Exponentielle de matrice

1 Réduction de matrices : valeurs propres et vecteurs propres

1.1 Introduction

Il n'est pas question ici de traiter de façon complète de la question de la réduction de matrices. Nous nous plaçons résolument dans $\mathbb{K}^n$, ou $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$, et choisissons une approche naive de la notion d'espace vectoriel. Certains énoncés fondamentaux seront, par cette raison, parfois admis.

L'objectif du chapitre suivant est de savoir déterminer les n-uples de fonctions $(x_1, \cdots, x_n)$, définies sur une partie I de $\mathbb{R}$, solutions des systèmes de n équations différentielles d'ordre 1 à n inconnues et à coefficients constants, C'est-à-dire les systèmes de la forme

$$\begin{cases} x_1'(t) = a_{1,1}x_1(t) + a_{1,2}x_2(t) + \cdots + a_{1,n}x_n(t) + y_1(t) \\ x_2'(t) = a_{2,1}x_1(t) + a_{2,2}x_2(t) + \cdots + a_{2,n}x_n(t) + y_2(t) \\ \vdots \\ x_n'(t) = a_{n,1}x_1(t) + a_{n,2}x_2(t) + \cdots + a_{n,n}x_n(t) + y_n(t) \end{cases}$$

où $y_1, \cdots, y_n$ sont n fonctions définies sur I et les $a_{i,j}$ des constantes dans $\mathbb{R}$. Lorsque $n = 1$ on sait que les solutions sont de la forme $t \mapsto x_1(t) = \alpha e^{a_{1,1}t} + f(t)$ où α est une constante arbitraire et f est une solution particulière de l'équation. Nous allons généraliser ce résultat et pour cela nous allons avoir recours à la notion d'*exponentielle de matrice*. Pour pouvoir définir cette exponentielle de matrice, nous devons commencer par quelques résultats de base d'algèbre linéaire. C'est l'objet du présent paragraphe.

Dans tout le chapitre, on fixe un entier $n \geq 1$. On rappelle que la notation $\mathbb{K}$, désigne $\mathbb{R}$ ou $\mathbb{C}$. Les éléments de $\mathbb{K}^n$ seront identifiés avec leurs matrices coordonnées respsectives, et seront donc notés en colonne sous la forme $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

1.2 Rappels

Nous commençons par quelques rappels sur les matrices et sur les systèmes linéaires.

1.2.1 Rappels sur $\mathbb{K}^n$ et sur les matrices

Commençons par rappeler que que $\mathbb{K}^n$ est muni d'une addition :

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

et d'une loi externe

$$\lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix} \text{ où } \lambda \text{ est dans } \mathbb{K}.$$

Rappelons également que si $X, X_1, \cdots, X_r$ sont dans $\mathbb{K}^n$, on dit que X est une combinaison linéaire des $X_1, \cdots, X_r$ s'il existe des éléments $\lambda_1, \cdots, \lambda_r$ de $\mathbb{K}$ tels que $X = \lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_r X_r$.

Nous noterons $M_{k,n}(\mathbb{K})$ l'ensemble des matrices à k lignes et n colonnes. Lorsque $n = k$, on écriera simplement $M_n(\mathbb{K})$. On rappelle que la matrice A de $M_n(\mathbb{K})$ est inversible si et seulement si son déterminant $\det(A)$ est différent de 0. Si A est dans $M_{k,n}(\mathbb{K})$, on note $\mathrm{Ker}(A)$ le noyau de A c'est-à-dire le sous-ensemble de $\mathbb{K}^n$ défini par $\mathrm{Ker}(A) = \{X \in \mathbb{K}^n \mid AX = 0\}$, où 0 désigne la matrice nulle de $\mathbb{K}^k$. On rappelle qu'une matrice A de $M_n(\mathbb{K})$ est inversible si et seulement si son noyau ne contient aucun élément non nul de $\mathbb{K}^n$, c'est-à-dire si $\mathrm{Ker}(A) = \{0\}$. Enfin, on note Id_n la matrice **identité** de $M_n(\mathbb{K})$ définie par $\mathrm{Id}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$, c'est-à-dire avec des 1 sur la diagonale et des 0 ailleurs.

1.2.2 Rappels sur les systèmes linéaires

Définition 1 On appelle **système linéaire sur $\mathbb{K}$ de k équations et n inconnues à coefficients constants** tout système de la forme :

$$\left\{ \begin{array}{l} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{k,1}x_1 + a_{k,2}x_2 + \cdots + a_{k,n}x_n = b_k \end{array} \right. \qquad (*)$$

où les $a_{i,j}$ et les b_j sont dans $\mathbb{K}$.
On dira que le système $(*)$ est homogène lorsque tous les b_i sont égaux à 0.

Rappelons que le rang du système $(*)$ est son nombre d'équations indépendantes. Rappelons également que si k est le rang du système $(*)$, alors on peut exprimer k inconnues en fonctions des $n-k$ autres. On notera rang$(*)$ le rang du système $(*)$.

On appelle **ensemble des solutions** sur $\mathbb{K}^n$ du système $(*)$ l'ensemble de n-uples de $\mathbb{K}^n$ qui vérifient les k équations du système $(*)$.

1.3 Espace Vectoriel

Définition 2 On appelle **sous-espace vectoriel** de $\mathbb{K}^n$ tout ensemble de solutions sur $\mathbb{K}^n$ d'un système linéaire sur $\mathbb{K}$ d'équations homogènes de k équations et à n inconnues $x_1, \cdots, x_n$. Dans ce cas, on dira que le système est **associé** à l'espace vectoriel. On dira que deux systèmes sont **équivalents** s'ils définissent le même sous-espace vectoriel de $\mathbb{K}^n$.

On dira que $\mathbb{K}^n$ est l'espace vectoriel ambiant (sur le corps $\mathbb{K}$), et ses éléments seront appelés des vecteurs.

