

Le théorème de Cayley-Hamilton

Marc Lorenzi

26 février 2025

Dans tout l'article, E est un $\mathbb{K}$ -espace vectoriel de dimension finie n sur un corps $\mathbb{K}$.

1 Polynômes d'endomorphismes

1.1 Un morphisme d'algèbres

Étant donné un polynôme $P \in \mathbb{K}[X]$, on a l'écriture unique

$$P = \sum_{k=0}^{\infty} a_k X^k$$

où les $a_k \in \mathbb{K}$ sont presque tous nuls. Posons alors pour tout $f \in \mathcal{L}(E)$,

$$P(f) = \sum_{k=0}^{\infty} a_k f^k$$

Nous définissons ainsi pour tout $f \in \mathcal{L}(E)$ l'application $\Phi_f : \mathbb{K}[X] \longrightarrow \mathcal{L}(E)$ définie par $\Phi_f(P) = P(f)$.

Proposition 1. *Pour tout $f \in \mathcal{L}(E)$, Φ_f est un morphisme d'algèbres.*

Ce sont de « simples » vérifications. Notons que pour tout $f \in \mathcal{L}(E)$, pour tous $P, Q \in \mathbb{K}[X]$, pour tout $\lambda \in \mathbb{K}$,

- $(P + Q)(f) = P(f) + Q(f)$.
- $(\lambda P)(f) = \lambda P(f)$.
- $(PQ)(f) = P(f) \circ Q(f)$.
- $1(f) = \text{id}_E$.

L'image de Φ_f est l'ensemble des « polynômes en f ». Cet ensemble est une sous-algèbre de $\mathcal{L}(E)$. Comme deux puissances de f commutent entre elles, on a pour tous $P, Q \in \mathbb{K}[X]$, $P(f) \circ Q(f) = Q(f) \circ P(f)$. Ainsi, l'image de Φ_f est une sous-algèbre commutative de $\mathcal{L}(E)$. Sa dimension est évidemment inférieure ou égale à n^2 , qui est la dimension de $\mathcal{L}(E)$. Le théorème de Cayley-Hamilton nous montrera bien mieux, à savoir que sa dimension est inférieure ou égale à n . Parlons maintenant du noyau de Φ_f .

1.2 Polynômes annulateurs

Pour tout $f \in \mathcal{L}(E)$, notons

$$\mathcal{A}(f) = \ker \Phi_f$$

Les polynômes $P \in \mathcal{A}(f)$ sont les *polynômes annulateurs* de f .

Proposition 2. Pour tout $f \in \mathcal{L}(E)$, $\mathcal{A}(f)$ est un idéal non nul de $\mathbb{K}[X]$.

Démonstration. Le noyau de Φ_f est en particulier le noyau d'un morphisme d'anneaux. C'est donc un idéal de $\mathbb{K}[X]$. De plus, $\mathcal{L}(E)$ est un $\mathbb{K}$ -espace vectoriel de dimension finie n^2 . La famille $(f^k)_{k \in \llbracket 0, n^2 \rrbracket}$ est donc liée puisqu'elle est de cardinal $n^2 + 1$. Il existe donc $a_0, \dots, a_{n^2} \in \mathbb{K}$ pas tous nuls tels que

$$\sum_{k=0}^{n^2} a_k f^k = 0$$

Soit

$$P = \sum_{k=0}^{n^2} a_k X^k$$

Le polynôme P n'est pas le polynôme nul et $P \in \mathcal{A}(f)$, donc $\mathcal{A}(f) \neq \{0\}$. $\square$

L'anneau $\mathbb{K}[X]$ est un anneau principal. Pour tout $f \in \mathcal{L}(E)$ il existe donc un polynôme non nul μ tel que $\mathcal{A}(f) = \mu \mathbb{K}[X]$. En choisissant μ unitaire on a également unicité.

Définition 1. Pour tout $f \in \mathcal{L}(E)$, le *polynôme minimal* de f est l'unique polynôme unitaire $\mu_f \in \mathbb{K}[X]$ tel que

$$\mathcal{A}(f) = \mu_f \mathbb{K}[X]$$

Autrement dit, pour tout polynôme $P \in \mathbb{K}[X]$, $P(f) = 0$ si et seulement si μ_f divise P .

