

LICENCE MATHÉMATIQUES ET INFORMATIQUE
DEUXIÈME ANNÉE

Unité d'enseignement LMI 3.02

ALGÈBRE ET GÉOMÉTRIE 2

Françoise GEANDIER

Université Henri Poincaré Nancy I
Département de Mathématiques

Table des matières

I Rappels d'algèbre linéaire	1
1. Espaces vectoriels.	
2. Applications linéaires - Matrices.	
3. Déterminants.	
4. Espaces vectoriels de polynômes.	
II Réduction des endomorphismes	13
1. Sous-espaces propres.	
2. Diagonalisation.	
3. Polynômes d'endomorphismes.	
4. Trigonalisation.	
III Exponentielles de matrices - Systèmes d'équations différentielles linéaires	27
1. Normes sur K^n et $\mathcal{M}_n(K)$.	
2. Exponentielles de matrices.	
3. Systèmes d'équations différentielles linéaires.	
4. Equations différentielles linéaires homogènes d'ordre n à coefficients constants.	
IV Espaces euclidiens	37
1. Formes bilinéaires symétriques - Formes quadratiques.	
2. Espaces euclidiens.	
3. Orthogonalité.	
4. Bases orthogonales et orthonormées.	
5. Projection orthogonale - Symétrie orthogonale.	
6. Problèmes de moindres carrés.	

V Espaces hermitiens 51

1. Formes sesquilinéaires hermitiennes - Formes quadratiques.
2. Espaces hermitiens.
3. Orthogonalité.
4. Bases orthogonales et orthonormées.

VI Réduction des matrices normales 57

1. Matrice adjointe.
2. Matrices normales.
3. Cas des matrices réelles.

I RAPPELS D'ALGÈBRE LINÉAIRE

1. Espaces vectoriels

1.1 \mathbb{R}^n et \mathbb{C}^n

On considère \mathbb{R} le corps des nombres réels et \mathbb{C} le corps des nombres complexes.

On appelle **vecteur** à n composantes réelles tout n -uplet $x = (x_1, x_2, \dots, x_n)$ où x_1, x_2, \dots, x_n sont des réels et on note \mathbb{R}^n l'ensemble de ces vecteurs. Le plus souvent, on utilisera la notation dite vecteur colonne suivante :

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

De la même façon, on définit \mathbb{C}^n l'ensemble des vecteurs à n composantes complexes. Dans la suite, K désignera l'un des corps \mathbb{R} ou \mathbb{C} et K^n l'ensemble des vecteurs à n composantes dans K . On notera 0 le vecteur nul de K^n , i.e le vecteur dont toutes les composantes sont nulles.

On définit deux opérations sur K^n

$$* \text{ l'addition : si } x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ et } y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \text{ on définit } z = x+y \text{ par } z = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} ;$$

$$* \text{ la multiplication par un scalaire } \lambda, \text{ i.e un élément de } K, \lambda x = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

On remarque que pour tout vecteur x de E , on peut écrire

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

On a ainsi exhibé n vecteurs particuliers de E :

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

et ainsi x s'écrit sous la forme $x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$.

On peut généraliser ces propriétés de K^n en définissant la notion d'espace vectoriel sur K :

1.2 Espaces vectoriels

Soit E un ensemble non vide muni d'une addition et d'une multiplication par les scalaires ; on dit que E est un **espace vectoriel** sur K s'il vérifie les conditions suivantes :

Si x, y et z sont des éléments de E et si λ et μ sont des scalaires

- a) l'addition est commutative : $x + y = y + x$;
- b) l'addition est associative : $(x + y) + z = x + (y + z)$;
- c) il existe un élément 0_E de E tel que $x + 0_E = x$ pour tout $x \in E$;
- d) tout élément x de E possède un opposé $-x$ dans E : $x + (-x) = x - x = 0_E$;
- e) $\lambda(x + y) = \lambda x + \lambda y$;
- f) $(\lambda + \mu)x = \lambda x + \mu x$ et $\lambda(\mu x) = (\lambda\mu)x$;
- g) $1x = x$

Les éléments d'un espace vectoriel E sont appelés **vecteurs**.

Exemples

K^n est évidemment un espace vectoriel sur K . Le singleton $\{0\}$ est un espace vectoriel sur K .

L'ensemble des applications d'un ensemble A à valeurs dans \mathbb{R} est un espace vectoriel sur \mathbb{R} :

Si f et g sont deux applications de A dans \mathbb{R} et si $\lambda \in \mathbb{R}$, on définit $f + g$ et λf par

$$\forall x \in A, (f + g)(x) = f(x) + g(x) \text{ et } (\lambda f)(x) = \lambda f(x).$$

Soit E un espace vectoriel sur K et soit F un sous-ensemble non vide de E ; on dit que F est un **sous-espace vectoriel** de E s'il vérifie les propriétés suivantes :

- * si x et $y \in F$, alors $x + y \in F$;
- * si $x \in F$ et $\lambda \in K$ alors $\lambda x \in F$.

Ces deux conditions sont équivalentes à l'unique condition suivante :

- * si $x, y \in F$ et si $\lambda, \mu \in K$, alors $\lambda x + \mu y \in F$.

Considérons p vecteurs u_1, u_2, \dots, u_p de E et notons $Vect(u_1, u_2, \dots, u_p)$ l'ensemble des combinaisons linéaires de u_1, u_2, \dots, u_p , i.e

$$Vect(u_1, u_2, \dots, u_p) = \{\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_p u_p / \lambda_1, \lambda_2, \dots, \lambda_p \in K\}.$$

Alors $Vect(u_1, u_2, \dots, u_p)$ est un sous-espace vectoriel de E .

Soit F un sous-espace vectoriel de E et soient p vecteurs u_1, u_2, \dots, u_p de E ; on dit que u_1, u_2, \dots, u_p **engendrent** F ou forment un **système générateur** de F si $F = Vect(u_1, u_2, \dots, u_p)$.

On dit que p vecteurs u_1, u_2, \dots, u_p de E sont **linéairement indépendants** ou forment un **système libre** si pour tous $\lambda_1, \lambda_2, \dots, \lambda_p \in K$, on a :

$$\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_p u_p = 0 \implies \lambda_1 = \lambda_2 = \dots = \lambda_p = 0.$$

On appelle **base** de E toute famille de vecteurs linéairement indépendants qui engendrent E . On montre que tout espace vectoriel non réduit à $\{0\}$ possède au moins une base ; si E possède une base comprenant un nombre fini n de vecteurs, alors toute base de E comprend également n vecteurs : cet entier n est appelé dimension de E et noté $\dim E$, et on dit que E est de dimension finie. Si E est de dimension finie, alors tout sous-espace vectoriel F de E est lui aussi de dimension finie et $\dim F \leq \dim E$.

En particulier, l'espace vectoriel K^n est de dimension n : la famille (e_1, e_2, \dots, e_n) définie ci-dessus est une base de E appelée base canonique. Par convention, on dit que la dimension de l'espace vectoriel $\{0\}$ est nulle.

On appelle **rang** d'un système (u_1, u_2, \dots, u_p) de vecteurs de E la dimension du sous-espace vectoriel $\text{Vect}(u_1, u_2, \dots, u_p)$: ce rang est $\leq p$.

Si (u_1, u_2, \dots, u_n) est une base de E , alors tout vecteur x de s'écrit de manière unique sous la forme

$$x = x_1 u_1 + x_2 u_2 + \dots + x_n u_n$$

Les éléments x_1, x_2, \dots, x_n de K sont appelés **coordonnées** de x dans la base (u_1, u_2, \dots, u_n) .

En outre, si E est de dimension n , toute famille libre (u_1, u_2, \dots, u_p) de p vecteurs de E avec $p < n$ peut être complétée en une base de E , i.e il existe $n - p$ vecteurs $(u_{p+1}, u_{p+2}, \dots, u_n)$ de E tels que (u_1, u_2, \dots, u_n) est une base de E .

Soit E de dimension n et soient F et G deux sous-espaces vectoriels de E ; si $F \subset G$ et si $\dim F = \dim G$, alors on a $F = G$.

