Équations différentielles

Cours de L3 par Frédéric Hélein[1], janvier–avril 2021

Mardi 9 février 2021

2.4 La solution des équations linéaires à coefficients constants non homogènes

A l'aide de l'exponentielle des matrices nous pouvons exprimer la solution d'une équation différentielle de la forme

$$\frac{dx}{dt}(t) = A\ x(t) + B(t), \tag{1}$$

où $A \in M(n, \mathbb{R})$ et $B \in \mathcal{C}^0(I, \mathbb{R}^n)$. Nous utilisons pour cela la méthode de la *variation de la constante*, due à Joseph-Louis Lagrange. L'idée est de chercher les solutions de (1) comme étant des déformations des solutions des solutions de l'équation différentielle linéaire homogène associée $\frac{dy}{dt} = Ay$. Nous savons que les solutions de cette équations sont toutes de la forme $y(t) = e^{tA}y_0$, où $y_0 \in \mathbb{R}^n$. Nous y remplaçons la constante y_0 par une fonction inconnue $z \in \mathcal{C}^1(I, \mathbb{R}^n)$.

Ainsi nous posons $z(t) := e^{-tA}x(t)$, de sorte que

$$\begin{array}{rrcl} & x(t) & = & e^{tA}z(t) \\ \text{et donc :} & \dfrac{dx}{dt}(t) & = & e^{tA}\left(\dfrac{dz}{dt}(t) + Az(t)\right) \end{array}$$

En remplaçant dans l'équation (1) nous obtenons

$$e^{tA}\left(\frac{dz}{dt} + Az\right) = A\ \left(e^{tA}z\right) + B$$

et donc en simplifiant, on obtient $e^{tA}\frac{dz}{dt} = B$, qui équivaut à

$$\frac{dz}{dt}(t) = e^{-tA}B(t).$$

La solution de cette équation est obtenue par une simple intégration

$$z(t) - z(t_0) = \int_{t_0}^{t} e^{-\tau A}B(\tau)d\tau$$

Si nous imposons la condition initiale $x(t_0) = x_0$, qui équivaut à $z(t_0) = e^{-t_0A}x_0$, nous obtenons $z(t) = e^{-t_0A}x_0 + \int_{t_0}^{t} e^{-\tau A}B(\tau)d\tau$ et donc

$$x(t) = e^{(t-t_0)A}x_0 + \int_{t_0}^{t} e^{(t-\tau)A}B(\tau)d\tau \tag{2}$$

1. Université de Paris, Licence 3 de Mathématiques, helein@math.univ-paris-diderot.fr

En particulier nous en concluons que l'ensemble $\mathcal{S}$ des solutions de (1) est un sous-espace affine de $\mathcal{C}^1(I,\mathbb{R}^n)$ de dimension n. En effet $\mathcal{S}$ est l'image de l'application affine

$$\begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathcal{C}^1(I,\mathbb{R}^n) \\ x_0 & \longmapsto & [t \longmapsto u(t) + e^{(t-t_0)A}x_0] \end{array}$$

où $u(t) := \int_{t_0}^t e^{(t-\tau)A}B(\tau)d\tau$.

2.5 Exemples d'exponentielles de matrice

Nous avons besoin de savoir calculer l'exponentielle d'une matrice A dans des cas pratiques. Nous connaissons déjà le principe dans le cas où A est diagonalisable : il est alors possible de décomposer $A = P\Delta P^{-1}$, où $P \in GL(n,\mathbb{C})$ et $\Delta \in M(n,\mathbb{C})$ est une matrice diagonale. Nous exploitons le fait que

$$A^2 = (P\Delta P^{-1})^2 = P\Delta P^{-1}\ P\Delta P^{-1} = P\Delta^2 P^{-1}$$

puis $A^3 = (P\Delta P^{-1})^3 = P\Delta^3 P^{-1}$ et, d'une façon générale, $A^k = P\Delta^k P^{-1}$, $\forall k \in \mathbb{N}$ (avec la convention $A^0 = 1_n$). Donc

$$e^{tA} = \sum_{k=0}^{+\infty} \frac{t^k A^k}{k!} = \sum_{k=0}^{+\infty} \frac{t^k P\Delta^k P^{-1}}{k!} = P \sum_{k=0}^{+\infty} \frac{t^k \Delta^k}{k!} P^{-1} = Pe^{t\Delta}P^{-1}$$

Voici quelques exemples.