1.3 Exemples élémentaires

Un cas particulier. Supposons μ_f de degré 0, c'est à dire, puisque μ_f est unitaire, $\mu_f = 1$. On a $\mu_f(f) = f^0 = \text{id}_E = 0$. Ceci ne peut arriver que si $n = \dim E = 0$. Ainsi, si $n \geq 1$ alors $\deg(\mu_f) \geq 1$. Dans les exemples qui suivent, nous supposons que c'est le cas.

Exemple 1. Soit $f \in \mathcal{L}(E)$. On a $\deg(\mu_f) = 1$ si et seulement si f est une homothétie. Précisément, si $f = \lambda \text{id}_E$ où $\lambda \in \mathbb{K}$, alors $\mu_f = X - \lambda$.

Exemple 2. Soit f un projecteur de E différent de 0 et de id_E . On a

$$(X^2 - X)(f) = f^2 - f = 0$$

Ainsi, μ_f divise $X^2 - X$. Comme f n'est pas une homothétie, $\deg(\mu_f) \geq 2$ et donc $\mu_f = X^2 - X$.

Exemple 3. De même, si f est une symétrie de E différente de $\pm \text{id}_E$, alors $\mu_f = X^2 - 1$.

2 Polynôme caractéristique

2.1 Matrices à coefficients polynomiaux

Nous allons considérer l'algèbre $\mathcal{M}_n(\mathbb{K}(X))$ des matrices à coefficients dans le corps $\mathbb{K}(X)$ des fractions rationnelles à une indéterminée X . En identifiant les fractions rationnelles constantes aux éléments de $\mathbb{K}$, cette algèbre contient $\mathcal{M}_n(\mathbb{K})$ comme sous-algèbre.

L'algèbre $\mathcal{M}_n(\mathbb{K}(X))$ contient la sous-algèbre $\mathcal{M}_n(\mathbb{K}[X])$ des matrices dont les coefficients sont des polynômes. On a la double inclusion

$$\mathcal{M}_n(\mathbb{K}) \subset \mathcal{M}_n(\mathbb{K}[X]) \subset \mathcal{M}_n(\mathbb{K}(X))$$

Proposition 3. Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K}[X])$, on a $\det A \in \mathbb{K}[X]$.

Démonstration. Soit $A \in \mathcal{M}_n(\mathbb{K}[X])$. On a

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) A_{\sigma(1),1} \dots A_{\sigma(n),n}$$

Les coefficients $A_{\sigma(k),k}$ étant des polynômes, $\det(A)$ en est un aussi. $\square$

2.2 Notion de polynôme caractéristique

Définition 2. Soit $A \in \mathcal{M}_n(\mathbb{K})$. Le *polynôme caractéristique* de A est le polynôme

$$\chi_A = \det(XI_n - A)$$

Le polynôme caractéristique de A est donc le déterminant d'une matrice de $\mathcal{M}_n(\mathbb{K}[X])$. Par la proposition précédente, c'est bien un polynôme.

Proposition 4. Soit $f \in \mathcal{L}(E)$. Le polynôme caractéristique de la matrice de f dans une base $\mathcal{B}$ de E ne dépend pas de la base $\mathcal{B}$ choisie.

Démonstration. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E . Notons A et A' les matrices de f dans ces deux bases. Notons également P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. On a

$$XI_n - A' = P^{-1}(XI_n - A)P$$

De là,

$$\chi_{A'} = (\det P)^{-1} \chi_A \det P = \chi_A$$

□

Définition 3. Soit $f \in \mathcal{L}(E)$. Le *polynôme caractéristique* de f est le polynôme caractéristique de la matrice de f dans une base quelconque de E .

Nous noterons χ_f le polynôme caractéristique de l'endomorphisme f .

Proposition 5. Soit $f \in \mathcal{L}(E)$. Le polynôme χ_f est unitaire et de degré n .

Démonstration. Soit $\mathcal{B}$ une base de E . Soit A la matrice de f dans la base $\mathcal{B}$. Soit $B = XI_n - A$. Remarquons que pour tous $i, j \in \llbracket 1, n \rrbracket$, tels que $i \neq j$, $B_{i,j}$ est un polynôme constant. Également, pour tout $i \in \llbracket 1, n \rrbracket$, $B_{i,i} = X - A_{i,i}$. Le polynôme $B_{i,i}$ est donc un polynôme unitaire de degré 1. On a

$$\begin{aligned} \chi_f &= \det B \\ &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) B_{\sigma(1),1} \dots B_{\sigma(n),n} \\ &= B_{1,1} \dots B_{n,n} + \sum_{\sigma \neq \text{id}} \varepsilon(\sigma) B_{\sigma(1),1} \dots B_{\sigma(n),n} \end{aligned}$$

