Théorème de Cayley-Hamilton : quatre démonstrations

Marc SAGE

26 septembre 2008 (màj 15 mars 2018)

Table des matières

1 Introduction **2**
- 1.1 Polynôme caractéristique d'une matrice . . . 2
- 1.2 Polynôme caractéristique d'un endomorphisme . . . 2

2 Première méthode : les matrices compagnes **2**
- 2.1 Lemme de la matrice compagne . . . 2
- 2.2 Démonstration de Cayley-Hamilton . . . 3

3 Seconde méthode : la comatrice **3**
- 3.1 Comatrice et $M_n(K[X])$ . . . 3
- 3.2 Démonstration de Cayley-Hamilton . . . 4

4 Troisème méthode : densité des matrices diagonalisables dans $M_n(\mathbb{C})$ **5**
- 4.1 Cayley-Hamilton pour les matrices diagonalisables . . . 5
- 4.2 Cayley-Hamilton pour les matrices complexes . . . 6

5 Quatrième méthode : par une représentation intégrale des puissances d'une matrices **6**

Résumé

On démontre ici le théorème de Cayley-Hamilton de quatre manières différentes : par les matrices compagnons, en utilisant la comatrice, en invoquant la densité des matrices diagonalisables dans $\mathcal{M}_n(\mathbb{C})$ puis écrivant les puissances entières d'une matrice sous forme intégrale.

1 Introduction

On rappelle qu'un corps K s'injecte canoniquement dans les K-algèbres $M_n(K)$, un scalaire λ étant identifié à la matrice scalaire[1] λI_n, laquelle nous noterons plutôt λ par soucis de clarté.

1.1 Polynôme caractéristique d'une matrice

Soit K un corps commutatif et $n \in \mathbb{N}^*$ un entier non nul.
On définit le *polynôme caractéristique* d'une matrice $A \in M_n(K)$ comme le polynôme

$$\chi_A = \det(X - A).$$

Notre but est de montrer que χ_A annule A, *i.e.*

$$\chi_A(A) = 0.$$

Attention au raisonnement naïf (et faux) qui consiste à dire : on remplace X par A dans la formule du polynôme χ_A, ce qui donne

$$\chi_A(A) = \det(A - A) = \det(A - A) = \det 0 = 0.$$

Déjà, écrire $\chi_A(A) = \det(A - A)$ fournit un calcul de déterminant dans le corps K, ce qui donne un scalaire et non une matrice comme ça devrait, d'où problème.
L'erreur réside en fait dans l'ordre de étapes « évaluation du déterminant » et « subtitution de A à X ».
Quand on écrit proprement

$$\chi_A(A) = [\det(X - A)](A),$$

on calcule le déterminant dans l'anneau $K[X]$, *puis* on évalue ce polynôme en la matrice A dans l'anneau $M_n(K)$.

1.2 Polynôme caractéristique d'un endomorphisme

On rappelle[2] que pour deux matrices A et B on a $\chi_{AB} = \chi_{BA}$. Ceci montre en particulier que $\chi_{PAP^{-1}} = \chi_A$ pour toute matrice P inversible et permet donc de définir le polynôme caractéristique d'un endomorphisme $f \in \mathcal{L}(E)$ où E est un K-espace vectoriel de dimension finie non nulle par

$$\chi_f = \chi_{\mathrm{Mat}_{\mathcal{B}} f}$$

où $\mathcal{B}$ est une base quelconque de E.

2 Première méthode : les matrices compagnes

2.1 Lemme de la matrice compagne

Pour $P = X^n + a_{n-1}X^{n-1} + ... + a_1X + a_0$ polynôme unitaire, on définit la *matrice compagne* de P :

$$\mathcal{C}(P) := \begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -a_{n-2} \\ & & 1 & -a_{n-1} \end{pmatrix};$$

pour $n = 1$, noter que P s'écrit $X + a_0$ et $\mathcal{C}(P) = \mathcal{C}(X + a_0) = (-a_0)$.

