Nous avons testé pour vous 30 démonstrations du théorème de Cayley-Hamilton

Michel Coste

Février 2008

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 $P_M(M) = 0$.

Le but de cettte note est de présenter diverses démonstrations de ce théorème rencontrées dans des ouvrages variés, en les regroupant par types (je ne mettrai pas de notes aux démonstrations !). Je ne donne pas des démonstrations très détaillées, mais simplement les grandes lignes qui devraient permettre de reconstituer les démonstrations complètes. Il est recommandé de se reporter aux sources citées. Prière de me signaler toute erreur, ou toute démonstration différente de celles présentées en écrivant à `michel.coste@univ-rennes1.fr`

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

1 Par diagonalisation ou trigonalisation

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

1.1 Trigonalisation et variantes

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 ; la démonstration procède par récurrence ([Goua] 1e démonstration p. 175, [Gri], [Gob]).

Une façon de faire la récurrence sans vraiment passer par la trigonalisation est présentée dans [Mne], p.19 : si $P_M(X) = (X - \lambda_1) \cdots (X - \lambda_n)$, on applique l'hypothèse de récurrence à un hyperplan contenant l'image de $M - \lambda_n I_n$; cet hyperplan est stable par M et le polynôme caractéristique de la restriction de M à cet hyperplan est $(X - \lambda_1) \cdots (X - \lambda_{n-1})$.

Dans [Pra] (p.74), la récurrence est menée en utilisant un quotient : on applique l'hypothèse de récurrence au quotient par une droite propre D de valeur propre associée λ_1, sur lequel M induit un endomorphisme de polynôme caractéristique $Q = (X - \lambda_2) \cdots (X - \lambda_n)$; l'image de $Q(M)$ est contenue dans D.

1.2 Démonstration générique

Cayley-Hamilton est vrai pour la matrice générique : c'est la matrice carrée $G = (T_{i,j})_{i,j=1,\dots,n}$ 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 Cayley-Hamilton pour G pour l'obtenir pour n'importe quelle matrice à coefficients dans un anneau commutatif.

Or G vérifie Cayley-Hamilton 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.

Cette démonstration m'a été transmise par A. Ducros. Elle figure, avec un argument un peu différent pour vérifier que G a des valeurs propres distinctes, dans [Mer], p.229

Cet argument peut aussi se présenter de la manière suivante, comme un "prolongement d'identité" : soit $C_{k,\ell} \in \mathbb{Z}[T_{i,j}]$ le coefficient sur la ligne k et la colonne ℓ de $P_G(G)$ (avec G la matrice générique comme ci-dessus). La fonction polynomiale associée à $C_{k,\ell}$ est nulle sur l'ouvert non vide de $\mathfrak{M}_n(\mathbb{C})$ formé des matrices à valeurs propres distinctes. Donc $C_{k,\ell}$ est le polynôme nul.

Une version réduite du précédent argument peut aussi être utilisée pour démontrer Cayley-Hamilton sur $\mathbb{C}$ par prolongement d'identité : la fonction continue $M \mapsto P_M(M)$ est nulle sur l'ensemble des matrices diagonalisables qui est dense dans $\mathfrak{M}_n(\mathbb{C})$; cette démonstration figure dans [BMP], p.221.

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 $x, Mx, \dots, M^{k-1}x$ 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, [Rom], [Cog] p.295.

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

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

où $\operatorname{com} N$ 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 $\mathfrak{M}_n(A[X])$:

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

3.1 Matrices à coefficients dans $A[M]$

On transpose (†) pour obtenir

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

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 dans [Lan] p. 574; elle est détaillée dans [LFAr] p. 344, 1e démonstration et aussi dans [BrMa], p.80. La démonstration dans [Mer] p.372 est erronée : la matrice $\epsilon_M(XI_n - {}^tM)$ peut s'appliquer à un vecteur à composantes dans A^n, pas à un élément de A^n.

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 $X^d Q_d + \cdots + XQ_1 + Q_0$, 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 $X^d Q_d + \cdots + XQ_1 + Q_0$ est

divisible à gauche par $XI_n - M$, alors $M^d Q_d + \cdots + MQ_1 + Q_0 = 0$ (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 $P_M(M) = 0$. Cette preuve, due paraît-il à Corentin Hémery (le compère de Lebossé), peut se lire dans [Tau] p. 190, [Ser] p. 20, [ArFr], [Fre] p. 133, [LFAr] p. 345 2e démonstration, [RDO] p. 407).

Dans [Cog], p.298, l'argument d'identification est remplacé de la manière suivante (à peu près) : on utilise (†) et le fait qu'on peut mettre en facteur $XI_n - M$ dans $P_M(X)I_n - P_M(M)$ pour montrer qu'il existe U polynôme à coefficients dans $\mathfrak{M}_n(A)$ tel que $P_M(M) = (XI_n - M)U$; la comparaison des degrés en X donne $U = P_M(M) = 0$.

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 $a_d X^d + \cdots + a_1 X + a_0 \in A[X]$ et $x \in A^n$,

$$(a_d X^d + \cdots + a_1 X + a_0)x = a_d M^d x + \cdots + a_1 Mx + a_0 x \ .$$

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 $X^d x_d + \cdots + Xx_1 + x_0$ avec $x_i \in A^n$. On pose $\varphi(X^d x_d + \cdots + Xx_1 + x_0) = M^d x_d + \cdots + Mx_1 + x_0$. 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 Cayley-Hamilton 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 Cayley-Hamilton 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 $i = 1, \ldots, n-1$ (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 $Q_r, \ldots, Q_n$ 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 $Q_r, \ldots, Q_n$. Le résultat du 1) ci-dessus implique que deux matrices sont semblables si et seulement si elles ont mêmes invariants de similitude. On remarque aussi que $P_M(X)$ annule tout $K[X]$-module $K[X]/Q_i$, et donc aussi le $K[X]$-module K^n. Cette démonstration de Cayley-Hamilton pour un corps figure dans [Jac] p.201, [Gan] p. 200 et [Mer] p.272.

Références

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

[BMP] V. Beck, J. Malik et G. Peyré, Objectif agrégation, H&K 2004.

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

[BrMa] J. Briançon et P. Maisonobe, Éléments d'algèbre commutative, Ellipses 2004.

[Cog] M. Cognet, Algèbre linéaire, Bréal 2000

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

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

[Gob] R. Goblot, Algèbre linéaire, Ellipses 2005.

[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.

[Jac] N. Jacobson, Basic algebra I, W.H. Freeman 1985

[Lan] S. Lang, Algebre, Dunod 2002.

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

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

[Mer] J-Y. Mérindol, Nombres et algèbre, EDP Sciences 2006

[Mne] R. Mneimné, Réduction des endomorphismes, Calvage et Mounet 2006

[Pra] V.V. Prasolov, Problems and theorems in linear algebra, Amer. Math. Soc. 1994.

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

[Rom] J-E. Rombaldi, Analyse matricielle, EDP Sciences 1999

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

[Tau] P. Tauvel, Cours d'algèbre, Dunod 2005.