Diagonalisation des endomorphismes et des matrices

Peter Haïssinsky, Université de Paul Sabatier

2014–2015

Ce chapitre traite de la diagonalisabilité des endomorphismes et des matrices en proposant différentes caractérisations et en donnant quelques applications. Un chapitre ultérieur devrait traiter de compléments importants sur le sujet.

Introduction

On se donne un espace vectoriel E de dimension finie sur un corps $\mathbb{K}$ (que l'on peut penser comme étant $\mathbb{R}$ ou $\mathbb{C}$) et un endomorphisme $u \in \mathcal{L}(\mathrm{E})$. La question traitée dans ce chapitre la suivante: existe-t-il une base $\mathcal{B}$ de E telle que l'expression de u dans cette base soit particulièrement simple? Plus précisément, existe-t-il une base $\mathcal{B}$ telle que l'expression de u soit une matrice diagonale? On dira alors que u est *diagonalisable*.

1 Motivations

Ce paragraphe présente quelques applications liées à la diagonalisabilité des matrices/endomorphismes.

On suppose que $\mathrm{D} = (d_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ est diagonale, c'est-à-dire que l'on a $d_{i,j} = 0$ si $i \neq j$. Il est alors pratique d'écrire $\mathrm{D} = \mathrm{diag}\,(d_{1,1}, \dots, d_{n,n})$.

1.1 Calculs d'invariants

On peut facilement calculer quelques invariants de similitude d'un endomorphisme. Par exemple, supposons que D soit diagonale ($d_{i,j} = 0$ si $i \neq j$).

Il vient

1. $\det \mathrm{D} = \prod \lambda_j$,

2. $\mathrm{tr}\,\mathrm{D} = \sum \lambda_j$,

3. Si aucun terme diagonal n'est nul, alors D^{-1} est la matrice diagonale de termes diagonaux $1/d_{j,j}$.

Par conséquent, si A est diagonalisable, prenant les valeurs $\lambda_1, \dots, \lambda_n$ sur la diagonale, alors

1. $\det \mathrm{A} = \prod \lambda_j$,

2. $\mathrm{tr}\,\mathrm{A} = \sum \lambda_j$,

1.2 Calcul de puissances

Soient $\mathrm{D} = (d_{i,j})$ et $\mathrm{D}' = (d'_{i,j})$ deux matrices diagonales. Notons $(c_{ij}) = \mathrm{DD}'$. On a

$$c_{ik} = \sum_{j=1}^{n} d_{ij} d'_{jk} = d_{ii} d'_{ik}\,.$$

Donc $c_{ii} = d_{ii} d'_{ii}$ et $c_{ij} = 0$ sinon. On en déduit que DD' est diagonale et plus précisément

$$\mathrm{diag}\,(d_{1,1}, \dots, d_{n,n}) \times \mathrm{diag}\,(d'_{1,1}, \dots, d'_{n,n}) = \mathrm{diag}\,(d_{1,1} d'_{1,1}, \dots, d_{n,n} d'_{n,n})\,.$$

Par conséquent, la matrice D^k est une matrice diagonale dont le $i^{\text{ème}}$ coefficient de la diagonale est d_{ii}^k. C'est vrai pour $k=1$, et si vrai au rang k, alors, d'après le calcul précédent, comme $\mathrm{D}^{k+1} = \mathrm{D}\cdot\mathrm{D}^k$, on en déduit que D^{k+1} est diagonale et le $i^{\text{ème}}$ coefficient de la diagonale est $d_{ii}\cdot d_{ii}^k = d_{ii}^{k+1}$.

Maintenant, supposons que A est semblable à D de sorte que $\mathrm{A} = \mathrm{PDP}^{-1}$. Montrons que $\mathrm{A}^k = \mathrm{PD}^k\mathrm{P}^{-1}$. C'est vrai pour $k=1$, et si c'est vrai au rang k, alors

$$\mathrm{A}^{k+1} = \mathrm{A}^k\mathrm{A} = (\mathrm{PD}^k\mathrm{P}^{-1})(\mathrm{PDP}^{-1}) = \mathrm{PD}^k(\mathrm{P}^{-1}\mathrm{P})\mathrm{DP}^{-1} = \mathrm{PD}^{k+1}\mathrm{P}^{-1}$$

donc c'est vrai au rang $k+1$. Donc c'est vrai pour tout $k \geq 1$ par récurrence.

