Chapitre 7

Valeurs propres et vecteurs propres

7.1 Définitions

Définition 7.1 *Soit $(K,+,\cdot)$ un corps et $A \in M_{n\times n}(K)$. Un scalaire $\lambda \in K$ s'appelle une* ***valeur propre*** *[eigenvalue] de A s'il existe un vecteur $x \in K^n$, $x \neq 0$, tel que*

$$Ax = \lambda x.$$

Le vecteur x s'appelle un ***vecteur propre*** *[eigenvector] de A associé à la valeur propre λ.*

Définition 7.2 *Soit V un K-espace vectoriel et $F \in L(V,V)$. Un scalaire $\lambda \in K$ s'appelle une valeur propre de F s'il existe un vecteur $v \in V$, $v \neq 0$, tel que $f(v) = \lambda v$. Le vecteur v s'appelle un vecteur propre de F associé à la valeur propre λ.*

Définition 7.3 *L'ensemble des valeurs propres de A (resp. de F) s'appelle le* ***spectre*** *[spectrum] de A (resp. de f), noté* $\mathsf{spec}(A)$ *(resp.* $\mathsf{spec}(F)$*).*

Exemple 7.4

1. Soit $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$. Alors
   - $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ est un vecteur propre associé à la valeur propre $\lambda_1 = 1$,
   - $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ est un vecteur propre associé à la valeur propre $\lambda_2 = 0$,
   - $v_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ n'est pas un vecteur propre.
2. Soit $A = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$ pour $\phi \in \mathbb{R}$.
   - Si $\phi \neq k\pi$, $k \in \mathbb{N}$, alors A n'a pas de valeur propre (réelle).
   - Si $\phi = (2k+1)\pi$, $k \in \mathbb{N}$, alors $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ a une valeur propre $\lambda = -1$ et tous les vecteurs non-nuls $x \in \mathbb{R}^2$ sont des vecteurs propres associés à λ.
   - Si $\phi = 2k\pi$, $k \in \mathbb{N}$, alors $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ a une valeur propre $\lambda = 1$ et encore tous les vecteurs non-nuls $x \in \mathbb{R}^2$ sont des vecteurs propres associés à λ.

   On va voire que si on considère A comme une matrice complexe, alors on a toujours les valeurs propres $\cos\phi + \mathrm{i}\sin\phi$ et $\cos\phi - \mathrm{i}\sin\phi$.

3. Soit $V = C^\infty(\mathbb{R})$ et $D^2 : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$, $f \mapsto D^2 f := f''$. Pour tout $k \in \mathbb{Z}$ les fonctions $g_k(\xi) = \cos(k\xi)$ et $h_k(\xi) = \sin(k\xi)$ (si $k \neq 0$) sont des vecteurs (fonctions) propres de D^2 associés à la valeur propre $\lambda_k = -k^2$.

♦

Si x est un vecteur propre de A et $\beta \in K \setminus \{0\}$, alors

$$Ax = \lambda x \quad \Rightarrow \quad A(\beta x) = \beta Ax = \beta \lambda x = \lambda(\beta x),$$

c-à-d βx est aussi un vecteur propre. De même pour un vecteur propre v d'un endomorphisme F.

7.2 Le polynôme caractéristique

Soit $A \in M_{n\times n}(K)$ et soit $x \in K^n$ un vecteur propre de A associé à la valeur propre $\lambda \in K$. Alors, puisque $x \neq 0$,

$$\begin{aligned} & Ax = \lambda x \\ \Leftrightarrow\ & (\lambda I - A)x = 0 \\ \Leftrightarrow\ & \lambda I - A \text{ n'est pas inversible} \\ \Leftrightarrow\ & \det(\lambda I - A) = 0. \end{aligned}$$

Donc les valeurs propres sont les zéros de la fonction $t \mapsto \det(tI - A)$ dans K.

La matrice $tI - A$ prend la forme

$$tI_n - A = \begin{pmatrix} t - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & t - a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & \cdots & -a_{n,n-1} & t - a_{nn} \end{pmatrix} \in M_{n\times n}(K[t]).$$

Comme $K[t]$ est un anneau, le déterminant de $tI - A$ est aussi dans $K[t]$.

Définition 7.5 *Le **polynôme caractéristique** [characteristic polynomial] d'une matrice $A \in M_{n\times n}(K)$ est le polynôme $p_A(t) := \det(tI - A)$.*

Pour $n = 1$ le polynôme caractéristique de $A = (a_{11})$ est donné par $p_A(t) = t - a_{11}$. Pour $n = 2$

$$p_A(t) = \det \begin{pmatrix} t - a_{11} & a_{12} \\ a_{21} & t - a_{22} \end{pmatrix} = t^2 - (a_{11} + a_{22})t + (a_{11}a_{22} - a_{12}a_{21}).$$

En général on n'a pas des formules simples pour les coefficients de ce polynôme. On peut néanmoins caractériser quelques termes.

Lemme 7.6 *Soit $A \in M_{n\times n}(K)$. Alors, le polynôme caractéristique p_A de A est de degré n et*

$$p_A(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_{n-1} t^{n-1} + t^n,$$

où

$$\alpha_{n-1} = -(a_{11} + a_{22} + \cdots + a_{nn}), \qquad \alpha_0 = (-1)^n \det(A).$$

DÉMONSTRATION. On définit le **symbole de Kronecker**

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

On a alors

$$p_A(t) = \sum_{\sigma\in S_n} \mathsf{sgn}(\sigma) \prod_{i=1}^{n} \left(\delta_{i,\sigma(i)} t - a_{i,\sigma(i)}\right) \tag{7.1}$$

et donc

$$\alpha_0 = p_A(0) = (-1)^n \sum_{\sigma\in S_n} \mathsf{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} = (-1)^n \det(A).$$

Ensuite on remarque que l'on peut écrire (7.1) comme

$$p_A(t) = \prod_{i=1}^{n} (t - a_{ii}) + \sum_{\substack{\sigma\in S_n \\ \sigma\neq \mathsf{id}}} \mathsf{sgn}(\sigma) \prod_{i=1}^{n} \left(\delta_{i,\sigma(i)} t - a_{i,\sigma(i)}\right), \tag{7.2}$$

où le premier terme prend la forme

$$\prod_{i=1}^{n} (t - a_{ii}) = t^n - (a_{11} + \cdots + a_{nn}) t^{n-1} + \text{polynôme de degré} \leq n-2.$$

Le deuxième terme à droite de (7.2) est un polynôme de degré inférieur à $n-2$, car si $\sigma \neq \mathsf{id}$ il existe au moins un couple i,j tel que $i \neq j$, $\sigma(i) \neq i$ et $\sigma(j) \neq j$. Alors on peut choisir au plus $n-2$ éléments diagonaux de $tI - A$. Donc

$$p_A(t) = t^n - (a_{11} + \cdots + a_{nn}) t^{n-1} + \text{polynôme de degré} \leq n-2.$$

■

La somme des termes diagonaux d'une matrice $A \in M_{n\times n}(K)$ s'appelle la **trace** de A, i.e.,

$$\mathsf{trace}(A) := a_{11} + a_{22} + \cdots + a_{nn}.$$

Le lemme suivant va dans la direction opposé. Etant donné un polynôme unitaire p il trouve une matrice A telle que p est le polynôme caractéristique de A.