(i) $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ Dans ce cas le polynôme caractéristique de A est $P_A(\lambda) = \det(\lambda 1_2 - A) = (\lambda - 1)(\lambda + 1)$. Les valeurs propres -1 et 1 sont distinctes, donc A est diagonalisable. Des vecteurs propres pour ces valeurs propres sont respectivement

$$u_{-1} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad \text{et} \quad u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Nous en déduisons la matrice de passage

$$P = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{avec} \quad P^{-1} = \frac{1}{2}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

Donc

$$A = \frac{1}{2}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

et

$$e^{tA} = \frac{1}{2}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} e^{-t} & 0 \\ 0 & e^t \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}$$

(ii) $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ Le polynôme caractéristique de A est $P_A(\lambda) = (\lambda - i)(\lambda + i)$. Les valeurs propres $-i$ et i sont distinctes, donc A est diagonalisable *dans* $\mathbb{C}$. Des vecteurs propres pour ces valeurs propres sont respectivement

$$u_{-i} = u := \begin{pmatrix} 1 \\ i \end{pmatrix} \quad \text{et} \quad u_1 = \bar{u} = \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

Nous en déduisons la matrice de passage

$$P = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \quad \text{avec} \quad P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

Donc

$$A = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

et

$$e^{tA} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\begin{pmatrix} e^{-it} & 0 \\ 0 & e^{it} \end{pmatrix}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

(iii) $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ Dans ce cas A n'est pas diagonalisable. En effet le polynôme caractéristique est $P_A(\lambda) = (\lambda - 2)^3$ (ce que l'on peut déduire directement de la forme triangulaire) et donc 2 est l'unique valeur propre de A, avec multiplicité 3. Si A était diagonalisable, on aurait $A = P\Delta P^{-1}$, avec $\Delta = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, mais comme Δ commute avec toutes les matrices, cela entraînerait que $A = \Delta$, ce qui est impossible.

En revanche on a $A = \Delta + N$, avec $N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, et on remarque que

— $[A, \Delta] = 0$ ou $\Delta A = A\Delta$, c'est à dire A et Δ *commutent* ;

— N est nilpotente, plus précisément,

$$N^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad N^3 = 0.$$

La première remarque a pour conséquence que $e^{tA} = e^{t\Delta + tN} = e^{t\Delta}e^{tN}$. La deuxième remarque implique que

$$e^{tN} = \sum_{k=0}^{+\infty} \frac{t^k N^k}{k!} = 1 + tN + \frac{1}{2}t^2N^2 = \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}$$

De plus, on a bien évidemment $e^{t\Delta} = e^{2t}1_3$. En conclusion on a donc

$$e^{tA} = e^{2t}\begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

2.6 Le calcul de l'exponentielle d'une matrice dans le cas général

2.6.1 Quelques rappels d'algèbre linéaire

Lorsqu'une matrice n'est diagonalisable ni dans $\mathbb{R}$, ni dans $\mathbb{C}$, parce que certaines de ses valeurs propres sont multiples, il est possible de la réduire sous la forme triangulaire. Cependant il sera très utile de le faire de manière à obtenir des blocs diagonaux qui ont une structure analogue à celle de la matrice A étudiée au paragraphe (iii) précédemment. Dans ce qui suit, nous considérons un endomorphisme φ d'un espace vectoriel E de dimension n.

Le polynôme caractéristique de φ dans une base $\mathcal{B}$ de E est le polynôme P_φ défini par $P_\varphi(\lambda) = \det(\lambda 1_n - A)$, où A est la matrice de φ dans la base $\mathcal{B}$. Ce polynôme ne dépend pas du choix de la base.