Du coup, on trouve l'expression

$$\mathrm{A}^k = \mathrm{P}\operatorname{diag} d_{11}^k, \ldots, d_{nn}^k)\mathrm{P}^{-1}\,.$$

1.3 Suites définies par des relations de récurrence linéaires

Dans un premier temps, on se donne des suites $(u_k^1)_{k\geq 0}, \ldots, (u_k^n)_{k\geq 0}$ et on suppose qu'elles vérifient, pour $k \geq 0$,

$$\left\{\begin{array}{lcl} u_{k+1}^1 & = & a_{1,1}u_k^1 + \ldots + a_{1,n}u_k^n \\ \ldots & = & \ldots \\ u_{k+1}^n & = & a_{n,1}u_k^1 + \ldots + a_{n,n}u_k^n \end{array}\right.$$

ce qui s'écrit sous forme matricielle $\mathrm{X}_{k+1} = \mathrm{AX}_k$ où $\mathrm{A} = (a_{i,j})$ et

$$\mathrm{X}_k = \begin{pmatrix} u_k^1 \\ \vdots \\ u_k^n \end{pmatrix}.$$

Du coup, on trouve $\mathrm{X}_k = \mathrm{A}^k\mathrm{X}_0$ pour tout $k \geq 0$. Si A est diagonalisable, on a $\mathrm{D} = \mathrm{P}^{-1}\mathrm{AP}$, donc $\mathrm{A}^k = \mathrm{PD}^k\mathrm{P}^{-1}$ et $\mathrm{X}_k = \mathrm{PD}^k\mathrm{P}^{-1}\mathrm{X}_0$. Si on ne s'intéresse qu'à la forme des solutions, alors, en écrivant $\mathrm{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, on trouve

$$\mathrm{X}_k = \mathrm{P}\operatorname{diag}(c_1\lambda_1^k, \ldots, c_n\lambda_n^k)$$

pour des coefficients c_j à déterminer.

Supposons maintenant que l'on ait une suite $(u_k)_{k\geq 0}$ telle que, pour $n \geq 1$ et des scalaires $a_1, \ldots, a_n$, on ait

$$u_{k+n} = a_1u_k + \ldots + a_nu_{k+n-1}\,.$$

Alors en posant

$$\mathrm{X}_k = \begin{pmatrix} u_k \\ \vdots \\ u_{k+n-1} \end{pmatrix}$$

et

$$\mathrm{A} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & & 0 \\ & & \ddots & \ddots & 0 \\ 0 & \ldots & & 0 & 1 \\ a_1 & a_2 & \ldots & a_{n-1} & a_n \end{pmatrix}$$

on se ramène à la situation précédente: $\mathrm{X}_{k+1} = \mathrm{AX}_k$.

Remarque 1.1. — *L'ensemble de ces suites forme un espace vectoriel sur $\mathbb{R}$ de dimension n.*

1.4 Systèmes d'équations différentielles linéaires

On s'intéresse ici à un système d'équations différentielles linéaires de la forme

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

où $(x_1, \ldots, x_n)$ sont des fonctions de classe $\mathcal{C}^1$ inconnues et les $(a_{i,j})$ sont des scalaires. On obtient sous forme matricielle $\mathrm{X}' = \mathrm{AX}$ où $\mathrm{A} = (a_{i,j})$ et

$$\mathrm{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} .$$

Si A est une matrice diagonale, alors chaque équation est indépendante des autres et toute solution s'écrit sous la forme

$$\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \begin{pmatrix} c_1\mathrm{e}^{a_{1,1}t} \\ \vdots \\ c_n\mathrm{e}^{a_{n,n}t} \end{pmatrix}$$

où les c_i sont des constantes.

Supposons maintenant A diagonalisable avec $\mathrm{D} = \mathrm{P}^{-1}\mathrm{AP} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Si on pose $\mathrm{Y} = \mathrm{P}^{-1}\mathrm{X}$, les règles de dérivation des fonctions impliquent $\mathrm{Y}' = \mathrm{P}^{-1}\mathrm{X}'$ et donc

$$\mathrm{Y}' = \mathrm{P}^{-1}\mathrm{X}' = \mathrm{P}^{-1}\mathrm{AX} = \mathrm{DY} ,$$

ce qui implique que les solutions sont de la forme

$$\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} c_1\mathrm{e}^{\lambda_1 t} \\ \vdots \\ c_n\mathrm{e}^{\lambda_n t} \end{pmatrix}$$

et

$$\mathrm{X} = \mathrm{PY} = \mathrm{P} \begin{pmatrix} c_1\mathrm{e}^{\lambda_1 t} \\ \vdots \\ c_n\mathrm{e}^{\lambda_n t} \end{pmatrix}$$

Il suffit donc, dans cette situation, de calculer une base de vecteurs propres.

Comme pour les suites, on peut aussi étudier une équation de la forme

$$y^{(n)} = a_0 y + a_1 y' + \ldots + a_{k-1}y^{(n-1)} .$$

On se ramène à la situation précédente en posant

$$\mathrm{X} = \begin{pmatrix} y \\ \vdots \\ y^{(n-1)} \end{pmatrix}$$

pour obtenir $\mathrm{X}' = \mathrm{AX}$ avec

$$\mathrm{A} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & & 0 \\ & & \ldots & & 0 \\ 0 & \ldots & & 0 & 1 \\ a_0 & a_1 & \ldots & a_{n-1} & a_n \end{pmatrix}$$

Exponentielle de matrices. Si $\mathrm{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ est diagonale, on pose

$$\exp \mathrm{D} = \mathrm{diag}(\exp \lambda_1, \ldots, \exp \lambda_n) = \mathrm{diag}\left(\sum_{k\geq 0} \frac{\lambda_1^k}{k!}, \ldots, \sum_{k\geq 0} \frac{\lambda_1^k}{k!}\right) = \sum_{k\geq 0} \frac{1}{k!}\mathrm{D}^k .$$

On peut en fait donner un sens à l'expression

$$\exp \mathrm{A} = \sum_{k\geq 0} \frac{1}{k!}\mathrm{A}^k$$

pour n'importe quelle matrice $\mathrm{A} \in \mathrm{Mat}_n(\mathbb{K})$, et on l'appelle l'exponentielle de la matrice A. En général, $\exp(\mathrm{A}+\mathrm{B}) \neq \exp \mathrm{A} \exp \mathrm{B}$; cependant, on a égalité si $\mathrm{AB} = \mathrm{BA}$.

