

Polynôme minimal - Théorème de Cayley-Hamilton

Exercice 3* — 1. Soient E_1, \dots, E_n les vecteurs de la base canonique de \mathbb{C}^n . Par définition, $JE_1 = E_n, JE_2 = E_1, \dots, JE_n = E_{n-1}$; on a donc $J^2E_1 = E_{n-1}, J^2E_2 = E_n, J^2E_3 = E_2, \dots, J^2E_n = E_{n-2}$ et, en raisonnant par récurrence, on vérifie que

$$J^p E_i = \begin{cases} E_{n-p+i} & \text{si } 1 \leq i \leq p-1 \\ E_n & \text{si } i = p \\ E_{i-p} & \text{si } p+1 \leq i \leq n \end{cases}$$

pour tout $p \in \{1, \dots, n\}$. La traduction matricielle est fort simple : J^p est la matrice obtenue en séparant les $p-1$ premières colonnes de J des $n-p+1$ dernières et en permutant ces deux paquets.

Lorsque $p = n$, les formules précédentes donnent $J^n E_i = E_i$ pour tout $i \in \{1, \dots, n\}$ et donc $J^n = I_n$. La matrice J est ainsi annihilée par le polynôme $T^n - 1$; comme ce dernier est scindé à racines simples (ce sont les racines n -èmes de l'unité), il en est de même du polynôme minimal de J et la matrice J est donc diagonalisable.

Étant donnés des scalaires $\lambda_0, \lambda_1, \dots, \lambda_{n-1} \in \mathbb{C}$ tels que la matrice $\lambda_0 I_n + \lambda_1 J + \dots + \lambda_{n-1} J^{n-1}$ soit nulle,

$$0 = (\lambda_0 I_n + \lambda_1 J + \dots + \lambda_{n-1} J^{n-1}) E_1 = \lambda_0 E_1 + \lambda_1 E_n + \dots + \lambda_{n-1} E_2$$

et donc $\lambda_0 = \lambda_1 = \dots = \lambda_{n-1} = 0$. Cela établit que les matrices I_n, J, \dots, J^{n-1} constituent une famille libre dans $\mathcal{M}_n(\mathbb{C})$.

Soit m_J le polynôme minimal de J . D'une part, ce polynôme est de degré au moins n car, ainsi qu'on vient de le voir, il n'existe pas de polynôme non nul et de degré inférieur ou égal à $n-1$ annihilant J ; d'autre part, m_J divise le polynôme $T^n - 1$ puisque ce dernier annule J . On a donc $m_J = \lambda(T^n - 1)$ avec $\lambda \in \mathbb{C}^\times$, puis $m_J = T^n - 1$ car le polynôme minimal d'un endomorphisme est *unitaire* (par définition).

Les valeurs propres de J sont les racines du polynôme $T^n - 1$, à savoir les n racines n -èmes de l'unité dans \mathbb{C} : $1, \omega, \omega^2, \dots, \omega^{n-1}$ avec $\omega = e^{2i\pi/n}$.

Le vecteur $E(1) = E_1 + \dots + E_n$ est manifestement un vecteur propre pour J associé à la valeur propre 1. Plus généralement, si ξ est une racine n -ème de l'unité, le vecteur $E(\xi) = E_1 + \xi E_2 + \xi^2 E_3 + \dots + \xi^{n-1} E_n$ est un vecteur propre pour J associé à la valeur propre ξ car

$$\begin{aligned} JE(\xi) &= J(E_1 + \xi E_2 + \xi^2 E_3 + \dots + \xi^{n-1} E_n) = E_n + \xi E_1 + \xi^2 E_2 + \dots + \xi^{n-1} E_{n-1} \\ &= \xi(E_1 + \xi E_2 + \dots + \xi^{n-1} E_n) = \xi E(\xi). \end{aligned}$$