Lemme 7.7 *Soit $p \in K[t]$ de la forme $p = \alpha_0 + \alpha_1 t + \cdots + \alpha_{n-1} t^{n-1} + t^n \in K[t]$. Alors p est le polynôme caractéristique de la matrice*

$$A = \begin{pmatrix} 0 & & & -\alpha_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -\alpha_{n-2} \\ & & 1 & -\alpha_{n-1} \end{pmatrix}.$$

DÉMONSTRATION. Exercice. ■

La matrice A du lemme 7.7 s'appelle la **matrice compagnon** [*companion matrice*] de p.

D'après la caractérisation

$$\lambda \text{ v.p. de } A \in M_{n\times n}(K) \quad \Leftrightarrow \quad p_A(\lambda I - A) = 0,$$

pour déterminer le spectre de A il faut déterminer les zéros de son polynôme caractéristique. En général ($K = \mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$), il n'est pas possible de trouver une formule explicite pour les zéros d'un polynôme de degré 5 ou plus (théorie de Galois) et on doit recourir aux méthodes numériques.

Par le théorème fondamental de l'algèbre (voir le théorème 2.46) une matrice complexe de taille $n \times n$ a n valeurs propres (comptées avec leurs multiplicités).

Corollaire 7.8 *Soit $A \in M_{n\times n}(K)$ une matrice triangulaire (inférieure ou supérieure). Alors les valeurs propres de A sont les éléments diagonaux de A.*

DÉMONSTRATION. Découle directement du lemme 6.5 :

$$p_A(t) = \det(tI - A) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn}).$$

■

7.2.1 Théorème de Hamilton-Cayley

Soit K un anneau commutatif et soit $p \in K[t]$ un polynôme à coefficients dans R,

$$p(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_n t_n, \qquad \alpha_0, \ldots, \alpha_n \in R.$$

Soit $A \in M_{n\times n}(K)$. On peut alors évaluer p en A de la manière suivante :

$$p(A) = \alpha_0 I_n + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_n A^n \in M_{n\times n}(R),$$

où

$$A^j = \underbrace{A \cdot A \cdots A}_{j \text{ fois}}.$$

On pose $A^0 := I_n$. Comme le déterminant est defini pour des matrices sur un anneau, on peut définir le polynôme charactéristique de $A \in M_{n\times n}(R)$ comme $p_A(t) = \det(tI_n - A)$ et donc $p_A \in R[t]$.

Théorème 7.9 (Hamilton-Cayley) *Soit $A \in M_{n\times n}(R)$ et p_A le polynôme charactéristique de A. Alors $p_A(A) = 0$.*

DÉMONSTRATION.[9] Pour $n = 1, A = \alpha \in R$ et $p_A(A) = \det(\alpha \cdot 1 - \alpha) = 0$. Soit alors $n \geq 2$. Pour une matrices fixe $A \in M_{n\times n}(R)$ on considère l'ensemble

$$R[A] := \{p(A) : p \in R[t]\} \subset M_{n\times n}(R),$$

contenant tous les polynômes evalués en A. On vérifie que cet ensemble (muni de l'addition et la multiplication matricielle) est un anneau commutatif. On définit

$$\mathscr{A} := \begin{pmatrix} A - a_{11}I_n & -a_{21}I_n & \cdots & -a_{n1}I_n \\ -a_{12}I_n & A - a_{22}I_n & \cdots & -a_{n2}I_n \\ \vdots & \vdots & & \vdots \\ -a_{1n}I_n & -a_{2n}I_n & \cdots & A - a_{nn}I_n \end{pmatrix} \in M_{n\times n}\big(M_{n\times n}(R)\big).$$

Le déterminant de $\mathscr{A}$ (par rapport à l'anneau $R[A]$!) est donné par

$$\det(\mathscr{A}) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \big(\delta_{i,\sigma(i)}A - a_{\sigma(i),i}I_n\big) = p_{A^\mathsf{T}}(A) = p_A(A) \in R[A].$$

Soit $x = \begin{pmatrix} e_1^\mathsf{T} & e_2^\mathsf{T} & \cdots & e_n^\mathsf{T} \end{pmatrix}^\mathsf{T} \in R^{n^2}$. Alors

$$\begin{aligned} \begin{pmatrix} A - a_{11}I_n & -a_{21}I_n & \cdots & -a_{n1}I_n \end{pmatrix} x &= (A - a_{11}I_n)e_1 - a_{21}I_n e_2 + \cdots + -a_{n1}I_n e_n \\ &= Ae_1 - a_{11}e_1 - a_{21}e_2 + \cdots + a_{n1}e_n \\ &= \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} - \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} = 0 \end{aligned}$$

9. [Higham, N. Functions of matrices. SIAM, 2008] (S. 7): *It is incorrect to prove the Cayley-Hamilton theorem by $p_A(A) = \det(A * I - A) = 0$.*

et, ainsi, la première ligne bloc de $\mathscr{A}$ multipliée par x donne 0. De façon similaire, la i-ième ligne bloc de $\mathscr{A}$ multipliée par x donne 0 pour $i = 2,\dots,n$. Alors

$$\mathscr{A}x = 0, \tag{7.3}$$

mais on doit faire attention qu'on voit $\mathscr{A}$ comme élément de $M_{n^2\times n^2}(R)$ (et non de $M_{n\times n}\big(M_{n\times n}(R)\big)$) dans le produit $\mathscr{A}x$. Par le Théorème 6.17, on a la formule

$$\mathsf{com}(\mathscr{A})^{\mathsf{T}}\mathscr{A} = \begin{pmatrix} \det(\mathscr{A}) & & \\ & \ddots & \\ & & \det(\mathscr{A}) \end{pmatrix}.$$

En regardant les deux côtés de cette égalité comme des éléments de $M_{n^2\times n^2}(R)$ et utilisant (7.3) on obtient

$$\begin{pmatrix} \det(\mathscr{A})e_1 \\ \vdots \\ \det(\mathscr{A})e_n \end{pmatrix} = \mathsf{com}(\mathscr{A})^{\mathsf{T}}\,\mathscr{A}x = 0.$$

Alors chaque colonne de la matrice $\det(\mathscr{A})$ est nulle et, ainsi, $p_A(A) = \det(\mathscr{A}) = 0$. ■

Exemple 7.10 Le polynôme caractéristique de $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ est $p_A(t) = (t-1)(t-2)$. On a

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

Pour la matrices $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, on a bien sur que $p_A(A) = 0$ pour $p_A(t) = (t-2)^2$. Mais il existe un polynôme de degré 1, $q(t) = t - 2$, tel que $q(A) = 0$. Le **polynôme minimal** de A est le polynôme unitaire de degré minimum parmi ceux qui annulent A, c-à-d $q(A) = 0$. ♦

Les resultats suivants donnent des utilisations typiques du théorème 7.9.

Corollaire 7.11 *Soit $A \in M_{n\times n}(R)$.*

(i) Toute puissance A^k avec $k \in \mathbb{N}$ peut s'écrire comme une combinaison linéaire des puissances $I, A, A^2, \dots, A^{n-1}$.

(ii) Si A est inversible, alors l'inverse A^{-1} peut s'écrire comme une combinaison linéaire des puissances $I, A, A^2, \dots, A^{n-1}$.

DÉMONSTRATION. (i). Trivialement, l'assertion est vraie pour $k = 0, 1, \dots, n-1$. On montre le cas $k = n$. Par le théorème 7.9 :

$$0 = p_A(A) = \alpha_0 I + \alpha_1 A \cdots + \alpha_{n-1}A^{n-1} + A^n \quad \Rightarrow \quad A^n = -\alpha_0 I - \alpha_1 A \cdots - \alpha_{n-1}A^{n-1}.$$

De façon similaire, on montre le cas $k > n$ par récurrence, utilisant $0 = A^{k-n}p_A(A)$.

