SUR LE POLYNÔME MINIMAL

Paul BOREL [1]

1 Introduction.

Dans ce qui suit, nous ne parlerons que de matrices à coefficients dans un corps commutatif K (ou un anneau commutatif R). La traduction en termes d'endomorphismes d'espaces vectoriels de dimension finie sur K (voire de R-modules libres de type fini) est immédiate. Soit $A \in \mathbf{M}_n(K)$ une matrice $n \times n$ à coefficients dans un corps K ; à la matrice A sont associés deux polynômes remarquables (entre autres) de $K[X]$, à savoir :

- le polynôme caractéristique $P_A(X) = \text{dét}\,(XI_n - A)$;
- le polynôme minimal $m_A(X)$.

Rappelons la définition de $m_A(X)$: nous avons un homomorphisme de K-algèbres $K[X] \longrightarrow M_n(K)$ associé à A et défini par $\varphi(f) = a_m X^m + \cdots + a_1 X + a_0 \mathbf{1}_n$. Comme $dim_K K[X] = +\infty$ et que $dim_K M_n(K) = n^2$, φ n'est certainement pas injectif ; donc $Ker\varphi$ est un idéal $\neq \{0\}$ de $K[X]$. Puisque $K[X]$ est un anneau principal, il existe un polynôme unique unitaire qui engendre $Ker\varphi$: c'est le polynôme minimal $m_A(X)$; $m_A(X)$ est donc le polynôme unitaire de plus bas degré qui s'annule pour la matrice A : $m_A(A) = 0$. Si $f(X) \in K[X]$ est tel que $f(A) = 0$, $f(X)$ est un multiple de $m_A(X)$. Par définition même du polynôme caractéristique $P_A(X)$, nous disposons d'un algorithme pour son calcul. Nous allons voir que nous avons aussi un algorithme pour le calcul de $m_A(X)$.

2 Calcul du polynôme minimal

Soit R un anneau commutatif et $M \in M_n(R)$ une matrice $n \times n$ à coefficients dans R. Nous pouvons calculer $\widetilde{M}$ = (ajointe de M)= (transposée de la matrice des cofacteurs de M) et nous avons alors $M\widetilde{M} = \widetilde{M}M = \text{dét}\ M$. Nous appliquons cela au cas où $R = K[X]$ et $M = X\mathbf{1}_n - A$; nous avons alors :

$$(X\mathbf{1}_n - A) \cdot \widetilde{(X\mathbf{1}_n - A)} = \widetilde{(X\mathbf{1}_n - A)} \cdot (X\mathbf{1}_n - A) = P_A(X).\mathbf{1}_n.$$

Remarquons que cette égalité peut être considérée comme un égalité dans $M_n(K[X])$ et comme une égalité dans $M_n(K)[X]$ (dans ces deux anneaux non commutatifs, $\mathbf{1}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ est l'élément unité et $\mathbf{1}_n D = 0$ implique $D = 0$; c'est une façon de prouver le théorème de Hamilton-Cayley).

Soit maintenant $h(X)$ le PGCD (calculé dans $K[X]$ des éléments de $\widetilde{(X\mathbf{1}_n - A)}$, c'est-à-dire le PGCD des mineurs $(n-1)\times(n-1)$ de $(X\mathbf{1}_n - A)$. Clairement $h(X)$ est

un polynôme unitaire de $K[X]$ qui divise $P_A(X)$ (développement d'un déterminant suivant une ligne ou une colonne) de sorte que nous pouvons écrire
$P_A(X) = h(X)g(X)$où $g(X)$ est un certain polynôme unitaire de $K[X]$.

Proposition. 1 : $g(X) = m_A(X)$ =*(polynôme minimal de A).*

Démonstration : Écrivons $\widetilde{(X\mathbf{1}_n - A)} = h(X)B(X)$ où $B(X) \in M_n(K[X])$ est une matrice dont les coefficients sont premiers entre eux. Nous avons alors :

$$h(X)g(X).\mathbf{1}_n = h(X)B(X)(X\mathbf{1}_n - A).$$

Comparant les coefficients des deux membres et puisque $K[X]$ est intègre, nous en déduisons : $g(X).\mathbf{1}_n = B(X)(X\mathbf{1}_n - A)$ (relation que nous pouvons considérer dans$M_n(K[X])$ tout comme dans $M_n(K)[X]$). Alors $g(A).\mathbf{1}_n = B(A)(A - A) = 0$ et donc $g(A) = 0$. Par conséquent $m_A(X)$ divise $g(X)$.