En désignant par ξ_1, \dots, ξ_n les n racines n -èmes de l'unité dans \mathbb{C} , les vecteurs $E(\xi_1), E(\xi_2), \dots, E(\xi_n)$ constituent donc une base de \mathbb{C}^n diagonalisant J et la matrice de passage correspondante est

$$P = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \xi_1 & \xi_2 & \dots & \xi_n \\ \xi_1^2 & \xi_2^2 & \dots & \xi_n^2 \\ \vdots & \vdots & \dots & \vdots \\ \xi_1^{n-1} & \xi_2^{n-1} & \dots & \xi_n^{n-1} \end{pmatrix}$$

2. En vertu du calcul des puissances de J à la question 1, $A = a_1 I_n + a_2 J + \dots + a_n J^{n-1}$.

Soit P la matrice de passage considérée à la fin de la question 1. Puisque $J = PDP^{-1}$ avec $D = \text{diag}(\xi_1, \xi_2, \dots, \xi_n)$ (ξ_1, \dots, ξ_n sont les n racines n -èmes de l'unité dans \mathbb{C}), $J^p = PD^p P^{-1}$ pour tout entier naturel p et donc

$$Q(J) = PQ(D)P^{-1}$$

pour tout polynôme $Q \in \mathbb{C}[T]$. La matrice $Q(D)$ étant diagonale, égale à $\text{diag}(Q(\xi_1), \dots, Q(\xi_n))$, la matrice $Q(J)$ est diagonalisable et ses valeurs propres sont les évaluations de Q en les valeurs propres de J :

$$\text{Sp}(Q(J)) = \{Q(\lambda), \lambda \in \text{Sp}(J)\}.$$

Comme $A = Q(J)$ avec $Q = a_1 + a_2 T + \dots + a_n T^{n-1}$, cette matrice est diagonalisable et ses valeurs propres sont les nombres complexes $a_1 + a_2 \xi + \dots + a_n \xi^{n-1}$, ξ parcourant l'ensemble des racines n -èmes de l'unité dans \mathbb{C} .

Le déterminant de A est le produit des valeurs propres comptées avec leur multiplicité, soit

$$\det(A) = \prod_{i=1}^n Q(\xi_i).$$

Exercice 4* — 1. Un endomorphisme v de E admet 0 pour valeur propre si et seulement si $\text{Ker}(v)$ contient un vecteur non nul, donc si et seulement si v n'est pas inversible. Comme l'endomorphisme u considéré est inversible, $0 \notin \text{Sp}(u)$.

2. Soit $P = \det(u - T \text{id}_E)$ le polynôme caractéristique de u . Les valeurs propres de u étant précisément les racines de P , l'observation faite à la question 1 se traduit par $P(0) \neq 0$ ⁽¹⁾.

En vertu du théorème de Cayley-Hamilton, $P(u) = 0$. Écrivant P sous la forme $P = TQ + P(0)$ avec $Q \in K[T]$, ceci conduit à l'identité

$$uQ(u) + P(0)\text{id}_E = 0,$$

soit, comme $P(0) \neq 0$,

$$u\tilde{Q}(u) = \tilde{Q}(u)u = \text{id}_E$$

en posant $\tilde{Q} = -\frac{1}{P(0)}Q$. On obtient ainsi $u^{-1} = \tilde{Q}(u)$, ce qui prouve que l'inverse de u s'exprime sous la forme d'un polynôme en u .

Exercice 5* — 1. Soit $P \in K[T]$ un polynôme premier au polynôme minimal m_u de u ; en vertu du théorème de Bézout, il existe des polynômes $R, S \in K[T]$ tels que $RP + Sm_u = 1$. Substituant u à T dans l'identité précédente, nous obtenons $R(u)P(u) + S(u)m_u(u) = \text{id}_E$, donc $R(u)P(u) = \text{id}_E$ puisque $m_u(u) = 0$, et en concluons que l'endomorphisme $P(u)$ est inversible.

