Nous avons testé pour vous 18 démonstrations du théorème de Hamilton-Cayley

Michel Coste

Juin 2002

Pour une matrice $M \in \mathfrak{M}_n(A)$ (où A est un anneau commutatif) et $P_M(X) = \det(XI_n - M)$ son polynôme caractéristique, on a $P_M(M) = 0$.

Dans ce qui suit on identifiera M à l'endomorphisme $x \mapsto Mx$ de A^n.

1 Par diagonalisation ou trigonalisation

Le théorème de Hamilton-Cayley est à peu près trivial pour une matrice diagonale, donc aussi pour une matrice diagonalisable. Ceci conduit à deux démonstrations.

1.1 Trigonalisation

Ceci se passe sur un corps. A défaut de diagonalisation, on utilise la trigonalisation (ceci nécessite de passer au corps de décomposition du polynôme caractéristique). On se ramène à vérifier le théorème pour une matrice triangulaire, ce qui est tout de même un peu moins simple que pour une matrice diagonale ([Goua] 1e démonstration p. 177, [Gri], [Gob]).

1.2 Démonstration générique

Hamilton-Cayley est vrai pour la matrice générique : c'est la matrice carrée $G = (T_{i,j})_{i,j=1,\ldots,n}$ que l'on considère comme matrice à coefficients dans l'anneau de polynômes $R = \mathbb{Z}[T_{i,j}]$ que l'on peut plonger dans son corps de fractions $\mathbb{Q}(T_{i,j})$. Si $M \in \mathfrak{M}_n(A)$, il existe un unique morphisme $\theta : R \to A$ qui envoie G sur M (on spécialise G en M par θ). Le morphisme θ envoie toute identité algébrique vérifiée par les coefficients de G sur l'identité correspondante pour M. En particulier, il suffit de vérifier Hamilton-Cayley pour G pour l'obtenir pour n'importe quelle matrice à coefficients dans un anneau commutatif.

Or G vérifie Hamilton-Cayley parce qu'elle est diagonalisable sur le corps de décomposition de son polynôme caractéristique. Il suffit de vérifier que les valeurs propres de G sont distinctes, c'est à dire que le discriminant $\Delta \in R$ de son polynôme caractéristique est non nul. Pour le montrer on peut spécialiser G en une matrice diagonale M à éléments diagonaux distincts sur $\mathbb{C}$, par exemple. Alors Δ s'envoie sur le discriminant du polynôme caractéristique de M qui est non nul.

Pas de référence connue pour cette démonstration transmise par A. Ducros.

2 Utilisation de sous-espaces M-monogènes

On travaille sur un corps K. On prend un vecteur x non nul de K^n, P_x le polynôme unitaire engendrant l'idéal des polynômes P tels que $P(M)x = 0$. Alors P_x est le polynôme caractéristique de la restriction de M au sous-espace M monogène E_x qui est le plus petit contenant x et stable par M. En effet la matrice de la restriction de M dans la base $x, Mx, \ldots, M^{k-1}x$ de E_x (où $k = \dim E_x = \deg P_x$) est la matrice compagnon de P_x. Par ailleurs le polynôme caractéristique de la restriction de M à un sous-espace stable divise P_M, ce qui se voit en complétant une base. Donc $P_M(M)(x) = 0$. Voir [Goua], 2e démonstration p. 177, [Fre] p. 125, [Gos] P. 361.

3 Formule de la comatrice

Il y a une idée très simple mais qui ne marche pas : faire $X = M$ dans $P_M(X) = \det(XI_n - M)$. Ceci ne marche pas parce que le morphisme d'évaluation $A[X] \to A[M]$ envoie le terme de droite sur le déterminant de la matrice de coefficients $\delta_{i,j}M - m_{i,j}I_n$ dans $A[M]$ (où $\delta_{i,j}$ est le symbole de Kronecker), et pas sur le déterminant de la matrice 0 !