1.3 Somme de sous-espaces vectoriels

Soit E un espace vectoriel sur K de dimension finie et soient F et G deux sous-espaces vectoriels de E ; on appelle **somme** de F et G et on note $F + G$ l'ensemble défini par

$$F + G = \{x + y \mid x \in F, y \in G\}.$$

on vérifie facilement que $F + G$ est un sous-espace vectoriel de E ; si (u_1, u_2, \dots, u_p) est un système générateur de F et (v_1, v_2, \dots, v_k) un système générateur de G , alors la réunion des deux systèmes $(u_1, u_2, \dots, u_p, v_1, v_2, \dots, v_k)$ est un système générateur de $F + G$.

On dit que F et G sont en **somme directe** si $F \cap G = \{0\}$; on note alors

$$F + G = F \oplus G.$$

Si F et G sont en somme directe, alors si (u_1, u_2, \dots, u_p) est une base de F et (v_1, v_2, \dots, v_k) est une base de G , $(u_1, u_2, \dots, u_p, v_1, v_2, \dots, v_k)$ est une base de $F \oplus G$; on en déduit que $\dim(F \oplus G) = \dim F + \dim G$.

On a deux autres définitions équivalentes pour la somme directe de deux sous-espaces vectoriels :

* F et G sont en somme directe si et seulement si pour tout $z \in F + G$, il existe un **unique** couple de vecteurs (x, y) où $x \in F$ et $y \in G$ tel que $z = x + y$. Autrement dit la condition $F \cap G = \{0\}$ est équivalente à l'unicité d'écriture.

* F et G sont en somme directe si et seulement si $\dim(F + G) = \dim F + \dim G$.

Si F et G sont en somme directe et si $F \oplus G = E$, on dit que F et G sont **supplémentaires** dans E ; en d'autres termes, F et G sont supplémentaires dans E si et seulement si

$$\dim F + \dim G = \dim E \text{ et } F \cap G = \{0\}.$$

2. Applications linéaires-matrices

2.1 Applications linéaires

Soient E un espace vectoriel de dimension n et F un espace vectoriel de dimension p sur K ; une application f de E dans F est dite **linéaire** si elle vérifie les deux conditions suivantes

- * $\forall x, y \in E$, on a $f(x + y) = f(x) + f(y)$;
- * $\forall x \in E, \forall \lambda \in K$, on a $f(\lambda x) = \lambda f(x)$.

Ces deux conditions pouvant être rassemblées en une seule :

- * $\forall x, y \in E, \forall \lambda, \mu \in K$, on a $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$.

On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F ; on peut définir une addition et une multiplication par les scalaires sur $\mathcal{L}(E, F)$ de la façon suivante :

- * si f et $g \in \mathcal{L}(E, F)$, on pose $(f + g)(x) = f(x) + g(x)$ pour tout $x \in E$.
- * si $f \in \mathcal{L}(E, F)$ et si $\lambda \in K$, on pose $(\lambda f)(x) = \lambda f(x)$ pour tout $x \in E$.

$\mathcal{L}(E, F)$ est alors un espace vectoriel sur K , de dimension $\dim E \times \dim F$.

D'autre part, si f est une application linéaire de E dans F et g est une application linéaire de F dans G , alors l'application composée $g \circ f$ de E dans G est linéaire.

Soit f une application linéaire de E dans F ; on appelle **noyau** de f et on note $\ker f$ le sous-ensemble de E suivant

$$\ker f = \{x \in E / f(x) = 0\}.$$

On a

$$f \text{ est injective} \iff \ker f = \{0\}.$$

On appelle **image** de f et on note $\text{Im} f$ le sous-ensemble de F suivant

$$\text{Im} f = f(E) = \{f(x) / x \in E\}.$$

On a

$$f \text{ est surjective} \iff \text{Im} f = F.$$

De plus $\ker f$ est un sous-espace vectoriel de E , $\text{Im} f$ est un sous-espace vectoriel de F ; on appelle **rang** de f et on note $rg(f)$ la dimension de $\text{Im} f$, i.e le rang du système $(f(e_1), f(e_2), \dots, f(e_n))$ pour toute base (e_1, e_2, \dots, e_n) de E . On a le théorème suivant, appelé **théorème du rang**

$$\dim(\ker f) + rg(f) = \dim E = n.$$

En particulier, si $n = p$, on a les équivalences

$$f \text{ est injective} \iff f \text{ est surjective} \iff f \text{ est bijective} \iff \ker f = \{0\} \iff rg(f) = n.$$

2.2 Matrices

Soit f une application linéaire de E dans F et choisissons une base $\mathcal{B} = (e_1, e_2, \dots, e_n)$ de E et une base $\mathcal{B}' = (u_1, u_2, \dots, u_p)$ de F ; alors, pour tout $x = x_1e_1 + x_2e_2 + \dots + x_n e_n$, on a :

$$f(x) = f(x_1e_1 + x_2e_2 + \dots + x_n e_n) = x_1f(e_1) + x_2f(e_2) + \dots + x_nf(e_n).$$

En d'autres termes, pour calculer $f(x)$ en fonction des coordonnées de x dans la base (e_1, e_2, \dots, e_n) , il suffit de connaître la valeur de $f(e_1), f(e_2), \dots, f(e_n)$. Or $f(e_1), f(e_2), \dots, f(e_n)$ sont des éléments de F donc s'écrivent de manière unique dans la base (u_1, u_2, \dots, u_p)

$$f(e_j) = \sum_{i=1}^p a_{ij}u_i \text{ pour } 1 \leq j \leq n.$$

où les éléments a_{ij} sont dans le corps K .

Ainsi on associe à f un tableau à p lignes et n colonnes d'éléments de K

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix}$$

appelé **matrice** représentative de l'application linéaire f dans les bases \mathcal{B} et \mathcal{B}' et notée $M_{\mathcal{B},\mathcal{B}'}(f)$. Les p composantes du j ème vecteur colonne de la matrice $A = M_{\mathcal{B},\mathcal{B}'}(f)$ sont donc les coordonnées du vecteur $f(e_j)$ dans la base (u_1, u_2, \dots, u_p) .

Pour tout vecteur $x = x_1e_1 + x_2e_2 + \dots + x_n e_n$ de E , on calcule facilement les coordonnées de $y = f(x)$ dans la base (u_1, u_2, \dots, u_p) : $y = y_1u_1 + y_2u_2 + \dots + y_pu_p$ où

$$\begin{cases} y_1 & = & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 & = & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots & & \vdots \\ y_p & = & a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n \end{cases}$$

ce que l'on note $Y = AX$ où

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} \text{ et } X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

On note $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$: a_{ij} est appelé **terme d'indice** (i, j) de la matrice A .

On notera $\mathcal{M}_{p,n}(K)$ l'ensemble des matrices à p lignes et n colonnes à coefficients dans K et simplement $\mathcal{M}_n(K)$ l'ensemble des matrices à n lignes et n colonnes à coefficients dans K , appelées matrices **carrées** d'ordre n à coefficients dans K .

Lorsque f est un endomorphisme de E , c'est-à-dire une application linéaire de E dans lui-même, alors on prend la même base $\mathcal{B} = (e_1, e_2, \dots, e_n)$ pour l'espace de départ et l'espace

d'arrivée et on considère la matrice de f dans la base \mathcal{B} , notée $M_{\mathcal{B}}(f)$, qui est alors une matrice carrée d'ordre n ; en particulier, l'application identique

$$\begin{aligned} Id: E &\rightarrow E \\ x &\mapsto x \end{aligned}$$

a pour matrice dans n'importe quelle base de E , la matrice

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

appelée matrice identité d'ordre n . On note $I_n = I$ lorsqu'il n'y a pas d'ambiguïté sur l'ordre de la matrice.

Soit A une matrice de $\mathcal{M}_{p,n}(K)$: on appelle **noyau** de A et on note $\ker A$ le sous-espace vectoriel de K^n défini par

$$\ker A = \{X \in K^n / AX = 0\}.$$

et on appelle **image** de A et on note $\text{Im}A$ le sous-espace vectoriel de K^p engendré par les vecteurs colonnes de A . On appelle **rang** de A et on note $\text{rg}(A)$ le rang du système formé par les vecteurs colonnes de A , i.e la dimension de $\text{Im}A$.

