Compléments d'algèbre linéaire

Céline Esser

Université de Liège, Année académique 2021–2022

Table des matières

Chapitre 1

Réduction d'endomorphismes

1.1 Rappels sur la diagonalisation

Soit V un espace vectoriel de dimension finie m sur $\mathbb{C}$, et f un endomorphisme de V, c'est-à-dire une application linéaire de V dans lui-même. Le but de la diagonalisation est de trouver une base de V dans laquelle la matrice représentant f est la plus "simple" possible. Rappelons que si $\mathcal{B} = (b_1, \ldots, b_m)$ est une base de V, la matrice représentant f dans la base $\mathcal{B}$ est la matrice $[f] = [f]_{\mathcal{B}}$ dont la $j^{\text{ième}}$ colonne est formée des composantes de $f(b_j)$ dans $\mathcal{B}$. Dans ce cas, si $x_1, \ldots, x_m$ sont les composantes d'un vecteur x dans la base $\mathcal{B}$, alors le vecteur colonne des composantes de $f(x)$ dans $\mathcal{B}$ est donné par

$$[f]_{\mathcal{B}} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}.$$

Définition 1.1.1 On dit que f est *diagonalisable* s'il existe une base $\mathcal{B}$ de V telle que $[f]_{\mathcal{B}}$ soit diagonale.

Si f est diagonalisable, diagonaliser f revient donc à trouver une base $\mathcal{B}$ dans laquelle $[f]_{\mathcal{B}}$ est diagonale et calculer cette matrice diagonale, c'est-à-dire déterminer ses éléments diagonaux. On peut mener de front ces deux tâches.

Définition 1.1.2 Un nombre complexe λ est une *valeur propre* de f s'il existe $x \in V \setminus \{0\}$ tel que $f(x) = \lambda x$. L'ensemble des valeurs propres de f s'appelle le *spectre* de f. Si λ est une valeur propre de f, tout élément $x \in V$ vérifiant $f(x) = \lambda x$ est un *vecteur propre* associé à la valeur propre λ. Le *sous-espace propre* associé à la valeur propre λ est défini par

$$E_\lambda = E_\lambda(f) = \{x \in V : f(x) = \lambda x\} = \ker(f - \lambda \mathrm{id}),$$

c'est-à-dire par l'ensemble des vecteurs propres de f associé à la valeur propre λ.

Proposition 1.1.3 *Des espaces propres distincts de f sont toujours en position de somme directe. Autrement dit, des vecteurs propres non-nuls associés à des valeurs propres distinctes sont linéairement indépendants.*

Dans le cas d'un endomorphisme diagonalisable, les espaces propres fournissent une décomposition de l'espace.

Théorème 1.1.4 *L'endomorphisme f est diagonalisable si et seulement si V est la somme directe de ses sous-espaces propres, autrement dit si V possède une base de vecteurs propres.*

Les valeurs propres d'un endomorphisme peuvent être obtenues à partir du polynôme caractéristique de celui-ci.

Définition 1.1.5 Le *polynôme caractéristique* de f est défini par

$$p_f(\lambda) = \det(f - \lambda \mathrm{id}).$$

Il s'agit d'un polynôme de degré égal à la dimension m de l'espace V. Rappelons que le déterminant d'un endomorphisme est le déterminant de la matrice qui représente cet endomorphisme dans une base donnée. Cette définition est légitime car le déterminant ne dépend pas de la base choisie pour le calculer : en effet, si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de V et si S est la matrice de changement de base de $\mathcal{B}$ à $\mathcal{B}'$, alors $[f]_{\mathcal{B}'} = S^{-1}[f]_{\mathcal{B}}S$ et $\det[f]_{\mathcal{B}'} = \det[f]_{\mathcal{B}}$.

Le théorème suivant donne une propriété importante du polynôme caractéristique.

Théorème 1.1.6 (Caley-Hamilton) *Tout endomorphisme annule son polynôme caractéristique, c'est-à-dire*

$$p_f(f) = 0.$$

Rappelons d'autres résultats classiques concernant le polynôme caractéristique d'un endomorphisme.

Proposition 1.1.7 *Les valeurs propres de f sont exactement les racines de son polynôme caractéristique. De plus, la multiplicité de λ comme zéro du polynôme caractéristique p_f est toujours supérieure ou égale à la dimension du sous-espace propre associé E_λ.*

Cette proposition justifie l'introduction de deux types de multiplicité pour une valeur propre.

Définition 1.1.8 Soit λ une valeur propre de f.
- La *multiplicité algébrique* de λ est sa multiplicité comme zéro du polynôme caractéristique p_f.
- La *multiplicité géométrique* de λ est la dimension du sous-espace propre E_λ.

Le résultat suivant donne une caractérisation des endomorphismes diagonalisables à partir du polynôme caractéristique.

Théorème 1.1.9 *Un endomorphisme est diagonalisable si et seulement si les multiplicités algébrique et géométrique de chaque valeur propre de f coïncident.*

Ces résultats permettent de donner une méthode simple pour diagonaliser un endomorphisme f lorsque c'est possible.

1. Fixer une base $\mathcal{B} = (b_1, \ldots, b_m)$ de V sur $\mathbb{C}$.
2. Calculer le polynôme caractéristique p_f.
3. Factoriser $p_f = (-1)^m \prod_{j=1}^{p} (\lambda - \lambda_j)^{\mu_j}$.
4. Pour chaque racine λ_j de p_f, étudier l'espace propre E_{λ_j} en résolvant le système homogène associé, et calculer la dimension du sous-espace propre E_{λ_j} formée des solutions.
5. S'il existe j tel que $\dim E_{\lambda_j} < \mu_j$, f n'est pas diagonalisable. Sinon, continuer.

6. Pour tout j, choisir une base $(x_{j,1}, \dots, x_{j,\mu_j})$ de E_{λ_j}. Alors
$$\mathcal{B}' = (x_{1,1}, \dots, x_{1,\mu_1}, x_{2,1}, \dots, x_{2,\mu_2}, \dots, x_{p,1}, \dots, x_{p,\mu_p})$$
est une base de V dans laquelle f se représente par une matrice diagonale.
7. Si S est la matrice dont les colonnes sont les composantes des éléments de $\mathcal{B}'$ dans la base $\mathcal{B}$, alors
$$[f]_{\mathcal{B}'} = S^{-1}[f]_{\mathcal{B}}S = \operatorname{diag}\big(\underbrace{\lambda_1, \dots, \lambda_1}_{\mu_1}, \dots, \underbrace{\lambda_p, \dots, \lambda_p}_{\mu_p}\big).$$

Il ne faut pas croire que tous les endomorphismes sont diagonalisables : la matrice
$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$
n'est pas diagonalisable. On est donc amené à trouver des affaiblissements convenables de la diagonalisabilité.

Nous allons montrer que l'on peut obtenir une forme qui se rapproche nettement de la forme diagonale, la forme de Jordan. Il est d'ailleurs à noter que la forme de Jordan est la “meilleure” possible, en ce sens qu'une matrice est diagonalisable si et seulement si sa forme de Jordan est diagonale. La réduction à la forme de Jordan nécessite plusieurs étapes intermédiaires.

1.2 Forme de Jordan, cas nilpotent

Les endomorphismes nilpotents sont fondamentaux dans l'étude de la réduction des endomorphismes à la forme de Jordan : ils permettent d'appréhender le cas des endomorphismes non-diagonalisables.

Définition 1.2.1 Un endomorphisme g est *nilpotent* s'il existe un entier positif n tel que $g^n = 0$. Le plus petit entier n vérifiant cette égalité est appelé l'*indice de nilpotence* de g.

Rappelons que la notation g^n désigne la composition $\underbrace{g \circ \cdots \circ g}_{n \text{ fois}}$. La matrice qui représente g^n dans une base $\mathcal{B}$ est donnée par $[g^n]_{\mathcal{B}} = [g]_{\mathcal{B}}^n$.

Remarque 1.2.2 Soit g un endomorphisme nilpotent. Son indice de nilpotence est égal à n si et seulement si $\ker g^n = V$ et $\ker g^{n-1} \subsetneq V$.

Proposition 1.2.3 *Soit g un endomorphisme nilpotent.*

1. *Son indice de nilpotence est toujours inférieur ou égal à la dimension de V.*
2. *g n'est pas inversible.*
3. *Le spectre de g est réduit à $\{0\}$.*
4. *Si g est non-nul, alors g n'est pas diagonalisable.*
5. *Si k est strictement inférieur à l'indice de nilpotence de g, alors $\ker g^k \subsetneq \ker g^{k+1}$.*

Démonstration : 1. Notons n l'indice de nilpotence de g. Alors il existe $x \in V$ tel que $g^{n-1}(x) \neq 0$. Montrons que les vecteurs non-nuls

$$x, g(x), \ldots, g^{n-1}(x)$$

sont linéairement indépendents. On en déduira que $n \leq \dim V$. Supposons donc que

$$\sum_{j=0}^{n-1} \alpha_j g^j(x) = 0$$

où les $\alpha_j \in \mathbb{C}$ sont non tous nuls. Soit j_0 le plus petit entier tel que $\alpha_{j_0} \neq 0$. Alors

$$0 = g^{n-j_0-1}\left(\sum_{j=j_0}^{n-1} \alpha_j g^j(x)\right) = \sum_{j=j_0}^{n-1} \alpha_j \underbrace{g^{n+j-(j_0+1)}(x)}_{=0 \text{ si } j \geq j_0+1} = \alpha_{j_0} g^{n-1}(x),$$

ce qui est impossible car $\alpha_{j_0} \neq 0$ et $g^{n-1}(x) \neq 0$.

2. Si g était inversible, on aurait $g^{n-1} = g^{-1} \circ g^n = g^{-1} \circ 0 = 0$, ce qui contredit la définition de l'indice de nilpotence n.

3. Considérons à nouveau $x \in V$ tel que $g^{n-1}(x) \neq 0$. On a $g(g^{n-1}(x)) = g^n(x) = 0$ et $g^{n-1}(x)$ est donc un vecteur propre non-nul de g associé à 0. Donc 0 est bien une valeur propre de g. De plus, si λ est une valeur propre de g, alors il existe $y \in V \setminus \{0\}$ tel que $g(y) = \lambda y$. On a $0 = g^n(y) = \lambda^n y$, d'où $\lambda = 0$.

4. Si g était diagonalisable, puisque son spectre est réduit à 0, il existerait une base dans laquelle $[g]$ est la matrice diagonale nulle. Donc $g = 0$, d'où une contradiction.

5. On a bien sûr que $\ker g^k \subseteq \ker g^{k+1}$ pour tout $k \geq 0$. Supposons par l'absurde que $k < n$ est tel que $\ker g^k = \ker g^{k+1}$. Soit $x \in \ker g^{k+2}$. Alors $g^{k+1}(g(x)) = 0$ et donc $g(x) \in \ker g^{k+1} = \ker g^k$. On obtient alors $g(x) \in \ker g^k$ c'est-à-dire $x \in \ker g^{k+1}$. Ainsi $\ker g^k = \ker g^{k+1} = \ker g^{k+2}$ et en procédant de la sorte, on obtient $\ker g^k = \ker g^n = V$, ce qui contredit la définition de l'indice de nilpotence n. ■

Dans la preuve précédente, on a obtenu que si g est un endomorphisme nilpotent d'indice de nilpotence n et si $x \in V$ est tel que $g^{n-1}(x) \neq 0$, alors $x, g(x), \ldots, g^{n-1}(x)$ sont linéairement indépendents. Si $n = \dim V$, on obtient donc une base de V. La proposition suivante utilise la forme particulière de celle-ci.

Proposition 1.2.4 *Soit g un endomorphisme nilpotent d'indice de nilpotence $n = \dim V$. Alors il existe une base $\mathcal{B} = (b_1, \ldots, b_n)$ de V telle que*

$$g(b_1) = 0 \quad \text{et} \quad g(b_j) = b_{j-1},\ j \in \{2, \ldots, n\}.$$

En particulier,

$$[g]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}.$$

Démonstration : Si x est tel que $g^{n-1}(x) \neq 0$, il suffit de prendre la base

$$\mathcal{B} = (g^{n-1}(x), g^{n-2}(x), \dots, g(x), x).$$

■

Le résultat précédent nous amène à la définition suivante.

Définition 1.2.5 Soit g un endomorphisme nilpotent. On appelle *chaîne engendrée par g* toute suite finie de la forme

$$x, g(x), \dots, g^{l-1}(x)$$

où $g^{l-1}(x) \neq 0$ et $g^l(x) = 0$. On dit que x est la *tête* de la chaîne et que $g^{l-1}(x)$ en est la *queue*. L'entier l est la *longueur* de la chaîne.

Bien sûr, la longueur d'une chaîne est toujours inférieure ou égale à l'indice de nilpotence de g. On peut cependant trouver des chaînes de longueur plus petite (il suffit de prendre une tête dans $\ker g^l \setminus \ker g^{l-1}$).

Exemple 1.2.6 Considérons l'endomorphisme $g : \mathbb{C}^4 \to \mathbb{C}^4 : x \mapsto Ax$, où A est la matrice

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

On vérifie facilement que g est un endormorphisme nilpotent en remarquant que

$$A^2 = \begin{pmatrix} 0 & 0 & -2 & 11 \\ 0 & 0 & 0 & -8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 & -8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad A^4 = 0.$$

On a

$$e_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad Ae_4 = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 0 \end{pmatrix}, \quad A^2e_4 = \begin{pmatrix} 11 \\ -8 \\ 0 \\ 0 \end{pmatrix}, \quad A^3e_4 = \begin{pmatrix} -8 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{et} \quad A^4e_4 = 0.$$

Ainsi, e_4 est la tête d'une chaîne de longueur 4 engendrée par f. Ainsi, la chaîne engendrée par e_4 nous donne une base dans laquelle la matrice qui représente g a la forme

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

En prenant comme tête de chaîne Ae_4, on obtient une chaîne de longueur 3. On a également

$$e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad Ae_3 = \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \end{pmatrix}, \quad A^2e_4 = \begin{pmatrix} -2 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{et} \quad A^3e_3 = 0,$$

et e_3 est la tête d'une autre chaîne de longueur 3.

On a déjà obtenu le résultat suivant au sein de la preuve de la Proposition 1.2.3.

Proposition 1.2.7 *Les éléments d'une chaîne sont linéairement indépendants.*

On peut avoir un résultat plus général, qui est très utile en pratique.

Proposition 1.2.8 *Les éléments d'un ensemble C de chaînes engendrées par un endomorphisme nilpotent g sont linéairement indépendants si et seulement si les queues de ces chaînes sont linéairement indépendantes.*

Démonstration : La condition est évidemment nécessaire. Montrons qu'elle est suffisante. Soit L la longueur maximale des chaînes de C. Pour tout $l \in \{1, \ldots, L\}$, soit N_l le nombre de chaînes de longueur l et soient $x_{l,1}, \ldots, x_{l,N_l}$ les têtes de ces chaînes. On a donc

$$C = \left\{ g^{l-i}(x_{l,j}) : l \in \{1, \ldots, L\}, j \in \{1, \ldots, N_l\}, i \in \{1, \ldots, l\} \right\}.$$

Supposons que

$$\sum_{l=1}^{L} \sum_{j=1}^{N_l} \sum_{i=1}^{l} \alpha_{l,j,i} g^{l-i}(x_{l,j}) = 0$$

où les coefficients $\alpha_{l,j,i} \in \mathbb{C}$ sont non tous nuls. Soit i_0 le plus grand i pour lequel il existe $l \in \{1, \ldots, L\}$ et $j \in \{1, \ldots, N_l\}$ tel que $\alpha_{l,j,i} \neq 0$. Alors

$$0 = g^{i_0-1} \left(\sum_{l=1}^{L} \sum_{j=1}^{N_l} \sum_{i=1}^{i_0} \alpha_{l,j,i} g^{l-i}(x_{l,j}) \right) = \sum_{l=1}^{L} \sum_{j=1}^{N_l} \sum_{i=1}^{i_0} \alpha_{l,j,i} \underbrace{g^{i_0-1+l-i}(x_{l,j})}_{=0 \text{ si } i \leq i_0-1}.$$

Ainsi, on obtient

$$\sum_{l=1}^{L} \sum_{j=1}^{N_l} \alpha_{l,j,i_0} g^{l-1}(x_{l,j}) = 0.$$

Or, les queues de chaînes sont linéairement indépendantes, ce qui implique que $\alpha_{l,j,i_0} = 0$ pour tout $l \in \{1, \ldots, L\}$ et tout $j \in \{1, \ldots, N_l\}$. Cela contredit la définition de i_0. ∎

Par conséquent, on dit que des chaînes sont *linéairement indépendantes* si les queues de ces chaînes le sont. La réduction à la forme de Jordan d'un endomorphisme nilpotent consiste à trouver une *base répartie en chaînes*, c'est-à-dire une base formée par les éléments de chaînes linéairement indépendantes.

Théorème 1.2.9 (Forme de Jordan des endomorphismes nilpotents) *Si g est un endomorphisme nilpotent, il existe une base dans laquelle $[g]$ a la forme*

$$[g] = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & J_s \end{pmatrix}$$

où chaque bloc diagonal J_i est une matrice de la forme

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

Démonstration : Notons n l'indice de nilpotence de g. Construisons la suite croissante $V_0, \dots, V_n$ de sous-espaces de V de la manière suivante : on pose

$$V_0 = \ker g^0 = \{0\} \quad \text{et} \quad V_i = \ker g^i \,,\, i \in \{1, \dots, n\}.$$

En particulier, $V_n = V$. Par la Proposition 1.2.3, on sait que la suite $V_0, \dots, V_n$ est strictement croissante. De plus, remarquons que pour tout $i \in \{1, \dots, n\}$, on a

$$V_i = \{x \in V : g^i(x) = 0\} = \{x \in V : g^{i-1}(g(x)) = 0\} = \{x \in V : g(x) \in V_{i-1}\} = g^{-1}(V_{i-1})\,.$$

Construisons à présent de proche en proche la suite $W_0, \dots, W_n$ de sous-espaces de V comme suit : on pose $W_n = \{0\}$. Pour construire W_{n-1}, il suffit de choisir un supplémentaire de V_{n-1} dans V_n. Ensuite, si $W_n, W_{n-1}, \dots, W_{i+1}$ ont été construits, on choisit W_i afin d'avoir

$$W_i \oplus V_i = V_{i+1} \quad \text{et} \quad g(W_{i+1}) \subseteq W_i.$$

Observons qu'une telle construction est possible : pour cela, il faut montrer que $g(W_{i+1})$ est un sous-espace de V_{i+1} d'intersection nulle avec V_i. Par induction, on sait que W_{i+1} est un supplémentaire de V_{i+1} dans V_{i+2}. Dès lors,

$$g(W_{i+1}) \subseteq g(V_{i+2}) = g(g^{-1}(V_{i+1})) \subseteq V_{i+1}.$$

De plus, si $x \in V_i \cap g(W_{i+1})$, alors $x = g(y)$ avec $y \in W_{i+1}$. Or $x \in V_i$, et donc on a $g^i(x) = 0$. Par conséquent, $g^{i+1}(y) = g^i(x) = 0$, d'où $y \in V_{i+1} \cap W_{i+1}$. Comme V_{i+1} et W_{i+1} sont en somme directe, on obtient $y = 0$, d'où $x = g(y) = 0$. Ainsi, $V_i \cap g(W_{i+1}) = \{0\}$.

Au vu de la construction des suites $V_0, \ldots, V_n$ et $W_0, \ldots, W_n$, on a

$$V = V_n = W_{n-1} \oplus V_{n-1} = W_{n-1} \oplus W_{n-2} \oplus V_{n-2} = \cdots = W_{n-1} \oplus W_{n-2} \oplus \cdots \oplus W_0.$$

Par conséquent, afin d'obtenir une base de V, il suffit de construire une base de W_i pour tout $i \in \{0, \ldots, n-1\}$. Commençons par remarquer que pour tout $i \in \{1, \ldots, n\}$, g est injectif sur W_i car

$$W_i \cap \ker g = W_i \cap V_1 \subseteq W_i \cap V_i = \{0\}.$$

Par conséquent, si $(y_1, \ldots, y_r)$ est base de W_i, les éléments $g(y_1), \ldots, g(y_r)$ sont linéairement indépendants (en effet, si $\sum_{j=1}^r \alpha_j g(y_j) = 0$, alors $g\left(\sum_{j=1}^r \alpha_j y_j\right) = 0$; l'injectivité de g donne alors $\sum_{j=1}^r \alpha_j y_j = 0$, d'où $\alpha_1 = \cdots = \alpha_r = 0$). Comme $g(W_i) \subseteq W_{i-1}$, ils peuvent être complétés en une base de W_{i-1}. Ceci permet de construire une base de V de la façon suivante.

1. On prend une base $(x_1, \ldots, x_{r_1})$ de W_{n-1}.
2. On construit une base de W_{n-2} en complétant $g(x_1), \ldots, g(x_{r_1})$ par $x_{r_1+1}, \ldots, x_{r_2}$.
3. On itère le processus et on construit une base de W_{n-i} en complétant les vecteurs

$$g^{i-1}(x_1), \ldots, g^{i-1}(x_{r_1}), g^{i-2}(x_{r_1+1}), \ldots, g^{i-2}(x_{r_2}), \ldots, g(x_{r_{i-2}+1}), \ldots, g(x_{r_{i-1}})$$

par $x_{r_{i-1}+1}, \ldots, x_{r_i}$.

Au total, on obtient une base de V :

$$
\begin{array}{ccccccc}
W_0 & \dots & W_{n-i} & \dots & W_{n-3} & W_{n-2} & W_{n-1} \\
\hline
g^{n-1}(x_1) & \dots & g^{i-1}(x_1) & \dots & g^2(x_1) & g(x_1) & x_1 \\
g^{n-1}(x_2) & \dots & g^{i-1}(x_2) & \dots & g^2(x_2) & g(x_2) & x_2 \\
\vdots & & \vdots & & \vdots & \vdots & \vdots \\
g^{n-1}(x_{r_1}) & \dots & g^{i-1}(x_{r_1}) & \dots & g^2(x_{r_1}) & g(x_{r_1}) & x_{r_1} \\
g^{n-2}(x_{r_1+1}) & \dots & g^{i-2}(x_{r_1+1}) & \dots & g(x_{r_1+1}) & x_{r_1+1} & \\
\vdots & & \vdots & & \vdots & \vdots & \\
g^{n-2}(x_{r_2}) & \dots & g^{i-2}(x_{r_2}) & \dots & g(x_{r_2}) & x_{r_2} & \\
g^{n-3}(x_{r_2+1}) & \dots & g^{i-3}(x_{r_2+1}) & \dots & x_{r_2+1} & & \\
\vdots & & \vdots & & \vdots & & \\
g^{n-3}(x_{r_3}) & \dots & g^{i-3}(x_{r_3}) & \dots & x_{r_3} & & \\
\vdots & & \vdots & & & & \\
\vdots & & \vdots & & & & \\
g^{n-i}(x_{r_{i-1}+1}) & \dots & x_{r_{i-1}+1} & & & & \\
\vdots & & \vdots & & & & \\
g^{n-i}(x_{r_i}) & \dots & x_{r_i} & & & & \\
\vdots & & & & & & \\
\vdots & & & & & & \\
x_{r_{n-1}+1} & & & & & & \\
\vdots & & & & & & \\
x_{r_n} & & & & & &
\end{array}
\tag{1.1}
$$

Les éléments d'une même ligne $g^l(x_j), \dots, x_j$ sont les éléments de la chaîne engendrée par g de tête x_j. Par construction, $g^l(x_j) \in W_0 \subseteq V_1$, et donc $g(g^l(x_j)) = 0$. Ainsi, si b est un élément de la base ci-dessus, $g(b)$ est l'élément suivant de la même chaîne si b n'est pas queue de chaîne, et $g(b) = 0$ sinon. Dans cette base répartie en chaînes, la matrice $[g]$ a donc la forme souhaitée.

■

La base dont il est question dans le théorème précédent n'est pas unique. Cependant, ces bases ont un certain nombre de propriétés communes.

Supposons que $\mathcal{B}$ est une base de V formée par les éléments de chaînes linéairement indépendantes engendrées par g. On forme un tableau en ordonnant les chaînes par longueur décroissante ; les lignes du tableau sont formées des chaînes écrites en commençant par la queue et on convient d'aligner verticalement les queues. Par exemple, si $\mathcal{B}$ est formée de deux chaînes de longueur 5, d'une chaîne de longueur 4, d'une chaîne de longueur 3 et d'une chaîne de longueur 1, on obtient le tableau

$$\begin{array}{lllll} g^4(x_1) & g^3(x_1) & g^2(x_1) & g(x_1) & x_1 \\ g^4(x_2) & g^3(x_2) & g^2(x_2) & g(x_2) & x_2 \\ g^3(x_3) & g^2(x_3) & g(x_3) & x_3 & \\ g^2(x_4) & g(x_4) & x_4 & & \\ x_5 & & & & \end{array}$$

ce que l'on représente symboliquement par

$$\begin{array}{|c|c|c|c|c|} \hline * & * & * & * & * \\ \hline * & * & * & * & * \\ \hline * & * & * & * \\ \cline{1-4} * & * & * \\ \cline{1-3} * \\ \cline{1-1} \end{array}$$

La forme du tableau obtenu ne dépend pas de la base répartie en chaînes choisie et on peut la déterminer en étudiant les rangs des puissances $g, g^2, \ldots, g^n$ de l'endomorphisme g, comme le montre le résultat suivant.

Proposition 1.2.10 *Considérons une base de V formée de chaînes linéairement indépendantes et engendrée par un endomorphisme nilpotent g. On place les éléments de la chaîne dans un tableau.*

1. *Les k premières colonnes du tableau forment une base de $\ker g^k$.*
2. *La longueur de la plus longue chaîne est l'indice de nilpotence de g.*
3. *Le nombre d'éléments dans la $k^{\text{ème}}$ colonne vaut $\mathrm{rg}(g^{k-1}) - \mathrm{rg}(g^k)$.*

Démonstration : 1. Puisque tous les éléments du tableau sont linéairement indépendants, il suffit de montrer que tout vecteur x de $\ker g^k$ est combinaison linéaire des vecteurs des k premières colonnes. Comme on a une base de V, on peut décomposer x sous la forme

$$x = \sum_{r=1}^{u} \alpha_r y_r + \sum_{s=1}^{v} \beta_s z_s$$

où $y_1, \ldots, y_u$ sont les éléments des k premières colonnes du tableau et $z_1, \ldots, z_v$ les éléments des autres colonnes. On a

$$0 = g^k(x) = \sum_{r=1}^{u} \alpha_r \underbrace{g^k(y_r)}_{=0} + \sum_{s=1}^{v} \beta_s g^k(z_s).$$

Par construction du tableau, les éléments $g^k(z_1), \ldots, g^k(z_v)$ sont des éléments distincts du tableau, et sont donc linéairement indépendants. Par conséquent, on obtient que $\beta_1 = \cdots = \beta_v = 0$, d'où la conclusion.

2. Vu le point 1., si le tableau est formé de n colonnes, alors le tableau forme une base de $\ker g^n$. On en tire que $V = \ker g^n$, et donc l'indice de nilpotence de g est plus petit ou égal à n. Puisque $\ker g^{n-1} \subsetneq \ker g^n$ (ils n'ont pas la même dimension vu le point 1.), on obtient la conclusion.