Ce que l'on a vu ci-dessus est un cas particulier de la situation suivante: on s'intéresse au système d'équations différentielles linéaires $\mathrm{X}' = \mathrm{AX}$. Alors ces solutions sont données par $\mathrm{X}(t) = \exp(\mathrm{A}t)\mathrm{X}_0$ où X_0 est un vecteur constant. Lorsque A est diagonalisable, alors on a une expression de $\exp \mathrm{A}$: $\exp \mathrm{A} = \mathrm{P}^{-1}(\exp \mathrm{D})\mathrm{P}$ si $\mathrm{A} = \mathrm{P}^{-1}\mathrm{DP}$.

2 Valeurs et vecteurs propres

Soit E un espace de dimension finie et soit $u \in \mathcal{L}(\mathrm{E})$. On dit que $\lambda \in \mathbb{K}$ est une *valeur propre* de u s'il existe une vecteur **non nul** $v \in \mathrm{E}$ tel que $u(v) = \lambda v$. On dit que v est un *vecteur propre* associé à la valeur propre λ si $v \neq 0$ et $u(v) = \lambda v$. Si λ est une valeur propre, son *espace propre* associé E_λ est l'ensemble des vecteurs propres associés à λ auquel on rajoute le vecteur nul. Le *spectre* $\mathcal{S}(u)$ de u désigne l'ensemble de toutes les valeurs propres de u.

Remarque 2.1. — *Si u est non injective, alors 0 est valeur propre, puisqu'alors $\mathrm{Ker}\, u$ est non trivial et il existe donc $x \neq 0$ tel que $u(x) = 0 = 0 \cdot x$. Dans ce cas, $\mathrm{E}_0 = \mathrm{Ker}\, u$.*

Lemme 2.2. — *Soit λ une valeur propre. Son espace propre E_λ est un sous-espace vectoriel.*

Démonstration. — On remarque que $v \in \mathrm{E}_\lambda$ si et seulement si $u(v) = \lambda v$. En particulier, $0 \in \mathrm{E}_\lambda$. Soient $v, v' \in \mathrm{E}_\lambda$ et $\mu \in \mathbb{K}$. On calcule

$$u(v + \mu v') = u(v) + \mu u(v') = \lambda u + \mu\lambda v' = \lambda(v + \mu v')$$

donc $(v + \mu v') \in \mathbb{K}$. ■

Proposition 2.3. — *Un endomorphisme est diagonalisable si et seulement si il existe une base de vecteurs propres de u.*

Démonstration. — Supposons u diagonalisable. Soit $\mathcal{B}$ une base de E qui rend u diagonale. Autrement dit, $\mathrm{Mat}_{\mathcal{B}}(u) = (a_{i,j})$ est diagonale c'est-à-dire $a_{i,j} = 0$ si $i \neq j$. Par définition, cela signifie que $u(\mathrm{e}_i) = a_{i,i}\mathrm{e}_i$, donc $\mathcal{B}$ est une base de vecteurs propres.

Réciproquement, si $\mathcal{B}$ est une base de vecteurs propres de u, alors, pour chaque j, il existe un scalaire λ_j tel que $u(\mathrm{e}_j) = \lambda_j \mathrm{e}_j$, donc $\mathrm{Mat}_{\mathcal{B}}(u)$ est diagonale. ■

Sommes directes de plusieurs sous-espaces. La notion de somme directe généralise aux sous-espaces vectoriels celle de famille libre pour les vecteurs. Soient E un espace vectoriel et $\mathrm{F}_1, \ldots, \mathrm{F}_k$ des sous-espaces de E. On peut munir le produit cartésien $\mathrm{F}_1 \times \mathrm{F}_2 \times \ldots \times \mathrm{F}_k$ d'une structure d'espace vectoriel comme pour $\mathbb{K}^n$ en définissant

$$\begin{pmatrix} f_1 \\ \vdots \\ f_k \end{pmatrix} + \begin{pmatrix} f_1' \\ \vdots \\ f_k' \end{pmatrix} = \begin{pmatrix} f_1 + f_1' \\ \vdots \\ f_k + f_k' \end{pmatrix}$$

et

$$\lambda \times \begin{pmatrix} f_1 \\ \vdots \\ f_k \end{pmatrix} = \begin{pmatrix} \lambda f_1 \\ \vdots \\ \lambda f_k \end{pmatrix} .$$

Considérons maintenant

$$\begin{array}{rccc} \Phi : & \mathrm{F}_1 \times \ldots \times \mathrm{F}_k & \to & \mathrm{E} \\ & \begin{pmatrix} f_1 \\ \vdots \\ f_k \end{pmatrix} & \mapsto & \sum_{j=1}^k f_j \end{array}$$

On vérifie que Φ est une application linéaire dont l'image définit la somme $\mathrm{F}_1 + \ldots + \mathrm{F}_k$.

