Polynômes d'endomorphismes

Le but de ce chapitre est de démontrer le théorème de Cayley-Hamilton. C'est un résultat théorique important, qui affirme que le polynôme caractéristique d'une matrice annule cette matrice.

1. Polynôme de matrice, polynôme d'endomorphisme

On note $M_n(\mathbb{K})$ l'ensemble des matrices de taille $n \times n$ à coefficients dans $\mathbb{K}$ ($\mathbb{K} = \mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$). Pour un $\mathbb{K}$-espace vectoriel E, on note $\mathscr{L}(E)$ l'ensemble des applications linéaires de E dans E. Un élément $f \in \mathscr{L}(E)$ est un *endomorphisme* de E. Dans ce chapitre, E sera de dimension finie.

1.1. Définition

Polynôme de matrice.

Soit $A \in M_n(\mathbb{K})$ une matrice. À X^k, on associe A^k ; à 1, on associe la matrice identité I_n. Plus généralement, pour un polynôme

$$P(X) = a_0 + a_1X + a_2X^2 + \cdots + a_mX^m \in \mathbb{K}[X],$$

on définit la matrice :

$$P(A) = a_0I_n + a_1A + a_2A^2 + \cdots + a_mA^m \in M_n(\mathbb{K})$$

Exemple 1.

Soient $A = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ et $P(X) = X^4 - 5X^2 - 2X + 1$. On calcule

$$A^2 = \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix} \qquad A^4 = (A^2)^2 = \begin{pmatrix} 13 & -21 \\ -21 & 34 \end{pmatrix} \qquad P(A) = A^4 - 5A^2 - 2A + I_2 = \begin{pmatrix} 2 & -4 \\ -4 & 6 \end{pmatrix}$$

Polynôme d'endomorphisme.

Soit $f \in \mathscr{L}(E)$. À X^k, on associe f^k, c'est-à-dire $\underbrace{f \circ f \circ \cdots \circ f}_{k \text{ occurrences}}$. À 1, on associe l'application identité id_E. Plus généralement, pour un polynôme

$$P(X) = a_0 + a_1X + a_2X^2 + \cdots + a_mX^m \in \mathbb{K}[X],$$

on définit l'endomorphisme :

$$P(f) = a_0 \,\mathrm{id}_E + a_1 f + a_2 f^2 + \cdots + a_m f^m \in \mathscr{L}(E)$$

Exemple 2.
Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle $\theta = \frac{\pi}{6}$ centrée à l'origine. Soit $P(X) = X^{11}$. Calculons $P(f)$. Pour $k \in \mathbb{Z}$, f^k est la rotation d'angle $k\theta = \frac{k\pi}{6}$. Donc $P(f) = f^{11}$ est la rotation d'angle $\frac{11\pi}{6}$, qui est aussi la rotation d'angle $-\frac{\pi}{6}$. Ainsi $P(f) = f^{11} = f^{-1}$.

Les opérations avec les polynômes de matrices se comportent sans surprise.

Proposition 1.
Soient $A \in M_n(\mathbb{K})$ et $P, Q \in \mathbb{K}[X]$. Alors
$$(P \times Q)(A) = P(A) \times Q(A).$$

De même, pour $f \in \mathscr{L}(E)$, $(P \times Q)(f) = P(f) \circ Q(f)$. (Noter la composition.)
Sachant en plus que, pour tous $\lambda, \mu \in \mathbb{K}$, $(\lambda P + \mu Q)(A) = \lambda P(A) + \mu Q(A)$, alors on dit en termes savants que l'application
$$\begin{array}{rccc} \Phi_A : & \mathbb{K}[X] & \longrightarrow & M_n(\mathbb{K}) \\ & P(X) & \longmapsto & P(A) \end{array}$$
est un ***morphisme d'algèbres*** ($A \in M_n(\mathbb{K})$ est fixée).

Démonstration. Si $P(X) = a_0 + a_1X + \cdots + a_mX^m$ et $Q(X) = b_0 + b_1X + \cdots + b_\ell X^\ell$, alors
$$(PQ)(X) = a_0b_0 + (a_0b_1 + a_1b_0)X + \cdots$$
Donc :
$$\begin{aligned} P(A)Q(A) &= (a_0I_n + a_1A + \cdots)(b_0I_n + b_1A + \cdots) \\ &= a_0b_0I_n + (a_0b_1 + a_1b_0)A + \cdots \\ &= (PQ)(A). \end{aligned}$$
□

Remarque (importante)**.**
En particulier, pour tous $P, Q \in \mathbb{K}[X]$, les matrices $P(A)$ et $Q(A)$ commutent :
$$P(A)Q(A) = Q(A)P(A).$$
De même, les endomorphismes $P(f)$ et $Q(f)$ commutent.

