le théorème 6.16, le polynôme minimal $\pi_f(X)$ de f divise $P(X)$, ce qui implique que $\pi_f(X)$ est également scindé à racines simples.

Montrons que (iii) implique (i). En supposant (iii), on peut écrire $\pi_f(X) = (X - \lambda_1) \cdots (X - \lambda_r)$ avec les λ_i distincts. En particulier les polynômes $X - \lambda_i$ sont deux à deux distincts. Comme $\pi_f(f) = 0_{\mathcal{L}(E)}$, le lemme des noyaux implique alors qu'on a

$$E = \operatorname{Ker}(\pi_f(f)) = \bigoplus_{i=1}^{r} \operatorname{Ker}(f - \lambda_i \mathrm{Id}_E) = \bigoplus_{i=1}^{r} E_{\lambda_i}(f).$$

On en conclut, en utilisant le théorème 4.24, que f est diagonalisable. □

Corollaire 6.24. *Soit $A \in \mathcal{M}_n(K)$ une matrice carrée. Les assertions suivantes sont équivalentes :*

(i) la matrice A est diagonalisable sur K ;

(ii) la matrice A est annulé par un polynôme scindé à racines simples de $K[X]$;

(iii) le polynôme minimal de A est scindé à racines simples dans $K[X]$.

Proposition 6.25. *Les racines du polynôme minimal π_f sont exactement les valeurs propres de f.*

Démonstration. Supposons que $\lambda \in \mathrm{Sp}(f)$. Il existe alors $v \in E$, $v \neq 0_E$ tel que $f(v) = \lambda v$. On a alors, pour tout $n \geqslant 1$, $f^n(v) = \lambda^n v$ et donc, pour tout $P \in K[X]$, $P(f)(v) = P(\lambda)v$. En particulier, comme $\pi_f(f) = 0_{\mathcal{L}(E)}$, on a $\pi_f(\lambda)v = 0_E$ et donc, puisque $v \neq 0_E$, $\pi_f(\lambda) = 0$.

Réciproquement supposons que $\pi_f(\lambda) = 0$. On peut écrire $\pi_f(X) = (X - \lambda)Q(X)$ avec $\deg Q = \deg \pi_f - 1$. On a alors

$$0_{\mathcal{L}(E)} = \pi_f(f) = (f - \lambda \mathrm{Id}_E) \circ Q(f).$$

Comme $\deg Q < \deg \pi_f$ et $Q \neq 0$, on a $Q(f) \neq 0_{\mathcal{L}(E)}$. En particulier l'endomorphisme $f - \lambda \mathrm{Id}_E$ n'est pas inversible. On déduit alors du théorème du rang que $E_\lambda(f) = \operatorname{Ker}(f - \lambda \mathrm{Id}_E) \neq 0$. Ainsi λ est valeur propre de f. □

Corollaire 6.26. *Soit $A \in \mathcal{M}_n(A)$. Les racines du polynôme minimal π_A dans K sont exactement les valeurs propres de A dans K.*

Corollaire 6.27. *Soit P un polynôme annulateur de f, alors les valeurs propres de f sont des racines de P.*

7 Trigonalisation

7.1 Endomorphismes trigonalisables

Soit E un espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$ un endomorphisme de E. On dit que f est *trigonalisable* s'il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(f)$ est triangulaire supérieure.

Soit $n \geqslant 1$ un entier et soit $A \in \mathcal{M}_n(K)$ une matrice carrée de taille n. On dit que A est *trigonalisable* sur K si et seulement s'il existe une matrice inversible $P \in \mathcal{M}_n(K)$ telle que $P^{-1}AP$ est triangulaire supérieure.

Remarque 7.1. Soit $\mathcal{B}$ une base de E. Alors f est trigonalisable si et seulement si $\mathrm{Mat}_{\mathcal{B}}(f)$ est trigonalisable sur K.

Théorème 7.2. *Une matrice $A \in \mathcal{M}_n(K)$ est trigonalisable sur K si et seulement si son polynôme caractéristique est scindé sur K.*

Démonstration. Supposons que A est trigonalisable. Il existe alors une matrice inversible P et une matrice triangulaire supérieure $T \in \mathcal{M}_n(K)$ telle que $A = PTP^{-1}$. On a alors $\chi_A = \chi_T$. Soit $t_{1,1}, \ldots, t_{n,n}$ les entrées diagonales de T. Le calcul du polynôme caractéristique d'une matrice triangulaire supérieure nous donne

$$\chi_A = \chi_T = \prod_{i=1}^{n}(X - t_{i,i})$$

ce qui montre que χ_A est scindé sur K.