Si f est une application linéaire de E dans F et si A est la matrice de f dans des bases de E et F , alors on a $\dim \ker(f) = \dim \ker A$ et $\text{rg}(f) = \text{rg}(A)$ d'où le théorème du rang pour les matrices: si $A \in \mathcal{M}_{p,n}(K)$, on a:

$$\dim \ker A + \text{rg}(A) = \dim(E) = n.$$

De plus, on a

$$f \text{ est injective} \iff \ker f = \{0\} \iff \ker A = \{0\}.$$

et

$$f \text{ est surjective} \iff \text{Im}f = F \iff \text{Im}A = F \iff \text{rg}(f) = p \iff \text{rg}(A) = p.$$

et dans le cas où $n = p$ on a les équivalences

$$f \text{ est bijective} \iff \ker A = \{0\} \iff \text{rg}(A) = n.$$

2.3 Opérations sur les matrices

On peut définir une addition et une multiplication par les scalaires sur $\mathcal{M}_{p,n}(K)$ de la manière suivante:

* si $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ et $B = (b_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ sont dans $\mathcal{M}_{p,n}(K)$, on pose

$$A + B = C = (c_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$$

avec $c_{ij} = a_{ij} + b_{ij}$ pour tous $1 \leq i \leq p$, $1 \leq j \leq n$.

* si $\lambda \in K$, on pose $\lambda A = D = (d_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$

avec $d_{ij} = \lambda a_{ij}$ pour tous $1 \leq i \leq p, 1 \leq j \leq n$.

Muni de ces deux opérations, $\mathcal{M}_{p,n}(K)$ est alors un espace vectoriel sur K , de dimension $p \times n$.

De plus on peut définir la multiplication d'une matrice $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq m}}$ de $\mathcal{M}_{p,m}(K)$ par une matrice $B = (b_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ de $\mathcal{M}_{m,n}(K)$ de la manière suivante :

* on pose $AB = C$ où $C = (c_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ est la matrice de $\mathcal{M}_{p,n}(K)$ définie par

$$c_{ij} = \sum_{k=1}^m a_{ik} b_{kj} \text{ pour tous } 1 \leq i \leq p, 1 \leq j \leq n.$$

En particulier, si A et B sont deux matrices carrées d'ordre n , le produit $C = AB$ est une matrice carrée d'ordre n . On a de plus $AI = IA = A$.

Soit $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ une matrice de $\mathcal{M}_{p,n}(K)$: on appelle matrice **transposée** de A et on note tA la matrice de $\mathcal{M}_{n,p}(K)$ dont le terme d'indice (i,j) est donné par a_{ji} ; en d'autres termes tA a pour vecteurs colonnes les vecteurs lignes de A et réciproquement. On dit que A est symétrique si $A = {}^tA$ (ce qui nécessite que A soit une matrice carrée). On a les propriétés suivantes :

- * ${}^t(A + B) = {}^tA + {}^tB$;
- * ${}^t(\lambda A) = \lambda {}^tA$;
- * ${}^t(AB) = {}^tB {}^tA$.

On dit qu'une matrice carrée A d'ordre n est **diagonale** si tous ses termes en dehors de la diagonale principale sont nuls, i.e

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

On dit qu'une matrice $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ de $\mathcal{M}_{p,n}(K)$ est **triangulaire**

- **supérieure** si $a_{ij} = 0$ pour $i > j$;
- **inférieure** si $a_{ij} = 0$ pour $i < j$.

Ainsi, si A est carrée d'ordre n , A est triangulaire supérieure si

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ 0 & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

et A est triangulaire inférieure si

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \cdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{pmatrix}$$

On dit qu'une matrice carrée A d'ordre n est **inversible** s'il existe une matrice carrée B d'ordre n vérifiant $AB = BA = I$; alors la matrice B est unique et on la note $B = A^{-1}$. On montre qu'une matrice carrée A d'ordre n est inversible si et seulement si $rg(A) = n$ ou, ce qui est équivalent, $\ker A = \{0\}$.

Ces trois opérations définies sur les matrices correspondent à celles définies précédemment sur les applications linéaires, plus précisément, on a le théorème suivant :

Théorème

* Soient f et g deux applications linéaires de E dans F , λ un scalaire, \mathcal{B} une base de E et \mathcal{B}' une base de F ; alors

$$M_{\mathcal{B},\mathcal{B}'}(f + g) = M_{\mathcal{B},\mathcal{B}'}(f) + M_{\mathcal{B},\mathcal{B}'}(g).$$

et

$$M_{\mathcal{B},\mathcal{B}'}(\lambda f) = \lambda M_{\mathcal{B},\mathcal{B}'}(f).$$

* Soient f une application linéaire de E dans F , g une application linéaire de F dans G , \mathcal{B} une base de E , \mathcal{B}' une base de F et \mathcal{B}'' une base de G , alors

$$M_{\mathcal{B},\mathcal{B}''}(g \circ f) = M_{\mathcal{B}',\mathcal{B}''}(g)M_{\mathcal{B},\mathcal{B}'}(f).$$

En particulier, si f est un endomorphisme bijectif de E , considérons la matrice A de f dans une base $\mathcal{B} = (e_1, e_2, \dots, e_n)$ de E ; alors la matrice B de la bijection réciproque f^{-1} dans la même base \mathcal{B} vérifie évidemment

$$AB = BA = I$$

puisque $f \circ f^{-1} = f^{-1} \circ f = Id$. Donc $B = A^{-1}$, i.e $M_{\mathcal{B}}(f^{-1}) = (M_{\mathcal{B}}(f))^{-1}$.

Réciproquement, si la matrice A d'un endomorphisme f de E dans une base \mathcal{B} de E est inversible, alors $rg(f) = rg(A) = n$ donc f est bijective et alors $M_{\mathcal{B}}(f^{-1}) = (M_{\mathcal{B}}(f))^{-1}$; pour résumer, si f est un endomorphisme de E , alors on a :

$$f \text{ est bijective} \iff \exists \mathcal{B}, M_{\mathcal{B}}(f) \text{ est inversible} \iff \forall \mathcal{B}, M_{\mathcal{B}}(f) \text{ est inversible.}$$

2.4 Changement de base

Soient $\mathcal{B} = (e_1, e_2, \dots, e_n)$ et $\mathcal{B}' = (e'_1, e'_2, \dots, e'_n)$ deux bases de E ; on cherche à calculer les coordonnées d'un vecteur x de E dans la base \mathcal{B}' à partir des coordonnées de x dans la base \mathcal{B} : Soit P la matrice carrée d'ordre n obtenue en prenant comme vecteurs colonnes les coordonnées des vecteurs e'_1, e'_2, \dots, e'_n dans la base $\mathcal{B} = (e_1, e_2, \dots, e_n)$ i.e

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

$$e'_1 \quad e'_2 \quad \cdots \quad e'_n$$

Alors $rg(P) = rg(e'_1, e'_2, \dots, e'_n) = n$ donc P est inversible: P est appelée la matrice de passage de la base \mathcal{B} à la base \mathcal{B}' .

De plus, si x a pour coordonnées $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ dans la base \mathcal{B} et $X' = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$ dans la

base \mathcal{B}' , on a

$$X = PX' \text{ ou encore } X' = P^{-1}X.$$

Considérons maintenant une application linéaire f de E dans F , A sa matrice dans les bases \mathcal{B} et \mathcal{B}_1 et A' sa matrice dans les bases \mathcal{B}' et \mathcal{B}'_1 . Notons P la matrice de passage de la base \mathcal{B} à la base \mathcal{B}' et Q la matrice de passage de la base \mathcal{B}_1 à la base \mathcal{B}'_1 , alors on a

$$A' = Q^{-1}AP.$$

En effet, si X et X' désignent les vecteurs colonnes des coordonnées de x dans les bases \mathcal{B} et \mathcal{B}' et Y et Y' désignent les vecteurs colonnes des coordonnées de $f(x)$ dans les bases \mathcal{B}_1 et \mathcal{B}'_1 , alors $Y = AX$ et $Y' = A'X'$; or $X = PX'$ et $Y' = Q^{-1}Y$, donc $Y' = Q^{-1}Y = Q^{-1}AX = Q^{-1}APX'$ donc on a bien $A' = Q^{-1}AP$. On dit alors que les matrices A et A' sont **équivalentes**.

En particulier dans le cas où $n = p$, si A est la matrice de f dans la base \mathcal{B} et A' est la matrice de f dans la base \mathcal{B}' alors, si P désigne la matrice de passage de la base \mathcal{B} à la base \mathcal{B}' , on a

$$A' = P^{-1}AP.$$

On dit alors que les matrices A et A' sont **semblables**.