[1] d'où la terminologie
[2] *cf.* feuille *Une identité classique sur le polynôme caractéritique* pour plusieurs démonstrations

$\mathcal{C}(P)$ s'appelle matrice *compagne* pour une bonne raison : son polynôme caractéristique se lit dedans :

$$\chi_{\mathcal{C}(P)} = P.$$

Pour montrer cela, on développe $\chi_{\mathcal{C}(P)} = \begin{vmatrix} X & & & a_0 \\ -1 & \ddots & & \vdots \\ & \ddots & X & a_{n-2} \\ & & -1 & X + a_{n-1} \end{vmatrix}$ selon sa dernière colonne ; en effet, le cofacteur associé à a_i est simple à calculer car il **doit** valoir X^i (on veut trouver $P = X^n + a_{n-1}X^{n-1} + ... + a_1X + a_0$). De fait, le mineur associé à a_i est formé d'une matrice triangulaire supérieure de taille $n-1$ de diagonale $(X, ..., X, -1, ..., -1)$ où X apparaît i fois, donc vaut $(-1)^{n-1-i} X^i$, d'où le cofacteur X^i en multipliant par le signe[3] $(-1)^{n+(i+1)}$. On vérifiera bien que le dernier mineur, associé à $X + a_{n-1}$, vaut X^{n-1} et nous donne ce que l'on veut.

On pourrait également profiter de la présence des zéros et développer selon la première ligne ; le résultat s'obtiendrait aisément par récurrence.

2.2 Démonstration de Cayley-Hamilton

Soit E un K-espace vectoriel de dimension $n \geq 1$ et $f \in \mathcal{L}(E)$. On se fixe un v dans K^n, et on va montrer que $\chi_f(f)(v) = 0$. En faisant varier v dans E, on obtiendra ainsi $\chi_f(f) = 0$ comme voulu.

Pour $v = 0$, c'est évident par linéarité. Pour $v \neq 0$, *i.e.* (v) libre, on considère

$$\nu = \min \left\{ k \geq 2 \; ; \left(v, f(v), ..., f^k(v) \right) \text{ liée} \right\},$$

de sorte que $f^\nu(v)$ est lié avec $\left(v, ..., f^{\nu-1}(v)\right)$ qui est libre, donc se décompose en

$$f^\nu(v) = a_0 v + a_1 f(v) + ... + a_{\nu-1} f^{\nu-1}(v).$$

On complète ensuite la famille libre $\left(v, f(v), ..., f^{\nu-1}(v)\right)$ en une base $\mathcal{B}$ de K^n, dans laquelle f s'écrit

$$\operatorname*{Mat}_{\mathcal{B}} f = \begin{pmatrix} 0 & & & a_0 & * \\ 1 & \ddots & & \vdots & * \\ & \ddots & 0 & a_{\nu-2} & * \\ & & 1 & a_{\nu-1} & * \\ 0 & \cdots & \cdots & 0 & M \end{pmatrix} = \begin{pmatrix} \mathcal{C}(P) & * \\ 0 & M \end{pmatrix}$$

où l'on a noté $P = v^\nu - a_{\nu-1} v^{\nu-1} - ... - a_1 v - a_0$. Observer que $P(f)(v) = 0$. En utilisant le lemme de la matrice compagne, on obtient

$$\chi_f = \chi_{\begin{pmatrix} \mathcal{C}(P) & * \\ 0 & M \end{pmatrix}} = \chi_{\mathcal{C}(P)} \times \chi_M = P \times \chi_M,$$

d'où en appliquant en v :

$$\chi_f(f)(v) = [\chi_M \times P](f)(v) = \chi_M \left(\underbrace{P(f)(v)}_{=0} \right) = 0, \textit{ CQFD}$$

3 Seconde méthode : la comatrice

3.1 Comatrice et $M_n(K[X])$

Pour une matrice M sur un anneau commutatif unitaire – par exemple $K[X]$ –, on rappelle que l'on a

$$M\ {}^t\text{com}\, M = \det M.$$

[3] Attention, le coefficient a_i est situé à la $(i+1)$-ième ligne...