(ii). Si A est inversible alors $\alpha_0 = (-1)^n \det(A)$ est inversible. De $0 = p_A(A)$ on obtient que

$$I = -\frac{\alpha_1}{\alpha_0}A \cdots - \frac{\alpha_{n-1}}{\alpha_0}A^{n-1} - \frac{1}{\alpha_0}A^n = A\Big(-\frac{\alpha_1}{\alpha_0}I \cdots - \frac{\alpha_{n-1}}{\alpha_0}A^{n-2} - \frac{1}{\alpha_0}A^{n-1}\Big)$$

et donc $A^{-1} = -\frac{\alpha_1}{\alpha_0}I \cdots - \frac{\alpha_{n-1}}{\alpha_0}A^{n-2} - \frac{1}{\alpha_0}A^{n-1}$. ■

7.2.2 Matrices semblables, endomorphismes

Deux matrices semblables ont le même spectre.

Théorème 7.12 *Soit $A \in M_{n\times n}(K)$ et $P \in M_{n\times n}(K)$ inversible. Alors*

(i) les spectres de A et celui de $P^{-1}AP$ sont identiques,

(ii) $x \in K^n$ est un vecteur propre de A si et seulement si $P^{-1}x$ est un vecteur propre de $P^{-1}AP$.

DÉMONSTRATION. (i). Posons $B = P^{-1}AP$,

$$\begin{aligned} p_B(t) &= \det(tI - B) = \det(tI - P^{-1}AP) = \det\left(P^{-1}(tI-A)P\right) \\ &= \underbrace{\det\left(P^{-1}\right)}_{=1/\det P} \det(tI-A)\det P = \det(tI-A) = p_A(t), \end{aligned}$$

donc les racines de p_B et p_A sont les mêmes et donc les spectres de A et B sont identiques.

(ii). Comme P^{-1} est inversible, on remarque que x est non nul si et seulement si $P^{-1}x$ est non nul.

Si $x \in K^n \setminus \{0\}$ est un vecteur propre de A, il existe $\lambda \in K$ tel que

$$Ax = \lambda x \quad \text{donc} \quad P^{-1}A\underbrace{PP^{-1}}_{=I}x = \lambda P^{-1}x \quad \text{donc} \quad BP^{-1}x = \lambda P^{-1}x.$$

Réciproquement si $P^{-1}x$ est un vecteur propre de $P^{-1}AP$, alors $P^{-1}Ax = \lambda P^{-1}x$ et donc $Ax = \lambda x$. ■

Définition 7.13 *Soit V un K-espace vectoriel de dimension finie et $F \in L(V,V)$. On définit le polynôme caractéristique de f comme le polynôme caractéristique de la matrice de l'application linéaire dans une certaine base.*

D'après le théorème 7.12 la définition 7.13 est bien indépendante du choix de la base : Soient $\mathscr{B}_V$, $\tilde{\mathscr{B}}_V$ deux bases de V. Comme

$$[F]_{\tilde{\mathscr{B}}_V,\tilde{\mathscr{B}}_V} = [I]_{\mathscr{B}_V,\tilde{\mathscr{B}}_V} \cdot [F]_{\mathscr{B}_V,\mathscr{B}_V} \cdot [I]^{-1}_{\mathscr{B}_V,\tilde{\mathscr{B}}_V},$$

$[F]_{\tilde{\mathscr{B}}_V,\tilde{\mathscr{B}}_V}$ et $[F]_{\mathscr{B}_V,\mathscr{B}_V}$ ont le même polynôme caractéristique.

Théorème 7.14 *Soit V un K-espace vectoriel de dimension finie et $F \in L(V,V)$. Soit $\mathscr{B}_V$ une base de V. Alors si $\lambda \in \mathsf{spec}(F)$ associé au vecteur propre $v \in V$ alors $\lambda \in \mathsf{spec}\big([F]_{\mathscr{B}_V,\mathscr{B}_V}\big)$ associé au vecteur propre $[v]_{\mathscr{B}_V} \in K^n$. Réciproquement si $\lambda \in \mathsf{spec}\big([F]_{\mathscr{B}_V,\mathscr{B}_V}\big)$ associé au vecteur propre $x \in K^n$, alors $\lambda \in \mathsf{spec}(F)$ associé au vecteur propre $[x]^{-1}_{\mathscr{B}_V} \in V$.*

DÉMONSTRATION. Découle des équivalences suivantes avec $[v]_{\mathscr{B}_V} = x$:

$$\begin{aligned} & F(v) = \lambda v \\ \Leftrightarrow\ & \left(F \circ [\cdot]^{-1}_{\mathscr{B}_V}\right)(x) = \lambda [x]^{-1}_{\mathscr{B}_V} \\ \Leftrightarrow\ & \left([\cdot]_{\mathscr{B}_V} \circ F \circ [\cdot]^{-1}_{\mathscr{B}_V}\right)(x) = \lambda x \\ \Leftrightarrow\ & [F]_{\mathscr{B}_V,\mathscr{B}_V} x = \lambda x. \end{aligned}$$

■

Exemple 7.15 Soit $V = \mathbb{R}_3[t]$ et $D \in L(V,V)$, $D : p \mapsto p'$. Par l'exemple 5.18 on connaît déjà la matrice de l'application linéaire D par rapport à $\mathscr{B} = \{1,t,t^2,t^3\}$:

$$[D]_{\mathscr{B},\mathscr{B}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

La seule valeur propre de cette matrice est $\lambda = 0$ associé au vecteur propre $x = (1,0,0,0)^\mathsf{T}$ (et à ses multiples scalaires). Par le théorème 7.14, $\lambda = 0$ est la seule valeur propre de D et tout vecteur propre et un polynôme constant non nul. ♦

Définition 7.16 *Soit $A \in M_{n\times n}(K)$ et $\lambda \in \mathsf{spec}(A)$. **L'espace propre** [eigenspace] associé à λ est défini par*

$$E_\lambda(A) = \{x \in K^n \,|\, Ax = \lambda x\}.$$

Soit V un K-espace vectoriel de dimension finie, $F \in L(V,V)$, et $\lambda \in \mathsf{spec}(F)$. L'espace propre associé à λ est défini par

$$E_\lambda(F) = \{v \in V \,|\, F(v) = \lambda v\}.$$

On voit facilement que $E_\lambda(A)$ et $E_\lambda(F)$ sont des sous-espaces vectoriels de K^n et V respectivement. Par définition

$$E_\lambda(A) = \mathrm{Ker}(\lambda I_n - A), \qquad E_\lambda(F) = \mathrm{Ker}(\lambda \cdot \mathsf{id} - F),$$

donc, par le théorème du rang,

$$\dim E_\lambda(A) = n - \mathsf{rang}(\lambda I_n - A), \qquad \dim E_\lambda(F) = \dim V - \mathsf{rang}(\lambda \cdot \mathsf{id} - F).$$

7.3 Diagonalisation

Le but de cette section est de trouver une matrice P inversible telle que $P^{-1}AP$ est diagonale. Ce n'est pas toujours possible, même si $K = \mathbb{C}$. Par exemple, soit

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_{2\times 2}(\mathbb{C}), \quad P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \in M_{2\times 2}(\mathbb{C}), \tag{7.4}$$

où P est inversible. Supposons qu'il existe une matrice diagonale $D = \mathsf{diag}\,(d_{11},d_{22})$ telle que $P^{-1}AP = D$. Alors, $d_{11} = d_{22} = 0$ car D et A ont les mêmes valeurs propres par le lemme 7.12. Mais c'est une contradiction : $0 \neq A = PDP^{-1} = 0$.