Proposition 3 *(1)* $\left\{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\right\}$ *et $\mathbb{K}^n$ sont des sous-espaces vectoriels de $\mathbb{K}^n$.*

(2) Un sous-espace vectoriel de $\mathbb{K}^n$ contient toujours $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

(3) Un sous-espace vectoriel est stable par combinaisons linéaires (Si X, Y sont dans V et λ, μ sont dans $\mathbb{K}$, alors $\lambda X + \mu Y$ est dans V).
(4) L'intersection de deux sous-espaces vectoriels est un sous-espace vectoriel.

Preuve : (1) $\left\{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\right\}$ est l'ensemble des solutions du système S_1 de n equations

$$S_1 : x_i = 0 \text{ pour } i \in \{1, \cdots, n\}$$

L'ensemble $\mathbb{K}^n$ est l'ensemble des solutions du système S_2 à une équation

$$S_2 : 0x_1 + \cdots 0x_n = 0.$$

(2) $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ est solution de tout système linéaire homogène et à coefficients constants.

(3) Si X et Y sont deux solutions d'un système S linéaire homogène et à coefficients constants et λ, μ sont dans $\mathbb{K}$ alors $\lambda X + \mu Y$ est aussi solution du système S.
(4) Si V_1 et V_2 sont deux sous-espaces vectoriels associés respectivement à deux systèmes S_1 et S_2 linéaires homogènes à coefficients constants, à k_1 et k_2 équations respectivement, alors $V_1 \cap V_2$ est l'ensemble des solutions du système S linéaire homogène à coefficients constants, à $k_1 + k_2$ équations, obtenu en regroupant les équations de S_1 et les équations de S_2. $\square$

Proposition 4 *(Admis) Soit V est un sous-espace vectoriel de $\mathbb{K}^n$. Tous les systèmes associés à V ont le même rang (et sont équivalents).*

Définition 5 Soit V est un sous-espace vectoriel de $\mathbb{K}^n$. On appelle **dimension** de V l'entier égal à $n - k$ où k est le rang de n'importe quel système associé à V. On notera cet entier $\dim(V)$.

Proposition 6 *(Admis) Soit V est un sous-espace vectoriel de $\mathbb{K}^n$. Si la dimension r de V est non nulle, alors il existe r éléments (non nuls) , $X_1, \cdots, X_r$, dans V tels que V est l'ensemble des combinaisons linéaires des $X_1, \cdots, X_r$. De plus chaque élément de V s'écrit alors de manière unique comme une combinaison linéaire des $X_1, \cdots, X_r$.*

Le r-uple $(X_1, \cdots, X_r)$ de la proposition s'appelle une **base** de V.

Notons qu'un espace vectoriel non nul possède toujours de nombreuses bases différentes, mais toutes de même cardinal (égal à la dimension de l'espace vectoriel).

Exemple 1 $\mathbb{K}^n$ est de dimension n et

$$\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right)$$

en est une base. Cette base s'appelle la base canonique de $\mathbb{K}^n$.

Traitons un exemple :
Soit $n = 4$, $\mathbb{K} = \mathbb{R}$, et

$$S : \left\{ \begin{array}{ccccccccc} x_1 & + & x_2 & + & x_3 & & & = & 0 \\ & & x_2 & & & - & x_4 & = & 0 \\ x_1 & - & x_2 & + & x_3 & + & 2x_4 & = & 0 \end{array} \right. .$$

Notons V le sous-espace vectoriel de $\mathbb{R}^4$ associé à S. Le système S est équivalent au système $\left\{\begin{array}{ccrr} x_3 & = & -x_1 & -x_2 \\ x_4 & = & & x_2 \end{array}\right.$, qui est clairement de rang 2. L'espace vectoriel V est donc de dimension $4-2=2$. L'ensemble V des solutions de S est

$$V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ -x_1 - x_2 \\ x_2 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R} \right\} = \left\{ x_1 \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R} \right\}.$$

Donc $\left(\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix} \right)$ est une base de V.

Le résultat suivant est immédiat mais essentiel :

Proposition 7 *Soit V un sous-ensemble de $\mathbb{K}^n$. Le sous-ensemble V est un sous-espace vectoriel de $\mathbb{K}^n$ si et seulement et si il existe un entier non nul k et une matrice A dans $M_{k,n}(\mathbb{K})$ tel que $V = \mathrm{Ker}(A)$.*

Dans le cas où V est associé au système $(*)$ avec $y_1 = \cdots = y_k = 0$, la matrice A est

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & a_{i,j} & \vdots \\ a_{k,1} & \cdots & a_{k,n} \end{pmatrix}.$$

1.4 Valeurs propres et vecteurs propres

Définition 8 Soit A dans $M_n(\mathbb{K})$. Soit λ dans $\mathbb{K}$.
(1) On dit que λ est une **valeur propre** de A s'il existe un vecteur <u>non nul</u> X de $\mathbb{K}^n$ tel que $AX = \lambda X$. Dans ce cas, X est appelé un **vecteur propre** de A pour la valeur propre λ.
(2) Si λ est une valeur propre de A, alors on appelle **espace propre de A associé à la valeur propre** λ l'ensemble $V_\lambda(A)$ des solutions de l'équation matricielle $AX = \lambda X$.

Autrement dit, λ est une valeur propre de la matrice $A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & a_{i,j} & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$

s'il existe $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ tel que

$$\left\{ \begin{array}{l} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = \lambda x_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = \lambda x_2 \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = \lambda x_n \end{array} \right. \qquad (\dagger)$$

Proposition 9 *Soit A dans $M_n(\mathbb{K})$. Soit λ dans $\mathbb{K}$ une valeur propre de A. Alors,*
(1) $V_\lambda(A)$ est un sous-espace vectoriel de $\mathbb{K}^n$. En fait, $V_\lambda(A) = \mathrm{Ker}(A - \lambda \, \mathrm{Id}_n)$.
(2) Si X est dans $V_\lambda(A)$, alors X est nul ou X est un vecteur propre associé à la valeur propre λ.