2. Soit réciproquement $P \in K[T]$ un polynôme tel que l'endomorphisme $P(u)$ soit inversible et soit $\delta = \text{pgcd}(P, m_u)$. Écrivant P et m_u sous la forme $P = \delta\tilde{P}$ et $m_u = \delta\tilde{m}_u$ avec $\tilde{P}, \tilde{m}_u \in K[T]$, on observe tout d'abord que l'inversibilité de $P(u)$ implique celle de l'endomorphisme $\delta(u)$ (puisque alors $\delta(u)\tilde{P}(u)P(u)^{-1} = \tilde{P}(u)P(u)^{-1}\delta(u) = \text{id}_E$), puis que l'identité $\delta(u)\tilde{m}_u(u) = m_u(u) = 0$ et l'inversibilité de $\delta(u)$ entraînent $\tilde{m}_u(u) = 0$.

Comme m_u est le polynôme minimal de u , m_u doit alors diviser le polynôme annulateur \tilde{m}_u . D'autre part, \tilde{m}_u divise m_u par définition; on a donc $m_u = \lambda\tilde{m}_u$ avec $\lambda \in K^\times$. Le polynôme δ est ainsi une constante non nulle, ce qui signifie précisément que les polynômes P et m_u sont premiers entre eux.

Exercice 6 — Quel que soit le sous-espace vectoriel u -stable F de E , la restriction de u à F définit un endomorphisme $u|_F$ de F et $P(u|_F) = P(u)|_F$ pour tout polynôme $P \in K[T]$.

Appliquant cette observation au polynôme minimal m_u de u , nous obtenons $m_u(u|_F) = m_u(u)|_F = 0$; m_u est donc un polynôme annulateur de $u|_F$ et c'est par conséquent un multiple du polynôme minimal $m_{u|_F}$ de $u|_F$.

Si l'endomorphisme u est diagonalisable, son polynôme minimal m_u est scindé à racines simples et il en est de même pour tout polynôme divisant m_u ; en particulier, le polynôme minimal $m_{u|_F}$ de $u|_F$ est scindé à racines simples et l'endomorphisme $u|_F$ de F est donc diagonalisable.

Exercice 7* — Comme dans l'exercice précédent, $m_u(u|_F) = m_u(u)|_F = 0$ et $m_u(u|_G) = m_u(u)|_G = 0$ donc m_u est un multiple commun des polynômes minimaux m_F et m_G de $u|_F$ et $u|_G$ respectivement; de manière équivalente, $\text{ppcm}(m_F, m_G) | m_u$.

Soit $\delta = \text{pgcd}(m_F, m_G)$. En écrivant les polynômes m_F et m_G sous la forme $m_F = \delta\tilde{m}_F$ et $m_G = \delta\tilde{m}_G$ avec $\tilde{m}_F, \tilde{m}_G \in K[T]$, nous avons

$$\begin{aligned} \text{ppcm}(m_F, m_G) &= \delta\tilde{m}_F\tilde{m}_G \\ &= \tilde{m}_G m_F \\ &= \tilde{m}_F m_G \end{aligned}$$

et il est maintenant facile de voir que ce polynôme annule l'endomorphisme u .

Posons $v = \text{ppcm}(m_F, m_G)$ et considérons un vecteur x dans E , que l'on écrit sous la forme $x = x_F + x_G$ avec $x_F \in F$ et $x_G \in G$. Comme $v = \tilde{m}_G(u)m_F(u)$,

$$v(x_F) = \tilde{m}_G(u)(m_F(u)(x_F)) = \tilde{m}_G(u)(m_F(u|_F)(x_F)) = 0;$$

1. Il est connu *a priori* que $P(0) = \det(u)$, de sorte que l'inversibilité de u équivaut à la non-nullité du terme constant $P(0)$ de son polynôme caractéristique.

en utilisant l'identité $v = \tilde{m}_F(u)m_G(u)$, on prouve de même que $v(x_G) = 0$. Nous obtenons ainsi que le polynôme $\text{ppcm}(m_F, m_G)$ annule u , donc qu'il est divisible par m_u , et la conclusion $m_u = \text{ppcm}(m_F, m_G)$ découle des relations de divisibilité $m_u | \text{ppcm}(m_F, m_G)$ et $\text{ppcm}(m_F, m_G) | m_u$ une fois que l'on a observé que tous les polynômes considérés sont unitaires.

