Théorème de Cayley-Hamilton

Jean-François Burnol, 8 octobre 2018

Je remercie M. O. Serman pour des discussions relatives à ce texte. Je renvoie par ailleurs à son riche polycopié

http://math.univ-lille1.fr/~serman/agreg/alglineaireCours.pdf

pour aller (beaucoup) plus loin.

Soit E un espace vectoriel de dimension finie d sur un corps $\mathbb{K}$. Soit f un endomorphisme sur E et P_f le polynôme caractéristique associé. Si $\mathcal{B}$ est une base de E et $M_{\mathcal{B}}(f)$ la matrice de f dans cette base, on a $P_f = \det(XI_d - M_{\mathcal{B}}(f))$ et P_f est unitaire de degré d. [1]

Pour les preuves je supposerai connus les faits suivants :

— si $x \in E$, le plus petit espace E_x stable par f et contenant x est isomorphe au quotient $\mathbb{K}[X]/(P_x)$ avec P_x le polynôme unitaire de plus petit degré tel que $P_x(f)(x) = 0$,

— cet isomorphisme $\mathbb{K}[X]/(P_x) \to E_x$ se fait par ϕ qui envoie une classe $\widetilde{Q}$ sur $Q(f)(x)$,

— on a $f(\phi(\widetilde{Q})) = \phi(\widetilde{XQ})$, et plus généralement $Q_1(f)(\phi(\widetilde{Q_2})) = \phi(\widetilde{Q_1Q_2})$, car ce dernier est $(Q_1Q_2)(f)(x) = Q_1(f)(Q_2(f)(x))$ (on connaît les règles de manipulation des polynômes d'endomorphismes),

— le polynôme caractéristique de la restriction de f à E_x (qui est celui de la multiplication par $\widetilde{X}$ sur $\mathbb{K}[X]/(P_x)$) [2] est exactement P_x. On le vérifie par le calcul du polynôme caractéristique de la « matrice compagnon ». Je ne rentre pas dans les détails, tout cela est supposé connu.

Il faut l'avoir fait au moins une fois dans sa vie en détail.

Bien sûr, si on connaît ce qui précède il y a de bonnes chances qu'on sache déjà qu'on peut l'utiliser pour le théorème 1 que voici :

Théorème 1 (Cayley-Hamilton)**.** *On a* $P_f(f) = 0$ *dans* $\mathrm{End}(E)$.

Théorème 2. *Pour tout* $Q \in \mathbb{K}[X]$, $Q(f)$ *est inversible si et seulement si* Q *est premier avec* P_f.

Théorème 3. *Le polynôme caractéristique* P_f *divise une puissance du polynôme minimal* π_f.

Théorème 4. *Pour toute décomposition* $P_f = P_1P_2$ *en deux polynômes unitaires* premiers entre eux*,* P_1 *est le polynôme caractéristique de la restriction de* f *à* $\mathrm{Ker}\, P_1(f)$*, et en particulier la dimension de* $\mathrm{Ker}\, P_1(f)$ *est égale au degré de* P_1.

Preuve du théorème 1. Soit x dans E et E_x l'espace « cyclique » que ses itérés sous f engendrent (voir les remarques préliminaires). Le polynôme P_x divise P_f (raisonnement de matrice par blocs avec en haut à gauche la matrice compagnon) et $P_x(f)(x) = 0$ donc $P_f(f)(x) = 0$.

Comme x est quelconque $P_f(f) = 0$. □

Preuve du théorème 2. Si Q est premier avec P_f on a une identité de Bézout $UQ + VP_f = 1$. Donc $U(f)Q(f) = \mathrm{Id}_E$ par Cayley-Hamilton et $Q(f)$ est inversible.

Si Q n'est pas premier avec P_f on va faire une récurrence sur la dimension. Le théorème est vrai en dimension nulle (je laisse à vérifier. . .).

1er cas E n'admet aucun sous-espace V stable sous f avec $\{0\} \subsetneq V \subsetneq E$. Donc pour x non nul, $E = E_x$. L'action de $Q(f)$ sur E_x est comme la multiplication par Q sur $\mathbb{K}[X]/(P_x)$. On sait bien que cette multiplication est inversible si et seulement si Q est premier avec P_x. [3] Or $P_x = P_f$.

1. On trouve parfois la convention $\det(M_{\mathcal{B}}(f) - XI_d)$ mais il est bien plus commode d'avoir des polynômes unitaires.
2. Bien sûr je ne vais pas garder du tout cette notation de $\widetilde{X}$, pour la classe de X dans $K[X]/(P_x)$, j'écrirai comme tout le monde X si j'en ai besoin.
3. On n'en a pas vraiment besoin mais à propos ce P_x est un irréductible de $\mathbb{K}[X]$ vu l'hypothèse faite.