3 Déterminant

3.1 Permutations

On appelle **permutation** de $\{1, 2, \dots, n\}$ toute bijection de $\{1, 2, \dots, n\}$ sur lui-même et on note S_n l'ensemble des permutations de $\{1, 2, \dots, n\}$: S_n possède $n!$ éléments.

Si $\sigma \in S_n$, on note

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

On dit que σ est **paire** (resp. **impaire**) si un nombre pair (resp. impair) d'échanges entre deux éléments est nécessaire pour remettre $\sigma(1), \sigma(2), \dots, \sigma(n)$ dans l'ordre naturel.

Exemples :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \text{ est une permutation impaire de } \{1, 2, 3\}.$$

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} \text{ est une permutation paire de } \{1, 2, 3, 4\}.$$

La **signature** $\varepsilon(\sigma)$ d'une permutation $\sigma \in S_n$ est définie par :

$$\begin{aligned} \varepsilon(\sigma) &= 1 \text{ si } \sigma \text{ est paire;} \\ \varepsilon(\sigma) &= -1 \text{ si } \sigma \text{ est impaire.} \end{aligned}$$

3.2 Déterminants

Soit $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(K)$; on appelle **déterminant** de A l'élément de K défini par

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

ou encore

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$$

Exemple :

Si $n = 2$, S_2 possède deux éléments : la permutation identique qui est paire et la permutation échangeant 1 et 2 qui est impaire ; ainsi, si $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, on a

$$\det(A) = (+1)a_{11}a_{22} + (-1)a_{12}a_{21} = a_{11}a_{22} - a_{12}a_{21}.$$

$$\text{Si } A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \text{ on note } \det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

3.3 Propriétés

* Si on multiplie une colonne de la matrice A par un scalaire λ , $\det(A)$ est multiplié par λ ; en conséquence

$$\det(\lambda A) = \lambda^n \det(A).$$

* Si deux colonnes (ou deux lignes) de la matrice A sont proportionnelles, $\det(A) = 0$.

* Si on modifie l'ordre des colonnes (ou des lignes) par une permutation $\sigma \in S_n$, $\det(A)$ est multipliée par $\varepsilon(\sigma)$.

* Si on ajoute à une colonne (resp. une ligne) une combinaison linéaire des autres colonnes (resp. lignes), on ne change pas la valeur de $\det(A)$.

* $\det({}^tA) = \det(A)$ et $\det(AB) = \det(A)\det(B)$.

* Le déterminant d'une matrice triangulaire A est égal au produit de ses éléments diagonaux :

$$\det(A) = \prod_{i=1}^n a_{ii}.$$

3.4 Développement selon une colonne ou une ligne

Soit $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathcal{M}_n(K)$; on note A_{ij} la matrice d'ordre $n - 1$ obtenue en supprimant la i ème ligne et la j ème colonne de A , alors on a la formule :

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$$

appelée développement de $\det(A)$ selon la j ème colonne.

De même on a la formule du développement selon la i ème ligne :

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}).$$

3.5 Déterminants et matrices inversibles

Théorème

Soit $A \in \mathcal{M}_n(K)$; alors

$$A \text{ est inversible} \iff \det(A) \neq 0$$

et dans ce cas

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

On en déduit que deux matrices semblables ont même déterminant : en effet si $A = P^{-1}BP$, alors

$$\det(A) = \det(P^{-1}) \det(B) \det(P) = \frac{1}{\det(P)} \det(B) \det(P) = \det(B).$$

On peut donc définir le déterminant d'un endomorphisme f de E : c'est la valeur commune du déterminant de toute matrice représentant f dans une base quelconque de E .

4 Espaces vectoriels de polynômes

4.1 Polynômes à coefficients dans K

On appelle **polynôme** à coefficients dans K toute suite $P = (a_0, a_1, \dots, a_n, \dots)$ d'éléments de K nulle à partir d'un certain rang, i.e

$$\exists p \in \mathbb{N} \text{ tel que } k > p \implies a_k = 0.$$

Pour tout $i \in \mathbb{N}$, a_i est appelé **coefficient de degré i** de P .

On écrit alors le polynôme P sous la forme

$$P = a_0 + a_1X + \cdots + a_pX^p$$

en adoptant la convention $X^0 = 1$.

On appelle polynôme nul et on note 0 le polynôme dont tous les coefficients sont nuls ; si P n'est pas le polynôme nul, alors il peut s'écrire sous la forme

$$P = a_0 + a_1X + \cdots + a_pX^p \text{ avec } a_p \neq 0.$$

On dit alors que le polynôme P est de **degré p** . Par convention, on dit que le polynôme nul est de degré $-\infty$.

On note $K[X]$ l'ensemble des polynômes à coefficients dans K et $K_p[X]$ l'ensemble des polynômes à coefficients dans K de degré $\leq p$.

4.2 Opérations sur $K[X]$

On définit trois opérations sur $K[X]$

* l'addition : si $P = a_0 + a_1X + \cdots + a_pX^p$ et $Q = b_0 + b_1X + \cdots + b_nX^n$ avec par exemple $n \leq p$, on pose

$$P + Q = (a_0 + b_0) + (a_1 + b_1)X + \cdots + (a_n + b_n)X^n + a_{n+1}X^{n+1} + \cdots + a_pX^p$$

autrement dit on additionne les coefficients degré par degré.

* la multiplication : si $P = a_0 + a_1X + \cdots + a_pX^p$ et $Q = b_0 + b_1X + \cdots + b_nX^n$ avec par exemple $n \leq p$, on pose

$PQ = a_0b_0 + (a_0b_1 + a_1b_0)X + \cdots + (a_0b_k + a_1b_{k-1} + \cdots + a_{k-1}b_1 + a_kb_0)X^k + \cdots + a_nb_pX^{n+p}$
c'est-à-dire qu'on développe le produit en appliquant la règle $X^kX^m = X^{k+m}$.

* la multiplication par un scalaire : si $\lambda \in K$ et si $P = a_0 + a_1X + \cdots + a_pX^p$, on pose

$$\lambda P = \lambda a_0 + \lambda a_1X + \cdots + \lambda a_pX^p.$$

L'ensemble $K[X]$ muni de l'addition et de la multiplication par un scalaire est alors un espace vectoriel sur K et pour tout $p \in \mathbb{N}$, $K_p[X]$ est un sous-espace vectoriel de $K[X]$.

Si $p \in \mathbb{N}$, le système $(1, X, X^2, \dots, X^p)$ est libre, en effet, si $\lambda_0, \lambda_1, \dots, \lambda_p$ sont des éléments de K tels que

$$\lambda_0 + \lambda_1X + \cdots + \lambda_pX^p = 0$$

alors le polynôme $P = \lambda_0 + \lambda_1X + \cdots + \lambda_pX^p$ est le polynôme nul, i.e tous ses coefficients sont nuls : $\lambda_0 = \lambda_1 = \cdots = \lambda_p = 0$; d'autre part tout polynôme de degré $\leq p$ est combinaison linéaire des polynômes $1, X, X^2, \dots, X^p$ et ainsi le système $(1, X, X^2, \dots, X^p)$ engendre $K_p[X]$; par conséquent $(1, X, X^2, \dots, X^p)$ est une base de $K_p[X]$, appelée **base canonique** et ainsi $K_p[X]$ est de dimension $p + 1$.

Pour construire une base de $K[X]$, il faut prendre par contre une infinité de polynômes ; en effet, tout polynôme P non nul a un certain degré p donc est combinaison linéaire de $1, X, X^2, \dots, X^p$, mais p peut être aussi grand que l'on veut, par conséquent il faut prendre tous les polynômes X^n , n décrivant \mathbb{N} , pour avoir une base de $K[X]$: on dit que $K[X]$ est un espace vectoriel de dimension **infinie**.

II REDUCTION DES ENDOMORPHISMES

Dans tout le chapitre, E désigne un espace vectoriel de dimension finie $n \geq 1$ sur $K = \mathbb{R}$ ou \mathbb{C} .

1 Sous-espaces propres

1.1 Définition

Soit u un endomorphisme de E ; on dit qu'un scalaire λ est **valeur propre** de u s'il existe un vecteur non nul x de E tel que $u(x) = \lambda x$: un tel vecteur x est appelé **vecteur propre** de u associé à la valeur propre λ .