Pour démontrer la réciproque nous considérons le polynôme $m_A(X) - m_A(Y)$ de $K[X, Y]$. Nous pouvons écrire dans $K[X, Y]$, $m_A(X) - m_A(Y) = (X - Y)R(X, Y)$; en substituant $X\mathbf{1}_n$ à X et A à Y nous obtenons :
$m_A(X\mathbf{1}_n) - m_A(A) = (X\mathbf{1}_n - A)R(X\mathbf{1}_n, A)$; compte tenu de $m_A(A) = 0$ et puisque $m_A(X\mathbf{1}_n) = m_A(X).\mathbf{1}_n$, nous avons $m_A(X).\mathbf{1}_n = (X\mathbf{1}_n - A)R(X\mathbf{1}_n, A)$; en multipliant à gauche par $\widetilde{(X\mathbf{1}_n - A)}$ nous obtenons :

$$\widetilde{(X\mathbf{1}_n - A)}m_A(X).\mathbf{1}_n = m_A(X)\widetilde{(X\mathbf{1}_n - A)} =$$
$$= P_A(X).\mathbf{1}_n R(X\mathbf{1}_n, A) = P_A(X)R(X\mathbf{1}_n, A) = h(X)g(X)R(X\mathbf{1}_n, A) =$$
$$= h(X)B(X)m_A(X) = h(X)m_A(X)B(X)$$

Comme précédemment, nous pouvons simplifier par $h(X)$, d'où :
$g(X)R(X\mathbf{1}_n, A) = m_A(X)B(X)$. Donc $g(X)$ divise tous les coefficients de la matrice $m_A(X)B(X)$; mais les coefficients de $B(X)$ sont premiers entre eux.

Lemme. 2 : *Soit R un anneau principal; $b_1, \cdots, b_m$des éléments de R premiers entre eux, $a \in R$ et $x \in R$. Si x divise tous les ab_i alors x divise a.*

Démonstration : $ab_i = xc_i (1 \leq i \leq m)$. Comme $PGCD(b_1, \cdots, b_m) = 1$, il existe des r_i tels que $r_1b_1 + \cdots + r_mb_m = 1$. Nous avons alors $ar_ib_i = xr_ic_i$, donc :
$a = a(r_1b_1 + \cdots + r_mb_m) = x(r_1c_1 + \cdots + r_mc_m)$ et par conséquent x divise a.
Revenons à la proposition : il résulte du lemme que $g(X)$ divise $m_A(X)$; puisque $g(X)$ est unitaire on a $g(X) = m_A(X)$. Nous avons a que $m_A(X)$ divise $P_A(X)$; il y a une réciproque.

Proposition. 3 : $P_A(X)$ *divise* $m_A(X)^n$.

Démonstration : Soit $\overline{K}$ une clôture algébrique de K. Les racines de $P_A(X)$ sont les valeurs propres de A. Si λ est une valeur propre de A, on a $m_A(\lambda) = 0$; soit $\overrightarrow{V} \neq \overrightarrow{0}$ un vecteur propre associé (dans $\overline{K}^n$) de sorte que $A\overrightarrow{V} = \lambda\overrightarrow{V}$ et donc $A^k\overrightarrow{V} = \lambda^k\overrightarrow{V}$. Si $m_A(X) = X^r + a_{r-1}X^{r-1} + \cdots + a_1X + a_0$ on a
$m_A(A) = A^r + a_{r-1}A^{r-1} + \cdots + a_1A + a_0\mathbf{1}_n = 0$ donc
$m_A(A).\overrightarrow{V} = A^r\overrightarrow{V} + \cdots + a_0\overrightarrow{V} = \lambda^r\overrightarrow{V} + a_{r-1}\lambda^{r-1}\overrightarrow{V} + \cdots + a_1\lambda\overrightarrow{V} + a_0\overrightarrow{V} = m_A(\lambda)\overrightarrow{V} =$

0. Puisque $\overrightarrow{V} \neq \overrightarrow{0}$, cela implique que $m_A(\lambda) = 0$. Dans $\overline{K}[X]$ nous pouvons écrire $P_A(X) = \prod_{i=1}^{m} (X - \lambda_i)^{\alpha_i}$ $(\sum_{i=1}^{m} \alpha_i = n)$ et $m_A(X) = \prod_{i=1}^{m} (X - \lambda_i)^{\beta_i}$ car $m_A(X) \mid P_A(X)$. On vient de voir que $1 \leq \beta_i$ donc $1 \leq \beta_i \leq \alpha_i \leq n$, d'où $\alpha_i \leq n\beta_i$; $m_A(X)^n = \prod_{i=1}^{m} (X - \lambda_i)^{n\beta_i}$. Donc $P_A(X) \mid m_A(X)^n$.

