REMARQUES SUR L'EXPONENTIELLE DE MATRICE

JEAN-FRANÇOIS BURNOL

Soit $n \geqslant 1$ et $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et $V = \mathrm{Mat}_n(\mathbb{K})$. On s'intéresse à l'exponentielle de matrice $\exp(A) = \sum_{k=0}^{\infty} A^k/k!$. On a parfois besoin d'une norme sur $\mathrm{Mat}_n(\mathbb{K})$, par exemple la norme opératorielle (une norme ayant été choisie préalablement sur $\mathbb{K}^n$) qui, avec d'autres, a l'avantage que $\|AB\| \leqslant \|A\|\|B\|$.

1. L'exponentielle de A est un polynôme de A

Cet énoncé est hyper-standard mais je n'en aime pas la démonstration habituellement donnée, ou disons mieux, je la trouve incomplète et je vais donc donner ici une version enrichie, puis reprendre le problème autrement et exhiber au passage une solution comme polynôme d'interpolation (que A soit ou non diagonalisable).

1.1. **L'approche bateau.**
Soit $V_A = \{P(A), P \in \mathbb{K}[X]\}$ le sous-espace vectoriel de $\mathrm{Mat}_n(\mathbb{K})$ des polynômes en A. Les sommes partielles $S_k = \sum_{i \leqslant k} A^i/i!$ de la série exponentielle sont dans V_A donc aussi leur limite : en effet V_A est fermé car il est défini par un nombre fini de conditions linéaires (égal à sa co-dimension) ; ou encore on explique que toutes les normes étant équivalentes en dimension finie, V_A muni de la norme induite de celle de $\mathrm{Mat}_n(\mathbb{K})$ est complet, et donc fermé dans $\mathrm{Mat}_n(\mathbb{K})$ (ce deuxième argument n'a pas besoin de la dimension finie de $\mathrm{Mat}_n(\mathbb{K})$).

Ça, c'était l'argument standard, mais on veut un P avec $P(A) = \exp(A)$.

1.2. **L'approche bateau améliorée.**
Soit π_A le polynôme minimal de A, de degré p, et soit R_k le reste dans la division euclidienne de S_k par π_A. On a $R_k(A) = S_k(A)$ donc les $R_k(A)$ convergent vers $\exp(A)$. Mais l'application $\mathbb{K}[X]_{<p} \to \mathrm{Mat}_n(\mathbb{K})$ qui à P associe $P(A)$ est injective. Donc $\|P\|_A \coloneqq \|P(A)\|$ est une norme sur $\mathbb{K}[X]_{<p}$, et les R_k forment une suite de Cauchy pour cette norme ; toutes les normes sont équivalentes en dimension finie, donc les R_k convergent, c'est-à-dire chaque coefficient converge. La limite R, par continuité de $P \mapsto P(A)$ est telle que $R(A)$ est la limite de $R_k(A)$, c'est-à-dire $\exp(A)$.

Pour moi, l'argument standard sans ce second paragraphe n'a aucun intérêt. Je vais maintenant reprendre avec une approche faisant un peu d'algèbre et au passage expliquer le lien avec un problème d'interpolation.

1.3. **Reboot.**
Soit $\chi_A = \det(XI_n - A)$ le polynôme caractéristique et M sa matrice compagnon. Sur l'espace vectoriel $E = \mathbb{K}[X]/(\chi_A)$, M est la matrice de l'endomorphisme θ de multiplication par X, dans la base canonique $(1, X, \ldots, X^{n-1})$.

Considérons l'image de 1 par l'endomorphisme $\exp(\theta)$ dont la matrice est $\exp(M)$ dans cette même base : $\exp(\theta)(1) = d_0 \cdot 1 + d_1 X + \cdots + d_{n-1} X^{n-1}$ et notons T le polynôme avec les mêmes coefficients. La matrice $T(M)$ a comme endomorphisme associé la multiplication par T modulo (χ_A), qui est aussi le polynôme d'endomorphisme $T(\theta)$. Donc

$$\exp(\theta)(1) = T(\theta)(1)$$

Date: v1 7 décembre 2018, v2 19 décembre 2018.

En composant sur la gauche par des puissances de θ et en utilisant la commutatitivé, $\exp(\theta)(X^k) = T(\theta)(X^k)$ pour tout k et donc finalement

$$\exp(\theta) = T(\theta)$$

Soit maintenant f l'endomorphisme sur $\mathbb{K}^n$ de matrice A. Soit $x \in \mathbb{K}^n$ quelconque alors on sait qu'il y a un morphisme $\phi : E \to \mathbb{K}^n$ qui envoie une classe de polynôme $\widetilde{U}$ sur $U(f)(x)$ et que la relation fondamentale

$$\phi \circ \theta = f \circ \phi$$

est satisfaite.