Définition 7.17 *Soit $A \in M_{n\times n}(K)$. On dit que A est **diagonalisable** s'il existe une matrice $P \in M_{n\times n}(K)$ inversible telle que $P^{-1}AP$ est diagonal.*

Soit V un K-espace vectoriel de dimension finie et $F \in L(V,V)$. On dit que F est diagonalisable si $[F]_{\mathscr{B}_V,\mathscr{B}_V}$ est diagonalisable, où $\mathscr{B}_V$ est une base quelconque de V.

Si A est diagonalisable les éléments diagonaux de $P^{-1}AP$ sont les valeurs propres de A : $P^{-1}AP = \mathsf{diag}\,(\lambda_1,\lambda_2,\ldots,\lambda_n)$.

Dans ce qui suit on va déduire une caractérisation des matrices diagonalisables. Le résultat suivant joue un rôle clé.

Théorème 7.18 *$A \in M_{n\times n}(K)$ est diagonalisable si et seulement s'il existe une base de K^n formée de vecteurs propres de A.*

DÉMONSTRATION. Supposons A est diagonalisable, alors il existe $P \in M_{n\times n}(K)$ inversible telle que $P^{-1}AP = \mathrm{diag}\,(\lambda_1,\lambda_2,\dots,\lambda_n)$. En notant $P = \big(x_1,x_2,\dots,x_n\big)$, cette rélation est équivalente à

$$\begin{aligned} & AP = P\cdot \mathrm{diag}\,(\lambda_1,\lambda_2,\dots,\lambda_n) \\ \Leftrightarrow\ & A\big(x_1,x_2,\dots,x_n\big) = \big(x_1,x_2,\dots,x_n\big)\cdot \mathrm{diag}\,(\lambda_1,\lambda_2,\dots,\lambda_n) \\ \Leftrightarrow\ & Ax_1 = \lambda_1 x_1,\, Ax_2 = \lambda_2 x_2,\ \dots,\, Ax_n = \lambda_n x_n. \end{aligned}$$

Comme $x_1,x_2,\dots,x_n$ est une famille linéairement indépendante, on a bien une base de vecteurs propres de K^n.

Réciproquement, supposons qu'il existe une base de K^n $\{x_1,x_2,\dots,x_n\}$ constituée de vecteurs propres de A, i.e., $Ax_i = \lambda_i x_i$ pour $i = 1,\dots,n$. Alors on pose $P = \big(x_1,x_2,\dots,x_n\big)$. P est inversible et, par les équivalences ci-dessus, on obtient que $P^{-1}AP = \mathrm{diag}\,(\lambda_1,\lambda_2,\dots,\lambda_n)$. ■

Lemme 7.19 *Soit $A \in M_{n\times n}(K)$ et $\lambda_1,\dots,\lambda_m \in K$ des valeurs propres distinctes de A associées aux vecteurs propres $x_1,\dots,x_m \in K^n$. Alors $(x_1,\dots,x_m)$ est une famille libre de K^n.*

DÉMONSTRATION. Par récurrence sur m. Pour $m = 1$ le résultat est vrai car un vecteur propre est non nul. Soit alors $m \geq 2$ et supposons l'assertion vraie pour $m-1$. Considérons une combinaison linéaire

$$0 = \alpha_1 x_1 + \dots + \alpha_{m-1}x_{m-1} + \alpha_m x_m, \quad \alpha_1,\dots,\alpha_m \in K. \tag{7.5}$$

Alors

$$0 = \alpha_1 Ax_1 + \dots + \alpha_{m-1}Ax_{m-1} + \alpha_m Ax_m = \alpha_1\lambda_1 x_1 + \dots + \alpha_{m-1}\lambda_{m-1}x_{m-1} + \alpha_m\lambda_m x_m.$$

En ajoutant la relation (7.5) multipliée par $-\lambda_m$

$$0 = \alpha_1(\lambda_1 - \lambda_m)x_1 + \dots + \alpha_{m-1}(\lambda_{m-1} - \lambda_m)x_{m-1}.$$

Par hypothèse de récurrence, $(x_1,\dots,x_{m-1})$ est libre et donc $\alpha_i(\lambda_i - \lambda_m) = 0$ pour $i = 1,\dots,m-1$. Comme $\lambda_i \neq \lambda_m$ on obtient

$$\alpha_1 = \dots = \alpha_{m-1} = 0.$$

Par suite comme $x_m \neq 0$, l'équation (7.5) montre que $\alpha_m = 0$. ■

La même démonstration montre aussi le résultat suivant : Soit V un K-espace vectoriel de dimension finie, $f \in L(V,V)$ et $\lambda_1,\dots,\lambda_r \in K$ des valeurs propres distinctes de f associées aux vecteurs propres $v_1,\dots,v_m \in V$. Alors $(v_1,\dots,v_m)$ est une famille libre de V.

Corollaire 7.20 *Soit $A \in M_{n\times n}(K)$ et $\lambda_1,\dots,\lambda_r$ les valeurs propres distinctes de A. Alors A est diagonalisable si et seulement si*

$$E_{\lambda_1}(A) \oplus E_{\lambda_2}(A) \oplus \dots \oplus E_{\lambda_r}(A) = K^n.$$

DÉMONSTRATION. Par le théorème 4.38, $E_{\lambda_1}(A) \oplus \dots \oplus E_{\lambda_r}(A) = K^n$ si et seulement si il existe une base de K^n qui est la réunion des bases de $E_{\lambda_i}(A)$, $i = 1,\dots,r$. Alors l'assertion découle du théorème 7.18. ■

Corollaire 7.21 *Soit $A \in M_{n\times n}(K)$ et supposons que A possède n valeurs propres distinctes. Alors A est diagonalisable.*

DÉMONSTRATION. Découle directement du lemme 7.19 et du théorème 7.18. ■

Avoir n valeurs propres distinctes est une condition suffisante pour être diagonalisable mais *pas* nécéssaire. Par exemple, la matrice identité I_n possède une seule valeur propre et elle est évidemment diagonalisable.

Soit $A \in M_{n\times n}(K)$ et $p_A(t) = \det(tI_n - A)$. Pour $\lambda \in \mathrm{spec}(A)$ on a $p_A(\lambda_i) = 0$, alors on peut écrire

$$p_A(t) = (t-\lambda)^m g(t) \quad \text{avec} \quad g(\lambda) \neq 0.$$

L'entier m s'appelle la multiplicité du zéro λ de p_A.

Définition 7.22 *La **multiplicité algébrique** d'une valeur propre $\lambda \in \mathrm{spec}(A)$ est la multiplicité de λ comme zéro de p_A. On la note $m_{\text{alg}}(\lambda)$.*

*La **multiplicité géometrique** d'une valeur propre $\lambda \in \mathrm{spec}(A)$ est la dimension de $E_\lambda(A)$. On la note $m_{\text{géom}}(\lambda)$.*

On a les mêmes définitions pour un endomorphisme $f \in L(V,V)$.