Cette idée de "faire $X = M$" marche de manière plus subtile, à partir de la formule de la comatrice

$$N\,{}^t(\mathrm{com}\,N) = {}^t(\mathrm{com}\,N)\,N = \det(N) I_n ,$$

où $\mathrm{com}\,N$ est la comatrice de la matrice carrée N de taille n. Cette formule est valable sur un anneau commutatif quelconque. (Si on l'a sur un corps, on l'obtient sur un anneau commutatif quelconque par le truc de la spécialisation de la matrice générique ci-dessus.) On applique cette formule à la matrice $N = XI_n - M$ à coefficients dans $A[X]$ pour obtenir une égalité entre éléments de $\mathfrak{M}_n(A[X])$:

$$(\dagger) \qquad (XI_n - M)\,{}^t(\mathrm{com}(XI_n - M)) = P_M(X) I_n .$$

3.1 Matrices à coefficients dans $A[M]$

On transpose (†) pour obtenir

$$\mathrm{com}(XI_n - M)\,(XI_n - {}^tM) = P_M(X) I_n .$$

On transporte cette identité par le morphisme de A-algèbres $\epsilon_M : A[X] \to A[M] \subset \mathfrak{M}_n(A)$ qui envoie X sur M (on notera encore $\epsilon_M : \mathfrak{M}_n(A[X]) \to \mathfrak{M}_n(A[M])$ le morphisme induit). La règle habituelle de multiplication ligne-colonne nous permet de multiplier à droite les matrices de $\mathfrak{M}_n(A[M])$ par le vecteur colonne $\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$, où $(e_1, \ldots, e_n)$ est la base canonique de A^n. On obtient ainsi

$$\epsilon_M(P_M(X) I_n) \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} P_M(M)e_1 \\ \vdots \\ P_M(M)e_n \end{pmatrix} .$$

Comme

$$\epsilon_M(XI_n - {}^tM) \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} Me_1 - \sum_{j=1}^n m_{j,1}e_j \\ \vdots \\ Me_n - \sum_{j=1}^n m_{j,n}e_j \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} ,$$

on en déduit $P_M(M)e_1 = \ldots = P_M(M)(e_n) = 0$ et donc $P_M(M) = 0$. Cette démonstration figure avec une erreur dans [Lan] p. 400 (il manque la transposition), et elle est détaillée dans [LFAr] p. 344, 1e démonstration.

3.2 Identification des coefficients des puissances de X

Une matrice de $\mathfrak{M}_n(A[X])$ s'écrit de manière unique sous la forme $X^dQ_d + \cdots + XQ_1 + Q_0$, où $Q_i \in \mathfrak{M}_n(A)$. Comme l'anneau $\mathfrak{M}_n(A)$ n'est pas commutatif, il est un peu délicat d'évaluer ces “polynômes à coefficients matriciels” en $X = M$. Voir cependant [Gan] p. 82, qui montre que si $X^dQ_d + \cdots + XQ_1 + Q_0$ est divisible à gauche par $XI_n - M$, alors $M^dQ_d + \cdots + MQ_1 + Q_0 = 0$ (dans le même esprit : [Lax] p. 51). Ce qui se fait le plus souvent (et revient en fait au même), c'est d'identifier les coefficients des puissances de X des deux côtés de (†) pour en déduire $P_M(M) = 0$ (voir [Tau] p. 186, [Ser] p. 20, [ArFr], [Fre] p. 133, [LFAr] p. 345 2e démonstration, [RDO] p. 407).

3.3 Suite exacte de $A[X]$-modules

La matrice $XI_n - M$ (comme toute matrice de $\mathfrak{M}_n(A[X])$) s'identifie à un endomorphisme de $A[X]$-module $(A[X])^n \to (A[X])^n$. Rappelons que la matrice M permet de munir A^n d'une structure de $A[X]$-module : pour $a_dX^d + \cdots + a_1X + a_0 \in A[X]$ et $x \in A^n$,

$$(a_dX^d + \cdots + a_1X + a_0)x = a_dM^dx + \cdots + a_1Mx + a_0x \ .$$

Dans ce cadre, “faire $X = M$” se traduit par le morphisme de $A[X]$-modules $\varphi : (A[X])^n \to A^n$ défini de la manière suivante. Tout élément de $(A[X])^n$ s'écrit de manière unique sous la forme $X^dx_d + \cdots + Xx_1 + x_0$ avec $x_i \in A^n$. On pose $\varphi(X^dx_d + \cdots + Xx_1 + x_0) = M^dx_d + \cdots + Mx_1 + x_0$. Alors “faire $X = M$ dans $XI_n - M$” se traduit par le fait que le composé $\varphi \circ (XI_n - M) : (A[X])^n \to A^n$ est nul (vérification facile). On en déduit Hamilton-Cayley en utilisant la formule de la comatrice (†) : pour tout $x \in A^n$, on a

$$\begin{aligned} P_M(M)x &= \varphi(P_M(X)x) = (\varphi \circ P_M(X)I_n)(x) \\ &= (\varphi \circ (XI_n - M) \circ {}^t(\mathrm{com}(XI_n - M)))(x) = 0 \ . \end{aligned}$$

C'est cette démonstration de Hamilton-Cayley qui est donnée dans [Bou] A III 107 et dans [Hou], p. 526.