On va montrer par récurrence sur $n \geqslant 1$ que si χ_A est scindé sur K, alors A est trigonalisable sur K. Le cas où $n = 1$ est trivial car toute matrice est alors trigonalisable et tout polynôme de degré 1 est scindé. Supposons le résultat démontré au rang n et soit $A \in \mathcal{M}_{n+1}(K)$ telle que χ_A est scindé. Soit $\lambda \in K$ une racine de χ_A, c'est-à-dire une valeur propre de A. Soit $v \in K^{n+1}$ un vecteur propre de A, c'est-à-dire un vecteur non nul vérifiant $Av = \lambda v$. Soit $\mathcal{B} = (v, v_2, \ldots, v_{n+1})$ une base de K^{n+1} et soit P la matrice de passage de la base canonique à la base $\mathcal{B}$. La matrice $P^{-1}AP$ est alors une matrice par blocs de la forme

$$P^{-1}AP = \begin{pmatrix} \lambda & B \\ 0_{n,1} & D \end{pmatrix}$$

avec $B \in \mathcal{M}_{1,n}(K)$ et $D \in \mathcal{M}_{n,n}(K)$. De plus on a une factorisation de polynôme caractéristiques $\chi_A(X) = (X - \lambda)\chi_D(X)$. Ainsi χ_D est scindé également. Par récurrence il existe une matrice inversible $Q \in \mathcal{M}_n(K)$ telle que $Q^{-1}DQ$ soit triangulaire supérieure. Posons

$$Q' = \begin{pmatrix} 1 & 0_{1,n} \\ 0_{n,1} & Q \end{pmatrix} \in \mathcal{M}_{n+1}(K).$$

Il s'agit d'une matrice inversible et

$$(Q')^{-1}P^{-1}APQ' = \begin{pmatrix} \lambda & B \\ 0_{n,1} & Q^{-1}DQ \end{pmatrix}$$

est une matrice triangulaire supérieure. Ainsi A est trigonalisable, ce qui achève la récurrence. $\square$

Corollaire 7.3. *Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$ un endomorphisme. Alors f est trigonalisable si et seulement si χ_f est scindé sur K.*

Corollaire 7.4. *Si $K = \mathbb{C}$, tout endomorphisme de E est trigonalisable et toute matrice de $\mathcal{M}_n(\mathbb{C})$ est trigonalisable sur $\mathbb{C}$.*

Démonstration. Il suffit de remarquer que tout polynôme de $\mathbb{C}[X]$ est scindé sur $\mathbb{C}$ d'après le théorème de d'Alembert–Gauss. $\square$

7.2 Le théorème de Cayley–Hamilton

Théorème 7.5. *Soit $A \in \mathcal{M}_n(K)$ une matrice carrée. Alors $\chi_A(A) = 0$.*

Démonstration. Comme $\mathcal{M}_n(\mathbb{Q}) \subset \mathcal{M}_n(\mathbb{R}) \subset \mathcal{M}_n(\mathbb{C})$, il suffit de prouver le théorème lorsque $K = \mathbb{C}$. On suppose donc que $K = \mathbb{C}$. La matrice A est alors trigonalisable, cela signifie qu'il existe une matrice inversibel $P \in \mathcal{M}_n(\mathbb{C})$ telle que $P^{-1}AP$ est triangulaire supérieure. Comme $\chi_A = \chi_{P^{-1}AP}$, il suffit de démontrer par récurrence sur $n \geqslant 1$ que $\chi_A(A) = 0$ pour $A \in \mathcal{M}_n(K)$ matrice triangulaire supérieure. Si $n = 1$, c'est évident car $A = aI_1$ et $\chi_A = X - a$. Supposons donc le résultat démontré au rang n et démontrons-le au rang $n+1$. Soit $A \in \mathcal{M}_{n+1}(K)$ triangulaire supérieure. On écrit A comme une matrice triangulaire par blocs

$$A = \begin{pmatrix} \lambda & B \\ 0_{n,1} & D \end{pmatrix}$$

avec $B \in \mathcal{M}_{1,n}(K)$ et $D \in \mathcal{M}_n(K)$ triangulaire supérieure. Par récurrence on a $\chi_D(D) = 0$. De plus $\chi_A = (X - \lambda)\chi_D$. On a alors

$$\chi_A(A) = (A - \lambda I_{n+1})\chi_D(A) = \begin{pmatrix} 0 & B \\ 0_{n,1} & D - \lambda I_n \end{pmatrix} \cdot \begin{pmatrix} \chi_D(\lambda) & B' \\ 0_{n,1} & 0_{n,n} \end{pmatrix} = \begin{pmatrix} 0 & 0_{1,n} \\ 0_{n,1} & 0_{n,n} \end{pmatrix}. \quad \square$$

Corollaire 7.6. *Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$. Alors $\chi_f(f) = 0_{\mathcal{L}(E)}$.*

7.3 Endomorphismes nilpotents

Soit E un K-espace vectoriel. Un endomorphisme f de E est dit *nilpotent* s'il existe un entier $n \geqslant 0$ tel que $f^n = 0_{\mathcal{L}(E)}$.