Exercice 8 — Le polynôme $T^3 - T$ est scindé à racines simples : $T^3 - T = T(T-1)(T+1)$. Étant donnée une matrice $A \in \mathcal{M}_n(\mathbb{R})$ telle que $A^3 = A$, son polynôme minimal m_A divise $T^3 - T$ et est donc scindé à racines simples ; les valeurs propres de A appartiennent à l'ensemble $\{-1, 0, 1\}$ et il existe une matrice $P \in \mathcal{M}_n(\mathbb{R})$ inversible telle que $P^{-1}AP$ soit une matrice diagonale ayant a coefficients égaux à 0, b coefficients égaux à 1 et c coefficients égaux à -1 , a, b, c étant des entiers naturels soumis à la seule condition : $a + b + c = n$. Réciproquement, le polynôme minimal d'une matrice diagonale D de cette forme est $T^{\varepsilon_a}(T-1)^{\varepsilon_b}(T+1)^{\varepsilon_c}$ avec, pour tout $x \in \{a, b, c\}$, $\varepsilon_x = 0$ si $x = 0$ et $\varepsilon_x = 1$ si $x \geq 1$, et la matrice PDP^{-1} est une solution de l'équation considérée quelle que soit la matrice inversible $P \in \mathcal{M}_n(\mathbb{R})$. Nous avons ainsi décrit toutes les solutions de l'équation $X^3 - X$ dans $\mathcal{M}_n(\mathbb{R})$.

Exercice 9 — 1. L'identité $(T^2 + 1) - T^2 = 1$ est une relation de Bézout entre les polynômes premiers entre eux T et $T^2 + 1$. En substituant A à T , nous obtenons la relation

$$I_3 = (A^2 + I_3) - A^2$$

dans $\mathcal{M}_3(\mathbb{R})$, laquelle implique immédiatement la décomposition $\mathbb{R}^3 = \text{Ker}(A) \oplus \text{Ker}(A^2 + I_3)$. En effet, quel que soit le vecteur $X \in \mathbb{R}^3$,

$$X = (A^2 + I_3)X - A^2X$$

et $(A^2 + I_3)X \in \text{Ker}(A)$, $A^2X \in \text{Ker}(A^2 + I_3)$ puisque $A^3 + A = 0$; si d'autre part X appartient simultanément à $\text{Ker}(A)$ et $\text{Ker}(A^2 + I_3)$, alors $X = (A^2 + I_3)X - A^2X = 0$.

2. Le polynôme minimal m_A de A est un diviseur de $T^3 + T = T(T^2 + 1)$, soit T , $T^2 + 1$ ou $T(T^2 + 1)$. Le cas $m_A = T$ est exclu puisque $A \neq 0$. Le cas $m_A = T^2 + 1$ est également exclu : en effet, si $m_A = T^2 + 1$, alors le polynôme caractéristique P_A de A est une puissance de $(T^2 + 1)$ car les polynômes m_A et P_A ont les mêmes facteurs irréductibles et ceci est impossible car P_A est de degré 3. Nous obtenons donc $m_A = T(T^2 + 1)$.

3. Soit x un vecteur de \mathbb{R}^3 n'appartenant pas à $\text{Ker}(A)$; nous avons donc par hypothèse $x \neq 0$ et $Ax \neq 0$. S'il existait deux nombres réels $\lambda, \mu \in \mathbb{R}$ non tous les deux nuls et tels que $\lambda x + \mu Ax = 0$, alors λ et μ seraient tous les deux non-nuls puisque $x, Ax \neq 0$; on aurait alors $Ax = -\frac{\lambda}{\mu}x$, ce qui signifierait, comme $x \neq 0$, que $-\frac{\lambda}{\mu}$ serait une valeur propre de A , c'est-à-dire une racine réelle du polynôme $m_A = T^3 + T$, et ceci contredirait la non-nullité de λ .