Remarquer aussi que l'on dispose d'un isomorphisme d'algèbres évident

$$M_n(K[X]) \cong M_n(K)[X]$$

qui permet d'écrire par exemple

$$\begin{pmatrix} X^3+X & 2 \\ 7X^3 & -X \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 7 & 0 \end{pmatrix} X^3 + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} X + \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$$

Cet isomorphisme étant signalé, montrons à présent pourquoi la démarche naïve exposée en introduction ne pouvait pas fonctionner (sauf cas trivial à préciser). Il s'agissait de voir si le diagramme suivant est commutatif :

$$\begin{array}{ccccc} M_n(K)[X] & \cong & M_n(K[X]) & \stackrel{\det}{\longrightarrow} & K[X] \\ \downarrow \text{eval}_A & & & & \downarrow \text{eval}_A \\ M_n(K) & \stackrel{\det}{\longrightarrow} & K & \hookrightarrow & M_n(K) \end{array}.$$

Supposons qu'il l'est et soit λ une valeur propre de A. Le polynôme $I_n(X-\lambda)$ a alors pour image 0 (chemin du haut) et $(A-\lambda)^n$ (chemin du bas), d'où une nilpotente N telle que $A=\lambda+N$. Son polynôme minimal est donc de la forme X^ν pour un certain naturel ν. Le polynôme $I_n(X+1-\lambda)$ a alors pour image $(N+1)^n=1$, donc ce polynôme minimal X^ν divise $(X+1)^n-1$, d'où $\nu=1$, $N=0$ et $A=\lambda$. Réciproquement, si A est scalaire, tout développer (les vecteurs selon la base $(E_{i,j}X^d)$, le déterminant selon son expression polynomiale) établirait la commutativité du diagramme.

Conclusion : la démarche naïve fonctionne ssi A est scalaire, auquel cas la nullité de $\chi_A(A)$ était déjà immédiate à vérifier.

Remarque subtile[4]. L'anneau $M_n(K)$ des coeffcients polynomiaux ci-dessus n'étant *pas* commutatif, même si l'indéterminée X de $M_n(K)[X]$ commute toujours avec chaque "coefficient" de l'anneau $M_n(K)$, le morphisme de groupes additif $\text{eval}_A := \left\{ \begin{array}{ccc} M_n(K)[X] & \longrightarrow & M_n(K) \\ \sum_{i\geq 0} m_i X^i & \longmapsto & \sum_{i\geq 0} m_i A^i \end{array} \right.$ ne préserve plus forcément la multiplication ! En effet, les égalités $\forall P,Q\in M_n(K)[X]$, $P(A)\,Q(A)=[PQ](A)$ forceraient pour chaque matrice m l'égalité (imposer $Q=m$ constant et $P=X$)

$$X(A)\,m(A) = [Xm](A) \stackrel{Xm=mX}{=} [mX](A)\text{, d'où } Am=mA$$

et A serait donc dans le *centre* de l'anneau des coefficients. Réciproquement, étant donnés deux polynômes $P,Q\in M_n(K)[X]$, si A commute avec chaque coefficient de Q, on pourra écrire

$$\begin{aligned} P(A)\,Q(A) &= \sum_{i\in\mathbf{N}} p_i A^i \sum_{j\in\mathbf{N}} q_j A^j = \sum_{i,j\in\mathbf{N}} p_i A^i q_j A^j = \sum_{i,j\in\mathbf{N}} p_i q_j A^i A^j \\ &= \sum_{i,j\in\mathbf{N}} p_i q_j A^{i+j} = \sum_{k\in\mathbf{N}} \left(\sum_{i+j=k} p_i q_j \right) A^k = [PQ](A). \end{aligned}$$

C'est cette dernière réciproque – plutôt fine – que nous utiliserons, la condition "être dans le centre" étant trop forte pour nos besoins.

3.2 Démonstration de Cayley-Hamilton

Soit $A\in M_n(K)$ où n est un entier ≥ 1.

On a vu que la méthode naïve (et fausse) qui consistait à remplacer $X\leftarrow A$ dans $\det(X-A)$ bloquait car l'on évaluait suivant $X\leftarrow A$ avant de calculer le déterminant. Afin de pouvoir remplacer $X\leftarrow A$ en toute impunité, nous allons nous débarrasser de ce dernier à l'aide de la comatrice, tout en gardant le précieux facteur annulateur $X-A$.