3. Par le point 1., on sait que le nombre d'éléments dans les k premières colonnes vaut $\dim\big(\ker g^k\big)$. Ainsi, si $k \geq 2$, le nombre d'éléments dans la $k^{\text{ème}}$ colonne vaut

$$\begin{aligned} \dim\big(\ker g^k\big) - \dim\big(\ker g^{k-1}\big) &= \dim V - \mathrm{rg}(g^k) - \dim V + \mathrm{rg}(g^{k-1}) \\ &= \mathrm{rg}(g^{k-1}) - \mathrm{rg}(g^k). \end{aligned}$$

Si $k = 1$, le nombre d'élements dans la première colonne est donné par

$$\dim \ker g = \dim V - \operatorname{rg}(g) = \operatorname{rg}(g^0) - \operatorname{rg}(g).$$

■

Dans la pratique, on peut donc trouver la forme du tableau de base répartie en chaînes grâce à la proposition précédente. On utilise la Proposition 1.2.8 pour trouver des chaînes linéairement indépendantes et former la base recherchée.

Remarque 1.2.11 Par conséquent, la forme de Jordan de g ne dépend pas de la base répartie en chaînes choisie, à une permutation des blocs près. En particulier, l'indice de nilpotence de g est la dimension du plus gros bloc. En regardant le tableau (1.1), on voit que le nombre de chaînes de longueur k correspond au nombre d'éléments nécessaires pour compléter une base de W_k pour obtenir une base de W_{k-1}, c'est-à-dire

$$\begin{aligned} \dim W_{k-1} - \dim W_k &= \dim\left(\ker g^k\right) - \dim\left(\ker g^{k-1}\right) - \dim\left(\ker g^{k+1}\right) + \dim\left(\ker g^k\right) \\ &= 2\dim\left(\ker g^k\right) - \dim\left(\ker g^{k-1}\right) - \dim\left(\ker g^{k+1}\right). \end{aligned}$$

Cela donne le nombre de blocs de dimension k dans la forme de Jordan de g.

1.3 Exemple

Considérons l'opérateur $f : \mathbb{C}^5 \to \mathbb{C}^5 : x \mapsto Ax$, où A est la matrice

$$A = \begin{pmatrix} -1 & 0 & -1 & 1 & 0 \\ -4 & -1 & -3 & 2 & 1 \\ -2 & -1 & -2 & 1 & 1 \\ -3 & -1 & -3 & 2 & 1 \\ -8 & -2 & -7 & 5 & 2 \end{pmatrix}$$

Il s'agit d'un endomorphisme nilpotent d'indice 3 car

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

Cherchons la forme du tableau de base répartie en chaînes pour A. En calculant les rangs des matrices A et A^2, on trouve

- nombre d'éléments dans la colonne 1 = $5 - \operatorname{rg}(A) = 5 - 3 = 2$
- nombre d'éléments dans la colonne 2 = $\operatorname{rg}(A) - \operatorname{rg}(A^2) = 3 - 1 = 2$
- nombre d'éléments dans la colonne 3 = $\operatorname{rg}(A^2) - \operatorname{rg}(A^3) = 1 - 0 = 1$.

Le tableau a donc la forme

*	*	*
*	*	

Il faut donc construire deux chaînes linéairement indépendantes de longueurs respectives 2 et 3. La tête d'une chaîne de longueur 3 est un vecteur x tel que $A^2x \neq 0$. On peut prendre

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

La chaîne est donnée par

$$e_1\,,\ Ae_1 = \begin{pmatrix} -1 \\ -4 \\ -2 \\ -3 \\ -8 \end{pmatrix},\ A^2e_1 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ -1 \\ -1 \end{pmatrix}.$$

On cherche à présent une chaîne de longueur 2, c'est-à-dire un vecteur y tel que $A^2y = 0$ et $Ay \neq 0$. Il faut de plus s'assurer que les queues des chaînes A^2e_1 et Ay sont linéairement indépendantes pour que les chaînes le soient. On peut prendre

$$y = e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

et la chaîne

$$e_2\,,\ Ae_2 = \begin{pmatrix} 0 \\ -1 \\ -1 \\ -1 \\ -2 \end{pmatrix}.$$

La matrice S de changement de base est donnée par

$$S = \begin{pmatrix} 0 & -1 & 1 & 0 & 0 \\ 0 & -4 & 0 & -1 & 1 \\ -1 & -2 & 0 & -1 & 0 \\ -1 & -3 & 0 & -1 & 0 \\ -1 & -8 & 0 & -2 & 0 \end{pmatrix}$$

et on sait que

$$S^{-1}AS = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

1.4 Sous-espaces caractéristiques et forme de Jordan, cas général

Nous allons à présent traîter la réduction à la forme de Jordan d'un endomorphisme quelconque f. Pour cela, pour chaque valeur propre de f, nous commençons par introduire des sous-espaces vectoriels stables pour f, et nous montrons qu'ils donnent une décomposition en somme directe de l'espace V. L'idée est ensuite de se ramener au cas nilpotent sur chacun de ces sous-espaces.

Définition 1.4.1 Soit f un endomorphisme de V. Un sous-espace vectoriel W de V est *stable pour* f si $f(W) \subseteq W$.

Exemple 1.4.2 Remarquons que les sous-espaces propres de f sont stables pour f. En effet, si $f(x) = \lambda x$, alors $f(f(x)) = f(\lambda x) = \lambda f(x)$, d'où $f(x) \in E_\lambda$.

L'intérêt des sous-espaces stables est le suivant : si $V = W_1 \oplus \cdots \oplus W_l$ où les sous-espaces $W_1, \ldots, W_l$ sont stables pour f, alors dans toute base de V obtenue en juxtaposant des bases de $W_1, \ldots, W_l$, la représentation matricielle de f est de la forme

$$[f] = \operatorname{diag}(B_1, \ldots, B_l),$$

où chaque B_i est une matrice de taille $\dim W_i$.

Nous savons que dans le cas où l'endomorphisme f n'est pas diagonalisable, l'espace V n'est pas égal à la somme directe de ses sous-espaces propres. Nous allons à présent introduire d'autres sous-espaces propres pour f qui donneront une décomposition de l'espace en une somme directe.

Définition 1.4.3 Soit f un endomorphisme de V. Si λ est une valeur propre de f de multiplicité algébrique μ, on définit le *sous-espace caractéristique* de f relatif à λ par

$$F_\lambda(f) = F_\lambda = \ker(f - \lambda \mathrm{id})^\mu \,.$$

Attention à ne pas confondre les sous-espaces propres $E_\lambda = \ker(f - \lambda\mathrm{id})$ et les sous-espaces caractéristiques $F_\lambda = \ker(f - \lambda\mathrm{id})^\mu$. Notons cependant que $E_\lambda \subseteq F_\lambda$.

Lemme 1.4.4 *Soient f un endomorphisme de V et λ une valeur propre de f. Le sous-espace caractéristique F_λ est stable pour f.*

Démonstration : Soit $x \in F_\lambda$. Si μ est la multiplicité algébrique de λ, alors

$$(f - \lambda\mathrm{id})^\mu(f(x)) = f\big((f - \lambda\mathrm{id})^\mu(x)\big) = f(0) = 0$$

puisque deux polynômes d'un endomorphisme commutent. Par conséquent, $f(x) \in F_\lambda$. ■

La clé de la réduction à la forme canonique de Jordan est contenue dans le résultat suivant.

Proposition 1.4.5 *Soit f un endomorphisme de V dont le polynôme caractéristique est donné par*

$$p_f(\lambda) = (-1)^m \prod_{j=1}^{p} (\lambda - \lambda_j)^{\mu_j} \,,$$

où $\lambda_1, \ldots, \lambda_p \in \mathbb{C}$ sont distincts. Alors, pour tout $i \in \{1, \ldots, p\}$, on a

$$\dim F_{\lambda_i} = \mu_i.$$

De plus,

$$V = F_{\lambda_1} \oplus \cdots \oplus F_{\lambda_p}.$$

Démonstration : Pour tout $i \in \{1, \ldots, p\}$, posons $f_i = (f - \lambda_i \mathrm{id})|_{F_{\lambda_i}}$. Par le Lemme 1.4.4, on sait que f_i est un endomorphisme de F_{λ_i}. De plus, on a

$$f_i^{\mu_i} = (f - \lambda_i \mathrm{id})^{\mu_i}|_{F_{\lambda_i}} = 0$$

par définition du sous-espace caractéristique, et donc f_i est nilpotent sur F_{λ_i}. Le Théorème 1.2.9 nous fournit alors une base $\mathcal{B}_i$ de F_{λ_i} dans laquelle $[f_i]_{\mathcal{B}_i}$ est réduite à la forme de Jordan, i.e.

$$[f_i]_{\mathcal{B}_i} = \begin{pmatrix} J_{i,1} & 0 & \ldots & 0 \\ 0 & J_{i,2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & \ldots & J_{i,s} \end{pmatrix}$$

où chaque bloc diagonal $J_{i,r}$ est une matrice de la forme

$$J_{i,r} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \ldots & \ldots & \ldots & 0 \end{pmatrix}.$$

Puisque $f|_{F_{\lambda_i}} = f_i + \lambda_i \mathrm{id}_{F_{\lambda_i}}$, on en tire que

$$[f|_{F_{\lambda_i}}]_{\mathcal{B}_i} = \begin{pmatrix} \widetilde{J}_{i,1} & 0 & \ldots & 0 \\ 0 & \widetilde{J}_{i,2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & \ldots & \widetilde{J}_{i,s} \end{pmatrix}$$

où chaque bloc diagonal $\widetilde{J}_{i,r}$ est de la forme

$$\widetilde{J}_{i,r} = \begin{pmatrix} \lambda_i & 1 & 0 & \ldots & 0 \\ \vdots & \lambda_i & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \ldots & \ldots & \ldots & \lambda_i \end{pmatrix}.$$

Considérons une base $\mathcal{B}$ de V obtenue en complétant la base $\mathcal{B}_i$ de F_{λ_i}. Dans cette base, f est représenté par une matrice de la forme

$$[f]_{\mathcal{B}} = \left(\begin{array}{c|c} [f|_{F_{\lambda_i}}]_{\mathcal{B}_i} & C \\ \hline 0 & D \end{array} \right),$$

et par conséquent,

$$\begin{aligned} p_f(\lambda) &= \det\left([f|_{F_{\lambda_i}}]_{\mathcal{B}_i} - \lambda I_{\dim F_{\lambda_i}}\right) \det\left(D - \lambda I_{m-\dim F_{\lambda_i}}\right) \\ &= (\lambda_i - \lambda)^{\dim F_{\lambda_i}} \det\left(D - \lambda I_{m-\dim F_{\lambda_i}}\right). \end{aligned}$$

En particulier, le polynôme $(\lambda_i - \lambda)^{\dim F_{\lambda_i}}$ divise p_f, d'où

$$\dim F_{\lambda_i} \leq \mu_i.$$

Montrons à présent que les sous-espaces caractéristiques $F_{\lambda_1}, \ldots, F_{\lambda_p}$ engendrent V (c'est-à-dire $V = F_{\lambda_1} + \cdots + F_{\lambda_p}$). Pour cela, posons

$$p_i(\lambda) = \prod_{j=1, j\neq i}^{p} (\lambda_j - \lambda)^{\mu_j}.$$

En décomposant $\frac{1}{p_f}$ en fractions rationnelles propres, on obtient

$$\frac{1}{p_f(\lambda)} = \sum_{j=1}^{p} \frac{A_j(\lambda)}{(\lambda_j - \lambda)^{\mu_j}}$$

où chaque A_j est un polynôme de degré strictement inférieur à μ_j. Il s'ensuit que

$$1 = \sum_{j=1}^{p} A_j(\lambda) p_j(\lambda)$$

et en appliquant ces polynômes à l'endomorphisme f, il vient

$$\mathrm{id} = \sum_{j=1}^{p} A_j(f) p_j(f).$$

Ainsi, pour tout $x \in V$, on a

$$x = \sum_{j=1}^{p} \left(A_j(f) p_j(f)\right)(x) = \sum_{j=1}^{p} x_j$$

où pour tout $j \in \{1, \ldots, p\}$, on a posé $x_j = \left(A_j(f) p_j(f)\right)(x)$. Il suffit alors de montrer que $x_j \in F_{\lambda_j}$. C'est immédiat car

$$\begin{aligned} (f - \lambda_j \mathrm{id})^{\mu_j}(x_j) &= (f - \lambda_j \mathrm{id})^{\mu_j}\left(A_j(f) p_j(f)\right)(x) \\ &= (f - \lambda_j \mathrm{id})^{\mu_j} p_j(f) A_j(f)(x) \\ &= \underbrace{p_f(f)}_{=0} A_j(f)(x) = 0 \end{aligned}$$

en utilisant le Théorème 1.1.6 de Caley-Hamilton et le fait que des polynômes d'un endomorphisme commutent. Ainsi, on a bien que $V = F_{\lambda_1} + \cdots + F_{\lambda_p}$ et donc

$$\dim V \leq \sum_{j=1}^{p} \dim F_{\lambda_j} \leq \sum_{j=1}^{p} \mu_j = \deg(p_f) = \dim V.$$

On en tire que

$$V = F_{\lambda_1} \oplus \cdots \oplus F_{\lambda_p},$$

et $\dim F_{\lambda_i} = \mu_i$. ∎

La preuve précédente nous fournit directement tous les outils et la méthode permettant d'obtenir la forme de Jordan d'un endomorphisme dans le cas général. Commençons par donner une définition.

Définition 1.4.6 On appelle *bloc de Jordan* une matrice carrée du type

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

Théorème 1.4.7 (Réduction à la forme de Jordan) *Soit f un endomorphisme de V. Il existe une base de V dans laquelle $[f]$ a la forme de Jordan, c'est-à-dire la forme*

$$[f] = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & J_s \end{pmatrix}$$

où chaque J_i est un bloc de Jordan.

Démonstration : Soient $\lambda_1, \dots, \lambda_p$ les valeurs propres distinctes de f. Par la Proposition 1.4.5, on sait que

$$V = F_{\lambda_1} \oplus \cdots \oplus F_{\lambda_p}.$$

De plus, en imitant le début de la preuve de la Proposition 1.4.5, on sait que pour tout $i \in \{1, \dots, p\}$, il existe une base $\mathcal{B}_i$ de F_{λ_i} dans laquelle

$$[f|_{F_{\lambda_i}}]_{\mathcal{B}_i} = \begin{pmatrix} \widetilde{J}_{i,1} & 0 & \dots & 0 \\ 0 & \widetilde{J}_{i,2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & \widetilde{J}_{i,s} \end{pmatrix}$$

où chaque bloc diagonal $\widetilde{J}_{i,r}$ est un bloc de Jordan associé à la valeur propre λ_i. Soit $\mathcal{B}$ la base de V formée des bases $\mathcal{B}_1, \dots, \mathcal{B}_p$. Dans cette base,

$$[f]_{\mathcal{B}} = \begin{pmatrix} [f|_{F_{\lambda_1}}]_{\mathcal{B}_1} & 0 & \dots & 0 \\ 0 & [f|_{F_{\lambda_2}}]_{\mathcal{B}_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & [f|_{F_{\lambda_p}}]_{\mathcal{B}_p} \end{pmatrix},$$

d'où la conclusion. ■

Remarque 1.4.8 Deux endomorphismes f et $\widetilde{f}$ ont la même forme de Jordan (à une permutation des blocs près) s'ils ont les mêmes valeurs propres et si les formes des tableaux des bases réparties en chaînes des espaces caractéristiques F_{λ_i} engendrées par les restrictions $f - \lambda_i \mathrm{id}|_{F_{\lambda_i}}$ et $\widetilde{f} - \lambda_i \mathrm{id}|_{F_{\lambda_i}}$ coïncident.

Le résultat suivant montre que la forme de Jordan est la meilleure possible.

Proposition 1.4.9 *Soit f un endomorphisme de valeurs propres distinctes $\lambda_1, \ldots, \lambda_p$. Les assertions suivantes sont équivalentes.*

1. *f est diagonalisable.*
2. *$E_{\lambda_j} = F_{\lambda_j}$ pour tout $j \in \{1, \ldots, p\}$.*
3. *La forme de Jordan de f est une matrice diagonale (c'est-à-dire tous les blocs de Jordan sont de dimension 1).*

Démonstration : Montrons que la première assertion implique la deuxième. Notons $\mu_1, \ldots, \mu_p$ les multiplicités algébriques de $\lambda_1, \ldots, \lambda_p$ respectivement. Si f est diagonalisable, on sait que $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$. On a aussi toujours $V = F_{\lambda_1} \oplus \cdots \oplus F_{\lambda_p}$ et $E_{\lambda_i} \subseteq F_{\lambda_i}$. Par conséquent,

$$\dim V = \sum_{j=1}^{p} \dim E_{\lambda_j} \leq \sum_{j=1}^{p} \dim F_{\lambda_j} = \dim V,$$

d'où $\dim E_{\lambda_j} = \dim F_{\lambda_j}$ pour tout j. On a donc deux espaces de même dimension finie dont l'un est inclus dans l'autre. Ils sont donc égaux, c'est-à-dire $E_{\lambda_j} = F_{\lambda_j}$.

Montrons que la deuxième assertion implique la troisième. Si $E_{\lambda_j} = F_{\lambda_j}$, on a $(f - \lambda_j \mathrm{id})|_{F_{\lambda_j}} = 0$ et l'indice de nilpotence de $(f - \lambda_j \mathrm{id})|_{F_{\lambda_j}}$ vaut donc 1. Les bases réparties en chaînes sont donc toutes de longueur 1, ce qui signifie que que les blocs correspondants sont de dimension 1.

Enfin, il est clair que la troisième assertion implique la première. ∎

1.5 Polynôme minimum

Dans cette section, nous allons associer à tout endomorphisme f un second polynôme. Nous montrerons ensuite le lien entre ce polynôme, les sous-espaces caractéristiques de f et sa forme de Jordan.

Proposition 1.5.1 *Soit f un endomorphisme de V. Il existe un unique polynôme monique $\mathcal{M}_f$ (c'est-à-dire de coefficient principal égal à 1) tel que*

$$\big\{P \in \mathbb{C}[z] : P(f) = 0\big\} = \big\{Q\mathcal{M}_f : Q \in \mathbb{C}[z]\big\}.$$

Démonstration : Considérons un polynôme non-nul monique $\mathcal{M}_f$ de degré minimum tel que $\mathcal{M}_f(f) = 0$ (un tel polynôme existe par le Théorème 1.1.6 de Caley-Hamilton). Il est clair que

$$\big\{Q\mathcal{M}_f : Q \in \mathbb{C}[z]\big\} \subseteq \big\{P \in \mathbb{C}[z] : P(f) = 0\big\}.$$

Montrons l'autre inclusion. Soit donc $P \in \mathbb{C}[z]$ tel que $P(f) = 0$. En effectuant la division euclidienne de P par $\mathcal{M}_f$, on trouve des polynômes Q et R tels que

$$P = Q\mathcal{M}_f + R,$$

où $\deg R < \deg \mathcal{M}_f$. Ainsi $0 = P(f) = Q(f)\mathcal{M}_f(f) + R(f) = R(f)$, d'où $R(f) = 0$. Or, le degré de R est strictement inférieur à celui de $\mathcal{M}_f$, et la définition de $\mathcal{M}_f$ implique donc que $R = 0$. Ainsi $P = Q\mathcal{M}_f$ et la deuxème inclusion s'ensuit.

Supposons à présent que les polynômes $\mathcal{M}_f$ et $\mathcal{N}_f$ conviennent. Alors, il existe $P, Q \in \mathbb{C}[z]$ tels que $\mathcal{M}_f = P\mathcal{N}_f$ et $\mathcal{N}_f = Q\mathcal{M}_f$. Donc $\mathcal{M}_f = PQ\mathcal{M}_f$ et on en tire que P et Q sont des constantes. Comme les polynômes sont moniques, on a $P = Q = 1$. ■

Définition 1.5.2 Le polynôme $\mathcal{M}_f$ dont il est question dans la Proposition 1.5.1 est appelé le *polynôme minimum* de f.

Proposition 1.5.3 *Soit f un endomorphisme de V. Les polynômes caractéristiques et minimum de f ont les mêmes zéros.*

Démonstration : En utilisant le Théorème 1.1.6 de Caley Hamilton et la Proposition 1.5.1 , on sait que le polynôme minimum divise le polynôme caractéristique de f. Donc tout zéro de $\mathcal{M}_f$ est un zéro de p_f. Réciproquement, soit λ un zéro de p_f. Alors λ est une valeur propre de f et il existe $x \in V \setminus \{0\}$ tel que $f(x) = \lambda x$. Par conséquent, on a

$$f^k(x) = f^{k-1}(f(x)) = f^{k-1}(\lambda x) = \lambda f^{k-1}(x) = \cdots = \lambda^k x$$

pour tout $k \geq 1$. Il s'ensuit que

$$0 = \mathcal{M}_f(f)(x) = \mathcal{M}_f(\lambda)x,$$

d'où $\mathcal{M}_f(\lambda) = 0$. ■

Si $\lambda_1, \ldots, \lambda_p$ sont les valeurs propres de f, on peut donc écrire

$$p_f(\lambda) = (-1)^m \prod_{j=1}^{p} (\lambda - \lambda_j)^{\mu_j}$$

et

$$\mathcal{M}_f(\lambda) = \prod_{j=1}^{p} (\lambda - \lambda_j)^{m_j}$$

avec $1 \leq m_j \leq \mu_j$.

Théorème 1.5.4 (Stabilisation des images et des noyaux) *Soit f un endomorphisme de V de polynôme mimimum*

$$\mathcal{M}_f(\lambda) = \prod_{j=1}^{p} (\lambda - \lambda_j)^{m_j}$$

où $\lambda_1, \ldots, \lambda_p$ sont les valeurs propres distinctes de f. Pour tout $j \in \{1, \ldots, p\}$, on a

$$\ker(f - \lambda_j id)^0 \subsetneq \ker(f - \lambda_j id) \subsetneq \ker(f - \lambda_j id)^2 \subsetneq \cdots \subsetneq \ker(f - \lambda_j id)^{m_j} = \ker(f - \lambda_j id)^{m_j+1} = \ldots$$

et

$$\mathrm{im}(f - \lambda_j id)^0 \supsetneq \mathrm{im}(f - \lambda_j id) \supsetneq \mathrm{im}(f - \lambda_j id)^2 \supsetneq \cdots \supsetneq \mathrm{im}(f - \lambda_j id)^{m_j} = \mathrm{im}(f - \lambda_j id)^{m_j+1} = \ldots$$

De plus,

$$V = \ker(f - \lambda_j id)^{m_j} \oplus \mathrm{im}(f - \lambda_j id)^{m_j}.$$

Démonstration : Pour simplifier les notations, on note $g_j = f - \lambda_j \mathrm{id}$ pour tout $j \in \{1, \ldots, p\}$. Il est clair que $\ker g_j^k \subseteq \ker g_j^{k+1}$ pour tout $k \geq 0$. Montrons que si k est tel que $\ker g_j^k = \ker g_j^{k+1}$, alors $\ker g_j^{k+1} = \ker g_j^{k+2}$. Il suffit de montrer que $\ker g_j^{k+1} \supseteq \ker g_j^{k+2}$. Soit donc $x \in \ker g_j^{k+2}$. Alors $0 = g_j^{k+2}(x) = g_j^{k+1}(g_j(x))$, d'où $g_j(x) \in \ker g_j^{k+1} = \ker g_j^k$. Par conséquent, on obtient $g_j^k(g_j(x)) = g_j^{k+1}(x) = 0$ et donc $x \in \ker g_j^{k+1}$. Ainsi, si la suite des noyaux se stabilise pour une valeur de k, elle devient stationnaire à partir de k.

Montrons à présent que cette stabilisation a lieu en m_j. Commençons par montrer que

$$\ker g_j^{m_j-1} \subsetneq \ker g_j^{m_j}.$$

Le polynôme

$$P(\lambda) = (\lambda - \lambda_j)^{m_j-1} \prod_{i=1, i\neq j}^{p} (\lambda - \lambda_i)^{m_i}$$

est un polynôme de degré strictement inférieur à $\mathcal{M}_f$. Par conséquent, $P(f) \neq 0$ et il existe $x \in V$ tel que

$$P(f)(x) = g_j^{m_j-1} \prod_{i=1, i\neq j}^{p} g_i^{m_i}(x) \neq 0.$$

Posons

$$y = \prod_{i=1, i\neq j}^{p} g_i^{m_i}(x).$$

Alors $y \notin \ker g_j^{m_j-1}$ mais $g_j^{m_j}(y) = 0$ par définition du polynôme minimum.

En particulier, on a aussi $\ker g_j^k \subsetneq \ker g_j^{k+1}$ pour tout $k \leq m_j$ car sinon, vu le début de la preuve, on aurait $\ker g_j^{m_j-1} = \ker g_j^{m_j}$.

Nous allons à présent prouver que $\ker g_j^{m_j} = \ker g_j^{m_j+1}$. Pour cela, il suffit bien sûr de montrer que $\ker g_j^{m_j} \supseteq \ker g_j^{m_j+1}$. Considérons donc $x \in \ker g_j^{m_j+1}$. Par définition de g_j, on a

$$0 = g_j(g_j^{m_j})(x) = (f - \lambda_j \mathrm{id}) g_j^{m_j}(x)$$

et donc $g_j^{m_j}(x)$ est un vecteur propre de f de valeur propre λ_j. En particulier, on a

$$g_k(g_j^{m_j})(x) = (f - \lambda_k \mathrm{id})(g_j^{m_j})(x) = (\lambda_j - \lambda_k)(g_j^{m_j})(x)$$

pour tout $k \in \{1, \ldots, p\}$. Par conséquent,

$$0 = \mathcal{M}_f(f) = g_1^{m_1} \ldots g_j^{m_j} \ldots g_p^{m_p}(x) = \prod_{i=1, i\neq j}^{p} (\lambda_j - \lambda_i)^{m_i} g_j^{m_j}(x),$$

où l'égalité à zéro provient de la définition du polynôme minimum. Puisque les valeurs propres sont distinctes, on obtient que $g_j^{m_j}(x) = 0$, d'où $x \in \ker g_j^{m_j}$. Ceci conclut la première partie de la preuve.

La deuxième partie de la preuve découle du théorème de la dimension qui donne

$$\dim V = \dim \ker g_j^k + \dim \operatorname{im} g_j^k$$

pour tout $k \geq 0$ et du fait que $\operatorname{im} g_j^k \supseteq \operatorname{im} g_j^{k+1}$ pour tout $k \geq 0$.
Pour conclure, il reste à montrer que

$$V = \ker g_j^{m_j} \oplus \operatorname{im} g_j^{m_j}.$$

On sait déjà que $\ker g_j^{m_j} + \operatorname{im} g_j^{m_j} \subseteq V$. Le théorème de la dimension nous donne alors l'égalité et il suffit de montrer que $\ker g_j^{m_j} \cap \operatorname{im} g_j^{m_j} = \{0\}$. Considérons donc $x \in \ker g_j^{m_j} \cap \operatorname{im} g_j^{m_j}$. Alors, il existe $y \in V$ tel que $x = g_j^{m_j}(y)$. On a $g_j^{2m_j}(y) = g_j^{m_j}(x) = 0$, donc $y \in \ker g_j^{2m_j} = \ker g_j^{m_j}$. On en tire que $x = g_j^{m_j}(y) = 0$, ce qui conclut la preuve. ■

Rappelons que le sous-espace caractéristique associé à λ_j est défini par $\ker(f - \lambda_j \mathrm{id})^{\mu_j}$, où μ_j est la multiplicité algébrique de λ_j. Puisque $\mu_j \geq m_j$, le théorème précédent donne directement le résultat suivant.