La famille (x, Ax) est donc libre.

4. Comme $A \neq 0$, $\text{Ker}(A) \neq \mathbb{R}^3$ et il existe un vecteur x non-nul dans le sous-espace vectoriel $\text{Ker}(A^2 + I_3)$ en vertu de la question 1. Ce sous-espace étant en outre stable par A puisque $(A^2 + I_3)A = A^3 + A = 0$, il contient la famille (x, Ax) et est donc de dimension au moins 2 d'après la question précédente. Cette discussion conduit à la majoration $\dim \text{Ker}(A) = 3 - \dim \text{Ker}(A^2 + I_3) \leq 1$. D'un autre côté, le polynôme minimal de A admettant 0 comme racine, 0 est une valeur propre de A et donc $\dim \text{Ker}(A) \geq 1$. Nous avons ainsi prouvé que le noyau de A est de dimension 1.

Soit e un vecteur non-nul dans $\text{Ker}(A)$ et soit x un vecteur non-nul dans $\text{Ker}(A^2 + I_3)$. La famille (x, Ax) étant libre en vertu de la question 3, elle constitue une base du sous-espace bidimensionnel $\text{Ker}(A^2 + I_3)$ et (e, x, Ax) est alors une base de \mathbb{R}^3 en vertu de la question 1. Désignant par P la matrice de passage de la base canonique de \mathbb{R}^3 à la base (e, x, Ax) , $P^{-1}AP$ est la matrice dont les colonnes sont les coordonnées des vecteurs Ae, Ax et AAx dans la base (e, x, Ax) ; il est manifeste que les deux premières colonnes sont celles que l'on souhaite, et il en est de même pour la dernière puisque, x appartenant à $\text{Ker}(A^2 + I_3)$, $A^2x = -x$.

Exercice 10* — 1. Il suffit de calculer $\det(A - XI_n)$. En développant ce déterminant par rapport à la première colonne, on obtient

$$\det(A - XI_n) = -X \det(A_1 - XI_{n-1}) - (-1)^n a_0,$$

où A_1 est la matrice compagnon du polynôme $X^{n-1} - a_{n-1}X^{n-2} - \dots - a_1$. À partir de cette identité et du cas évident $n = 1$, on prouve par récurrence sur n que $P_u = P_A$ n'est autre que le polynôme $(-1)^n P$.

2. Soient E_1, \dots, E_n les vecteurs de la base canonique de K^n . Par définition de A , $AE_1 = E_2$, $AE_2 = E_3, \dots, AE_{n-1} = E_n$ et $AE_n = a_0E_1 + \dots + a_{n-1}E_n$, donc $A^p E_1 = E_{p+1}$ pour tout entier $p \in \{1, \dots, n-1\}$.

Nous allons vérifier que la matrice $P(A)$ est nulle. Cela revient à prouver $P(A)E_i = 0$ pour tout i et, comme $P(A)E_i = P(A)A^{i-1}E_1 = A^{i-1}P(A)E_1$, il suffit de s'assurer que $P(A)E_1 = 0$. Le calcul est facile :

$$\begin{aligned} P(A)E_1 &= A^n E_1 - \sum_{i=0}^{n-1} a_i A^i E_1 \\ &= AA^{n-1} E_1 - \sum_{i=0}^{n-1} a_i E_{i+1} \\ &= AE_n - \sum_{i=0}^{n-1} a_i E_{i+1} \\ &= 0. \end{aligned}$$

Nous avons donc bien $P(A) = 0$, d'où $P_u(u) = 0$ d'après la question précédente.