De même si $A \in \mathcal{M}_n(K)$ est une matrice carrée, on dit que A est *nilpotente* s'il existe un entier $n \geqslant 0$ tel que $A^n = 0$.

Remarque 7.7. Soit $\mathcal{B}$ une base de E. Il est clair que f est nilpotent si et seulement si $\mathrm{Mat}_{\mathcal{B}}(f)$ est nilpotente.

Théorème 7.8. *Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$ un endomorphisme de E. Les assertions suivantes sont équivalentes :*

(i) l'endomorphisme f est nilpotent ;

(ii) le polynôme minimal de f est de la forme X^n pour un entier $1 \leqslant n \leqslant \dim E$;

(iii) il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(f)$ est triangulaire supérieure avec des zéros sur la diagonale ;

(iv) le polynôme caractéristique de f est égal à $X^{\dim E}$.

Démonstration. Montrons que (i) implique (ii). Si f est nilpotent, il existe un entier m tel que $f^m = 0_{\mathcal{L}(E)}$. Cela signifie que f est annulé par X^m. Ainsi le polynôme minimal de f divise X^m. Il doit donc être de la forme X^k avec $k \leqslant m$.

L'implication (ii) implique (i) est évidente car le polynôme minimal est un polynôme annulateur de f.

Montrons que (i) implique (iii). Soit $n = \dim E$ et soit m l'indice de nilpotence de f. On note $i_r = \dim \operatorname{Ker} f^r$ pour $1 \leqslant r \leqslant m$. On construit alors une base de E de la façon suivante : on choisit une base $\mathcal{B} = (e_1, \ldots, e_{i_1})$ de $\operatorname{Ker} f$ que l'on complète en une base $(e_1, \ldots, e_{i_2})$ de $\operatorname{Ker} f^2$ etc. jusqu'à obtenir une base $(e_1, \ldots, e_n)$ de $E = \operatorname{Ker} f^m$ telle que, pour tout $1 \leqslant j \leqslant r$, $e_i \in \operatorname{Ker} f^j$ pour $i \leqslant i_j$. En particulier, si $i_{j-1} < i \leqslant i_j$, on a $e_i \in \operatorname{Ker}(f^j)$ et donc

$$f(e_i) \in \operatorname{Ker} f^{j-1} = \mathrm{Vect}(e_1, \ldots, e_{i_{j-1}}) \subset \mathrm{Vect}(e_1, \ldots, e_{i-1}).$$

Ainsi $\mathrm{Mat}_{\mathcal{B}}(f)$ est triangulaire supérieure avec des 0 sur la diagonale.

Montrons que (iii) implique (iv). Soit $\mathcal{B}$ une base de E telle que $\mathrm{Mat}_{\mathcal{B}}(f)$ est triangulaire supérieure avec des zéros sur la diagonale. Alors

$$\chi_f = \chi_{\mathrm{Mat}_{\mathcal{B}}(f)} = X^{\dim E}.$$

Enfin (iv) implique (i) en utilisant le théorème de Cayley–Hamilton. $\square$

Corollaire 7.9. *Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$. Supposons qu'il existe $\lambda \in K$ et $n \geqslant 1$ tels que $(f - \lambda \mathrm{Id}_E)^n = 0_{\mathcal{L}(E)}$. Alors f esr trigonalisable.*

Démonstration. En effet on déduit du théorème 7.16 que $f - \lambda \mathrm{Id}_E$ est nilpotent et donc trigonalisable. Ainsi f est trigonalisable. $\square$

7.4 Sous-espaces caractéristiques

Soit E un K-espace vectoriel et soit $f \in \mathcal{L}(E)$. Soit $\lambda \in K$ un scalaire. On appelle *sous-espace caractéristique* de f associé à λ le sous-espace vectoriel

$$F_\lambda(f) = \{v \in E \mid \exists n \geqslant 0,\ (f - \lambda \mathrm{Id}_E)^n(v) = 0_E\} = \bigcup_{n \geqslant 0} \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^n).$$

On remarque que l'on a $(f - \lambda \mathrm{Id}_E)^n = P(f)$ avec $P(X) = (X - \lambda)^n$, ainsi chaque sous-espace $\mathrm{Ker}((f - \lambda \mathrm{Id}_E)^n)$ est stable par f et $F_\lambda(f)$ est donc aussi un sous-espace stable par f.

