

Chapitre 1

Exponentielle d'une matrice

1.1 Introduction

Dans le chapitre "séries entières", nous avons étudié les fonctions de la forme $f(z) = \sum_{n \geq 0} a_n z^n$, $a_n, z \in \mathbb{C}$. Ces fonctions sont limite des sommes partielles $S_N(z) = \sum_{n=0}^N a_n z^n$. Ces sommes partielles sont des polynômes en z . Puisque pour une matrice A donnée, nous savons calculer $P(A)$ pour tout polynôme P . Peut-on donner un sens à $f(A)$, pour f une fonction définie par une série entière? cela revient à définir "la notion de la limite" pour une suite de matrices. Pour cela on aura besoin d'une norme d'algèbre sur l'espace des matrices, c'est -à-dire, une norme, $\|\cdot\|$, d'espace vectoriel vérifiant en plus, $\|AB\| \leq \|A\| \|B\|$ pour toutes matrices A, B .

Dans ce chapitre, on s'intéresse particulièrement à la fonction exponentielle d'une matrice. Comme application, nous utilisons l'exponentielle d'une matrice, pour donner des méthodes de résolution des systèmes différentiels linéaires.

Dans toute la suite, \mathbb{K} désigne \mathbb{R} ou \mathbb{C} .

Si $n \geq 1$, \mathbb{K}^n désigne l'espace vectoriel sur \mathbb{K} des $x = (x_1, x_2, \dots, x_n)$ où $x_i \in \mathbb{K}$. Ainsi $\mathbb{K}^n = \mathbb{R}^n$ ou \mathbb{C}^n .

L'espace \mathbb{K}^n sera muni du produit scalaire : pour $x = (x_1, x_2, \dots, x_n)$ et $y = (y_1, y_2, \dots, y_n)$ dans \mathbb{K}^n ,

$$\langle x, y \rangle = \sum_{k=1}^n x_k \overline{y_k}.$$

Alors, l'espace \mathbb{K}^n sera muni de la norme euclidienne :

$$\|x\| = \|(x_1, x_2, \dots, x_n)\| = \left(\sum_{k=1}^n |x_k|^2 \right)^{\frac{1}{2}},$$

et de la distance (ou métrique) associée

$$d(x, y) = \|x - y\|, \quad x, y \in \mathbb{K}^n$$

1.2 Séries de vecteurs dans \mathbb{K}^n

Soit $(x_k) = ((x_{k1}, x_{k2}, \dots, x_{kn}))$, où $k \geq 0$, une suite de vecteurs dans \mathbb{K}^n . On dira que la série de vecteurs $\sum_{k \geq 0} x_k$ converge, et a pour somme s , si la suite des

"sommes partielles" $s_K = \sum_{k=0}^K x_k$ tend vers s quand K tend vers $+\infty$. Pour cela, il faut et il suffit (critère de Cauchy) que

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N} \text{ tel que } \forall k' \geq k \geq N_\varepsilon, \|x_k + x_{k+1} + \dots + x_{k'}\| \leq \varepsilon.$$

On dit que la série de vecteurs $\sum_{k \geq 0} x_k$ est absolument (ou normalement) convergente si la série de nombres réels positifs $\sum_{k \geq 0} \|x_k\|$ converge. ce qui implique la série de vecteurs $\sum_{k \geq 0} x_k$ elle-même converge, et on a $\|\sum_{k \geq 0} x_k\| \leq \sum_{k \geq 0} \|x_k\|$.

Notons que la série de vecteurs $\sum_{k \geq 0} x_k$ converge, et a pour somme s , si et seulement si pour $i = 1, \dots, n$, chaque la série (numérique) composante $\sum_{k \geq 0} x_{ki}$ converge vers la composante s_i de s . Ainsi l'Analyse dans \mathbb{K}^n se ramène à faire de l'Analyse dans $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

1.3 Norme d'une matrice carrée

Soit $\mathcal{M}_n(\mathbb{K})$, $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , l'algèbre (= espace vectoriel muni du produit de matrices) des matrices carrées $A = (a_{ij})$ à n lignes et n colonnes, à coefficients $a_{ij} \in \mathbb{K}$ où $i = 1, \dots, n$; $j = 1, \dots, n$.

Définition 1.1. Si $A \in \mathcal{M}_n(\mathbb{K})$, posons

$$\|A\| = \sup_{x \in \mathbb{K}^n, \|x\|=1} \|Ax\|,$$

nous appelons ce nombre la norme de la matrice A (associée à la norme vectorielle euclidienne $\|x\|$)

Rappelons que d'après l'inégalité de Cauchy-Schwarz, on a

$$\left| \sum_{j=1}^n a_{ij} x_j \right|^2 \leq \left(\sum_{j=1}^n |a_{ij}|^2 \right) \left(\sum_{j=1}^n |x_j|^2 \right),$$

donc si $A = (a_{ij})$, pour tout $x \in \mathbb{K}^n$ tel que $\|x\| = \left(\sum_j |x_j|^2 \right)^{\frac{1}{2}} = 1$:

$$\|Ax\| = \left(\sum_i \left| \sum_{j=1}^n a_{ij} x_j \right|^2 \right)^{\frac{1}{2}} \leq \left(\sum_i \sum_j |a_{ij}|^2 \right)^{\frac{1}{2}} \cdot \left(\sum_j |x_j|^2 \right)^{\frac{1}{2}} \leq M,$$

où M est la constante $\left(\sum_i \sum_j |a_{ij}|^2 \right)^{\frac{1}{2}}$. Donc le sup de la définition existe.

Théorème 1.1. Si $A \in \mathcal{M}_n(\mathbb{K})$, alors on a

1. Pour tout $x \in \mathbb{K}^n$, on a $\|Ax\| \leq \|A\| \|x\|$.
2. Si $A, B \in \mathcal{M}_n(\mathbb{K})$, $\lambda \in \mathbb{K}$, on a les propriétés suivantes (fondamentales d'une norme) :
 - (i) $\|A\|$ est un réel positif et $\|A\| = 0 \iff A = 0$;
 - (ii) $\|A + B\| \leq \|A\| + \|B\|$ et $\|\lambda A\| = |\lambda| \|A\|$.

3. $\|AB\| \leq \|A\| \|B\|$ et donc pour tout entier $k \geq 0$, on a

$$\|A^k\| \leq \|A\|^k \quad \text{et} \quad \|I_d\| = 1 \quad \text{où } I_d \text{ est la matrice identité.}$$

4. Soit S une matrice unitaire (orthogonale si $\mathbb{K} = \mathbb{R}$), c-à-d S est inversible et $S^{-1} = S^*$. Alors

$$\|S\| = 1 \quad \text{et} \quad \|S^{-1}AS\| = \|A\| \quad \forall A \in \mathcal{M}_n(\mathbb{K}).$$

5. $\|A\|^2 = \|A^*A\| = \|AA^*\| = \|A^*\|^2$.

6. $\|A^*A\| = \max\{\lambda; \lambda \in \sigma_p(A^*A)\}$. D'où
 $\|A\| = \text{la racine carrée de la plus grande valeur propre de la matrice } A^*A$.

Démonstration (1) Pour $x = 0$ l'inégalité est évidente. Supposons que $x \neq 0$, posons $y = \frac{x}{\|x\|}$. Alors $\|y\| = 1$ et donc $\|Ay\| \leq \|A\|$. Par conséquent :

$$\|Ax\| = \|A(\|x\|y)\| = \|\|x\|Ay\| = \|x\| \|Ay\| \leq \|A\| \|x\|.$$

(2) (i) Si $\|A\| = 0$, d'après (1), pour tout $x \in \mathbb{K}^n$, $\|Ax\| = 0$ et donc $Ax = 0$.