1.2. Exemples

Exemple 3 (Polynôme d'une matrice diagonale)**.**
Pour
$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$
on a :
$$P(D) = \begin{pmatrix} P(\lambda_1) & 0 & \cdots & 0 \\ 0 & P(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & P(\lambda_n) \end{pmatrix}$$
quel que soit le polynôme $P(X) \in \mathbb{K}[X]$.

Exemple 4.
Montrer plus généralement que pour une matrice triangulaire

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

on a :

$$P(T) = \begin{pmatrix} P(\lambda_1) & * & \cdots & * \\ 0 & P(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & P(\lambda_n) \end{pmatrix}$$

pour tout polynôme $P(X) \in \mathbb{K}[X]$. Les coefficients au-dessus de la diagonale peuvent avoir une expression compliquée, mais les coefficients diagonaux sont obtenus simplement en leur appliquant le polynôme P.

Mini-exercices.

1. Soit $A = \begin{pmatrix} -2 & 1 \\ 0 & 3 \end{pmatrix}$. Pour $P(X) = X^2 - X$, calculer $P(A)$. Idem avec $P(X) = X^3 - X$, puis $P(X) = X^4 - X$.
2. Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (\frac{y}{2}, -x, 3z)$. Pour $P(X) = X^n$, calculer $P(f)$ en fonction de $n \geqslant 1$ (commencer par les petites valeurs de n : $n = 1, 2, 3, 4, \ldots$).
3. Soient $A = \begin{pmatrix} 2 & -3 \\ -4 & 1 \end{pmatrix}$, $P(X) = X^2 - 3X$. Montrer que $P(A) = 10I_2$. Factoriser $P(X)$ et en déduire A^{-1}. Faire un travail similaire pour $A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & -4 & 1 \\ 1 & 0 & -1 \end{pmatrix}$, $P(X) = X^3 + 4X^2 + X$.
4. Soient A, A', B des matrices (avec B inversible) telles que $A' = BAB^{-1}$. Montrer que, pour tout polynôme $P(X) \in \mathbb{K}[X]$, $P(A') = BP(A)B^{-1}$.
5. Trouver une matrice A de taille 3×3 telle que $A^2 \neq 0$, mais $A^3 = 0$. Trouver une matrice B de taille 3×3 telle que $B^2 \neq I_3$, mais $B^3 = I_3$.

2. Sous-espaces stables

2.1. Définition

Définition 1.
Soit E un $\mathbb{K}$-espace vectoriel. Soit $f \in \mathcal{L}(E)$. Le sous-espace vectoriel F de E est ***stable*** par f si :

$$\forall x \in F \quad f(x) \in F.$$

Autrement dit, F est stable par f si $f(F) \subset F$.
Un premier exemple : les sous-espaces propres de f sont stables par f. En effet, si $F = \operatorname{Ker}(f - \lambda \operatorname{id}_E)$ alors, pour $x \in F$, $f(x) = \lambda x \in F$.

Exemple 5.
Soit (e_1, e_2, e_3) la base canonique de $\mathbb{R}^3$. Soit r_θ la rotation d'axe vertical e_3 et d'angle θ. L'endo-

morphisme r_θ de $\mathbb{R}^3$ laisse invariant deux sous-espaces :

$$F_1 = \mathrm{Vect}(e_1, e_2) = \mathbb{R}e_1 \oplus \mathbb{R}e_2 \quad \text{et} \quad F_2 = \mathrm{Vect}(e_3) = \mathbb{R}e_3$$

La matrice de f dans cette base (e_1, e_2, e_3) est la matrice

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

La matrice de cet exemple a une structure particulière. Voyons pourquoi.

Effet sur les matrices.

Supposons que E est de dimension n, que f est un endomorphisme de E, et que F est un sous-espace de E stable par f. Notons

$$(e_1, \ldots, e_p)$$

une base de F. On la complète en une base de E :

$$\mathscr{B} = (e_1, \ldots, e_p, e_{p+1}, \ldots, e_n).$$

La matrice de f dans la base $\mathscr{B}$ est triangulaire par blocs :

$$\mathrm{Mat}_{\mathscr{B}}(f) = \left(\begin{array}{ccc|ccc} a_{1,1} & \cdots & a_{1,p} & b_{1,1} & \cdots & b_{1,n-p} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{p,1} & \cdots & a_{p,p} & b_{p,1} & \cdots & b_{p,n-p} \\ \hline 0 & \cdots & 0 & d_{1,1} & \cdots & d_{1,n-p} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & d_{n-p,1} & \cdots & d_{n-p,n-p} \end{array}\right) = \left(\begin{array}{c|c} A & B \\ \hline 0 & D \end{array}\right)$$

où $A = (a_{i,j})_{1 \leqslant i,j \leqslant p} \in M_p(\mathbb{K})$ est la matrice de $f_{|F}$ dans la base $(e_1, \ldots, e_p)$ de F.

Remarque.