Preuve : (1) Soit X dans $\mathbb{K}^n$. Alors on a :

$$\begin{aligned} X \in V_\lambda(A) \iff AX = \lambda X \iff AX = \lambda \mathrm{Id}_n X \iff AX - \lambda \mathrm{Id}_n X = 0 \\ \iff (A - \lambda \mathrm{Id}_n)X = 0 \iff X \in \mathrm{Ker}(A - \lambda \mathrm{Id}_n). \end{aligned}$$

D'où, $V_\lambda(A) = \mathrm{Ker}(A - \lambda \, \mathrm{Id}_n)$, qui est un sous-espace vectoriel d'après la proposition 7.
Le (2) est vrai par définition de $V_\lambda(A)$. □

Ainsi, s'il on veux savoir si le réel λ est une valeur propre de A, il suffit de résoudre le système matriciel $AX = \lambda X$ et de voir si ce système possède une solution non nulle. Néanmoins, cette approche ne permet pas de déterminer toutes les valeurs propres de A car on ne peut pas tester tous les nombres λ de $\mathbb{K}$ un par un. Nous allons voir dans le paragraphe suivant un moyen de determiner toutes les valeurs propres de A. Toutefois, on peut parfois déterminer certaines valeurs propres en regardant la matrice A.

Nous regardons quelques exemples ci-dessous.

Exemple 2 (1) Soit $A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 2 & 0 & 0 \end{pmatrix}$ Alors $\det(A) = 0$ donc A n'est pas inversible et il existe X non nul dans $\mathbb{R}^3$ tel que $AX = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0X$. Donc 0 est valeur propre de A.

(2) Soit $B = \begin{pmatrix} 3 & 2 & -1 \\ 0 & -1 & 0 \\ 0 & 1 & 3 \end{pmatrix}$. La première colonne de la matrice B est l'image par B du premier vecteur de la base canonique (plus généralement, la colonne i correspond à l'image du $i^{\text{ème}}$ vecteur de la base canonique). Donc

$$B \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

D'où, 3 est une valeur propre, et $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ est un vecteur propre de B associé à la valeur propre 3.

(3) Soit $C = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 3 \end{pmatrix}$. La somme des deux premières colonnes de la matrice est $\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. Donc, 2 est une valeur propre de C, et $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ est un vecteur propre de C associée à la valeur propre 2.

(4) Soit $D = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$. On a

$$D \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ et } D \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 3 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Donc 3 est une valeur propre de D et les deux vecteurs $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ sont deux vecteurs propres de D associés à la valeur propre 3. De plus, ils sont clairement indépendants. En résolvant l'équation matricielle

$$D \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 3 \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

on voit que $\dim(V_3(D)) = 2$ et que les deux vecteurs $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ forment une base de $V_3(D)$.

1.5 Polynôme caractéristique et polynôme minimal

Nous allons maintenant associer à toute matrice deux polynômes qui vont nous permettre de calculer son exponentielle.

1.5.1 Polynôme caractéristique

Définition 10 Soit A dans $M_n(\mathbb{K})$. On appelle **polynôme caractéristique de la matrice** A le polynôme, noté $\chi(A)$, défini par $\chi_A = \det(X\,\mathrm{Id}_n - A)$.

Proposition 11 *Soit A dans $M_n(\mathbb{K})$ et λ dans $\mathbb{K}$. Alors, λ est une valeur propre de A si et seulement si $\chi_A(\lambda) = 0$.*

Preuve : λ est une valeur propre de A si et seulement si $Ker(A - \lambda\,\mathrm{Id}_n) \neq \{0\}$ si et seulement si $\det(A - \lambda\,\mathrm{Id}_n) = 0$ si et seulement si $\det(\lambda\,\mathrm{Id}_n - A) = 0$ si et seulement si $\chi_A(\lambda) = 0$. □

Pour déterminer les valeurs propres de A, il suffit donc de calculer son polynôme caractéristique χ_A et de déterminer ses racines. On peut alors déterminer chaque espace caractéristique. Remarquons que si l'on veut trouver les valeurs propres, il faut être capable de trouver les racines de χ_A, et donc de le factoriser. Les méthodes de type Sarrus sont donc à proscrire. Il faut calculer χ_A de façon " intelligente ".

Exemple 3 Considérons les matrices A, B, C et D de l'exemple 2.
(1) On a

$$\chi_A(X) = \det(X\mathrm{Id}_n - A) = \begin{vmatrix} X-1 & -1 & -1 \\ 1 & X-1 & -1 \\ -2 & 0 & X \end{vmatrix} = \begin{vmatrix} X-1 & 0 & -1 \\ 1 & X & -1 \\ -2 & -X & X \end{vmatrix}$$

$$= X \begin{vmatrix} X-1 & 0 & -1 \\ 1 & 1 & -1 \\ -2 & -1 & X \end{vmatrix} = X \begin{vmatrix} X-1 & 0 & -1 \\ 1 & 1 & -1 \\ -1 & 0 & X-1 \end{vmatrix} = X \begin{vmatrix} X-1 & -1 \\ -1 & X-1 \end{vmatrix}$$

$$= X((X-1)^2 + 1) = X(X^2 - 2X) = X^2(X-2).$$

Les valeurs propres de A sont donc 0 et 2.
(2) On a

$$\chi_B(X) = \det(X\mathrm{Id}_n - B) = \begin{vmatrix} X-3 & -2 & 1 \\ 0 & X+1 & 0 \\ 0 & -1 & X-3 \end{vmatrix}$$

$$= (X+1) \begin{vmatrix} X-3 & 1 \\ 0 & X-3 \end{vmatrix} = (X+1)(X-3)^2.$$

Les valeurs propres de B sont donc 1 et 3.
(3) On a

$$\chi_C(X) = \det(X\mathrm{Id}_n - C) = \begin{vmatrix} X-1 & -1 & 1 \\ 0 & X-2 & -1 \\ -1 & -1 & X-3 \end{vmatrix}$$

$$= \begin{vmatrix} X-2 & -1 & 1 \\ X-2 & X-2 & -1 \\ 0 & -1 & X-3 \end{vmatrix} = (X-2) \begin{vmatrix} 1 & -1 & 1 \\ 1 & X-2 & -1 \\ 0 & -1 & X-3 \end{vmatrix}$$

$$= (X-2)\begin{vmatrix} 1 & -1 & 1 \\ 0 & X-1 & -2 \\ 0 & -1 & X-3 \end{vmatrix} = (X-2)(X^2-4X+1)$$
$$= (X-2)(X-(2+\sqrt{3}))(X-(2-\sqrt{3})).$$

Les valeurs propres de C sont donc 2, $2+\sqrt{3}$ et $2-\sqrt{3}$.
(4) On trouve $\chi_D(X) = (X-3)^2(x-2)$. Les valeurs propres de la matrice D sont donc 3 et 2.