(ii) D'après l'inégalité triangulaire de la norme dans \mathbb{K}^n , on a $\|Ax + Bx\| \leq \|Ax\| + \|Bx\| \quad \forall x \in \mathbb{K}^n$. Donc

$$\|A+B\| = \sup_{\|x\|=1} \|(A+B)x\| = \sup_{\|x\|=1} \|Ax+Bx\| \leq \sup_{\|x\|=1} \|Ax\| + \sup_{\|x\|=1} \|Bx\| = \|A\| + \|B\|.$$

L'égale $\|\lambda A\| = |\lambda| \|A\|$ est facile à vérifier.

(3) On a

$$\|AB\| = \sup_{\|x\|=1} \|(AB)x\| = \sup_{\|x\|=1} \|A(Bx)\| \leq \|A\| \sup_{\|x\|=1} \|Bx\| = \|A\| \|B\|.$$

En particulier $\|A^2\| \leq \|A\|^2$ puis par récurrence sur k ,

$$\|A^k\| = \|A^{k-1}A\| \leq \|A^{k-1}\| \|A\| \leq \|A\|^{k-1} \|A\| = \|A\|^k.$$

D'autre part, $\|I_d\| = \sup_{\|x\|=1} \|I_dx\| = \sup_{\|x\|=1} \|x\| = 1$.

(4) Si S est unitaire, c'est une isométrie (c-à-d, $\|Sx\| = \|x\|$, $\forall x \in \mathbb{K}^n$), donc $\|S\| = \sup_{\|x\|=1} \|Sx\| = \sup_{\|x\|=1} \|x\| = 1$. De plus S^{-1} est aussi unitaire, donc $\|S^{-1}\| = 1$. Par conséquent,

$$\|S^{-1}AS\| \leq \|S^{-1}\| \|A\| \|S\| = \|A\|.$$

Inversement

$$\|A\| = \|S(S^{-1}AS)S^{-1}\| \leq \|S\| \|S^{-1}AS\| \|S^{-1}\| = \|S^{-1}AS\|.$$

Donc $\|S^{-1}AS\| = \|A\|$.

(5) Puisque (5) est vérifié si $A = 0$, on suppose que $A \neq 0$. Alors on a $\|Ax\|^2 = \langle Ax, Ax \rangle = \langle A^*Ax, x \rangle \leq \|A^*Ax\| \|x\|$. Donc

$$\|A\|^2 = \sup_{\|x\|=1} \|Ax\|^2 \leq \sup_{\|x\|=1} \|A^*Ax\| = \|A^*A\| \leq \|A^*\| \|A\|, \quad (*)$$

donc $\|A\|^2 \leq \|A^*\| \|A\|$ et donc $\|A\| \leq \|A^*\|$. Dans cette dernière inégalité, si on remplace A par A^* , donc A^* par $(A^*)^* = A$, on trouve $\|A\| = \|A^*\|$. Maintenant, (*) implique (5).

(6) Puisque la matrice A^*A est égale à son adjointe, il existe une matrice unitaire S telle que $S^{-1}(A^*A)S = D$ soit diagonale. D'autre part, les valeurs propres de A^*A , sont des nombres réels ≥ 0 . En effet si $\lambda \in \sigma_p(A^*A)$, alors il existe $v \in \mathbb{K}^n$ non nul, tel que $A^*Av = \lambda v$. Alors

$$0 \leq \|Av\|^2 = \langle Av, Av \rangle = \langle A^*Av, v \rangle = \langle \lambda v, v \rangle = \lambda \langle v, v \rangle = \lambda \|v\|^2$$

ce qui implique que $\lambda \geq 0$. Donc il existe $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ tel que $S^{-1}A^*AS = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. En utilisant, (4), on obtient $\|A^*A\| = \|D\|$. Mais

$$\|D\|^2 = \sup_{|x_1|^2 + \dots + |x_n|^2 = 1} \{\lambda_1^2 |x_1|^2 + \dots + \lambda_n^2 |x_n|^2\},$$

ce sup est atteint quand $x_1 = 1, x_2 = x_3 = \dots = x_n = 0$. Donc

$$\|A^*A\| = \|D\| = \lambda_1 = \max\{\lambda; \lambda \in \sigma_p(A^*A)\}.$$

Exemple 1.1. Soit

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Alors

$$A^*A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

a pour polynôme caractéristique $-\lambda^3 + 4\lambda^2 - 2\lambda$, donc les valeurs propres de A^*A :

$$\sigma_p(A^*A) = \{0, 2 - \sqrt{2}, 2 + \sqrt{2}\}, \quad \text{par conséquent } \|A\| = \sqrt{2 + \sqrt{2}}.$$

Exercice 1.1. Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ où $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

Montrer que pour tout $i = 1, \dots, n$ et $j = 1, \dots, n$

$$|a_{ij}| \leq \|A\| \leq \left(\sum_{i=1}^n \sum_{j=1}^n |a_{i,j}|^2 \right)^{\frac{1}{2}}.$$

Solution. Soit $\{e_i, i = 1, \dots, n\}$ la base canonique de \mathbb{K}^n . Alors pour tout $i, j = 1, \dots, n$,

$$|a_{ij}| = |\langle Ae_j, e_i \rangle| \leq \|Ae_j\| \|e_i\| \leq \|A\|.$$

Pour l'autre inégalité : Pour tout $x \in \mathbb{K}^n$ tel que $\|x\| = \left(\sum_j |x_j|^2 \right)^{\frac{1}{2}} = 1$:

$$\|Ax\| = \left(\sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right|^2 \right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}} \cdot \left(\sum_j |x_j|^2 \right)^{\frac{1}{2}} = \left(\sum_{i=1}^n \sum_{j=1}^n |a_{i,j}|^2 \right)^{\frac{1}{2}}.$$

D'où

$$\|A\| = \sup_{\|x\|=1} \|Ax\| \leq \left(\sum_{i=1}^n \sum_{j=1}^n |a_{i,j}|^2 \right)^{\frac{1}{2}}.$$

1.4 Analyse matricielle

(1) Soit $(A_k)_{k \geq 0}$ une suite de matrices dans $\mathcal{M}_n(\mathbb{K})$ et $A \in \mathcal{M}_n(\mathbb{K})$. On dira que $\lim_{k \rightarrow +\infty} A_k = A$ si $\lim_{k \rightarrow +\infty} \|A_k - A\| = 0$. D'après l'Exercice 1.1, il faut et il suffit, pour cela, que, pour chaque $i, j = 1, \dots, n$ fixés, la suite numérique des coefficients $(a_k)_{ij}$ de A_k tende vers le coefficient a_{ij} de A quand $k \rightarrow +\infty$. Donc faire de l'Analyse matricielle, revient à faire de l'Analyse ordinaire sur les nombres, coefficient par coefficient. On a alors,

$$\text{si } \lim_{k \rightarrow +\infty} A_k = A; \quad \lim_{k \rightarrow +\infty} B_k = B \text{ et si } \lambda \in \mathbb{K}, \text{ alors}$$

$$\lim_{k \rightarrow +\infty} (A_k + B_k) = A + B; \quad \lim_{k \rightarrow +\infty} (\lambda A_k) = \lambda A; \quad \lim_{k \rightarrow +\infty} A_k B_k = AB$$

(2) On dira que la série de matrices $\sum_{k \geq 0} A_k$ converge, et a pour somme A , si la suite des "sommations partielles" $S_N = \sum_{k=0}^N A_k$ tend vers A quand N tend vers $+\infty$. Pour cela, il faut et il suffit (**critère de Cauchy**) que

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N} \text{ tel que } \forall k' \geq k \geq N_\varepsilon, \quad \|A_k + A_{k+1} + \dots + A_{k'}\| \leq \varepsilon.$$

(3) On dira que la série de matrices $\sum_{k \geq 0} A_k$ est absolument (ou normalement) convergente si la série de nombres réels positifs $\sum_{k \geq 0} \|A_k\|$ converge. ce qui implique la série de matrices $\sum_{k \geq 0} A_k$ elle-même converge, et on a $\|\sum_{k \geq 0} A_k\| \leq \sum_{k \geq 0} \|A_k\|$.