Exemple 7.23 Soient

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}), \quad B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}),$$

$$C = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}).$$

Pour la valeur propre 2 de A, on a $m_{\text{alg}}(2) = 3$, $m_{\text{géom}}(2) = 1$. Pour la valeur propre 2 de B, on a $m_{\text{alg}}(2) = 3$, $m_{\text{géom}}(2) = 2$. Pour la valeur propre 2 de C, on a $m_{\text{alg}}(2) = 3$, $m_{\text{géom}}(2) = 3$. ♦

Lemme 7.24 *Soit $\lambda \in \mathrm{spec}(A)$. Alors $m_{\text{géom}}(\lambda) \leq m_{\text{alg}}(\lambda)$.*

DÉMONSTRATION. Soit $m = m_{\text{géom}}(\lambda)$ et $(y_1, \dots, y_m) \subset K^n$ une base de $E_\lambda(A)$. On complète cette base en une base de K^n : $(y_1, \dots, y_m, y_{m+1}, \dots, y_n)$. On pose la matrice inversible $Y = \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix}$. Alors

$$\begin{aligned} AY &= A\begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} = \begin{pmatrix} Ay_1 & \cdots & Ay_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda y_1 & \cdots & \lambda y_m & Ay_{m+1} & \cdots & Ay_n \end{pmatrix} \end{aligned}$$

et par suite

$$P^{-1}AP = \begin{pmatrix} \lambda I_m & B_{12} \\ 0 & B_{22} \end{pmatrix}.$$

Par le théorème 7.12 et le corollaire 6.15, on obtient que

$$\det(tI - A) = \det(tI_m - \lambda I_m)\det(tI - B_{22}) = (t-\lambda)^m \det(tI - B_{22}).$$

Donc $m_{\text{alg}}(\lambda) \geq m$. ■

Théorème 7.25 (théorème de diagonalisation) *Soit $A \in K^{n\times n}$. Alors A est diagonalisable si et seulement si*

(i) p_A est scindé (sur K), et

(ii) $m_{\text{alg}}(\lambda) = m_{\text{géom}}(\lambda)$ pour tout $\lambda \in \mathrm{spec}(A)$.

DÉMONSTRATION. Supposons A diagonalisable, alors il existe P inversible telle que $P^{-1}AP$ est diagonale. En regroupant les valeurs propres identiques on peut supposer que

$$P^{-1}AP = \mathrm{diag}\,\big(\underbrace{\lambda_1,\dots,\lambda_1}_{m_{\mathrm{alg}}(\lambda_1)\text{ fois}},\underbrace{\lambda_2,\dots,\lambda_2}_{m_{\mathrm{alg}}(\lambda_2)\text{ fois}},\dots,\underbrace{\lambda_r,\dots,\lambda_r}_{m_{\mathrm{alg}}(\lambda_r)\text{ fois}}\big), \qquad \lambda_i \neq \lambda_j \text{ pour } i \neq j.$$

Alors

$$m_{\text{géom}}(\lambda_i) = \dim \mathrm{Ker}(\lambda_i I - A) = \dim P\cdot \mathrm{Ker}(\lambda_i - P^{-1}AP) = \dim \mathrm{Ker}(\lambda_i - P^{-1}AP) = m_{\mathrm{alg}}(\lambda_i).$$

Réciproquement, supposons que (i) et (ii) sont vraies, alors

$$p_A(t) = \prod_{i=1}^{r}(t-\lambda_i)^{m_{\mathrm{alg}}(\lambda_i)}.$$

Posons $W := E_{\lambda_1}(A)+\cdots+E_{\lambda_r}(A)$, on obtient $W := E_{\lambda_1}(A)\oplus\cdots\oplus E_{\lambda_r}(A)$ par le lemme 7.19. Comme $m_{\mathrm{alg}}(\lambda_i) = m_{\text{géom}}(\lambda_i)$ on a

$$\dim W = \sum_{i=1}^{r}\dim E_{\lambda_i}(A) = \sum_{i=1}^{r} m_{\text{géom}}(\lambda_i) = \sum_{i=1}^{r} m_{\mathrm{alg}}(\lambda_i) = n,$$

et donc $W = K^n$. Par le corollaire 7.20, ceci implique que A est diagonalisable. ■

Le resultat (et la preuve) du théorème 7.25 est anlogue pour $F \in L(V,V)$, V un K-espace vectoriel de dimension finie.

Exemple 7.26 (i). Soit

$$A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}).$$

On calcule $p_A(t) = \det(tI-A) = \cdots = (\lambda-1)(\lambda+2)^2$, alors les valeurs propres sont $\lambda_1 = 1$, $\lambda_2 = -2$, et p_A est scindé. Par le lemme 7.24, $m_{\text{géom}}(1) = m_{\mathrm{alg}}(1) = 1$. $m_{\mathrm{alg}}(-2) = 2$ et il reste à calculer $m_{\text{géom}}(-2)$. On cherche

$$Ax = -2x \quad\Leftrightarrow\quad (-2I-A)x = 0 \quad\Leftrightarrow\quad \begin{pmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$\Leftrightarrow \quad x_1+x_2+x_3 = 0,$$

c-à-d deux variables sont libres et $m_{\text{géom}}(-2) = \dim \mathrm{Ker}(-2I-A) = 2 = m_{\mathrm{alg}}(-2)$. Alors A est diagonalisable.

Si on veut calculer une matrice P telle que $P^{-1}AP$ soit sous forme diagonale, on calcule une base des espaces propres :

$$Ax = \lambda_1 x \quad\Leftrightarrow\quad \begin{pmatrix} 0 & -3 & -3 \\ 3 & 6 & 3 \\ -3 & -3 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0.$$

Ceci donne $x_1 = x_3 = -x_2$. Par exemple, on peut choisir $x_2 = -1$, $x_1 = x_3 = 1$, et donc $E_1(A) = \mathrm{span}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. Par le calcul ci-dessus, $E_{-2}(A) = \mathrm{span}\left(\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right)$. Finalement, en posant

$$P = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \text{on a bien} \quad P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

(ii). Soit

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad \det(t-A) = (t-1)(t-2)^2.$$

Les valeurs propres sont $\lambda_1 = 1$, $\lambda_2 = 2$, et $m_{\text{géom}}(1) = m_{\text{alg}}(1) = 1$, $m_{\text{alg}}(2) = 2$. Pour la multiplicité géometrique de 2 on calcule

$$Ax = 2x \quad \Leftrightarrow \quad (2I-A)x = 0 \quad \Leftrightarrow \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \quad \Leftrightarrow \quad \begin{array}{l} x_1 = 0 \\ x_2 \text{ libre} \\ x_3 = 0. \end{array}$$

Alors $m_{\text{géom}}(2) = \dim \operatorname{Ker}(2I-A) = 1 \neq m_{\text{alg}}(2)$ et donc A n'est pas diagonalisable. ♦

Algorithme pour la diagonalisation *à la main* de $A \in M_{n\times n}(K)$:

1. On calcule $p_A(\lambda)$ et on cherche les racines de ce polynôme.
2. Si on trouve n racines (comptées avec leurs multiplicités) on va à l'étape 3, sinon A n'est pas diagonalisable.
3. Soient $\lambda_1, \dots, \lambda_r$ les racines distinctes du polynôme scindé $p_A(t) = \prod_{i=1}^r (t-\lambda_i)^{m_i}$m où $m_i = m_{\text{alg}}(\lambda_i)$. On calcule de bases de $E_{\lambda_i}(A) = \operatorname{Ker}(\lambda_i I - A)$ et $m_{\text{géom}}(\lambda_i) = \dim E_{\lambda_i}(A)$. Si $m_{\text{géom}}(\lambda_i) < m_{\text{alg}}(\lambda_i)$ pour un $i \in \{1, \dots, n\}$, alors A n'est pas diagonalisable.
4. En posant

$$P = \big(\underbrace{p_1, \dots, p_{m_1}}_{\text{base de } E_{\lambda_1}(A)}, \underbrace{p_{m_1+1}, \dots, p_{m_1+m_2}}_{\text{base de } E_{\lambda_2}(A)}, \dots, \underbrace{p_{n-m_r+1}, \dots, p_n}_{\text{base de } E_{\lambda_r}(A)} \big),$$

on obtient que P est inversible et

$$P^{-1}AP = \mathsf{diag}\,\big(\underbrace{\lambda_1, \dots, \lambda_1}_{m_{\text{alg}}(\lambda_1) \text{ fois}}, \underbrace{\lambda_2, \dots, \lambda_2}_{m_{\text{alg}}(\lambda_2) \text{ fois}}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{m_{\text{alg}}(\lambda_r) \text{ fois}} \big)$$

Le même algorithme s'applique par la diagonalisation de $F \in L(V,V)$, V de dimension finie. D'abord on choisit une base $\mathscr{B}_V$ de V et on calcule $A = [F]_{\mathscr{B}_V, \mathscr{B}_V}$. Ensuite, on applique l'algorithme ci-dessus à A.