Remarques :

- Si $A = \lambda \mathbf{1}_n$ alors on a $P_A(X) = m_A(X)^n$ puisque $P_A(X) = (X - \lambda)^n$ et $m_A(X) = X - \lambda$.

- La division qui a priori à lieu dans $\overline{K}[X]$, a en fait déjà lieu dans $K[X]$.

3 Remarques concernant les matrices diagonalisables

Une matrice $A \in M_n(K)$ est dite diagonalisable si il existe un corps L tel que $K \subseteq L \subseteq \overline{K}$(clôture algébrique de K) et dans $M_n(L)$ une matrice inversible S vérifiant la condition : $S^{-1}AS$ est une matrice diagonale $D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$

Alors les λ_i sont exactement les valeurs propres de A comptées avec leurs multiplicités. On a alors la caractérisation : A est diagonalisable si et seulement si toutes les racines du polynôme minimal $m_A(X)$ sont simples (c'est-à-dire sans facteur carré). Comme on a un algorithme pour calculer $m_A(X)$ dans $K[X]$, nous pouvons calculer $m'_A(X)$ =la dérivée du polynôme minimal et, par l'algorithme de la division, le PGCD $D_A(X)$ de $m_A(X)$ et $m'_A(X)$: alors A est diagonalisable si et seulement si $D_A(X) \neq 1$.

On a donc un algorithme qui permet de décider si une matrice est diagonalisable sans calculer aucune valeur propre.

Dans $M_n(K)$, deux matrices A et B sont dites semblables s'il existe une matrice inversible S telle que $B = S^{-1}AS$.

Proposition. 4 *: Si A et B sont semblables, alors $m_A(X) = m_B(X)$.*

Démonstration : Si $B = S^{-1}AS$ (alors $A = SBS^{-1}$) on a :
$m_A(B) = S^{-1}m_A(A)S = S^{-1}0S = 0$ donc $m_A(X) \mid m_B(X)$; de la même façon $m_B(X) \mid m_A(X)$. Donc $m_A(X) = m_B(X)$.

Corollaire. 5 *: Supposons A diagonalisable. Alors B est semblable à A si et seulement si*
$m_A(X) = m_B(X)$ en supposant que toutes les valeurs propres sont dans le corps de base.

Démonstration : On vient de voir qu'en toute généralité, si A et B sont semblables, $m_A(X) = m_B(X)$. Réciproquement, si $m_B(X) = m_A(X)$ comme A est diagonalisable $m_A(X)$ est sans facteur carré donc aussi $m_B(X)$, donc B est diagonalisable. Soit $L = K(\lambda_1, \cdots, \lambda_n)$ le corps engendré sur K par les valeurs propres de A et B. Dans une base $\{\overrightarrow{v_1}, \cdots, \overrightarrow{v_n}\}$ A devient $D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$ i.e.

$D = S^{-1}AS$, $S \in GL(n, L)$. De même $D = T^{-1}BT$, $T \in GL(n, L)$. Donc $T^{-1}BT = S^{-1}AS$ et $B = (ST^{-1})^{-1}AST^{-1}$ donc A et B sont semblables, la similitude ayant lieu dans L (corps contenant les valeurs propres).

Remarques :

- En général la situation est plus compliquée.

Par exemple si $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$,

on a $P_A(X) = P_B(X) = (X-1)^4$, $m_A(X) = m_B(X) = (X-1)^2$ mais A et B ne sont pas semblables (A possède trois vecteurs propres linéairement indépendants tandis que B n'en possède que deux).

- Prenons pour corps de base $\mathbb{R}$ ou $\mathbb{C}$. L'application $M_n(K) \longrightarrow K[X]$: $A \longmapsto P_A(X)$ est continue puisque polynomiale. Pour le polynôme minimal ça n'est pas vrai : voici un exemple simple. Dans $M_2(\mathbb{C})$ on considère la matrice $A_\alpha = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$. Alors $\begin{cases} (m_{A_\alpha}(X) = (X-\alpha)(X-1) = X^2 - (\alpha+1)X + \alpha \text{ si } \alpha \neq 1 \\ m_{A_1}(X) = X - 1 \end{cases}$

donc $\lim\limits_{\alpha \longrightarrow 1} m_{A_\alpha}(X) \neq m_{A_1}(X)$.