L'ensemble des valeurs propres d'un endomorphisme u est appelé **spectre** de u et noté $\text{Sp}(u)$.

1.2 Proposition

Soient u un endomorphisme de E et λ un scalaire ; alors λ est valeur propre de u si et seulement si $\ker(u - \lambda Id) \neq \{0\}$.

Preuve :

Par définition, λ est valeur propre de u si et seulement si il existe un vecteur x non nul de E tel que $u(x) = \lambda x$ ce qui s'écrit encore $(u - \lambda Id)(x) = 0$, autrement dit $\ker(u - \lambda Id) \neq \{0\}$.

□

1.3 Définition

Soient u un endomorphisme de E et λ une valeur propre de u . Le sous-espace vectoriel $\ker(u - \lambda Id)$ de E est appelé **sous-espace propre** de u associé à la valeur propre λ ; tous ses éléments, hormis le vecteur nul, sont des vecteurs propres de u associés à la valeur propre λ .

1.4 Proposition

Soient u un endomorphisme de E et $(\lambda_1, \dots, \lambda_p)$ une famille finie de valeurs propres deux à deux distinctes de u . Alors les sous-espaces propres $\ker(u - \lambda_1 Id), \dots, \ker(u - \lambda_p Id)$ sont en somme directe, i.e tout vecteur x de $\ker(u - \lambda_1 Id) + \dots + \ker(u - \lambda_p Id)$ s'écrit de manière unique sous la forme :

$$x = a_1 + a_2 + \dots + a_p$$

où $a_i \in \ker(u - \lambda_i Id)$ pour tout $i \in [1, p]$. En particulier, des vecteurs propres associés à des valeurs propres distinctes sont linéairement indépendants.

Preuve :

Si on a deux écritures $x = a_1 + a_2 + \dots + a_p = b_1 + b_2 + \dots + b_p$ où $a_i, b_i \in \ker(u - \lambda_i Id)$ pour tout $i \in [1, p]$, alors

$$(a_1 - b_1) + (a_2 - b_2) + \dots + (a_p - b_p) = 0$$

avec $a_i - b_i \in \ker(u - \lambda_i Id)$; il suffit donc de montrer que si $x_1 + \dots + x_p = 0$ avec $x_i \in \ker(u - \lambda_i Id)$ pour tout $i \in [1, p]$, alors $x_1 = \dots = x_p = 0$. On va faire un raisonnement par récurrence sur p :

Considérons la proposition (\mathcal{P}_p) suivante:

(\mathcal{P}_p) : si $x_1 + \dots + x_p = 0$ avec $x_i \in \ker(u - \lambda_i Id)$ pour tout $i \in [1, p]$, alors $x_1 = \dots = x_p = 0$.

(\mathcal{P}_1) est trivialement vraie.

(\mathcal{P}_2) : si $x_1 + x_2 = 0$ avec $x_1 \in \ker(u - \lambda_1 Id)$ et $x_2 \in \ker(u - \lambda_2 Id)$, alors $u(x_1 + x_2) = 0$ i.e $u(x_1) + u(x_2) = 0$ ou encore $\lambda_1 x_1 + \lambda_2 x_2 = 0$; or en multipliant l'égalité $x_1 + x_2 = 0$ par λ_2 , on obtient $\lambda_2 x_1 + \lambda_2 x_2 = 0$, d'où par différence $(\lambda_1 - \lambda_2)x_1 = 0$ et ainsi $x_1 = 0$ puisque $\lambda_1 \neq \lambda_2$, puis $x_2 = 0$.

Supposons (\mathcal{P}_{p-1}) vérifiée;

Soient $x_i \in \ker(u - \lambda_i Id)$ pour $i \in [1, p]$ tels que $x_1 + \dots + x_p = 0$. Alors, on a

$$0 = u(x_1 + \dots + x_p) = u(x_1) + \dots + u(x_p) = \lambda_1 x_1 + \dots + \lambda_p x_p$$

D'autre part, on a aussi $0 = \lambda_p(x_1 + \dots + x_p)$ d'où par différence

$$0 = (\lambda_p - \lambda_1)x_1 + \dots + (\lambda_p - \lambda_{p-1})x_{p-1}.$$

Avec l'hypothèse de récurrence (\mathcal{P}_{p-1}) , on en déduit alors

$$(\lambda_p - \lambda_1)x_1 = \dots = (\lambda_p - \lambda_{p-1})x_{p-1} = 0.$$

Or les valeurs propres $\lambda_1, \dots, \lambda_{p-1}, \lambda_p$ sont distinctes deux à deux, donc $x_1 = \dots = x_{p-1} = 0$. On en déduit aussitôt que $x_p = 0$ également et ainsi (\mathcal{P}_p) est vérifiée.

□

1.5 Proposition

Soit u un endomorphisme de E ; alors, pour tout scalaire λ , $\ker(u - \lambda Id)$ est stable par u . En particulier, tout sous-espace propre de u est stable par u .

Preuve:

Soit $x \in \ker(u - \lambda Id)$, alors on a $u(x) = \lambda x$, d'où $u(u(x)) = u(\lambda x) = \lambda u(x)$ et ainsi $u(x) \in \ker(u - \lambda Id)$.

□

2 Diagonalisation

2.1 Proposition

Soient u un endomorphisme de E et λ un scalaire; alors λ est valeur propre de u si et seulement si $\det(u - \lambda Id) = 0$.

Preuve:

D'après 1.2, λ est valeur propre de u si et seulement si $\ker(u - \lambda Id) \neq \{0\}$, autrement dit $u - \lambda Id$ est non injective; en dimension finie, cela équivaut à $u - \lambda Id$ est non bijective, ou encore $\det(u - \lambda Id) = 0$.

□

2.2 Définition et proposition

Soit A une matrice de $\mathcal{M}_n(K)$. On appelle **polynôme caractéristique** de A et on note $P_A(X)$ le déterminant de la matrice à coefficients dans $K[X]$ donnée par $A - XI$; $P_A(X)$ est un polynôme de degré n , plus précisément

$$P_A(X) = \det(A - XI) = (-1)^n X^n + (-1)^{n-1}(\text{tr}(A))X^{n-1} + \dots + \det(A).$$

Preuve:

Notons $a_{i,j}$ le terme général de la matrice A ; alors, la formule du déterminant nous donne

$$\det(A - XI) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{j=1}^n (a_{\sigma(j),j} - \delta_{\sigma(j),j} X).$$

où les $\delta_{i,j}$ sont les symboles de Kronecker : $\delta_{i,j} = 1$ si $i = j$ et $\delta_{i,j} = 0$ si $i \neq j$. Autrement dit, chaque terme du développement de $\det(A - XI)$ est obtenu, au signe près, en faisant le produit de n coefficients de la matrice choisis de façon à prendre un et un seul coefficient par ligne et par colonne. Par conséquent, le terme de plus haut degré en X s'obtient en faisant le produit des n termes diagonaux $(a_{1,1} - X) \cdots (a_{n,n} - X)$ (ce qui correspond à la permutation identique): le terme de plus haut degré de $P_A(X)$ est donc $(-1)^n X^n$. De la même façon, pour obtenir un terme en X^{n-1} , il est nécessaire de prendre $n - 1$ termes de la diagonale; mais alors, pour respecter la règle d'un et d'un seul terme par ligne et par colonne, le dernier facteur est forcément lui aussi sur la diagonale. Donc le terme de degré $n - 1$ de $P_A(X)$ provient du produit $(a_{1,1} - X) \cdots (a_{n,n} - X)$ et de lui seul: il est donc égal à $(-1)^{n-1}(a_{1,1} + \dots + a_{n,n})X^{n-1} = (-1)^{n-1}(\text{tr}(A))X^{n-1}$. Enfin le coefficient constant de $P_A(X)$ est $P_A(0) = \det(A)$.

□

2.3 Théorème et définition

Soit u un endomorphisme de E . Alors le polynôme caractéristique d'une matrice représentative de u dans une base de E est indépendant de la base choisie; on l'appelle le polynôme caractéristique de u et on le note P_u .

Les valeurs propres de u sont exactement les racines sur K de P_u ; l'ordre de multiplicité d'une racine λ de P_u est appelé multiplicité de la valeur propre λ de u et on la note $m(\lambda)$.