1.4.1 Série géométrique de matrices et inversibilité

Le Théorème suivant est connu sous le nom "Lemme de Neumann"

Théorème 1.2. Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que $\|A\| < 1$, alors

- (1) la série $\sum_{k \geq 0} A^k$ est normalement convergente.
- (2) La matrice $I_d - A$ est inversible, et on a

$$(I_d - A)^{-1} = \sum_{k \geq 0} A^k.$$

Démonstration (1) Puisque $\|A^k\| \leq \|A\|^k$, qui est le terme général d'une série numérique géométrique de raison $\|A\| < 1$, donc convergente. Par le théorème de comparaison, la série à terme positifs $\sum_{k \geq 0} \|A^k\|$ converge.

(2) Soit $S_N = \sum_{k=0}^N A^k$. Alors

$$(I_d - A)S_N = S_N - AS_N = I_d - A^{N+1}.$$

Comme $\|A^{N+1}\| \leq \|A\|^{N+1} \rightarrow 0$, on déduit que

$$(I_d - A) \sum_{k \geq 0} A^k = I_d \quad \text{et donc} \quad (I_d - A)^{-1} = \sum_{k \geq 0} A^k.$$

Corollaire 1.1. Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que $\|I_d - A\| < 1$, alors A est inversible et on a

$$A^{-1} = \sum_{k \geq 0} (I_d - A)^k.$$

Démonstration Il suffit de remarquer que $A = I_d - (I_d - A)$ et appliquer le Théorème.

Remarque 1.1. Le Corollaire précédent, dit que toute matrice "suffisamment proche de I_d est inversible.

Plus généralement, on a

Corollaire 1.2. Soit $A_0 \in \mathcal{M}_n(\mathbb{K})$ inversible. Alors toute matrice $A \in \mathcal{M}_n(\mathbb{K})$ telle que $\|A\| < \frac{1}{\|A_0^{-1}\|}$, alors la matrice $A_0 - A$ est inversible.

Démonstration En effet, puisque $\|A_0^{-1}A\| \leq \|A_0^{-1}\| \|A\| < 1$, la matrice $I_d - A_0^{-1}A$ est inversible. Donc $A_0 - A = A_0(I_d - A_0^{-1}A)$ est inversible comme produit de deux matrices inversibles. De plus on a,

$$(A_0 - A)^{-1} = (I_d - A_0^{-1}A)^{-1}A_0^{-1}.$$

Remarque 1.2. Comme conséquence directe du Corollaire précédent : $GL_n(\mathbb{K})$ le groupe des matrices inversibles est ouvert dans $\mathcal{M}_n(\mathbb{K})$. Plus précisément, si $A_0 \in GL_n(\mathbb{K})$ alors la boule de centre A_0 et de rayon $\frac{1}{\|A_0^{-1}\|}$ est incluse dans $GL_n(\mathbb{K})$:

$$B(A_0, \frac{1}{\|A_0^{-1}\|}) := \{A \in \mathcal{M}_n(\mathbb{K}); \|A_0 - A\| < \frac{1}{\|A_0^{-1}\|}\} \subset GL_n(\mathbb{K}).$$

Exercice 1.2. Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que $\|A\| < 1$. Montrer que

$$\frac{1}{1 + \|A\|} \leq \|(I_d - A)^{-1}\| \leq \frac{1}{1 - \|A\|}.$$

Solution Puisque $I_d = (I_d - A)(I_d - A)^{-1}$, on a

$$1 = \|(I_d - A)(I_d - A)^{-1}\| \leq \|(I_d - A)\| \|(I_d - A)^{-1}\| \leq (1 + \|A\|) \|(I_d - A)^{-1}\|.$$

D'où la première inégalité

$$\frac{1}{1 + \|A\|} \leq \|(I_d - A)^{-1}\|.$$

D'autre part,

$$\|(I_d - A)^{-1}\| = \left\| \sum_{k \geq 0} A^k \right\| \leq \sum_{k \geq 0} \|A^k\| \leq \sum_{k \geq 0} \|A\|^k = \frac{1}{1 - \|A\|}.$$

D'où la seconde inégalité

$$\|(I_d - A)^{-1}\| \leq \frac{1}{1 - \|A\|}.$$

Remarque 1.3. Le Théorème précédent dit que, sous l'hypothèse $\|A\| < 1$, le système linéaire

$$x = Ax + b$$

admet une seule solution, donnée par

$$x = (I_d - A)^{-1}b = \sum_{k \geq 0} A^k b.$$

Cette série donne une approximation de la solution x .

Compléments :

(I) Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $\lambda \in \sigma_p(A)$ une valeur propre de A et $0 \neq v \in \mathbb{C}^n$, un vecteur propre associé à λ . Alors

$$Av = \lambda v \implies A^2v = \lambda^2v \implies \dots \implies A^kv = \lambda^kv; \quad \forall k \geq 1.$$

D'où pour tout $\lambda \in \sigma_p(A)$ et $k \geq 1$,

$$|\lambda| \leq \|A^k\|^{\frac{1}{k}}.$$

Donc

$$r(A) := \max\{|\lambda|; \lambda \in \sigma_p(A)\} \leq \|A^k\|^{\frac{1}{k}}, \quad \forall k \geq 1.$$

$r(A)$ est dit le **rayon spectral** de A .

On montre que $\lim_{k \rightarrow +\infty} \|A^k\|^{\frac{1}{k}}$ existe et on a la fameuse formule "Beurling-Gelfand"

$$r(A) = \lim_{k \rightarrow +\infty} \|A^k\|^{\frac{1}{k}}.$$

Avec ces notations, le Théorème 1.2, reste encore vrai si on remplace l'hypothèse $\|A\| < 1$ par $r(A) < 1$.

En général, on a, pour tout $A \in \mathcal{M}_n(\mathbb{C})$, $r(A) \leq \|A\|$. L'inégalité peut être stricte. Par exemple une matrice nilpotente $A \neq 0$, a pour spectre $\sigma_p(A) = \{0\}$, donc $r(A) = 0$, par contre $\|A\| \neq 0$.

Il existe toute une classe de matrices (par exemple : matrice auto-adjoint $A = A^*$, matrice normale $AA^* = A^*A$), telle que $r(A) = \|A\|$.

(II) Soit $A_0 \in \mathcal{M}_n(\mathbb{C})$ inversible. Posons $\|A_0^{-1}\| = \frac{1}{a}$.

Soit $A \in \mathcal{M}_n(\mathbb{C})$ telle que $\|A\| = b < a$. Alors par le Corollaire 1.2, la matrice $A_0 + A$ est inversible et $(A_0 + A)^{-1} = (I_d + A_0^{-1}A)^{-1}A_0^{-1}$. On a

$$\|(A_0 + A)^{-1} - A_0^{-1}\| \leq \frac{b^2}{a^2(a - b)}.$$

En effet,

$$\begin{aligned} (A_0 + A)^{-1} &= (I_d + A_0^{-1}A)^{-1}A_0^{-1} \\ &= \left(\sum_{k \geq 0} (-1)^k (A_0^{-1}A)^k \right) A_0^{-1} \\ &= A_0^{-1} - A_0^{-1}AA_0^{-1} + \left(\sum_{k \geq 2} (-1)^k (A_0^{-1}A)^k \right) A_0^{-1} \\ &= A_0^{-1} - A_0^{-1}AA_0^{-1} + (A_0^{-1}A)^2 \left(\sum_{k \geq 0} (-1)^k (A_0^{-1}A)^k \right) A_0^{-1} \end{aligned}$$

Donc

$$\begin{aligned} \|(A_0 + A)^{-1} - A_0^{-1}\| &\leq \|A_0^{-1}\|^2 \|A\|^2 \sum_{k \geq 0} \|A_0^{-1}\|^k \|A\|^k \|A_0^{-1}\| \\ &= \frac{b^2}{a^2} \frac{1}{1 - \frac{b}{a}} \frac{1}{a} \\ &= \frac{b^2}{a^2(a - b)}. \end{aligned}$$

Donc

$$\|(A_0 + A)^{-1} - A_0 + A_0^{-1}AA_0^{-1}\| \leq \frac{b^2}{a^2(a-b)} = b\varepsilon(b) \quad \text{avec} \quad \lim_{b \neq 0, b \rightarrow 0} \varepsilon(b) = 0.$$

Il résulte alors que l'application $\Phi : GL_n(\mathbb{C}) \rightarrow GL_n(\mathbb{C})$ définie par $\Phi(X) = X^{-1}$ est continue et même différentiable et sa différentielle en un point A_0 de $GL_n(\mathbb{C})$, étant l'application linéaire $D_{A_0}\Phi : \mathcal{M}_n(\mathbb{C}) \rightarrow \mathcal{M}_n(\mathbb{C})$

$$X \mapsto D_{A_0}\Phi(X) = -A_0^{-1}XA_0^{-1}.$$

Ceci généralise le fait que la dérivée de la fonction $x \mapsto \frac{1}{x}$, $x \neq 0$, vaut $-\frac{1}{x^2}$ en tout point $x_0 \neq 0$.