Lemme 12 *(Admis) Soit A dans $M_n(\mathbb{K})$ et λ une valeur propre de A. Supposons que $\chi_A(X) = (X-\lambda)^m P(X)$ avec $P(\lambda) \neq 0$. Alors, $1 \leq \dim(V_\lambda(A)) \leq m$.*

Ainsi, dans l'exemple 2 on a $\dim(V_{2+\sqrt{3}}(C)) = 1$ et $1 \leq dim(V_0(A)) \leq 2$.

Définition 13 (1) On dit qu'un polynôme P est **unitaire** si son coefficient dominant est égal à 1, c'est-à-dire si $P(X) = \sum_{j=0}^k a_j X^j$ avec $a_k = 1$.
(2) On dit qu'un polynôme est **scindé**, s'il s'écrit comme un produit de polynômes de degré 1.

Rappelons que dans $\mathbb{C}[X]$, tout polynôme est scindé, mais que ce n'est pas le cas dans $\mathbb{R}[X]$. Par exemple, $X^2+1 = (X-i)(X+i)$, mais n'est pas scindé dans $\mathbb{R}[X]$.

Proposition 14 *(Admis) Soit A dans $M_n(\mathbb{K})$. Le polynôme χ_A est unitaire, de degré n et de la forme $\chi_A = X^n - trace(A)X^{n-1} + \sum_{j=1}^{n-2} a_j X^j + (-1)^n det(A)$. En particulier, si χ_A est scindé, alors $trace(A)$ est égal à la somme des valeurs propres de A comptées avec leur mutiplicité dans χ_A.*

Exemple 4 Dans l'exemple 2, on a $trace(A) = 1+1+0 = 2 = 0+0+2$ et $trace(C) = 1+2+3 = 6 = 2+(2-\sqrt{3})+(2+\sqrt{3})$.

La proposition suivante va nous permettre d'énoncer l'un des théorèmes fondamentaux de l'algèbre linéraire : le théorème de Cayley-Hamilton.

Proposition 15 *(Admis) Pour toute matrice A dans $M_n(\mathbb{K})$, l'application de $\mathbb{K}[X]$ dans $M_n(\mathbb{K})$ qui associe au polynôme $P(X) = \sum_{i=0}^k a_i X^i$ la matrice $\sum_{i=0}^k a_i A^i$ est un morphisme de $\mathbb{K}$-algèbres unitaires. L'image du polynôme P par cette application s'appelle* **l'évaluation de P en A** *et se note $P(A)$.*

Le vocabulaire abscon de cette proposition veut simplement dire que si on remplace la variable X par la matrice A, tout se passe bien : si P et Q sont des polynômes et α est dans $\mathbb{K}$, alors $(P+Q)(A) = P(A)+Q(A)$; $(PQ)(A) = P(A)Q(A)$ et $(\alpha P)(A) = \alpha P(A)$; de plus, $1(A) = \mathrm{Id}_n$. Attention toutefois à ne pas oublier cette dernière propriété. Par exemple $(2X+1)(A) = 2A + Id_n$. Puisque $(X-2)(X+1) = X^2-X-2$, pour toute matrice A, on a $(A-2\,\mathrm{Id}_n)(A+\mathrm{Id}_n) = A^2 - A - 2\,\mathrm{Id}_n$.

Théorème 16 (de Cayley-Hamilton) *(admis)*
Soit A dans $M_n(\mathbb{K})$. Alors $\chi_A(A)$ est nul.

Exemple 5 Reprenons l'exemple 2.
On a $\chi_C(X) = (X-2)(X^2-4X+1)$, donc $\chi_C(C) = (C-\mathrm{Id}_n)(C^2-4C+\mathrm{Id}_n) =$

$$\left(\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 3 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)\left(\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 3 \end{pmatrix}\right.$$
$$\left.-4\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right) =$$
$$\begin{pmatrix} -1 & 1 & -1 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix}\left(\begin{pmatrix} 2 & 2 & -3 \\ -1 & 5 & 5 \\ -4 & 4 & 11 \end{pmatrix} - 4\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)$$
$$= \begin{pmatrix} -1 & 1 & -1 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} -1 & -2 & 1 \\ -1 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

1.5.2 Polynôme minimal

Proposition 17 *Soit A dans $M_n(\mathbb{K})$. L'ensemble $Ann(A)$ des* **polynômes annulateurs de** A*, défini par $Ann(A) = \{P \in \mathbb{K}[X] \mid P \neq 0$ et $P(A) = 0\}$, est non vide et possède un unique polynôme μ_A de degré minimal dans $Ann(A)$ et unitaire. De plus, $P \in \mathbb{K}[X]$, non nul, est dans $Ann(A)$ si et seulement si μ_A divise P.*