Si $E = F_1 \oplus F_2$ et que F_1 et F_2 sont tous les deux stables par f, alors la matrice de f est diagonale par blocs :

$$\mathrm{Mat}_{\mathscr{B}}(f) = \left(\begin{array}{ccc|ccc} a_{1,1} & \cdots & a_{1,p} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{p,1} & \cdots & a_{p,p} & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & d_{1,1} & \cdots & d_{1,n-p} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & d_{n-p,1} & \cdots & d_{n-p,n-p} \end{array}\right) = \left(\begin{array}{c|c} A & 0 \\ \hline 0 & D \end{array}\right)$$

Voir l'exemple 5 ci-dessus.

2.2. Polynôme d'endomorphisme

> **Lemme 1.**
> *Si F est un sous-espace vectoriel stable par f alors, pour tout polynôme $P \in \mathbb{K}[X]$, F est stable par $P(f)$.*

Démonstration. Si $x \in F$, alors $f(x) \in F$ et donc $f(f(x)) \in F$. Par récurrence sur k, on montre que $f^k(x) \in F$, pour tout $k \geqslant 0$. Maintenant, si $P(X) = \sum_{k=0}^m a_k X^k$, alors $P(f)$ est l'endomorphisme défini par

$$P(f) = a_0 \,\mathrm{id}_E + a_1 f + a_2 f^2 + \cdots + a_m f^m.$$

Donc
$$P(f)(x) = a_0 x + a_1 f(x) + a_2 f^2(x) + \cdots + a_m f^m(x).$$
Chaque terme $a_k f^k(x) \in F$, donc $P(f)(x) \in F$ car F est un espace vectoriel. Conclusion : F est stable par $P(f)$. □

Une autre proposition souvent utile est la suivante :

Proposition 2.
Soient f et g deux endomorphismes de E qui commutent, c'est-à-dire tels que $f \circ g = g \circ f$. Alors $\operatorname{Ker} g$ et $\operatorname{Im} g$ sont stables par f.

Démonstration.
- Soit $x \in \operatorname{Ker} g$. On a $g(x) = 0$, d'où $g(f(x)) = f(g(x)) = f(0) = 0$, donc $f(x) \in \operatorname{Ker} g$.
- Soit $y \in \operatorname{Im} g$. Il existe $x \in E$ tel que $y = g(x)$, d'où $f(y) = f(g(x)) = g(f(x))$, donc $f(y) \in \operatorname{Im} g$.

□

2.3. Polynôme caractéristique

Soit F un sous-espace stable par un endomorphisme $f : E \to E$. Dans ce cas, on note $f_{|F} : F \to F$, $x \in F \mapsto f(x) \in F$, la *restriction* de f à F. L'application $f_{|F}$ est un endomorphisme de F.

Lemme 2.
Soit f un endomorphisme de E (de dimension finie). On suppose aussi qu'il existe un sous-espace F de E laissé stable par f. Notons $\chi_{f_{|F}}$ le polynôme caractéristique de la restriction de f à F. Alors :
$$\chi_{f_{|F}}(X) \ \textit{divise} \ \chi_f(X) \ \textit{dans} \ \mathbb{K}[X].$$

Démonstration. On considère une base $(e_1, \dots, e_p)$ de F, et on la complète en une base $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ de E. La matrice de f dans cette base est de la forme
$$M = \left(\begin{array}{c|c} A & B \\ \hline 0 & D \end{array}\right)$$
où $A \in M_p(\mathbb{K})$ est la matrice de $f_{|F}$ dans la base $(e_1, \dots, e_p)$.
On a alors
$$\begin{aligned} \chi_f(X) &= \det(M - XI_n) \\ &= \left|\begin{array}{c|c} A - XI_p & B \\ \hline 0 & D - XI_{n-p} \end{array}\right| \\ &= \det(A - XI_p) \times \det(D - XI_{n-p}) \\ &= \chi_{f_{|F}}(X) \times Q(X). \end{aligned}$$
Cela prouve que $\chi_{f_{|F}}(X)$ divise $\chi_f(X)$. □

Mini-exercices.

1. Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (2x - y, 3x - 2y, \frac{1}{3}z)$. Calculer la matrice de f dans la base canonique et déterminer le polynôme caractéristique de f. En déduire les sous-espaces stables de l'application f.