1.4.2 Exponentielle d'une matrice

Si $A \in \mathcal{M}_n(\mathbb{K})$, on pose

$$e^A = \exp(A) = \sum_{k \geq 0} \frac{1}{k!} A^k = I_d + A + \frac{1}{2}A^2 + \dots + \frac{1}{k!}A^k + \dots$$

cette série est normalement convergente, puisque que

$$\left\| \frac{1}{k!} A^k \right\| \leq \frac{1}{k!} \|A\|^k.$$

Définition 1.2. $\exp(A)$ est dit l'exponentielle de A .

Remarque 1.4. Conséquence directe de la définition, on a

$$\exp(0) = I_d,$$

et

$$\|\exp(A)\| \leq e^{\|A\|} \quad \forall A \in \mathcal{M}_n(\mathbb{K}).$$

Exemple 1.2. Si

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{alors} \quad A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{et donc} \quad A^k = 0, \quad \forall k \geq 2.$$

Par conséquent

$$\exp(A) = I_d + A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Exemple 1.3. Si

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{alors} \quad A^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_d \quad \text{et donc} \quad A^{2k} = I_d \quad \text{et} \quad A^{2k+1} = A \quad \forall k \geq 0.$$

Donc

$$\begin{aligned}
\exp(A) &= \left(I_d + \frac{1}{2}A^2 + \cdots + \frac{1}{(2k)!}A^{2k} + \cdots \right) + \left(A + \frac{1}{3!}A^3 + \cdots + \frac{1}{(2k+1)!}A^{2k+1} + \cdots \right) \\
&= \left(1 + \frac{1}{2} + \cdots + \frac{1}{(2k)!} + \cdots \right) I_d + \left(1 + \frac{1}{3!} + \cdots + \frac{1}{(2k+1)!} + \cdots \right) A \\
&= \cosh(1)I_d + \sinh(1)A.
\end{aligned}$$

Donc

$$\exp(A) = \begin{pmatrix} \cosh(1) & \sinh(1) \\ \sinh(1) & \cosh(1) \end{pmatrix}.$$

Proposition 1.1.

1. Soit $A, B \in \mathcal{M}_n(\mathbb{K})$. Supposons que A et B commutent ($AB = BA$), alors

$$\exp(A + B) = \exp(A) \exp(B).$$

2. Soit $A, S \in \mathcal{M}_n(\mathbb{K})$. Supposons que S inversible, alors

$$\exp(S^{-1} A S) = S^{-1} \exp(A) S.$$

Corollaire 1.3.

1. Pour tout $A \in \mathcal{M}_n(\mathbb{K})$, $\exp(A)$ est inversible et on a, $(\exp(A))^{-1} = \exp(-A)$.

2. Pour tout $A, S \in \mathcal{M}_n(\mathbb{K})$ avec S inversible on a,

$$\exp(A) = S \exp(S^{-1} A S) S^{-1}.$$

Remarque 1.5. Sans l'hypothèse A et B commutent, le résultat de la Proposition 1.1 (1) est faux.

Contre-exemple, Si

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

alors

$$\exp(A) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad \exp(B) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{donc} \quad \exp(A) \exp(B) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

D'autre part,

$$A + B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{donc} \quad \exp(A + B) = \begin{pmatrix} \cosh(1) & \sinh(1) \\ \sinh(1) & \cosh(1) \end{pmatrix} \neq \exp(A) \exp(B).$$

Notations : Notons par $\mathbb{K}[X]$ l'espace des polynômes à coefficients dans $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , et $\mathbb{K}_{n-1}[X]$ le sous espace des polynômes de degré $\leq n - 1$.

Pour $A \in \mathcal{M}_n(\mathbb{K})$, soit

$$\mathbb{K}[A] = \{P(A); P \in \mathbb{K}[X]\} \quad \text{et} \quad \mathbb{K}_{n-1}[A] = \{P(A); P \in \mathbb{K}_{n-1}[X]\}.$$

Alors

$$\mathbb{K}[A] = \mathbb{K}_{n-1}[A].$$

En effet, Soit χ_A le polynôme caractéristique de A , on sait que le degré de $\chi_A \leq n$ et par le Théorème de Cayley-Hamilton, $\chi_A(A) = 0$.

Soit $P \in \mathbb{K}[X]$ de degré $\geq n$. Alors $P = Q\chi_A + R$ avec degré de $R < n$ (division euclidienne). Il résulte alors $P(A) = R(A) \in \mathbb{K}_{n-1}[A]$. D'où $\mathbb{K}[A] \subseteq \mathbb{K}_{n-1}[A]$ et donc $\mathbb{K}[A] = \mathbb{K}_{n-1}[A]$.

Puisque pour tout $N \geq 0$, $S_N = \sum_{k=0}^N \frac{1}{k!} A^k \in \mathbb{K}[A] = \mathbb{K}_{n-1}[A]$, la limite de la suite S_N , on a $\exp(A) \in \mathbb{K}_{n-1}[A]$ ($\mathbb{K}_{n-1}[A]$ est fermé dans $\mathcal{M}_n(\mathbb{K})$). D'où le résultat suivant

Théorème 1.3. *Pour tout $A \in \mathcal{M}_n(\mathbb{K})$ il existe $P \in \mathbb{K}_{n-1}[X]$ tel que $\exp(A) = P(A)$.*

Remarque 1.6. *Puisque $\exp(A)$ est toujours inversible, $P(0) \neq 0$.*

Proposition 1.2. *Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors*

- (i) $\sigma_p(\exp(A)) = \exp(\sigma_p(A))$.
- (ii) Si $A = D + N$ avec $DN = ND$, la décomposition de Dunford de A , alors $\exp(A) = \exp(D) + \exp(D)(\exp(N) - I_d)$ est la décomposition de Dunford de $\exp(A)$.
- (iii) Si χ_A est scindé alors

A est diagonalisable si et seulement si $\exp(A)$ est diagonalisable.

Démonstration. Rappelons que toute matrice A est trigonalisable sur \mathbb{C} . Donc il existe $S \in \mathcal{M}_n(\mathbb{C})$ inversible telle que si $\sigma_p(A) = \{\lambda_1, \dots, \lambda_n\}$, alors

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & \star & \dots & \star \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

Donc $\exp(A)$ est semblable à

$$\exp(S^{-1}AS) = \begin{pmatrix} e^{\lambda_1} & \star & \dots & \star \\ 0 & e^{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \dots & 0 & e^{\lambda_n} \end{pmatrix}$$

D'où $\sigma_p(\exp(A)) = \{e^{\lambda_1}, \dots, e^{\lambda_n}\} = \exp(\sigma_p(A))$ et donc (i) est démontré.

(ii) est facile à vérifier.