Preuve : $Ann(A)$ est non vide d'après le théorème de Cayley-Hamilton. Soit Q dans $Ann(A)$ de degré minimal et c son coefficient principal. Posons $\mu_A = \frac{1}{c}Q$. Alors, μ_a est unitaire et dans $Ann(A)$, puisque $\mu_A(A) = \frac{1}{c}Q(A) = 0$. Soit P dans $\mathbb{K}[X]$ non nul. Si μ_A divise P, alors il existe R dans $\mathbb{K}[X]$ tel que $P = \mu_A R$. Donc $P(A) = \mu_A(A)R(A) = 0$; donc P est dans $Ann(A)$. Réciproquement supposons que P est dans $Ann(A)$. Il existe R, S dans $\mathbb{K}[X]$ avec $d^\circ S < d^\circ \mu_A$ tel que $P = \mu_A R + S$. On a alors $S(A) = P(A) - R(A)\mu_A(A) = 0$. Mais S n'est pas dans $Ann(A)$ puisque $d^\circ S < d^\circ \mu_A$. Donc $S = 0$ et $P = \mu_A R$. Supposons de plus que P est de degré minimum dans $Ann(A)$. Alors, $d^\circ P = d^\circ \mu_A$, et R appartient à $\mathbb{K}$. Puisque μ_A est unitaire, R est le coefficient principal de P. Si on suppose que P est aussi unitaire, alors on doit avoir $R = 1$ et $P = \mu_A$. $\square$

Définition 18 Soit A dans $M_n(\mathbb{K})$. Alors le polynôme μ_A de la proposition 17 s'appelle le **polynôme minimal** de A.

Remarquons que d'après le théorème de Cayley-Hamilton μ_A divise χ_A, par conséquent, toute racine de μ_A est une racine de χ_A. Le proposition suivante montre que l'on a un résultat bien plus fort :

Proposition 19 *(admis) Soit A dans $M_n(\mathbb{K})$. Le polynôme μ_A possède les mêmes racines que χ_A.*

Corollaire 20 *Soit A dans $M_n(\mathbb{K})$. Supposons que χ_A est scindé, et que*

$$\chi_A = (X-\alpha_1)^{n_1}(X-\alpha_2)^{n_2}\cdots(X-\alpha_k)^{n_k}$$

avec $\alpha_i \neq \alpha_j$ pour $i \neq j$ et $n_i \geq 1$ pour $i \in \{1, \cdots, k\}$. Alors

$$\mu_A = (X-\alpha_1)^{n_1}(X-\alpha_2)^{n_2}\cdots(X-\alpha_k)^{n_k}$$

avec $1 \leq m_i \leq n_i$ pour tout i dans $\{1, \cdots, k\}$.

Nous ne verrons pas de méthode générale permettant de calculer μ_A. Pour le calculer, on calculera χ_A et on utilisera la proposition précédente ainsi que le fait que μ_A est le polynôme unitaire de degré minimal qui annule A.

Exemple 6 Reprenons les matrices de l'exemple 2.
(1) Puisque $\chi_A = X^2(x-2)$, on a $\mu_A = X(X-2)$ ou bien $\mu_A = X^2(X-2)$. Mais $(X(X-2))(A) = A(A-2\,\mathrm{Id}_n) = \begin{pmatrix} 0 & 0 & 0 \\ 2 & -2 & -2 \\ -2 & 2 & 2 \end{pmatrix}$. Donc $\mu_A = X^2(X-2)$.
(2) $\chi_B = (X-3)^2(X-1)$, donc $\mu_B = (X-3)(X-1)$ ou bien $\mu_B = (X-3)^2(X+1)$. Mais, $(B-3\,\mathrm{Id}_n)(B+\mathrm{Id}_n)$ est différent de 0 donc $\mu_A = (X-3)^2(X+1)$.
(3) Puisque toutes les racines de χ_C sont simples, on a $\chi_A = \mu_A$.
(4) $\chi_D = (X-3)^2(X-2)$, donc $\mu_D = (X-3)(X-2)$ ou bien $\mu_D = (X-3)^2(x-2)$. Mais, $(D-3\mathrm{Id}_n)(D-2\mathrm{Id}_n) =$

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

Donc $\mu_D = (X-3)(X-2)$. Bien entendu, on a $(D-3\,\mathrm{Id}_n)^2(D-2\,\mathrm{Id}_n)$ est aussi la matrice nulle, mais le degré de $(X-3)(X-2)$ est plus petit que celui de $(X-3)^2(X-2)$.

2 Exponentielle de matrice

2.1 Définition et propriétés

On rappelle que l'on dit qu'une suite complexe $(x_k)_{k\geq 0}$ converge vers x dans $\mathbb{C}$ si $\lim_{k\to+\infty} x_k = x$, c'est-à-dire si

$$\forall \epsilon > 0 \exists M \in \mathbb{N} / \forall k \geq M, |x_k - x| \leq \epsilon.$$

Rappelons également que pour tout x dans $\mathbb{C}$, on a $e^z = \lim_{k\to+\infty} \sum_{i=0}^{k} \frac{x^i}{i!}$.

Définition 21 Soit $(A_k)_{k\geq 0}$ une suite de matrices de $M_n(\mathbb{K})$ avec

$$A_k = \begin{pmatrix} a_{k,1,1} & \cdots & a_{k,1,n} \\ \vdots & a_{k,i,j} & \vdots \\ a_{k,n,1} & \cdots & a_{k,n,n} \end{pmatrix}.$$

Soit $A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$ dans $M_n(\mathbb{K})$. On dit que la suite $(A_k)_{k\geq 0}$ **converge vers la matrice** A si pour tout i, j, la suite $(a_{k,i,j})_{k\geq 0}$ converge dans $\mathbb{K}$ vers $a_{i,j}$. Dans ce cas, la matrice A s'appelle **la limite** de la suite.

Notons que les règles classiques sur les sommes et les produits de limites de suites s'appliquent dans ce contexte.

Théorème 22 *(Admis) Soit A dans $M_n(\mathbb{K})$. La suite de matrice $\left(\sum_{j=0}^{k} \frac{A^j}{j!}\right)_{k\geq 0}$ converge dans $M_n(\mathbb{K})$.*

Définition 23 Soit A dans $M_n(\mathbb{K})$. On appelle **exponentielle de** A la limite de la suite $\left(\sum_{j=0}^{k} \frac{A^j}{j!}\right)_{k\geq 0}$. On note e^A ou bien $exp(A)$ cette matrice.

