1 Réduction des endomorphismes, I

Soit $K = \mathbb{R}$ ou $\mathbb{C}$.

1.1 Motivation

Soit $n \geq 1$, $A \in M_n(K)$. On considère les vecteurs colonnes $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ et $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ dans K^n. On s'intéresse aux solutions de l'équation $y = Ax$ pour $x \in K^n$.

Un cas particulièrement simple se présente lorsque la matrice A est diagonalisable.

Definition 1.1. Une matrice $A \in M_n(K)$ est dite diagonalisable s'il existe une matrice inversible $P \in GL_n(K)$ telle que $P^{-1}AP = D$ soit une matrice diagonale, c'est-à-dire de la forme $D = Diag(\lambda_1, \ldots, \lambda_n)$, où $\lambda_i \in K$ pour $i = 1, \ldots, n$.

Dans ce cas, l'équation $y = Ax$ est équivalente à un système plus simple. Posons $y' = P^{-1}y$ et $x' = P^{-1}x$. Alors $y = Py'$ et $x = Px'$. L'équation $y = Ax$ devient $Py' = A(Px) = (AP)x = (APP^{-1})Px = (PDP^{-1})Px = PDx'$. En multipliant à gauche par P^{-1}, on obtient $y' = Dx'$.

L'ensemble des solutions $\{x' \in K^n \mid y' = Dx'\}$ est facile à déterminer. Si toutes les valeurs diagonales λ_i sont non nulles ($\lambda_i \neq 0$ pour $i = 1, \ldots, n$), alors pour chaque composante i, on a $y'_i = \lambda_i x'_i$, ce qui implique $x'_i = \frac{y'_i}{\lambda_i}$. Ainsi, $x' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} y'_1/\lambda_1 \\ \vdots \\ y'_n/\lambda_n \end{pmatrix}$.

Definition 1.2. Une matrice $A \in M_n(K)$ est dite trigonalisable s'il existe une matrice inversible $P \in GL_n(K)$ telle que $P^{-1}AP = T$ soit une matrice triangulaire supérieure.

Si $A = (a_{ij})$ est triangulaire supérieure, alors résoudre $Ax = y$ se fait facilement par substitution en commençant par la première ligne si les éléments diagonaux sont non nuls. Par exemple, si $A = (a_{ij})$ est triangulaire supérieure, alors la première équation est $a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1$. Si $a_{11} \neq 0$, on peut exprimer x_1 en fonction des autres variables et de y_1: $x_1 = \frac{1}{a_{11}}(y_1 - a_{12}x_2 - \cdots - a_{1n}x_n)$. En particulier, si on cherche à résoudre $Ax = y$ lorsque A est triangulaire supérieure et $a_{ii} \neq 0$ pour tout i, on peut trouver $x_1 = y_1/a_{11}$ (si $a_{11} \neq 0$), puis $x_2 = \frac{1}{a_{22}}(y_2 - a_{21}x_1)$ (si $a_{22} \neq 0$), et ainsi de suite.

Le but de la réduction d'endomorphismes est, étant donnée une matrice $A \in M_n(K)$, de trouver une matrice inversible $P \in GL_n(K)$ telle que la matrice $B = P^{-1}AP$, qui est semblable à A, ait une forme particulièrement simple, comme une matrice diagonale ou triangulaire supérieure.

Nous allons énoncer les théorèmes suivants (que nous allons démontrer dans la suite du cours) :

Theorem 1.3. Soit $K = \mathbb{R}$ et $A \in M_n(\mathbb{R})$ une matrice symétrique. Alors A est diagonalisable.

Theorem 1.4. Soit $K = \mathbb{C}$ et $A \in M_n(\mathbb{C})$. Alors A est trigonalisable.

Remark 1.5. Si une matrice A est diagonalisable, on peut calculer ses puissances A^i pour $i \geq 0$ de manière efficace. Si A est diagonalisable, il existe $P \in GL_n(K)$ telle que $P^{-1}AP = D =$

$Diag(\lambda_1, \ldots, \lambda_n)$ est une matrice diagonale. Alors $A = PDP^{-1}$. Pour calculer A^i, on a :

$$\begin{aligned} A^i = (PDP^{-1})^i &= \underbrace{(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})}_{i \text{ fois}} \\ &= PD\underbrace{(P^{-1}P)}_{I}D\underbrace{(P^{-1}P)}_{I}\cdots\underbrace{(P^{-1}P)}_{I}DP^{-1} \\ &= PD^iP^{-1} \end{aligned}$$

où $D^i = Diag(\lambda_1^i, \ldots, \lambda_n^i)$.