(iii) Il est clair que si A est diagonalisable alors $\exp(A)$ est diagonalisable. Supposons que $\exp(A)$ est diagonalisable et soit $A = D + N$ avec $DN = ND$, la décomposition de Dunford de A . Alors $\exp(A) = \exp(D)\exp(N)$. Donc $\exp(N) = \exp(-D)\exp(A)$ est diagonalisable (comme produit de deux matrices diagonalisables). Puisque $\sigma_p(N) = \{0\}$, d'après (i), $\sigma_p(\exp(N)) = \{1\}$. D'où $\exp(N) = I_d$ (puisque $\exp(N)$ est diagonalisable). D'autre part, N nilpotente implique que $\exp(N) = I_d + N + \frac{N^2}{2} + \dots + \frac{N^{d-1}}{(d-1)!}$. D'où $N + \frac{N^2}{2} + \dots + \frac{N^{d-1}}{(d-1)!} = 0$. Il résulte alors que le polynôme $P(X) = X + \frac{X^2}{2} + \dots + \frac{X^{d-1}}{(d-1)!}$ est annulateur pour N et donc divisible par son polynôme minimal X^d . Par conséquent $\chi_N(X) = X$ et donc $N = 0$ et A est bien diagonalisable.

Notations : Pour $M \in \mathcal{M}_n(\mathbb{K})$ notons par $\lambda(M)$ la valeur propre λ de M et $\det(M)$ désigne le déterminant de la matrice $M = (a_{ij})$ et $\text{tr}(M) = a_{11} + a_{22} + \dots + a_{nn}$ la trace de M .

Proposition 1.3. *Soit $A, B \in \mathcal{M}_n(\mathbb{K})$ (ne commutent pas nécessairement), alors on a*

$$\det(\exp(A + B)) = e^{\text{tr}(A) + \text{tr}(B)} = \det(\exp(A) \exp(B)).$$

En particulier (pour $B = 0$),

$$\det(\exp(A)) = \exp(\text{tr}(A))$$

Démonstration On a

$$\begin{aligned} \det(\exp(A + B)) &= \prod_{i=1}^n \lambda_i(\exp(A + B)) = \prod_{i=1}^n e^{\lambda_i(A+B)} = e^{\text{tr}(A+B)} = \\ &= e^{\text{tr}(A) + \text{tr}(B)} = e^{\text{tr}(A)} e^{\text{tr}(B)} = \det e^A \det e^B = \det(e^A e^B). \end{aligned}$$

Le Théorème suivant complète le Théorème 1.2 et donne une méthode pour calculer $\exp(A)$, sans savoir si A est diagonalisable ou non.

Théorème 1.4. *Si $A \in \mathcal{M}_n(\mathbb{K})$, alors*

$$(1) \exp(tA) = \alpha_0 I_d + \alpha_1 tA + \dots + \alpha_{n-1} t^{n-1} A^{n-1}.$$

Les α_i sont des fonctions de t .

$$(2) \text{ Posons } P(\lambda) = \alpha_0 + \alpha_1 \lambda + \dots + \alpha_{n-1} \lambda^{n-1} - 1.$$

Si λ_i est une valeur propre de tA , alors $e^{\lambda_i} = P(\lambda_i)$.

De plus, si λ_i est de multiplicité $k_i > 1$, alors

$$e^{\lambda_i} = \frac{dP}{d\lambda}(\lambda_i) = \frac{d^2 P}{d\lambda^2}(\lambda_i) = \dots = \frac{d^{k_i-1} P}{d\lambda^{k_i-1}}(\lambda_i)$$

Remarque 1.7. (1) $\lambda \in \sigma_p(A) \iff t\lambda \in \sigma_p(tA)$.

(2) Les équations formées par les dérivées, permettent de calculer α_i et donc $\exp(A)$.

(3) Pour des exemples d'application de ce Théorème, voir le Poly, Equ. Diff.

Calcul pratique de l'exponentielle d'une matrice.

(1) si $D = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ alors

$$\exp(D) = \text{diag}\{e^{\lambda_1}, \dots, e^{\lambda_n}\}.$$

(2) Si A est diagonalisable alors $A = S^{-1} D S$ avec D diagonale, alors

$$\exp(A) = S \exp(D) S^{-1}.$$

(3) Si A est nilpotente ($A^p = 0$), alors

$$\exp(A) = I_d + A + \frac{1}{2} A^2 + \dots + \frac{1}{(p-1)!} A^{p-1}$$

(4) Si $A = \lambda I_d + N$ avec N matrice nilpotente, alors

$$\exp(A) = e^\lambda \exp(N).$$

En particulier si A est un bloc de Jordan i.e. $A = \lambda I_d + N$.

(5) Si $A = D + N$ la décomposition de Dunford, alors

$$\exp(A) = \exp(D) \exp(N).$$

(6) Par trigonalisation. Toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, il existe $S \in \mathcal{M}_n(\mathbb{C})$ inversible telle que $S^{-1}AS$ soit diagonale par blocs :

$S^{-1}AS = \text{diag}\{B_1, \dots, B_p\}$ où chaque bloc B_i est une matrice triangulaire supérieure à diagonale constante. Alors $\exp(B_i)$ se calcule par (4). Finalement, $\exp(A) = S \text{diag}\{\exp(B_1), \dots, \exp(B_p)\} S^{-1}$.

Exemples

(1) Exemple d'application du Théorème 1.3 :

Soit la matrice

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

Calcul de $\exp(A)$? Par le Théorème 1.3, on a $\exp(tA) = \alpha_0 I_d + \alpha_1 tA + \alpha_2 t^2 A^2$. Donc

$$\exp(tA) = \begin{pmatrix} \alpha_0 & \alpha_1 t & \alpha_2 t^2 \\ 0 & \alpha_0 - \alpha_2 t^2 & \alpha_1 + 2\alpha_2 t^2 \\ 0 & -\alpha_1 t - 2\alpha_2 t^2 & \alpha_0 + 2\alpha_1 t + 3\alpha_2 t^2 \end{pmatrix}.$$

On a $P(\lambda) = \alpha_0 + \alpha_1 \lambda + \alpha_2 \lambda^2$. D'autre part, $\sigma_P(A) = \{0, 1\}$ donc $\sigma_P(tA) = \{0, t\}$. Puisque $\lambda_1 = 0$ est de multiplicité un et λ_2 est de multiplicité deux, il résulte alors du Théorème 1.3, que $e^0 = P(0)$ et $e^t = P(t)$, $e^t = P'(t) = \alpha_1 + 2\alpha_2 t$. Ces équations donnent le système d'équation dont les inconnues sont les α_i :

$$\begin{cases} \alpha_0 & = e^0 = 1 \\ \alpha_1 + 2\alpha_2 t & = e^t \\ \alpha_0 + \alpha_1 t + \alpha_2 t^2 & = e^t \end{cases}$$

dont la solution est

$$\alpha_0 = 1; \quad \alpha_1 = \frac{-2 + 2e^t - tet}{t}; \quad \alpha_2 = \frac{1 - e^t + te^t}{t^2}.$$

En utilisant la matrice de $\exp(tA)$ et en simplifiant, on trouve

$$\exp(tA) = \begin{pmatrix} 1 & -2 + 2e^t - tet & 1 - e^t + te^t \\ 0 & e^t - te^t & te^t \\ 0 & -te^t & e^t + te^t \end{pmatrix}.$$

(2) Exemple d'une matrice diagonalisable. Soit

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

Le polynôme caractéristique de A est : $\chi_A(\lambda) = -\lambda(1 - \lambda)(2 - \lambda)$.