Exemple 7 Considérons la matrice $A = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \lambda & 0 \\ 0 & \cdots & 0 & \lambda \end{pmatrix}$. Alors, pour tout entier j, on a $A^j = \begin{pmatrix} \lambda^j & 0 & \cdots & 0 \\ 0 & \lambda^j & \ddots & \vdots \\ \vdots & \ddots & \lambda^j & 0 \\ 0 & \cdots & 0 & \lambda^j \end{pmatrix}$ et

$$\left(\sum_{j=0}^{k} \frac{A^j}{j!}\right) = \begin{pmatrix} \sum_{j=0}^{k} \frac{\lambda^j}{j!} & 0 & \cdots & 0 \\ 0 & \sum_{j=0}^{k} \frac{\lambda^j}{j!} & 0 & \vdots \\ \vdots & 0 & \sum_{j=0}^{k} \frac{\lambda^j}{j!} & 0 \\ 0 & \cdots & 0 & \sum_{j=0}^{k} \frac{\lambda^j}{j!} \end{pmatrix}.$$

Donc, $Exp(A) = \begin{pmatrix} e^\lambda & 0 & \cdots & 0 \\ 0 & e\lambda & \ddots & \vdots \\ \vdots & \ddots & e^\lambda & 0 \\ 0 & \cdots & 0 & e^\lambda \end{pmatrix}$.

Ainsi, l'exponentielle de la matrice nulle est la matrice Id_n.

Nous énonçons maintenant les résultats principaux concernant l'exponentielle d'une matrice.

Proposition 24 *(Admis)*
(1) Soit A et B dans $M_n(\mathbb{K})$. Si $AB = BA$ alors $Exp(A+B) = Exp(A)Exp(B)$.
(2) Pour toute matrice A de $M_n(\mathbb{K})$, la matrice $Exp(A)$ est inversible et

$$(Exp(A))^{-1} = Exp(-A).$$

Remarquons que le (2) est une conséquence de (1) puisque, d'une part, A et $-A$ commutent et, d'autre part, l'exponentielle de la matrice nulle est Id_n.

Proposition 25 *Soit A dans $M_n(\mathbb{K})$. Soit ℓ le degré de μ_A. Alors, pour tout polynôme P de $\mathbb{R}[X]$ il existe un (unique) ℓ-uple $(\beta_0, \cdots, \beta_{\ell-1})$ d'éléments de $\mathbb{K}$ tels que $P(A) = \beta_{\ell-1}A^{\ell-1} + \cdots \beta_1 A + \beta_0 \mathrm{Id}_n$.*

Preuve : Soit P dans $\mathbb{K}[X]$. Il existe deux polynômes S, R avec $d^\circ S < d^\circ \mu_A$ tels que $P = R\mu_A + S$. Donc $P(A) = R(A)\mu_A(A) + S(A) = S(A)$. D'où, $P(A) = \beta_{\ell-1}A^{\ell-1} + \cdots \beta_1 A + \beta_0 \mathrm{Id}_n$ avec $S(X) = \beta_{\ell-1}X^{\ell-1} + \cdots \beta_1 X + \beta_0$. □

De cette proposition on peut déduire les résultats suivants. Nous les admettrons cependant car ils nous manquent certaines notions mathématiques pour les prouver.

Proposition 26 *(Admis) Soit A dans $M_n(\mathbb{K})$ et ℓ le degré de μ_A. Alors il existe un (unique) ℓ-uple $(\beta_0, \cdots, \beta_{\ell-1})$ d'éléments de $\mathbb{K}$ tels que*

$$exp(A) = \beta_{\ell-1}A^{\ell-1} + \cdots \beta_1 A + \beta_0 Id_n.$$

Proposition 27 *(Admis) Soit A dans $M_n(\mathbb{K})$ et ℓ le degré de μ_A. Alors il existe un (unique) ℓ-uple $(\beta_0, \cdots \beta_{\ell-1})$ de fonctions de classe C^∞ de $\mathbb{K}$ dans $\mathbb{K}$ telles que pour tout t de $\mathbb{K}$ on a*

$$exp(tA) = \beta_{\ell-1}(t)A^{\ell-1} + \cdots \beta_1(t)A + \beta_0(t)Id_n.$$

2.2 Détermination d'une exponentielle de matrice

Le résultat ci-desous va nous permettre de calculer l'exponentielle de n'importe quelle matrice de $M_n(\mathbb{K})$.

Proposition 28 *Soit A dans $M_n(\mathbb{K})$ et ℓ le degré du polynôme minimal μ_A. Notons $V_t(X) = \beta_{\ell-1}(t)X^{\ell-1} + \cdots \beta_1 X + \beta_0$ le polynôme dans $\mathbb{K}[X]$ tel que $V_t(A) = Exp(tA)$. Supposons que μ_A est scindé avec*

$$\mu_A = (x - \lambda_1)^{m_1} \cdots (X - \lambda_k)^{m_k}$$

avec $\lambda_i \neq \lambda_j$ si $i \neq j$ et $m_i \geq 1$ pour i dans $\{1, \cdots, k\}$.
Alors les fonctions $\beta_0, \cdots, \beta_{\ell-1}$ sont les uniquement déterminées par les système de ℓ équations :

$$t^j e^{\lambda_i t} = V_t^{(j)}(\lambda_i)$$

où i parcours l'ensemble $\{1, \cdots, k\}$, et pour chaque i l'entier j parcours l'ensemble $\{0, \cdots m_i - 1\}$.
Pour obtenir $Exp(A)$, il suffit alors de prendre $t = 1$.

Remarquons que dans $\mathbb{C}$, tous les polynômes sont scindés.