Théorème 2.1 (de Cayley–Hamilton) $P_\varphi(\varphi) = 0$.

Soit $r \in \{1, \cdots, n\}$ le nombre de valeurs propres distinctes, soit $\lambda_1, \cdots, \lambda_r \in \mathbb{C}$ les valeurs propres de φ (comptées sans répétition) et, pour tout $j \in \{1, \cdots, r\}$, soit m_j la multiplicité de λ_j dans le polynôme P_φ. On a alors $m_1 + \cdots + m_j = n$.

Rappelons que, lorsqu'elle est possible, la diagonalisation de φ repose sur l'introduction des *sous-espaces propres* $E_j := \mathrm{Ker}(\lambda_j 1_E - \varphi)$ de φ, le fait que chaque E_j est stable par φ, i.e. $\forall u \in E_j$, $\varphi(u) \in E_j$ et que la somme $E_1 + \cdots + E_r$ est directe. Alors on a $E_1 \oplus \cdots \oplus E_r = E$ ssi φ est diagonalisable. Si cela n'est pas le cas, on fait appel aux sous-espaces caractéristiques.

Définition 2.1 *Pour tout $j \in \{1, \cdots, r\}$ on définit le* ***sous-espace caractéristique*** *de φ pour la valeur propre λ_j comme étant $C_j := Ker(\lambda_j 1_E - \varphi)^{m_j}$.*

On a bien sûr $E_j \subset C_j$ et les résultats :

Lemme 2.1 *Pour tout $j \in \{1, \cdots, r\}$, C_j est stable par φ, i.e. $\forall u \in C_j$, $\varphi(u) \in C_j$.*

Proposition 2.1 *La somme $C_1 + \cdots + C_r$ est directe et on a toujours $C_1 \oplus \cdots \oplus C_r = E$.*

Comme, pour tout $j \in \{1, \cdots, r\}$, C_j est stable par φ, la restriction de φ à C_j est un endomorphisme de C_j. Nous notons $\varphi_j = \varphi|_{C_j} \in \mathrm{End}(C_j)$ et, étant donnée une base $\mathcal{B}_j = \{u_{j,1}, \cdots, u_{j,m_j}\}$ de C_j, nous notons A_j la matrice de φ_j dans cette base. Alors comme $C_1 \oplus \cdots \oplus C_r = E$, la famille $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ est une base de E et, en respectant l'ordre indiqué des vecteurs de cette base, la matrice de A dans cette base est la matrice diagonale par bloc :

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}$$

et donc

$$e^{tA} = \begin{pmatrix} e^{tA_1} & 0 & \cdots & 0 \\ 0 & e^{tA_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{tA_r} \end{pmatrix}$$

Il nous reste à trouver, pour chaque j, une base $\mathcal{B}_j$ et donc une matrice A_j telle que le calcul de e^{tA_j} soit pratiquable. Posons $\beta_j := \varphi_j - \lambda_j 1_{C_j} = (\varphi - \lambda_j 1_E)|_{C_j} \in \mathrm{End}(C_j)$, de sorte que

$$(\beta_j)^{m_j} = (\varphi - \lambda_j 1_E)^{m_j}|_{C_j} = 0$$

Alors $\varphi_j = \lambda_j 1_{C_j} + \beta_j$ et, comme 1_{C_j} commute avec tous les éléments de $\mathrm{End}(C_j)$, $[\lambda_j 1_{C_j}, \beta_j] = 0$ et donc

$$e^{t\varphi_j} = e^{t\lambda_j} e^{t\beta_j} = e^{t\lambda_j} \sum_{k=0}^{m_j-1} \frac{t^k (\beta_j)^k}{k!}$$

En effet la série qui définit $e^{t\beta_j}$ est un polynôme car β_j est nilpotent.

Enfin il est possible de trouver une base $\mathcal{B}_j$ dans laquelle la matrice de β_j est triangulaire supérieure, avec que des zéros sur la diagonale. Dans le lemme qui suit, nous omettons l'indice j pour alléger l'écriture.