4 Questions de rationalité

Dans ce dernier paragraphe nous allons examiner brièvement ce qui subsiste lorsque le corps K est remplacé par un anneau. Soit donc R un anneau commutatif et $A \in M_n(R)$ une matrice carrée $n \times n$ à coefficients dans R. Nous pouvons bien sûr considérer $P_A(X) = \text{dét}\,(X\mathbf{1}_n - A)$ et sans aucune autre hypothèse, le théorème de Hamilton-Cayley reste vrai (cf par exemple [1] p.441).

La formule $(X\mathbf{1}_n - A).\widetilde{(X\mathbf{1}_n - A)} = \widetilde{(X\mathbf{1}_n - A)}.(X\mathbf{1}_n - A) = P_A(X).\mathbf{1}_n$ est encore valable ; mais la détermination de $h(X) = [\text{PGCD des coefficients de l'adjointe } \widetilde{(X\mathbf{1}_n - A)}\ ]$ ainsi que la formule $P_A(X) = h(X)g(X)$ posent des problèmes en général.

D'autre part nous pouvons encore considérer l'homomorphisme de R-algèbres $R[X] \xrightarrow{\varphi} M_n(R) : \varphi(f(X)) = f(A)$; posons alors $I_A = \text{Ker}\varphi$. En général $R[X]$ n'est pas un anneau principal de sorte qu'on ne peut rien dire a priori des générateurs de I_A .

La première hypothèse raisonnable à faire est de supposer R intègre ; désormais R est un anneau intègre et posons $K = \text{Frac}(R) = $ (corps des fractions de R) . Nous pouvons alors calculer $h(X)$ dans $K[X]$ et écrire $P_A(X) = h(X)g(X) \text{dans} K[X]$: le problème est de savoir à quelles conditions cette écriture valable dans $K[X]$ l'est en fait dans $R[X]$.

Soit $R \subset S$ des anneaux commutatifs ; un élément $x \in S$ est dit entier sur R s'il vérifie une équation du type $x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0 = 0$ où les $a_i \in R$ (équation de dépendance intégrale pour x sur R).

Un anneau commutatif intègre R est dit intégralement clos si tout élément du corps des fractions de R, entier sur R, est en fait dans R. Voici des exemples d'an-

neaux intégralement clos : $\mathbb{Z}$; tout anneau principal ; tout anneau factoriel ; tout anneau de valuation ; tout anneau de Dedekind (en particulier l'anneau des entiers d'un corps de nombres ou l'anneau des fonctions régulières sur une courbe non singulière) etc. (voir [1] pp 623-624 exercices 41-50).

Théorème. 6 *Soit R un anneau intégralement clos, $K = Frac(R)$ le corps des fractions de R et $A \in M_n(R)$ une matrice à coefficients dans R. Alors :*

i. le polynôme minimal $m_A(X)$ calculé dans $K[X]$ est dans $R[X]$;

ii. l'écriture $P_A(X) = h(X)m_A(X)$ a lieu dans $R[X]$;

iii. l'idéal I_A défini ci-dessus est principal et engendré par $m_A(X)$.

Démonstration : Soit $\overline{K}$ une clôture algébrique de K. Dans $\overline{K}[X]$ nous pouvons écrire $P_A(X) = \prod_{i=1}^{m} (X - \lambda_i)^{a_i}$ et $m_A(X) = \prod_{i=1}^{m} (X - \lambda_i)^{b_i}$ avec des b_i tels que $1 \leq b_i \leq a_i$; puisque $P_A(X)$ est unitaire les λ_i sont entiers sur R.

Cela montre que les coefficients du polynôme minimal $m_A(X)$ sont entiers sur R (le produit et la somme d'entiers sont encore des entiers). Mais $m_A(X) \in K[X]$; puisque R est intégralement clos, nous avons en fait $m_A(X) \in R[X]$.

Puisque $m_A(X)$ est unitaire, nous pouvons diviser $P_A(X)$ par $m_A(X)$ et écrire $P_A(X) = m_A(X)q(X) + r(X)$ avec deg $r <$ deg m_A. Comme $P_A(A) = m_A(A) = 0$, nous avons $r(A) = 0$ donc $r(X) = 0$ (dans $K[X]$ donc aussi dans $R[X]$). Par suite l'écriture $P_A(X) = h(X)m_A(X)$ a lieu dans $R[X]$.

Clairement $m_A(X) \in I_A$. Si $f(X) \in I_A$, comme $m_A(X)$ est unitaire, nous pouvons diviser et comme ci-dessus on voit que $f(X) = m_A(X)q(X)$ donc $m_A(X)$ engendre I_A .

On vient de voir que dans le cas des anneaux intégralement clos une belle partie de la théorie du polynôme minimal reste vraie.

Référence : [1] R. GODEMENT *Cours d'algèbre*