Exemple 8 Reprenons les matrices de l'exemple 2.
(1) Puisque $\mu_A = X^2(X-2)$ est de degré 3, la matrice $Exp(tA)$ est égale à $V_t(A)$ où $V_t(X) = \beta_2(t)X^2 + \beta_1(t)A + \beta_0(t)$ tel que β_0, β_1 et β_2 sont déterminés par le système :

$$\left\{ \begin{array}{ccc|l} 1 & = & \beta_0(t) & (\lambda = 0, m_i = 2, j = 0) \\ t & = & \beta_1(t) & (\lambda = 0, m_i = 2, j = 1) \\ e^{2t} & = & 4\beta_2(t) + 2\beta_1(t) + \beta_0(t) & (\lambda = 2, m_i = 1, j = 0) \end{array} \right.$$

On trouve donc $\beta_0(t) = 1$; $\beta_1(t) = t$, $\beta_2(t) = \frac{1}{4}(e^{2t} - 2t - 1)$ et

$Exp(tA) =$

$$\frac{1}{4}(e^{2t} - 2t - 1) \begin{pmatrix} 2 & 2 & 2 \\ 0 & 0 & 0 \\ 2 & 2 & 2 \end{pmatrix} + t \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 2 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} =$$

$$\frac{1}{2}e^{2t} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} + t \begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 & -1 & -1 \\ 0 & 2 & 0 \\ -1 & -1 & 1 \end{pmatrix}.$$

(2) Puisque $\mu_B = (X-3)^2(X-1)$ est de degré 3, la matrice $Exp(tB)$ est égale à $V_t(B)$ où $V_t(X) = \beta_2(t)X^2 + \beta_1(t)A + \beta_0(t)$ tel que β_0, β_1 et β_2 sont déterminés par le système :

$$\left\{ \begin{array}{ccc|l} e^{3t} & = & 9\beta_2(t) + 3\beta_1(t) + \beta_0(t) & (\lambda = 3, m_i = 2, j = 0) \\ te^{3t} & = & 6\beta_2(t) + \beta_1(t) & (\lambda = 3, m_i = 2, j = 1) \\ e^t & = & \beta_2(t) + \beta_1(t) + \beta_0(t) & (\lambda = 1, m_i = 1, j = 0) \end{array} \right.$$

(3) Puisque $\mu_C = (X-2)(X-(2+\sqrt{3}))(X-(2-\sqrt{3}))$ est de degré 3, la matrice $Exp(tC)$ est égale à $V_t(C)$ où $V_t(X) = \beta_2(t)X^2 + \beta_1(t)A + \beta_0(t)$ tel que β_0, β_1 et β_2 sont déterminés par le système :

$$\left\{ \begin{array}{ccc|l} e^{2t} & = & 4\beta_2(t) + 2\beta_1(t) + \beta_0(t) & (\lambda = 2, m_i = 1, j = 0) \\ e^{(2+\sqrt{3})t} & = & (2+\sqrt{3})^2\beta_2(t) + (2+\sqrt{3})\beta_1(t) + \beta_0(t) & (\lambda = 2+\sqrt{3}, m_i = 1, j = 0) \\ e^{(2-\sqrt{3})t} & = & (2-\sqrt{3})^2\beta_2(t) + (2-\sqrt{3})\beta_1(t) + \beta_0(t) & (\lambda = 2-\sqrt{3}, m_i = 1, j = 0) \end{array} \right.$$

(4) Puisque $\mu_D = (X-3)(X-2)$ est de degré 2, on a $Exp(D) = \beta_1(t)D + \beta_0(t)\,\mathrm{Id}_n$ tel que le polynôme $\beta_1(t)X+\beta_0(t)$ est déterminé par les deux équations $e^{3t} = 3\beta_1(t)+\beta_0(t)$ et $e^{2t} = 2\beta_1(t)+\beta_0(t)$. En résolvant, on trouve $\beta_1(t) = e^{3t}-e^{2t}$ et $\beta_0(t) = 3e^{2t}-2e^{3t}$. Donc,

$$Exp(tD) = (e^{3t}-e^{2t})\begin{pmatrix} 9 & 5 & -5 \\ 0 & 4 & 5 \\ 0 & 0 & 9 \end{pmatrix} + (3e^{2t}-2e^{3t})\begin{pmatrix} 3 & 1 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} =$$

$$3e^{3t}\begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} + 2e^{2t}\begin{pmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

(5) Considérons la matrice $E = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. On a $\chi_E = X^2+1$ qui n'est pas scindé dans $\mathbb{R}$. Pour calculer $Exp(tE)$, on considère la matrice E dans $M_2(\mathbb{C})$. On a alors $\chi_E = (X-i)(X+i)$ et par conséquence $\mu_E = \chi_E$. Donc $Exp(tE) = \beta_1(t)E+\beta_0(t)\,\mathrm{Id}_n$ tel que le polynôme $\beta_1(t)X+\beta_0(t)$ est déterminé par les deux équations $e^{it} = i\beta_1(t)+\beta_0(t)$ et $e^{-it} = -i\beta_1(t)+\beta_0(t)$. On trouve donc $\beta_0(t) = cos(t)$ et $\beta_1(t) = sin(t)$. Donc $Exp(tE) = \begin{pmatrix} cos(t) & -sin(t) \\ sin(t) & cos(t) \end{pmatrix}$.
En particulier $Exp(E) = \begin{pmatrix} cos(1) & -sin(1) \\ sin(1) & cos(1) \end{pmatrix}$

Remarquons que l'exponentielle de la matrice nulle est la matrice Id_n. Donc pour toute matrice A et pour $t = 0$, on a $Exp(tA)$ est la matrice Id_n. On voit facilement dans l'exemple précédent qu'en faisant $t = 0$ dans les formules de $Exp(tA)$, $Exp(tD)$ et $Exp(tE)$ on retrouve bien la matrice Id_n. C'est un moyen de détecter une éventuelle erreur de calcul.