La diagonalisation de matrices (lorsque c'est possible) est un outil puissant. La caractérisation suivante de la commutativité de matrices donne un premièr aperçu de cette puissance.

Lemme 7.27 *Soient $A, B \in M_{n\times n}(K)$ des matrices diagonalisable. Alors $AB = BA$ si et seulement s'il existe $P \in M_{n\times n}(K)$ inversible telle que $P^{-1}AP$ et $P^{-1}BP$ soient diagonales (on dit dans ce cas que A et B sont* ***simultanément diagonalisables****).*

DÉMONSTRATION. (i). Supposons qu'il existe $P \in M_{n\times n}(K)$ inversible telle que $P^{-1}AP = \Lambda_A$ et $P^{-1}BP = \Lambda_B$, où Λ_A, Λ_B sont diagonales. Alors

$$AB = P\Lambda_A P^{-1} P\Lambda_B P^{-1} = P\Lambda_A\Lambda_B P^{-1} = P\Lambda_B\Lambda_A P^{-1} = P\Lambda_B P^{-1} P\Lambda_A P^{-1} = BA,$$

où l'on utilisé que deux matrices diagonales commutent.

(ii). Supposons que A et B soient diagonalisables et $AB = BA$. Alors il existe $\tilde{P}$ inversible telle que

$$\tilde{P}^{-1}A\tilde{P} = \begin{pmatrix} \lambda_1 I_{n_1} & & \\ & \ddots & \\ & & \lambda_k I_{n_k} \end{pmatrix} =: \Lambda_A, \quad \text{avec } \lambda_i \neq \lambda_j \text{ pour } i \neq j.$$

Alors comme $\tilde{P}\Lambda_A\tilde{P}^{-1}B = AB = BA = B\tilde{P}\Lambda_A\tilde{P}^{-1}$ on obtient

$$\Lambda_A\tilde{B} = \tilde{B}\Lambda_A, \quad \text{avec} \quad \tilde{B} = \tilde{P}^{-1}B\tilde{P}. \tag{7.6}$$

On écrit $\tilde{B}$ sous forme partitionée $\tilde{B} = \left(B_{ij}\right)_{i,j=1}^{k}$ avec $B_{ij} \in M_{n_i\times n_j}(K)$. Alors d'après (7.6)

$$\begin{pmatrix} \lambda_1 B_{11} & \cdots & \lambda_1 B_{1k} \\ \vdots & & \vdots \\ \lambda_k B_{k1} & \cdots & \lambda_k B_{kk} \end{pmatrix} = \begin{pmatrix} \lambda_1 B_{11} & \cdots & \lambda_k B_{1k} \\ \vdots & & \vdots \\ \lambda_1 B_{k1} & \cdots & \lambda_k B_{kk} \end{pmatrix},$$

donc comme $\lambda_i \neq \lambda_j$ pour $i \neq j$ on obtient $B_{ij} = 0$ pour $i \neq j$. Donc $\tilde{B}$ est une matrice diagonale par bloc. Comme B est diagonalisable, $\tilde{B} = \tilde{P}^{-1}B\tilde{P}$ est diagonalisable et donc les blocs diagonales B_{ii}, $i = 1,\ldots,r$, sont diagonalisables. Soit $P_i \in M_{n_i\times n_i}(K)$ inversible telle que $P_i^{-1}B_{ii}P_i$ soit diagonale pour $i = 1,\ldots,k$. On pose

$$P := \tilde{P}\begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_k \end{pmatrix}.$$

Alors $P^{-1}BP$ est diagonale par construction et

$$P^{-1}AP = \begin{pmatrix} P_1^{-1} & & \\ & \ddots & \\ & & P_k^{-1} \end{pmatrix}\tilde{P}^{-1}A\tilde{P}\begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_k \end{pmatrix} = \Lambda_A.$$

■

7.4 Dynamique discrète

Dans la section 1.5.2 on a vu la suite de Fibonacci

$$f_0 = 1, \qquad f_1 = 1, \qquad f_{k+2} = f_{k+1} + f_k, \quad k \geq 0, \tag{7.7}$$

qui peut s'écrire comme produit matrice-vecteur. En général, on considère une **suite récurrente linéaire d'ordre n** :

$$f_{k+n} = \alpha_{n-1}f_{k+n-1} + \cdots + \alpha_1 f_{k+1} + \alpha_0 f_k \tag{7.8}$$

où $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ sont des scalaries fixés d'un corps K. Pour définir une suite à partir de (7.8) il faut fixer n conditions initiales

$$f_0 = s_0, \quad f_1 = s_1, \quad \ldots, \quad f_{n-1} = s_{n-1}, \tag{7.9}$$

où $s_0, s_1, \ldots, s_{n-1} \in K$ sont des scalaries. En définissant

$$u_k := \begin{pmatrix} f_k \\ f_{k+1} \\ \vdots \\ f_{k+n-1} \end{pmatrix},$$

on voit que (7.8)–(7.9) est équivalente à

$$u_{k+1} = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ \alpha_0 & \cdots & \alpha_{n-2} & \alpha_{n-1} \end{pmatrix} u_k, \qquad u_0 = \begin{pmatrix} s_0 \\ s_1 \\ \vdots \\ s_{n-1} \end{pmatrix}. \tag{7.10}$$

À part les signes des coefficients α_i, la matrice dans (7.10) est la transposée de la matrice compagnon du lemme 7.7 !

Lemme 7.28 *Soit $A \in M_{n\times n}(K)$ une matrice comme dans (7.10) avec $\alpha_0, \alpha_1, \ldots, \alpha_{n-1} \in K$. Alors :*

(i) le polynôme caractéristique de A est $p_A = t^n - \alpha_{n-1}t^{n-1} - \cdots - \alpha_1 t - \alpha_0$.

(ii) Si $\lambda \in K$ est une valeur propre de A alors $\big(1, \lambda, \lambda^2, \ldots, \lambda^{n-1}\big)^{\mathsf{T}}$ est un vecteur propre associé.

(iii) A est diagonalisable si et seulement si A possède n valeurs propres distinctes.

DÉMONSTRATION. (i) découle du lemme 7.7.

(ii).

$$\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ \alpha_0 & \cdots & \alpha_{n-2} & \alpha_{n-1} \end{pmatrix} \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{pmatrix} = \begin{pmatrix} \lambda \\ \vdots \\ \lambda^{n-1} \\ \lambda^n - p_A(\lambda) \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ \vdots \\ \lambda^{n-2} \\ \lambda^{n-1} \end{pmatrix}.$$

(iii). Par le corollaire 7.21, une matrice avec n valeurs propres distinctes est diagonalisable. Réciproquement, supposons que A soit diagonalisable. Trivialement, les $n-1$ dernières colonnes de la matrice $tI - A$ sont toujours linéairement indépendantes. Si $\lambda \in K$ est une valeur propre on obtient alors que $m_{\text{géom}}(\lambda) = \dim \operatorname{Ker}(\lambda I - A) = 1$. Par le théorème 7.25, $m_{\text{alg}}(\lambda) = m_{\text{géom}}(\lambda) = 1$. Alors les n valeurs propres de A sont distinctes. ■

En notant A la matrice $n \times n$ dans (7.10) supposons que A possède n valeur propres distinctes $\lambda_1, \ldots, \lambda_n \in K$. Alors la matrice des vecteurs propres associés,

$$P = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}, \tag{7.11}$$

est inversible.

