Théorème de Cayley-Hamilton

Alexey Muranov

29 avril 2025

1 Introduction

Théorème (Théorème de Cayley-Hamilton)**.** *Soit R un anneau commutatif unitaire (c'est-à-dire, avec l'unité 1_R). Soient A une matrice carrée à coefficients dans R et χ_A le polynôme caractéristique de A. Alors $\chi_A(A)$ est la matrice nulle.*

Il y a une démonstration très courte de ce théorème basée sur les deux ingrédients suivants :

(1) l'*isomorphisme* canonique $R^{n\times n}[X] \cong R[X]^{n\times n}$, qui permet d'« identifier » les matrices $n \times n$ à coefficients dans $R[X]$ avec les polynômes en X à coefficients dans $R^{n\times n}$ (on « identifie » matrices à coefficients polynômes avec polynômes à coefficients matrices),

(2) certaines *actions à gauche* de $R^{n\times n}[X]$ (et, donc, aussi de $R[X]^{n\times n}$) sur $R^{n\times n}$, associées aux éléments de $R^{n\times n}$.

Cette démonstration suit essentiellement celle de Serge Lang dans son livre *Algebra*, volume 1 (2002, 3e édition, en anglais).

2 Préliminaires

Pour tout matrice carrée A à coefficients dans un anneau commutatif unitaire,

$$({}^{\mathsf{t}}\operatorname{com} A)A = I \det A,$$

où I est la matrice identité (de la même taille que A).

Soit R est un anneau commutatif unitaire. Alors $R[X]$ l'est aussi. Soient A une matrice carrée à coefficients dans R et I la matrice identité de la même taille. Considérons la matrice carrée à coefficients dans $R[X]$ donnée par l'expression $A - IX$. On a :

$$\big({}^{\mathsf{t}}\operatorname{com}(A - IX)\big)(A - IX) = I \det(A - IX) = I\chi_A.$$

3 Démonstration

Dans cette section, R est un anneau commutatif unitaire, n est un nombre naturel, A est une matrice $n \times n$ à coefficients dans R, I est la matrice identité $n \times n$, O est la matrice nulle $n \times n$ ($A, I, O \in R^{n\times n}$, $n \in \mathbf{N}$).

Quelles que soient $\boldsymbol{B} \in R[X]^{n\times n}$ et $M \in R^{n\times n}$, avec

$$\begin{aligned}\boldsymbol{B} &= B_0 + B_1X + \cdots + B_pX^p \\ &= B_0X^0 + B_1X^1 + \cdots + B_pX^p, \qquad B_0, \ldots, B_p \in R^{n\times n},\end{aligned}$$

posons

$$\begin{aligned}\boldsymbol{B} \triangleleft_A M &\overset{\text{déf}}{=} B_0MI + B_1MA + \cdots + B_pMA^p \\ &= B_0MA^0 + B_1MA^1 + \cdots + B_pMA^p.\end{aligned}$$

Ainsi, en particulier,

$$\begin{aligned}\boldsymbol{B} \triangleleft_A I &= B_0II + B_1IA + \cdots + B_pIA^p \\ &= B_0I + B_1A + \cdots + B_pA^p = B_0 + B_1A + \cdots + B_pA^p.\end{aligned}$$

En vue de la démonstration du théorème de Cayley-Hamilton, observons que

$$(A - IX) \triangleleft_A I = AI - IA = A - A = O,$$

et que pour tout $P = \alpha_0 + \alpha_1X + \cdots + \alpha_pX^p \in R[X]$,

$$\begin{aligned}IP \triangleleft_A I &= (\alpha_0I + \alpha_1IX + \cdots + \alpha_pIX^p) \triangleleft_A I \\ &= \alpha_0II + \alpha_1IA + \cdots + \alpha_pIA^p \\ &= \alpha_0I + \alpha_1A + \cdots + \alpha_pA^p \\ &= P(A).\end{aligned}$$

On aura besoin de quelques propriétés de l'opération $(\triangleleft_A)$, présentées comme les lemmes suivants.