2. Soit $A = \begin{pmatrix} 1 & -1 & 0 \\ -2 & -1 & -1 \\ -1 & -2 & 1 \end{pmatrix}$. Trouver une valeur propre de A et un vecteur propre associé. Montrer que les vecteurs $w_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ et $w_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ engendrent un sous-espace stable de dimension 2 de cette matrice. En déduire une matrice P telle que $P^{-1}AP$ soit une matrice diagonale par blocs.
3. Soit f l'application linéaire définie par la matrice $A = \begin{pmatrix} -2 & 0 & -2 \\ -1 & 1 & -1 \\ 4 & 0 & 4 \end{pmatrix}$. Soit g l'application linéaire définie par la matrice $B = \begin{pmatrix} 2 & 0 & 1 \\ 3 & 3 & 3 \\ -2 & 0 & -1 \end{pmatrix}$. Montrer que $f \circ g = g \circ f$. Calculer $\operatorname{Ker} g$ et $\operatorname{Im} g$, et vérifier qu'ils sont stables par f. Calculer $\operatorname{Ker} f$ et $\operatorname{Im} f$, et vérifier qu'ils sont stables par g.

3. Théorème de Cayley-Hamilton

3.1. Énoncé

Théorème 1 (de Cayley-Hamilton)**.**
Soit $A \in M_n(\mathbb{K})$. Alors

$$\boxed{\chi_A(A) = 0.}$$

De même, soit $f \in \mathscr{L}(E)$, avec E de dimension finie. Alors

$$\boxed{\chi_f(f) = 0.}$$

L'égalité $\chi_A(A) = 0$ signifie que le polynôme caractéristique appliqué à A donne la matrice nulle. L'égalité $\chi_f(f) = 0$ signifie que le polynôme caractéristique appliqué à f donne l'application nulle.

Exemple 6.
Soit $A = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix} \in M_2(\mathbb{R})$. Le polynôme caractéristique de A est

$$\chi_A(X) = \det(A - XI_2) = \begin{vmatrix} 1-X & -2 \\ 1 & -1-X \end{vmatrix} = (1-X)(-1-X) + 2 = X^2 + 1.$$

Vérifions le théorème de Cayley-Hamilton sur cet exemple, en calculant $\chi_A(A)$:

$$\chi_A(A) = A^2 + I_2 = -I_2 + I_2 = 0.$$

Exemple 7.
Plus généralement, en dimension 2, posons

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}).$$

On sait que

$$\chi_A(X) = X^2 - (a+d)X + ad - bc.$$

Vérifions que $\chi_A(A) = 0$:

$$\begin{aligned}\chi_A(A) &= A^2-(a+d)A+(ad-bc)I_2\\ &= \begin{pmatrix} a^2+bc & ab+bd\\ ac+cd & bc+d^2\end{pmatrix}-(a+d)\begin{pmatrix} a & b\\ c & d\end{pmatrix}+\begin{pmatrix} ad-bc & 0\\ 0 & ad-bc\end{pmatrix}\\ &= \begin{pmatrix} 0 & 0\\ 0 & 0\end{pmatrix}\end{aligned}$$

Exemple 8.
Soient les deux matrices de $M_n(\mathbb{K})$ suivantes :

$$N=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ \vdots & & \ddots & \ddots & 0\\ \vdots & & & \ddots & 1\\ 0 & \cdots & \cdots & \cdots & 0\end{pmatrix} \qquad \text{et} \qquad J=\begin{pmatrix} 0 & \cdots & \cdots & 0 & 1\\ 1 & 0 & \cdots & \cdots & 0\\ 0 & 1 & \ddots & & \vdots\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ 0 & \cdots & 0 & 1 & 0\end{pmatrix}.$$

D'une part, $\chi_N(X)=(-1)^nX^n$ et on a bien $N^n=0$. D'autre part, $\chi_J(X)=(-1)^n(X^n-1)$ et on a bien $J^n=I_n$.

3.2. Preuve

Démonstration du théorème de Cayley-Hamilton.
On suppose E de dimension finie n. Soit f un endomorphisme de E. Soit x un vecteur non nul de E. Soit $1\leqslant p\leqslant n$ le plus grand entier tel que la famille

$$\big(x, f(x),\ldots,f^{p-1}(x)\big)$$

soit libre. Alors, forcément, la famille

$$\big(x, f(x),\ldots,f^{p-1}(x), f^p(x)\big)$$

est liée et, plus précisément,

$$c_0x+c_1f(x)+\cdots+c_{p-1}f^{p-1}(x)+f^p(x)=0 \tag{1}$$

pour certains coefficients $c_0,\ldots,c_{p-1}\in\mathbb{K}$.
Posons $F=\mathrm{Vect}(x,f(x),\ldots,f^{p-1}(x))$. C'est un sous-espace vectoriel de E (de dimension p) stable par f. En effet, notons $v_0=x$, $v_1=f(x)$, …, $v_{p-1}=f^{p-1}(x)$. Alors, pour $0\leqslant k\leqslant p-2$, on a $f(v_k)=v_{k+1}\in F$; et, par la relation (1),

$$f(v_{p-1})=f^p(x)=-c_0v_0-c_1v_1-\cdots-c_{p-1}v_{p-1}\in F.$$

De plus, la matrice de la restriction $f_{|F}$ dans la base

$$\big(x, f(x),\ldots,f^{p-1}(x)\big)$$

est la matrice

$$A=\begin{pmatrix} 0 & \cdots & \cdots & 0 & -c_0\\ 1 & \ddots & & \vdots & \vdots\\ 0 & \ddots & \ddots & \vdots & \vdots\\ \vdots & \ddots & \ddots & 0 & \vdots\\ 0 & \cdots & 0 & 1 & -c_{p-1}\end{pmatrix}.$$