En général un **système dynamique linéaire discret** (homogène) prend la forme

$$u_{k+1} = Au_k, \qquad u_0 = s, \tag{7.12}$$

où $A \in M_{n\times n}(K)$, $u_k \in K^n$, et $s \in K^n$ est le vecteur des conditions initiales. Si A est diagonalisable, on trouve une formule explicite des vecteurs $u_k = A^k s$ définies par (7.12). Soit $P = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \in M_{n\times n}(K)$ inversible tel que

$$P^{-1}AP = \operatorname{diag}\big(\lambda_1, \ldots, \lambda_n\big) =: \Lambda.$$

Alors

$$A^k = \big(P\Lambda P^{-1}\big)\big(P\Lambda P^{-1}\big)\cdots\big(P\Lambda P^{-1}\big) = P\Lambda^k P^{-1}.$$

En posant $c := P^{-1}u_0$ la solution de (7.12) peut s'écrire comme suit :

$$u_k = P\Lambda^k c = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = c_1\lambda_1^k x_1 + \cdots + c_n\lambda_n^k x_n. \quad (7.13)$$

On note que le vecteur c représente les coordonnées des conditions initiales s par rapport à la base des vecteurs propres (les colonnes de P). Si s est un vecteur propre x_j (ou un multiple) on obtient la solution particulière $\lambda_j^k x_j$. D'après (7.13) on a que u_k est en fait une combinaison linéaire de telles solutions.

Exemple 7.29 On considère le système associé à la suite de Fibonacci (7.7) :

$$u_{k+1} = Au_k, \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \qquad u_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Les valeurs propres de A sont données par

$$0 = \det(\lambda I - A) = \lambda^2 - \lambda - 1 \quad \Rightarrow \quad \lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_2 = \frac{1-\sqrt{5}}{2},$$

et d'après (7.11)

$$P = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \quad \Rightarrow \quad c = P^{-1}u_0 = \frac{1}{\lambda_1 - \lambda_2}\begin{pmatrix} 1-\lambda_2 \\ \lambda_1 - 1 \end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix} \lambda_1 \\ -\lambda_2 \end{pmatrix}.$$

Par (7.13) on obtient

$$u_k = \frac{\lambda_1}{\sqrt{5}}\lambda_1^k \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} - \frac{\lambda_2}{\sqrt{5}}\lambda_2^k \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}.$$

Le premier coefficient donne la formule explicite

$$f_k = \frac{1}{\sqrt{5}}(\lambda_1^{k+1} - \lambda_2^{k+1}) = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{k+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k+1}\right).$$

♦

Dans ce qui suit on analyse le comportement asymptotiqe de la solution u_k de (7.12) lorsque $k \to \infty$.

Définition 7.30 *Le* ***rayon spectral*** *[spectral radius]* $\rho(A)$ *d'une matrice* $A \in \mathbb{C}^{n\times n}$ *est définie par*

$$\rho(A) := \max\big\{|\lambda| : \lambda \in \mathbb{C} \text{ est une valeur propre de } A\big\}.$$

Supposons que A soit diagonalisable. On ordonne les valeurs propres $\lambda_1, \dots, \lambda_n$ comme suit :

$$\rho(A) = |\lambda_1| = \cdots = |\lambda_m| > |\lambda_{m+1}| \geq \cdots \geq |\lambda_n|. \quad (7.14)$$

Par (7.13),

$$u_k = \rho(A)^k \left(c_1 \frac{\lambda_1^k}{\rho(A)^k}x_1 + \cdots + c_m\frac{\lambda_m^k}{\rho(A)^k}x_m + c_{m+1}\frac{\lambda_{m+1}^k}{\rho(A)^k}x_{m+1} + \cdots + c_n\frac{\lambda_n^k}{\rho(A)^k}x_n\right). \quad (7.15)$$

Les valeur absolues des m premiers termes de cette somme sont constantes (par rapport à k). Les $n-m$ derniers termes convergent vers 0 quand $k \to \infty$. Ainsi la convergence de u_k est determinée par $\rho(A)$.

Lemme 7.31 *Soit* $A \in M_{n\times n}(\mathbb{C})$ *diagonalisable. Alors les vecteurs* u_k *définis par* (7.12) *convergent vers* 0 *pour toute condition initiale* u_0 *si et seulement si* $\rho(A) < 1$.

L'assertion du lemme (7.31) est aussi vraie si A est non diagonalisable; mais il nous manque les outils pour prouver ce fait. Si $\rho(A) > 1$ la formule (7.15) implique que les vecteur u_k divergent (sauf si u_0 est tel que $c_1 = \cdots = c_m = 0$). Dans le cas limite $\rho(A) = 1$ le vecteur u_k peut converger ou diverger. En particulier, si

$$1 = \lambda_1 = \cdots = \lambda_m \tag{7.16}$$

alors

$$u_k \overset{k\to\infty}{\to} u_\infty = c_1 x_1 + \cdots + c_m x_m, \tag{7.17}$$

à condition que A soit diagonalisable. La situation est plus compliquée si A n'est pas diagonalisable. Par exemple,

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad A^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix},$$

ce qui implique que la suite $u_k = A^k u_0$ peut être non bornée.

7.5 Matrices stochastiques

Soit $P \in M_{n\times n}(\mathbb{R})$ une matrice telle que $p_{ij} \geq 0$ pour $i,j = 1,\ldots,n$ et $\sum_{i=1}^n p_{ij} = 1$ pour $j = 1,\ldots,n$. De telles matrices sont appelées **matrices stochastiques** (colonnes). Ces matrices jouent un rôle important en la théorie de probabilité. Soit $\{X_k\}_{k\geq 0}$ une suite de variables aléatoires discrètes définies sur l'ensemble des éventualités $E = \{1,2,\ldots,n\}$. Alors la **matrice de transition de probabilité** $P^{(k)} \in M_{n\times n}(\mathbb{R})$, donnée par

$$\left(P^{(k)}\right)_{ij} = \mathbb{P}(X_{k+1} = i \,|\, X_k = j),$$

est une matrice stochastique.

Une matrice $A \in M_{n\times n}(\mathbb{R})$ est dite **strictement positive**, notée $A > 0$, si $a_{ij} > 0$ pour tous $i,j = 1,\ldots,n$. Pour deux matrices $A,B \in M_{n\times n}(\mathbb{R})$ on dit que $A > B$ si $A - B > 0$. Pour un vecteur $x \in \mathbb{R}^n$ on dit que x est (strictement) positif, noté $x \geq 0$ ($x > 0$), si $x_i \geq 0$ ($x_i > 0$) pour tous $i = 1,\ldots,n$.

Théorème 7.32 (Perron) *Soit $A \in \mathbb{R}^{n\times n}$ une matrice positive. Alors :*

(i) Le rayon spectral $r = \rho(A)$ est strictement positif.

(ii) r est une valeur propre de A.

(iii) r est la seule valeur propre[10] *de A de module r.*

(iv) $m_{\text{alg}}(r) = 1$.

