L'exponentielle de matrice comme polynôme

Jean-François BURNOL, février 2017

Soit $V = \mathbf{C}^n$ muni de sa base canonique et M une matrice $n \times n$ à coefficients complexes et ϕ l'endomorphisme associé agissant sur V.

Cette fiche va faire un usage intensif de la notion de polynôme d'endomorphisme et de matrice.

Soit $T \in \mathbf{C}[X]$ le polynôme minimal unitaire de M, de degré t, soit λ_i, $1 \leqslant i \leqslant p$ ses racines, qui sont les valeurs propres de M, et $m_1, \dots, m_p$ leurs multiplicités dans T.

On a nécessairement $t \geqslant 1\dots$ sauf si $n = 0$, et on suppose donc par la suite que V n'est pas l'espace vectoriel nul !

Théorème 1. *Un polynôme* L *vérifie*
$$e^{M} = L(M)$$
si et seulement si il satisfait aux conditions d'interpolation
$$1 \leqslant i \leqslant p \implies L^{(k)}(\lambda_i) = e^{\lambda_i}, 0 \leqslant k < m_i$$
Il existe un unique polynôme L_0 *de degré au plus* $t-1$ *vérifiant ces conditions.*

Les autres polynômes satisfaisant $e^M = L(M)$ *sont ceux qui sont congrus à* L_0 *modulo le polynôme minimal* T *de la matrice* M.

Je vais commencer par une preuve intelligente qui montre comme on découvre ce genre de théorème. Puis je ferai la preuve idiote plus courte et qu'il faudra connaître à l'avenir pour réussir les oraux de concours de bachotage.

Preuve. Je laisse en exercice le fait que le problème d'interpolation a bien une (unique) solution de degré strictement inférieur à t et que les autres polynômes interpolateurs lui sont congrus modulo T.

Il existe une décomposition en éléments simples :
$$\frac{1}{\prod_i (X-\lambda_i)^{m_i}} = \sum_i \frac{U_i}{(X-\lambda_i)^{m_i}}$$
avec les U_i des polynômes de degrés au plus les $m_i - 1$. En fait ils sont pour cette fraction rationnelle de degrés exactement égaux aux $m_i - 1$ (exercice).

Posons $B_i = L/(X-\lambda_i)^{m_i}$, on obtient l'identité polynomiale :
$$1 = \sum_i U_i B_i$$
Posons $P_i = U_i B_i$ (son degré est exactement $t-1$) et considérons maintenant les polynômes d'endormorphismes $P_i(\phi)$.

Comme il est bien connu, V est la somme directe des espaces caractéristiques V_i associés aux valeurs propres λ_i. L'endomorphisme $(\phi - \lambda_1 \mathrm{Id})^{m_1}$ est identiquement nul sur V_1, puisque

$T(\phi) = 0$ or les facteurs de T autres que $(\phi - \lambda_1 \mathrm{Id})^{m_1}$ créent sur V_1 des endomorphismes injectifs, puisque $\phi - \lambda_j \mathrm{Id}$ est injectif sur V_1, puisque $\lambda_j \neq \lambda_1$ et que λ_1 est la seule valeur propre sur V_1. Ainsi sur V_1 les $P_i(\phi)$, $i > 1$ sont identiquement nuls, car $(X - \lambda_1)^{m_1}$ divise P_i pour $i > 1$.

Bref, il en resulte que sur V_1, alors $P_1(\phi)$ agit comme l'identité. Et il agit comme l'endormorphisme nul sur les autres espaces caractéristiques donc $P_1(\phi)$ n'est autre que la projection sur V_1 parallèlement aux autres et est un idempotent ($(P_1(\phi))^2 = P_1(\phi)$), et de même pour les $P_i(\phi)$. Par ailleurs $P_i(\phi)P_j(\phi) = 0$ pour $i \neq j$ car $P_i P_j$ est un multiple de T pour $i \neq j$.

On a la décomposition

$$\mathrm{Id} = \sum_i P_i(\phi)$$

et donc aussi

$$\phi = \sum_i \phi P_i(\phi)$$

puis en revenant à la matrice M

$$M = \sum_i M P_i(M)$$

Toutes ces matrices commutent donc

$$e^M = \prod_i \exp(M P_i(M))$$

$$e^M = \prod_i \exp\big((M - \lambda_i) P_i(M)\big) \exp(\lambda_i P_i(M))$$

Bon je reviens aux endomorphismes et je note que $(\phi - \lambda_i \mathrm{Id}) P_i(\phi)$ est nilpotent, et, en rappelant que $P_i(\phi)$ est un idempotent, que son exponentielle vaut

$$e^{(\phi - \lambda_i \mathrm{Id}) P_i(\phi)} = \mathrm{Id} + \sum_{1 \leqslant k < m_i} \frac{(\phi - \lambda_i \mathrm{Id})^k}{k!} P_i(\phi)$$

Par ailleurs $\exp(\lambda_i P_i(\phi)) = \mathrm{Id} + (e^{\lambda_i} - 1) P_i(\phi)$ donc $P_i(\phi) \exp(\lambda_i P_i(\phi)) = e^{\lambda_i} P_i(\phi)$, puis

$$\begin{aligned}
\exp\big((\phi - \lambda_i \mathrm{Id}) P_i(\phi)\big) \cdot \exp(\lambda_i P_i(\phi)) &= \exp(\lambda_i P_i(\phi)) + \sum_{1 \leqslant k < m_i} \frac{(\phi - \lambda_i \mathrm{Id})^k}{k!} e^{\lambda_i} P_i(\phi) \\
&= \mathrm{Id} + \left(-\mathrm{Id} + e^{\lambda_i} \sum_{0 \leqslant k < m_i} \frac{(\phi - \lambda_i \mathrm{Id})^k}{k!} \right) P_i(\phi)
\end{aligned}$$