Preuve:

Si A et B sont deux matrices représentatives de u , il existe une matrice inversible P telle que $A = P^{-1}BP$; alors $A - XI = P^{-1}BP - XI = P^{-1}(B - XI)P$, d'où $\det(A - XI) = \det(P^{-1}(B - XI)P) = \det(P^{-1}) \det(B - XI) \det(P) = \det(B - XI)$.

On en déduit alors que les valeurs propres de u sont exactement les racines sur K de P_u en utilisant 2.1.

□

2.4 Corollaire

Si E est un espace vectoriel sur \mathbb{C} , tout endomorphisme de E possède au moins une valeur propre.

Preuve :

En effet, le polynôme caractéristique est de degré ≥ 1 et à coefficients dans \mathbb{C} donc possède au moins une racine dans \mathbb{C} .

□

Remarque : le corollaire précédent est en défaut si le corps de base est \mathbb{R} , puisqu'un polynôme à coefficients réels ne possède pas toujours de racines réelles.

2.5 Proposition

Une matrice et sa transposée ont même polynôme caractéristique.

Preuve :

En effet $\det({}^t A - XI) = \det({}^t(A - XI)) = \det(A - XI)$.

□

2.6 Proposition

Soit u un endomorphisme de E et soit F un sous-espace vectoriel de E non réduit à $\{0\}$ et stable par u . Soit u' la restriction de u à F ; $u' : F \rightarrow F$ est alors un endomorphisme de F et le polynôme caractéristique de u' divise le polynôme caractéristique de u .

Preuve :

Choisissons une base $\mathcal{B}' = (e_1, \dots, e_p)$ de F et complétons-la en une base $\mathcal{B} = (e_1, \dots, e_p, \dots, e_n)$ de E ; soient A la matrice de u dans la base \mathcal{B} et A' la matrice de u' dans la base \mathcal{B}' . Comme $u(F) \subset F$, les vecteurs $u(e_1), \dots, u(e_p)$ sont combinaisons linéaires de (e_1, \dots, e_p) seulement, autrement dit

$$A = \begin{pmatrix} A' & B \\ 0 & C \end{pmatrix}$$

où B est une matrice $(p, n-p)$ et C est une matrice carrée d'ordre $n-p$. Donc

$$P_A(X) = \begin{vmatrix} A' - XI_p & B \\ 0 & C - XI_{n-p} \end{vmatrix} = \det(A' - XI_p) \det(C - XI_{n-p}) = P_{A'}(X) \det(C - XI_{n-p}).$$

□

2.7 Proposition

Soit u un endomorphisme de E et λ une valeur propre de u ; alors la multiplicité de λ est supérieure ou égale à la dimension du sous-espace propre de u associé à λ :

$$1 \leq \dim \ker(u - \lambda Id) \leq m(\lambda).$$

En particulier, si λ est racine simple du polynôme caractéristique, alors le sous-espace propre de u associé à λ est de dimension 1.

Preuve:

Le sous-espace propre $\ker(u - \lambda I)$ est non réduit à $\{0\}$ par définition et est stable par u d'après 1.5, donc la proposition 2.6 s'applique à la restriction u' de u à $\ker(u - \lambda I)$ et ainsi $P_{u'}$ divise P_u ; or u' n'est autre que l'homothétie λId sur $\ker(u - \lambda I)$, donc son polynôme caractéristique $P_{u'}(X) = (\lambda - X)^p$ si p désigne la dimension de $\ker(u - \lambda I)$. On en déduit aussitôt que $m(\lambda) \geq p$.

□

2.8 Théorème et définition

Soit u un endomorphisme de E . On dit que u est **diagonalisable** s'il vérifie une des quatre conditions équivalentes suivantes :

- Il existe une base de E dans laquelle la matrice représentative de u est une matrice diagonale ;
- Il existe une base de E formée de vecteurs propres de u ;
- E est la somme directe des sous-espaces propres $\ker(u - \lambda Id)$ quand λ décrit le spectre de u ;
- Le polynôme caractéristique P_u de u est scindé sur le corps K (i.e P_u est produit de polynômes du premier degré à coefficients dans K) et pour toute valeur propre λ de u , la multiplicité $m(\lambda)$ est égale à la dimension du sous-espace propre associé à λ .

Preuve:

a) \implies b) Soit \mathcal{B} une base de E dans laquelle la matrice représentative de u est une matrice diagonale ; alors de façon évidente chaque vecteur de la base \mathcal{B} est un vecteur propre de u .

b) \implies c) Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E formée de vecteurs propres de u , alors chaque vecteur e_i appartient à un sous-espace propre de u ; de plus, d'après 1.4, les sous-espaces propres de u sont en somme directe, on en déduit aussitôt que E est la somme directe des sous-espaces propres $\ker(u - \lambda Id)$ quand λ décrit le spectre de u .

c) \implies d) Si E est la somme directe des sous-espaces propres $\ker(u - \lambda Id)$ quand λ décrit le spectre de u , alors, en prenant une base de chaque sous-espace propre de u et en réunissant ces bases, on obtient une base \mathcal{B} de E formée de vecteurs propres de u . La matrice représentative de u dans cette base est alors diagonale. Notons $\lambda_1, \dots, \lambda_p$ les éléments (distincts deux à deux) du spectre, et pour tout $i \in [1, p]$, $m_i = \dim \ker(u - \lambda_i Id)$; alors on a

$$M_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 I_{m_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{m_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p I_{m_p} \end{pmatrix}$$

Donc $P_u(X) = \det(M_{\mathcal{B}}(u) - XI) = (\lambda_1 - X)^{m_1} \cdots (\lambda_p - X)^{m_p}$; on déduit aussitôt que P_u est scindé et que pour toute valeur propre λ de u , la multiplicité $m(\lambda)$ est égale à la dimension du sous-espace propre associé à λ .

d) \implies a) Si $P_u(X) = (\lambda_1 - X)^{m_1} \cdots (\lambda_p - X)^{m_p}$ avec $m_i = \dim \ker(u - \lambda_i Id)$ pour tout $i \in [1, p]$, alors comme P_u est de degré n , on a $n = m_1 + \cdots + m_p$; notons $F = \bigoplus_{i=1}^p \ker(u - \lambda_i Id)$, on a donc $\dim F = m_1 + \cdots + m_p = n$ et ainsi $F = E$. Prenons alors une base de chaque sous-espace propre $\ker(u - \lambda_i Id)$ et réunissons ces bases : on obtient de cette façon une base \mathcal{B} de E formée de vecteurs propres de u ; alors $M_{\mathcal{B}}(u)$ est évidemment diagonale. □

Exemple

Considérons l'endomorphisme f dont la matrice dans la base canonique de \mathbb{R}^3 est donnée par

$$A = \begin{pmatrix} -3 & 2 & 4 \\ -2 & 1 & 4 \\ -2 & 2 & 3 \end{pmatrix}$$

Le calcul de son polynôme caractéristique donne $P_A(X) = -(X + 1)^2(X - 3)$.

Le sous-espace propre associé à la valeur propre -1 , $\ker(A + I)$ est un plan d'équation $-x + y + 2z = 0$ et de base (u_1, u_2) où $u_1 = (1, 1, 0)$ et $u_2 = (2, 0, 1)$. Donc $\dim \ker(A + I) = 2 = m(-1)$.

D'autre part, la valeur propre 3 étant racine simple de P_A , le sous-espace propre associé $\ker(A - 3I)$ est une droite, de base $u_3 = (1, 1, 1)$; $\mathcal{B} = (u_1, u_2, u_3)$ est alors une base de \mathbb{R}^3 dans laquelle la matrice de f est diagonale :

$$M_{\mathcal{B}}(f) = D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

2.9 Corollaire

Tout endomorphisme u d'un espace E de dimension n qui admet n valeurs propres distinctes deux à deux est diagonalisable.

Cette condition suffisante n'est évidemment nullement nécessaire : l'endomorphisme Id admet une seule valeur propre 1 et est diagonalisable (sa matrice dans n'importe quelle base est diagonale !)

2.10 Définition

On dit qu'une matrice A de $M_n(K)$ est **diagonalisable** si l'endomorphisme de K^n dont A est la matrice dans la base canonique est diagonalisable, ou, ce qui est équivalent, s'il existe une matrice diagonale semblable à A .