Proposition 7.10. *Si E est de dimension finie, il existe $n \geqslant 0$ tel que $F_\lambda(f) = \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^n)$.*

Démonstration. En effet on a $\mathrm{Ker}((f - \lambda \mathrm{Id}_E)^n) \subset \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{n+1})$ pour tout $n \geqslant 0$. La suite de sous-espaces vectoriels $(\mathrm{Ker}((f - \lambda \mathrm{Id}_E)^n))_{n \geqslant 0}$ est donc croissante, de même que la suite d'entiers $(\dim \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^n))_{n \geqslant 0}$. Or cette dernière suite est bornée par $\dim E$. Elle est donc constante pour n assez grand, ce qui implique que la suite de sous-espaces $(\mathrm{Ker}((f - \lambda \mathrm{Id}_E)^n))_{n \geqslant 0}$ est constante pour n assez grand. Ceci implique le résultat. $\square$

Supposons désormais E est de dimension finie et notons q_λ le plus petit entier tel que $F_\lambda(f) = \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{q_\lambda})$.

Nous allons à présent donner une autre caractérisation de l'entier q_λ.

Lemme 7.11. *Soit E un espace vectoriel et soit f un endomorphisme de E. Soit $n \geqslant 0$ un entier. Supposons que $\mathrm{Ker}(f^n) = \mathrm{Ker}(f^{n+1})$. Alors $\mathrm{Ker}(f^m) = \mathrm{Ker}(f^n)$ pour tout $m \geqslant n$.*

Démonstration. En raisonnant par récurrence, on se ramène à montrer que $\mathrm{Ker}(f^{n+2}) = \mathrm{Ker}(f^{n+1})$. Soit $v \in \mathrm{Ker}(f^n n + 2)$ et posons $w = f(v)$. Alors $w \in \mathrm{Ker}(f^{n+1}) = \mathrm{Ker}(f^n)$. Ainsi $f^n(w) = 0_E$, ce qui implique $f^{n+1}(v) = 0_E$ et donc $v \in \mathrm{Ker}(f^{n+1})$. $\square$

Proposition 7.12. *L'entier q_λ est le plus petit entier tel que $\mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{q_\lambda}) = \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{q_\lambda + 1})$.*

Démonstration. On applique simplement le lemme 7.11 à $f - \lambda \mathrm{Id}_E$. $\square$

Proposition 7.13. *On a $E_\lambda(f) \neq \{0_E\}$ si et seulement si $F_\lambda(f) \neq \{0_E\}$.*

Démonstration. On a $E_\lambda(f) \subset F_\lambda(f)$ pour tout $\lambda \in K$ donc $E_\lambda(f) \neq \{0_E\}$ implique $F_\lambda(f) \neq 0_E$. Réciproquement supposons $E_\lambda(f) = \{0_E\}$. Alors $\{0_E\} = \mathrm{Ker}(f - \lambda \mathrm{Id}_E)^0 = \mathrm{Ker}(f - \lambda \mathrm{Id}_E)$ et donc $\mathrm{Ker}(f - \lambda \mathrm{Id}_E)^n = \{0_E\}$ pour tout $n \geqslant 0$, donc $F_\lambda(f) = \{0_E\}$. □

Ainsi les sous-espaces caractéristiques $F_\lambda(f)$ sont non nuls exactement pour $\lambda \in \mathrm{Sp}(f)$.

Proposition 7.14. *Les sous-espaces caractéristiques de f sont en somme directe.*

Démonstration. Pour tout $\lambda \in \mathrm{Sp}(f)$, il existe un entier q_λ tel que $F_\lambda(f) = \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{q_\lambda})$. Comme les polynômes $(X - \lambda)^{q_\lambda}$ sont premiers entre eux deux-à-deux, on déduit du lemme des noyaux (corollaire 6.22) que

$$\mathrm{Ker}(\prod_{\lambda \in \mathrm{Sp}(f)} (f - \lambda \mathrm{Id}_E)^{q_\lambda}) = \bigoplus_{\lambda \in \mathrm{Sp}(f)} \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{q_\lambda}).$$

Ainsi les sous-espaces caractéristiques sont en somme directe. □

7.5 Multiplicité d'une valeur propre

Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$. Si λ est une valeur propre de f, on appelle *multiplicité algébrique* de λ la multiplicité de λ comme racine du polynôme caractéristique de f. On la note $m_a(\lambda)$

On appelle *multiplicité géométrique* de la valeur propre λ l'entier $\dim E_\lambda(f)$. On la note $m_g(\lambda)$.