Lemme 2.2 *Soit C un espace vectoriel de dimension m et soit $\beta \in \mathrm{End}(C)$ tel que $\beta^m = 0$. Alors il existe une base $(u_1, \cdots, u_m)$ dans laquelle la matrice de β est triangulaire supérieure et tous les coefficients sur la diagonale sont nuls.*

(Bien entendu on peut remplacer « triangulaire supérieure » par « triangulaire supérieure » sans difficulté.)

Démonstration — Nous avons une suite d'inclusions au sens large

$$\mathrm{Ker}\beta \ \subset \ \mathrm{Ker}\beta^2 \ \subset \ \cdots \ \subset \ \mathrm{Ker}\beta^m = C$$

Pour $k \in \{1, \cdots, m\}$, notons $j_k = \dim(\mathrm{Ker}\beta^k)$, de sorte que $0 \leq j_1 \leq j_2 \leq \cdots \leq j_m = m$.

— choisissons une base $\mathcal{B}^{(1)} = \{u_1^{(1)}, \cdots, u_{j_1}^{(1)}\}$ de $\mathrm{Ker}\beta$;

— complétons $\mathcal{B}^{(1)}$ en une base $\mathcal{B}^{(1)} \cap \mathcal{B}^{(2)}$ de $\mathrm{Ker}\beta^2$, où $\mathcal{B}^{(2)} = \{u_1^{(2)}, \cdots, u_{j_2-j_1}^{(2)}\}$;

— complétons $\mathcal{B}^{(1)} \cap \mathcal{B}^{(2)}$ en une base $\mathcal{B}^{(1)} \cap \mathcal{B}^{(2)} \cap \mathcal{B}^{(3)}$ de $\mathrm{Ker}\beta^3$, où $\mathcal{B}^{(3)} = \{u_1^{(3)}, \cdots, u_{j_3-j_2}^{(3)}\}$;

— etc.

— à la fin nous complétons $\mathcal{B}^{(1)} \cap \cdots \cap \mathcal{B}^{(m-1)}$ en une base $\mathcal{B}^{(1)} \cap \cdots \cap \mathcal{B}^{(m-1)} \cap \mathcal{B}^{(m)}$ de C, où $\mathcal{B}^{(m)} = \{u_1^{(m)}, \cdots, u_{j_m - j_{m-1}}^{(m)}\}$.

Bien entendu, si à une étape k, on a $j_k = j_{k-1}$, il n'y a rien à faire. Comme β envoie $\mathrm{Ker}\beta^{k+1}$ sur $\mathrm{Ker}\beta^k$, l'image de chaque vecteur dans $\mathcal{B}^{(k+1)}$ est une combinaison linéaire des vecteurs dans $\mathcal{B}^{(1)} \cap \cdots \cap \mathcal{B}^{(k)}$, ce qui fait que la matrice de β dans la base (ordonnée)

$$\left(u_1^{(1)}, \cdots, u_{j_1}^{(1)}, u_1^{(2)}, \cdots, u_{j_2 - j_2}^{(2)}, \cdots \quad \cdots, u_1^{(m)}, \cdots, u_{j_m - j_{m-1}}^{(m)}\right)$$

a la structure annoncée. □

Compléments d'algèbre linéaire

Sur la preuve du théorème de Cayley–Hamilton

Il en existe plusieurs, en voici une possible. Soit E un espace vectoriel sur $\mathbb{R}$ de dimension n et soit $\varphi \in \mathrm{End}(E)$. On note P_φ son polynôme caractéristique.

Définition *Un vecteur $u \in E$ est dit **cyclique** pour φ si la famille $(u, \varphi(u), \varphi^2(u), \cdots, \varphi^{n-1}(u))$ est libre.*

a) Nous supposons dans un premier temps qu'il existe un vecteur cyclique u pour φ. Alors $(u, \varphi(u), \varphi^2(u), \cdots, \varphi^{n-1}(u))$ est une base de E et donc nous pouvons décomposer $\varphi^n(u)$ dans cette base : $\exists a_0, a_1, \cdots, a_{n-1} \in \mathbb{R}$,

$$\varphi^n(u) = -a_0 u - a_1 \varphi(u) - \cdots - a_{n-1}\varphi^{(n-1)}(u) \tag{3}$$