Corollaire 1.5.5 *Soit f un endomorphisme de V. Si λ_j est une valeur propre de f dont la multiplicité comme zéro du polynôme minimum est m_j, alors*

$$F_{\lambda_j} = \ker(f - \lambda_j id)^{m_j} = \ker(f - \lambda_j id)^l$$

pour tout $l \geq m_j$.

Remarque 1.5.6 Le Théorème 1.5.4 implique que l'indice de nilpotence de $(f - \lambda_j \mathrm{id})|_{F_{\lambda_j}}$ est égal à m_j. Par conséquent, la longueur de la plus longue chaîne de tout tableau formant une base répartie en chaînes de F_{λ_j} est égal à m_j. La taille de ce tableau est égale à μ_j, la dimension de F_{λ_j}.

Proposition 1.5.7 *Un endomorphisme f est diagonalisable si et seulement si son polynôme minimum ne contient que des zéros simples.*

Démonstration : Rappelons que

$$E_{\lambda_j} = \ker(f - \lambda_j \mathrm{id})\,, \quad F_{\lambda_j} = \ker(f - \lambda_j \mathrm{id})^{m_j} \quad \text{et} \quad E_{\lambda_j} \subseteq F_{\lambda_j}.$$

Par le Théorème 1.5.4, on a l'égalité entre les deux espaces si et seulement si $m_j = 1$. On conclut en utilisant la Proposition 1.4.9. ■

1.6 Exemple

Considérons la matrice A définie par

$$A = \begin{pmatrix} 5 & -1 & -3 & 2 & -5 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & -2 \\ 0 & -1 & 0 & 3 & 1 \\ 1 & -1 & -1 & 1 & 1 \end{pmatrix}.$$

Cette matrice est associée à l'endomorphisme $f : \mathbb{C}^5 \to \mathbb{C}^5 : x \mapsto Ax$. Afin de la réduire à la forme de Jordan, on commence par calculer son polynôme caractéristique

$$\det(A - \lambda I) \quad = \quad \begin{vmatrix} 5-\lambda & -1 & -3 & 2 & -5 \\ 0 & 2-\lambda & 0 & 0 & 0 \\ 1 & 0 & 1-\lambda & 1 & -2 \\ 0 & -1 & 0 & 3-\lambda & 1 \\ 1 & -1 & -1 & 1 & 1-\lambda \end{vmatrix} = -(\lambda-3)^2(\lambda-2)^3.$$

Ainsi, les valeurs propres de A sont 2 de multiplicité algébrique 3 et 3 de multiplicité algébrique 2. Déterminons les sous-espaces propres correspondants.

Pour la valeur propre 2, étudions la matrice

$$A - 2I = \begin{pmatrix} 2 & -1 & -3 & 2 & -5 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 1 & -2 \\ 0 & -1 & 0 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 \end{pmatrix}.$$

On vérifie que le rang de cette matrice est 3. Par conséquent,

$$\dim E_2 = \dim \ker(A - 2I) = 5 - 3 = 2 < 3.$$

La matrice A n'est donc pas diagonalisable. Vu le théorème de stabilisation des noyaux et puisque $\dim F_2 = 3$, on sait donc que

$$F_2 = \ker(A - 2I)^2.$$

On calcule donc

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

Le tableau d'une base répartie en chaînes pour F_2 est de la forme

$$\begin{array}{|c|c|} \hline * & * \\ \hline * & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array}$$

Remarquons que le vecteur $e_2 + e_4$ est la tête d'une chaîne de longueur 2 engendrée par $A - 2I$ car

$$(A - 2I)^2(e_2 + e_4) = 0 \quad \text{et} \quad (A - 2I)(e_2 + e_4) = e_1 + e_3.$$

Il faut encore trouver un vecteur qui est la tête d'une chaîne de longueur 1 (donc dans le noyau de $A - 2I$) et linéairement indépendant avec la queue $e_1 + e_3$. On vérifie aisément que le vecteur $e_2 - 2e_3 + e_5$ convient.

Pour la valeur propre 3, on s'intéresse à la matrice

$$A - 3I = \begin{pmatrix} 2 & -1 & -3 & 2 & -5 \\ 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & -2 & 1 & -2 \\ 0 & -1 & 0 & 0 & 1 \\ 1 & -1 & -1 & 1 & -2 \end{pmatrix}.$$

On vérifie facilement que son rang est 4. Par conséquent, le sous-espace propre associé à la valeur propre 3 est de dimension 1. Le rang de la matrice

$$(A-3I)^2 = \begin{pmatrix} -4 & 2 & 5 & -4 & 8 \\ 0 & 1 & 0 & 0 & 0 \\ -2 & 0 & 3 & -2 & 4 \\ 1 & 0 & -1 & 1 & -2 \\ -1 & 1 & 1 & -1 & 2 \end{pmatrix}$$

est nécessairement 3 puisque $\dim \ker(A-3I)^2 = \dim F_3 = 2$. Ainsi, le tableau d'une base répartie en chaînes pour F_3 est de la forme

*	*

On vérifie facilement que le vecteur $2e_1 + e_5$ est la tête d'une chaîne de longueur 2 engendrée par $A - 3I$, c'est-à-dire

$$(A-3I)^2(2e_1 + e_5) = 0 \text{ et } (A-3I)(2e_1+e_5) = -e_1 + e_4.$$

Finalement, on construit la matrice

$$S = \begin{pmatrix} 1 & 0 & 0 & -1 & 2 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & -2 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

qui est telle que

$$S^{-1}AS = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

Notons également, vu la stabilisation des noyaux obtenue, que nous savons que

$$\mathcal{M}_A(\lambda) = (\lambda-2)^2(\lambda-3)^2.$$

Illustrons à présent ce que nous venons de calculer dans Mathematica. On définit la matrice A par

A = {{5, −1, −3, 2, −5}, {0, 2, 0, 0, 0}, {1, 0, 1, 1, −2}, {0, −1, 0, 3, 1}, {1, −1, −1, 1, 1}}

A//MatrixForm

$$\begin{pmatrix} 5 & -1 & -3 & 2 & -5 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & -2 \\ 0 & -1 & 0 & 3 & 1 \\ 1 & -1 & -1 & 1 & 1 \end{pmatrix}$$

La réduction à la forme de Jordan est donnée par

J = JordanDecomposition[A]

$\{\{\{2, 1, 0, -1, 2\}, \{1, 0, 1, 0, 0\}, \{0, 1, 0, 0, 0\}, \{0, 0, 1, 1, 0\}, \{1, 0, 0, 0, 1\}\},$

$\{\{2, 0, 0, 0, 0\}, \{0, 2, 1, 0, 0\}, \{0, 0, 2, 0, 0\}, \{0, 0, 0, 3, 1\}, \{0, 0, 0, 0, 3\}\}\}$

Map[MatrixForm, J]

$$\left\{\begin{pmatrix} 2 & 1 & 0 & -1 & 2 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}\right\}$$

La première entrée donne la matrice de changement de base (remarquez que celle-ci n'est pas unique). La deuxième entrée donne la forme de Jordan. Vérifions à présent que les valeurs propres de A sont bien 2 (triple) et 3 (double).

Eigenvalues[A]

$\{3, 3, 2, 2, 2\}$

On peut également voir que la matrice n'est pas diagonalisable en cherchant ses vecteurs propres.

Eigenvectors[A]

$\{\{-1, 0, 0, 1, 0\}, \{0, 0, 0, 0, 0\}, \{2, 1, 0, 0, 1\}, \{1, 0, 1, 0, 0\}, \{0, 0, 0, 0, 0\}\}$

Vérifions à présent qu'on a bien obtenu la forme de Jordan de A. On définit la matrice S par

S = J[[1]]

$\{\{2, 1, 0, -1, 2\}, \{1, 0, 1, 0, 0\}, \{0, 1, 0, 0, 0\}, \{0, 0, 1, 1, 0\}, \{1, 0, 0, 0, 1\}\}$

Inverse[S].A.S//MatrixForm

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

% == $\boldsymbol{J}$[[2]]

True

Pour finir, illustrons le calcul du polynôme caractéristique de A et sa factorisation.

CharacteristicPolynomial[$\boldsymbol{A, x}$]

$72 - 156x + 134x^2 - 57x^3 + 12x^4 - x^5$

Factor[%]

$-(-3+x)^2(-2+x)^3$

1.7 Applications

Dans beaucoup d'applications, il peut être utile de calculer la puissance $n^{\text{ème}}$ d'une matrice carrée A. Le Théorème 1.4.7 nous donne une matrice inversible S telle que $S^{-1}AS = J$ où J est une matrice diagonale par blocs, dont tous les blocs de la diagonale sont des blocs de Jordan. Alors $A = SJS^{-1}$ et pour tout $n \in \mathbb{N}$, on a

$$A^n = (SJS^{-1})(SJS^{-1})\dots(SJS^{-1}) = SJ^nS^{-1}.$$

On s'est donc ramené à un problème de calcul des puissances de J. Remarquons que si

$$J = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & J_s \end{pmatrix},$$

alors

$$J^n = \begin{pmatrix} J_1^n & 0 & \dots & 0 \\ 0 & J_2^n & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & J_s^n \end{pmatrix}.$$

Il reste donc à trouver un formule pour les puissances d'un bloc de Jordan.

Lemme 1.7.1 *Soit J_λ un bloc de Jordan de dimension k. Pour tout $n \geq 0$, on a*

$$J_\lambda^n = \begin{pmatrix} \lambda^n & C_n^1 \lambda^{n-1} & \dots & C_n^{k-2} \lambda^{n-k+2} & C_n^{k-1} \lambda^{n-k+1} \\ 0 & \lambda^n & \dots & C_n^{k-3} \lambda^{n-k+3} & C_n^{k-2} \lambda^{n-k+2} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \lambda^n & C_n^1 \lambda^{n-1} \\ 0 & \dots & \dots & 0 & \lambda^n \end{pmatrix}$$

où on utilise la convention que $C_n^j = 0$ si $j > n$.

Démonstration : Notons N la matrice carrée de dimension k définie par

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

La matrice N est nilpotente d'indice k. De plus, $J_\lambda = \lambda I + N$ et en utilisant la formule du binôme de Newton (qui est valide puisque les matrices λI et N commutent), on obtient

$$\begin{aligned} J_\lambda^n = (\lambda I + N)^n &= \sum_{j=0}^{n} C_n^j \lambda^{n-j} N^j \\ &= \sum_{j=0}^{\min\{k-1,n\}} C_n^j \lambda^{n-j} N^j . \end{aligned}$$

Pour tous $i, j \in \{1, \dots, k\}$, notons $E_{i,j}$ la matrice carrée de dimension k qui a un 1 à la position (i, j) et des zéros ailleurs. Alors, pour tout $j \in \{0, \dots, k-1\}$

$$N^j = \sum_{i=1}^{k-1} E_{i,i+j}$$

et on en tire la conclusion. ∎

1.7.1 Suites linéaires récurrentes homogènes à coefficients constants

Définition 1.7.2 Une suite de complexes $(u_n)_{n \in \mathbb{N}}$ est dite *linéaire récurrente* si elle satisfait une équation de récurrence linéaire homogène à coefficients constants de la forme

$$u_{n+k} = a_{k-1} u_{n+k-1} + \cdots + a_0 u_n \quad \text{pour tout } n \in \mathbb{N}, \tag{1.2}$$

avec $a_0, \dots, a_{k-1} \in \mathbb{C}$ et $a_0 \neq 0$. On dit que k est l'*ordre* de la relation.

Remarquons que l'ensemble des suites vérifiant l'équation (1.2) est un sous-espace vectoriel E de $\mathbb{C}^{\mathbb{N}}$: si $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ satisfont (1.2), alors $(u_n+v_n)_{n\in\mathbb{N}}$ et $(\lambda u_n)_{n\in\mathbb{N}}$ vérifient également (1.2) pour tout $\lambda \in \mathbb{C}$. De plus, le sous-espace E des solutions de (1.2) est isomorphe à $\mathbb{C}^k$: à toute condition initiale $U_0 := (u_0, \dots, u_{k-1}) \in \mathbb{C}^k$ correspond une unique suite linéaire récurrente, et inversément.

Notre but est de donner une forme explicite pour u_n en fonction de n.

Définition 1.7.3 La *matrice compagnon* de l'équation linéaire récurrente (1.2) est définie par

$$A = \begin{pmatrix} a_{k-1} & a_{k-2} & \dots & \dots & a_0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

Remarquons qu'en posant

$$U_n = \begin{pmatrix} u_{n+k-1} \\ \vdots \\ u_{n+1} \\ u_n \end{pmatrix}$$

pour tout $n \in \mathbb{N}$, l'équation (1.2) se réécrit

$$U_{n+1} = AU_n \quad \text{pour tout } n \in \mathbb{N}, \tag{1.3}$$

où U_0 est donné. En particulier, pour tout $n \in \mathbb{N}$, on a

$$U_n = A^n U_0.$$

La difficulté est donc réduite au calcul des puissances de A. Comme expliqué précédemment, la réduction à la forme de Jordan fournit un outil idéal pour ce calcul. Commençons par remarquer que le polynôme caractéristique d'une matrice compagnon a toujours la même forme, ce qui facilite le calcul de ses valeurs propres.

Lemme 1.7.4 *Le polynôme caractéristique de la matrice compagnon A est donné par*

$$p_A(\lambda) = (-1)^k(\lambda^k - a_{k-1}\lambda^{k-1} - \dots - a_1\lambda - a_0).$$

En particulier, comme $a_0 \neq 0$, 0 n'est pas une valeur propre de A.

Démonstration : Il suffit de développer le déterminant de $A - \lambda I$ par rapport à sa première ligne. Il vient

$$
\begin{aligned}
\det(A-\lambda I) &= \begin{vmatrix} a_{k-1}-\lambda & a_{k-2} & \dots & \dots & a_0 \\ 1 & -\lambda & 0 & \dots & 0 \\ 0 & 1 & -\lambda & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -\lambda \end{vmatrix} \\
&= (a_{k-1}-\lambda) \begin{vmatrix} -\lambda & 0 & \dots & \dots & 0 \\ 1 & -\lambda & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & -\lambda \end{vmatrix} - a_{k-2} \begin{vmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & -\lambda & 0 & \dots & 0 \\ 0 & 1 & -\lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & -\lambda \end{vmatrix} \\
&\quad + \dots + (-1)^k a_1 \begin{vmatrix} 1 & -\lambda & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\lambda & 0 \\ \vdots & & \ddots & 1 & 0 \\ 0 & \dots & \dots & 0 & -\lambda \end{vmatrix} + (-1)^{k+1} a_0 \begin{vmatrix} 1 & -\lambda & 0 & \dots & 0 \\ 0 & 1 & -\lambda & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & -\lambda \\ 0 & \dots & \dots & 0 & 1 \end{vmatrix} \\
&= (a_{k-1}-\lambda)(-\lambda)^{k-1} - a_{k-2}(-\lambda)^{k-2} + \dots + (-1)^k(-\lambda)a_1 + (-1)^{k+1}a_0 \\
&= (-1)^k\bigl(\lambda^k - a_{k-1}\lambda^{k-1} - \dots - a_1\lambda - a_0\bigr)
\end{aligned}
$$

ce qui conclut la preuve. ■

Par le Théorème 1.4.7, il existe une matrice de changement de base S et une matrice

$$
J = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & J_s \end{pmatrix}
$$

formée de blocs de Jordan telle que $J = S^{-1}AS$. Si $(u_n)_{n\in\mathbb{N}}$ satisfait l'équation (1.2), on obtient

$$
\begin{pmatrix} u_{n+k-1} \\ \vdots \\ u_{n+1} \\ u_n \end{pmatrix} = SJ^nS^{-1} \begin{pmatrix} u_{k-1} \\ \vdots \\ u_1 \\ u_0 \end{pmatrix} \tag{1.4}
$$

pour tout $n \in \mathbb{N}$, où

$$
J^n = \begin{pmatrix} J_1^n & 0 & \dots & 0 \\ 0 & J_2^n & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & J_s^n \end{pmatrix}.
$$

Soient $\lambda_1, \dots, \lambda_p$ les valeurs propres distinctes de la matrice compagnon A de l'équation linéaire récurrente (1.2). Notons $\mu_1, \dots, \mu_p$ leurs multiplicités respectives. Fixons un bloc de Jordan J_i

associé à la valeur propre λ_j de multiplicité algébrique μ_j. Par le Lemme 1.7.1, on sait que les entrées de la matrice J_i^n sont de la forme

$$C_n^m \lambda^{n-m} = \frac{1}{m!\lambda_j^m} n(n-1)\dots(n-m+1)\lambda_j^n,$$

où m est strictement inférieur à la taille du bloc, donc en particulier à la multiplicité μ_j. Ainsi, les entrées de la matrice J_i^n appartiennent à l'espace vectoriel sur $\mathbb{C}$ engendré par

$$\lambda_j^n, n\lambda_j^n, n^2\lambda_j^n, \dots, n^{\mu_j-1}\lambda_j^n.$$

Au total, les entrées de la matrice J^n appartiennent donc à l'espace vectoriel engendré par

$$\left\{n^l \lambda_j^n : j \in \{1, \dots, p\}, l \in \{0, \dots, \mu_j - 1\}\right\}.$$

Proposition 1.7.5 *Soient $\lambda_1, \dots, \lambda_p$ les valeurs propres distinctes de la matrice compagnon A de l'équation linéaire récurrente* (1.2). *Notons $\mu_1, \dots, \mu_p$ leurs multiplicités respectives. Pour tout $j \in \{1, \dots, p\}$ et tout $l \in \{0, \dots, \mu_j - 1\}$, la suite*

$$u^{j,l} := (n^l \lambda_j^n)_{n \in \mathbb{N}}$$

est une solution de l'équation (1.2).

Démonstration : Il faut montrer que pour tout $n \in \mathbb{N}$, on a

$$(n+k)^l \lambda_j^{n+k} = \sum_{i=0}^{k-1} a_i (n+i)^l \lambda_j^{n+i}. \tag{1.5}$$

Par hypothèse, on sait que λ_j est une racine de p_A de multiplicité $\mu_j > l$. Par conséquent, $D^r p_A(\lambda_j) = 0$ pour tout $r \leq l$. Par le Lemme 1.7.4, on sait que

$$D^r p_A(z) = (-1)^k \left(\frac{k!}{(k-r)!} z^{k-r} - \sum_{i=r}^{k-1} a_i \frac{i!}{(i-r)!} z^{i-r} \right)$$

et en évaluant cette dérivée en λ_j, on obtient

$$\frac{k!}{(k-r)!} \lambda_j^{k-r} = \sum_{i=r}^{k-1} a_i \frac{i!}{(i-r)!} \lambda_j^{i-r}.$$

En multipliant par λ_j^r, il vient

$$\frac{k!}{(k-r)!} \lambda_j^{k} = \sum_{i=r}^{k-1} a_i \frac{i!}{(i-r)!} \lambda_j^{i}. \tag{1.6}$$

Considérons à présent pour tout $r \in \{1, \dots, l\}$, le polynôme P_r de degré r défini par

$$P_r(z) = z(z-1)\dots(z-r+1).$$

Si $i \geq r$, on a

$$P_r(i) = \frac{i!}{(i-r)!}.$$

Il est clair que $P_0 = 1, P_1, \ldots, P_l$ forment une base de l'espace des polynômes de degré au plus l (il suffit de montrer qu'ils forment une partie génératrice puisqu'ils sont en nombre égal à la dimension ; c'est direct en procédant de proche en proche). Par conséquent, il existe des constantes $c_1, \ldots, c_l$ telles que

$$(n+\lambda)^l = \sum_{r=0}^{l} c_r P_r(\lambda).$$

En particulier, pour tout $i \leq k$, on a

$$(n+i)^l = \sum_{r=0}^{l} c_r P_r(i). \tag{1.7}$$

De plus, (1.6) se réécrit

$$P_r(k)\lambda_j^k = \sum_{i=0}^{k-1} a_i P_r(i) \lambda_j^i \tag{1.8}$$

puisque $P_r(i) = 0$ si $i < r$.

En combinant (1.7) et (1.8), on obtient

$$\begin{aligned}
(n+k)^l \lambda_j^{n+k} &= \sum_{r=0}^{l} c_r \left(P_r(k)\lambda_j^k\right) \lambda_j^n \\
&= \sum_{r=0}^{l} c_r \left(\sum_{i=0}^{k-1} a_i P_r(i)\lambda_j^i\right) \lambda_j^n \\
&= \sum_{i=0}^{k-1} a_i \left(\sum_{r=0}^{l} c_r P_r(i)\right) \lambda_j^{i+n} \\
&= \sum_{i=0}^{k-1} a_i (n+i)^l \lambda_j^{i+n},
\end{aligned}$$

d'où (1.5). ■

Ce résultat montre que les suites $u^{j,l}$, $j \in \{1, \ldots, p\}$ et $l \in \{0, \ldots, \mu_j - 1\}$, sont $\mu_1 + \cdots + \mu_p = k$ solutions de l'équation (1.2). Comme l'espace E des solutions est un espace de dimension k, il est légitime d'espérer que ces solutions forment une base de E.

Remarquons que puisque l'ensemble E des solutions de (1.2) est un sous-espace vectoriel, toute suite de la forme

$$u_n = \sum_{j=1}^{p} P_j(n) \lambda_j^n, \quad n \in \mathbb{N}, \tag{1.9}$$

où pour tout $j \in \{1, \ldots, p\}$, P_j est un polynôme de degré strictement inférieur à μ_j, est une solution de (1.2).

Proposition 1.7.6 *Les k solutions $u^{l,j}$, $j \in \{1,\ldots,p\}$ et $l \in \{0,\ldots,\mu_j-1\}$ de (1.2) données dans la Proposition 1.7.5 sont linéairement indépendantes. Elles forment donc une base des solutions de* (1.2).

Démonstration : Commençons par fixer une valeur propre λ_j et montrons que les solutions $u^{l,j}$, $l \in \{0,\ldots,\mu_j-1\}$ sont linéairement indépendantes. Supposons qu'il existe des constantes $c_0,\ldots,c_{\mu_j-1}$ telles que

$$\sum_{l=0}^{\mu_j-1} c_l u^{l,j} = 0.$$

Alors, pour tout $n \in \mathbb{N}$, on a

$$\sum_{l=0}^{\mu_j-1} c_l n^l \lambda_j^n = 0$$

et puisque $\lambda_j \neq 0$, cela implique directement que $c_1 = \cdots = c_{\mu_l-1} = 0$.

Montrons à présent que les vecteurs de conditions initiales de solutions de (1.2) de la forme $(P_j(n)\lambda_j^n)_{n\in\mathbb{N}}$, où P_l est un polynôme de degré strictement inférieur à μ_j appartiennent à l'espace caractéristique F_{λ_j} de la matrice compagnon A.

Soit donc P un polynome de degré $l < \mu_j$, et u la solution de (1.2) définie par

$$u = \left(P(n)\lambda_j^n\right)_{n\in\mathbb{N}}.$$

On a $(A-\lambda_j I)U_0 = AU_0 - \lambda_j U_0 = U_1 - \lambda_j U_0$ par la relation (1.3). Cette relation se réécrit

$$(A-\lambda_j I)U_0 = \begin{pmatrix} u_k - \lambda_j u_{k-1} \\ \vdots \\ u_2 - \lambda_j u_1 \\ u_1 - \lambda_j u_0 \end{pmatrix}.$$

Pour tout $r \in \{1,\ldots,k\}$, on a $u_r - \lambda_j u_{r-1} = \big(P(r)-P(r-1)\big)\lambda_j^r = Q(r)\lambda_j^r$ où Q est un polynôme de degré strictement inférieur à l, et indépendant de r. En particulier, la suite $v = \left(Q(n)\lambda_j^n\right)_{n\in\mathbb{N}}$ est encore une solution de (1.3). On réitère le processus pour calculer $(A-\lambda_j I)^2 U_0 = (A-\lambda_j I)V_0$, et en faisant diminer le degré du polynôme étape par étape, on obtient finalement

$$(A-\lambda_j I)^{\mu_j} U_0 = 0.$$

Comme les espaces caractéristiques $F_{\lambda_1},\ldots,F_{\lambda_p}$ sont en position de somme directe, des solutions associées à des valeurs propres distinctes sont linéairement indépendantes, ce qui conclut la preuve. ∎

Au total, on a obtenu le résultat suivant.

Théorème 1.7.7 *Soient $\lambda_1,\ldots,\lambda_p$ les valeurs propres distinctes de la matrice compagnon A de l'équation linéaire récurrente* (1.2). *Notons $\mu_1,\ldots,\mu_p$ leurs multiplicités respectives.*

La solution générale de l'équation (1.2) *est donnée par la suite* $(u_n)_{n\in\mathbb{N}}$ *de terme général*

$$u_n = \sum_{j=1}^{p} P_j(n)\lambda_j^n, \tag{1.10}$$

où pour tout $j \in \{1, \ldots, p\}$, P_j *est un polynôme de degré strictement inférieur à* μ_j.

Exemple 1.7.8 Considérons la suite $(u_n)_{n\in\mathbb{N}}$ définie par l'équation linéaire récurrente

$$u_{n+2} = 2u_{n+1} - 2u_n$$

et les conditions initiales $u_0 = 0$, $u_1 = 1$. On peut obtenir les 25 premiers éléments de la suite en les calculant de proche en proche, ou en utilisant la commande

LinearRecurrence[2, −2, 0, 1, 25]

$\{0, 1, 2, 2, 0, -4, -8, -8, 0, 16, 32, 32, 0, -64, -128, -128, 0, 256, 512, 512, 0, -1024, -2048, -2048, 0\}$

La matrice compagnon de l'équation est

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

et son polynôme caractéristique est donné par $p_A(\lambda) = \lambda^2 - 2\lambda + 2$. Ses racines sont $1+i$ et $1-i$, toutes les deux de multiplicité 1. Ainsi, il existe des constantes C_1 et C_2 telles que

$$u_n = C_1(1+i)^n + C_2(1-i)^n.$$

Les conditions initiales donnent le système d'équations

$$\left\{ \begin{array}{l} 0 = C_1 + C_2 \\ 1 = C_1(1+i) + C_2(1-i), \end{array} \right.$$

d'où $C_1 = -i/2$ et $C_2 = i/2$. Au total,

$$u_n = \frac{i}{2}\big((1-i)^n - (1+i)^n\big) = \frac{i}{2}(\sqrt{2})^n \left(e^{-in\pi/4} - e^{in\pi/4}\right) = (\sqrt{2})^n \sin\left(\frac{n\pi}{4}\right).$$

Remarquons que cette suite est bien, comme on s'y attendait initialement, une suite d'entiers.

1.7.2 Exponentielle matricielle et équations différentielles

Définition 1.7.9 Un système d'équations différentielles d'ordre 1 linéaires homogènes à coefficients constants est un système de la forme

$$\left\{ \begin{array}{rcl} u_1'(t) & = & a_{1,1}u_1(t) + \cdots + a_{1,n}u_n(t) \\ u_2'(t) & = & a_{2,1}u_1(t) + \cdots + a_{2,n}u_n(t) \\ & \vdots & \\ u_n'(t) & = & a_{n,1}u_1(t) + \cdots + a_{n,n}u_n(t) \end{array} \right. \tag{1.11}$$

où $u_1, \ldots, u_n$ sont des fonctions dérivables sur $\mathbb{R}$ et les coefficients $a_{i,j}$, $i, j \in \{1, \ldots, n\}$, sont des complexes.