Plus généralement $\phi \circ \theta^k = f^k \circ \phi$ pour toutes les puissances et donc aussi pour les sommes partielles de la série exponentielle et donc en passant à la limite $\phi \circ \exp(\theta) = \exp(f) \circ \phi$.

On a donc $\phi(\exp(\theta)(1)) = \exp(f)(x)$ mais on a vu précédemment $\exp(\theta) = T(\theta)$, et $\phi(T(\theta)(1)) = T(f)(x)$, donc $\exp(f)(x) = T(f)(x)$ et x est arbitraire donc $\exp(f) = T(f)$.

Dans la preuve ci-dessus j'ai utilisé un peu plus que les exponentielles de matrices, j'ai aussi utilisé les exponentielles d'endomorphismes car c'est le vocabulaire naturel.

À propos de ce T, considérons la division euclidienne

$$\sum_{k \leqslant N} \frac{X^k}{k!} = Q_N \chi_A + R_N, \quad \deg(R_N) < n$$

Si on évalue cela en le morphisme θ sur E, et puisque $\chi_A(\theta) = 0_{\mathrm{End}(E)}$, puis si on applique sur $1 \in E$ on obtient simplement R_N vu comme combinaison linéaire de la base canonique de E.

Comme la série $\sum_{k \leqslant N} \frac{\theta^k}{k!}$ converge dans $\mathrm{End}(E)$, on vient de prouver que les polynômes R_N, vus comme éléments de E convergeaient. Mais on connaît la limite, c'est T. Par conséquent les polynômes R_N vus maintenant dans $\mathbb{K}[X]$ (muni disons de la norme sup, attention bien sûr il n'est pas complet) convergent vers T vu comme un élément de $\mathbb{K}[X]$. Et on sait que $T(A)$ est aussi $\exp(A)$.

Il n'était pas totalement évident a priori que la série des restes R_N dans la division euclidienne convergeait (ce qui signifie simplement que chacun des n coefficients converge dans $\mathbb{K}$). Nos raisonnements ci-dessus nous en convainquent. De plus on a une vision concrète de ce polynôme limite comme la première colonne de l'exponentielle de la matrice compagnon de χ_A.

En partant de l'égalité polynomiale vue dans $\mathbb{C}$

$$S_N = \sum_{k \leqslant N} \frac{X^k}{k!} = Q_N \chi_A + R_N, \quad \deg(R_N) < n$$

il est clair que $S_N^{(j)}(\lambda) = R_N^{(j)}(\lambda)$ pour toute valeur propre λ (dans $\mathbb{C}$) et tout j inférieur strictement à la multiplicité. En passant à la limite lorsque $N \to \infty$ on en déduit

$$\forall \lambda, \forall j < \mathrm{mult}(\lambda), \quad e^\lambda = T^{(j)}(\lambda)$$

Il y a un seul polynôme de degré $< n$ qui vérifie ces propriétés d'interpolation car la différence de deux solutions est divisible par χ_A qui est de degré n. Donc on a aussi prouvé que la solution au problème d'interpolation nous donne un polynôme T avec $T(A) = \exp(A)$.

Attention que le polynôme minimal de A est peut-être de degré $< n$, donc il peut y avoir d'autres solutions (pas au problème d'interpolation avec n conditions, mais à l'équation $\exp(A) = T(A)$). Mais on peut recommencer tout le raisonnement en remplaçant le polynôme caractéristique χ_A par le polynôme minimal : il y a un polynôme T de degré strictement inférieur au polynôme minimal et tel que

$T(A) = \exp(A)$, et il est unique par définition du polynôme minimal. Ce polynôme est la limite des restes dans la division euclidienne des sommes partielles de la série exponentielle par le polynôme minimal de A ; et ses coefficients sont la première colonne de l'exponentielle de la matrice compagnon du polynôme minimal. Et ce polynôme est aussi caractérisé uniquement par le fait d'être une solution de degré minimal au problème d'interpolation (dans le plan complexe).

2. Différentielle de l'exponentielle de matrice

Je note $\|\cdot\|$ une norme sur $\mathrm{Mat}_n(\mathbb{K})$ telle que $\|AB\| \leqslant \|A\|\|B\|$ toujours.

Posons $S_n(A,H) = \sum_{k=0}^{n-1} A^k H A^{n-k}$. Il est clair que

$$\|(A+H)^n - A^n - S_n(A,H)\| \leqslant (\|A\| + \|H\|)^n - \|A\|^n - S_n(\|A\|, \|H\|)$$

puisque $(A+H)^n - A^n - S_n(A,H)$ est une somme (de coefficients tous 1) de monômes (non-commutatifs) en A et H. On remarque au passage

$$S_n(\|A\|, \|H\|) = n\|A\|^{n-1}\|H\|$$

La série de terme général $(A+H)^n/n!$ est absolument convergente, aussi celle avec $A^n/n!$, donc aussi celle avec $S_n(A,H)/n!$ ($S_0(A,H) = 0$). Et nous obtenons

$$\left\| \exp(A+H) - \exp(A) - \sum_{n=1}^{\infty} \frac{S_n(A,H)}{n!} \right\| \leqslant e^{\|A\|+\|H\|} - e^{\|A\|} - e^{\|A\|}\|H\| = \mathcal{O}(\|H\|^2)$$

Ceci prouve que l'exponentielle de matrices est une application différentiable au point A et que $\mathrm{Dexp}_A(H) = \sum_{n=1}^{\infty} \frac{S_n(A,H)}{n!}$.