2.3 Dérivation

Définition 29 Soit I un intervalle de $\mathbb{R}$ et

$$\begin{array}{rccc} \varphi : & I & \to & M_n(\mathbb{K}) \\ & t & \mapsto & \begin{pmatrix} a_{1,1}(t) & \cdots & a_{1,n}(t) \\ \vdots & a_{i,j}(t) & \vdots \\ a_{k,1}(t) & \cdots & a_{k,n}(t) \end{pmatrix} \end{array}$$ une application.

(1) On dira que l'application φ est **dérivable** (*resp.* **de classe** $\mathcal{C}^k$) sur I si pour chaque paire (i,j) l'application $a_{i,j} : I \to \mathbb{K}, t \mapsto a_{i,j}(t)$ est dérivable (*resp.* de classe $\mathcal{C}^k$) sur I. On dira que φ est **de classe** $\mathcal{C}^\infty$ si φ est de classe $\mathcal{C}^k$ pour tout entier positif k.
(2) Si φ est dérivable sur I, on appelle **dérivée de** φ l'application φ' donnée par :

$$\forall t \in I,\ \varphi'(t) = \begin{pmatrix} a'_{1,1}(t) & \cdots & a_{1,n}(t) \\ \vdots & a'_{i,j}(t) & \vdots \\ a'_{k,1}(t) & \cdots & a'_{k,n}(t) \end{pmatrix}.$$

(3) Si φ est de classe $\mathcal{C}^k$ sur I, on appelle dérivée $k$ème de φ la dérivée de la dérivée $(k-1)$ème de φ. On notera $\varphi^{(k)}$ cette dérviée $k$ème de φ. Ainsi on a

$$\forall t \in I,\ \varphi^{(k)}(t) = \begin{pmatrix} a_{1,1}^{(k)}(t) & \cdots & a_{1,n}^{(k)}(t) \\ \vdots & a_{i,j}^{(k)}(t) & \vdots \\ a_{k,1}^{(k)}(t) & \cdots & a_{k,n}^{(k)}(t) \end{pmatrix}.$$

Remarque 1 *Soit I un intervalle de $\mathbb{R}$. Soit $A_1, \cdots, A_r$ des matrices de $M_n(\mathbb{K})$ et $\beta_1, \cdots, \beta_r$ des application de I dans $\mathbb{K}$. Alors l'application*

$$\begin{array}{rccl} \varphi : & I & \to & M_n(\mathbb{K}) \\ & t & \mapsto & \beta_1(t)A_1 + \cdots \beta_r(t)A_r \end{array}$$

est dérivable (resp. de classe $\mathcal{C}^k$, de classe $\mathcal{C}^\infty$) si et seulement si chaque application β_i est dérivable (resp. de classe $\mathcal{C}^k$, de classe $\mathcal{C}^\infty$). De plus, dans ce cas, on a

$$\varphi^{(k)}(t) = \beta_1^{(k)}(t)A_1 + \cdots \beta_r^{(k)}(t)A_r$$

pour tout t dans I.

Remarque 2 *les régles classiques de dérivation d'un produit ou d'une somme restent valablent dans ce contexte.*

Proposition 30 *(Admise) Pour toute matrice A de $M_n(\mathbb{K})$, l'application $t \mapsto Exp(tA)$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$. De plus, sa dérivée $k$ème est $t \mapsto A^k Exp(tA)$.*

Remarque 3 *D'après la règle de dérivation d'un produit, il suffit de savoir montrer le résultat pour $k = 1$ pour déduire la proposition par récurrence sur k : la matrice $A^{(k-1)}$ est constante. Pour tout entier m, on a*

$$\sum_{j=0}^{m} \left(\frac{t^j A^j}{j!}\right)' = \sum_{j=1}^{m} \frac{t^{j-1}A^j}{(j-1)!} = A \sum_{j=0}^{m-1} \frac{t^j A^j}{j!}.$$

Donc, on a $\lim_{m\to+\infty} \sum_{j=0}^{m} (\frac{t^j A^j}{j!})' = AExp(tA)$. Il faut donc montrer que $\lim_{m\to+\infty} \sum_{j=0}^{m} (\frac{t^j A^j}{j!})' = \lim_{m\to+\infty} (\sum_{j=0}^{m} \frac{t^j A^j}{j!})'$. C'est ce point que nous admettons car nous n'avons pas les outils mathématiques pour le faire.

Exemple 9 Reprenons les matrices de l'exemple 2 et le calcul de l'exemple 8.
(1) On a $(Exp(tA))' = e^{2t} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix}$ et pour $k \geq 2$,

$$(Exp(tA))^{(k)} = 2^{k-1}e^{2t} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

(2) Pour tout entier k, on a

$$Exp(tD)^{(k)} = 3^{k+1}e^{3t}\begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} + 2^{k+1}e^{2t}\begin{pmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Corollaire 31 *Soit A dans $M_n(\mathbb{K})$. Alors, pour tout entier k, on a*

$$A^k = exp(tA)^{(k)}|_{t=0},$$

c'est-à-dire que la dérivée k-ème de l'application $t \mapsto exp(tA)$ évaluée en $t = 0$ est égale à la matrice A^k.

Ainsi, si

$$exp(tA) = \beta_{\ell-1}(t)A^{\ell-1} + \cdots \beta_1(t)A + \beta_0(t)\mathrm{Id}_n,$$

alors

$$\forall k \in \mathbb{N},\ A^k = \beta_{\ell-1}^{(k)}(0)A^{\ell-1} + \cdots \beta_1^{(k)}A + \beta_0^{(k)}(0)\mathrm{Id}_n.$$

Exemple 10 Poursuivons l'exemple 9 :

(1) Pour tout entier $k \geq 2$, on a $A^k = 2^{k-1}\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

Notez que l'on a bien $A = (exp(tA))'|_{t=0}$.

(2) Pour tout entier $k \geq 0$, on a

$$D^k = 3^{k+1}\begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} + 2^{k+1}\begin{pmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Notez que pour $k = 0$, on retrouve la matrice Id_n, et pour $k = 1$, on retrouve la matrice D.