Lemme. *L'opération $(\triangleleft_A)$ est linéaire en second argument* :

$$\boldsymbol{B} \triangleleft_A (M + N) = \boldsymbol{B} \triangleleft_A M + \boldsymbol{B} \triangleleft_A N \qquad \textit{et} \qquad \boldsymbol{B} \triangleleft_A \alpha M = \alpha(\boldsymbol{B} \triangleleft_A M)$$

pour tous $\boldsymbol{B} \in R[X]^{n\times n}$, $M, N \in R^{n\times n}$, et $\alpha \in R$.

Exercice. Démontrer ce lemme.

Lemme. *L'opération $(\triangleleft_A)$ est linéaire en premier argument* :

$$(\boldsymbol{B} + \boldsymbol{C}) \triangleleft_A M = \boldsymbol{B} \triangleleft_A M + \boldsymbol{C} \triangleleft_A M \qquad \textit{et} \qquad \alpha\boldsymbol{B} \triangleleft_A M = \alpha(\boldsymbol{B} \triangleleft_A M)$$

pour tous $M \in R^{n\times n}$, $\boldsymbol{B}, \boldsymbol{C} \in R[X]^{n\times n}$, et $\alpha \in R$.

Exercice. Démontrer ce lemme.

Lemme. *Pour tous* $M \in R^{n\times n}$ *et* $\boldsymbol{B}, \boldsymbol{C} \in R[X]^{n\times n}$,

$$\boldsymbol{BC} \triangleleft_A M = \boldsymbol{B} \triangleleft_A (\boldsymbol{C} \triangleleft_A M).$$

Esquisse d'une démonstration. Il suffit de traiter le cas où $\boldsymbol{B} = BX^p$ et $\boldsymbol{C} = CX^q$, avec $B, C \in R^{n\times n}$ et $p, q \in \mathbf{N}$. Le cas général en résulte par bilinéarité. Or,

$$BX^pCX^q \triangleleft_A M = BCX^pX^q \triangleleft_A M = BCX^{p+q} \triangleleft_A M = BCMA^{p+q},$$

et

$$BX^p \triangleleft_A (CX^q \triangleleft_A M) = BX^p \triangleleft_A CMA^q = BCMA^qA^p = BCMA^{q+p}. \qquad \square$$

Démonstration du théorème de Cayley-Hamilton.

$$\begin{aligned}
\chi_A(A) &= I\chi_A \triangleleft_A I \\
&= I \det(A - IX) \triangleleft_A I \\
&= \big({}^{\mathbf{t}}\mathrm{com}(A - IX)\big)(A - IX) \triangleleft_A I \\
&= \big({}^{\mathbf{t}}\mathrm{com}(A - IX)\big) \triangleleft_A \big((A - IX) \triangleleft_A I\big) \\
&= \big({}^{\mathbf{t}}\mathrm{com}(A - IX)\big) \triangleleft_A O \\
&= O. \qquad \square
\end{aligned}$$

4 Généralisation

La méthode utilisée ci-dessus pour démontrer le théorème de Cayley-Hamilton permet de prouver d'autres énoncés d'apparence plus générale, comme la proposition suivante :

Proposition. *Soient* $A_1, \ldots, A_m, B_1, \ldots, B_m \in R^{n\times n}$ *telles que* :

(1) $B_1A_1 + \cdots + B_mA_m = O$,

(2) $A_1, \ldots, A_m$ *commutent deux à deux.*

Posons

$$\boldsymbol{B} = B_1X_1 + \cdots + B_mX_m \in R[X_1, \ldots, X_m]^{n\times n}$$

et

$$P = \det \boldsymbol{B} \in R[X_1, \ldots, X_m].$$

Alors

$$P(A_1, \ldots, A_m) = O.$$

Exemple. Soient $A, B \in R^{n\times n}$ telles que $BA = AB$. Posons

$$P = \det(BX - AY) \in R[X, Y].$$

Alors $P(A, B) = O$.

Exercice. Soient $A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}$ $(A, B \in \mathbf{Z}^{2\times 2})$. Posons

$$P = \det(BX - AY) \in \mathbf{Z}[X, Y].$$

(1) Calculer AB et BA.

(2) Calculer $P(A, B)$.