1.2 Application aux systèmes d'équations différentielles linéaires

Considérons une application aux systèmes d'équations différentielles linéaires. Soit $I \subseteq \mathbb{R}$ un intervalle. Si $x : I \to \mathbb{C}$ est une fonction de classe C^1, c'est-à-dire dérivable avec dérivée continue, on note $x' = \frac{dx}{dt}$.

Remark 1.6. Soit $\lambda \in \mathbb{C}$. L'équation différentielle $x' = \lambda x$ pour une fonction $x : I \to \mathbb{C}$ a pour solutions les fonctions de la forme $x(t) = ce^{\lambda t}$, où $c \in \mathbb{C}$ est une constante arbitraire. En effet, si $x(t) = ce^{\lambda t}$, alors $x'(t) = c\lambda e^{\lambda t} = \lambda(ce^{\lambda t}) = \lambda x(t)$. Réciproquement, si $x'(t) = \lambda x(t)$, considérons $y(t) = x(t)e^{-\lambda t}$. Alors, en dérivant par rapport à t:

$$\begin{aligned} y'(t) &= \frac{d}{dt}(x(t)e^{-\lambda t}) = x'(t)e^{-\lambda t} + x(t)(-\lambda)e^{-\lambda t} \\ &= (x'(t) - \lambda x(t))e^{-\lambda t} = 0 \end{aligned}$$

car $x'(t) = \lambda x(t)$. Donc, $y : I \to \mathbb{C}$ est une fonction constante, disons $y(t) = c$ pour une constante $c \in \mathbb{C}$. Alors $x(t)e^{-\lambda t} = c$, d'où $x(t) = ce^{\lambda t}$.

Considérons un système d'équations différentielles linéaires de la forme:

$$\begin{cases} x_1'(t) = a_{11}x_1(t) + \cdots + a_{1n}x_n(t) \\ \vdots \\ x_n'(t) = a_{n1}x_1(t) + \cdots + a_{nn}x_n(t) \end{cases}$$

où les coefficients $a_{ij} \in \mathbb{C}$ sont constants, et les $x_i : I \to \mathbb{C}$ sont des fonctions de classe C^1 pour $i \in \{1, \ldots, n\}$.

Posons $A = (a_{ij})_{1 \le i,j \le n} \in M_n(\mathbb{C})$, et $X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$, $X'(t) = \begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix}$. Alors, $X : I \to \mathbb{C}^n$ est une fonction vectorielle de classe C^1 avec dérivée $X'(t)$, et le système d'équations différentielles s'écrit sous forme matricielle:

$$X'(t) = AX(t)$$

Supposons que la matrice A est diagonalisable. Alors il existe une matrice inversible $P \in GL_n(\mathbb{C})$ et une matrice diagonale $D = Diag(\lambda_1, \ldots, \lambda_n)$ telles que $A = PDP^{-1}$. Posons $Y(t) = P^{-1}X(t)$. Alors $X(t) = PY(t)$, et $X'(t) = PY'(t)$ puisque P est une matrice constante (ne dépend pas de t). En substituant dans l'équation $X'(t) = AX(t)$, on obtient $PY'(t) = A(PY(t)) = (AP)Y(t) = (PDP^{-1})(PY(t)) = PDY(t)$. En multipliant à gauche par P^{-1}, on obtient $Y'(t) = DY(t)$.

Si $Y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}$, et $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, alors l'équation $Y'(t) = DY(t)$ se décompose en n équations différentielles indépendantes:

$$\begin{cases} y_1'(t) = \lambda_1 y_1(t) \\ \vdots \\ y_n'(t) = \lambda_n y_n(t) \end{cases}$$

On sait que la solution générale de chaque équation $y_i'(t) = \lambda_i y_i(t)$ est $y_i(t) = c_i e^{\lambda_i t}$, où $c_i \in \mathbb{C}$ est une constante arbitraire. Donc, $Y(t) = \begin{pmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix}$.