A nouveau, remarquons que l'ensemble des n-uples de fonctions vérifiant l'équation (1.11) est un espace vectoriel. De plus, si deux n-uples de fonctions vérifiant l'équation (1.11) partagent les mêmes conditions initiales, i.e. prennent les mêmes valeurs en 0, alors elles sont égales. Enfin, notons que cet espace est isomorphe à l'ensemble $\mathbb{C}^n$ des conditions initiales en 0 : à toute condition initiale $\big(u_1(0), \ldots, u_n(0)\big) \in \mathbb{C}^n$ correspond une unique solution $u_1, \ldots, u_n$ (l'existence sera prouvée dans la Proposition 1.7.13), et inversément.

Notons que le système (1.11) se réécrit

$$U'(t) = AU(t)$$

où A est la matrice définie par $A = (a_{i,j})_{i\in\{1,\ldots,n\},j\in\{1,\ldots,n\}}$ et

$$U(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_n(t) \end{pmatrix}.$$

Dans le cas où $n = 1$, nous savons que l'équation différentielle $u'(t) = au(t)$ admet pour solution générale $u(t) = Ce^{at}$, où $C = u(0)$. Nous allons voir qu'on un résultat analogue pour $n > 1$. Pour cela, nous devons d'abord définir l'exponentielle matricielle.

Définition 1.7.10 Soit $A \in \mathbb{C}^n_n$. L'*exponentielle matricielle* de A est la matrice définie par

$$\exp(A) = \sum_{k=0}^{+\infty} \frac{A^k}{k!}.$$

Montrons tout d'abord que cette formule a du sens (on considère la convergence entrée par entrée). Pour cela, il faut montrer que pour tout $i, j \in \{1, \ldots, n\}$, la série

$$\sum_{k=0}^{+\infty} \frac{(A^k)_{i,j}}{k!}$$

converge. Si on note

$$\|A\| = \max_{i\in\{1,\ldots,n\}} \sum_{j=1}^{n} |a_{i,j}|,$$

alors pour toutes matrices A et B, on a $\|AB\| \leq \|A\| \, \|B\|$. En effet, pour tout $i \in \{1, \ldots, n\}$, on a

$$\begin{aligned} \sum_{j=1}^{n} |(AB)_{i,j}| = \sum_{j=1}^{n} \left| \sum_{l=1}^{n} a_{i,l} b_{l,j} \right| \quad &\leq \quad \sum_{j=1}^{n} \sum_{l=1}^{n} |a_{i,l}||b_{l,j}| \\ &= \quad \sum_{l=1}^{n} |a_{i,l}| \underbrace{\sum_{j=1}^{n} |b_{l,j}|}_{\leq \max_{l\in\{1,\ldots,n\}} \sum_{j=1}^{n} |b_{l,j}|} \\ &\leq \quad \left(\sum_{l=1}^{n} |a_{i,l}| \right) \|B\| \leq \|A\| \, \|B\|. \end{aligned}$$

Par conséquent, $|(A^k)_{i,j}| \leq \|A^k\| \leq \|A\|^k$. Puisque la série $\sum_{k=0}^{+\infty} \frac{\|A\|^k}{k!}$ converge, on obtient la conclusion.

Remarque 1.7.11 Dans Mathematica, l'exponentielle d'une matrice A s'obtient aisément en utilisant la commande **MatrixExp[*A*]**.

Proposition 1.7.12 *Soient $A, B \in \mathbb{C}_n^n$, $a \in \mathbb{C}$ et $S \in \mathbb{C}_n^n$ une matrice inversible.*

1. $\exp(0) = I$ *et* $\exp(aI) = e^a I$.
2. *Si* $A = \mathrm{diag}(a_1, \ldots, a_n)$, *alors* $\exp(A) = \mathrm{diag}(e^{a_1}, \ldots, e^{a_n})$. *De plus, si A est diagonale par blocs, alors* $\exp(A)$ *l'est aussi et ses blocs sont les exponentielles des blocs de A.*
3. *Si A et B commutent, alors* $\exp(A + B) = \exp(A)\exp(B)$.
4. $\exp(A)$ *est inversible d'inverse* $\exp(-A)$.
5. $\exp(S^{-1}AS) = S^{-1}\exp(A)S$.
6. *Si* $U(t) = \exp(tA)$, $t \in \mathbb{R}$, *alors* $U'(t) = A\exp(tA)$.

Démonstration : 1. et 2. Il s'agit de simples vérifications.

3. Si A et B commutent, on a

$$(A + B)^k = \sum_{j=0}^{k} C_k^j A^j B^{k-j},$$

et on procède alors comme dans le cas de l'exponentielle classique.

4. Cela découle immédiatemment des points précédents puisque $\exp(I) = \exp(A - A)$.

5. Pour tout $K \in \mathbb{N}$, on a

$$\sum_{k=0}^{K} \frac{(S^{-1}AS)^k}{k!} = S^{-1} \sum_{k=0}^{K} \frac{A^k}{k!} S$$

et il suffit ensuite de faire tendre K vers l'infini.

6. Montrons tout d'abord que pour tous $i, j \in \{1, \ldots n\}$, la série définissant $U_{i,j}$ converge uniformément (même normalement) sur tout compact. Fixons donc $[-N, N] \subseteq \mathbb{R}$ et $t \in [-N, N]$. Alors, en procédant comme précédemment, on a pour tout $k \geq 0$

$$\left| \frac{(tA)_{i,j}^k}{k!} \right| \leq \frac{|t|^k \, \|A\|^k}{k!} \leq \frac{N^k \, \|A\|^k}{k!}$$

qui est le terme général de la série définissant l'exponentielle de $N\|A\|$. Pour tous $i, j \in \{1, \ldots, n\}$, on a donc

$$\begin{aligned}(U'(t))_{i,j} = \sum_{k=0}^{+\infty}\left(\frac{t^k(A^k)_{i,j}}{k!}\right)' = \sum_{k=1}^{+\infty} k\frac{t^{k-1}(A^k)_{i,j}}{k!} &= \sum_{k=0}^{+\infty}\frac{t^k(A^{k+1})_{i,j}}{k!}\\ &= \sum_{k=0}^{+\infty}\frac{t^k}{k!}\sum_{l=1}^{n} A_{i,l}(A^k)_{l,j}\\ &= \sum_{l=1}^{n} A_{i,l}\sum_{k=0}^{+\infty}\frac{t^k}{k!}(A^k)_{l,j}\\ &= \sum_{l=1}^{n} A_{i,l}\exp(tA)_{l,j}\\ &= \big(A\exp(tA)\big)_{i,j}.\end{aligned}$$

■

Pour tout vecteur $V \in \mathbb{C}^n$, le point 6 de la Proposition 1.7.12 nous donne que $\exp(tA)V$ est une solution du système (1.11). Par le point 1, on a également que le vecteur de conditions initiales de cette solution est donné par V. Comme les vecteurs $e_1, \ldots, e_n \in \mathbb{C}^n$ forment une base de $\mathbb{C}^n$, on en tire qu'une base des solutions du système (1.11) est donnée par les fonctions

$$\exp(tA)e_1, \ldots, \exp(tA)e_n$$

(il suffit d'exploiter l'isomorphisme entre $\mathbb{C}^n$ et l'ensemble des solutions). On a donc obtenu le résultat suivant.

Proposition 1.7.13 *Soit $A = (a_{i,j})_{i\in\{1,\ldots,n\},j\in\{1,\ldots,n\}}$. La solution la plus générale du système* (1.11) *est donnée par les fonctions $u_1, \ldots, u_n$ définies sur $\mathbb{R}$ par*

$$\begin{pmatrix} u_1(t)\\ u_2(t)\\ \vdots\\ u_n(t)\end{pmatrix} = \exp(tA)\begin{pmatrix} v_1\\ v_2\\ \vdots\\ v_n\end{pmatrix} \tag{1.12}$$

où $v_1, \ldots, v_n \in \mathbb{C}$.

Le résultat suivant est à mettre en lien avec le Théorème 1.7.7 obtenu lors de l'étude des suites linéaires récurrentes.

Proposition 1.7.14 *Soient $\lambda_1, \ldots, \lambda_p$ les valeurs propres distinctes de la matrice A de l'équation différentielle linéaire* (1.11). *Notons $\mu_1, \ldots, \mu_p$ leurs multiplicités respectives. Les solutions de l'équation* (1.11) *sont de la forme*

$$u_k(t) = \sum_{j=1}^{p} P_j^k(t)e^{\lambda_j t}, \quad k \in \{1, \ldots, n\} \tag{1.13}$$

où pour tout $j \in \{1, \dots, p\}$, P_j^k est un polynôme de degré strictement inférieur à μ_j.

Démonstration : Le Théorème 1.4.7 nous donne l'existence d'une matrice S telle que la matrice $J = S^{-1}AS$ est de la forme

$$J = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & J_s \end{pmatrix}$$

où chaque J_i est un bloc de Jordan. Alors

$$\exp(tJ) = \begin{pmatrix} \exp(tJ_1) & 0 & \dots & 0 \\ 0 & \exp(tJ_2) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & \exp(tJ_s) \end{pmatrix}.$$

Fixons un bloc de Jordan J_i associé à la valeur propre λ_j de multiplicité algébrique μ_j. On peut écrire $J_i = \lambda_j I + N$ où N est une matrice nilpotente d'ordre inférieur ou égal à la dimension l du bloc. En utilisant les propriétés de la Proposition 1.7.12, on a

$$\exp(tJ_i) = \exp(t\lambda_j I + tN) = e^{t\lambda_j} \exp(tN).$$

On calcule ensuite

$$\exp(tN) = \sum_{k=0}^{+\infty} \frac{(tN)^k}{k!} = \sum_{k=0}^{l-1} \frac{(tN)^k}{k!}$$

et en procédant comme dans la preuve du Lemme 1.7.1, on a

$$\exp(tN) = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{l-1}}{(l-1)!} \\ 0 & 1 & t & \dots & \frac{t^{l-2}}{(l-2)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & t \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Par conséquent, si $U(t) = \exp(tA)V$ où V est un vecteur de $\mathbb{C}^n$, puisque $\exp(tA) = S\exp(tJ)S^{-1}$, alors chaque u_k est combinaison linéaire des fonctions

$$t^l e^{\lambda_j t}, \quad j \in \{1, \dots, p\}, \, l \in \{0, \dots, \mu_j - 1\}.$$

Ainsi, toute solution s'écrit bien sous la forme (1.13). ∎

Chapitre 2

Algèbre bilinéaire

2.1 Rappels sur l'espace dual

Soit V un espace vectoriel de dimension m sur un champ $\mathbb{K}$. Rappelons que l'espace dual de V est l'ensemble des formes linéaires sur V, c'est-à-dire des applications linéaires $f : V \to \mathbb{K}$. On le note V^*. Il s'agit d'un espace vectoriel pour les opérations

$$\big(f_1 + f_2\big)(x) = f_1(x) + f_2(x)$$

et

$$\big(\lambda f\big)(x) = \lambda f(x).$$

Théorème 2.1.1 *Soit $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V. Pour tout $i \in \{1, \ldots, m\}$, on considère la forme linéaire $b_i^* : V \to \mathbb{K}$ telle que $b_i^*(b_j) = \delta_{ij}$. Alors $\mathcal{B}^* = (b_1^*, \ldots, b_m^*)$ est une base de V^*, appelée la* base duale *de $\mathcal{B}$. En particulier,*

$$\dim V = \dim V^*.$$

Démonstration : Remarquons que b_i^* est l'opérateur de projection selon la $i^{\text{ème}}$ composante dans la base $\mathcal{B}$, c'est-à-dire

$$b_i^*(x) = x_i \quad \text{si} \quad x = \sum_{j=1}^{m} x_j b_j.$$

Montrons tout d'abord que les formes $b_1^*, \ldots, b_m^*$ sont linéairement indépendantes. Supposons que

$$\sum_{i=1}^{m} \lambda_i b_i^* = 0.$$

En évaluant cette relation en b_j, il vient

$$0 = \sum_{i=1}^{m} \lambda_i b_i^*(b_j) = \lambda_j,$$

ce qui suffit.

Montrons à présent que les formes $b_1^*, \ldots, b_m^*$ constituent une partie génératrice de V^*. Soit $f \in V^*$. Pour tout $x \in V$, en décomposant x dans la base $\mathcal{B}$, on obtient

$$f(x) = f\left(\sum_{j=1}^m x_j b_j\right) = \sum_{j=1}^m x_j f(b_j) = \sum_{j=1}^m b_j^*(x) f(b_j).$$

Il s'ensuit que

$$f = \sum_{j=1}^m f(b_j) b_j^*$$

d'où la conclusion. ■

Ainsi, à toute base de V correspond une base de V^*. Nous allons à présent montrer que toute base de V^* permet de construire une base de V.

Lemme 2.1.2 *Soit $x \in V \setminus \{0\}$. Alors il existe $f \in V^*$ tel que $f(x) = 1$.*

Démonstration : Soit $(b_1, \ldots, b_m)$ une base de V. Si $x = \sum_{i=1}^m x_i b_i$, alors il existe un i_0 tel que $x_{i_0} \neq 0$ puisque $x \neq 0$. Il suffit alors de poser $f = \frac{1}{x_{i_0}} b_{i_0}^*$. ■

Proposition 2.1.3 *Toute base de V^* est la base duale d'une unique base de V, appelée* base préduale.

Démonstration : Soit $\mathcal{F} = (\varphi_1, \ldots, \varphi_m)$ une base de V^*. Considérons l'application linéaire

$$\Phi : V \to \mathbb{K}^m : x \mapsto \big(\varphi_1(x), \ldots, \varphi_m(x)\big).$$

Montrons que cette application est injective : supposons qu'il existe $x \in V \setminus \{0\}$ tel que $\Phi(x) = 0$, i.e. $\varphi_i(x) = 0$ pour tout $i \in \{1, \ldots, m\}$. Par le Lemme 2.1.2, il existe $f \in V^*$ tel que $f(x) = 1$. Puisque $\mathcal{F}$ est une base de V^*, on peut décomposer f sous la forme

$$f = \sum_{i=1}^m \alpha_i \varphi_i.$$

En évaluant cette relation en x, on en tire que $f(x) = 0$, ce qui est impossible. Ainsi, $\ker \Phi = \{0\}$ et Φ est injective. Comme il s'agit d'une application linéaire entre espaces de même dimension, Φ est un isomorphime et l'image inverse de toute base de $\mathbb{K}^m$ donne une base de V. Considérons la base canonique $(e_1, \ldots, e_m)$ de $\mathbb{K}^m$ et posons $\mathcal{B} = (\Phi^{-1}(e_1), \ldots, \Phi^{-1}(e_m))$. Pour conclure, il reste à montrer que $\mathcal{F}$ est la base duale de $\mathcal{B}$. C'est évident puisque $\varphi_i(\Phi^{-1}(e_j)) = \delta_{ij}$.

■

Illustrons à présent le calcul de bases préduales.

Exemple 2.1.4 Considérons les trois formes linéaires définies sur $\mathbb{R}^3$ par

$$\varphi_1(x_1, x_2, x_3) = x_1 - x_2 + 2x_3, \quad \varphi_2(x_1, x_2, x_3) = x_2 + 2x_3 \quad \text{et} \quad \varphi_3(x_1, x_2, x_3) = x_3.$$

On vérifie facilement qu'elles sont linéairement indépendantes et qu'elles forment donc une base du dual de $\mathbb{R}^3$. On cherche $b_1, b_2, b_3 \in \mathbb{R}^3$ tels que

$$\varphi_i(b_j) = \delta_{ij}$$

pour tous $i, j \in \{1, 2, 3\}$. Considérons pour cela le système

$$\left\{\begin{array}{rcl} \varphi_1(x_1,x_2,x_3) & = & a \\ \varphi_2(x_1,x_2,x_3) & = & b \\ \varphi_3(x_1,x_2,x_3) & = & c \end{array}\right. \iff \left\{\begin{array}{rcl} x_1 - x_2 + 2x_3 & = & a \\ x_2 + 2x_3 & = & b \\ x_3 & = & c \end{array}\right.$$

$$\iff \left\{\begin{array}{rcl} x_1 & = & a+b-4c \\ x_2 & = & b-2c \\ x_3 & = & c \end{array}\right. .$$

En prenant pour a, b, c les valeurs $1, 0, 0$, on obtient $b_1 = (1, 0, 0) = e_1$. En prenant les valeurs $0, 1, 0$ pour a, b, c, il vient $b_2 = (1, 1, 0) = e_1 + e_2$. Enfin, les valeurs $0, 0, 1$ pour a, b, c donnent $b_3 = (-4, -2, 1) = -4e_1 - 2e_2 + e_3$.

Exemple 2.1.5 Considérons les trois formes linéaires définies sur $\mathbb{R}^3$ par

$$\varphi_1(x_1,x_2,x_3) = x_1 + x_2 + 3x_3, \quad \varphi_2(x_1,x_2,x_3) = x_1 - x_2 + x_3 \quad \text{et} \quad \varphi_3(x_1,x_2,x_3) = x_3.$$

On vérifie facilement qu'elles sont linéairement indépendantes et qu'elles forment donc une base du dual de $\mathbb{R}^3$. On cherche $b_1, b_2, b_3 \in \mathbb{R}^3$ tels que

$$\varphi_i(b_j) = \delta_{ij}$$

pour tous $i, j \in \{1, 2, 3\}$. On résoud le système

$$\left\{\begin{array}{rcl} \varphi_1(x_1,x_2,x_3) & = & a \\ \varphi_2(x_1,x_2,x_3) & = & b \\ \varphi_3(x_1,x_2,x_3) & = & c \end{array}\right. \iff \left\{\begin{array}{rcl} x_1 + x_2 + 3x_3 & = & a \\ x_1 - x_2 + x_3 & = & b \\ x_3 & = & c \end{array}\right.$$

$$\iff \left\{\begin{array}{rcl} x_1 & = & \frac{a+b}{2} - 2c \\ x_2 & = & \frac{a-b}{2} - c \\ x_3 & = & c \end{array}\right. .$$

En prenant pour a, b, c successivement les valeurs $1, 0, 0$, ensuite $0, 1, 0$ et finalement $0, 0, 1$, on obtient $b_1 = (\frac{1}{2}, \frac{1}{2}, 0) = \frac{e_1+e_2}{2}$, $b_2 = (\frac{1}{2}, -\frac{1}{2}, 0) = \frac{e_1-e_2}{2}$ et $b_3 = (-2, -1, 1) = -2e_1 - e_2 + e_3$.

2.2 Formes bilinéaires

On suppose que V est un espace vectoriel sur un champ $\mathbb{K}$.

Définition 2.2.1 Une *forme bilinéaire* est une application $g : V \times V \to \mathbb{K}$ linéaire en chaque variable, c'est-à-dire que pour tous $x, y \in V$, les applications

$$x_g : V \to \mathbb{K} : z \mapsto g(x,z) \quad \text{et} \quad y_d : V \to \mathbb{K} : z \mapsto g(z,y)$$

sont linéaires.

On parle de *forme* car l'application g est à valeurs dans $\mathbb{K}$. Le terme *bilinéaire* vient du fait que l'application possède deux variables et est linéaire en chacune d'elles.

Remarquons que l'ensemble des formes bilinéaires sur V est un espace vectoriel pour les opérations

$$\big(\alpha g\big)(x,y) = \alpha g(x,y)$$

et

$$\big(g+h\big)(x,y) = g(x,y) + h(x,y).$$

Présentons à présent quelques exemples classiques de formes bilinéaires.

Exemple 2.2.2

1. L'archétype des formes linéaires est le produit scalaire euclidien sur $\mathbb{R}^m$ donné par

$$(x,y) \in \mathbb{R}^m \times \mathbb{R}^m \mapsto \sum_{i=1}^{m} x_i y_i.$$

   Le produit scalaire sur $\mathbb{C}^m$ n'est pas une forme bilinéaire car il n'est pas linéaire sur la deuxième variable.

2. Si $A \in \mathbb{K}^m_m$, alors l'application

$$(x,y) \in \mathbb{K}^m \times \mathbb{K}^m \mapsto x^T A y$$

   est une forme bilinéaire.

3. Un exemple de forme bilinéaire sur $C_0([0,1])$ est donné par l'application

$$(f,g) \in C_0([0,1]) \times C_0([0,1]) \mapsto \int_{[0,1]^2} f(s)K(s,t)g(t)ds\,dt,$$

   où $K : [0,1] \times [0,1] \to [0,1]$ est une fonction continue.

4. Si $f \in C^2(\Omega)$ où Ω est un ouvert de $\mathbb{R}^m$, sa différentielle d'ordre 2 en $x \in \Omega$ est la forme bilinéaire

$$d^2 f_x : (h,\ell) \in \mathbb{R}^m \times \mathbb{R}^m \mapsto \sum_{i,j=1}^{m} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} h_i \ell_j.$$

5. Si φ et ψ sont des formes linéaires sur V, alors

$$(x,y) \in V \times V \mapsto \varphi(x)\psi(y)$$

   est une forme bilinéaire.

Nous verrons dans la section suivante que toute forme bilinéaire sur un espace de dimension finie peut en fait se ramener aux formes bilinéaires présentées dans le deuxième exemple.

Proposition 2.2.3 *Soit g une forme bilinéaire sur V.*

1. *Soient $x,y \in V$. Si x ou y est nul, alors $g(x,y) = 0$.*
2. *Soient $n,l \in \mathbb{N}_0$. Pour tous $x^{(1)},\ldots,x^{(n)},y^{(1)},\ldots,y^{(l)} \in V$ et $\alpha_1,\ldots,\alpha_n,\beta_1,\ldots,\beta_l \in$*

$\mathbb{K}$, on a

$$g\left(\sum_{i=1}^{n}\alpha_i x^{(i)}, \sum_{j=1}^{l}\beta_j y^{(j)}\right) = \sum_{i=1}^{n}\sum_{j=1}^{l}\alpha_i\beta_j g(x^{(i)}, y^{(j)}).$$

3. Les applications

$$\cdot_g : V \to V^* : x \mapsto x_g \quad \text{et} \quad \cdot_d : V \to V^* : y \mapsto y_d.$$

sont linéaires.

Démonstration : 1. Si $x = 0$, alors $g(0, y) = y_d(0) = 0$ par linéarité de l'application y_d. De même, si $y = 0$, alors $g(x, 0) = x_g(0) = 0$ par linéarité de x_g.

2. Posons $x = \sum_{i=1}^{n}\alpha_i x^{(i)}$. Par linéarité de l'application x_g, on a

$$g\left(\sum_{i=1}^{n}\alpha_i x^{(i)}, \sum_{j=1}^{l}\beta_j y^{(j)}\right) = g\left(x, \sum_{j=1}^{l}\beta_j y^{(j)}\right) = \sum_{j=1}^{l}\beta_j g\left(x, y^{(j)}\right).$$

Vu la linéarité des applications $(y^{(j)})_d$, on a aussi

$$g(x, y^{(j)}) = g\left(\sum_{i=1}^{n}\alpha_i x^{(i)}, y^{(j)}\right) = \sum_{i=1}^{n}\alpha_i g(x^{(i)}, y^{(j)}),$$

d'où l'égalité annoncée.

3. Il s'agit de simples vérifications. Pour tout $\lambda \in \mathbb{K}$ et tout $x \in V$, on a

$$\cdot_g(\lambda x)(y) = g(\lambda x, y) = \lambda g(x, y) = \lambda x_g(y) = \lambda\big(\cdot_g(x)\big)(y), \quad \forall y \in V,$$

et donc $\cdot_g(\lambda x) = \lambda \cdot_g (x)$. De même, pour tous $x_1, x_2 \in V$, on a

$$\cdot_g(x_1 + x_2)(y) = g(x_1 + x_2, y) = g(x_1, y) + g(x_2, y) = \big(\cdot_g(x_1)\big)(y) + \big(\cdot_g(x_2)\big)(y), \quad \forall y \in V,$$

d'où $\cdot_g(x_1 + x_2) = \cdot_g(x_1) + \cdot_g(x_2)$. ■

Dans ce cours, lorsqu'on n'a qu'une seule forme bilinéaire g à considérer et que le contexte est donc clair, on utilise parfois les notations génériques suivantes : $g(x, y)$ est noté $x \cdot y$ et $g(x, x)$ est noté x^2.

2.3 Représentation matricielle

On suppose dès à présent que l'espace vectoriel V est de dimension finie m. Nous allons définir la matrice associée à une forme bilinéaire dans une base : elle va nous permettre de réduire les calculs à des multiplications matricielles.

Soit $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V. Si $x, y \in V$, on peut les décomposer sous la forme

$$x = \sum_{j=1}^{m} x_j b_j \quad \text{et} \quad y = \sum_{j=1}^{m} y_j b_j\,.$$

La Proposition 2.2.3 nous donne alors

$$g(x,y) = \sum_{i=1}^{m}\sum_{j=1}^{m} x_i y_j g(b_i, b_j).$$

On remarque donc que la forme bilinéaire g est connue dès que les valeurs de g sur tous les couples (b_i, b_j) d'éléments de base sont connues. Cela nous conduit à la définition suivante.

Définition 2.3.1 Soit g une forme bilinéaire sur V et soit $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V. La *matrice associée* à g dans la base $\mathcal{B}$ est la matrice $G \in \mathbb{K}^m_m$ définie par

$$G_{ij} = g(b_i, b_j), \quad \forall i, j \in \{1, \ldots, m\}.$$

On la note également $[g]_{\mathcal{B}}$. On dit aussi que c'est la matrice qui *représente* g dans la base $\mathcal{B}$.

Le raisonnement précédent nous donne le résultat suivant.

Proposition 2.3.2 *Soient $\mathcal{B}$ une base de V et g une forme bilinéaire sur V. Si $x, y \in V$, alors*

$$g(x,y) = [x]^T_{\mathcal{B}} G [y]_{\mathcal{B}}$$

où $[x]_{\mathcal{B}}$, $[y]_{\mathcal{B}}$ sont les vecteurs colonnes des composantes dans $\mathcal{B}$ de x et y respectivement et G est la matrice associée à g dans $\mathcal{B}$.

Démonstration : Il suffit de remarquer que si $x = \sum_{j=1}^{m} x_j b_j$ et $y = \sum_{j=1}^{m} y_j b_j$, alors

$$\sum_{i=1}^{m}\sum_{j=1}^{m} x_i y_j g(b_i, b_j) = \sum_{i=1}^{m} x_i \sum_{j=1}^{m} G_{ij} y_j = \sum_{i=1}^{m} x_i (G[y]_{\mathcal{B}})_i = [x]^T_{\mathcal{B}} G [y]_{\mathcal{B}}.$$

■

Du point de vue local (passage aux composantes dans une base donnée), les formes bilinéaires ne sont donc rien d'autre que des matrices. Réciproquement, nous avons déjà signalé dans l'exemple 2.2.2 que toute matrice définissait une forme bilinéaire. En fait, la bijection présentée ci-dessus entre les formes bilinéaires et les matrices est linéaire comme le montre le résultat suivant. On a donne un isomorphisme entre l'espace des formes bilinéaires sur V et l'espace des matrices $\mathbb{K}^m_m$.