C'est une matrice compagnon, donc

$$\chi_A(X) = \pm(X^p + c_{p-1}X^{p-1} + \cdots + c_0).$$

D'après le lemme 2, $\chi_A(X)$ divise $\chi_f(X)$, c'est-à-dire

$$\chi_f(X) = Q(X)\chi_A(X)$$

pour un certain polynôme $Q(X) \in \mathbb{K}[X]$. On a alors :

$$\begin{aligned}
\chi_f(f)(x) &= \big(Q(f) \circ \chi_A(f)\big)(x) \\
&= Q(f)\big(\chi_A(f)(x)\big) \\
&= Q(f)\big(\pm(f^p(x) + c_{p-1}f^{p-1}(x) + \cdots + c_0 x)\big) \\
&= Q(f)(0) \\
&= 0
\end{aligned}$$

Finalement, $\chi_f(f)(x) = 0$ pour tout vecteur x de E, et donc $\chi_f(f) = 0$. □

3.3. Polynôme annulateur

Définition 2.
On dit qu'un polynôme $P(X)$ est un ***polynôme annulateur*** de la matrice A (ou de l'endomorphisme f) si $P(A) = 0$ (ou $P(f) = 0$).

Exemple 9.

- Soit $p : E \to E$ tel que $p^2 = p$ (c'est une ***projection***). Alors $X^2 - X$ est un polynôme annulateur de p.
- Soit $r : E \to E$ tel que $r^2 = \mathrm{id}_E$ (c'est une ***réflexion***). Alors $X^2 - 1$ est un polynôme annulateur de r.

Reformulation du théorème de Cayley-Hamilton :

> Le polynôme caractéristique de la matrice A (resp. de l'endomorphisme f) est un polynôme annulateur de A (resp. de f).

Où chercher les valeurs propres, connaissant un polynôme annulateur mais ne connaissant pas le polynôme caractéristique ?

Proposition 3.
Si P est un polynôme annulateur de f, alors

$$sp(f) \subset \{ \text{ racines de } P \}.$$

Le même énoncé est bien sûr vrai pour les matrices.

Démonstration. Soit x est vecteur propre de f, associé à une valeur propre λ. Comme $f(x) = \lambda x$, on a :

$$\forall k \geqslant 0 \quad f^k(x) = \lambda^k x$$

et plus généralement, pour tout polynôme $Q(X)$:

$$Q(f)(x) = Q(\lambda)x$$

En particulier, comme $P(f)(x) = 0$, alors $P(\lambda)x = 0$, ce qui implique $P(\lambda) = 0$ car $x \neq 0$. □

Mini-exercices.

1. Soit $A = \begin{pmatrix} -5 & -3 & 3 \\ 7 & 3 & -3 \\ 1 & -2 & 2 \end{pmatrix}$. Calculer le polynôme caractéristique de A. Vérifier que $\chi_A(A) = 0$.
2. Soit $f : \mathbb{R}^4 \to \mathbb{R}^4$ défini par $f(x_1, x_2, x_3, x_4) = (3x_1 + 5x_2, -2x_1 - 4x_2, -x_4, x_3 + 2x_4)$. Calculer le polynôme caractéristique de f. Vérifier que $\chi_f(f) = 0$.
3. Soit $A = \begin{pmatrix} -3 & -2 & 0 & -2 \\ 2 & 1 & 0 & 2 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & -1 \end{pmatrix}$. Montrer que $X^3 - 3X - 2$ est un polynôme annulateur de A. En déduire les valeurs propres de A.

4. Polynôme minimal

Nous venons de démontrer que si f est un endomorphisme et χ_f est son polynôme caractéristique alors $\chi_f(f) = 0$ (et de même $\chi_A(A) = 0$ pour $A \in M_n(\mathbb{K})$). Nous allons démontrer qu'il existe un plus petit polynôme ayant cette propriété, ce polynôme n'étant pas toujours le polynôme caractéristique.

4.1. Définition

Proposition 4.
Soit f un endomorphisme de E. Il existe un unique polynôme $\mu_f(X) \in \mathbb{K}[X]$ qui vérifie les trois conditions suivantes :

- *$\mu_f(X)$ est un polynôme annulateur pour f ;*
- *$\mu_f(X)$ est unitaire ;*
- *si $P(X) \in \mathbb{K}[X]$ est un polynôme annulateur de f alors $\deg \mu_f(X) \leqslant \deg P(X)$.*

Ce polynôme est appelé le ***polynôme minimal*** de f. C'est donc le polynôme unitaire de degré le plus petit qui annule f. On définit de même le polynôme minimal $\mu_A(X)$ d'une matrice $A \in M_n(\mathbb{K})$.