Comme $X(t) = PY(t)$, on obtient la solution générale $X(t) = P \begin{pmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix} = \sum_{j=1}^n c_j P_j e^{\lambda_j t}$, où $P = (P_1 | \cdots | P_n)$ et P_j est la j-ème colonne de la matrice P.

Remark 1.7. Soit $d \geq 1$ et soient $a_0, \dots, a_{d-1} \in \mathbb{C}$. Considérons une équation différentielle linéaire d'ordre d à coefficients constants pour une fonction $f : I \to \mathbb{C}$:

$$f^{(d)}(t) + \sum_{i=0}^{d-1} a_i f^{(i)}(t) = 0$$

où $f^{(0)} = f, f^{(1)} = f', \dots, f^{(d)} = (f^{(d-1)})'$. Posons $X(t) = \begin{pmatrix} f^{(0)}(t) \\ f^{(1)}(t) \\ \vdots \\ f^{(d-1)}(t) \end{pmatrix} = \begin{pmatrix} f(t) \\ f'(t) \\ \vdots \\ f^{(d-1)}(t) \end{pmatrix}$. Alors, $X'(t) = \begin{pmatrix} f'(t) \\ f''(t) \\ \vdots \\ f^{(d)}(t) \end{pmatrix}$. Le système d'équations différentielles $X'(t) = AX(t)$ pour la matrice compagnon

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{d-1} \end{pmatrix}$$

est équivalent à l'équation différentielle d'ordre d. En effet, si $X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_d(t) \end{pmatrix}$, le système $X'(t) = AX(t)$ s'écrit:

$$\begin{cases} x_1'(t) = x_2(t) \\ x_2'(t) = x_3(t) \\ \vdots \\ x_{d-1}'(t) = x_d(t) \\ x_d'(t) = -a_0 x_1(t) - a_1 x_2(t) - \cdots - a_{d-1} x_d(t) \end{cases}$$

Si on pose $x_1(t) = f(t)$, alors $x_2(t) = x_1'(t) = f'(t)$, $x_3(t) = x_2'(t) = f''(t)$, ..., $x_d(t) = f^{(d-1)}(t)$, et la dernière équation devient $x_d'(t) = f^{(d)}(t) = -a_0 f(t) - a_1 f'(t) - \cdots - a_{d-1} f^{(d-1)}(t)$, ce qui est précisément l'équation différentielle d'ordre d.

Example 1.8. Cherchons les solutions de l'équation différentielle $f''(t) - f(t) = 0$, ou $f''(t) + (-1)f(t) = 0$. Ici $d = 2$, $a_0 = -1$, $a_1 = 0$. La matrice compagnon est $A =$

$\begin{pmatrix} 0 & 1 \\ -a_0 & -a_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Considérons la matrice de passage $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in GL_2(\mathbb{R})$. Son inverse est $P^{-1} = \frac{1}{\det(P)} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{-2} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. On calcule $P^{-1}AP = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = D = Diag(1, -1)$.

Donc, chaque solution $X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} f(t) \\ f'(t) \end{pmatrix}$ est une combinaison linéaire de $P \begin{pmatrix} e^t \\ 0 \end{pmatrix} = \begin{pmatrix} e^t \\ e^t \end{pmatrix}$ et $P \begin{pmatrix} 0 \\ e^{-t} \end{pmatrix} = \begin{pmatrix} e^{-t} \\ -e^{-t} \end{pmatrix}$. La solution générale est donc de la forme $X(t) = P \begin{pmatrix} c_1 e^t \\ c_2 e^{-t} \end{pmatrix} = c_1 \begin{pmatrix} e^t \\ e^t \end{pmatrix} + c_2 \begin{pmatrix} e^{-t} \\ -e^{-t} \end{pmatrix} = \begin{pmatrix} c_1 e^t + c_2 e^{-t} \\ c_1 e^t - c_2 e^{-t} \end{pmatrix}$ pour des constantes $c_1, c_2 \in \mathbb{C}$. La première composante donne $f(t) = c_1 e^t + c_2 e^{-t}$, qui est la solution générale de $f''(t) - f(t) = 0$. Les solutions sont des combinaisons linéaires de e^t et e^{-t}.

2 Bases adaptées

2.1 Définition

Soit E un K-espace vectoriel de dimension finie n.