On dit que la somme est *directe* et l'on note $\mathrm{F}_1 \oplus \ldots \oplus \mathrm{F}_k = \oplus_{j=1}^k \mathrm{F}_j$ si Φ est injective. Cela signifie que tout vecteur $v \in (\mathrm{F}_1 + \ldots + \mathrm{F}_k)$ admet une unique écriture de la forme $v = \sum f_j$ avec $f_j \in \mathrm{F}_j$.

Lemme 2.4. — *Avec les notations ci-dessus, la somme des* F_j *est directe si et seulement si, pour tout* $1 \leq j \leq k-1$,

$$(\mathrm{F}_1 + \ldots + \mathrm{F}_j) \cap \mathrm{F}_{j+1} = \{0\} .$$

Démonstration. — Supposons la somme directe et prenons $v \in (\mathrm{F}_1 + \ldots + \mathrm{F}_j) \cap \mathrm{F}_{j+1}$. Cela signifie que l'on peut trouver $(f_1, \ldots, f_j) \in (\mathrm{F}_1 \times \ldots \times \mathrm{F}_j)$ tel que $v = \sum f_i$. Du coup,

$$\Phi \begin{pmatrix} f_1 \\ \vdots \\ f_j \\ -v \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0 .$$

Par injectivité, cela signifie que tous les vecteurs sont nuls, et en particulier $v = 0$.

Réciproquement, prenons

$$\begin{pmatrix} f_1 \\ \vdots \\ f_k \end{pmatrix} \in \operatorname{Ker} \Phi .$$

On a donc

$$f_k = -\sum_{j=1}^{k-1} f_j$$

donc $f_k \in \mathrm{F}_k \cap (\mathrm{F}_1 + \ldots + \mathrm{F}_{k-1})$ et $f_k = 0$. Par récurrence descendante, on montre ainsi que tous les f_j sont nuls, montrant que la somme est directe. ∎

On obtient ainsi deux méthodes pour montrer que des espaces sont somme directe:

1. on applique le lemme 2.4;
2. on montre que si $x_j \in \mathrm{F}_j$ pour tous $j \in \{1, \ldots, k\}$ vérifient $\sum x_j = 0$, alors tous les x_j sont nuls.

Propriété 2.5. — *Soient* E *un espace vectoriel et* $\mathrm{F}_1, \ldots, \mathrm{F}_k$ *des sous-espaces de* E. *Les propriétés suivantes sont équivalentes.*

1. $\mathrm{F}_1 + \ldots + \mathrm{F}_k = \oplus_{j=1}^k \mathrm{F}_j$.
2. $\dim(\mathrm{F}_1 + \ldots + \mathrm{F}_k) = \dim \mathrm{F}_1 + \ldots + \dim \mathrm{F}_k$,
3. *La réunion de familles libres des* F_j *forme une famille libre.*

On obtient ainsi la proposition suivante:

Proposition 2.6. — *Soient $\lambda_1, \ldots, \lambda_k$ des valeurs propres de u deux à deux distinctes.*

1. *Leurs sous-espaces propres sont en somme directe.*
2. *Si, pour tout $j \in \{1, \ldots, k\}$, v_j est un vecteur propre associé à λ_j, alors $(v_1, \ldots, v_k)$ est libre.*

Démonstration. — La seconde propriété découle aisément de la première.

On procède par récurrence pour montrer que les sous-espaces sont en somme directe. Si $k = 2$, prenons $v \in \mathrm{E}_{\lambda_1} \cap \mathrm{E}_{\lambda_2}$, on a donc

$$\lambda_1 v = u(v) = \lambda_2 v$$

donc $(\lambda_1 - \lambda_2)v = 0$ et comme $\lambda_1 \neq \lambda_2$, on obtient $v = 0$.

Supposons maintenant que pour tout $1 \leq j \leq k-2$,

$$(\mathrm{E}_{\lambda_1} \oplus \ldots \oplus \mathrm{E}_{\lambda_j}) \cap \mathrm{E}_{\lambda_{j+1}} = \{0\}\,.$$

Traitons le cas $j = k-1$ et prenons donc $v \in (\mathrm{E}_{\lambda_1} + \ldots + \mathrm{E}_{\lambda_{k-1}}) \cap \mathrm{E}_{\lambda_k}$. On écrit $v = \sum_{j=1}^{k-1} v_j$ avec $v_j \in \mathrm{E}_{\lambda_j}$ et on obtient comme ci-dessus

$$\lambda_k v = \sum_{j=1}^{k-1} \lambda_j v_j \quad \text{donc} \quad \sum_{j=1}^{k-1} (\lambda_j - \lambda_k) v_j = 0$$

et comme les $(k-1)$ premiers espaces sont en somme directe, cela signifie que $(\lambda_j - \lambda_k)v_j = 0$ pour tout j donc $v_j = 0$ et $v = 0$. ■