2e cas E admet un sous-espace stable sous f, notons-le V, avec $0 < \dim V < \dim E$. Par un raisonnement de matrices par blocs le polynôme P_f se factorise en $P_{f,V} P_{f,E/V}$. Donc soit Q n'est pas premier avec $P_{f,V}$ soit il n'est pas premier avec $P_{f,E/V}$. Par l'hypothèse de récurrence soit $Q(f)$ n'est pas inversible sur V soit $Q(f)$ passé au quotient n'est pas inversible sur le quotient E/V. Or le déterminant de $Q(f)$ est le produit de celui de sa restriction à V par celui de son passage au quotient sur E/V. Donc $Q(f)$ n'est pas inversible.

On peut organiser la preuve un peu différemment, par exemple en expliquant qu'il y a une trigonalisation par blocs avec des blocs diagonaux (de tailles variables) qui sont des matrices compagnons. □

Preuve du théorème 3. Par l'absurde : soit $P \mid P_f$ un facteur irréductible de P_f qui ne divise pas π_f. Par Bézout $UP + V\pi_f = 1$, et donc $P(f)$ est inversible, ce qui contredit le théorème 2. □

Preuve du théorème 4. Donc, $P = P_1 P_2$ avec P_1 et P_2 premiers entre eux. On a une identité de Bézout $UP_1 + VP_2 = 1$. On définit $E_1 = \operatorname{Ker} P_1(f)$ et $E_2 = \operatorname{Ker} P_2(f)$. Je rappelle (Lemme des noyaux) que l'on montre facilement à partir de $UP_1 + VP_2 = 1$ que E est la somme directe $E_1 \oplus E_2$.

Certainement E_1 et E_2 sont stables par f. Soit f_1 la restriction de f à E_1, et Q_1 son polynôme caractéristique. De même f_2 et Q_2.

Un vecteur de E_1 dans le noyau de $P_2(f_1)$ est dans $E_1 \cap E_2 = \{0\}$ donc $P_2(f_1)$ est inversible. Par le théorème 2, P_2 est premier avec Q_1.

Or Q_1 divise $P = P_1 P_2$, donc par le Lemme de Gauss, Q_1 divise P_1. Par un raisonnement analogue on montre que Q_2 divise P_2.

Or $E = E_1 \oplus E_2$, donc $P = Q_1 Q_2$. En comparant les degrés et en rappelant que tous nos polynômes sont unitaires on conclut $Q_1 = P_1$ et $Q_2 = P_2$. □

§§§

On aura compris qu'on a voulu éviter de parler de valeurs propres qui auraient peut-être nécessité l'utilisation d'un passage à une clôture algébrique de $\mathbb{K}$ (ou au moins une « extension des scalaires »).

Je signale au passage que même sur les nombres complexes, et même après avoir trigonalisé une matrice M de sorte qu'on peut sans perte de généralité la supposer de la forme :

$$M = \begin{pmatrix} \lambda_1 & * & \dots & \dots & * \\ 0 & \lambda_2 & * & \dots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \dots & \ddots & * \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix},$$

il n'est pas tout-à-fait trivial que $P_M(M) = 0$ (bien sûr on a noté $P_M = \det(XI_n - M)$). Une façon de le voir est de passer à l'endomorphisme associé ϕ sur $\mathbb{C}^n$ et de remarquer que $(\phi - \lambda_n)(\mathbb{C}^n) \subset V_{n-1}$, $(\phi - \lambda_{n-1})(V_{n-1}) \subset V_{n-2}$, …, avec V_j l'espace vectoriel engendré par les j premiers vecteurs de la base canonique et $V_0 = \{0\}$. Ainsi

$$P_\phi(\phi)(\mathbb{C}^n) = (\phi - \lambda_1 \operatorname{Id})\Big(\underbrace{(\phi - \lambda_2 \operatorname{Id})\big(\underbrace{\dots \underbrace{(\phi - \lambda_n \operatorname{Id})(\mathbb{C}^n)}_{\subset V_{n-1}} \dots}_{\subset V_2}\big)}_{\subset V_1}\Big) = \{0\}$$

Bien sûr si, mieux que trigonaliser, on utilise la décomposition en somme directe $\mathbb{C}^n = \oplus E_\lambda$ d'espaces caractéristiques, alors Cayley-Hamilton se voit plus immédiatement sur une matrice par blocs que dans le cas précédent triangulaire général.