Tout part de la remarque élémentaire suivante : le polynôme caractéristique χ_A est le déterminant de la matrice $M:=X-A$ de $M_n(K[X])$. Nous pouvons donc écrire, dans $M_n(K[X])$, l'identité de la comatrice sous la forme

$$\chi_A I_n = (\det M)\,I_n = M\, \underbrace{{}^t\text{com}\, M}_{=:P} = P\times(X-A)\,,$$

l'envoyer de $M_n(K[X])$ dans $M_n(K)[X]$, puis l'évaluer suivant $X\leftarrow A$. À gauche, en notant $\chi_A =: \sum_{i=0}^n c_i X^i$, le polynôme $\chi_A I_n = \sum_{i=0}^n (c_i I_n) X^i$ est envoyé sur $\sum_{i=0}^n (c_i I_n) A^i = \sum_{i=0}^n c_i A^i = \chi_A(A)$. À droite, le facteur $I_n X - A$ a tous ses coefficients (I_n et A) qui commutent à A, donc (par la réciproque fine

[4] Je remercie ici Silvain Dupertuis pour m'avoir signalé en mars 2018 une erreur que je traînais depuis 2005 : quand l'anneau de base n'est plus commutatif, l'évaluation polynomiale perd sa multiplicativité ! À lui également de m'avoir indiqué comment "préserver" cette multiplicativité dans le cas particulier qui se présente.

précédente[5]) le produit $P \times (X - A)$ est envoyé sur $P(A) \times [X - A](A)$ dont le facteur de droite est nul, ce qui conclut.

Remarque. Dans le cas complexe, on peut expliciter le (?) ci-dessus et donner une preuve un tantinet différente. Pour $z \neq 0$ assez petit, la matrice $1 - zA$ s'inverse en $\sum_{k \geq 0} (zA)^k$, ce qui permet d'écrire

$$\det(1 - zA) \sum_{k \geq 0} (zA)^k = \frac{\det(1 - zA)}{1 - zA} = \operatorname{Com}(1 - zA).$$

En voyant l'égalité précédente dans $M_n(\mathbb{C})[X]$, on voit que la matrice-coefficient en z^n est nulle à droite (ses coordonnées sont des cofacteurs !), donc celui à gauche aussi. En notant $\chi_A =: \sum a_i X^i$, le facteur $\det(1 - zA)$ à gauche se réécrit $z^n \chi_A\left(\frac{1}{z}\right) = \sum a_i z^{n-i}$, d'où la matrice-coefficient en z^n à gauche :

$$\operatorname*{Coef}_{z^n}\left(\sum_{i=0}^{n} a_i z^{n-i} \sum_{i \geq 0} A^i z^i\right) = \sum a_i A^i = \chi_A(A), \textit{ CQFD.}$$

4 Troisème méthode : densité des matrices diagonalisables dans $M_n(\mathbb{C})$

Dans cette partie, K est pris égal à $\mathbb{C}$.

4.1 Cayley-Hamilton pour les matrices diagonalisables

Montrons que Cayley-Hamilton est déjà vrai pour les matrices diagonalisables.
Soit A une telle matrice. On peut écrire

$$A = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix} P^{-1},$$

de sorte que $\chi_A = \prod_{i=1}^{n} (X - \lambda_i)$ et donc

$$\begin{aligned} P^{-1} \chi_A(A) P &= \chi_A \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix} = \prod_{i=1}^{n} \left[\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix} - \lambda_i \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}\right] \\ &= \prod_{i=1}^{n} \begin{pmatrix} \ddots & & & & \\ & \lambda_{i-1} - \lambda_i & & & \\ & & 0 & & \\ & & & \lambda_{i+1} - \lambda_i & \\ & & & & \ddots \end{pmatrix} = 0 \end{aligned}$$

car un zéro apparaît dans chaque produit formant un coefficient diagonal.

Remarquer que l'on n'utilise nullement ici les propriétés topologiques de $\mathbb{C}$, le résultat est valable sur un corps quelconque.

D'ailleurs, en adoptant le point de vue des endomorphismes, si v est un vecteur propre associé à une valeur propre λ d'un endomorphisme f, on peut écrire (en dimension finie)

$$\chi_f(f)(v) = \left[\frac{\chi_f}{v - \lambda}(f)\right] \underbrace{(f(v) - \lambda v)}_{=0} = 0,$$

donc $\chi_f(f)$ s'annule sur l'engendré des vecteurs propres, lequel forme tout l'espace si f est diagonalisable.

[5] Si l'on voulait directement utiliser le caractère *multiplicatif* de l'évaluation en A, il aurait fallu (d'après la précédente remarque subtile) se restreindre aux matrices *centrales*, *i. e.* scalaires, ce qui aurait fortement nuit à la généralité. D'où l'intérêt de la réciproque fine mise à jour.