Corollaire 2.7. — *Soit u un endomorphisme sur un espace de dimension finie. Les propriétés suivantes sont équivalentes.*

1. *L'endomorphisme u est diagonalisable.*
2. *On a $\oplus_{\lambda \in \mathcal{S}(u)} \mathrm{E}_\lambda = \mathrm{E}$.*
3. *On a $\sum_{\lambda \in \mathcal{S}(u)} \dim \mathrm{E}_\lambda = \dim \mathrm{E}$.*

Démonstration. — Si u est diagonalisable, alors il existe une base de vecteurs propres. Cela signifie que les espaces sont en somme directe et que leur somme vaut E. Si on a $\oplus_{\lambda \in \mathcal{S}(u)} \mathrm{E}_\lambda = \mathrm{E}$, alors il vient $\sum_{\lambda \in \mathcal{S}(u)} \dim \mathrm{E}_\lambda = \dim \mathrm{E}$.

D'après la proposition précédente, les sous-espaces propres sont en somme directe, donc si on suppose $\sum_{\lambda \in \mathcal{S}(u)} \dim \mathrm{E}_\lambda = \dim \mathrm{E}$, alors on obtient $\mathrm{E} = \oplus_{\lambda \in \mathcal{S}(u)} \mathrm{E}_\lambda$. En prenant une base de chaque espace, on obtient une base de E de vecteurs propres. ■

3 Polynôme caractéristique

On suppose maintenant E de dimension finie n et on se donne un endomorphisme $u \in \mathcal{L}(\mathrm{E})$. Si $\lambda \in \mathbb{K}$ est valeur propre, il existe donc par définition un vecteur non nul $x \in \mathrm{E}$ tel que $u(x) - \lambda x = 0$, c'est-à-dire l'application $u_\lambda := u - \lambda \mathrm{Id}$ est non injective. Réciproquement, si, pour un scalaire λ, l'application u_λ est non injective, alors cela signifie que λ est valeur propre. Par ailleurs, on obtient ainsi $\mathrm{E}_\lambda = \operatorname{Ker} u_\lambda$ montrant que les espaces propres sont bien des sous-espaces vectoriels.

On peut exprimer la non-injectivité de u_λ en disant que $\det u_\lambda = 0$.

Définition 3.1. — *Le polynôme caractéristique de u est*

$$\chi_u(\lambda) = \det(u - \lambda \mathrm{Id})\,.$$

Propriété 3.2. — *Le polynôme caractéristique d'un endomorphisme est un polynôme à coefficients dans $\mathbb{K}$ de degré n en λ dont les racines forment le spectre de l'endomorphisme. On a*

$$\chi_u(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} (\operatorname{tr} u) \lambda^{n-1} + \ldots + \det u\,.$$

Démonstration. — Prenons une base $\mathcal{B}$ de E et calculons χ_u en considérant $\mathrm{Mat}_{\mathcal{B}}(u) = (a_{i,j})$. D'après la formule du déterminant d'une matrice, on constate que χ_u est un polynôme de degré au plus n. Les termes de plus haut degré s'obtiennent dans le produit des termes de la diagonale. En effet, si $\sigma \neq \mathrm{Id}$, alors nous aurons deux facteurs hors de la diagonale donc le produit sera de degré au plus $n-2$. On obtient ainsi

$$\prod_{j=1}^{n}(a_{j,j}-\lambda) = (-\lambda)^n + (-\lambda)^{n-1}\sum_{j=1}^{n} a_{j,j} + \dots$$

Cette formule peut s'établir par récurrence sur n.

Enfin, pour le terme constant, on peut par exemple calculer $\chi_u(0) = \det u$. ■

Proposition 3.3. — *Soient* E *un espace de dimension* n*,* $u \in \mathcal{L}(\mathrm{E})$*,* $\mu \in \mathbb{K}$ *et* $k \geq 1$*. On suppose que* $\chi_u(\lambda) = (\lambda-\mu)^k \mathrm{P}(\lambda)$ *avec* $\mathrm{P}(\mu) \neq 0$*, c'est-à-dire* μ *est une racine de multiplicité* k *de* χ_u*. Alors* $\dim \mathrm{E}_\mu \leq k$*.*

La multiplicité d'une valeur propre λ en tant que racine d'un polynôme caractéristique (la valeur m_λ) est appelée la *multiplicité algébrique* de λ. La dimension de son espace propre est *la multiplicité géométrique*.