Donc $\sigma_P(A) = \{0, 1, 2\}$. On a $A = SDS^{-1}$ où

$$S = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}; \quad S^{-1} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & 0 \\ 2 & -1 & -1 \\ -\frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \quad \text{et} \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Donc $\exp(A) = S \exp(D) S^{-1}$ et

$$\begin{aligned} \exp(tA) &= \begin{pmatrix} 1 & 1 & 1 \\ 3 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^{2t} \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & 0 \\ 2 & -1 & -1 \\ -\frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{2} + 2e^t - \frac{1}{2}e^{2t} & \frac{1}{2} - e^t + \frac{1}{2}e^{2t} & -e^t + e^{2t} \\ -\frac{3}{2} + 2e^t - \frac{1}{2}e^{2t} & \frac{3}{2} - e^t + \frac{1}{2}e^{2t} & -e^t + e^{2t} \\ \frac{1}{2} - \frac{1}{2}e^{2t} & -\frac{1}{2} + \frac{1}{2}e^{2t} & e^{2t} \end{pmatrix}. \end{aligned}$$

(3) Soit $A \in \mathcal{M}_n(\mathbb{K})$ avec $n = 4$, sous forme de bloc de Jordan.

$$A = \begin{pmatrix} a & 1 & 0 & 0 \\ 0 & a & 1 & 0 \\ 0 & 0 & a & 1 \\ 0 & 0 & 0 & a \end{pmatrix} = aI_d + J,$$

avec $J^k = 0$, $\forall k > 3$. Comme I_d et J commutent, on a pour tout $k \geq 0$,

$$A^k = (aI_d + J)^k = a^k I_d + k a^{k-1} J + \frac{k(k-1)}{2} a^{k-2} J^2 + \frac{k(k-1)(k-2)}{2 \cdot 3} a^{k-3} J^3$$

D'où, après calcul, on obtient : $\exp(A) = e^a I_d + e^a J + \frac{1}{2} e^a J^2 + \frac{1}{6} e^a J^3$ et donc

$$\exp(A) = e^a \begin{pmatrix} 1 & 1 & \frac{1}{2} & \frac{1}{6} \\ 0 & 1 & 1 & \frac{1}{2} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

1.5 Dérivée d'une fonction à valeurs matricielles

Soit $A(\cdot) : I \rightarrow \mathcal{M}_n(\mathbb{K})$ où I est intervalle ouvert de \mathbb{R} . Alors, pour $t \in I$, on a $A(t) = (a_{ij}(t))$.

Définition 1.3. On dira que la fonction $A(\cdot)$ a pour dérivée $C = (c_{ij})$ au point $t_0 \in I$, si

$$\lim_{h \rightarrow 0} \left\| \frac{A(t_0 + h) - A(t_0)}{h} - C \right\| = 0.$$

Dans ce cas, on écrit $C = A'(t_0) = \frac{dA}{dt}(t_0)$.

Propriétés de la dérivée. Si $t \mapsto A(t)$ et $t \mapsto B(t)$ sont dérivables et si $\lambda \in \mathbb{K}$, on a

$$\frac{d}{dt}(A + B) = \frac{dA}{dt} + \frac{dB}{dt}; \quad \frac{d(\lambda A)}{dt} = \lambda \frac{dA}{dt} \quad \text{et} \quad \frac{d(A(t)B(t))}{dt} = \frac{dA}{dt}B(t) + A(t)\frac{dB}{dt}.$$

Si $x = x(t)$ une fonction à valeurs vectorielles dans \mathbb{K}^n , dérivable, alors on a

$$\frac{d}{dt}(A(t)x(t)) = \frac{dA}{dt}x(t) + A(t)\frac{dx}{dt}.$$

Proposition 1.4. Soit $A \in \mathcal{M}_n(\mathbb{K})$ fixée. La fonction $t \mapsto \exp(tA)$ est dérivable sur \mathbb{R} et on a

$$\frac{d}{dt}(\exp(tA)) = A \exp(tA) = \exp(tA) A.$$

Proposition 1.5. Si $A(t) \in \mathcal{M}_n(\mathbb{C})$ dérivable dans \mathbb{R} telle que

$$(i) \quad A(0) = 1$$

$$(2) \quad \forall t_1, t_2 \in \mathbb{R}, \quad A(t_1 + t_2) = A(t_1)A(t_2).$$

Alors il existe $A \in \mathcal{M}_n(\mathbb{C})$ fixe (à savoir $A = A'(0)$) telle que

$$A(t) = \exp(tA) \quad \forall t \in \mathbb{R}.$$

Proposition 1.6. Si $A(t) \in \mathcal{M}_n(\mathbb{K})$ dérivable et $A(t)A'(t) = A'(t)A(t)$, alors

$$\frac{d}{dt}(\exp(A(t))) = A'(t) \exp(A(t)).$$

Remarque 1.8. Sans l'hypothèse de commutation le résultat de la proposition précédente est faux. En effet, si

$$A(t) = \begin{pmatrix} 1 & t \\ 0 & 0 \end{pmatrix},$$

alors pour tout $n \geq 1$, $A(t)^n = A(t)$ donc

$$\exp(A(t)) = I_d + (e-1)A(t) = \begin{pmatrix} e & (e-1)t \\ 0 & 1 \end{pmatrix} \quad \text{donc} \quad \frac{d}{dt}(\exp(A(t))) = \begin{pmatrix} 0 & e-1 \\ 0 & 0 \end{pmatrix}$$

par contre le produit $A'(t) \exp(A(t))$ est la matrice nulle.

1.6 Systèmes différentiels linéaires sans second membre

Le problème : Pour $1 \leq i \leq n$, $1 \leq j \leq n$, soit n^2 nombres complexes, a_{ij} donnés. On cherche à déterminer n fonctions numériques inconnues $y_1(t), y_2(t), \dots, y_n(t)$ dérivables sur \mathbb{R} , à valeurs dans \mathbb{C} , et telles que, pour tout $t \in \mathbb{R}$,

$$\begin{cases} \frac{dy_1}{dt} = a_{11}y_1(t) + a_{12}y_2(t) + \dots + a_{1n}y_n(t) \\ \frac{dy_2}{dt} = a_{21}y_1(t) + a_{22}y_2(t) + \dots + a_{2n}y_n(t) \\ \dots\dots\dots \\ \frac{dy_n}{dt} = a_{n1}y_1(t) + a_{n2}y_2(t) + \dots + a_{nn}y_n(t) \end{cases} \quad (1.1)$$

Notons A la matrice (a_{ij}) et posons

$$y(t) = (y_1(t), y_2(t), \dots, y_n(t))$$

Alors le système d'équations (1,1) s'écrit :

$$\frac{dy}{dt}(t) = Ay(t) \quad (1.2)$$

et il s'agit de trouver les solutions de (1,1) c-à-d les fonction $y(\cdot) : \mathbb{R} \rightarrow \mathbb{C}^n$ dérivables dans \mathbb{R} qui satisfont l'égalité (1.2). On pourra s'intéresser à :

(1) la solution générale de (1.2), qui, on le verra, dépend de n constantes dans \mathbb{C} arbitraires, ou chercher

(2) la solution particulière $y(t)$ passant par des conditions initiales précises, par exemple telle que $y(0) = y_0$ où y_0 est un vecteur de \mathbb{C}^n donné.

Remarque 1.9. *Rappel : Pour $n = 1$, l'équation $\frac{dy}{dt}(t) = ay(t)$ admet, pour tout y_0 réel donné, une solution $y(t)$ et une seule telle que $y(0) = y_0$, à savoir $y(t) = y_0 e^{at} = e^{at} y_0$.*

Théorème 1.5. *Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $y_0 \in \mathbb{C}^n$ donné. Alors il existe une solution et une seule de*

$$\frac{dy}{dt}(t) = Ay(t), \quad \text{telle que } y(0) = y_0,$$

donnée par

$$y(t) = \exp(tA) y_0.$$

Remarque 1.10. *Soit $t_0 \in \mathbb{R}$ et $v_0 \in \mathbb{C}^n$. Il existe une solution et une seule de (1.2) $\frac{dy}{dt}(t) = Ay(t)$, telle que $y(t_0) = v_0$, à savoir $y(t) = \exp((t - t_0)A) v_0$.*

Exemple 1.4. *Soit le système d'équations*

$$\begin{cases} 9 \frac{dy_1}{dt} &= 4y_1 + 5y_2 - 2y_3 \\ 9 \frac{dy_2}{dt} &= 2y_1 - 2y_2 + 8y_3 \\ 9 \frac{dy_3}{dt} &= -5y_1 - 4y_2 - 2y_3 \end{cases}$$

telles que $y_1(0) = 1; y_2(0) = y_3(0) = 0$.