Proposition 2.3.3 *Soient $\alpha, \beta \in \mathbb{K}$ et soient g, h des formes bilinéaires sur V. Alors, on a*

$$[\alpha g + \beta h]_{\mathcal{B}} = \alpha[g]_{\mathcal{B}} + \beta[h]_{\mathcal{B}}\,.$$

Démonstration : C'est évident puisque $\big(\alpha g + \beta h\big)(b_i, b_j) = \alpha g(b_i, b_j) + \beta h(b_i, b_j)$. ■

2.4 Rang et noyau

Regardons ce qu'il se passe lorsque l'on effectue un changement de base. Rappelons que si $\mathcal{B}$ et $\mathcal{C}$ sont deux bases de V, la matrice de changement de base de $\mathcal{B}$ à $\mathcal{C}$ est la matrice dont les colonnes sont les composantes des vecteurs de $\mathcal{B}$ dans la base $\mathcal{C}$.

Proposition 2.4.1 *Soient $\mathcal{B}$ et $\mathcal{C}$ deux bases de V et g une forme bilinéaire sur V. Si S est la matrice de changement de base de $\mathcal{B}$ à $\mathcal{C}$, alors*

$$[g]_{\mathcal{B}} = S^T[g]_{\mathcal{C}}S.$$

Démonstration : Si $\mathcal{B} = (b_1, \ldots, b_m)$, la Proposition 2.3.2 donne

$$\big([g]_{\mathcal{B}}\big)_{ij} = g(b_i, b_j) = [b_i]_{\mathcal{C}}^T[g]_{\mathcal{C}}[b_j]_{\mathcal{C}} = \big(S^T[g]_{\mathcal{C}}S\big)_{ij}$$

puisque par définition, $S = ([b_1]_{\mathcal{C}}, \ldots, [b_m]_{\mathcal{C}})$. ■

Remarque 2.4.2 Notons la différence entre cette formule de changement de base pour une forme bilinéaire et la formule de changement de base pour une application linéaire f, donnée par $S^{-1}[f]_{\mathcal{C}}S$.

Puisque les matrices S et S^T sont inversibles, le rang de la matrice qui représente la forme bilinéaire g ne dépend pas de la base choisie. Cela nous mène à la définition suivante.

Définition 2.4.3 Soit g une forme bilinéaire sur V. Le *rang de* g est le rang de n'importe quelle matrice associée à g dans une base de V. On le note $rg(g)$.

On peut également introduire la notion de noyaux d'une forme bilinéaire.

Définition 2.4.4 Soit g une forme bilinéaire sur V. Le *noyau à gauche* $\ker_g(g)$ de g est le noyau de l'application linéaire $\cdot_g$, c'est-à-dire

$$\ker_g(g) = \{x \in V : x_g = 0\} = \{x \in V : g(x, y) = 0\,,\ \forall y \in V\}$$

où le premier 0 est le zéro de V^* (la fonction nulle) et le deuxième 0 est le zéro de $\mathbb{K}$. Le *noyau à droite* $\ker_d(g)$ de g est le noyau de l'application linéaire $\cdot_d$, c'est-à-dire

$$\ker_d(g) = \{y \in V : y_d = 0\} = \{y \in V : g(x, y) = 0\,,\ \forall x \in V\}.$$

Bien sûr, les noyaux sont des sous-espaces vectoriels de V. A priori, les noyaux à gauche et à droite d'une forme bilinéaire ne coïncident pas. On peut néanmoins montrer qu'ils ont la même dimension. Commençons par remarquer que la matrice qui représente la forme bilinéaire g est liée aux matrices représentant les applications linéaires $\cdot_g$ et $\cdot_d$.

Proposition 2.4.5 *Soient g une forme bilinéaire sur V et G sa matrice associée dans une base $\mathcal{B}$ de V. Si $\mathcal{B}^*$ est la base duale de $\mathcal{B}$, alors les matrices représentants les applications linéaires $\cdot_g$ et $\cdot_d$ dans les bases $\mathcal{B}$ et $\mathcal{B}^*$ sont données par*

$$[\cdot_g]_{\mathcal{B},\mathcal{B}^*} = G^T \quad \textit{et} \quad [\cdot_d]_{\mathcal{B},\mathcal{B}^*} = G.$$

Démonstration : Considérons le cas de l'application linéaire $\cdot_g : V \to V^* : x \mapsto x_g$. Notons $\mathcal{B} = (b_1, \ldots, b_m)$ et $\mathcal{B}^* = (b_1^*, \ldots, b_m^*)$. On a

$$(b_j)_g = \sum_{i=1}^{m} g\big(b_j, b_i\big)b_i^*.$$

Comme $[\cdot_g]_{\mathcal{B},\mathcal{B}^*}$ est la matrice dont la $j^{\text{ème}}$ colonne est formée des composantes de $(b_j)_g$ dans la base $\mathcal{B}^*$, on en tire que

$$\big([\cdot_g]_{\mathcal{B},\mathcal{B}^*}\big)_{ij} = g(b_j, b_i) = G^T_{ij}.$$

On procède de la même manière pour $\cdot_d$. ■

Corollaire 2.4.6 *Pour toute forme bilinéaire g sur V, on a*

$$\dim\big(\ker_g(g)\big) = \dim\big(\ker_d(g)\big) = \dim(V) - \mathrm{rg}(g).$$

Démonstration : Par la proposition précédente, on sait que le rang de l'application linéaire $\cdot_g$ est égal au rang de la forme bilinéaire g, et par définition, les noyaux sont égaux. La conclusion s'ensuit. On procède de même pour $\cdot_d$. ■

Définition 2.4.7 Une forme bilinéaire g sur V est *non-dégénérée* si l'une des conditions équivalentes suivantes est satisfaite :

1. $rg(g) = \dim(V)$,
2. la matrice qui représente g dans n'importe quelle base est inversible,
3. $\ker_g(g) = \{0\}$, c'est-à-dire pour tout $x \in V \setminus \{0\}$, il existe $y \in V$ tel que $g(x, y) \neq 0$,
4. l'application linéaire $\cdot_g : V \to V^*$ est un isomorphisme,
5. $\ker_d(g) = \{0\}$ c'est-à-dire pour tout $y \in V \setminus \{0\}$, il existe $x \in V$ tel que $g(x, y) \neq 0$,
6. l'application linéaire $\cdot_d : V \to V^*$ est un isomorphisme.

Les équivalences entre les points 1, 2, 3 et 5 se déduisent directement du Corollaire 2.4.6. Pour montrer l'équivalence entre les points 3 et 4, il suffit d'utiliser le Théorème de la dimension et le fait que $\dim(V) = \dim(V^*)$. L'équivalence entre les points 5 et 6 s'obtient de manière similaire.

Remarque 2.4.8 Le deuxième point de la définition précédente est équivalent au fait qu'il existe une base dans laquelle la matrice qui représente g est inversible.

2.5 Formes bilinéaires symétriques et formes quadratiques

Le produit scalaire sur $\mathbb{R}^m$ transforme l'espace vectoriel $\mathbb{R}^m$ en l'espace euclidien $\mathbb{R}^m$, en dotant $\mathbb{R}^m$ d'une distance (définie via une norme) donc d'une topologie, prémisses de l'analyse. Le concept de forme bilinéaire permet de généraliser ces notions à n'importe quel espace vectoriel. La théorie se développe agréablement lorsque l'on impose à la forme bilinéaire certaines conditions.

Définition 2.5.1 Soit g une forme bilinéaire sur V. On dit que g est *symétrique* si

$$g(x, y) = g(y, x)\,, \quad \forall x, y \in V.$$

Proposition 2.5.2 *Une forme bilinéaire g est symétrique si et seulement si la matrice représentant g dans n'importe quelle base de V est symétrique.*

Démonstration : Soit G la matrice représentant g dans une base $\mathcal{B} = (b_1, \ldots, b_m)$ donnée. Si g est symétrique, alors

$$G_{ij} = g(b_i, b_j) = g(b_j, b_i) = G_{ji}$$

pour tous $i, j \in \{1, \ldots, m\}$. Réciproquement, si la matrice G est symétrique alors pour tous $x = \sum_{j=1}^m x_j b_j$ et $y = \sum_{j=1}^m y_j b_j$, on a

$$g(x, y) = \sum_{j=1}^m \sum_{i=1}^m x_i y_j G_{ij} = \sum_{i=1}^m \sum_{j=1}^m x_i y_j G_{ji} = g(y, x).$$

■

Si g est une forme bilinéaire symétrique, les applications x_g et x_d coïncident : on les notera simplement $x\cdot$. De même, les applications $\cdot_g$ et $\cdot_d$ coïncident : on les notera $\cdot$. Remarquons que

$$x \cdot y = g(x, y)$$

ce qui justifie les notations introduites précédemment. Néanmoins, les applications g et $\cdot$ ne sont pas définies sur les mêmes espaces ! On utilisera aussi la notation suivante : $g(x, x) = x^2$.

Dans le cas d'une forme bilinéaire symétrique, les noyaux à gauche et à droite coïncident avec le noyau de l'application $\cdot$. On adopte donc la définition suivante.

Définition 2.5.3 Soit g une forme bilinéaire symétrique sur V. Le *noyau* de g est défini par

$$\ker g = \ker \cdot\,.$$

Enfin, des résultats précédents, on tire que la forme g est non-dégénérée si et seulement si l'application $\cdot : V \to V^*$ est un isomorphisme. De plus, les applications g et $\cdot$ sont représentées par la même matrice (dans la base $\mathcal{B}$ pour g et dans les bases $\mathcal{B}$ et $\mathcal{B}^*$ pour $\cdot$).

Dans la suite, nous ferons régulièrement l'hypothèse que le champ $\mathbb{K}$ est de caractéristique différente de 2. Rappelons que la caractéristique d'un champ de neutre 1 est le plus petit $n \in \mathbb{N}_0$ tel que

$$\underbrace{1 + \cdots + 1}_{n \text{ fois}} = 0.$$

Si un tel naturel non-nul n'existe pas, on dit que le champ est de caractéristique nulle.

Proposition 2.5.4 *Supposons que la caractéristique de $\mathbb{K}$ est différente de 2. Soient g et h deux formes bilinéaires symétriques sur V telles que $g(x, x) = h(x, x)$ pour tout $x \in V$. Alors $g = h$.*

Démonstration : Soient $x, y \in V$. On a

$$\begin{aligned} g(x+y, x+y) - g(x, x) - g(y, y) &= g(x, x) + g(x, y) + g(y, x) + g(y, y) - g(x, x) - g(y, y) \\ &= g(x, y) + g(y, x) \\ &= 2g(x, y) \end{aligned}$$

et de même

$$h(x+y, x+y) - h(x, x) - h(y, y) = 2h(x, y).$$

Comme $\mathbb{K}$ est de caractéristique différente de 2, on en tire que $g(x, y) = h(x, y)$. ■

Ce résultat et sa preuve justifient la définition suivante.

Définition 2.5.5 Une application $q : V \to \mathbb{K}$ est une *forme quadratique* sur V s'il existe une forme bilinéaire symétrique $g : V \times V \to \mathbb{K}$ telle que

$$q(x) = g(x, x)$$

pour tout $x \in V$. On dit que q est la forme quadratique *associée à* g. La forme bilinéaire symétrique g est unique : on l'appelle la *forme polaire* de q et on la note g_q.

La forme quadratique associée à une forme bilinéaire est un équivalent du carré de la norme d'un vecteur de $\mathbb{R}^m$, obtenue en prenant le produit scalaire de ce vecteur avec lui-même. Une forme quadratique n'est pas linéaire puisque

$$q(\lambda x) = g(\lambda x, \lambda x) = \lambda^2 g(x, x) = \lambda^2 q(x),$$

ce qui justifie le terme *quadratique*.

Bien sûr, l'ensemble des formes quadratiques sur V forme un espace vectoriel pour les opérations

$$\big(q_1 + q_2\big)(x) = q_1(x) + q_2(x)$$

et

$$\big(\lambda q\big)(x) = \lambda q(x).$$

La preuve de la Proposition 2.5.4 nous a fourni le résultat suivant.

Proposition 2.5.6 (Formule de polarisation) *Supposons que la caractéristique de* $\mathbb{K}$ *est différente de* 2. *Soit* q *une forme quadratique sur* V. *Alors*

$$g_q(x, y) = \frac{1}{2}\big(q(x + y) - q(x) - q(y)\big)$$

pour tous $x, y \in V$.

Il s'agit en fait d'une généralisation de la formule sur $\mathbb{R}$

$$ab = \frac{1}{2}\big((a + b)^2 - a^2 - b^2\big).$$

On tire de la formule de polarisation que l'application qui associe à toute forme quadratique q sa forme polaire g_q est un isomorphisme entre l'espace des formes quadratiques et celui des formes bilinéaires symétriques de V. On appliquera donc aux formes quadratiques tout le vocabulaire introduit pour les formes bilinéaires. En particulier, la matrice qui représente q dans une base est celle qui représente sa forme polaire, le rang d'une forme quadratique est celui de sa forme polaire et on dit qu'une forme quadratique est non-dégénérée quand sa forme polaire l'est.

Remarque 2.5.7 Attention, on pourrait définir le noyau d'une forme quadratique q comme l'ensemble $\{x \in V : q(x) = 0\}$. Remarquez d'une part que ce noyau n'est pas un sous-espace vectoriel (q n'est pas linéaire) et d'autre part, il ne coïncide pas avec celui de sa forme polaire. En effet, considérons la forme quadratique sur $\mathbb{R}^2$ définie par $q(x) = 2x_1x_2$. Alors le vecteur $(1, 0)$ appartient au noyau de q. Néanmoins, la forme polaire de q est donnée par $g(x, y) = x_1y_2 + x_2y_1$ et on vérifie facilement que $(1, 0)$ n'appartient pas à son noyau.

Proposition 2.5.8 (Formule du parallélogramme) *Supposons que la caractéristique de $\mathbb{K}$ est différente de 2. Soit q une forme quadratique sur V. Pour tous $x, y \in V$, on a*

$$q(x+y) + q(x-y) = 2\big(q(x) + q(y)\big).$$

Démonstration : Soit g la forme polaire de q. Par linéarité et symétrie de g, on a

$$\begin{aligned} q(x+y) + q(x-y) &= g(x+y, x+y) + g(x-y, x-y) \\ &= g(x,x) + g(x,y) + g(y,x) + g(y,y) + g(x,x) - g(x,y) - g(y,x) + g(y,y) \\ &= 2g(x,x) + 2g(y,y) \\ &= 2\big(q(x) + q(y)\big). \end{aligned}$$

■

Il s'agit évidemment d'une généralisation de la relation

$$(a+b)^2 + (a-b)^2 = 2(a^2 + b^2).$$

Si $\mathcal{B} = (b_1, \ldots, b_m)$ est une base de V et si G est la matrice associée à q dans cette base, alors

$$q(x) = [x]_{\mathcal{B}}^T G [x]_{\mathcal{B}}.$$

On en tire que

$$q(x) = \sum_{i=1}^{m} G_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq m} G_{ij} x_i x_j$$

pour tout $x \in V$, où $x_1, \ldots, x_m$ sont les composantes de x dans la base $\mathcal{B}$. Une forme quadratique s'exprime donc comme un polynôme homogène du second degré en fonction des variables $x_1, \ldots, x_m$: c'est une combinaison linéaire des monômes x_i^2 et $x_i x_j$.

Exemple 2.5.9 Considérons la matrice symétrique

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

La forme quadratique associée à G est donnée par

$$q(x) = 2x_1^2 - x_3^2 + 2\big(x_1 x_2 + 3x_1 x_3\big) = 2x_1^2 - x_3^2 + 2x_1 x_2 + 6x_1 x_3.$$

On sait que la forme locale d'une forme quadratique est un polynôme homogène de degré 2. Réciproquement, on a le résultat suivant.

Proposition 2.5.10 *Supposons que la caractéristique de $\mathbb{K}$ est différente de 2. Soit $\mathcal{B}$ une base de V et soit Q un polynôme homogène de degré 2 à m variables. Alors l'application*

$$q : V \to \mathbb{K} : x \mapsto Q(x_1, \ldots, x_m)$$

où $(x_1, \ldots, x_m)$ sont les composantes de x dans $\mathcal{B}$, est une forme quadratique.

Démonstration : Par hypothèse, on peut écrire Q sous la forme

$$Q(x_1, \ldots, x_m) = \sum_{i=1}^{m} a_{ii} x_i^2 + \sum_{1 \leq i < j \leq m} a_{ij} x_i x_j.$$

Clairement, on a donc

$$q(x) = [x]_{\mathcal{B}}^T G [x]_{\mathcal{B}}$$

où G est la matrice symétrique définie par

$$G_{ij} = \begin{cases} a_{ii} & \text{si} \quad i = j \\ \frac{1}{2} a_{ij} & \text{si} \quad i < j. \end{cases}$$

Puisque toute matrice symétrique définit une forme bilinéaire symétrique, donc une forme quadratique, on obtient la conclusion. ■

Exemple 2.5.11 Considérons le polynôme homogène de degré 2

$$Q(x_1, x_2, x_3) = x_1^2 + 7x_2^2 + 6x_1x_2 - 2x_1x_3 - 4x_2x_3.$$

La forme quadratique définie par q est associée à la forme bilinéaire donnée par la matrice symétrique

$$G = \begin{pmatrix} 1 & 3 & -1 \\ 3 & 7 & -2 \\ -1 & -2 & 0 \end{pmatrix}.$$

Remarque 2.5.12 Si

$$q(x) = \sum_{i=1}^{m} a_{ii} x_i^2 + \sum_{1 \leq i < j \leq m} a_{ij} x_i x_j,$$

alors il est facile de vérifier que

$$g_q(x, y) = \sum_{i=1}^{m} a_{ii} x_i y_i + \frac{1}{2} \sum_{1 \leq i < j \leq m} a_{ij} (x_i y_j + x_j y_i).$$

On appelle cette règle le *dédoublement des indices.*

Exemple 2.5.13 Reprenons la forme quadratique donnée par

$$q(x) = x_1^2 + 7x_2^2 + 6x_1x_2 - 2x_1x_3 - 4x_2x_3.$$

Sa forme polaire est donnée par

$$g_q(x, y) \quad = \quad x_1y_1 + 7x_2y_2 + 3(x_1y_2 + x_2y_1) - (x_1y_3 + x_3y_1) - 2(x_2y_3 + x_3y_2).$$

On retrouve donc bien la matrice

$$G = \begin{pmatrix} 1 & 3 & -1 \\ 3 & 7 & -2 \\ -1 & -2 & 0 \end{pmatrix}.$$

Terminons cette section par quelques remarques sur le caractère non-dégénéré de la forme polaire.

Lemme 2.5.14 *Si la caractéristique de $\mathbb{K}$ est différente de 2 et si q est une forme quadratique sur V telle que $q(x) \neq 0$ pour tout $x \neq 0$, alors sa forme polaire est non-dégénérée.*

Démonstration : On va montrer la contraposée : supposons que g_q est dégénérée. Alors il existe $x \in V \setminus \{0\}$ tel que $g_q(x, y) = 0$ pour tout $y \in V$. En particulier, $q(x) = g(x, x) = 0$. ■

Remarque 2.5.15 En général, la réciproque de ce résultat est fausse. Par exemple, sur $\mathbb{R}^2$, considérons la forme bilinéaire

$$g(x, y) = x_1 y_1 - x_2 y_2$$

et sa forme quadratique

$$q(x) = x_1^2 - x_2^2.$$

Alors g est non-dégénérée puisque $G = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ est inversible, mais $q\big((1, 1)\big) = 0$.

2.6 Orthogonalité

Les expressions permettant de calculer $g(x, y)$ se simplifient grandement lorsque la base $\mathcal{B}$ est bien choisie : en particulier, il peut être intéressant de se débarrasser des termes croisés. Pour cela, on va introduire la notion d'orthogonalité et s'intéresser à la recherche de bases orthogonales.

Définition 2.6.1 Soit g une forme bilinéaire symétrique sur V et soient $x, y \in V$. On dit que x et y sont *orthogonaux* si $g(x, y) = 0$. Dans ce cas, on écrit $x \perp y$.

Remarquons que, contrairement au produit scalaire classique sur $\mathbb{R}^m$, on peut trouver des formes bilinéaires g pour lesquelles il existe $x \in V$ avec $g(x, x) = 0$. On dit qu'un tel vecteur est *isotrope*. Par exemple, soit g la forme bilinéaire sur $\mathbb{R}^2$ définie par $g(x, y) = x_1 y_2 + x_2 y_1$. Alors $x = (0, 1)$ est tel que $g(x, x) = 0$.

Définition 2.6.2 Soit g une forme bilinéaire symétrique sur V et soient $S, T \subseteq V$. On dit que S et T sont *orthogonaux* si $s \perp t$ pour tous $s \in S, t \in T$. Dans ce cas, on écrit $S \perp T$. L'*orthogonal* de S est défini par

$$S^\perp = \{x \in V : g(x, s) = 0\,,\, \forall s \in S\}.$$

Puisqu'il peut exister des vecteurs isotropes, on n'a pas nécessairement $S \cap S^\perp = \{0\}$. Nous verrons des conditions qui garantissent cette relation.

Proposition 2.6.3 *Soit g une forme bilinéaire symétrique sur V. Pour tous $S, T \subseteq V$, on a*

1. *$S^\perp$ est un sous-espace vectoriel de V,*
2. *si $S \subseteq T$, alors $S^\perp \supseteq T^\perp$,*
3. *$S \subseteq S^{\perp\perp}$,*
4. *$S^\perp = (\rangle S \langle)^\perp$.*

Démonstration : 1. Il s'agit d'une simple vérification qui découle de la bilinéarité.

2. Supposons que $S \subseteq T$ et soit $x \in T^\perp$. Alors $g(x,y) = 0$ pour tout $y \in T$, d'où $g(x,y) = 0$ pour tout $y \in S$, ce qui signifie $x \in S^\perp$.

3. Soit $x \in S$. Alors $g(x,y) = 0$ pour tout $y \in S^\perp$. Donc $x \in S^{\perp\perp}$.

4. Comme $S \subseteq \rangle S \langle$, le point 2 nous donne $S^\perp \supseteq (\rangle S \langle)^\perp$. Pour l'autre inclusion, considérons $x \in S^\perp$. Si $y \in \rangle S \langle$, alors il existe $s_1, \ldots, s_n \in S$ et $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tels que $y = \sum_{j=1}^n \lambda_j s_j$. On en tire que

$$g(x,y) = \sum_{j=1}^n \lambda_j \underbrace{g(x,s_j)}_{=0} = 0.$$

■

Pour espérer avoir l'égalité entre S et $S^{\perp\perp}$, il faut évidemment que S soit un sous-espace vectoriel. Néanmoins, cette condition ne suffit pas comme le montre l'exemple suivant : soit g la forme bilinéaire sur $\mathbb{R}^2$ définie par $g(x,y) = x_1 y_1$ et soit $S = \rangle(1,0)\langle$. Alors on vérifie facilement que $S^\perp = \rangle(0,1)\langle$ et que $(S^\perp)^\perp = \mathbb{R}^2$.

Théorème 2.6.4 *Soit g une forme bilinéaire symétrique sur V et soit W un sous-espace vectoriel de V.*

1. *La forme g est non-dégénérée si et seulement si $V^\perp = \{0\}$.*
2. *Si g est non-dégénérée, alors $\dim V = \dim W + \dim W^\perp$ et $W = W^{\perp\perp}$.*

Démonstration : 1. Par définition, on a $V^\perp = \{0\}$ si et seulement si pour tout $x \in V \setminus \{0\}$, il existe $y \in V$ tel que $g(x,y) \neq 0$, ce qui suffit.

2. Supposons que g est non-dégénérée. Si $x \in V$, l'application

$$x \cdot |_W : W \to \mathbb{K} : y \mapsto g(x,y)$$

est une forme linéaire sur W. On peut donc considérer l'application

$$\varphi : V \to W^* : x \mapsto x \cdot |_W.$$

A nouveau, on vérifie facilement que φ est linéaire. Le Théoréme de la dimension implique que

$$\dim V = \dim \operatorname{im} \varphi + \dim \ker \varphi.$$

Or,

$$\ker \varphi = \big\{x \in V : x \cdot |_W = 0\big\} = \big\{x \in V : g(x,y) = 0, \forall y \in W\big\} = W^\perp.$$

Il reste donc à montrer que $\dim \operatorname{im} \varphi = \dim W$. Puisque $\dim W = \dim W^*$, il faut donc montrer que φ est surjectif. Soit $f \in W^*$. Soit $\mathcal{C}$ une base de W. On peut compléter $\mathcal{C}$ pour obtenir une base $\mathcal{B}$ de V. Alors, on peut prolonger f en une application linéaire

$$f' : V \to \mathbb{K}$$

en posant $f'(c) = f(c)$ pour tout $c \in \mathcal{C}$, et $f'(b) = 0$ pour tout $b \in \mathcal{B} \setminus \mathcal{C}$. Puisque g est non-dégénéré sur V, la Définition 2.4.7 implique que l'application $\cdot : V \to V^*$ est un isomorphisme et il existe donc $x \in V$ tel que $f' = x\cdot$. Dès lors, $f = x\cdot\,|_W \in \operatorname{im}\varphi$ et la conclusion de la première partie s'ensuit.

Pour la deuxième partie, on sait déjà par la Proposition 2.6.3 que $W \subseteq W^{\perp\perp}$. De plus, vu ce qui précède, on a $\dim W^{\perp\perp} = \dim V - \dim W^{\perp} = \dim W$, d'où l'égalité annoncée.

■

Les espaces vectoriels W satisfaisant le dernier point du théorème précédent ont la particularité que le restriction de g à W est non-dégénérée. Cela n'est pas toujours le cas : si g est la forme bilinéaire symétrique sur $\mathbb{R}^2$ définie par $g(x, y) = x_1y_1 - x_2y_2$ et si $W = \rangle(1, 1)\langle$, alors g est non-dégénérée mais sa restriction à W est nulle.

Définition 2.6.5 Soit g une forme bilinéaire symétrique sur V. Un sous-espace vectoriel W est *non-dégénéré* si la restriction

$$g|_W : W \times W \to \mathbb{K}$$

de g à W est non-dégénérée, c'est-à-dire pour tout $x \in W \setminus \{0\}$, il existe $y \in W$ tel que $g(x, y) \neq 0$.

Proposition 2.6.6 *Soit g une forme bilinéaire symétrique sur V et soit W un sous-espace vectoriel de V. Les assertions suivantes sont équivalentes :*

1. *W est non-dégénéré,*
2. *$W \cap W^{\perp} = \{0\}$,*
3. *$W \oplus W^{\perp} = V$.*

Démonstration : Si W est non-dégénée et si $x \in W \cap W^{\perp}$, alors $g(x, y) = 0$ pour tout $y \in W$, d'où $x = 0$. Ainsi, le point 1 implique le point 2. Réciproquement, si $W \cap W^{\perp} = \{0\}$, alors pour tout $x \in W \setminus \{0\}$, on a $x \notin W^{\perp}$ et donc il existe $y \in W$ tel que $g(x, y) \neq 0$. Donc le point 2 implique le point 1.