Le polynôme unitaire P annule u . D'autre part, si $\lambda_0, \dots, \lambda_{n-1}$ sont des nombres réels tels que le polynôme $\lambda_{n-1}X^{n-1} + \dots + \lambda_1 X \lambda_0$ annule u , alors

$$0 = \left(\sum_{i=0}^{n-1} \lambda_i A^i \right) E_1 = \sum_{i=0}^{n-1} \lambda_i E_{i+1}$$

et donc $\lambda_0 = \lambda_1 = \dots = \lambda_{n-1} = 0$, ce qui montre qu'il n'existe aucune polynôme non-nul de degré inférieur ou égal à $n-1$ annihilant u . Le polynôme minimal m_u de u est ainsi de degré au moins n et, puisqu'il divise P , $m_u = P$ car ces deux polynômes sont unitaires.

4. Pour établir que le sous-espace $E_x = \text{Vect}(x, v(x), \dots, v^p(x))$ est stable par v , il suffit de vérifier qu'il contient les images des vecteurs $x, v(x), \dots, v^p(x)$ par v et seul le cas de $v^p(x)$ n'est pas complètement évident. Par définition de p , la famille $(x, v(x), \dots, v^p(x), v^{p+1}(x))$ n'est pas libre et il existe donc des scalaires $\lambda_0, \dots, \lambda_{p+1}$ non tous nuls tels que

$$\lambda_0 x + \lambda_1 v(x) + \dots + \lambda_{p+1} v^{p+1}(x) = 0.$$

La famille $(x, v(x), \dots, v^p(x))$ étant libre par hypothèse, le coefficient λ_{p+1} n'est pas nul et on peut dès lors écrire $v^{p+1}(x)$ sous la forme

$$v^{p+1}(x) = -\frac{1}{\lambda_{p+1}} (\lambda_0 x + \dots + \lambda_p v^p(x)),$$

ce qui prouve que l'image de $v^p(x)$ par v est contenue dans E_x .

5 & 6. En utilisant les notations de la question précédente, la matrice de v dans la base \mathcal{B}_x de E_x est la matrice compagnon du polynôme

$$X^p + \frac{\lambda_p}{\lambda_{p+1}} X^{p-1} + \dots + \frac{\lambda_0}{\lambda_{p+1}}.$$

7 & 8. Soit P_v le polynôme caractéristique de v et soit x un vecteur non nul dans E . Comme le sous-espace E_x de E est stable par v , P_v s'écrit sous la forme QP_x où P_x est le polynôme caractéristique de la restriction de v à E_x et $Q \in K[X]$. En vertu de ce qui précède, P_x annule la restriction de v à E_x et donc, en particulier, $P_x(v)(x) = 0$. Comme $P(v)(x) = Q(v)(P_x(v)) = 0$, nous en déduisons que l'endomorphisme $P_v(v)$ est nul et avons ainsi redémontré le théorème de Cayley-Hamilton.

Exercice 11* — 1. On justifie comme dans la question 5 de l'exercice précédent qu'il existe des scalaires a_0, a_1, \dots, a_{n-1} dans K tels que

$$u^n(x) = a_0 x + a_1 u(x) + \dots + a_{n-1} u^{n-1}(x)$$

et on vérifie que la matrice de u dans la base \mathcal{B}_x est la matrice compagnon du polynôme $T^n - (a_{n-1}T^{n-1} + \dots + a_1 T + a_0)$.

2. L'unicité de la matrice compagnon d'un endomorphisme cyclique est une conséquence immédiate de la question 1 de l'exercice précédent : les coefficients de la dernière colonne d'une telle matrice sont en effet, au signe près, ceux du polynôme caractéristique de u et ne dépendent donc pas du choix du vecteur cyclique considéré.

3. Tout endomorphisme d'un espace vectoriel de dimension n est diagonalisable s'il possède n valeurs propres distinctes.

La réciproque est vraie dans le cas d'un endomorphisme cyclique. En effet, si u est un endomorphisme cyclique, son polynôme minimal et son polynôme caractéristique coïncident en vertu de la question 3 de l'exercice précédent ; si u est diagonalisable, cela implique donc que son polynôme caractéristique est scindé à racines simples, ce qui signifie précisément que u possède n valeurs propres distinctes.