Démonstration. — Soit $(\mathrm{e}_1, \dots, \mathrm{e}_m)$ une base de E_μ que l'on complète en une base de E. La matrice de u dans cette base est donc diagonale sur les m premières lignes et colonnes, donc, dans le calcul de χ_u, on pourra mettre en facteur $(\mu-\lambda)^m$. Du coup, $m \leq k$ et $\dim \mathrm{E}_\mu \leq k$. En effet, on peut écrire $\mathrm{Mat}\, u$ sous la forme

$$\mathrm{Mat}\, u = \begin{pmatrix} \mu \mathrm{I}_m & \mathrm{A} \\ 0 & \mathrm{C} \end{pmatrix}$$

donc

$$\chi_u(\lambda) = \det \begin{pmatrix} (\mu-\lambda)\mathrm{I}_m & \mathrm{A} \\ 0 & \mathrm{C}-\lambda \mathrm{I}_{n-m} \end{pmatrix} = (\mu-\lambda)^m \times \det(\mathrm{C}-\lambda\mathrm{I}_{n-m}) = (\mu-\lambda)^k \mathrm{P}(\lambda)\,.$$

■

Corollaire 3.4. — *Soient* E *un espace de dimension* $n \geq 1$*. Un endomorphisme* u *est diagonalisable si et seulement si*

1. χ_u *est scindé sur* $\mathbb{K}$*, c'est-à-dire il existe* $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ *deux à deux distincts et* $m_1, \dots, m_k$ *tels que*

$$\chi_u(\lambda) = \prod_{j=1}^{k}(\lambda_j - \lambda)^{m_j}$$

et

2. *pour chaque* $j \in \{1, \dots, k\}$*, on a* $\dim \mathrm{E}_{\lambda_j} = m_j$*.*

Démonstration. — Si u est diagonalisable, on considère la matrice de u dans une base de vecteurs propres et on calcule χ_u. Comme la matrice est diagonale, les conditions sont évidentes. Réciproquement, les hypothèses impliquent

$$\dim \mathrm{E} = \deg \chi_u = \sum_{j=1}^{k} m_j = \sum_{j=1}^{k} \dim \mathrm{E}_{\lambda_j}$$

donc le corollaire 2.7 permet de conclure. ■

On en tire un cas particulier important:

Corollaire 3.5. — *Soient* E *un espace de dimension* $n \geq 1$*. Soit* u *un endomorphisme de* E *tel que*

1. χ_u *est scindé sur* $\mathbb{K}$ *et*

2. *toutes les racines sont simples* *alors u est diagonalisable.*

Démonstration. — En effet on a pour chaque valeur propre λ, $1 \leq \dim \mathrm{E}_\lambda \leq 1$, donc le corollaire précédent s'applique. ■

Pour terminer ce paragraphe, rappelons le théorème fondamental suivant:

Théorème 3.6 (d'Alembert). — *Tout polynôme non constant à coefficients dans $\mathbb{C}$ est scindé. Si les coefficients sont réels, alors les racines complexes sont conjuguées. Tout polynôme non constant à coefficients réels se factorise en produit de monômes et de polynômes quadratiques dont les racines complexes sont conjuguées.*

4 Trigonalisation

On dit qu'un endomorphisme est *trigonalisable* s'il existe une base dans laquelle la matrice de l'endomorphisme est triangulaire supérieure. Si $\mathrm{Mat}(u, \mathcal{B}) = (a_{i,j})$ alors $a_{i,j} = 0$ pour $1 \leq j < i \leq n$.

Supposons qu'un endomorphisme u est trigonalisable et prenons une base $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ de sorte que $\mathrm{Mat}(u, \mathcal{B})$ est triangulaire supérieure. On peut remarquer que les valeurs propres sont les termes de la diagonale et que le polynôme caractéristique est scindé. De plus, si on note $\mathrm{F}_j = \mathrm{Vect}(\mathrm{e}_1, \ldots, \mathrm{e}_j)$, alors $u(\mathrm{F}_j) \subset \mathrm{F}_j$.

Théorème 4.1. — *Un endomorphisme u sur un espace vectoriel de dimension finie est trigonalisable si et seulement si son polynôme caractéristique est scindé.*

Démonstration. — On procède par récurrence sur la dimension. Si l'espace est de dimension 1, alors pas de problème. Supposons maintenant le théorème vrai pour tout endomorphisme d'un espace de dimension n et prenons u sur un espace E de dimension $n+1$. Comme χ_u est scindé, il existe une valeur propre λ_1 et un vecteur propre e_1. On complète e_1 en une base de E. Les derniers vecteurs engendrent un espace F de dimension n. Notons $p : \mathrm{E} \to \mathrm{F}$ la projection de E sur F parallèlement à $\mathbb{K}\mathrm{e}_1$ Notons $v = p \circ u|_{\mathrm{F}} : \mathrm{F} \to \mathrm{F}$. On a $\chi_u(\lambda) = (\lambda_1 - \lambda)\chi_v(\lambda)$. Donc χ_v est scindé et on peut trouver une autre base $\mathcal{B}$ de F telle que v soit triangulaire. Dans la base $\{\mathrm{e}_1\} \cup \mathcal{B}$, la matrice de u est aussi triangulaire. ■

Calculs d'invariants. On peut facilement calculer quelques invariants de similitude d'un endomorphisme trigonalisable. Par exemple, si T est triangulaire supérieure, alors les valeurs propres de T sont justement les termes diagonaux.