Triviallement, le point 3 implique le point 2 par définition de la somme directe. Il reste à montrer la réciproque. Supposons que $W \cap W^{\perp} = \{0\}$. Pour montrer que $V = W \oplus W^{\perp}$, il suffit donc de montrer que $\dim V = \dim W + \dim W^{\perp}$. On procède comme dans la preuve du Théorème 2.6.4 et on considère l'application linéaire

$$\varphi|_W : W \to W^* : x \mapsto x \cdot |_W.$$

Remarquons que

$$\ker \varphi|_W = \{x \in W : x \cdot |_W = 0\} = \big\{x \in W : g(x, y) = 0, \forall y \in W\big\} = W \cap W^{\perp} = \{0\}.$$

Par conséquent, l'application linéaire $\varphi|_W$ est injective, et puisque $\dim W = \dim W^*$, le Théorème de la dimension implique que $\varphi|_W$ est surjective. On en tire que l'application $\varphi : V \to W^*$ est également surjective, c'est-à-dire $\operatorname{im}\varphi = W^*$. En utilisant à nouveau le Théorème de la dimension, on trouve

$$\dim V = \dim W^* + \dim \ker \varphi = \dim W + \dim W^{\perp},$$

ce qui suffit. ■

Introduisons à présent la notion de somme orthogonale, qui généralise naturellement la notion de somme directe au cas des espaces vectoriels munis d'une forme bilinéaire symétrique.

Définition 2.6.7 Si $W_1, \ldots, W_n$ sont des sous-espaces vectoriels en position de somme directe qui sont deux à deux orthogonaux, on dit qu'ils sont en *somme orthogonale* et on note

$$W_1 \bigoplus \ldots \bigoplus W_n.$$

La Proposition 2.6.6 se résume en écrivant que W est non-dégénéré si et seulement si

$$V = W \bigoplus W^\perp.$$

Le concept de base dans un espace vectoriel se particularise dans le cas d'un espace vectoriel muni d'une forme bilinéaire symétrique en celui de base orthogonale.

Définition 2.6.8 Soit g une forme bilinéaire symétrique sur V. Une base $\mathcal{B} = (b_1, \ldots, b_m)$ de V est *orthogonale* (pour g) si $b_i \perp b_j$ pour tous $i \neq j$ appartenant à $\{1, \ldots, m\}$.

Un exemple de base orthogonale est donnée par la base canonique de $\mathbb{R}^m$ muni de son produit scalaire usuel. L'existence d'une base orthogonale va de pair avec une décomposition de l'espace en somme orthogonale.

Proposition 2.6.9 *Soit g une forme bilinéaire symétrique sur V. Si $\mathcal{B} = (b_1, \ldots, b_m)$ est une base orthogonale de V, alors*

1. $V = \rangle b_1 \langle \bigoplus \ldots \bigoplus \rangle b_m \langle$,
2. *la matrice représentant g dans $\mathcal{B}$ est* $[g]_{\mathcal{B}} = \mathrm{diag}(b_1^2, \ldots, b_m^2)$,
3. *pour tous $x, y \in V$, on a*
$$g(x, y) = \sum_{i=1}^m b_i^2 x_i y_i \quad \text{et} \quad x^2 = g(x, x) = \sum_{i=1}^m b_i^2 x_i^2,$$
où $x_1, \ldots, x_m$ et $y_1, \ldots, y_m$ sont les composantes de x et y dans $\mathcal{B}$,
4. *si $I = \{i \in \{1, \ldots, m\} : b_i^2 \neq 0\}$, alors* $\ker(g) = \rangle \{b_i : i \notin I\} \langle$ *et* $\mathrm{rg}(g) = \# I$.

Démonstration : 1. On sait que $V = \rangle b_1 \langle \oplus \cdots \oplus \rangle b_m \langle$ et il suffit donc de montrer que les espaces $\rangle b_i \langle$ sont deux à deux orthogonaux. C'est évident car si $x = \alpha b_i$ et $y = \beta b_j$ avec $i \neq j$, on a

$$g(x, y) = \alpha\beta g(b_i, b_j) = 0.$$

2. On a $\left([g]_{\mathcal{B}}\right)_{ij} = g(b_i, b_j) = \delta_{ij} b_i^2$.

3. C'est évident puisque

$$g(x, y) = \sum_{i=1}^m \sum_{j=1}^m x_i y_j g(b_i, b_j) = \sum_{i=1}^m \sum_{j=1}^m x_i y_j \delta_{ij} b_i^2 = \sum_{i=1}^m x_i y_i b_i^2.$$

4. Par le point 2, on sait que le rang de g est donné par $\# I$. Le Corollaire 2.4.6 donne que $\dim \ker g = m - \# I$. Comme il est clair que $b_i \in \ker g$ pour tout $i \notin I$, la conclusion s'ensuit.

■

L'existence de bases orthogonales n'est assurée que si le champ des scalaires $\mathbb{K}$ est de caractéristique différente de 2.

Théorème 2.6.10 *Supposons que la caractéristique de $\mathbb{K}$ est différente de 2. Soit g une forme bilinéaire symétrique sur V. Alors V possède une base orthogonale $\mathcal{B} = (b_1, \ldots, b_m)$.*

Démonstration : On procède par récurrence sur $\dim V$. Le cas de la dimension 1 est trivial. Supposons que le résultat est vérifié pour les espaces de dimension inférieure ou égale à m et montrons-le pour les espaces de dimension $m+1$. On peut supposer que $g \neq 0$, sinon le résultat est trivial. Ainsi, il existe $b_1 \in V$ tel que $b_1^2 \neq 0$. En effet, par la formule de polarisation, si $q = 0$, alors g serait également identiquement nul. Posons $W = \rangle b_1 \langle$. Alors W est non-dégénéré et la Proposition 2.6.6 implique que $V = W \oplus W^\perp$. Bien sûr, on a également $\dim W^\perp = m$. Par induction, $W^\perp$ admet une base orthogonale $(b_2, \ldots, b_{m+1})$. Alors $(b_1, \ldots, b_{m+1})$ est clairement une base orthogonale de V. ■

Par conséquent, quitte à réordonner les éléments de la base orthogonale $\mathcal{B}$, on peut supposer que la matrice qui représente g dans $\mathcal{B}$ a la forme

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & 0 & \ldots & 0 \\ \vdots & \ddots & \lambda_r & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}$$

où $r = \mathrm{rg}(g)$ et où $\lambda_i \neq 0$ pour tout $i \in \{1, \ldots, r\}$.

Regardons ce que donne l'existence de bases orthogonales dans le cas où le champ des scalaires est $\mathbb{C}$.

Théorème 2.6.11 *Soit g une forme bilinéaire symétrique de rang r sur un espace vectoriel complexe V de dimension finie. Alors il existe une base orthogonale $\mathcal{B}$ de V dans laquelle g est représentée par la matrice diagonale*

$$[g]_{\mathcal{B}} = \mathrm{diag}\big(\underbrace{1, \ldots, 1}_{r}, 0, \ldots, 0\big).$$

Démonstration : En utilisant le Théorème 2.6.10, on peut considérer une base orthogonale $\mathcal{B} = (b_1, \ldots, b_m)$ de V. Quitte à permutter les éléments de la base, on peut supposer que $b_i^2 \neq 0$ pour tout $i \leq r$. De plus, quitte à multiplier chaque b_i, $i \leq r$, par un nombre complexe convenable, on peut supposer que $b_i^2 = 1$, ce qui suffit. ■

En particulier, si $G \in \mathbb{C}_m^m$ est une matrice symétrique, il existe une matrice inversible $S \in \mathbb{C}_m^m$ telle que

$$S^T G S = \mathrm{diag}\big(\underbrace{1, \ldots, 1}_{r}, 0, \ldots, 0\big)$$

où r est le rang de la matrice G.

Regardons à présent ce qui se passe lorsque le champ considéré est celui des réels.

Théorème 2.6.12 (Sylvester) *Soit g une forme bilinéaire symétrique de rang r sur un espace vectoriel réel V de dimension finie. Alors il existe une base orthogonale $\mathcal{B}$ de V dans laquelle g est représentée par la matrice diagonale*

$$[g]_{\mathcal{B}} = \operatorname{diag}\big(\underbrace{1,\ldots,1}_{p},\underbrace{-1,\ldots,-1}_{s},0,\ldots,0\big)$$

où $p, s \in \mathbb{N}$ sont tels que $p + s = r$. En particulier, g est non-dégénéré si et seulement si $p + s = m$.

Démonstration : Le Théorème 2.6.10 fournit une base orthogonale $\mathcal{B} = (b_1, \ldots, b_m)$ de V. Ordonnons les éléments de la base en commençant par les b_i tels que $b_i^2 > 0$, ensuite en considérant les b_i tels que $b_i^2 < 0$ et en terminant les b_i tels que $b_i^2 = 0$. Enfin, en divisant chaque b_i tel que $b_i^2 \neq 0$ par $\sqrt{|b_i^2|}$, on obtient la conclusion. ∎

En particulier, si $G \in \mathbb{R}^m_m$ est une matrice symétrique, il existe une matrice inversible $S \in \mathbb{R}^m_m$ telle que

$$S^T G S = \operatorname{diag}\big(\underbrace{1,\ldots,1}_{p},\underbrace{-1,\ldots,-1}_{s},0,\ldots,0\big)$$

avec $p + s$ égal au rang de la matrice G.

Nous allons à présent montrer que les entiers p et s ne dépendent pas de la base choisie pour représenter g sous la forme $\operatorname{diag}\big(1,\ldots,1,-1,\ldots,-1,0,\ldots,0\big)$. Pour cela, travaillons avec la forme quadratique associée à g. Bien sûr, dans la base dont il est question dans le théorème précédent, q prend la forme

$$q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+s}^2.$$

On dit que q s'écrit comme une *somme de carrés* (même si l'expression locale de q contient également des différences).

Définition 2.6.13 Soit q une forme quadratique sur un espace vectoriel réel V et soit W un sous-espace vectoriel de V. On dit que q est *définie positive* sur W si $q(x) > 0$ pour tout $x \in W \setminus \{0\}$. De même, on dit que q est *définie négative* sur W si $q(x) < 0$ pour tout $x \in W \setminus \{0\}$.

Proposition 2.6.14 *Soit q une forme quadratique sur un espace vectoriel réel V. Si $\mathcal{B} = (b_1, \ldots, b_m)$ est une base de V dans laquelle*

$$q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+s}^2,$$

alors

$$p = \max\big\{\dim W : W \text{ sous-espace vectoriel de } V,\ q \text{ définie positive sur } W\big\}$$

et

$$s = \max\big\{\dim W : W \text{ sous-espace vectoriel de } V,\ q \text{ définie négative sur } W\big\}.$$

Par conséquent, le couple (p, s) ne dépend pas de la base choisie pour représenter q en somme de carrés. On l'appelle la signature de la forme *q.*

Démonstration : Il suffit de démontrer le résultat pour p, celui pour s s'obtiendra alors en considérant la forme $-q$. Soit W un sous-espace vectoriel de V sur lequel q est définie positive. Soit également T le sous-espace vectoriel de V engendré par les vecteurs $b_{p+1}, \ldots, b_m$. Il est clair que $q(x) \leq 0$ pour tout $x \in T$. Par conséquent, puisque $q(x) > 0$ pour tout $x \in W \setminus \{0\}$, on en tire que

$$W \cap T = \{0\}$$

et donc que $\dim W = \dim(W + T) - \dim T \leq m - (m - p) = p$. De plus, il est clair que q est définie positive sur l'espace de dimension p engendré par $b_1, \ldots, b_p$. La conclusion s'ensuit. ■

En général, on dira qu'une forme bilinéaire symétrique g sur V est *définie positive* si sa forme quadratique associée q est définie positive sur V, et est *définie négative* si q est définie négative sur V. Par le Lemme 2.5.14 ou en utilisant la Proposition 2.6.14, toute forme bilinéaire symétrique définie positive (resp. négative) est non-dégénérée.

2.7 Algorithme de Gauss

Dans cette section, nous allons décrire un algorithme basé sur l'utilisation de la forme quadratique associée à une forme bilinéaire permettant de construire une base orthogonale et de trouver la signature de la forme. L'idée est la suivante : on écrit la forme quadratique comme combinaison linéaire de carrés d'applications linéaires indépendantes. Ensuite, on cherche la base préduale correspondante.

Il est facile de construire des formes quadratiques à partir de formes linéaires : il suffit d'en prendre le carré. De manière générale, on a le résultat suivant.

Proposition 2.7.1 *Soient $n \geq 1$, $\varphi_1, \ldots, \varphi_n \in V^*$ et $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$. Alors*

$$q = \sum_{i=1}^{n} \alpha_i \varphi_i^2$$

est une forme quadratique sur V et sa forme polaire est donnée par

$$g_q(x, y) = \sum_{i=1}^{n} \alpha_i \varphi_i(x) \varphi_i(y).$$

Démonstration : Remarquons que pour tout $i \in \{1, \ldots, n\}$, l'application

$$(x, y) \in V \times V \mapsto \varphi_i(x)\varphi_i(y)$$

est une forme bilinéaire symétrique. La forme quadratique associée est φ^2. La linéarité de l'application $q \mapsto g_q$ permet de conclure. ■

Toute combinaison linéaire de carrés de formes linéaires est donc une forme quadratique. La proposition suivante permet de décomposer toute forme quadratique q en une combinaison linéaire de carrés de formes linéaires. Par le Théorème 2.6.10, on sait qu'il existe une base orthogonale $\mathcal{B} = (b_1, \ldots, b_m)$ pour q. Alors, si q est de rang r, on a

$$q = \sum_{i=1}^{r} b_i^2 \left(b_i^*\right)^2$$

puisque $q(x) = \sum_{i=1}^{r} x_i^2 b_i^2$. On a donc décomposé q en une combinaison linéaire de carrés de formes linéaires indépendantes. De plus, ces formes linéaires sont des éléments de la base duale de la base orthogonale $\mathcal{B}$. Afin de trouver une base orthogonale, nous allons effectuer la démarche inverse : nous allons décomposer q en une combinaison linéaire de carrés d'élements de V^*, et en prendre ensuite la base pré-duale. Une des méthodes pour trouver une telle décomposition repose sur une construction algorithmique. C'est la méthode de réduction de Gauss.

Proposition 2.7.2 (Algorithme de Gauss) *Soit V un espace vectoriel sur un champ $\mathbb{K}$ de caractéristique différente de 2. Si q est une forme quadratique sur V, on peut trouver de manière algorithmique un entier $n \geq 1$, des formes $\varphi_1, \ldots, \varphi_n \in V^*$ linéairement indépendantes et des scalaires $\alpha_1, \ldots, \alpha_n \in \mathbb{K} \setminus \{0\}$ tels que*

$$q = \sum_{i=1}^{n} \alpha_i \varphi_i^2.$$

Démonstration : Le cas où $q = 0$ est trivial. On peut donc supposer que $q \neq 0$. On procède par récurrence sur la dimension de V. Si $\dim V = 1$, il est clair que toute forme quadratique est de la forme $\alpha\varphi^2$. Supposons donc le résultat vérifié dans tout espace de dimension strictement inférieure à m et prouvons-le pour $\dim V = m$. Soit $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V et

$$q(x) = \sum_{i=1}^{m} G_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq m} G_{ij} x_i x_j$$

l'expression de q dans cette base. Deux cas peuvent se produire.

1. Il existe $i \in \{1, \ldots, m\}$ tel que $G_{ii} \neq 0$. Quitte à permuter les éléments de la base, on peut supposer que $G_{11} \neq 0$. Dans ce cas, il vient

$$\begin{aligned} q(x) &= G_{11} x_1^2 + 2 \sum_{j=2}^{m} G_{1j} x_1 x_j + \sum_{i=2}^{m} G_{ii} x_i^2 + 2 \sum_{2 \leq i < j \leq m} G_{ij} x_i x_j \\ &= G_{11} \left(x_1 + \sum_{j=2}^{m} \frac{G_{1j}}{G_{11}} x_j \right)^2 + \sum_{i=2}^{m} \widetilde{G}_{ii} x_i^2 + 2 \sum_{2 \leq i < j \leq m} \widetilde{G}_{ij} x_i x_j \end{aligned}$$

pour des $\widetilde{G}_{ij}$ bien choisis. Notre hypothèse de récurrence appliquée à la forme quadratique définie sur $\rangle b_2, \ldots, b_m \langle$ par

$$\widetilde{q}(x) = \sum_{i=2}^{m} \widetilde{G}_{ii} x_i^2 + 2 \sum_{2 \leq i < j \leq m} \widetilde{G}_{ij} x_i x_j$$

nous donne des formes linéaires indépendantes $\varphi_2, \ldots, \varphi_n$ et des scalaires non-nuls $\alpha_2, \ldots, \alpha_n$ tels que

$$\widetilde{q} = \alpha_2 \varphi_2^2 + \cdots + \alpha_n \varphi_n^2.$$

Considérons la forme linéaire

$$\varphi_1(x) = x_1 + \sum_{j=2}^{m} \frac{G_{1j}}{G_{11}} x_j$$

et posons $\alpha_1 = G_{11}$. Par construction, il est clair que

$$q = \sum_{i=1}^{p} \alpha_i \varphi_i^2.$$

Il reste à montrer que les formes $\varphi_1, \ldots, \varphi_n$ sont linéairement indépendantes. Supposons donc que

$$\sum_{i=1}^{n} \beta_i \varphi_i = 0.$$

On a $0 = \sum_{i=1}^{n} \beta_i \varphi_i(b_1) = \beta_1$ et il s'ensuit que

$$\sum_{i=2}^{n} \beta_i \varphi_i = 0.$$

Les formes $\varphi_2, \ldots, \varphi_n$ étant linéairement indépendantes, on obtient $\beta_2 = \cdots = \beta_n = 0$.

2. Si $G_{ii} = 0$ pour tout $i \in \{1, \ldots, m\}$, alors il existe $i \neq j$ tels que $G_{ij} \neq 0$. Quitte à permuter les éléments de la base, on peut supposer que $G_{12} \neq 0$. Il vient alors

$$\begin{aligned}
q(x) &= 2 \sum_{1 \leq i < j \leq m} G_{ij} x_i x_j \\
&= 2G_{12} x_1 x_2 + 2 \sum_{j=3}^{m} G_{1j} x_1 x_j + 2 \sum_{j=3}^{m} G_{2j} x_2 x_j + 2 \sum_{3 \leq i < j \leq m} G_{ij} x_i x_j \\
&= 2G_{12} \left(x_1 + \sum_{j=3}^{m} \frac{G_{2j}}{G_{12}} x_j \right) \left(x_2 + \sum_{j=3}^{m} \frac{G_{1j}}{G_{12}} x_j \right) + \sum_{j=3}^{m} \widetilde{G}_{jj} x_j^2 + 2 \sum_{3 \leq i < j \leq m} \widetilde{G}_{ij} x_i x_j
\end{aligned}$$

pour des $\widetilde{G}_{ij}$ bien choisis. Considérons la forme quadratique

$$\widetilde{q}(x) = \sum_{j=3}^{m} \widetilde{G}_{jj} x_j^2 + 2 \sum_{3 \leq i < j \leq m} \widetilde{G}_{ij} x_i x_j$$

sur l'espace $\rangle b_3, \ldots, b_m \langle$. Par hypothèse de récurrence, il existe alors des formes linéaires indépendantes $\varphi_3, \ldots, \varphi_n$ et des scalaires non-nuls $\alpha_3, \ldots, \alpha_n$ tels que

$$\widetilde{q} = \alpha_3 \varphi_3^2 + \cdots + \alpha_n \varphi_n^2.$$

Posons également $\alpha_1 = \frac{G_{12}}{2}$, $\alpha_2 = \frac{-G_{12}}{2}$,

$$\varphi_1(x) = x_1 + x_2 + \sum_{j=3}^{m} \frac{G_{2j} + G_{1j}}{G_{12}} x_j$$

et

$$\varphi_2(x) = x_1 - x_2 + \sum_{j=3}^{m} \frac{G_{2j} - G_{1j}}{G_{12}} x_j.$$

Puisque

$$ab = \frac{1}{4}\left((a+b)^2 - (a-b)^2\right),$$

on a

$$\left(x_1 + \sum_{j=3}^{m} \frac{G_{2j}}{G_{12}} x_j\right)\left(x_2 + \sum_{j=3}^{m} \frac{G_{1j}}{G_{12}} x_j\right) = \frac{1}{4}\left(\varphi_1^2(x) - \varphi_2^2(x)\right).$$

Il s'ensuit que

$$q = \alpha_1 \varphi_1^2 + \cdots + \alpha_n \varphi_n^2.$$

Pour conclure, il reste à montrer que les formes $\varphi_1, \ldots, \varphi_n$ sont linéairement indépendantes. Supposons donc que

$$\sum_{i=1}^{n} \beta_i \varphi_i = 0.$$

On a d'une part $0 = \sum_{i=1}^{n} \beta_i \varphi_i(b_1) = \beta_1 + \beta_2$ et d'autre part $0 = \sum_{i=1}^{n} \beta_i \varphi_i(b_2) = \beta_1 - \beta_2$. D'où $\beta_1 = \beta_2 = 0$. On procède alors comme dans le premier cas en utilisant l'indépendance linéaire de $\varphi_3, \ldots, \varphi_n$.

■

Exemple 2.7.3 Considérons la forme quadratique q donnée dans une base (b_1, b_2, b_3) par

$$q(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 8x_3^2 - 2x_1x_2 + 4x_1x_3.$$

Comme il y a un terme en x_1^2, nous sommes dans le premier cas de l'algorithme. On a successivement

$$\begin{aligned} q(x_1, x_2, x_3) &= (x_1 - x_2 + 2x_3)^2 - x_2^2 - 4x_3^2 + 4x_2x_3 + 2x_2^2 + 8x_3^2 \\ &= (x_1 - x_2 + 2x_3)^2 + 4x_2x_3 + x_2^2 + 4x_3^2 \\ &= (x_1 - x_2 + 2x_3)^2 + (x_2 + 2x_3)^2 \end{aligned}$$

et on pose $\varphi_1(x) = x_1 - x_2 + 2x_3$ et $\varphi_2(x) = x_2 + 2x_3$.

Exemple 2.7.4 Considérons la forme quadratique q donnée dans une base (b_1, b_2, b_3) par

$$q(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + 2x_2x_3.$$

Comme il n'y pas pas de termes de la forme x_i^2, on est dans le deuxième cas. On écrit

$$\begin{aligned} q(x_1, x_2, x_3) &= (x_1 + 2x_3)(x_2 + x_3) - 2x_3^2 \\ &= \frac{1}{4}\left((x_1 + 2x_3 + x_2 + x_3)^2 - (x_1 + 2x_3 - x_2 - x_3)^2\right) - 2x_3^2 \\ &= \frac{1}{4}(x_1 + x_2 + 3x_3)^2 - \frac{1}{4}(x_1 - x_2 + x_3)^2 - 2x_3^2. \end{aligned}$$

Il suffit alors de poser $\varphi_1(x) = x_1 + x_2 + 3x_3$, $\varphi_2(x) = x_1 - x_2 + x_3$ et $\varphi_3(x) = x_3$.

Expliquons à présent comment la Proposition 2.7.2 permet de trouver une base orthogonale. Grâce à l'algorithme de Gauss, on a obtenu des formes linéairement indépendantes $\varphi_1, \ldots, \varphi_n \in V^*$ et des scalaires $\alpha_1, \ldots, \alpha_n \in \mathbb{K} \setminus \{0\}$ tels que

$$q = \sum_{i=1}^{n} \alpha_i \varphi_i^2.$$

Complétons la partie libre $(\varphi_1, \ldots, \varphi_n)$ en une base $\mathcal{F} = (\varphi_1, \ldots, \varphi_n, \varphi_{n+1}, \ldots, \varphi_m)$ de V^*. Notons $\mathcal{B} = (b_1, \ldots, b_m)$ sa base préduale. On a donc $\varphi_i(B_j) = \delta_{ij}$ et si $x = \sum_{j=1}^{m} x_j b_j$, il vient

$$q(x) = \sum_{i=1}^{n} \alpha_i \left(\varphi_i(\sum_{j=1}^{m} x_j b_j) \right)^2 = \sum_{i=1}^{n} \alpha_i x_i^2.$$

Dans le cas où $\mathbb{K} = \mathbb{R}$, on obtient la décomposition en somme de carrés en réordonnant la base $\mathcal{B}$ et en divisant les vecteurs de base par des constantes adéquates.

Exemple 2.7.5 Reprenons l'exemple 2.7.3. On peut compléter les formes linéairement indépendantes $\varphi_1(x) = x_1 - x_2 + 2x_3$ et $\varphi_2(x) = x_2 + 2x_3$ par $\varphi_3(x) = x_3$ pour obtenir une base du dual. On cherche $\widetilde{b}_1, \widetilde{b}_2, \widetilde{b}_3$ tels que $\varphi_i(\widetilde{b}_j) = \delta_{ij}$. En procédant comme dans l'exemple 2.1.4, il suffit de prendre $\widetilde{b}_1 = b_1$, $\widetilde{b}_2 = b_1 + b_2$ et $\widetilde{b}_3 = -4b_1 - 2b_2 + b_3$. Dans cette base,

$$q(x) = \widetilde{x}_1^2 + \widetilde{x}_2^2.$$

Exemple 2.7.6 Reprenons l'exemple 2.7.4. En procédant comme dans l'exemple 2.1.5, on pose $\widetilde{b}_1 = \frac{1}{2}(b_1 + b_2)$, $\widetilde{b}_2 = \frac{1}{2}(b_1 - b_2)$ et $\widetilde{b}_3 = b_3$. Dans cette base,

$$q(x) = \widetilde{x}_1^2 - \widetilde{x}_2^2 - \widetilde{x}_3^2.$$

2.8 Espaces pré-hilbertiens réels

Soit V un espace vectoriel réel. Dans le cas où la forme bilinéaire symétrique est définie positive, on sait qu'elle est en particulier non-dégénérée. On parle de produit scalaire.

Définition 2.8.1 Soit V un espace vectoriel réel. Un *produit scalaire (réel)* sur V est une forme bilinéaire symétrique définie positive sur V. En général, on note le produit scalaire $\langle\ ,\ \rangle$. Un espace vectoriel réel V muni d'un produit scalaire est appelé un espace *pré-hilbertien (réel)*.

Lorsqu'on parlera d'un espace pré-hilbertien, on supposera donc implicitement disposer d'un produit scalaire.

Remarque 2.8.2 Dans le cas où V est de dimension finie, on parle également d'espace euclidien.

Exemple 2.8.3

1. Le produit scalaire euclidien sur $\mathbb{R}^m$ est évidemment un produit scalaire.

2. La forme bilinéaire définie par

$$(f,g) \in C_0([a,b],\mathbb{R}) \times C_0([a,b],\mathbb{R}) \mapsto \int_a^b f(x)g(x)dx$$

est un produit scalaire sur l'espace des fonctions continues sur $[a,b]$ et à valeurs réelles (rappelons que l'intégrale entre a et b d'une fonction continue positive est nulle si et seulement si f est identiquement nulle sur $[a,b]$).

Les propriétés du produit scalaire permettent d'introduire la notion de norme d'un vecteur. Comme le produit scalaire est défini positif, on a $\langle x,x\rangle \geq 0$ pour tout $x \in V$ et on peut donc en prendre la racine carrée.

Définition 2.8.4 Soit V un espace pré-hilbertien. Pour tout $x \in V$, la *norme* de x est définie par

$$\|x\| = \sqrt{\langle x,x\rangle}.$$

Par définition d'un produit scalaire, on a $\|x\| \geq 0$ et $\|x\| = 0$ si et seulement si $x = 0$.

Proposition 2.8.5 (Inégalité de Cauchy-Schwarz) *Soit V un espace pré-hilbertien. Alors pour tous $x, y \in V$, on a*

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

et on a l'égalité si et seulement si x et y sont linéairement dépendants.