4.2 Cayley-Hamilton pour les matrices complexes

Soit maintenant A quelconque dans $M_n(\mathbb{C})$. On sait[6] que l'on peut l'approcher par une suite de matrices diagonalisables $A = \lim D_k$ où chaque D_k vérifie par ce qui précède $\chi_{D_k}(D_k) = 0$. Or, on dispose du lemme suivant :

Lemme[7].
Soit $n \in \mathbb{N}$. Alors

$$\left\{ \begin{array}{l} P_k \overset{k\infty}{\longrightarrow} P \text{ dans } \mathbb{C}_n[X] \\ A_k \overset{k\infty}{\longrightarrow} A \text{ dans } M_n(\mathbb{C}) \end{array} \right. \Longrightarrow P_k(A_k) \longrightarrow P(A) \text{ dans } M_n(\mathbb{C}).$$

Pour conclure, on applique le lemme aux suites $\left\{ \begin{array}{c} \chi_{D_k} \longrightarrow \chi_A \\ D_k \longrightarrow A \end{array} \right.$ (invoquer la continuité de $M \mapsto \chi_M$), d'où $\chi_{D_k}(D_k) \longrightarrow \chi_A(A)$ et $\chi_A(A) = 0$ par unicité de la limite.

5 Quatrième méthode : par une représentation intégrale des puissances d'une matrices

On se donne une matrice A complexe. Pour z assez grand, la matrice $z - A$ est inversible, donc pour r assez grand l'intégrale $\int_{-\pi}^{\pi} \frac{\left(re^{i\theta}\right)^{k+1}}{re^{i\theta}-A} \frac{d\theta}{2\pi}$ a un sens pour tout entier $k \geq 0$. On peut la calculer en développant l'inverse en série entière :

$$\int_{-\pi}^{\pi} \frac{\left(re^{i\theta}\right)^{k+1}}{re^{i\theta} - A} \frac{d\theta}{2\pi} = \int_{-\pi}^{\pi} \frac{\left(re^{i\theta}\right)^{k}}{1 - \frac{A}{re^{i\theta}}} \frac{d\theta}{2\pi} = \int_{-\pi}^{\pi} \sum_{l \geq 0} \left(\frac{A}{re^{i\theta}}\right)^l \left(re^{i\theta}\right)^k \frac{d\theta}{2\pi} \overset{?}{=} \sum_{l \geq 0} \left(\int_{-\pi}^{\pi} \left(re^{i\theta}\right)^{k-l} \frac{d\theta}{2\pi}\right) A^l = A^k.$$

On peut intervertir car la somme en prenant les modules converge :

$$\int_{-\pi}^{\pi} \sum_{l \geq 0} \left\| A^l \left(re^{i\theta}\right)^{k-l} \right\| \frac{d\theta}{2\pi} \leq r^k \int_{-\pi}^{\pi} \sum_{l \geq 0} \left(\frac{\|A\|}{r}\right)^l \frac{d\theta}{2\pi} = \frac{r^k}{1 - \frac{\|A\|}{r}}.$$

Ainsi, en notant $\chi_A =: \sum a_k X^k$, on aura

$$\begin{aligned} \chi_A(A) &= \sum_{k \geq 0} a_k A^k = \sum a_k \int_{-\pi}^{\pi} \frac{\left(re^{i\theta}\right)^{k+1}}{re^{i\theta} - A} \frac{d\theta}{2\pi} = \int_{-\pi}^{\pi} re^{i\theta} \frac{\chi_A\left(re^{i\theta}\right)}{re^{i\theta} - A} \frac{d\theta}{2\pi} = \int_{-\pi}^{\pi} re^{i\theta} \frac{\det\left(re^{i\theta} - A\right)}{re^{i\theta} - A} \frac{d\theta}{2\pi} \\ &= \int_{-\pi}^{\pi} re^{i\theta} \; {}^t\mathrm{com}\left(re^{i\theta} - A\right) \frac{d\theta}{2\pi}. \end{aligned}$$

Tout les coefficients de la matrice-intégrande sont des poylnômes en $re^{i\theta}$ sans coef constant (à cause du $re^{i\theta}$ devant en facteur), donc l'intégrale est nulle, *CQFD*.

[6] *cf.* feuille d'exercices sur la topologie des matrices
[7] *cf.* la feuille d'exercices sur les espaces vectoriels normés