On montre en fait dans ces références plus de choses sur le morphisme φ, à savoir que la suite

$$(A[X])^n \xrightarrow{XI_n - M} (A[X])^n \xrightarrow{\varphi} A^n \to 0$$

est exacte. Ceci veut dire que φ est surjective (facile) et que $\ker \varphi$ est l'image de $XI_n - M$. Autrement dit, A^n avec sa structure de $A[X]$-module donnée par M est isomorphe au quotient de $(A[X])^n$ par l'image de $XI_n - M$. On peut en tirer les conséquences suivantes :

1. Deux matrices M et M' de $\mathfrak{M}_n(A)$ sont semblables si et seulement si $XI_n - M$ et $XI_n - M'$ sont équivalentes. En effet si $XI_n - M$ et $XI_n - M'$ sont équivalentes, les quotients de $(A[X])^n$ par les images de $XI_n - M$ et $XI_n - M'$ sont isomorphes, et donc les matrices M et M' induisent des structures de $A[X]$-modules isomorphes sur A^n ; ceci veut dire qu'elles sont semblables. L'autre implication est évidente. Ce résultat est montré directement (sur un corps) dans [Ser], p. 69 et dans [Gan], p. 147.
2. Dans le cas où l'anneau de base A est un corps K, on sait que la matrice $XI_n - M$ est équivalente sur $K[X]$ à une matrice diagonale

$$\begin{pmatrix} Q_1 & & \\ & \ddots & \\ & & Q_n \end{pmatrix} ,$$

avec Q_i polynôme unitaire et Q_i divisant Q_{i+1} pour $i = 1, \dots, n-1$ (ce sont les facteurs invariants de $XI_n - M$). Remarquer qu'aucun Q_i n'est nul puisque le déterminant de $XI_n - M$, c'est-à-dire P_M, n'est pas nul. Soient $Q_r, \dots, Q_n$ ceux de degrés strictement positifs. L'exactitude de la suite dit que K^n, avec sa structure de $K[X]$-module donnée par M, est isomorphe au quotient de $(K[X])^n$ par l'image de $XI_n - M$. Il est donc aussi isomorphe à la somme directe

$$K[X]/Q_r \oplus \cdots \oplus K[X]/Q_n .$$

Remarquer que le facteur $K[X]/Q_i$ est nul si $Q_i = 1$. Les invariants de similitude de la matrice M sont les $Q_r, \dots, Q_n$. Le résultat du 1) ci-dessus implique que deux matrices sont semblables si et seulement si elles ont mêmes invariants de similitude.

Références

[ArFr] J-M. Arnaudies et H. Fraysse, Cours de mathématiques 1 - Algèbre, Dunod.

[Bou] N. Bourbaki, Algèbre chap. 3, Hermann.

[Fre] J. Fresnel, Algèbre matricielle, Hermann.

[Gan] F.R. Gantmacher, Théorie des matrices, tome 1, Dunod.

[Gob] R. Goblot, Algèbre linéaire, Masson.

[Gos] Gostiaux, Cours de mathématiques spéciales 1 - Algèbre, P.U.F.

[Goua] X. Gourdon, Les maths en tête - Algèbre, Ellipses 1994.

[Gri] J. Grifone, Algèbre linéaire, Cepadues-Éditions 1990.

[Hou] C. Houzel, Analyse mathématique, Belin 1996.

[Lan] S. Lang, Algebra, Addison-Wesley 1965.

[Lax] P.D. Lax, Linear algebra, John Wiley & Sons 1997.

[LFAr] J. Lelong-Ferrand et J-M. Arnaudies, Cours de mathématiques 1 - Algèbre, Dunod.

[RDO] E. Ramis, C. Deschamps et J. Odoux, Mathématiques spéciales 1 - Algèbre, Masson.

[Ser] D. Serre, Les matrices, Dunod 2001.

[Tau] P. Tauvel, Mathématiques générales pour l'agrégation, Masson 1992.