Pour toute permutation $\sigma \neq \text{id}$, il existe au moins deux entiers $k, k' \in \llbracket 1, n \rrbracket$ distincts tels que $\sigma(k) \neq k$ et $\sigma(k') \neq k'$. On a donc

$$\deg B_{\sigma(1),1} \dots B_{\sigma(n),n} = \sum_{k=1}^n \deg B_{\sigma(k),k} \leq n - 2$$

On en déduit que le degré de la somme pour $\sigma \neq \text{id}$ est inférieur ou égal à $n - 2$. Par ailleurs,

$$\deg B_{1,1} \dots B_{n,n} = n$$

De là, $\deg(\chi_f) = n$. De plus, le coefficient dominant de χ_f est

$$\text{cd}(B_{1,1} \dots B_{n,n}) = \text{cd}((X - A_{1,1}) \dots (X - A_{n,n})) = 1$$

□

Remarque. Comme la somme pour $\sigma \neq \text{id}$ est de degré inférieur ou égal à $n - 2$, le coefficient de degré $n - 1$ de χ_f est aussi celui de degré $n - 1$ de $B_{1,1} \dots B_{n,n}$. Facilement, ce coefficient est

$$-(A_{1,1} + \dots + A_{n,n}) = -\text{trace}(f)$$

Notons aussi que le coefficient constant de χ_f est

$$\chi_f(0) = \det(-f) = (-1)^n \det(f)$$

2.3 Matrice compagnon

Soient $a_0, \dots, a_{n-1} \in \mathbb{K}$. Soit

$$P = X^n + \sum_{k=0}^{n-1} a_k X^k$$

La *matrice compagnon* du polynôme P est

$$C_P = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}$$

Proposition 6. $\chi_{C_P} = P$.

Démonstration. Montrons par récurrence sur n que pour tout $n \in \mathbb{N}$, pour tout polynôme P unitaire de degré n , $\chi_{C_P} = P$.

Pour $n = 0$, l'égalité est immédiate : on a dans ce cas $P = 1$. Par ailleurs, la matrice C_P est la matrice vide, de déterminant 1.

Soit $n \geq 1$. Supposons la propriété vérifiée pour $n-1$. Soit $P = X^n + \sum_{k=0}^{n-1} a_k X^k$ un polynôme unitaire de degré n . On a

$$XI_n - C_P = \begin{pmatrix} X & 0 & \dots & 0 & a_0 \\ -1 & X & \dots & 0 & a_1 \\ 0 & -1 & \dots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & X + a_{n-1} \end{pmatrix}$$

Développons le déterminant de $XI_n - C_P$ par rapport à la première ligne. Le mineur de a_0 est un déterminant triangulaire, égal à $(-1)^{n-1}$. Comme a_0 se situe ligne 1, colonne n , le cofacteur de a_0 est $(-1)^{n+1}(-1)^{n-1} = 1$. On a donc

$$\det(XI_n - C_P) = a_0 + X \begin{vmatrix} X & 0 & \dots & 0 & a_1 \\ -1 & X & \dots & 0 & a_2 \\ 0 & -1 & \dots & 0 & a_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & X + a_{n-1} \end{vmatrix} = a_0 + X \chi_{C_Q}$$

où

$$Q = X^{n-1} + \sum_{k=0}^{n-2} a_{k+1} X^k$$

Par l'hypothèse de récurrence, le déterminant ci-dessus est égal à Q . On a donc

$$\chi_{C_P} = \det(XI_n - C_P) = a_0 + XQ = P$$

□

3 Le théorème de Cayley-Hamilton

Soit $f \in \mathcal{L}(E)$. Pour tout $v \in E$, soit

$$\mathcal{A}_v(f) = \{P \in \mathbb{K}[X] : P(f)(v) = 0\}$$

Les polynômes qui appartiennent à $\mathcal{A}_v(f)$ sont les polynômes annulateurs de f relativement à v .

Proposition 7. Pour tout $v \in E$, l'ensemble $\mathcal{A}_v(f)$ est un idéal de $\mathbb{K}[X]$.

Démonstration. Soit $v \in E$. Le polynôme nul appartient à $\mathcal{A}_v(f)$, donc $\mathcal{A}_v(f) \neq \emptyset$. Soient $P, Q \in \mathcal{A}_v(f)$. On a

$$(P - Q)(f)(v) = P(f)(v) - Q(f)(v) = 0$$

donc $P - Q \in \mathcal{A}_{f,v}$. Enfin, pour tout $P \in \mathbb{K}[X]$ et tout $Q \in \mathcal{A}_v(f)$,

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

donc $PQ \in \mathcal{A}_{f,v}$. $\square$

Proposition 8. Pour tout $v \in E$, $\mathcal{A}_v(f) \neq \{0\}$.