Solution : La matrice

$$A = \frac{1}{9} \begin{pmatrix} 4 & 5 & -2 \\ 2 & -2 & 8 \\ -5 & -4 & -2 \end{pmatrix}.$$

a pour polynôme caractéristique $\chi_A(\lambda) = \lambda^3$. Donc $A^3 = 0$ (par le Théorème de Cayley-Hamilton), et $\exp(tA) = I_d + tA + \frac{t^2}{2}A^2$. Or

$$A^2 = \frac{1}{9} \begin{pmatrix} 4 & 2 & 4 \\ -4 & -2 & -4 \\ -2 & -1 & -2 \end{pmatrix}.$$

d'où

$$\exp(tA) = \begin{pmatrix} 1 + \frac{4t}{9} + \frac{2t^2}{9} & \frac{5t}{9} + \frac{t^2}{9} & -\frac{2t}{9} + \frac{2t^2}{9} \\ \frac{2t}{9} - \frac{2t^2}{9} & -1 - \frac{2t}{9} - \frac{t^2}{9} & \frac{8t}{9} - \frac{2t^2}{9} \\ -\frac{5t}{9} - \frac{t^2}{9} & -\frac{4t}{9} - \frac{t^2}{18} & 1 - \frac{2t}{9} - \frac{t^2}{9} \end{pmatrix}$$

Par le Théorème 1.4 la solution est $y(t) = \exp(A)y_0$ avec $y_0 = (1, 0, 0)$. Donc les solutions sont :

$$y_1(t) = 1 + \frac{4t}{9} + \frac{2t^2}{9}; \quad y_2(t) = \frac{2t}{9} - \frac{2t^2}{9} \quad \text{et} \quad y_3(t) = -\frac{5t}{9} - \frac{t^2}{9}.$$

Remarque 1.11. *L'ensemble des solutions de l'équation $\frac{dy}{dt}(t) = Ay(t)$ est un espace vectoriel et l'application linéaire $y_0 \mapsto \exp(tA)y_0$, est un isomorphisme de \mathbb{C}^n sur l'espace des solutions.*

Théorème 1.6. *Soit $A \in \mathcal{M}_n(\mathbb{C})$. Alors l'espace vectoriel des solutions de*

$$\frac{dy}{dt}(t) = Ay(t)$$

est de dimension n sur \mathbb{C} .

Par conséquent, si on connaît n solutions linéairement indépendantes

$$y^1(t), y^2(t), \dots, y^n(t),$$

la solution générale est donnée par

$$y(t) = C_1 y^1(t) + C_2 y^2(t) + \dots + C_n y^n(t),$$

où C_1, C_2, \dots, C_n sont des constantes complexes arbitraires.

Il n'est pas toujours facile de calculer l'exponentielle d'une matrice. Le théorème suivant fournit un moyen d'en trouver n solutions linéairement indépendant et donc la solution générale.

Théorème 1.7. *Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $v \in \mathbb{C}^n$ un vecteur propre de A associé à la valeur propre λ (c-à-d, $v \neq 0$ et $Av = \lambda v$).*

Alors la fonction vectorielle

$$y(t) = e^{\lambda t}v$$

est solution de

$$\frac{dy}{dt}(t) = Ay(t).$$

Démonstration

$$\frac{dy}{dt}(t) = \lambda e^{\lambda t}v = e^{\lambda t}\lambda v = e^{\lambda t}Av = A(e^{\lambda t}v) = Ay(t).$$

Corollaire 1.4. *Soit $A \in \mathcal{M}_n(\mathbb{C})$ diagonalisable et v_1, v_2, \dots, v_n une base formée par les vecteurs propres de A associés respectivement aux valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$ (pas forcément deux à deux distinctes).*

Alors la solution générale de

$$\frac{dy}{dt}(t) = Ay(t).$$

est donnée par

$$y(t) = C_1 e^{\lambda_1 t} v_1 + C_2 e^{\lambda_2 t} v_2 + \dots + C_n e^{\lambda_n t} v_n,$$

où C_1, C_2, \dots, C_n sont des constantes arbitraires.

Exemple 1.5. Soit le système d'équations

$$\begin{cases} \frac{dy_1}{dt} = 2y_1 + y_2 + y_3 \\ \frac{dy_2}{dt} = y_1 + 2y_2 + y_3 \\ \frac{dy_3}{dt} = y_1 + y_2 + 2y_3 \end{cases}$$

Solution : La matrice

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

a deux valeurs propres $\lambda_1 = 4$ et $\lambda_2 = 1$. Pour $\lambda_1 = 4$, un vecteur propre $v_1 = (1, 1, 1)$ et pour $\lambda_2 = 1$, deux vecteurs propres $v_2 = (1, -1, 0)$, $v_3 = (1, 1, -2)$. Alors la solution vectorielle est

$$y(t) = C_1 e^{4t} v_1 + C_2 e^t v_2 + C_3 e^t v_3$$

Donc

$$y_1(t) = C_1 e^{4t} + (C_2 + C_3) e^t; \quad y_2(t) = C_1 e^{4t} - (C_2 - C_3) e^t; \quad y_3(t) = C_1 e^{4t} - 2C_3 e^t.$$

Remarque 1.12. Lorsque A n'est pas diagonalisable, pour trouver n fonctions solutions linéairement indépendantes, on utilise le Théorème suivant

Théorème 1.8. Soit $A \in \mathcal{M}_n(\mathbb{C})$ et λ une valeur propre de A d'ordre d entant que racine du polynôme caractéristique de A . Alors le système,

$$\frac{dy}{dt}(t) = Ay(t)$$

admet d solutions linéairement indépendantes de la forme

$$y(t) = \begin{cases} y_1(t) &= P_1(t) e^{\lambda t} \\ y_2(t) &= P_2(t) e^{\lambda t} \\ \dots\dots\dots \\ y_n(t) &= P_n(t) e^{\lambda t} \end{cases}$$

où P_1, P_2, \dots, P_n sont des polynômes de degré $\leq d-1$ qui dépendent de d constantes pour les calculer on reporte $y(t)$ dans l'équation $\frac{dy}{dt}(t) = Ay(t)$.

Exemple 1.6. Soit le système d'équations

$$\begin{cases} \frac{dy_1}{dt} = y_1 + y_2 + y_3 \\ \frac{dy_2}{dt} = -y_1 + y_2 - y_3 \\ \frac{dy_3}{dt} = y_1 + 2y_3 \end{cases}$$

Solution : La matrice

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

a pour polynôme caractéristique $(1 - \lambda)^2(2 - \lambda)$.

La valeur propre $\lambda = 2$ est simple et admet $(0, 1, -1)$ comme vecteur propre. Le Théorème 1.6, fournit une première solution particulière $y_1(t) = 0$, $y_2(t) = e^{2t}$, $y_3(t) = -e^{2t}$.

La valeur propre $\lambda = 1$ est double et admet seulement $(1, 1, -1)$ comme vecteur propre.