On trouve ainsi

1. $\det \mathrm{T} = \prod \lambda_j$,

2. $\mathrm{tr}\, \mathrm{T} = \sum \lambda_j$,

3. $\chi_{\mathrm{T}}(\lambda) = \prod(\lambda_j - \lambda)$,

Par conséquent, si A est trigonalisable, comme c'est le cas sur $\mathbb{C}$, de valeurs propres $\lambda_1, \ldots, \lambda_n$ alors

1. $\det \mathrm{A} = \prod \lambda_j$,

2. $\mathrm{tr}\, \mathrm{A} = \sum \lambda_j$,

3. $\chi_{\mathrm{A}}(\lambda) = \prod(\lambda_j - \lambda)$.

Remarque 4.2. — *Si $\mathrm{A} \in \mathcal{M}_n(\mathbb{R})$, et*

$$\mathcal{S}_{\mathrm{A}} = \{\lambda_1, \ldots, \lambda_k, \mu_1, \overline{\mu_1}, \ldots, \mu_\ell, \overline{\mu_\ell}\}$$

où λ_j désigne les valeurs propres réelles et $\mu_j, \overline{\mu_j}$ les complexes (conjuguées deux à deux) comptées avec mulitplicité. On a alors

$$\det \mathrm{A} = \left(\prod\nolimits_{j=1}^{k} \lambda_j\right) \left(\prod\nolimits_{i=1}^{\ell} |\mu_i|^2\right) \quad \textit{et} \quad \mathrm{tr}\, \mathrm{A} = \sum_{j=1}^{k} \lambda_j + 2 \sum_{i=1}^{\ell} Re\, \mu_i \,.$$

5 Méthode pour diagonaliser un endomorphisme ou une matrice

Voici les grandes étapes pour étudier la diagonalisation d'un endormorphisme u d'un espace E de dimension finie ou d'une matrice A de taille n.

1. Si on part d'un endomorphisme, on lui attribue une matrice A en l'exprimant dans une base.

2. On calcule χ_A et on détermine ses racines et leurs multiplicités (λ_j, m_j).

3. Si χ_A n'est pas scindé, alors A n'est pas diagonalisable.

4. Pour chaque valeur propre λ, on résout le système $AX = \lambda X$ et on calcule la dimension des solutions E_λ. Si $\dim E_\lambda < m_\lambda$, alors A n'est pas diagonalisable.

5. Pour chaque λ, on calcule une base de E_λ.

6. Si A est diagonalisable, alors la réunion des bases de chaque espace propre fournit une base de $\mathbb{K}^n$ et on écrit la matrice de passage P ainsi que la matrice (diagonale) de u dans cette base.

Une méthode pédestre que je déconseille, mais qui peut marcher, consiste à écrire un système à paramètre $AX = \lambda X$ que l'on résout en fonction de λ. Les valeurs propres correspondent alors aux paramètres λ pour lesquels on a une infinité de solutions. Pour ceux-là, on détermine une base comme ci-dessus.

5.1 Matrices à coefficients réels

Soit A une matrice à coefficients réels que l'on suppose diagonalisable sur $\mathbb{C}$. Il est possible que le polynôme caractéristique soit scindé sur $\mathbb{C}$ mais pas sur $\mathbb{R}$. Dans ce cas nous aurons l'existence d'une valeur propre $\lambda = a + ib$ non réelle et d'une base de vecteurs propres $(v_1, \ldots, v_k)$ de E_λ.

Comme A est réelle, on déduit de $Av_j = \lambda v_j$ l'identité

$$A\overline{v_j} = \overline{\lambda}\overline{v_j}$$

ce qui implique que $\overline{\lambda}$ est aussi valeur propre de A et $(\overline{v_1}, \ldots, \overline{v_k})$ est une base de $E_{\overline{\lambda}}$.

Si on décompose $v_j = r_j + is_j$ en partie réelle et imaginaire, alors, en remarquant que

$$\begin{cases} r_j = (1/2)(v_j + \overline{v_j}) \\ s_j = (-i/2)(v_j - \overline{v_j}) \end{cases}$$

on constate que $(r_1, s_1, \ldots, r_k, s_k)$ est aussi une base de $E_\lambda \oplus E_{\overline{\lambda}}$ vu comme sous espace de $\mathbb{C}^n$ (c'est une famille génératrice de cardinal minimal). Par conséquent, c'est aussi une famille libre de $\mathbb{R}^n$ dont l'espace engendré est en somme directe avec les autres. En identifiant parties réelles et imaginaires, on obtient

$$\begin{cases} Ar_j = ar_j - bs_j \\ As_j = br_j + as_j \end{cases}$$

Si on écrit $\lambda = |\lambda| e^{i\theta}$, alors on a

$$\begin{cases} Ar_j = |\lambda|(\cos\theta r_j - \sin\theta s_j) \\ As_j = |\lambda|(\sin\theta r_j + \cos\theta s_j) \end{cases}$$

donc A opère comme une homothétie de rapport $|\lambda|$ suivie d'une rotation d'angle $-\theta$. Du coup, la matrice de $u_A \,|_{E_\lambda \oplus E_{\overline{\lambda}}}$ dans la base $(r_1, s_1, \ldots, r_k, s_k)$ sera la matrice par blocs

$$\begin{pmatrix} |\lambda| R_{-\theta} & 0 & \ldots & 0 \\ 0 & |\lambda| R_{-\theta} & 0 & 0 \\ 0 & \ldots & \ldots & 0 \\ 0 & \ldots & 0 & |\lambda| R_{-\theta} \end{pmatrix}$$

5.2 Exemples

1. On considère la matrice

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

à diagonaliser.

