

Une démonstration du théorème de Cayley-Hamilton

Dans tout ce qui suit, E est un espace vectoriel de dimension finie sur le corps $\mathbb{K} \subset \mathbb{C}$, et f un endomorphisme de E fixé. On note χ_f le polynôme caractéristique de f .

On veut démontrer que $\chi_f(f) = 0$.

1. Sous-espace stable associé à un vecteur

Soit $x \in E \setminus \{0_E\}$. Pour tout $p \in \mathbb{N}^*$, considérons la famille

$$\mathcal{F}_p = (f^k(x))_{0 \leq k \leq p-1} = (x, f(x), f^2(x), \dots, f^{p-1}(x))$$

Puisque $x \neq 0_E$, la famille $\mathcal{F}_1 = (x)$ est libre ; mais, pour $p > \dim E$, le cardinal de \mathcal{F}_p est supérieur à la dimension de E , donc la famille est liée. Il existe donc un entier $n \geq 1$, qui est le plus grand des entiers $k \in \mathbb{N}^*$ tels que \mathcal{F}_k soit libre.

Par définition de n , \mathcal{F}_n est libre, mais \mathcal{F}_{n+1} est liée. Puisque \mathcal{F}_{n+1} s'obtient en ajoutant le vecteur $f^n(x)$ à la famille \mathcal{F}_n , c'est que $f^n(x)$ est combinaison linéaire des vecteurs de \mathcal{F}_n ; posons

$$f^n(x) = \sum_{k=0}^{n-1} a_k f^k(x) = a_0 x + a_1 f(x) + a_2 f^2(x) + \dots + a_{n-1} f^{n-1}(x) \quad (1)$$

Notons d'autre part F_x le sous-espace de E engendré par la famille \mathcal{F}_n , qui en est donc une base. Les images par f des vecteurs de cette base sont $f(x), f^2(x), \dots, f^n(x)$, qui sont tous dans F_x d'après ce qui précède : F_x est donc un sous-espace de E stable par f , et qui contient x .

On peut même noter que, si un sous-espace G contient x et est stable par f , il doit contenir tous les $f^k(x)$, et donc $F_x \subset G$; par suite F_x est en fait le plus petit, au sens de l'inclusion, des sous-espaces contenant x et stables par f .

2. Polynôme minimal de x

Avec les notations précédentes, posons $P_x = X^n - \sum_{k=0}^{n-1} a_k X^k$. La relation (1) donne $[P_x(f)](x) = 0_E$.

De plus, le fait que la famille \mathcal{F}_n est libre montre que, si Q est un polynôme non nul de degré inférieur strictement à n , alors $[Q(f)](x) \neq 0_E$.

Le polynôme P_x est unitaire et de degré minimal parmi les polynômes non nuls Q vérifiant $[Q(f)](x) = 0_E$; on l'appellera le polynôme minimal de x .

3. Étude de l'endomorphisme induit

Soit g l'endomorphisme induit par f sur F_x . Pour tout $k \in \mathbb{N}$, notons e_k le vecteur $f^k(x)$; la base \mathcal{F}_n de F_x s'écrit alors $(e_0, e_1, \dots, e_{n-1})$. On a d'autre part, pour tout $k \in \llbracket 0, n-2 \rrbracket$, $g(e_k) = f^{k+1}(x) = e_{k+1}$ et $g(e_{n-1}) = f^n(x) = \sum_{k=0}^{n-1} a_k e_k$. La matrice de g dans la base \mathcal{F}_n est donc

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & & \vdots & a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & a_{n-1} \end{pmatrix} \quad \text{et, pour tout } \lambda, \quad \chi_A(\lambda) = \begin{vmatrix} \lambda & 0 & \cdots & 0 & -a_0 \\ -1 & \lambda & \ddots & \vdots & -a_1 \\ 0 & -1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & \lambda & -a_{n-2} \\ 0 & \cdots & 0 & -1 & \lambda - a_{n-1} \end{vmatrix}$$

Effectuons sur ce déterminant les opérations $L_{n-1} \leftarrow L_{n-1} + \lambda L_n$, puis $L_{n-2} \leftarrow L_{n-2} + \lambda L_{n-1}$, et ainsi de suite jusqu'à $L_1 \leftarrow L_1 + \lambda L_2$. On obtient

$$\chi_A(\lambda) = \begin{vmatrix} 0 & \cdots & \cdots & 0 & \lambda^n - a_{n-1}\lambda^{n-1} - \cdots - a_1\lambda - a_0 \\ -1 & \ddots & & \vdots & \vdots \\ 0 & -1 & \ddots & \vdots & \lambda^3 - a_{n-1}\lambda^2 - a_{n-2}\lambda - a_{n-3} \\ \vdots & \ddots & \ddots & 0 & \lambda^2 - a_{n-1}\lambda - a_{n-2} \\ 0 & \cdots & 0 & -1 & \lambda - a_{n-1} \end{vmatrix}$$

Ne reste qu'à développer par rapport à la première ligne : on obtient que $\chi_A(\lambda)$ est égal au coefficient en haut à droite, autrement dit $\chi_A(\lambda) = P_x(\lambda)$.

4. Le théorème de Cayley-Hamilton

Puisque g est l'endomorphisme induit par f sur un sous-espace stable par f , on sait que $\chi_g = \chi_A = P_x$ divise χ_f ; soit donc Q un polynôme tel que $\chi_f = QP_x$.

On a alors $\chi_f(f) = Q(f) \circ P_x(f)$; puisque $[P_x(f)](x) = 0_E$, on a donc

$$[\chi_f(f)](x) = [Q(f)](0_E) = 0_E$$

Ceci est vrai pour tout x non nul de E , et évidemment encore vrai si $x = 0_E$; on a donc bien $\chi_f(f) = 0$.