Idée d'une démonstration de la proposition. On peut définir une opération $(\triangleleft_{A_1,\dots,A_m})$ de sorte que pour tous $\boldsymbol{B} \in R[X_1, \dots, X_m]^{n\times n}$ et $M \in R^{n\times n}$, on ait $\boldsymbol{B} \triangleleft_{A_1,\dots,A_m} M \in R^{n\times n}$, et que

$$\begin{aligned}
&P(A_1, \dots, A_m) \\
&\quad = IP \triangleleft_{A_1,\dots,A_m} I \\
&\quad = I \det(B_1X_1 + \cdots + B_mX_m) \triangleleft_{A_1,\dots,A_m} I \\
&\quad = \big({}^{\mathbf{t}}\mathrm{com}(B_1X_1 + \cdots + B_mX_m)\big)(B_1X_1 + \cdots + B_mX_m) \triangleleft_{A_1,\dots,A_m} I \\
&\quad = \big({}^{\mathbf{t}}\mathrm{com}(B_1X_1 + \cdots + B_mX_m)\big) \triangleleft_{A_1,\dots,A_m} \big((B_1X_1 + \cdots + B_mX_m) \triangleleft_{A_1,\dots,A_m} I\big) \\
&\quad = \big({}^{\mathbf{t}}\mathrm{com}(B_1X_1 + \cdots + B_mX_m)\big) \triangleleft_{A_1,\dots,A_m} O \\
&\quad = O.
\end{aligned}$$

□

Proposition. *Soient $A_1, \dots, A_m, B_1, \dots, B_m \in R^{n\times n}$ telles que :*

(1) $A_1B_1 + \cdots + A_mB_m = O$,

(2) $A_1, \dots, A_m$ *commutent deux à deux.*

Posons

$$\boldsymbol{B} = B_1X_1 + \cdots + B_mX_m = X_1B_1 + \cdots + X_mB_m \in R[X_1, \dots, X_m]^{n\times n}$$

et

$$P = \det \boldsymbol{B} \in R[X_1, \dots, X_m].$$

Alors

$$P(A_1, \dots, A_m) = O.$$

Idée d'une démonstration. On peut définir une opération $(\triangleright_{A_1,\dots,A_m})$ de sorte que pour tous $\boldsymbol{B} \in R[X_1,\dots,X_m]^{n\times n}$ et $M \in R^{n\times n}$, on ait $M \triangleright_{A_1,\dots,A_m} \boldsymbol{B} \in R^{n\times n}$, et que

$$\begin{aligned}
&P(A_1,\dots,A_m)\\
&\quad= I \triangleright_{A_1,\dots,A_m} PI\\
&\quad= I \triangleright_{A_1,\dots,A_m} \det(X_1B_1+\cdots+X_mB_m)I\\
&\quad= I \triangleright_{A_1,\dots,A_m} (X_1B_1+\cdots+X_mB_m)\,{}^{\mathbf{t}}\mathrm{com}(X_1B_1+\cdots+X_mB_m)\\
&\quad= \big(I \triangleright_{A_1,\dots,A_m} (X_1B_1+\cdots+X_mB_m)\big) \triangleright_{A_1,\dots,A_m} {}^{\mathbf{t}}\mathrm{com}(X_1B_1+\cdots+X_mB_m)\\
&\quad= O \triangleright_{A_1,\dots,A_m} {}^{\mathbf{t}}\mathrm{com}(X_1B_1+\cdots+X_mB_m)\\
&\quad= O.
\end{aligned}$$

□

Si on veut chercher d'autres généralisations du théorème de Cayley-Hamilton, on peut se poser la question suivante, par exemple :

> Est-il vrai que si $A,B,C \in R^{n\times n}$, $BA-AC=O$, et qu'on pose
>
> $$P = \det(BX-XC) = \det(BX-CX) = \big(\det(B-C)\big)X^n \in R[X],$$
>
> alors $P(A)=O$?

Ceci n'est pas le cas, par exemple, pour $A=\begin{pmatrix}0&1\\1&0\end{pmatrix}$, $B=\begin{pmatrix}0&0\\1&0\end{pmatrix}$, $C=\begin{pmatrix}0&1\\0&0\end{pmatrix}$.