Definition 2.1. Soient $F_1, \dots, F_r \subseteq E$ des sous-espaces vectoriels tels que $E = F_1 \oplus \cdots \oplus F_r$ est une somme directe. Une base B de E est dite adaptée à la décomposition $E = F_1 \oplus \cdots \oplus F_r$ si elle est de la forme $B = B_1 \sqcup \cdots \sqcup B_r$, où B_i est une base de F_i pour chaque $i = 1, \dots, r$.

Lemma 2.2. Soit $f \in L(E)$ un endomorphisme de E, et soit $B = B_1 \sqcup \cdots \sqcup B_r$ une base adaptée à une décomposition $E = F_1 \oplus \cdots \oplus F_r$. Les propositions suivantes sont équivalentes :

1. Chaque sous-espace F_i est stable par f, c'est-à-dire $f(F_i) \subseteq F_i$ pour tout $i = 1, \dots, r$.
2. La matrice $A = Mat_B(f)$ de f dans la base B est une matrice diagonale par blocs.

Par matrice diagonale par blocs, on entend une matrice de la forme

$$A = \begin{pmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & A_{rr} \end{pmatrix}$$

où chaque bloc A_{ii} est une matrice carrée de taille $dim(F_i) \times dim(F_i)$.

Preuve (Démonstration (pour simplifier, supposons $r = 2$))**.** Soient F_1, F_2 des sous-espaces vectoriels tels que $E = F_1 \oplus F_2$. Soit $B = B_1 \sqcup B_2$ une base adaptée, où $B_1 = (e_1, \dots, e_{d_1})$ est une base de F_1 et $B_2 = (e_{d_1+1}, \dots, e_n)$ est une base de F_2, avec $d_i = dim(F_i)$, $d_1 + d_2 = n = dim(E)$. La matrice de f dans la base B est de la forme

$$A = Mat_B(f) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

où A_{11} est de taille $d_1 \times d_1$, A_{12} est de taille $d_1 \times d_2$, A_{21} est de taille $d_2 \times d_1$, et A_{22} est de taille $d_2 \times d_2$.

Le bloc A_{ij} est la matrice représentative dans les bases B_j et B_i de l'application linéaire $f_{ij} : F_j \xrightarrow{incl} E \xrightarrow{f} E \xrightarrow{pr_{F_i}} F_i$, où $pr_{F_i} : E = F_1 \oplus F_2 \to F_i$ est la projection sur F_i parallèlement à F_{3-i}.

Si F_1 est stable par f, alors pour tout $x \in F_1$, $f(x) \in F_1$. Dans ce cas, $pr_{F_2}(f(x)) = 0$ pour tout $x \in F_1$, donc $f_{21} = pr_{F_2} \circ f \circ incl_{F_1} = 0$. La matrice de l'application nulle est la matrice nulle, donc

$A_{21} = 0$. De même, F_2 est stable par f si et seulement si $f_{12} = 0$, c'est-à-dire $A_{12} = 0$.

Donc, si F_1 et F_2 sont stables par f, alors $A_{12} = 0$ et $A_{21} = 0$, et $A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$ est diagonale par blocs. Réciproquement, si $A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$, alors $A_{21} = 0$ et $A_{12} = 0$, donc $f_{21} = 0$ et $f_{12} = 0$. Ceci implique que $f(F_1) \subseteq F_1$ et $f(F_2) \subseteq F_2$. □

3 Diagonalisation et bases adaptées

Theorem 3.1. Soit $A \in M_n(K)$. On note $u_A : K^n \to K^n$ l'endomorphisme de K^n défini par $u_A(x) = Ax$. Alors $u_A \in L(K^n)$. Réciproquement, tout endomorphisme $f \in L(K^n)$ est de la forme $f = u_{Mat_B(f)}$, si B est la base canonique de K^n.

Une matrice $A \in M_n(K)$ est diagonalisable si et seulement s'il existe une décomposition de K^n en somme directe de sous-espaces vectoriels de dimension 1 stables par u_A, disons $K^n = F_1 \oplus \cdots \oplus F_n$ avec $dim(F_i) = 1$ et $u_A(F_i) \subseteq F_i$ pour tout $i = 1, \ldots, n$.