Démonstration : Le résultat est évident si $x = 0$ ou $y = 0$. On peut donc supposer que $x, y \in V \setminus \{0\}$. Pour tout $\lambda \in \mathbb{R}$, on a

$$\begin{aligned} \|x+\lambda y\|^2 = \langle x+\lambda y, x+\lambda y\rangle &= \langle x,x\rangle + \langle x,\lambda y\rangle + \langle \lambda y,x\rangle + \langle \lambda y,\lambda y\rangle \\ &= \|x\|^2 + 2\lambda\langle x,y\rangle + \lambda^2\|y\|^2. \end{aligned}$$

C'est une fonction quadratique de λ qui est toujours positive ou nulle. Son discriminant est donc plus petit ou égal à 0, c'est-à-dire

$$4(\langle x,y\rangle)^2 - 4\|x\|^2\|y\|^2 \leq 0,$$

ce qui nous donne

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

De plus, on a l'égalité si et seulement s'il existe $\lambda \in \mathbb{R}$ tel que $\|x+\lambda y\|^2 = 0$, c'est-à-dire si et seulement s'il existe $\lambda \in \mathbb{R}$ tel que $x + \lambda y = 0$. ■

Nous pouvons à présent montrer que l'application $\|\cdot\|$ est bien une norme, c'est-à-dire satisfait les propriétés suivantes.

Proposition 2.8.6 *Soit V un espace pré-hilbertien. Alors*

1. *$\|x\| \geq 0$ pour tout $x \in V$ et $\|x\| = 0$ si et seulement si $x = 0$,*
2. *$\|\lambda x\| = |\lambda|\|x\|$ pour tout $\lambda \in \mathbb{R}$ et tout $x \in V$,*
3. *$\|x+y\| \leq \|x\| + \|y\|$ pour tous $x, y \in V$.*

Démonstration : Les deux premiers points sont immédiats. Pour le troisième, en utilisant l'inégalité de Cauchy-Schwarz, on a

$$\|x+y\|^2 = \|x\|^2+2\langle x,y\rangle+\|y\|^2 \leq \|x\|^2+2|\langle x,y\rangle|+\|y\|^2 \leq \|x\|^2+2\|x\|\,\|y\|+\|y\|^2 = (\|x\|+\|y\|)^2.$$

Comme $\|x+y\| \geq 0$ et $\|x\|+\|y\| \geq 0$, on peut prendre la racine carrée des deux membres et obtenir la conclusion. ■

Remarque 2.8.7 Toute norme, et donc tout produit scalaire, sur un espace vectoriel V permet de définir une distance. Celle-ci est donnée par

$$d(x,y) = \|x-y\|\,, \quad \forall x,y \in V.$$

C'est une distance au sens où

1. $d(x,y) \geq 0$ pour tous $x,y \in V$ et $d(x,y) = 0$ si et seulement si $x = y$,
2. $d(x,y) = d(y,x)$ pour tous $x,y \in V$,
3. $d(x,y) \leq d(x,z) + d(z,y)$ pour tous $x,y,z \in V$.

Proposition 2.8.8 (Théorème de Pythagore) *Soit V un espace pré-hilbertien. Considérons des vecteurs $x_1,\ldots,x_n \in V$ deux à deux orthogonaux. Alors, on a*

$$\Big\|\sum_{j=1}^n x_j\Big\|^2 = \sum_{j=1}^n \|x_j\|^2.$$

Démonstration : C'est immédiat puisque

$$\Big\|\sum_{j=1}^n x_j\Big\|^2 = \Big\langle \sum_{j=1}^n x_j, \sum_{j=1}^n x_j \Big\rangle = \sum_{j=1}^n \sum_{i=1}^n \underbrace{\langle x_j, x_i\rangle}_{=0 \text{ si } i\neq j} = \sum_{j=1}^n \|x_j\|^2.$$

■

Proposition 2.8.9 *Soit V un espace pré-hilbertien. Si $x_1,\ldots,x_n \in V$ sont des vecteurs non-nuls deux à deux orthogonaux, alors ils sont linéairement indépendants.*

Démonstration : Soient $\lambda_1,\ldots,\lambda_n \in \mathbb{R}$ tels que

$$\sum_{j=1}^n \lambda_j x_j = 0.$$

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

$$0 = \Big\langle \sum_{j=1}^n \lambda_j x_j, x_i \Big\rangle = \sum_{j=1}^n \lambda_j \underbrace{\langle x_j, x_i\rangle}_{=0 \text{ si } i\neq j} = \lambda_i \|x_i\|^2$$

et puisque $\|x_i\| \neq 0$, on obtient la conclusion. ■

Dans le cas d'un espace préhilbertien, la notion de base orthogonale se particularise en base orthonormée.

Définition 2.8.10 Soient V un espace préhilbertien et $\mathcal{B}$ une base de V. On dit que la base $\mathcal{B}$ est *orthonormée* si $\langle b_i, b_j\rangle = \delta_{ij}$ pour tous $b_i, b_j \in \mathcal{B}$.

Le Théorème 2.6.12 et la Proposition 2.6.14 nous fournissent directement l'existence de bases orthonormées.

Proposition 2.8.11 *Soit V un espace pré-hilbertien de dimension finie. Alors V possède une base orthonormée.*

Lorsque l'on dispose d'une base orthonormée d'un espace pré-hilbertien, il est aisé de calculer les composantes d'un vecteur dans cette base. Elles sont simplement données par les produits scalaire du vecteur avec les vecteurs de base, comme le montre le résultat suivant.

Proposition 2.8.12 *Soit V un espace pré-hilbertien et soit $\mathcal{B} = (b_1, \dots, b_m)$ une base orthonormée de V. Pour tout $x, y \in V$, on a*

1. $x = \sum_{j=1}^{m} \langle x, b_j\rangle b_j$,
2. $\|x\|^2 = \sum_{j=1}^{m} \left(\langle x, b_j\rangle\right)^2$,
3. $\langle x, y\rangle = \sum_{j=1}^{m} \langle x, b_j\rangle\langle y, b_j\rangle$.

Démonstration : 1. Si $x = \sum_{j=1}^{m} x_j b_j$, alors on a $\langle x, b_i\rangle = x_i$.

2. On sait que

$$\|x\|^2 = \left\langle \sum_{j=1}^{m} x_j b_j, \sum_{j=1}^{m} x_j b_j \right\rangle = \sum_{j=1}^{m}\sum_{i=1}^{m} x_j x_i \langle b_j, b_i\rangle = \sum_{j=1}^{m} x_j^2 = \sum_{j=1}^{m} \left(\langle x, b_j\rangle\right)^2$$

par le point 1.

3. On procède comme précédemment et on a

$$\langle x, y\rangle = \sum_{j=1}^{m} x_j y_j = \sum_{j=1}^{m} \langle x, b_j\rangle\langle y, b_j\rangle.$$

■

Comme un produit scalaire est une forme bilinéaire définie positive, tout sous-espace vectoriel W de V est non-dégénéré. En particulier, lorsque V est de dimension finie, la Proposition 2.6.6 implique que

$$V = W \oplus W^{\perp}$$

et tout vecteur $x \in V$ se décompose de manière unique sous la forme

$$x = w + y \quad \text{avec} \quad w \in W \text{ et } y \in W^{\perp}.$$

Définition 2.8.13 Soit V un espace pré-hilbertien de dimension finie et W un sous-espace vectoriel de V. Pour tout $x = w + y \in V$ avec $w \in W$ et $y \in W^{\perp}$, on pose

$$P_W(x) = w.$$

Le vecteur $P_W(x) \in W$ est appelé la *projection orthogonale* de x sur W.

Montrons à présent que la projection orthogonale de x sur W est le point de W le plus proche de x.

Proposition 2.8.14 *Soit V un espace pré-hilbertien de dimension finie et W un sous-espace vectoriel de V.*

1. *L'application $P_W : V \to W$ est linéaire.*
2. *Pour tout $x \in V$, on a $x - P_W(x) \in W^\perp$.*
3. *Si $(b_1, \ldots, b_n)$ est une base orthonormée de W, alors*
$$P_W(x) = \sum_{j=1}^{n} \langle x, b_j \rangle b_j.$$
4. *Pour tout $x \in V$, le vecteur $P_W(x)$ est l'unique vecteur de W réalisant le minimum de la distance de x à W, c'est-à-dire*
$$\|x - P_W(x)\| \leq \|x - y\|$$
pour tout $y \in W$, et on a l'égalité si et seulement si $y = P_W(x)$.

Démonstration : Les deux premiers points sont immédiats. Pour le troisième point, en décomposant $P_W(x)$ dans la base $(b_1, \ldots, b_n)$ de W, on peut écrire
$$P_W(x) = \sum_{j=1}^{n} \lambda_j b_j.$$
Puisque $x - P_W(x) \in W^\perp$, on sait que pour tout $i \in \{1, \ldots, n\}$, $\langle x - P_W(x), b_i \rangle = 0$, c'est-à-dire
$$0 = \langle x, b_i \rangle - \left\langle \sum_{j=1}^{n} \lambda_j b_j, b_i \right\rangle = \langle x, b_i \rangle - \lambda_i,$$
où on a utilisé le fait que la base est orthonormée. Ainsi $\lambda_i = \langle x, b_i \rangle$, ce qui suffit.

Pour le quatrième point, on remarque tout d'abord que si $y \in W$, alors
$$x - y = \underbrace{\big(x - P_W(x)\big)}_{\in W^\perp} + \underbrace{\big(P_W(x) - y\big)}_{\in W}.$$
Comme les vecteurs $x - P_W(x)$ et $P_W(x) - y$ sont orthogonaux, le Théorème 2.8.8 donne
$$\|x - y\|^2 = \|x - P_W(x)\|^2 + \|P_W(x) - y\|^2 \geq \|x - P_W(x)\|^2.$$
De plus, on a l'égalité si et seulement si $\|P_W(x) - y\| = 0$, c'est-à-dire si et seulement si $y = P_W(x)$. ■

Terminons cette section par la présentation du procédé d'*orthogonalisation de Gram-Schmidt*, qui permet de construire algorithmiquement une base orthonormée $\widetilde{\mathcal{B}} = (\widetilde{b}_1, \ldots, \widetilde{b}_m)$ d'un espace pré-hilbertien V à partir d'une base $\mathcal{B} = (b_1, \ldots, b_m)$ de V.

1. On pose $\widetilde{b}_1 = \frac{b_1}{\|b_1\|}$. Alors $\widetilde{b}_1$ est bien de norme 1 et l'espace engendré par $\widetilde{b}_1$ est le même que celui engendré par b_1.
2. Supposons que $\widetilde{b}_1, \ldots, \widetilde{b}_{k-1}$ ont été construits et que le sous-espace W engendré par les vecteurs $\widetilde{b}_1, \ldots, \widetilde{b}_{k-1}$ est égal au sous-espace engendré par $b_1, \ldots, b_{k-1}$. On veut construire un vecteur $\widetilde{b}_k$ orthogonal à $\widetilde{b}_1, \ldots, \widetilde{b}_{k-1}$, donc à W. Par la Proposition 2.8.14, on sait que $b_k - P_W(b_k) \in W^\perp$. On pose donc
$$v_k = b_k - P_W(b_k) = b_k - \sum_{i=1}^{k-1} \langle b_k, \widetilde{b}_i \rangle \widetilde{b}_i.$$
Remarquons que $v_k \neq 0$ car sinon, on aurait $b_k \in W$, ce qui impliquerait que $b_1, \ldots, b_k$ sont linéairement dépendants.
3. Pour normer ce vecteur, il suffit de poser
$$\widetilde{b}_k = \frac{v_k}{\|v_k\|}.$$
4. On remarque que les vecteurs $\widetilde{b}_1, \ldots, \widetilde{b}_k$ engendrent bien le même espace que $b_1, \ldots, b_k$.
5. A la fin de la construction, on a obtenu des vecteurs $\widetilde{b}_1, \ldots, \widetilde{b}_m$ deux à deux orthogonaux, donc linéairement indépendants par la Proposition 2.8.9. Ils forment donc une base de V.

Exemple 2.8.15 On considère l'espace $\mathbb{R}_2[x]$ des polynômes de degré inférieur ou égal à 2 muni du produit scalaire
$$\langle P, Q \rangle = \int_0^1 P(x)Q(x)dx.$$
Une base de $\mathbb{R}_2[x]$ est donnée par les polynômes $P_1(x) = 1$, $P_2(x) = x$ et $P_3(x) = x^2$. On vérifie facilement que cette base n'est pas orthonormée. Appliquons le procédé d'orthonormalisation de Gram-Schmidt. Puisque $\|P_1\|^2 = \int_0^1 1^2 dx = 1$, on pose $\widetilde{P}_1(x) = 1$. Pour construire $\widetilde{P}_2$, on calcule
$$\langle P_2, \widetilde{P}_1 \rangle = \int_0^1 x\,dx = \frac{1}{2}$$
et
$$\|P_2 - \langle P_2, \widetilde{P}_1 \rangle \widetilde{P}_1\|^2 = \int_0^1 \big(x - \frac{1}{2}\big)^2 dx = \frac{1}{12}.$$
Ainsi
$$\widetilde{P}_2(x) = \sqrt{12}(x - \frac{1}{2}) = \sqrt{3}(2x - 1).$$
Enfin, afin de construire $\widetilde{P}_3$, on calcule tout d'abord
$$\langle P_3, \widetilde{P}_1 \rangle = \int_0^1 x^2 dx = \frac{1}{3} \quad \text{et} \quad \langle P_3, \widetilde{P}_2 \rangle = \int_0^1 \sqrt{3}(2x^3 - x^2)dx = \frac{\sqrt{3}}{6}.$$
On a $P_3(x) - \langle P_3, \widetilde{P}_1 \rangle \widetilde{P}_1(x) - \langle P_3, \widetilde{P}_2 \rangle \widetilde{P}_2(x) = x^2 - \frac{1}{3} - \frac{3}{6}(2x - 1) = x^2 - x + \frac{1}{6}$ et donc
$$\|P_3 - \langle P_3, \widetilde{P}_1 \rangle \widetilde{P}_1 - \langle P_3, \widetilde{P}_2 \rangle \widetilde{P}_2\|^2 = \int_0^1 \big(x^2 - x + \frac{1}{6}\big)^2 dx = \frac{1}{180}.$$
On en tire finalement que
$$\widetilde{P}_3(x) = \sqrt{180}\big(x^2 - x + \frac{1}{6}\big) = \sqrt{5}(6x^2 - 6x + 1).$$

2.9 Formes hermitiennes et espaces pré-hilbertiens complexes

Dans de nombreux cas, on a besoin d'utiliser des espaces vectoriels complexes. Néanmoins, les résultats présentés dans la section précédente ne se généralisent pas directement car aucune forme bilinéaire symétrique sur un espace vectoriel complexe ne peut être définie positive. En effet, si

$$g(x,x) > 0,$$

alors

$$g(ix,ix) = i^2 g(x,x) = -g(x,x) < 0.$$

Néanmoins, remarquons que pour tout complexe z, on a $z\overline{z} = |z|^2 \geq 0$. C'est la raison pour laquelle on travaille avec le produit scalaire sur $\mathbb{C}^m$ défini par

$$(x,y) \in \mathbb{C}^m \times \mathbb{C}^m \mapsto \sum_{j=1}^{m} x_j \overline{y_j}.$$

Comme mentionné auparavant, il ne s'agit pas d'une forme bilinéaire.

Définition 2.9.1 Soit V un espace vectoriel complexe. Une application $f : V \to \mathbb{C}$ est *antilinéaire* si pour tous $\lambda, \mu \in \mathbb{C}$ et tous $x, y \in V$, on a

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

Définition 2.9.2 Soit V un espace vectoriel complexe. Une *forme bilinéaire gauche* sur V est une application $h : V \times V \to \mathbb{C}$ linéaire en la première variable et antilinéaire en la deuxième variable. Si, de plus, elle satisfait $h(x,y) = \overline{h(y,x)}$, on dit qu'elle est *hermitienne*.

Remarquons que si h est une forme hermitienne, alors $h(x,x) = \overline{h(x,x)}$, c'est-à-dire $h(x,x)$ est réel pour tout $x \in V$.

Définition 2.9.3 Soit V un espace vectoriel complexe et h une forme hermitienne sur V. On dit que la forme h est *définie positive* si $h(x,x) > 0$ pour tout $x \in V \setminus \{0\}$. Un *produit scalaire complexe* sur V est une forme hermitienne définie positive. A nouveau, on note en général le produit scalaire $\langle\,,\rangle$. Un espace vectoriel complexe muni d'un produit scalaire complexe est appelé un *espace pré-hilbertien complexe*.

Exemple 2.9.4

1. Le produit scalaire sur $\mathbb{C}^m$ est un produit scalaire complexe.
2. La forme bilinéaire définie par

$$(f,g) \in C_0([a,b],\mathbb{C}) \times C_0([a,b],\mathbb{C}) \mapsto \int_a^b f(x)\overline{g(x)}dx$$

   est un produit scalaire complexe sur l'espace des fonctions continues sur $[a,b]$ et à valeurs complexes.

Tout ce que nous avons présenté pour les espaces pré-hilbertiens réels se généralise aisément pour les espaces pré-hilbertiens complexes.

Chapitre 3

Algèbre multilinéaire

3.1 Formes multilinéaires

Soit V un espace vectoriel sur un champ $\mathbb{K}$ et soit r un naturel plus grand que 1. On note

$$V^r = \underbrace{V \times \cdots \times V}_{r \text{ fois}}$$

l'ensemble des r-uples d'éléments de V. Un élément de V^r est donc un r-uple $(x_1, \ldots, x_r)$ où $x_i \in V$ pour tout $i \in \{1, \ldots, r\}$.

Définition 3.1.1 Une *forme r-multilinéaire* (ou r-forme) est une application $f : V^r \to \mathbb{K}$ linéaire en chaque variable, c'est-à-dire pour tout $i \in \{1, \ldots, r\}$, tous $x_1, \ldots, x_r, y_i \in V$ et tout $\alpha \in \mathbb{K}$

$$\begin{aligned} f(x_1, \ldots, x_{i-1}, x_i + y_i, x_{i+1}, \ldots, x_r) = f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_r) \\ + f(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_r) \end{aligned}$$

et

$$f(x_1, \ldots, x_{i-1}, \alpha x_i, x_{i+1}, \ldots, x_r) = \alpha f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_r).$$

Bien sûr, l'ensemble des formes r-multilinéaires est un espace vectoriel pour les opérations

$$\big(f_1 + f_2\big)(x_1, \ldots, x_r) = f_1(x_1, \ldots, x_r) + f_2(x_1, \ldots, x_r)$$

et

$$\big(\lambda f\big)(x_1, \ldots, x_r) = \lambda f(x_1, \ldots, x_r).$$

On sait que le produit de deux formes linéaires $\psi, \varphi \in V^*$ défini par

$$g : (x, y) \in V \times V \mapsto \psi(x)\varphi(y)$$

est une forme bilinéaire. De manière semblable, on peut définir le produit de r formes linéaires $\varphi_1, \ldots, \varphi_r \in V^*$.

Définition 3.1.2 Soient $\varphi_1, \ldots, \varphi_r \in V^*$. Le *produit tensoriel* de $\varphi_1, \ldots, \varphi_r$ est l'application définie par

$$(x_1, \ldots, x_r) \mapsto \varphi_1(x_1) \ldots \varphi_r(x_r).$$

On note cette application $\varphi_1 \otimes \cdots \otimes \varphi_r$.

Il est immédiat de montrer que le produit tensoriel de r formes linéaires est une forme r-multilinéaire. Nous allons montrer que toute forme r-multilinéaire peut se décomposer en combinaison linéaire de produits tensoriels. Pour cela, comme dans le cas des formes bilinéaires, on considère le résultat suivant.

Proposition 3.1.3 *Soit f une forme r-multilinéaire sur V.*

1. *Soient $x_1, \ldots, x_r \in V$. Si un des vecteurs $x_1, \ldots, x_r$ est nul, alors $f(x_1, \ldots, x_r) = 0$.*
2. *Soit $(I_k)_{k\in\{1,\ldots,r\}}$ une famille de r ensembles finis d'indices. Si pour tout $k \in \{1, \ldots, r\}$, $(u_{k,j})_{j\in I_k}$ est une famille d'éléments de V et $(\alpha_{k,j})_{j\in I_k}$ une famille de scalaires, alors*

$$f\left(\sum_{j_1\in I_1} \alpha_{1,j_1} u_{1,j_1}, \ldots, \sum_{j_r\in I_r} \alpha_{r,j_r} u_{r,j_r}\right) = \sum_{j_1\in I_1} \ldots \sum_{j_r\in I_r} \alpha_{1,j_1} \ldots \alpha_{r,j_r} f(u_{1,j_1}, \ldots, u_{r,j_r}).$$

En particulier, si V est de dimension finie m et si $\mathcal{B} = (b_1, \ldots, b_m)$ est une base de V, alors f est entièrement déterminée par les valeurs qu'elle prend sur toute famille de r vecteurs parmi $(b_1, \ldots, b_m)$, c'est-à-dire par $f(b_{i_1}, \ldots, b_{i_r})$ où $b_{i_1}, \ldots, b_{i_r} \in \mathcal{B}$ pour tous $i_1, \ldots, i_r \in \{1, \ldots, m\}$. Ainsi, si $(x_1, \ldots, x_r) \in V^r$ et si pour tout $i \in \{1, \ldots, r\}$, on a

$$x_i = \sum_{j=1}^{m} x_{i,j} b_j,$$

alors

$$\begin{aligned} f(x_1, \ldots, x_r) &= \sum_{j_1=1}^{m} \ldots \sum_{j_r=1}^{m} x_{1,j_1} \ldots x_{r,j_r} f(b_{j_1}, \ldots, b_{j_r}) \\ &= \sum_{(j_1,\ldots,j_r)\in\{1,\ldots,m\}^r} x_{1,j_1} \ldots x_{r,j_r} f(b_{j_1}, \ldots, b_{j_r}) \\ &= \sum_{(j_1,\ldots,j_r)\in\{1,\ldots,m\}^r} b_{j_1}^*(x_1) \ldots b_{j_r}^*(x_r) f(b_{j_1}, \ldots, b_{j_r}). \end{aligned}$$

Lemme 3.1.4 *Soit $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V. Les formes r-multilinéaires*

$$b_{j_1}^* \otimes \cdots \otimes b_{j_r}^*, \quad (j_1, \ldots, j_r) \in \{1, \ldots, m\}^r,$$

forment une base de l'espace des formes r-multilinéaires.

Démonstration : Montrons tout d'abord qu'elles forment une partie génératrice. Soit f une forme r-multilinéaire alternée. Alors vu ce qui précède, pour tout r-uple $(x_1, \ldots, x_r) \in V^r$, on a

$$\begin{aligned} f(x_1, \ldots, x_r) &= \sum_{(j_1,\ldots,j_r)\in\{1,\ldots,m\}^r} b_{j_1}^*(x_1) \ldots b_{j_r}^*(x_r) f(b_{j_1}, \ldots, b_{j_r}) \\ &= \sum_{(j_1,\ldots,j_r)\in\{1,\ldots,m\}^r} \left(b_{j_1}^* \otimes \cdots \otimes b_{j_r}^*\right)(x_1, \ldots, x_r) f(b_{j_1}, \ldots, b_{j_r}), \end{aligned}$$

c'est-à-dire
$$f = \sum_{(j_1,\ldots,j_r)\in\{1,\ldots,m\}^r} f(b_{j_1},\ldots,b_{j_r}) b^*_{j_1}\otimes\cdots\otimes b^*_{j_r}.$$
De plus, supposons que
$$\sum_{(j_1,\ldots,j_r)\in\{1,\ldots,m\}^r} \alpha_{j_1,\ldots,j_r} b^*_{j_1}\otimes\cdots\otimes b^*_{j_r} = 0.$$
Soit $(j'_1,\ldots,j'_r)\in\{1,\ldots,m\}^r$. Alors on a
$$\begin{aligned} 0 &= \sum_{(j_1,\ldots,j_r)\in\{1,\ldots,m\}^r} \alpha_{j_1,\ldots,j_r}\big(b^*_{j_1}\otimes\cdots\otimes b^*_{j_r}\big)(b_{j'_1},\ldots,b_{j'_r}) \\ &= \sum_{(j_1,\ldots,j_r)\in\{1,\ldots,m\}^r} \alpha_{j_1,\ldots,j_r} \underbrace{b^*_{j_1}(b_{j'_1})}_{=\delta_{j_1 j'_1}}\cdots\underbrace{b^*_{j_r}(b_{j'_r})}_{=\delta_{j_r j'_r}} \\ &= \alpha_{j'_1,\ldots,j'_r}, \end{aligned}$$
ce qui suffit pour obtenir l'indépendance linéaire. ∎

Le résultat précédent s'étend aisément pour une base quelconque de V^*.

Théorème 3.1.5 *Soit $(\varphi_1,\ldots,\varphi_m)$ une base de V^*. Les formes r-multilinéaires*
$$\varphi_{j_1}\otimes\cdots\otimes\varphi_{j_r}\,,\quad (j_1,\ldots,j_r)\in\{1,\ldots,m\}^r,$$
forment une base de l'espace des formes r-multilinéaires sur V.

Démonstration : Comme $(\varphi_1,\ldots,\varphi_m)$ est une base de V^*, il suffit de considérer sa base pré-duale $(b_1,\ldots,b_m)$ et d'appliquer le Lemme 3.1.4. ∎

3.2 Formes multilinéaires alternées

On va s'intéresser à présent à une classe particulière de formes multilinéaires.

Définition 3.2.1 On dit qu'une forme r-multilinéaire f est *alternée* lorsqu'elle satisfait la condition
$$f(x_1,\ldots,x_r)=0$$
dès qu'il existe $i\neq j$ tel que $x_i = x_j$.

Il est direct de vérifier que l'ensemble des formes r-multilinéaires alternées est un sous-espace vectoriel de l'espace des formes r-multilinéaires.

Exemple 3.2.2 Soit $V=\mathbb{R}^m$. Alors l'application
$$\det : (x_1,\ldots,x_m)\in V^m \mapsto \det(x_1,\ldots,x_m)$$
où $(x_1,\ldots,x_m)$ est la matrice dont la $i^{\text{ème}}$ colonne est donnée par le vecteur x_i, est une forme m-multilinéaire alternée. Plus généralement, si $r\le m$ et si $i_1,\ldots,i_r\in\{1,\ldots,m\}$, l'application
$$(x_1,\ldots,x_r)\in V^r\mapsto \det\Big((x_{i_k,j})_{j,k\in\{1,\ldots,r\}}\Big)$$

est une r-forme multilinéaire alternée. Elle associe aux vecteurs $x_1, \dots, x_r$ le déterminant de la matrice obtenue en choisissant les r lignes correspondant à $i_1, \dots, i_r$ dans la matrice dont la $i^{\text{ème}}$ colonne est donnée par le vecteur x_i.

La proposition suivante montre que les formes multilinéaires alternées sont les formes antisymétriques dans le cas d'un champ de caractéristique différente de 2.