On en déduit que la matrice M de φ dans la base $(u, \varphi(u), \varphi^2(u), \cdots, \varphi^{n-1}(u))$ a pour expression

$$M = \begin{pmatrix} 0 & 0 & \cdots & -a_0 \\ 1 & 0 & & -a_1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

Il n'est pas très difficile de calculer le polynôme caractéristique de cette matrice, et donc, en même temps, P_φ, ce qui donne :

$$P_\varphi(\lambda) = \det(\lambda 1_n - M) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

A posteriori la relation (3) signifie que $P_\varphi(\varphi)(u) = 0$. Cela entraîne aussi

$$\forall j \in \{1, \cdots, n-1\}, \quad P_\varphi(\varphi)\left(\varphi^j(u)\right) = \varphi^j\left(P_\varphi(\varphi)(u)\right) = 0$$

et donc, comme $(u, \varphi(u), \varphi^2(u), \cdots, \varphi^{n-1}(u))$ est une base de E, $P_\varphi(\varphi) = 0$. Cela le montre le théorème de Cayley–Hamilton dans le cas où il existe un vecteur cyclique pour φ.

b) Le cas général s'obtient en observant que :

(i) l'application

$$\begin{array}{ccc} \mathrm{End}(E) & \longrightarrow & \mathbb{R} \\ \varphi & \longmapsto & P_\varphi(\varphi) \end{array}$$

est un polynôme en n^2 variables et est donc une application continue.

(ii) cette application s'annule sur le sous-ensemble de $\mathrm{End}(n, \mathbb{R})$ des endomorphismes qui admettent un vecteur cyclique.

(iii) le sous-ensemble de $\mathrm{End}(n, \mathbb{R})$ des endomorphismes qui admettent un vecteur cyclique est dense dans $\mathrm{End}(n, \mathbb{R})$.

Les observations (ii) et (iii) entraînent effectivement qu'il existe un sous-ensemble dense de $\mathrm{End}(E)$ sur lequel $P_\varphi(\varphi)$ s'annule, ce qui entraîne, grâce au (i) que que $P_\varphi(\varphi) = 0$ partout.

L'assertion (i) est évidente, l'assertion (ii) a été montrée précédemment en a), il nous reste à montrer (ii). Fixons une base $(e_1, \cdots, e_n)$ de E et considérons l'endomorphisme $\chi \in \mathrm{End}(E)$ tel que $\chi(e_i) = e_{i+1}$, $\forall i \in \{1, \cdots, n-1\}$ et $\chi(e_n) = 0$. Cet endomorphisme admet bien entendu e_1 comme vecteur cyclique.

Soit, à présent, $\varphi \in \mathrm{End}(E)$ et montrons qu'on peut approcher φ par une suite d'endomorphismes qui admettent e_1 comme vecteur cyclique. Il suffit pour cela de considérer $\varphi_s := (1-s)\varphi + s\chi$, pour $s \in \mathbb{R}$. Cet endomorphisme admet e_1 comme vecteur cyclique ssi le rang de $(e_1, \varphi_s(e_1), \cdots, (\varphi_s)^{n-1}(e_1))$ est égal à n, c'est à dire ssi

$$\det_{(e_1,\cdots,e_n)} \left(e_1, \varphi_s(e_1), \cdots, (\varphi_s)^{n-1}(e_1)\right) \neq 0$$

L'expression à gauche est un polynôme en la variable s de degré $\frac{n(n-1)}{2}$ qui est non nul car il vaut 1 pour $s = 1$. Il admet donc au plus $\frac{n(n-1)}{2}$ racines réelles et est non nul en dehors de ces racines, donc pour s prenant ses valeurs dans un sous-ensemble dense de $\mathbb{R}$. En prenant une suite de valeurs de s qui évite ces racines et qui tend vers 0, on obtient une suite d'endomorphismes φ_s qui tend vers φ et qui admettent e_1 comme vecteur cyclique.