Exemple 10.
Soit

$$A = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 2 & -1 \\ 1 & 1 & 0 \end{pmatrix} \in M_3(\mathbb{R}).$$

Montrons que $\mu_A(X) = X^2 - X$.

- On vérifie que $A^2 - A = 0$, donc le polynôme $X^2 - X$ annule A.
- On vérifie que $A - \lambda I_3$ n'est jamais la matrice nulle (quel que soit $\lambda \in \mathbb{R}$). Donc aucun polynôme de degré 1 n'annule A.
- Le polynôme $X^2 - X$ est donc le polynôme unitaire de plus petit degré annulant A. Conclusion : $\mu_A(X) = X^2 - X$.

Autres exemples. Le polynôme minimal de la matrice identité I_n est $\mu_A(X) = X - 1$ (quel que soit $n \geqslant 1$). Le polynôme minimal de la matrice nulle est X.

Non seulement le polynôme minimal est le polynôme annulateur de degré le plus petit, mais c'est aussi le plus petit au sens de la division des polynômes.

Proposition 5.

Le polynôme minimal de f, $\mu_f(X)$, divise tous les polynômes annulateurs de f.

Rappels sur la division euclidienne.
Soient P, Q deux polynômes dans $\mathbb{K}[X]$. Si $Q \neq 0$, alors il existe un unique couple (B, R) tels que $B, R \in \mathbb{K}[X]$ et :

$$P = BQ + R \quad \text{et} \quad \deg R < \deg Q.$$

Le polynôme R peut éventuellement être nul. En plus, $R = 0$ si et seulement si Q divise P.
Nous prouvons les propositions 4 et 5 en même temps.

Démonstration.

- Notons $n = \dim E$. Soit f un endomorphisme de E et soit χ_f son polynôme caractéristique. Par le théorème de Cayley-Hamilton, le polynôme χ_f annule f. Ainsi, l'ensemble des polynômes P non nuls vérifiant $P(f) = 0$ n'est pas vide. Choisissons dans cet ensemble un polynôme Q de degré minimal.
- Il est clair que tout polynôme multiple de Q s'annule également en f. En effet, si $P = BQ$, alors $P(f) = B(f) \circ Q(f) = 0$ car $Q(f) = 0$.
- Réciproquement, soit $P \in \mathbb{K}[X]$ tel que $P(f) = 0$. On effectue la division euclidienne de P par Q : on obtient $P = BQ + R$ avec $\deg R < \deg Q$. Ainsi

  $$P(f) = B(f) \circ Q(f) + R(f) = 0.$$

  Comme de plus $Q(f) = 0$, alors on déduit de la division euclidienne que l'on a aussi $R(f) = 0$. Par l'absurde, si $R(X)$ n'est pas le polynôme nul, alors on a obtenu un polynôme non nul qui annule f et qui est de degré strictement inférieur à celui de Q ; c'est contradictoire avec le choix de Q. Bilan : $R = 0$. D'où $P = BQ$, c'est-à-dire Q divise P.
- Vérifions l'unicité d'un tel Q s'il est choisi unitaire. Supposons qu'il existe Q_1 et Q_2, tous deux de degré minimal, unitaires et annulant f. Alors, par ce qui précède, Q_1 divise Q_2 et de même Q_2 divise Q_1, ce qui prouve que $Q_1 = Q_2$.

□

4.2. Lien avec le polynôme caractéristique

Une conséquence immédiate de la section précédente (proposition 5) et du théorème de Cayley-Hamilton est que le polynôme minimal divise le polynôme caractéristique :

$$\boxed{\mu_f(X) \ \text{divise} \ \chi_f(X)}$$

Cela va impliquer que les racines du polynôme minimal sont exactement les valeurs propres :

Proposition 6.
Soit $\lambda \in \mathbb{K}$. Alors :

$$\boxed{\lambda \text{ valeur propre de } f \iff \mu_f(\lambda) = 0}$$

Démonstration.

- $\Longleftarrow$. On sait que μ_f divise χ_f. Comme $\chi_f(X) = B(X) \times \mu_f(X)$ alors, si λ est racine du polynôme minimal, $\mu_f(\lambda) = 0$ et donc $\chi_f(\lambda) = B(\lambda) \times \mu_f(\lambda) = 0$. Donc λ est racine du polynôme caractéristique : c'est donc une valeur propre de f.