Proposition 7.15. *Soit λ une valeur propre de f. On a alors*

$$F_\lambda(f) = \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{m_a(\lambda)}).$$

Autrement dit $m_a(\lambda) \geqslant q_\lambda$.

Démonstration. Par définition de $F_\lambda(f)$, on a $\mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{m_q(\lambda)}) \subset F_\lambda(f)$. Prouvons l'inclusion réciproque. On peut écrire $\chi_f(X) = (X - \lambda)^{m_a(\lambda)} Q(X)$ avec $Q(\lambda) \neq 0$. Le corollaire 6.22 et le corollaire 7.6 nous donnent une décomposition $E = \mathrm{Ker}(\chi_f(f)) = \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{m_a(\lambda)}) \oplus \mathrm{Ker}(Q(f))$ de sorte que $\dim E = \dim \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{m_a(\lambda)}) + \dim \mathrm{Ker}(Q(f))$. Posons $V = F_\lambda(f) \cap \mathrm{Ker}(Q(f))$. Comme $F_\lambda(f)$ et $\mathrm{Ker}(Q(f))$ sont stables par f, il en est de même de V. Considérons $f_V \in \mathcal{L}(V)$. On a alors $Q(f_V) = 0_{\mathcal{L}(V)}$, ce qui implique (par le corollaire 6.27) que λ n'est pas valeur propre de f_V. Ainsi $f_V - \lambda \mathrm{Id}_V$ est un endomorphisme inversible de V. Comme par ailleurs il existe un entier $N \geqslant 1$ tel que $(f_{F_\lambda(f)} - \lambda \mathrm{Id}_{F_\lambda(f)})^N = 0_{\mathcal{L}(F_\lambda(f)}$, on a $(f_V - \lambda \mathrm{Id}_V)^N = 0_{\mathcal{L}(V)}$, de sorte que $f_V - \lambda \mathrm{Id}_V$ n'est pas injective. Ceci n'est possible que si $V = \{0_E\}$. Ainsi les sous-espaces $F_\lambda(f)$ et $\mathrm{Ker}(Q(f))$ sont en somme directe. On a donc

$$\dim F_\lambda(f) = \leqslant \dim(E) - \dim \mathrm{Ker}(Q(f)) = \dim(\mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{m_a(\lambda)})).$$

Comme $\mathrm{Ker}((f-\lambda\mathrm{Id}_E)^{m_a(\lambda)} \subset F_\lambda(f)$, on a $\dim \mathrm{Ker}((f-\lambda\mathrm{Id}_E)^{m_a(\lambda)} = \dim F_\lambda(f)$ et donc $\mathrm{Ker}((f-\lambda\mathrm{Id}_E)^{m_a(\lambda)} = F_\lambda(f)$. □

Théorème 7.16. *Soit λ une valeur propre de f. Alors la multiplicité algébrique de f est égale à la dimension du sous-espace caractéristique de λ. En particulier on a*

$$m_g(\lambda) \leqslant m_a(\lambda).$$

Démonstration. Soit $F = F_\lambda(f)$. Il s'agit d'un sous-espace stable par f, soit f_F l'endomorphisme de F induit par f. Par définition de $F_\lambda(f)$, l'endomorphisme f_F est annulé par $(X-\lambda)^{q_\lambda}$. On en déduit que $\pi_F(X)$ divise $(X-\lambda)^{q_\lambda}$. Ainsi λ est l'unique valeur propre de f_F et $f_F - \lambda\mathrm{Id}_F$ est nilpotent. On déduit du théorème 7.8 que $\chi_{f_F-\lambda\mathrm{Id}_F} = X^{\dim F}$ et donc $\chi_{f_F} = (X-\lambda)^{\dim F}$. Soit m la multiplicité algébrique de λ. On peut alors écrire $\chi_f(X) = (X-\lambda)^m Q(X)$ avec $Q(\lambda) \neq 0$. Ainsi les polynômes $(X-\lambda)^m$ et $Q(X)$ sont premiers entre eux. Le théorème de Cayley–Hamilton, le lemme des noyaux et la proposition 7.15 impliquent donc que

$$E = \mathrm{Ker}(\chi_f(f)) = \mathrm{Ker}((f-\lambda\mathrm{Id}_E)^m) \oplus \mathrm{Ker}(Q(f)) = F_\lambda(f) \oplus \mathrm{Ker}(Q(f)).$$

On a donc prouvé que $E = F \oplus S$ où $S = \mathrm{Ker}(Q(f))$ est un supplémentaire stable par f tel que λ n'est pas valeur propre de f_S. On a donc

$$\chi_f = \chi_{f_F}\chi_{f_S}.$$

Comme $\chi_{f_S}(\lambda) \neq 0$ et $\chi_{f_F} = (X-\lambda)^{\dim F}$, on en déduit que $m = \dim F_\lambda(f)$. □

Corollaire 7.17. *Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$. Les assertions suivantes sont équivalentes.*

(i) L'endomorphisme f est diagonalisable.