3 Polynômes d'endomorphismes

3.1 Définition et proposition

Soit u un endomorphisme de E ; pour tout $n \in \mathbb{N}^*$ on note $u^n = \overbrace{u \circ u \circ \dots \circ u}^{n \text{ fois}}$ et $u^0 = Id$.

On a les propriétés suivantes, pour tous $p, q \in \mathbb{N}$

a) $u^p \circ u^q = u^{p+q} = u^q \circ u^p$;

b) $(u^p)^q = u^{pq} = (u^q)^p$.

On peut alors définir la notion de **polynôme d'endomorphisme** ;

pour tout polynôme $P \in K[X]$ qui s'écrit

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_0$$

on pose

$$P(u) = a_n u^n + a_{n-1} u^{n-1} + \dots + a_0 Id$$

$P(u)$ ainsi défini est un endomorphisme de E . De la même façon, si A est une matrice de $M_n(K)$, on pose

$$P(A) = a_n A^n + a_{n-1} A^{n-1} + \dots + a_0 I.$$

Preuve: Les propriétés a) et b) s'obtiennent facilement par récurrence.

3.2 Proposition

Soient u et v deux endomorphismes de E et P et Q deux polynômes de $K[X]$. Alors on a

a) $(P + Q)(u) = P(u) + Q(u)$ et $(PQ)(u) = P(u) \circ Q(u) = (QP)(u)$;

b) si $u \circ v = v \circ u$, alors $P(u) \circ Q(v) = Q(v) \circ P(u)$;

c) $M_{\mathcal{B}}(P(u)) = P(M_{\mathcal{B}}(u))$.

Preuve: immédiat.

3.3 Proposition

Soit u un endomorphisme de E ; alors on a la suite croissante des noyaux itérés

$$\{0\} \subset \ker u \subset \ker u^2 \subset \dots \subset \ker u^k \subset \ker u^{k+1} \subset \dots \subset E.$$

En particulier, si λ est valeur propre de u , on a

$$\{0\} \subset \ker(u - \lambda Id) \subset \ker(u - \lambda Id)^2 \subset \dots \subset \ker(u - \lambda Id)^k \subset \ker(u - \lambda Id)^{k+1} \subset \dots \subset E.$$

Preuve: soit $k \in \mathbb{N}^*$ et soit $x \in \ker u^k$, alors $u^{k+1}(x) = u(u^k(x)) = u(0) = 0$ et ainsi $x \in \ker u^{k+1}$.

□

3.4 Théorème de Cayley-Hamilton

Soit u un endomorphisme de E ; alors on a $P_u(u) = 0$. De même, si A est une matrice de $\mathcal{M}_n(K)$; alors $P_A(A) = 0$.

Preuve: admis.

3.5 Définition

Soit u un endomorphisme de E dont le polynôme caractéristique P_u est scindé sur K

$$P_u(X) = (-1)^n (X - \lambda_1)^{m_1} \cdots (X - \lambda_p)^{m_p}$$

où $\lambda_1, \dots, \lambda_p$ sont les valeurs propres distinctes de u de multiplicités respectives m_1, \dots, m_p . On appelle **sous-espace spectral (ou caractéristique)** associé à la valeur propre λ_i le sous-espace vectoriel $\ker(u - \lambda_i Id)^{m_i}$ (qui est stable par u).

3.6 Théorème

Soient u un endomorphisme de E , P_1, P_2, \dots, P_p des polynômes de $K[X]$ premiers entre eux deux à deux, et $P = \prod_{i=1}^p P_i$. Alors on a

$$\ker P(u) = \ker P_1(u) \oplus \ker P_2(u) \oplus \cdots \oplus \ker P_p(u).$$

Preuve:

Démontrons $\ker P(u) = \ker P_1(u) \oplus \ker P_2(u) \oplus \cdots \oplus \ker P_p(u)$ par récurrence sur l'entier p :

Si $p = 1$, le théorème est vérifié.

Si $p = 2$, $P = P_1 P_2$ avec P_1 et P_2 premiers entre eux ; alors, d'après le théorème de Bezout il existe deux polynômes U_1 et U_2 tels que

$$1 = U_1 P_1 + U_2 P_2$$

d'où

$$Id = U_1(u) \circ P_1(u) + U_2(u) \circ P_2(u). \quad (*)$$

Montrons que $\ker P_1(u) \cap \ker P_2(u) = \{0\}$:

soit $x \in E$ tel que $P_1(u)(x) = P_2(u)(x) = 0$ alors, comme $x = U_1(u)(P_1(u)(x)) + U_2(u)(P_2(u)(x))$ d'après (*), on a $x = U_1(u)(0) + U_2(u)(0) = 0$.

Montrons maintenant que $\ker P(u) \subset \ker P_1(u) + \ker P_2(u)$:

soit $x \in \ker P(u)$, alors d'après (*), on a

$$x = U_1(u)(P_1(u)(x)) + U_2(u)(P_2(u)(x)).$$

Or $U_1(u)(P_1(u)(x)) \in \ker P_2(u)$, en effet, d'après la proposition 3.2, on a

$$P_2(u)(U_1(u)(P_1(u)(x))) = (P_2(u) \circ U_1(u) \circ P_1(u))(x) = (U_1(u) \circ P_1(u) \circ P_2(u))(x)$$

et

$$U_1(u) \circ P_1(u) \circ P_2(u) = U_1(u) \circ (P_1 P_2)(u) = U_1(u) \circ P(u)$$

donc

$$P_2(u)(U_1(u)(P_1(u)(x))) = U_1(u)(P(u)(x)) = U_1(u)(0) = 0.$$

De la même façon, $U_2(u)(P_2(u)(x)) \in \ker P_1(u)$ et ainsi $\ker P \subset \ker P_1 + \ker P_2$.

Il reste à montrer que $\ker P_1(u) + \ker P_2(u)$ est contenu dans $\ker P(u)$. Soit $x \in \ker P_1(u)$, alors $P(u)(x) = (P_1 P_2)(u)(x) = P_2(u)(P_1(u)(x)) = P_2(u)(0) = 0$ et ainsi $\ker P_1(u)$ est contenu dans $\ker P(u)$.

De même $\ker P_2(u)$ est contenu dans $\ker P(u)$, d'où $\ker P_1(u) + \ker P_2(u) \subset \ker P(u)$.

Ainsi on a prouvé que $\ker P(u) = \ker P_1(u) \oplus \ker P_2(u)$.

Si $p \geq 3$, supposons le théorème démontré à l'ordre $p-1$; comme P_1, P_2, \dots, P_p sont premiers entre eux deux à deux, alors les deux polynômes P_1 et $P_2 \cdots P_p$ sont premiers entre eux, on peut donc leur appliquer le cas précédent

$$\ker P(u) = \ker P_1(u) \oplus \ker(P_2 \cdots P_p)(u)$$

Or par hypothèse de récurrence, on a $\ker(P_2 \cdots P_p)(u) = \ker P_2(u) \oplus \cdots \oplus \ker P_p(u)$, d'où $\ker P(u) = \ker P_1(u) \oplus \ker P_2(u) \oplus \cdots \oplus \ker P_p(u)$.

□

3.7 Théorème

Soit u un endomorphisme de E dont le polynôme caractéristique P_u est scindé sur K

$$P_u(X) = (-1)^n (X - \lambda_1)^{m_1} \cdots (X - \lambda_p)^{m_p}$$

avec $\lambda_1, \dots, \lambda_p$ distinctes deux à deux. Notons $E_i = \ker(u - \lambda_i Id)^{m_i}$ pour $1 \leq i \leq p$ les sous-espaces spectraux de u ; alors

a) on a la somme directe $E = E_1 \oplus E_2 \oplus \cdots \oplus E_p$;

b) pour tout $1 \leq i \leq p$, on a $\dim E_i = m_i$;

c) si u est diagonalisable, pour tout $1 \leq i \leq p$, $E_i = \ker(u - \lambda_i Id)^{m_i} = \ker(u - \lambda_i Id)$.