On a

$$A - \lambda I_3 = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 2 & 0 \\ -1 & 2 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} 1-\lambda & 2 & -1 \\ 0 & 2-\lambda & 0 \\ -1 & 2 & 1-\lambda \end{pmatrix}$$

Pour caluler le polynôme caractéristique, on développe selon la deuxième ligne:

$$\chi_A(\lambda) = (2-\lambda) \times \det \begin{pmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{pmatrix} .$$

Or

$$\begin{aligned} \det \begin{pmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{pmatrix} &= (1-\lambda)^2 - 1 \\ &= (1-\lambda-1)(1-\lambda+1) \\ &= (-1)\lambda(2-\lambda) \end{aligned}$$

Par suite, on trouve

$$\chi_A(\lambda) = (-1)\lambda(2-\lambda)^2$$

et le polynôme caractéristique est scindé, les valeurs propres étant $\lambda_1 = 0$ de multiplicité algébrique $m_0 = 1$ et $\lambda_2 = 2$ de multiplicité algébrique $m_2 = 2$.

$\underline{\lambda = 0}$. On cherche donc un vecteur du noyau: on constate que si on ajoute la première colonne de A à la troisième, alors on obtient une colonne nulle: donc

$$v_0 = e_1 + e_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

engendre E_0.

$\underline{\lambda = 2}$. On cherche à résoudre

$$\begin{pmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

On n'a qu'une seule équation:

$$-x_1 + 2x_2 - x_3 = 0$$

soit

$$x_2 = \frac{1}{2}(x_1 + x_3) .$$

Du coup, on a l'équation d'un plan (de dimension 2) et on en déduit que A est diagonalisable.

Si on prend $x_1 = 2$ et $x_3 = 0$, on trouve

$$v_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

et si on prend $x_1 = 1$ et $x_3 = -1$ alors on trouve

$$v_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} .$$

v_2 et v_3 ne sont pas colinéaires, donc ils forment une base de E_2.

On a trouvé une base de vecteurs propres dont la matrice de passage est

$$\mathrm{P} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} .$$

Du coup, on a

$$\mathrm{P}^{-1}\mathrm{AP} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} .$$

2. On considère la matrice

$$\mathrm{B} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

à diagonaliser.

On a

$$\mathrm{B} - \lambda \mathrm{I}_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix}$$

donc, comme cette matrice est triangulaire, on obtient

$$\chi_{\mathrm{B}}(\lambda) = (1-\lambda)^2$$

et $\lambda = 1$ est valeur propre de multiplicité algébrique 2. On remarque que si B était diagonalisable, alors il existerait une matrice inversible P telle que $\mathrm{P}^{-1}\mathrm{BP} = \mathrm{I}_2$. Or, en multipliant cette équation par P à gauche et P^{-1} à droite, on obtiendrait $\mathrm{B} = \mathrm{P} \times \mathrm{I}_2 \times \mathrm{P}^{-1} = \mathrm{I}_2$. Or $\mathrm{B} \neq \mathrm{I}_2$, donc B n'est pas diagonalisable.

3. On considère la matrice

$$\mathrm{C} = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}$$

à diagonaliser que l'on considère comme l'endomorphisme $u : \mathrm{X} \mapsto \mathrm{CX}$. On calcule

$$\mathrm{C} - \lambda \mathrm{I}_2 = \begin{pmatrix} 1-\lambda & 2 \\ -1 & -1-\lambda \end{pmatrix}$$

puis

$$\chi_u(\lambda) = (1-\lambda)(-1-\lambda) - 2(-1) = \lambda^2 + 1 \,.$$

Donc le polynôme n'est pas scindé sur $\mathbb{R}$, mais uniquement sur $\mathbb{C}$. Les valeurs propres sont $\pm i$. Calculons un vecteur propre associé à i: on résout

$$\left\{ \begin{array}{rrcl} x & +2y & = & ix \\ -x & -y & = & iy \end{array} \right.$$

On remplace L2 par $2\mathrm{L}2 + (1+i)\mathrm{L}1$

$$\left\{ \begin{array}{rrcl} (1-i)x & +2y & = & 0 \\ 0x & +0y & = & 0 \end{array} \right.$$

Donc $(1+i, -1)$ est un vecteur propre sur $\mathbb{C}^2$ associé à i et $(1-i, 1)$ est un vecteur propre associé à $(-i)$. Dand cette base $\mathcal{B}$, on a

$$\mathrm{Mat}(u, \mathcal{B}) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} .$$

Les vecteurs $(1, -1)$ et $(1, 0)$ forment une base $\mathcal{C}$ de $\mathbb{R}^2$ et aussi de $\mathbb{C}^2$. Dans cette base on obtient

$$\mathrm{Mat}(u, \mathcal{B}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

où l'on reconnaît la matrice $\sqrt{2}\mathrm{R}_{\pi/2}$.