(ii) Le polynôme caractéristique χ_f est scindé et, pour tout racine λ de χ_f, la multiplicité géométrique de λ est égale à sa multiplicité algébrique.

(iii) Le polynôme caractéristique χ_f est scindé et, pour tout racine λ de χ_f, on a $E_\lambda(f) = F_\lambda(f)$.

Démonstration. Supposons que χ_f est scindé. Alors le corollaire 6.22 et le corollaire 7.6 impliquent que

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(f)} F_\lambda(f).$$

De plus, comme $E_\lambda(f) \subset F_\lambda(f)$, la multiplicité algébrique de λ est égal à sa multiplicité géométrique si et seulement si $E_\lambda(f) = F_\lambda(f)$. On en déduit le résultat. □

Théorème 7.18. *L'endomorphisme f est trigonalisable si et seulement si*

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(f)} F_\lambda(f).$$

Démonstration. Supposons dans un premier temps que f est trigonalisable. D'après le corollaire 7.3, le polynôme $\chi_f(X)$ est scindé. On peut donc écrire $\chi_f(X) = \prod_\lambda (X - \lambda)^{m_a(\lambda}$. Le lemme des noyaux et le théorème de Cayley–Hamilton impliquent alors que

$$E = \mathrm{Ker}(\chi_f(f)) = \bigoplus_{\lambda \in \mathrm{Sp}(f)} \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{m_a(\lambda)}) = \bigoplus_{\lambda \in \mathrm{Sp}(f)} F_\lambda(f)$$

en utilisant la proposition 7.15.

Réciproquement supposons que E est égal à la somme directe des sous-espaces $F_\lambda(f)$. La restriction de f à $F_\lambda(f)$ est annulée par $(X - \lambda)^{m_a(\lambda)}$ d'après la proposition 7.15. On déduit du corollaire 7.9 que $f_{F_\lambda(f)}$. Alors E est somme directe de sous-espaces stables par f sur lesquels f est trigonalisable. On en déduit que f est trigonalisable. □

Corollaire 7.19. *L'endomorphisme f est trigonalisable si et seulement si il est annulé par un polynôme scindé.*

Démonstration. Si f est trigonalisable, alors f est annulé par $\chi_f(X)$ qui est scindé d'après le corollaire 7.3. Réciproquement si f est annulé par un polynôme scindé $P(X) = \prod_\lambda (X - \lambda)^{m_\lambda}$, alors le lemme des noyaux implique que

$$E = \mathrm{Ker}(P(f)) = \bigoplus_\lambda \mathrm{Ker}((f - \lambda \mathrm{Id}_E)^{m_\lambda}).$$

Ainsi E est égal à la somme de ses sous-espaces caractéristiques et donc f est trigonalisable d'après le théorème 7.18. □

7.6 Décomposition de Dunford–Jordan d'un endomorphisme trigonalisable

Théorème 7.20. *Soit E un K-espace vectoriel de dimension finie et soit $f \in \mathcal{L}(E)$ un endomorphisme trigonalisable. On peut alors décomposer f, de façon unique, sous la forme $f = d + n$ où d est un endomorphisme diagonalisable et n un endomorphisme nilpotent tels que $d \circ n = n \circ d$.*

Démonstration. On commence par décomposer E sous la forme

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(f)} F_\lambda(f).$$

Chaque sous-espace caractéristique $F_\lambda(f)$ est alors stable par f. L'endomorphisme $n_\lambda = f_{F_\lambda(f)} - \lambda \mathrm{Id}_{F_\lambda(f)}$ est alors un endomorphisme nilpotent de $F_\lambda(f)$.

Soit $v \in E$. On peut écrire v, de façon unique, sous la forme $v = \sum_{\lambda \in \mathrm{Sp}(f)} v_\lambda$ avec $v_\lambda \in F_\lambda(f)$. On pose alors

$$d(v) = \sum_{\lambda \in \mathrm{Sp}(f)} \lambda v_\lambda.$$

On vérifie facilement que d est un endomorphisme de E. De plus d est diagonalisable car tout vecteur de $F_\lambda(f)$ est vecteur propre de d et E est égal la somme des sous-espaces $F_\lambda(f)$.