On va l'exprimer d'une manière plus sympathique.

Notons $L \in \mathrm{End}(\mathrm{Mat}_n(\mathbb{K}))$ le morphisme de multiplication à gauche par A et R celui de multiplication à droite par A. Il est crucial que $LR = RL$, heureusement c'est évident.

Ainsi $S_n(A,H) = (\sum_{k=0}^{n-1} L^k R^{n-1-k})(H)$. On a une identité remarquable (grâce à la commutativité !) :

$$(L-R)\sum_{k=0}^{n-1} L^k R^{n-1-k} = -R^n + L^n$$

Malheureusement, $L - R$ n'est sûrement pas inversible puique A est dans son noyau (et si A est nulle $L - R$ aussi). Peu importe soit $0 < \epsilon$ suffisamment petit, alors $L - R + \epsilon$ est inversible. Il est certainement vrai (dans $\mathrm{End}(\mathrm{Mat}_n(\mathbb{K}))$) que

$$\lim_{\epsilon\to 0} \sum_{n=1}^{\infty} \frac{1}{n!}\left(\sum_{k=0}^{n-1} L^k (R+\epsilon)^{n-1-k}\right) = \sum_{n=1}^{\infty} \frac{1}{n!}\left(\sum_{k=0}^{n-1} L^k R^{n-1-k}\right)$$

par convergence normale donc uniforme en $\epsilon \in [0,1]$.

On écrit maintenant (tout commute !) :

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{n!}\left(\sum_{k=0}^{n-1} L^k (R+\epsilon)^{n-1-k}\right) &= (L-R+\epsilon)^{-1} \sum_{n=1}^{\infty} \frac{L^n - (R-\epsilon)^n}{n!} \\
&= (L-R+\epsilon)^{-1}(e^L - e^{R-\epsilon}) \\
&= (L-R+\epsilon)^{-1}(e^{L-R+\epsilon} - I)e^{R-\epsilon} \\
&= \sum_{n=1}^{\infty} \frac{(L-R+\epsilon)^{n-1}}{n!} \cdot e^{R-\epsilon}
\end{aligned}$$

Le passage à la limite (dans $\mathrm{End}(\mathrm{Mat}_n(\mathbb{K}))$) ne pose aucun problème à nouveau par convergence normale donc uniforme en la variable $\epsilon \in [0,1]$.

Et nous pouvons conclure, avec $\mathrm{ad}_A = L - R$, $\mathrm{ad}_A(H) = AH - HA$:

$$\mathrm{Dexp}_A = \sum_{n=1}^{\infty} \frac{\mathrm{ad}_A^{n-1}}{n!} \cdot e^R$$

$$\mathrm{Dexp}_A(H) = \sum_{n=1}^{\infty} \frac{\mathrm{ad}_A^{n-1}}{n!}(H) \cdot e^A$$

On peut aussi manipuler ainsi :

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^{n-1} L^k (R+\epsilon)^{n-1-k} \right) &= (L - R + \epsilon)^{-1} \sum_{n=1}^{\infty} \frac{L^n - (R-\epsilon)^n}{n!} \\
&= (L - R + \epsilon)^{-1} (e^L - e^{R-\epsilon}) \\
&= (L - R + \epsilon)^{-1} (I - e^{R-\epsilon-L}) e^L \\
&= (R - \epsilon - L)^{-1} (e^{R-\epsilon-L} - I) e^L \\
&= \sum_{n=1}^{\infty} \frac{(R - L - \epsilon)^{n-1}}{n!} \cdot e^L
\end{aligned}$$

ce qui donne comme limite

$$\mathrm{Dexp}_A = \sum_{n=1}^{\infty} \frac{(-\mathrm{ad}_A)^{n-1}}{n!} \cdot e^L$$

Et en rappelant que e^L commute avec L et avec R donc avec $L - R = \mathrm{ad}_A$:

$$\mathrm{Dexp}_A(H) = \sum_{n=1}^{\infty} \frac{(-\mathrm{ad}_A)^{n-1}}{n!} (e^A H) = e^A \cdot \sum_{n=1}^{\infty} \frac{(-\mathrm{ad}_A)^{n-1}}{n!} (H)$$

Cette belle formule nous explique que l'écart de Dexp_A à la multiplication à gauche par e^A est lié au défaut de commutativité de H avec A.