Preuve (Démonstration)**.** Supposons que A est diagonalisable. Alors il existe $P \in GL_n(K)$ telle que $P^{-1}AP = D = Diag(\lambda_1, \ldots, \lambda_n)$ est diagonale. Alors $AP = PD$. Soit $P = (P_1|\cdots|P_n)$, où P_i est la i-ème colonne de P, qui est un vecteur de K^n. Alors $AP = (AP_1|\cdots|AP_n)$. D'autre part, $PD = (P_1|\cdots|P_n)\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = (\lambda_1 P_1|\cdots|\lambda_n P_n)$. L'égalité $AP = PD$ implique donc $AP_i = \lambda_i P_i$ pour tout $i = 1, \ldots, n$.

Soit $F_i = Vect(P_i)$ le sous-espace vectoriel engendré par P_i. Puisque $AP_i = \lambda_i P_i$, on a $u_A(P_i) = \lambda_i P_i \in Vect(P_i) = F_i$. Donc $u_A(F_i) \subseteq F_i$, et F_i est stable par u_A. De plus, puisque $P \in GL_n(K)$ est inversible, ses colonnes $(P_1, \ldots, P_n)$ forment une famille libre, et donc une base de K^n. Ainsi, $K^n = Vect(P_1) \oplus \cdots \oplus Vect(P_n) = F_1 \oplus \cdots \oplus F_n$, et $dim(F_i) = dim(Vect(P_i)) = 1$ car $P_i \neq 0$ (colonnes d'une matrice inversible).

Réciproquement, supposons qu'il existe une décomposition $K^n = F_1 \oplus \cdots \oplus F_n$ avec $dim(F_i) = 1$ et $u_A(F_i) \subseteq F_i$ pour tout i. Puisque $dim(F_i) = 1$, on peut choisir un vecteur non nul $C_i \in F_i$ tel que $F_i = Vect(C_i)$. La condition $u_A(F_i) \subseteq F_i$ implique $u_A(C_i) \in F_i = Vect(C_i)$, donc il existe un scalaire $\lambda_i \in K$ tel que $u_A(C_i) = \lambda_i C_i$, c'est-à-dire $AC_i = \lambda_i C_i$. Soit $P = (C_1|\cdots|C_n)$ la matrice dont les colonnes sont les vecteurs $C_1, \ldots, C_n$. Puisque $K^n = F_1 \oplus \cdots \oplus F_n = Vect(C_1) \oplus \cdots \oplus Vect(C_n)$ est une somme directe et $dim(K^n) = n = \sum dim(F_i) = n$, la famille $(C_1, \ldots, C_n)$ est une base de K^n, et $P = (C_1|\cdots|C_n) \in GL_n(K)$ est inversible.

Considérons la matrice $P^{-1}AP$. On a $AP = (AC_1|\cdots|AC_n) = (\lambda_1 C_1|\cdots|\lambda_n C_n)$. Et $PD = (C_1|\cdots|C_n)Diag(\lambda_1, \ldots, \lambda_n) = (\lambda_1 C_1|\cdots|\lambda_n C_n)$. Donc $AP = PD$, et en multipliant à gauche par P^{-1}, on obtient $P^{-1}AP = D = Diag(\lambda_1, \ldots, \lambda_n)$. La matrice $P^{-1}AP$ est diagonale, donc A est diagonalisable.

□

4 Vecteurs propres et valeurs propres

Definition 4.1. Un vecteur propre d'un endomorphisme $f \in L(E)$ est un vecteur $v \in E$ vérifiant les deux propriétés suivantes :

1. $v \neq 0_E$ (non nul)
2. Il existe un scalaire $\lambda \in K$ tel que $f(v) = \lambda v$.

Ceci est équivalent à dire que $Vect(v)$ est un sous-espace de dimension 1 stable par f.

Remark 4.2. Si v est un vecteur propre de f, et si $f(v) = \lambda v$, alors le scalaire λ est appelé valeur propre de f associé au vecteur propre v. La valeur propre λ est uniquement déterminée par le vecteur propre $v \neq 0$.

Definition 4.3. Le spectre de f, noté $Sp(f)$, est l'ensemble des valeurs propres de f :

$$Sp(f) = \{\lambda \in K \mid \exists v \neq 0 \text{ vecteur propre de } f \text{ tel que } f(v) = \lambda v\}$$