- $\Longrightarrow$. Réciproquement, supposons que λ soit une valeur propre de f et notons x un vecteur propre, de sorte que $f(x) = \lambda x$. Alors, pour tout $k \geqslant 0$, $f^k(x) = \lambda^k x$ et plus généralement, pour tout polynôme $P \in \mathbb{K}[X]$, $P(f)(x) = P(\lambda)x$. Donc $P(\lambda)$ est valeur propre de l'endomorphisme $P(f)$. On applique ceci au polynôme minimal $\mu_f(X)$: $\mu_f(f)(x) = \mu_f(\lambda)x$. Or, par définition du polynôme minimal, $\mu_f(f)$ est l'endomorphisme identiquement nul. Ainsi $\mu_f(f)(x) = 0$, donc par l'égalité précédente $\mu_f(\lambda)x = 0$. Comme $x \neq 0$, alors $\mu_f(\lambda) = 0$.

□

Comment trouver le polynôme minimal d'une matrice ?

Comme le polynôme minimal divise le polynôme caractéristique, alors on a le résultat suivant :

Lemme 3.

Soit $A \in M_n(\mathbb{K})$. On suppose que le polynôme caractéristique de A est scindé :

$$\chi_A(X) = \pm(X-\lambda_1)^{m_1}\cdots(X-\lambda_r)^{m_r}$$

où $m_1,\dots,m_r \geqslant 1$, $\lambda_1,\dots,\lambda_r \in \mathbb{K}$ sont deux à deux distincts. Alors :

$$\mu_A(X) = (X-\lambda_1)^{k_1}\cdots(X-\lambda_r)^{k_r}$$

pour certains entiers $1 \leqslant k_i \leqslant m_i$ ($i = 1,\dots,r$).

Un cas particulier important : si le polynôme caractéristique est $\pm(X-\lambda)^n$, alors le polynôme minimal est $(X-\lambda)^d$ avec $1 \leqslant d \leqslant n$.

Exemple 11.

Voici trois exemples, à vous de faire les calculs.

A	$\begin{pmatrix} 0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0 \end{pmatrix}$	$\begin{pmatrix} 0&1&0&0\\0&0&1&0\\0&0&0&0\\0&0&0&0 \end{pmatrix}$	$\begin{pmatrix} 0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0 \end{pmatrix}$
$\chi_A(X)$	X^4	X^4	X^4
$\mu_A(X)$	X^2	X^3	X^4

Exemple 12 (Polynôme minimal d'une matrice diagonale)**.**

Soit

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

Alors $\mu_D(X) = \prod_{\lambda \in \mathrm{sp}(D)}(X-\lambda)$ où $\mathrm{sp}(D) = \{\lambda_1,\dots,\lambda_n\}$ et les valeurs propres sont comptées sans multiplicité. Par exemple, si

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -5 \end{pmatrix},$$

alors $\lambda_1 = 2$, $\lambda_2 = 2$ et $\lambda_3 = -5$. Donc $\mathrm{sp}(D) = \{2, -5\}$. Le polynôme minimal est $\mu_D(X) = (X-2)(X+5)$. Le polynôme caractéristique, quant à lui, est $\chi_D(X) = -(X-2)^2(X+5)$.

4.3. Diagonalisation

Nous arrivons à une nouvelle condition nécessaire et suffisante de diagonalisation.

> **Théorème 2.**
> *Une matrice $A \in M_n(\mathbb{K})$ (resp. un endomorphisme $f \in \mathscr{L}(E)$) est diagonalisable sur $\mathbb{K}$ si et seulement si son polynôme minimal est scindé à racines simples dans $\mathbb{K}$.*

On rappelle qu'un polynôme $P(X) \in \mathbb{K}[X]$ est ***scindé à racines simples dans*** $\mathbb{K}$ s'il se factorise en :

$$P(X) = a(X-\lambda_1)\cdots(X-\lambda_r)$$

où $a \in \mathbb{K}^*$ et $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ sont deux à deux distincts.

Exemple 13.
Soit

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

On a

$$\chi_A(X) = \begin{vmatrix} 2-X & 1 & -1 \\ -1 & -1-X & 3 \\ 0 & -1 & 3-X \end{vmatrix} = (2-X)(1-X)^2.$$

Le polynôme minimal μ_A de A divise χ_A et admet les mêmes racines : il est donc égal à $(X-1)(X-2)$ ou $(X-1)^2(X-2)$. Or

$$(A-I_3)(A-2I_3) = \begin{pmatrix} -1 & -1 & 1 \\ 2 & 2 & -2 \\ 1 & 1 & -1 \end{pmatrix} \neq 0,$$

donc le polynôme minimal de A n'est pas $(X-1)(X-2)$. Ainsi, le polynôme minimal est égal à $(X-1)^2(X-2)$, et en particulier il admet une racine double. Par conséquent, le théorème 2 implique que A n'est diagonalisable ni dans $\mathbb{R}$ ni dans $\mathbb{C}$.

Démonstration.