(v) Il y a un vecteur propre x associé à r tel que x est réel et strictement positif.

(vi) À part les multiples positives de x, il n'y a plus des vecteurs propres de A qui sont réels et positifs.

DÉMONSTRATION. On fait la preuve à condition que A soit diagonalisable. Les résultats du théorème sont vrais si A n'est pas diagonalisable mais il nous manque les moyens d'en prouver.

(i). Raisonnement par l'absurde : Soit $r = 0$. Alors toutes les valeurs propres de A sont nulles et le polynôme caractéristique est $p_A = t^n$. Par le théorème 7.9 on a que $A^n = 0$, mais ceci contredit que A^n est positive.

Dans ce qui suit on peut supposer que $r = 1$, en remplaçant A par $\frac{1}{\rho(A)}A$.

10. On regarde A comme une matrice complexe et prend toutes les valeurs propres complexes ou réelles.

(ii) et (v). Comme $r = \rho(A) = 1$, il y a une valeur propre $\lambda \in \mathbb{C}$ telle que $|\lambda| = 1$. Soit $y \in \mathbb{C}^{n\times n}$ le vecteur propre associé. Pour une matrice $C \in \mathbb{C}^{m\times n}$ on note $|C|$ la matrice dont les coefficients sont $(|C|)_{ij} := |c_{ij}|$. Par l'inégalité triangulaire, on obtient que

$$|y| = |\lambda y| = |Ay| \leq |A| \cdot |y| = A|y|. \tag{7.18}$$

Alors, en posant $z := A|y|$ le vecteur $b := z - |y|$ est positif. Supposons que $b \neq 0$, c-à-d au moins un coefficient de b est strictement positif. Alors $Ab > 0$. Comme $z = A|y| > 0$, il existe $\varepsilon > 0$ tel que $Ab > \varepsilon z$. En substituant $b = z - |y|$ on obtient

$$Bz > z, \quad \text{avec} \quad B := \frac{1}{1+\varepsilon} A.$$

Alors $B^k z > z$ pour tout $k \geq 0$. Mais, c'est une contradiction car $\rho(B) = \rho(A)/(1+\varepsilon) = 1/(1+\varepsilon) < 1$ et ainsi $B^k z \to 0$ par le lemme 7.31. Alors $b \neq 0$ et l'inegalité (7.18) est en fait une egalité. Ceci montre que $x = |y|$ est un vecteur propre réel et positif associé à la valeur propre $r = 1$. En outre, $x = |y| = A|y| > 0$.

(iii). Il reste à montrer que toute valeur propre λ de module 1 est en fait égale à 1. Soit $Ay = \lambda y$ alors $|y| = A|y| > 0$ et

$$\Big| \sum_{j=1}^{n} a_{ij} y_j \Big| = |y_i| = \sum_{j=1}^{n} |a_{ij} y_j|.$$

L'égalité implique que tous les termes de la somme de gauche ont le même signe. Alors il existe un vecteur $p = (1, p_2, \dots, p_n)^{\mathsf{T}}$ tel que $p > 0$ et $y = y_1 p$. De $Ay = \lambda y$, on obtient que

$$\lambda p = Ap = |Ap| = |\lambda p| = p \quad \Rightarrow \quad \lambda = 1.$$

(iv). Comme $\rho(A^{\mathsf{T}}) = \rho(A) = 1$ il existe $w \in \mathbb{R}^{n\times 1}$ tel que $w > 0$ et $A^{\mathsf{T}} w = w$. En multipliant w par un scalaire on peut supposer que $w^{\mathsf{T}} x = 1$. Soit $X_\perp \in \mathbb{R}^{n\times(n-1)}$ telle que les colonnes de $X_\perp$ forment une base de $\mathrm{Ker}(w^{\mathsf{T}})$. Comme $w^{\mathsf{T}} x \neq 0$ on a que $x \notin \mathrm{Ker}(w^\perp)$. Alors la matrice $P = (x, X_\perp)$ est inversible est l'inverse est de la forme $P^{-1} = \begin{pmatrix} w^{\mathsf{T}} \\ W_\perp^{\mathsf{T}} \end{pmatrix}$ avec une matrice $W_\perp \in \mathbb{R}^{n\times(n-1)}$ telle que $W_\perp^{\mathsf{T}} x = 0$. Donc

$$P^{-1}AP = \begin{pmatrix} w^{\mathsf{T}}Ax & w^{\mathsf{T}}AX_\perp \\ W_\perp^{\mathsf{T}}Ax & W_\perp^{\mathsf{T}}AX_\perp \end{pmatrix} = \begin{pmatrix} w^{\mathsf{T}}x & w^{\mathsf{T}}X_\perp \\ W_\perp^{\mathsf{T}}x & W_\perp^{\mathsf{T}}AX_\perp \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & W_\perp^{\mathsf{T}}AX_\perp \end{pmatrix}.$$

Supposons que $m_{\mathrm{alg}}(1) > 1$ alors $W_\perp^{\mathsf{T}}AX_\perp$ possède un valeur propre 1 avec un vecteur propre $\tilde{y}_2$. Alors A possède deux vecteurs propres x et $\tilde{x} = P^{-1}\begin{pmatrix} 1 \\ \tilde{y}_2 \end{pmatrix}$ qui sont linéairement indépendants. Choisissons i tel que $\tilde{x}_i \neq 0$. Le vecteur $y = x - \frac{x_i}{\tilde{x}_i}\tilde{x}$ est non nul (grâce à l'indépendance linéaire de x et $\tilde{x}$) et un vecteur propre associé à 1. Par la démonstration de la partie (ii), ceci implique que $|y|$ est un vecteur propre : $|y| = A|y| > 0$. Mais, c'est une contradiction : le i-ième coefficient de y est nul par construction. Alors $m_{\mathrm{alg}}(1) = 1$.

(vi). Soit (encore) $w > 0$ tel que $w^{\mathsf{T}}A = w$ et $w^{\mathsf{T}}x = 1$. Si $\tilde{x} \geq 0$ est un vecteur propre de A associé à une valeur propre λ, $A\tilde{x} = \lambda\tilde{x}$, on obtient que $\lambda = \lambda w^{\mathsf{T}}\tilde{x} = w^{\mathsf{T}}A\tilde{x} = w^{\mathsf{T}}\tilde{x} = 1$. D'après (iv), $m_{\mathrm{géom}}(1) = m_{\mathrm{alg}}(1) = 1$ et ainsi $\tilde{x}$ est un multiple positif de x. ∎

Corollaire 7.33 *Une matrice stochastique $P \in M_{n\times n}(\mathbb{R})$ a une valeur propre égale à 1 et $\rho(P) = 1$. Si, en plus, P est strictement positive alors $m_{\mathrm{alg}}(1) = 1$ il existe un vecteur propre strictement positif associé à 1.*

DÉMONSTRATION. Soit $e = (1, \dots, 1)^\mathsf{T}$ alors $P^\mathsf{T} e = e$ donc $1 \in \mathsf{spec}(P^\mathsf{T}) = \mathsf{spec}(P)$. Soit $\lambda \in \mathbb{C}$ une valeur propre de A^T et soit y un vecteur propre associé. Soit j tel que $y_j \neq 0$. Par $A^\mathsf{T} y = \lambda y$ on obtient que

$$|\lambda||y_j| = \left|\sum_{i=1}^{n} a_{ij} y_i\right| \leq \sum_{i=1}^{n} a_{ij}|y_i| \leq |y_j|$$

et donc $|\lambda| \leq 1$. Alors $\rho(A) = 1$. Le reste du corollaire découle directement du théorème 7.32. ■