Il s'agit maintenant de faire le produit sur tous les i, mais les termes croisés sont tous nuls par $P_i(\phi) P_j(\phi) = 0$, donc on obtient simplement :

$$e^{\phi} = \mathrm{Id} + \sum_i \left(-\mathrm{Id} + e^{\lambda_i} \sum_{0 \leqslant k < m_i} \frac{(\phi - \lambda_i \mathrm{Id})^k}{k!} \right) P_i(\phi)$$

À ce stade il est naturel de définir le polynôme [1]

$$L = 1 + \sum_i \left(-1 + e^{\lambda_i} \sum_{0 \leqslant k < m_i} \frac{(X - \lambda_i)^k}{k!} \right) P_i$$

1. comme $\sum_i P_i = 1$, il vaut aussi $\sum_i \left(\sum_{0 \leqslant k < m_i} e^{\lambda_i} \frac{(X - \lambda_i)^k}{k!} \right) P_i$ qui se comprend peut-être plus facilement.

et de rappeler que $\mathrm{P}_j = \mathrm{U}_j \frac{\mathrm{L}}{(\mathrm{X}-\lambda_j)^{m_j}}$ s'annule avec au moins multiplicité m_i en λ_i pour tout $j \neq i$, et que $\mathrm{P}_i = 1 - \sum_{j\neq i} \mathrm{P}_j = 1 + \mathscr{O}((\mathrm{X}-\lambda_i)^{m_i})$. Donc pour calculer L et ses dérivées jusqu'à l'ordre $m_i - 1$ en λ_i, on est ramené à étudier la fonction polynomiale

$$f_i(z) = 1 + \left(-1 + e^{\lambda_i} \sum_{0\leqslant k<m_i} \frac{(z-\lambda_i)^k}{k!}\right) = \sum_{0\leqslant k<m_i} \frac{(z-\lambda_i)^k}{k!} e^{\lambda_i}$$

qui n'est autre que le polynôme de Taylor à l'ordre m_i (ou $m_i - 1$ suivant la convention) de la fonction exponentielle, au point λ_i. Ainsi $f_i^{(k)}(\lambda_i) = e^{\lambda_i}$ pour $0 \leqslant k < m_i$ et par conséquent $\mathrm{L}^{(k)}(\lambda_i) = e^{\lambda_i}$, pour $0 \leqslant k < m_i$.

Donc ce L est à la fois solution du problème d'interpolation et tel que $e^{\mathrm{M}} = \mathrm{L}(\mathrm{M})$, mais il est peut-être de degré trop grand. Qu'à cela ne tienne, on le remplace par son reste L_0 dans la division euclidienne par le polynôme minimal T. Alors $\mathrm{L}(\mathrm{M}) = \mathrm{L}_0(\mathrm{M}) = e^{\mathrm{M}}$ et L_0 vérifie aussi les conditions d'interpolation.

Le résultat désiré a été obtenu. □

Voici maintenant la preuve idiote car pour la trouver il faut déjà connaître l'énoncé ... alors que l'autre preuve *construit* l'énoncé.

Seconde preuve. Soit L le polynôme d'interpolation défini dans l'énoncé du théorème. Il existe une série entière de rayon de convergence infini avec $e^z - \mathrm{L}(z) = \mathrm{T}(z)\mathrm{S}(z)$. En effet on sait que toute série entière (de rayon de convergence infini) est la somme de sa série de Taylor en tout point (avec rayon de convergence infini), et $e^z - \mathrm{L}(z)$ développée au point $z = \lambda_1$ n'a des termes qu'à partir de $(z-\lambda_1)^{m_1}$, donc est de la forme $(z-\lambda_1)^{m_1}\mathrm{S}_0(z-\lambda_1) = (z-\lambda_1)^{m_1}\mathrm{S}_1(z)$, puis on itère : car l'annulation de $e^z - \mathrm{L}(z)$ et de ses dérivées aux $\lambda_2, \ldots, \lambda_t$ se transfère à S_1.

Finalement on remplace z par la matrice M et on obtient

$$e^{\mathrm{M}} - \mathrm{L}(\mathrm{M}) = \mathrm{T}(\mathrm{M})\mathrm{S}(\mathrm{M}) = 0$$

et c'est fini! □

C'est si simple et évident!

Sauf qu'il fallait connaître la solution avant!

On peut aussi deviner l'énoncé dans le cas diagonalisable, mais il faut un peu d'imagination pour en extraire l'énoncé général avec des multiplicités.

Si l'on résout le problème d'interpolation avec les multiplicités du polynôme caractéristique, il y aura une solution L_1 de degré strictement inférieur à n. Cette solution sera a fortiori une solution (pas de degré minimal, peut-être) du problème d'interpolation n'utilisant que les multiplicités du polynôme minimal. Et donc en tout cas la formule $e^{\mathrm{M}} = \mathrm{L}_1(\mathrm{M})$ sera assurée. On peut d'ailleurs faire directement la seconde preuve en y remplaçant le polynôme minimal par le polynôme caractéristique.

Le point d'algèbre qui ressort du truc, c'est que la connaissance du polynôme caractéristique (plus précisément : de ses racines et multiplicités!) suffit à déterminer un polynôme L_1 avec $e^{\mathrm{M}} = \mathrm{L}_1(\mathrm{M})$, alors que le polynôme caractéristique ne donne pas tous les invariants de similitude de M.