- $\Longrightarrow$. Si A est diagonalisable, A est semblable à une matrice diagonale. Or deux matrices semblables ont le même polynôme minimal (à faire en exercice). Donc il suffit de calculer le polynôme minimal d'une matrice diagonale, ce qui est l'objet d'un exercice précédent (exemple 12) : ce polynôme est scindé à racines simples.
- $\Longleftarrow$. On suppose que le polynôme minimal de A est scindé à racines simples : $\mu_A(X) = (X-\lambda_1)\cdots(X-\lambda_r)$ avec $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ deux à deux distincts.
  — La décomposition en éléments simples de la fraction $\frac{1}{\mu_A(X)}$ donne :

  $$\frac{1}{\mu_A(X)} = \frac{a_1}{X-\lambda_1} + \cdots + \frac{a_r}{X-\lambda_r}$$

  où $a_i \in \mathbb{K}^*$ pour tout $1 \leqslant i \leqslant r$. On multiplie par $\mu_A(X)$ des deux côtés :

  $$1 = a_1\frac{\mu_A(X)}{X-\lambda_1} + \cdots + a_r\frac{\mu_A(X)}{X-\lambda_r}$$

  On définit les polynômes :

  $$Q_i(X) = \frac{\mu_A(X)}{X-\lambda_i} = \prod_{\substack{j=1,\ldots,r \\ j\neq i}} (X-\lambda_j)$$

Donc
$$1 = a_1 Q_1(X) + \cdots + a_r Q_r(X).$$

— Si on applique cette égalité à la matrice A, on trouve :
$$I_n = a_1 Q_1(A) + \cdots + a_r Q_r(A).$$
Soit un vecteur $v \in \mathbb{K}^n$. On a :
$$v = a_1 Q_1(A)(v) + \cdots + a_r Q_r(A)(v). \tag{2}$$

— Or, pour tout $1 \leqslant i \leqslant r$, $Q_i(A)(v) \in \operatorname{Ker}(A - \lambda_i I_n)$. En effet,
$$(A - \lambda_i I_n)(Q_i(A)(v)) = \mu_A(A)(v) = 0.$$

— Par conséquent, l'égalité (2) implique que $v \in \operatorname{Ker}(A-\lambda_1 I_n) + \cdots + \operatorname{Ker}(A-\lambda_r I_n)$. Ceci étant vrai quel que soit $v \in \mathbb{K}^n$, alors
$$\mathbb{K}^n = \operatorname{Ker}(A-\lambda_1 I_n) + \cdots + \operatorname{Ker}(A-\lambda_r I_n).$$
Mais, d'autre part, les valeurs propres étant distinctes, les sous-espaces propres sont en somme directe (voir le chapitre « Diagonalisation »). En conclusion :
$$\mathbb{K}^n = \operatorname{Ker}(A-\lambda_1 I_n) \oplus \cdots \oplus \operatorname{Ker}(A-\lambda_r I_n)$$
C'est exactement dire que A est diagonalisable.

□

Remarque : pour la réciproque, on peut aussi utiliser le lemme des noyaux (voir le chapitre « Décomposition de Dunford et réduction de Jordan »).

Corollaire 1.
Une matrice $A \in M_n(\mathbb{K})$ est diagonalisable sur $\mathbb{K}$ si et seulement si elle admet un polynôme annulateur scindé à racines simples dans $\mathbb{K}$.

Corollaire 2.
Soit f un endomorphisme de E, diagonalisable sur $\mathbb{K}$. Si F est un sous-espace vectoriel de E stable par f, alors la restriction $f_{|F}$ est encore diagonalisable sur $\mathbb{K}$.

Démonstration. En effet,
$$\mu_f(f) = 0 \implies \mu_f(f_{|F}) = 0 \implies \mu_{f_{|F}} \text{ divise } \mu_f.$$
Mais si μ_f est scindé à racines simples dans $\mathbb{K}$, tous ses diviseurs le sont aussi. En particulier, $\mu_{f_{|F}}$ est scindé à racines simples dans $\mathbb{K}$, donc $f_{|F}$ est diagonalisable. □

Mini-exercices.

1. Soit P un polynôme annulateur d'un endomorphisme f. Montrer que, pour tout $\lambda \in \operatorname{sp}(f)$, $P(\lambda) = 0$.
2. Montrer que deux matrices semblables ont le même polynôme minimal.
3. Trouver le polynôme minimal de $A = \begin{pmatrix} -1 & 1 \\ -4 & 3 \end{pmatrix}$. Cette matrice est-elle diagonalisable ? Même exercice avec $B = \begin{pmatrix} 12 & -5 \\ 30 & -13 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 0 & -1 \\ 3 & -2 & 5 \\ 0 & 0 & 1 \end{pmatrix}$.

Auteurs du chapitre

D'après un cours de Sandra Delaunay et un cours d'Alexis Tchoudjem.
Revu et augmenté par Arnaud Bodin.

Relu par Stéphanie Bodin et Vianney Combet.