Démonstration. Soit $v \in E$. La famille $(f^k(v))_{k \in \llbracket 0, n \rrbracket}$ est une famille de $n+1$ vecteurs de l'espace E , de dimension n . Il existe donc $a_0, \dots, a_n \in \mathbb{K}$ pas tous nuls tels que

$$\sum_{k=0}^n a_k f^k(v) = 0$$

Soit

$$P = \sum_{k=0}^n a_k X^k$$

Le polynôme P n'est pas le polynôme nul et $P(f)(v) = 0$ donc $P \in \mathcal{A}_v(f)$. $\square$

Une autre preuve consiste à remarquer que pour tout $P \in \mathcal{A}(f)$, polynôme annulateur de f , on a $P(f)(v) = 0(v) = 0$. Ainsi,

$$\{0\} \subsetneq \mathcal{A}(f) \subseteq \mathcal{A}_v(f)$$

L'anneau $\mathbb{K}[X]$ est principal. Pour tout $v \in E$, on a donc

$$\mathcal{A}_v(f) = \mu_{f,v} \mathbb{K}[X]$$

où $\mu_{f,v}$ est un polynôme unitaire. Le polynôme $\mu_{f,v}$ est le *polynôme minimal de f relativement à v* . Remarquons que $\deg(\mu_{f,v}) \in \llbracket 0, n \rrbracket$. Remarquons également que $\mu_{f,v}$ divise μ_f puisque $\mu_f \in \mathcal{A}_{f,v}$.

Proposition 9. [CAYLEY-HAMILTON]

Pour tout $f \in \mathcal{L}(E)$, $\chi_f(f) = 0$.

Démonstration. Soit $v \in E$. Soit $m \in \llbracket 0, n \rrbracket$ le degré de $\mu_{f,v}$. Posons

$$\mu_{f,v} = X^m + \sum_{k=0}^{m-1} a_k X^k$$

où les a_k appartiennent à $\mathbb{K}$. La famille $(v, f(v), \dots, f^{m-1}(v))$ est libre, de cardinal m . En effet, supposons qu'une combinaison linéaire de ces vecteurs est nulle. Ceci équivaut à se donner un polynôme P de degré inférieur ou égal à $m-1$ tel que $P(f)(v) = 0$ c'est à dire $P \in \mathcal{A}_v(f)$. Comme $\deg(P) < \deg(\mu_{f,v})$, la minimalité de $\mu_{f,v}$ montre que $P = 0$. Complétons cette famille en une base de E :

$$\mathcal{B} = (v, f(v), \dots, f^{m-1}(v), u_{m+1}, \dots, u_n)$$

Remarquons que

$$f(f^{m-1}(v)) = f^m(v) = - \sum_{k=0}^{m-1} a_k f^k(v)$$

La matrice de f dans la base $\mathcal{B}$ est triangulaire par blocs. Posons

$$A = \text{mat}_{\mathcal{B}} f = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

où $B \in \mathcal{M}_m(\mathbb{K})$, $C \in \mathcal{M}_{n-m,m}(\mathbb{K})$ et $D \in \mathcal{M}_{n-m}(\mathbb{K})$. De plus,

$$B = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_{m-1} \end{pmatrix}$$

Autrement dit, B est la matrice compagnon de $\mu_{f,v}$ et donc $\chi_B = \mu_{f,v}$. On a

$$\begin{aligned} \chi_f &= \det(XI_n - A) \\ &= \det(XI_n - B) \det(XI_n - D) \\ &= \chi_B \chi_D \\ &= \mu_{f,v} \chi_D \end{aligned}$$

Ainsi, $\mu_{f,v}$ divise χ_f . Il existe donc $Q \in \mathbb{K}[X]$ tel que $\chi_f = Q\mu_{f,v}$. De là,

$$\begin{aligned} \chi_f(f)(v) &= (Q(f) \circ \mu_{f,v}(f))(v) \\ &= Q(f)(\mu_{f,v}(f)(v)) \\ &= Q(f)(0) \\ &= 0 \end{aligned}$$

Ceci étant vrai pour tout $v \in E$, il en résulte que $\chi_f(f) = 0$. $\square$