On pose alors $n = f - d$. Montrons que n est nilpotent. Chaque sous-espace $F_\lambda(f)$ est stable par f et par d, il est donc stable par n et on a $n_{F_\lambda(f)} = n_\lambda$. Comme tout vecteur de E est somme de vecteurs appartenant aux différents $F_\lambda(f)$, il suffit de vérifier qu'il existe un entier $m \geqslant 1$ tel que, pour tout $\lambda \in \mathrm{Sp}(f)$ et pour tout $v \in F_\lambda(f)$, on a $n^m(v) = 0_E$. Si $v \in F_\lambda(f)$, on a donc $n(v) = n_\lambda(v)$ et, puisque n_λ est nilpotent, il existe $m_\lambda \geqslant 1$ tel que $n_\lambda^{m_\lambda} = 0_{\mathcal{L}(F_\lambda(f))}$. Posons $m = \max\{m_\lambda \mid \lambda \in \mathrm{Sp}(f)\}$. On a alors, pour tout $\lambda \in Sp(f)$ et pour tout $v \in F_\lambda(f)$, $n^m(v) = 0_E$, ce qu'il fallait démontrer.

Vérifions enfin que $d \circ n = n \circ d$. Comme tout vecteur de E est somme de vecteurs appartenant aux différents $F_\lambda(f)$, il suffit de vérifier que, pour tout $\lambda \in \mathrm{Sp}(f)$ et pour tout $v \in F_\lambda(f)$, on a

$$d(n(v)) = n(d(n)).$$

Or $F_\lambda(f)$ est stable par n, donc $n(v) \in F_\lambda(f)$. Ainsi

$$d(n(v)) = \lambda(n(v)) = n(\lambda v) = n(d(v)).$$

On admet l'unicité de la décomposition. □

Corollaire 7.21. *Soit $A \in \mathcal{M}_n(\mathbb{C})$. On peut alors décomposer A, de façon unique, sous la forme $A = B + N$ où B est une matrice diagonalisable et N une matrice nilpotente telles que $BN = NB$.*

8 Exponentielles de matrices

Dans cette partie on suppose que $K = \mathbb{R}$ ou $= \mathbb{C}$.

8.1 Définition

Si $A \in \mathcal{M}_n(\mathbb{R})$, on admet que la série suivante converge dans $\mathcal{M}_n(\mathbb{R})$ (en un sens qui sera rendu précis dans une UE ultérieure) :

$$\sum_{n \geqslant 0} \frac{A^n}{n!}.$$

On note $\exp(A)$ sa limite. Il s'agit de *l'exponentielle* de la matrice A.

Proposition 8.1. *Si $A, B \in \mathcal{M}_n(K)$ sont deux matrices qui commutent, alors*

$$(A+B)^k = \sum_{i=0}^{k} \binom{k}{i} A^i B^{k-i}$$

pour tout entier $k \geqslant 0$.

Démonstration. La preuve se fait de façon immédiate par récurrence. □

Remarque 8.2. L'hypothèse de commutation des matrices est indispensable. Par exemple, si $k = 2$, on a

$$(A + B)^2 = A^2 + AB + BA + B^2$$

de sorte que la formule est valable si et seulement si $AB + BA = 2AB$, c'est-à-dire si et seulement si $AB = BA$.

Proposition 8.3. *Si $A, B \in \mathcal{M}_n(K)$ sont deux matrices qui commutent, alors*

$$\exp(A + B) = \exp(A)\exp(B).$$

Démonstration. Nous en donnons une preuve en admettant que toutes les inversions de séries sont bien légitimes (les outils permettant de le vérifier sont hors programme).

$$\begin{aligned}
\exp(A)\exp(B) &= \left(\sum_{m\geqslant 0} \frac{A^m}{m!}\right)\left(\sum_{n\geqslant 0} \frac{B^n}{n!}\right) \\
&= \sum_{k\geqslant 0} \frac{1}{k!}\left(\sum_{m+n=k} \frac{k!}{m!n!} A^m B^n\right) \\
&= \sum_{k\geqslant 0} \frac{1}{k!}\left(\sum_{m+n=k} \binom{k}{m} A^m B^n\right) \\
&= \sum_{k\geqslant 0} \frac{1}{k!}(A + B)^k = \exp(A + B).
\end{aligned}$$

□

Corollaire 8.4. *Pour toute matrice $A \in \mathcal{M}_n(K)$, la matrice* $\exp(A)$ *est une matrice inversible, d'inverse* $\exp(-A)$*.*