Proposition 3.2.3 *Soit V un espace vectoriel sur un champ de caractéristique différente de 2. Une forme r-multilinéaire f sur V est alternée si et seulement si pour tous $x_1, \dots, x_r \in V$ et tous $1 \leq i < j \leq r$,*

$$f(x_1, \dots, x_i, \dots, x_j, \dots, x_r) = -f(x_1, \dots, x_j, \dots, x_i, \dots, x_r).$$

Démonstration : Supposons que f est alternée. Par multilinéarité, on a que la relation

$$f(x_1, \dots, x_i, \dots, x_j, \dots, x_r) = -f(x_1, \dots, x_j, \dots, x_i, \dots, x_r)$$

est équivalente à

$$f(2x_1, \dots, x_i + x_j, \dots, x_i + x_j, \dots, 2x_r) = 0.$$

Puisque la forme est alternée, cette dernière relation est satisfaite et la conclusion s'ensuit.

Réciproquement, supposons que $x_i = x_j$ et montrons que $f(x_1, \dots, x_i, \dots, x_j, \dots, x_r) = 0$. D'une part, vu l'hypothèse, on a

$$f(x_1, \dots, x_i, \dots, x_j, \dots, x_r) = -f(x_1, \dots, x_j, \dots, x_i, \dots, x_r).$$

D'autre part, on a également

$$f(x_1, \dots, x_i, \dots, x_j, \dots, x_r) = f(x_1, \dots, x_j, \dots, x_i, \dots, x_r)$$

puisque les vecteurs x_i et x_j sont égaux. On en tire que

$$2f(x_1, \dots, x_i, \dots, x_j, \dots, x_r) = 0,$$

d'où la conclusion. ■

Proposition 3.2.4 *Soient f une forme r-multilinéaire alternée sur V et $x_1, \dots, x_r \in V$. Si les vecteurs $x_1, \dots, x_r$ sont linéairement dépendants, alors $f(x_1, \dots, x_r) = 0$.*

Démonstration : Soit $j \in \{1, \dots, r\}$ tel que x_j est combinaison linéaire des autres vecteurs. Ainsi, il existe des scalaires α_i, $i \in \{1, \dots, r\} \setminus \{j\}$, tels que

$$x_j = \sum_{i \neq j} \alpha_i x_i.$$

En utilisant la multilinéarité de f, il vient que

$$\begin{aligned} f(x_1, \dots, x_j, \dots, x_r) &= f\left(x_1, \dots, \sum_{i \neq j} \alpha_i x_i, \dots, x_r\right) \\ &= \sum_{i \neq j} \alpha_i f(x_1, \dots, x_i, \dots, x_r) \\ &= 0 \end{aligned}$$

puisque pour tout $i \neq j$, la famille de vecteurs $(x_1, \ldots, x_i, \ldots, x_p)$ contient deux vecteurs égaux, à savoir le $i^{\text{ème}}$ et le $j^{\text{ème}}$. ■

Les deux résultats précédents nous permettent directement d'obtenir les propriétés suivantes.

Corollaire 3.2.5 *Soit V un espace vectoriel sur un champ de caractéristique différente de 2. Soient f une forme r-multilinéaire alternée sur V et $x_1, \ldots, x_r \in V$.*

1. *Si σ est une permutation de $\{1, \ldots, r\}$, alors*

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(r)}) = \operatorname{sign}(\sigma) f(x_1, \ldots, x_r).$$

2. *Si on ajoute à un des vecteurs de la famille $(x_1, \ldots, x_r)$ une combinaison linéaire des autres, la valeur que prend f reste inchangée, i.e. pour tout $j \in \{1, \ldots, r\}$ et tous scalaires α_i, $i \neq j$, on a*

$$f\left(x_1, \ldots, x_j + \sum_{i \neq j} \alpha_i x_i, \ldots, x_i, \ldots, x_r\right) = f(x_1, \ldots, x_r).$$

Démonstration : Pour le premier point, on sait que si σ est une permutation, alors il existe des transpositions $\tau_1, \ldots, \tau_q$ telles que $\sigma = \tau_q \ldots \tau_1$ et de plus, $\operatorname{sign}(\sigma) = (-1)^q$. Par conséquent, pour passer de la famille $(x_1, \ldots, x_r)$ à la famille $(x_{\sigma(1)}, \ldots, x_{\sigma(r)})$, il suffit d'appliquer successivement chacune des transpositions. Comme une transposition permute simplement deux éléments, et puisque la forme est alternée, la Proposition 3.2.3 donne successivement

$$\begin{aligned}
f(x_{\sigma(1)}, \ldots, x_{\sigma(r)}) &= f(x_{\tau_q \ldots \tau_1(1)}, \ldots, x_{\tau_q \ldots \tau_1(r)}) \\
&= -f(x_{\tau_{q-1} \ldots \tau_1(1)}, \ldots, x_{\tau_{q-1} \ldots \tau_1(r)}) \\
&= f(x_{\tau_{q-2} \ldots \tau_1(1)}, \ldots, x_{\tau_{q-2} \ldots \tau_1(r)}) \\
&= \ldots \\
&= (-1)^{q-1} f(x_{\tau_1(1)}, \ldots, x_{\tau_1(r)}) \\
&= (-1)^q f(x_1, \ldots, x_r).
\end{aligned}$$

Pour le deuxième point, en utilisant la multilinéarité de f, on a

$$\begin{aligned}
f\left(x_1, \ldots, x_j + \sum_{i \neq j} \alpha_i x_i, \ldots, x_r\right) &= f(x_1, \ldots, x_i, \ldots, x_r) + \sum_{i \neq j} \alpha_i f(x_1, \ldots, x_i, \ldots, x_r) \\
&= f(x_1, \ldots, x_j, \ldots, x_r)
\end{aligned}$$

puisque les termes de la somme sur les $i \neq j$ sont nuls (la famille de vecteurs $(x_1, \ldots, x_i, \ldots, x_p)$ contient deux vecteurs égaux). ■

Les résultats précédents montrent que les formes multilinéaires alternées partagent de nombreuses propriétés des déterminants. Montrons à présent qu'en fait, les seules formes m-multilinéaires d'un espace de dimension m sont données par les multiples du déterminant. Commençons par illustrer ce qui se passe en dimension 2.

Exemple 3.2.6 Soit V un espace vectoriel de dimension 2 sur un champ de caractéristique différente de 2. Considérons le cas d'une forme bilinéaire alternée f sur V. Si $\mathcal{B} = (b_1, b_2)$ est une base de V et si x et y sont deux éléments de V de composantes x_1, x_2 et y_1, y_2 respectivement dans $\mathcal{B}$, alors on a

$$\begin{aligned} f(x_1b_1 + x_2b_2, y_1b_1 + y_2b_2) &= x_1y_1 \underbrace{f(b_1, b_1)}_{=0} + x_1y_2 f(b_1, b_2) + x_2y_1 f(b_2, b_1) + x_2y_2 \underbrace{f(b_2, b_2)}_{=0} \\ &= x_1y_2 f(b_1, b_2) + x_2y_1 f(b_2, b_1) \\ &= (x_1y_2 - x_2y_1) f(b_1, b_2). \end{aligned}$$

On voit donc apparaître la quantité $x_1y_2 - x_2y_1$ qui n'est rien d'autre que le déterminant de la matrice dont les colonnes sont les composantes de x et y dans $\mathcal{B}$, i.e.

$$f(x_1b_1 + x_2b_2, y_1b_1 + y_2b_2) = \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} f(b_1, b_2).$$

Définition 3.2.7 Soient V un espace vectoriel de dimension m et $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V. On définit l'application $\det^{\mathcal{B}}$ via

$$det^{\mathcal{B}} : (x_1, \ldots, x_m) \in V^m \mapsto \det\big([x_1]_{\mathcal{B}}, \ldots, [x_m]_{\mathcal{B}}\big),$$

c'est-à-dire l'application qui associe à chaque famille $(x_1, \ldots, x_m) \in V^m$ le déterminant de la matrice dont la $j^{\text{ème}}$ colonne est donnée par les composantes de x_j dans la base $\mathcal{B}$.

Ainsi, si pour tout $j \in \{1, \ldots, m\}$, on a

$$x_j = \sum_{i=1}^{m} x_{j,i} b_i,$$

on prend le déterminant de la matrice de dimension $m \times m$ dont l'élément (i, j) est donné par $x_{j,i}$. On a donc

$$\det^{\mathcal{B}}(x_1, \ldots, x_m) = \sum_{\sigma \in \mathcal{S}_m} \operatorname{sign}(\sigma) x_{1,\sigma(1)} \ldots x_{m,\sigma(m)}$$

où $\mathcal{S}_m$ désigne l'ensemble des permutations de $\{1, \ldots, m\}$.

Proposition 3.2.8 *Soit V un espace vectoriel de dimension m sur un champ de caractéristique différente de 2 et soit $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V. Si f est une application m-multilinéaire alternée sur V, alors*

$$f(x_1, \ldots, x_m) = f(b_1, \ldots, b_m) \det^{\mathcal{B}}(x_1, \ldots, x_m).$$

En particulier, il existe un scalaire α tel que $f = \alpha \det^{\mathcal{B}}$.

Démonstration : Remarquons que puisque f est une forme m-multilinéaire alternée f, la représentation obtenue précédemment dans une base $\mathcal{B} = (b_1, \ldots, b_m)$ se simplifie en

$$\begin{aligned} f(x_1, \ldots, x_m) &= \sum_{(j_1, \ldots, j_m) \in \{1, \ldots, m\}^m} x_{1,j_1} \ldots x_{m,j_m} f(b_{j_1}, \ldots, b_{j_m}) \\ &= \sum_{\sigma \in \mathcal{S}_m} x_{1,\sigma(1)} \ldots x_{m,\sigma(m)} f(b_{\sigma(1)}, \ldots, b_{\sigma(m)}). \end{aligned}$$

En effet, se donner $(j_1, \ldots, j_m) \in \{1, \ldots, m\}^r$ revient à se donner l'application σ qui associe à tout $k \in \{1, \ldots, m\}$ l'élément j_k. De plus, si σ n'est pas une permutation de $\{1, \ldots, m\}$, deux composantes au moins de $(b_{\sigma(1)}, \ldots, b_{\sigma(m)})$ sont égales. En utilisant le Corollaire 3.2.5, il vient

$$\begin{aligned}\sum_{\sigma \in \mathcal{S}_m} x_{1,\sigma(1)} \ldots x_{m,\sigma(m)} f(b_{\sigma(1)}, \ldots, b_{\sigma(m)}) &= \sum_{\sigma \in \mathcal{S}_m} x_{1,\sigma(1)} \ldots x_{m,\sigma(m)} \mathrm{sign}(\sigma) f(b_1, \ldots, b_m) \\ &= f(b_1, \ldots, b_m) \left(\sum_{\sigma \in \mathcal{S}_m} \mathrm{sign}(\sigma) x_{1,\sigma(1)} \ldots x_{m,\sigma(m)} \right) \\ &= f(b_1, \ldots, b_m) \det{}^{\mathcal{B}}(x_1, \ldots, x_m)\end{aligned}$$

par définition du déterminant. ■

Par conséquent, l'espace des formes m-multilinéaires alternées sur un espace vectoriel de dimension m est un espace de dimension 1. Etant donnée une base $\mathcal{B} = (b_1, \ldots, b_m)$ de V, $\det^{\mathcal{B}}$ est l'unique forme m-multilinéaire alternée f sur V telle que $f(b_1, \ldots, b_m) = 1$. Cette propriété est parfois utilisée comme définition même du déterminant.

Remarque 3.2.9 Si $\mathcal{B}'$ est une autre base de V, alors par la Proposition 3.2.8, on a

$$\det{}^{\mathcal{B}'} = \det{}^{\mathcal{B}'}(b_1, \ldots, b_m) \det{}^{\mathcal{B}}.$$

Etudions à présent ce qui se passe si $r \neq m$.

Proposition 3.2.10 *Si $r > m$, toute forme r-multilinéaire alternée est identiquement nulle.*

Démonstration : Si $x_1, \ldots, x_r \in V$, alors ces vecteurs sont nécessairement linéairement dépendants puisque $r > m = \dim V$. La conclusion est alors obtenue par la Proposition 3.2.4. ■

Il reste à étudier le cas $r < m$. Commençons par introduire quelques notations. Soit A une matrice de dimension $m \times r$ à coefficients dans $\mathbb{K}$ et soit E un sous-ensemble de $\{1, \ldots, m\}$ constitué de r éléments. On peut prendre le déterminant de la sous-matrice extraite de A constituée des éléments $a_{i,j}$ où $i \in E$. On note $\det_E(A)$ ce déterminant. On a alors

$$\det{}_E(A) = \sum_{\nu \in \mathcal{S}_r} \mathrm{sign}(\nu) a_{i_1,\nu(1)} \ldots a_{i_r,\nu(r)} = \sum_{\sigma \in \mathcal{S}_r} \mathrm{sign}(\sigma) a_{i_{\sigma(1)},1} \ldots a_{i_{\sigma(r)},r}$$

si $E = \{i_1, \ldots, i_r\}$.

Définition 3.2.11 Soient V un espace vectoriel de dimension m, $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V et E un sous-ensemble de $\{1, \ldots, m\}$ constitué de r éléments. On définit l'application $\det_E^{\mathcal{B}}$ via

$$\det{}_E^{\mathcal{B}} : (x_1, \ldots, x_r) \in V^r \mapsto \det{}_E([x_1]_{\mathcal{B}}, \ldots, [x_r]_{\mathcal{B}}).$$

Ainsi, si pour tout $j \in \{1, \ldots, m\}$, on a

$$x_j = \sum_{i=1}^{m} x_{j,i} b_i,$$

on prend le déterminant d'une sous-matrice de la matrice de dimension $r \times m$ dont l'élément (i, j) est donné par $x_{j,i}$. Si $E = \{i_1, \ldots, i_r\}$, on a donc

$$\det{}_E^{\mathcal{B}}(x_1, \ldots, x_r) = \sum_{\nu \in \mathcal{S}_r} \operatorname{sign}(\nu) x_{\nu(1),i_1} \ldots x_{\nu(r),i_r} = \sum_{\sigma \in \mathcal{S}_r} \operatorname{sign}(\sigma) x_{1,i_{\sigma(1)}} \ldots x_{r,i_{\sigma(r)}}.$$

Cette application définit évidemment une forme r-multilinéaire alternée.

Nous allons voir que ces applications engendrent l'espace des formes r-multilinéaires alternées.

Proposition 3.2.12 *Soit V un espace vectoriel de dimension m sur un champ de caractéristique différente de 2, soit f une forme r-multilinéaire alternée sur V, avec $r < m$, et soit $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V. Pour tous $x_1, \ldots, x_r \in V$, on a*

$$f(x_1, \ldots, x_r) = \sum_{E \subseteq \{1,\ldots,m\},\, \#E = r} \det{}_E^{\mathcal{B}}(x_1, \ldots, x_r) f(\mathbf{b}_E)$$

où le r-uple $\mathbf{b}_E \in V^r$ est défini par $\mathbf{b}_E = (b_{j_1}, \ldots, b_{j_r})$ si $j_1 < \cdots < j_r$ sont les éléments ordonnés de E.

Démonstration : A nouveau, il s'agit d'une simple réécriture de la représentation obtenue pour f dans la base $\mathcal{B}$. Rappelons que

$$f(x_1, \ldots, x_r) = \sum_{(j_1,\ldots,j_r) \in \{1,\ldots,m\}^r} x_{1,j_1} \ldots x_{r,j_r} f(b_{j_1}, \ldots, b_{j_r}),$$

et puisque f est alternée, on peut restreindre la somme aux $(j_1, \ldots, j_r) \in \{1, \ldots, m\}^r$ tels que $j_k \neq j_l$ si $k \neq l$, puisque sinon on a $f(b_{j_1}, \ldots, b_{j_r}) = 0$. Cela revient à dire qu'on regarde les $(j_1, \ldots, j_r) \in \{1, \ldots, m\}^r$ tels que $\#\{j_1, \ldots, j_r\} = r$. De plus, parmi tous ces r-uples, on peut regrouper ceux qui donnent le même ensemble $\{j_1, \ldots, j_r\}$ contenant r éléments. Cela revient à considérer tous les r-uples $(j_{\sigma(1)}, \ldots, j_{\sigma(r)})$ pour $\sigma \in \mathcal{S}_r$. On a donc

$$f(x_1, \ldots, x_r) = \sum_{(j_1,\ldots,j_r) \in \{1,\ldots,m\}^r,\, \#\{j_1,\ldots,j_r\} = r} \left(\sum_{\sigma \in \mathcal{S}_r} x_{1,j_{\sigma(1)}} \ldots x_{r,j_{\sigma(r)}} f(b_{j_{\sigma(1)}}, \ldots, b_{j_{\sigma(r)}}) \right).$$

Pour chaque ensemble $E = \{j_1, \ldots, j_r\}$ avec $j_1 < \cdots < j_r$ et chaque $\sigma \in \mathcal{S}_r$, on réordonne ensuite les éléments de $f(b_{j_{\sigma(1)}}, \ldots, b_{j_{\sigma(r)}})$ et il vient

$$f(b_{j_{\sigma(1)}}, \ldots, b_{j_{\sigma(r)}}) = \operatorname{sign}(\sigma) f(b_{j_1}, \ldots, b_{j_r}).$$

On obtient alors

$$\begin{aligned} \sum_{\sigma \in \mathcal{S}_r} x_{1,j_{\sigma(1)}} \ldots x_{r,j_{\sigma(r)}} f(b_{j_{\sigma(1)}}, \ldots, b_{j_{\sigma(r)}}) &= \left(\sum_{\sigma \in \mathcal{S}_r} \operatorname{sign}(\sigma) x_{1,j_{\sigma(1)}} \ldots x_{r,j_{\sigma(r)}} \right) f(b_{i_1}, \ldots, b_{i_r}) \\ &= \det{}_E^{\mathcal{B}}(x_1, \ldots, x_r) f(\mathbf{b}_E) \end{aligned}$$

en reprenant les notations introduites précédemment, d'où la conclusion. ■

En particulier, puisque les applications $\det_E^{\mathcal{B}}$ engendrent l'espace des formes r-multilinéaires alternées sur V et puisqu'elles sont au nombre de C_n^r, la dimension de cet espace est inférieure ou égale à C_n^r. Elle définissent en fait une base de cet espace, comme le montre le résultat suivant.

Théorème 3.2.13 *Soit V un espace vectoriel de dimension m sur un champ de caractéristique différente de 2, soit $r < m$ et soit $\mathcal{B}$ une base de V. L'espace des formes r-multilinéaires alternées sur V est l'espace de dimension C_m^r engendré par les applications $\det_E^{\mathcal{B}}$, $E \subseteq \{1, \ldots, m\}$, $\#E = r$.*

Démonstration : Par la Proposition 3.2.12, on sait que les applications r-multilinéaires alternées $\det_E^{\mathcal{B}}$, $E \subseteq \{1, \ldots, m\}$, $\#E = r$, forment une partie génératrice de l'espace des formes r-multilinéaires alternées sur V. Il reste à montrer que ces formes sont linéairement indépendantes. Supposons donc qu'il existe des scalaires α_E tels que

$$\sum_{E\subseteq\{1,\ldots,m\},\#E=r} \alpha_E \det_E^{\mathcal{B}} = 0.$$

Fixons $E' \subseteq \{1, \ldots, m\}, \#E' = r$, et évaluons la relation précédente en $\mathbf{b}_{E'} = (b_{i_1}, \ldots, b_{i_r})$. Remarquons que

$$\det_E^{\mathcal{B}}(\mathbf{b}_{E'}) = \det_E\big([b_{i_1}]_{\mathcal{B}}, \ldots, [b_{i_r}]_{\mathcal{B}}\big) = \det_E A$$

où A est la matrice de dimension $m \times r$ dont l'élément (i, j) est donné par δ_{ij}. Par conséquent, si $E \neq E'$, comme on sélectionne dans le calcul du déterminant uniquement les lignes correspondantes aux éléments de E, on a forcément une ligne de 0 et donc $\det_E^{\mathcal{B}}(\mathbf{b}_{E'}) = 0$. Si $E = E'$, on a $\det_E^{\mathcal{B}}(\mathbf{b}_{E'}) = \det I = 1$. On en tire que

$$0 = \sum_{E\subseteq\{1,\ldots,m\},\#E=r} \alpha_E \det_E(\mathbf{b}_{E'}) = \alpha_{E'}$$

d'où la conclusion. ■

Dans le Théorème 3.1.5, nous avons utilisé le produit tensoriel $\varphi_1 \otimes \cdots \otimes \varphi_r$ de formes linéaires pour obtenir une base de l'espace des formes r-multilinéaires. Notons que celui-ci n'a en général aucune raison d'être alterné. On peut néanmoins construire une "antisymétrisation" de celui-ci. Par exemple, si $r = 2$, on peut définir l'application bilinéaire alternée

$$\varphi_1 \wedge \varphi_2 = \varphi_1 \otimes \varphi_2 - \varphi_2 \otimes \varphi_1,$$

c'est-à-dire

$$\varphi_1 \wedge \varphi_2(x_1, x_2) = \varphi_1(x_1)\varphi_2(x_2) - \varphi_2(x_1)\varphi_1(x_2).$$

De manière générale, on adopte la définition suivante.

Définition 3.2.14 Soient $\varphi_1, \ldots, \varphi_r \in V^*$. Le *produit extérieur* de $\varphi_1, \ldots, \varphi_r$ est l'application définie par

$$\varphi_1 \wedge \cdots \wedge \varphi_r = \sum_{\sigma\in\mathcal{S}_r} \operatorname{sign}(\sigma)\, \varphi_{\sigma(1)} \otimes \cdots \otimes \varphi_{\sigma(r)}.$$

Proposition 3.2.15 *Soient $\varphi_1, \ldots, \varphi_r \in V^*$. Le produit extérieur $\varphi_1 \wedge \cdots \wedge \varphi_r$ est une forme r-multilinéaire alternée.*

Démonstration : La multilinéarité est immédiate. Montrons que la forme est alternée. Supposons que $(x_1, \ldots, x_r) \in V^r$ est tel qu'il existe $i \neq j$ avec $x_i = x_j$. On a

$$\varphi_1 \wedge \cdots \wedge \varphi_r(x_1, \ldots, x_r) = \sum_{\sigma \in \mathcal{S}_r} \mathrm{sign}(\sigma)\, \varphi_{\sigma(1)}(x_1) \ldots \varphi_{\sigma(r)}(x_r).$$

Notons τ la transposition qui envoie i sur j. On peut décomposer la somme sur les $\sigma \in \mathcal{S}_r$ en la somme sur les σ qui contiennent, dans leur décomposition en transpositions, la transposition τ et ceux qui ne la contiennent pas. Notons $\mathcal{S}_r(\tau)$ les permutations de $\mathcal{S}_r$ qui contiennent τ dans leur décomposition en transpositions. On peut alors décomposer la somme sur les $\sigma \in \mathcal{S}_r$ en

$$\sum_{\sigma \in \mathcal{S}_r(\tau)} \mathrm{sign}(\sigma)\, \varphi_{\sigma(1)}(x_1) \ldots \varphi_{\sigma(r)}(x_r) + \sum_{\sigma \in \mathcal{S}_r \setminus \mathcal{S}_r(\tau)} \mathrm{sign}(\sigma)\, \varphi_{\sigma(1)}(x_1) \ldots \varphi_{\sigma(r)}(x_r).$$

Notons que toute permutation $\sigma \in \mathcal{S}_r(\tau)$ s'écrit $\sigma = \tau\nu$ où $\nu \in \mathcal{S}_r \setminus \mathcal{S}_r(\tau)$. De plus, on a $\mathrm{sign}(\sigma) = -\mathrm{sign}(\nu)$ et $\varphi_{\sigma(1)}(x_1) \ldots \varphi_{\sigma(r)}(x_r) = \varphi_{\nu(1)}(x_1) \ldots \varphi_{\nu(r)}(x_r)$ puisque les vecteurs x_i et x_j sont égaux. Enfin, lorsque l'on passe en revue tous les $\sigma \in \mathcal{S}_r(\tau)$, on obtient tous les $\nu \in \mathcal{S}_r \setminus \mathcal{S}_r(\tau)$. Au total, la somme précédente se réécrit

$$-\sum_{\nu \in \mathcal{S}_r \setminus \mathcal{S}_r(\tau)} \mathrm{sign}(\nu)\, \varphi_{\nu(1)}(x_1) \ldots \varphi_{\nu(r)}(x_r) + \sum_{\sigma \in \mathcal{S}_r \setminus \mathcal{S}_r(\tau)} \mathrm{sign}(\sigma)\, \varphi_{\sigma(1)}(x_1) \ldots \varphi_{\sigma(r)}(x_r) = 0,$$

ce qui prouve le résultat. ■

Lemme 3.2.16 *Soit V un espace vectoriel de dimension m sur un champ de caractéristique différente de 2, soit $r < m$ et soit $\mathcal{B} = (b_1, \ldots, b_m)$ une base de V. Les formes r-multilinéaires alternées*

$$b^*_{i_1} \wedge \cdots \wedge b^*_{i_r}, \quad 1 \leq i_1 < \cdots < i_r \leq m,$$

forment une base de l'espace des formes r-multilinéaires alternées sur V.

Démonstration : On a

$$\begin{aligned} b^*_{i_1} \wedge \cdots \wedge b^*_{i_r}(x_1, \ldots, x_r) &= \sum_{\sigma \in \mathcal{S}_r} \mathrm{sign}(\sigma)\, b^*_{i_{\sigma(1)}}(x_1) \ldots b^*_{i_{\sigma(r)}}(x_r) \\ &= \sum_{\sigma \in \mathcal{S}_r} \mathrm{sign}(\sigma) x_{1,i_{\sigma(1)}} \ldots x_{r,i_{\sigma(r)}} \\ &= \det{}^{\mathcal{B}}_E(x_1, \ldots, x_r) \end{aligned}$$

si $E = \{i_1, \ldots, i_r\}$. Pour conclure, il suffit d'utiliser le Théorème 3.2.13. ■

On obtient l'équivalent du Théorème 3.1.5.

Théorème 3.2.17 *Soit V un espace vectoriel de dimension m sur un champ de caractéristique différente de 2, soit $r < m$ et soit $(\varphi_1, \ldots, \varphi_m)$ une base de V^*. Alors les formes r-multilinéaires alternées*

$$\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_r}, \quad 1 \leq i_1 < \cdots < i_r \leq m,$$

forment une base de l'espace des formes r-multilinéaires alternées sur V.

Démonstration : Comme $(\varphi_1, \ldots, \varphi_m)$ est une base de V^*, il suffit de considérer sa base pré-duale $(b_1, \ldots, b_m)$ et d'appliquer le Lemme 3.2.16. ∎

Bibliographie

[1] Arnaudiès, J.M. et Fraysse, H. (1990) *Algèbre bilinéaire et géométrie* Dunod Université, Paris.

[2] Chambadal, L. et Ovaert, J.L. (1968) *Algèbre linéaire et algèbre tensorielle* Dunod Université, Paris.

[3] Hansoul, G. (2016–2017) *Algèbre II.* Université de Liège.

[4] Lang, S. (1976) *Algèbre linéaire 2.* InterEditions, Paris.

[5] Laubin, P. (1991) *Algèbre linéaire, première candidature en sciences mathématiques.* Université de Liège.

[6] Nering, E.D. (1970) *Linear algebra and matrix theory* John Wiley & Sons, Inc.

[7] Rigo, M. (2009–2010) *Algèbre linéaire.* Université de Liège.

[8] Roudier, H. (1998) *Algèbre linéaire : une introduction : cours et exercices corrigés.* Vuibert.

[9] Schneiders, J.P. (2007–2008) *Analyse III (1ère partie), troisième année de bachelier en sciences mathématiques.* Université de Liège.