La matrice A n'est pas diagonalisable. On va utiliser le Théorème précédent. Donc on cherche des solutions de la forme

$$y_1(t) = (at + b)e^t, \quad y_2(t) = (ct + d)e^t, \quad y_3(t) = (et + f)e^t$$

Pour calculer 6 constantes, on utilise l'équation $\frac{dy}{dt}(t) = Ay(t)$, on obtient le système suivant

$$\begin{cases} at + a + b &= (a + c + e)t + (b + d + f) \\ ct + c + d &= (-a + c - e)t + (-b + d - f) \\ et + e + f &= (a + 2e)t + (b + 2f) \end{cases}$$

d'où, en identifiant les coefficients, on trouve

$$c + e = 0; \quad a = d + f; \quad a + e = 0; \quad c = -b - f; \quad a + e = 0; \quad e = b + f.$$

En posant $C_1 = -e$ et $C_2 = -f$, on exprime les solutions du système par :

$$y_1(t) = C_1(t - 1)e^t + C_2e^t; \quad y_2(t) = C_1(t + 1)e^t + C_2e^t; \quad y_3(t) = -C_1te^t - C_2e^t.$$

Enfin pour trouver la solution générale, on ajoute les solutions fournies par la $\lambda = 2$, on obtient la solution générale du système

$$\begin{cases} y_1(t) &= [C_1(t - 1) + C_2]e^t \\ y_2(t) &= [C_1(t + 1) + C_2]e^t + C_3e^{2t} \\ y_3(t) &= -[C_1t + C_2]e^t - C_3e^{2t} \end{cases}$$

1.7 Systèmes différentiels linéaires avec second membre

Il s'agit d'étudier les systèmes différentiels de cette forme

$$\begin{cases} \frac{dy_1}{dt} = a_{11}y_1(t) + a_{12}y_2(t) + \cdots + a_{1n}y_n(t) + b_1(t) \\ \frac{dy_2}{dt} = a_{21}y_1(t) + a_{22}y_2(t) + \cdots + a_{2n}y_n(t) + b_2(t) \\ \dots\dots\dots \\ \frac{dy_n}{dt} = a_{n1}y_1(t) + a_{n2}y_2(t) + \cdots + a_{nn}y_n(t) + b_n(t) \end{cases} \quad (1.3)$$

Notons A la matrice (a_{ij}) et posons

$$Y(t) = (y_1(t), y_2(t), \dots, y_n(t)) \quad \text{et} \quad B(t) = (b_1(t), b_2(t), \dots, b_n(t))$$

Alors le système d'équations (1.3) s'écrit :

$$\frac{dY}{dt}(t) = AY(t) + B(t) \quad (1.4)$$

Pour résoudre ce système, trois étapes :

(1) on cherche les solutions Y_s de l'équation dite homogène (ou sans second membre) $\frac{dY}{dt}(t) = AY(t)$. Pour cela on utilise le paragraphe précédent.

(2) On cherche une solution particulière Y_p de l'équation générale $\frac{dY}{dt}(t) = AY(t) + B(t)$.

(3) La solution générale de $\frac{dY}{dt}(t) = AY(t) + B(t)$ est donnée par

$$Y(t) = Y_s(t) + Y_p(t).$$

Recherche de solution particulière Y_p :

(a) Superposition de second membre. Si $B(t) = B_1(t) + B_2(t)$, alors la solution de $\frac{dY}{dt}(t) = AY(t) + B(t)$ s'obtient en ajoutant une solution de $\frac{dY}{dt}(t) = AY(t) + B_1(t)$ à une solution de $\frac{dY}{dt}(t) = AY(t) + B_2(t)$.

(b) Méthode de variation de la constante. Supposons connues n solutions $Y_1(t), Y_2(t), \dots, Y_n(t)$ de $\frac{dY}{dt}(t) = AY(t)$, donc la solution générale est

$$Y(t) = C_1 Y_1(t) + C_2 Y_2(t) + \dots + C_n Y_n(t)$$

où C_1, C_2, \dots, C_n sont des constantes.

On suppose que les C_1, C_2, \dots, C_n soient des fonctions numériques et on cherche une solution de la forme

$$Y_p(t) = C_1(t)Y_1(t) + C_2(t)Y_2(t) + \dots + C_n(t)Y_n(t)$$

de l'équation $\frac{dY}{dt}(t) = AY(t) + B(t)$. On obtient alors un système d'équation dont les inconnues sont $C'_1(t), C'_2(t), \dots, C'_n(t)$, d'où l'on déduit explicitement $C_1(t), C_2(t), \dots, C_n(t)$. Alors une solution particulière sera donnée par. $Y_p(t) = C_1(t)Y_1(t) + C_2(t)Y_2(t) + \dots + C_n(t)Y_n(t)$.

Exemple 1.7. Soit le système d'équations

$$(E) \quad \begin{cases} \frac{dy_1}{dt} = y_1 + y_2 + y_3 + e^t \\ \frac{dy_2}{dt} = -y_1 + y_2 - y_3 - 2e^t \\ \frac{dy_3}{dt} = y_1 + 2y_3 \end{cases}$$

On va chercher la solution générale du système.

Solution : (1) L'équation homogène a été traitée dans l'exemple 1.6, précédent. La solution générale est donnée par

$$(*) \quad Y_s(t) = \begin{cases} y_1(t) = [C_1(t-1) + C_2]e^t \\ y_2(t) = [C_1(t+1) + C_2]e^t + C_3e^{2t} \\ y_3(t) = -[C_1t + C_2]e^t - C_3e^{2t} \end{cases}$$

(2) Recherche d'une solution particulière par la méthode de la variation de la constante. Faisons varier les constantes, c-à-d, supposons que C_1, C_2, C_3 sont des fonctions en t qu'on doit déterminer. Alors (*) est solution de (E) si et seulement si $C'_1(t), C'_2(t), C'_3(t)$ sont solutions du système (ici $B(t) = (e^t, -2e^t, 0)$) :

$$\begin{cases} ((t-1)C'_1(t) + C'_2(t))e^t = e^t \\ ((t+1)C'_1(t) + C'_2(t))e^t + C'_3(t)e^{2t} = -2e^t \\ -[tC'_1(t) + C'_2(t)]e^t - C'_3(t)e^{2t} = 0 \end{cases}$$

Ici $B(t) = (e^t, -2e^t, 0)$.

On trouve $C_1'(t) = -2$; $C_2'(t) = 2t - 1$; $C_3'(t) = e^{-1}$
 et donc $C_1(t) = -2t$; $C_2(t) = t^2 - t$; $C_3(t) = -e^{-1}$

Pour trouver une solution particulière, on remplace les constantes C_1, C_2, C_3 par les fonctions $C_1(t) = -2t$; $C_2(t) = t^2 - t$; $C_3(t) = -e^{-1}$ dans (*). On trouve

$$Y_p(t) = \begin{cases} y_1(t) &= [-2t(t-1) + t^2 - t]e^t = (-t^2 + t)e^t \\ y_2(t) &= [-2t(t+1) + t^2 - t]e^t - e^{-t}e^{2t} = -(t^2 + 3t + 1)e^t \\ y_3(t) &= -[-2t^2 + t^2 - t]e^t + e^{-t}e^{2t} = (t^2 + t + 1)e^t \end{cases}$$

(3) Finalement la solution générale du système est :

$$Y(t) = Y_s(t) + Y_p(t) = \begin{cases} y_1(t) &= \{(-t^2 + t) + C_1(t-1) + C_2\}e^t \\ y_2(t) &= \{-(t^2 + 3t + 1) + C_1(t+1) + C_2\}e^t + C_3e^{2t} \\ y_3(t) &= \{t^2 + t + 1 - C_1t - C_2\}e^t - C_3e^{2t} \end{cases}$$

où C_1, C_2, C_3 sont des constantes arbitraires.

Remarque 1.13. Si on cherche la solution vérifiant la condition initiale $y_1(0) = -1$, $y_2(0) = 0$, $y_3(0) = 1$. En remplaçant t par 0 dans la solution, on trouve

$$\begin{cases} -C_1 + C_2 &= -1 \\ -1 + C_1 + C_2 + C_3 &= 0 \\ 1 - C_2 - C_3 &= 1 \end{cases}$$

d'où $C_1 = -1$, $C_2 = 0$, $C_3 = 0$. Donc la solution recherchée est

$$Y(t) = \begin{cases} y_1(t) &= (-t^2 + t - 1)e^t \\ y_2(t) &= -(t^2 + 2t)e^t \\ y_3(t) &= (t^2 + 1)e^t \end{cases}$$

1.8 Equations différentielles linéaires d'ordre supérieur

VOIR LE POLY CHAPITRE 4 page 61