Proposition 8.5. *Si $A \in \mathcal{M}_n(K)$ et si $P \in \mathcal{M}_n(K)$ est inversible, alors*

$$\exp(PAP^{-1}) = P\exp(A)P^{-1}.$$

Démonstration.

$$\begin{aligned}
\exp(PAP^{-1}) &= \sum_{k\geqslant 0} \frac{1}{k!}(PAP^{-1})^k \\
&= \sum_{k\geqslant 0} P\frac{1}{k!}A^k P^{-1} \\
&= P\left(\sum_{k\geqslant 0} \frac{A^k}{k!}\right) P^{-1} \\
&= P\exp(A)P^{-1}
\end{aligned}$$

(une justification rigoureuse de la dernière étape est hors programme). □

Supposons à présent que la matrice $A \in \mathcal{M}_n(A)$ est trigonalisable. On peut donc l'écrire sous la forme $A = B + N$ avec B diagonalisable et N nilpotente telles que $BN = NB$. On a alors $\exp(A) = \exp(B)\exp(N)$.

La matrice $\exp(N)$ est simple à calculer. En effet, si m désigne l'indice de nilpotence de N, alors

$$\exp(N) = I_n + N + \cdots + \frac{N^{m-1}}{(m-1)!}.$$

On peut calculer $\exp(B)$ de la façon suivante. Comme B est diagonalisable, il existe une matrice P inversible telle que $D = P^{-1}BP$ est diagonale. On a alors $\exp(B) = P\exp(D)P^{-1}$ et le calcul de $\exp(D)$ se fait terme à terme :

$$\exp\left(\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}\right) = \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & e^{\lambda_n} \end{pmatrix}.$$

On peut aussi commencer par décomposer A sous la forme $A = PTP^{-1}$ avec T triangulaire supérieure et utiliser la formule

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

8.2 Application aux équations différentielles linéaires

On suppose toujours que $K = \mathbb{R}$ ou $K = \mathbb{C}$. On admet le théorème suivant.

Théorème 8.6. *Soit $A \in \mathcal{M}_n(K)$. Pour tout $X_0 \in K^n$, il existe un unique n-uplet $(x_1, \ldots, x_n)$ d'applications de classe $\mathcal{C}^1$ de $\mathbb{R}$ dans K tel que, pour tout $t \in \mathbb{R}$,*

$$\begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix} = A \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \qquad \begin{pmatrix} x_1(0) \\ \vdots \\ x_n(0) \end{pmatrix} = X_0.$$

De plus, on a, pour tout $t \in \mathbb{R}$,

$$\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \exp(tA)X_0.$$

Exemple 8.7. Supposons que $A \in \mathcal{M}_2(\mathbb{R})$ est trigonalisable et non diagonalisable avec une unique valeur propre λ. Il existe alors une matrice inversible $P \in \mathcal{M}_2(\mathbb{R})$ telle que

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

On a alors, pour tout $t \in \mathbb{R}$,

$$\exp(tB) = \exp\begin{pmatrix} t\lambda & 0 \\ 0 & t\lambda \end{pmatrix} \exp\begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & 0 \\ 0 & e^{\lambda t} \end{pmatrix} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}.$$

Soit X une application de $\mathbb{R}$ dans $\mathbb{R}^2$ telle que $X'(t) = AX(t)$ pour tout $t \in \mathbb{R}$ et $X(0) = X_0$. Posons $Y(t) = P^{-1}X(t)$ et $Y_0 = P^{-1}X_0$. Cette application vérifie l'équation différentielle $Y'(t) = BY(t)$ de sorte que

$$Y(t) = \exp(tB)Y_0 = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} e^{\lambda t}y_1 + te^{\lambda t}y_2 \\ e^{\lambda t}y_2 \end{pmatrix}.$$

9 Un petit complément sur la trace

Soit $A = (a_{i,j})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} \in \mathcal{M}_n(K)$. La *trace* de A est le scalaire

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

Pour A et B dans $\mathcal{M}_n(K)$, on a $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. En particulier si $P \in \mathrm{GL}_n(K)$, alors

$$\mathrm{Tr}(PAP^{-1}) = \mathrm{Tr}(A).$$

Théorème 9.1. *Soit $A \in \mathcal{M}_n(K)$ une matrice trigonalisable. On a alors*

$$\mathrm{Tr}(A) = \sum_{\lambda \in \mathrm{Sp}\, A} m_{\mathrm{alg}}(\lambda)\lambda, \quad \det(A) = \prod_{\lambda \in \mathrm{Sp}\, A} \lambda^{m_{\mathrm{alg}}(\lambda)}.$$