Preuve:

a) D'après le théorème de Cayley-Hamilton, $P_u(u) = 0$ donc $\ker P_u(u) = E$, de plus $(X - \lambda_1)^{m_1}, \dots, (X - \lambda_p)^{m_p}$ sont premiers entre eux deux à deux, donc $E = E_1 \oplus E_2 \oplus \cdots \oplus E_p$ d'après 3.6.

b) Notons $r_i = \dim E_i$ pour $1 \leq i \leq p$, alors d'après a) on a

$$n = \dim E = r_1 + \dots + r_p$$

D'autre part, on a

$$P_u(X) = (-1)^n (X - \lambda_1)^{m_1} \cdots (X - \lambda_p)^{m_p}$$

donc

$$n = d^{\circ} P = m_1 + m_2 + \cdots + m_p$$

Montrons que pour tout $1 \leq i \leq p$, $r_i \leq m_i$:

D'après 3.3 on a pour tout $1 \leq i \leq p$, la suite croissante

$$\ker(u - \lambda_i Id) \subset \ker(u - \lambda_i Id)^2 \subset \dots \subset \ker(u - \lambda_i Id)^k \subset \dots \ker(u - \lambda_i Id)^{m_i} = E_i$$

On construit alors une base de E_i de la manière suivante : on commence par calculer une base du sous-espace propre $\ker(u - \lambda_i Id)$; si $\ker(u - \lambda_i Id) = E_i$, le calcul est terminé. Si $\ker(u - \lambda_i Id) \neq E_i$ on calcule alors une base de $\ker(u - \lambda_i Id)^2$ en complétant la base de $\ker(u - \lambda_i Id)$ précédemment calculée puis une base de $\ker(u - \lambda_i Id)^3$ et ainsi de suite jusqu'à obtenir une base de E_i ; on voit alors facilement que la matrice de la restriction u_i de u à E_i dans cette base est de la forme

$$T_i = \begin{pmatrix} \lambda_i & * & \cdots & * & * \\ 0 & \lambda_i & * & \cdots & * \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i & * \\ 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}$$

On en déduit que le polynôme caractéristique $P_{u_i}(X)$ de u_i est égal à $(\lambda_i - X)^{r_i}$; or d'après 2.6, $P_{u_i}(X)$ divise $P_u(X)$ donc $r_i \leq m_i$.

Si on avait $r_j < m_j$ pour un entier $1 \leq j \leq p$, alors on aurait :

$$n = r_1 + \dots + r_j + \dots + r_p < m_1 + \dots + m_j + \dots + m_p = n$$

ce qui est absurde, donc pour tout $1 \leq i \leq p$, on a $r_i = m_i$ et ainsi $\dim E_i = m_i$.

c) Si u est diagonalisable, alors d'après 2.8, pour tout $1 \leq i \leq p$, on a

$$\dim \ker(u - \lambda_i Id) = m_i = \dim E_i$$

Or $\ker(u - \lambda_i Id) \subset E_i$, d'où $\ker(u - \lambda_i Id) = E_i$.

□

4 Trigonalisation

4.1 Définition

Soit u un endomorphisme de E ; on dit que u est **trigonalisable** si et seulement si il existe une base de E dans laquelle la matrice représentative de u est triangulaire.

On peut imposer que cette matrice soit triangulaire supérieure ; en effet, si la matrice de u dans une base (e_1, \dots, e_n) de E est triangulaire inférieure, alors la matrice de u dans la base (e_n, \dots, e_1) est triangulaire supérieure.

4.2 Théorème

Soit u un endomorphisme de E ; u est trigonalisable si et seulement si son polynôme caractéristique est scindé sur K .

Preuve :

Supposons u trigonalisable, alors il existe une base \mathcal{B} de E telle que la matrice $M_{\mathcal{B}}(u)$ est triangulaire ; si $a_{i,j}$ est le terme général de $M_{\mathcal{B}}(u)$, on a alors

$$P_u(X) = \det(M_{\mathcal{B}}(u) - XI) = \prod_{i=1}^n (a_{i,i} - X)$$

et $P_u(X)$ est scindé sur K .

Réciproquement, supposons que $P_u(X)$ est scindé sur K et montrons par récurrence sur la dimension n de E que u est trigonalisable.

Si $n = 1$, u est évidemment trigonalisable.

Soit $n \geq 2$ et supposons que tout endomorphisme d'un K -espace vectoriel de dimension $n - 1$ qui possède un polynôme caractéristique scindé sur K , est trigonalisable.

Le polynôme $P_u(X)$ est scindé sur K , donc P_u possède au moins une racine $\lambda \in K$ et ainsi u admet au moins un vecteur propre e_1 associé à λ : $u(e_1) = \lambda e_1$. Complétons le système libre e_1 en une base $\mathcal{B} = (e_1, \dots, e_n)$ de E , alors la matrice de u dans cette base est de la forme

$$A = \begin{pmatrix} \lambda & B \\ 0_{n-1} & A' \end{pmatrix}$$

où B est une matrice ligne et A' une matrice de $M_{n-1}(K)$. Considérons le sous-espace vectoriel F de E engendré par e_2, \dots, e_n et v l'endomorphisme de F dont la matrice dans la base (e_2, \dots, e_n) est donnée par A' . Il est clair que

$$P_u(X) = P_A(X) = (\lambda - X)P_{A'}(X) = (\lambda - X)P_v(X)$$

ainsi $P_v(X)$ est scindé sur K , donc, par hypothèse de récurrence, il existe une base (e'_2, \dots, e'_n) de F dans laquelle la matrice de v est triangulaire supérieure. Alors la matrice de u dans la base (e_1, e'_2, \dots, e'_n) est aussi triangulaire supérieure.

□

4.3 Définition et proposition

On dit qu'une matrice A de $M_n(K)$ est **trigonalisable** si elle vérifie l'une des conditions équivalentes suivantes :

- a) l'endomorphisme de K^n dont A est la matrice dans la base canonique est trigonalisable ;
- b) il existe une matrice triangulaire de $M_n(K)$ semblable à A ;
- c) le polynôme caractéristique de A est scindé sur le corps K .

En particulier, toute matrice à coefficients complexes est trigonalisable.

4.4 Pratique de la trigonalisation

Soit u un endomorphisme de E dont le polynôme caractéristique P_u est scindé sur K

$$P_u(X) = (-1)^n (X - \lambda_1)^{m_1} \cdots (X - \lambda_p)^{m_p}$$

avec $\lambda_1, \dots, \lambda_p$ distinctes deux à deux. alors, d'après 3.7 on a la somme directe

$$E = \ker(u - \lambda_1 Id)^{m_1} \oplus \cdots \oplus \ker(u - \lambda_p Id)^{m_p}$$

On construit une base de chaque sous-espace spectral $\ker(u - \lambda_i Id)^{m_i}$ comme dans la preuve de 3.7 :

on commence par calculer une base du sous-espace propre $\ker(u - \lambda_i Id)$; si $\ker(u - \lambda_i Id) = \ker(u - \lambda_i Id)^{m_i}$, le calcul est terminé. Si $\ker(u - \lambda_i Id) \neq \ker(u - \lambda_i Id)^{m_i}$, comme on a la suite croissante

$$\ker(u - \lambda_i Id) \subset \ker(u - \lambda_i Id)^2 \subset \dots \subset \ker(u - \lambda_i Id)^{m_i}$$

on calcule alors une base de $\ker(u - \lambda_i Id)^2$ en complétant la base de $\ker(u - \lambda_i Id)$ précédemment calculée et ainsi de suite jusqu'à obtenir une base de $\ker(u - \lambda_i Id)^{m_i}$; la matrice de la restriction de u à $\ker(u - \lambda_i Id)^{m_i}$ dans cette base est alors de la forme

$$T_i = \begin{pmatrix} \lambda_i & * & \cdots & * & * \\ 0 & \lambda_i & * & \cdots & * \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i & * \\ 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix} = \lambda_i I_{m_i} + N_i$$

donc est triangulaire supérieure; de plus la matrice N_i est nilpotente d'indice $\leq m_i$, i.e $N_i^{m_i} = 0$. Comme E est somme directe des sous-espaces spectraux, en réunissant toutes ces bases, on obtient une base dans laquelle la matrice de u est formée de blocs diagonaux donnés par les T_i donc est aussi triangulaire supérieure.

Exemple

Considérons l'endomorphisme f dont la matrice dans la base canonique de \mathbb{R}^3 est donnée par

$$A = \begin{pmatrix} -8 & 1 & 5 \\ 2 & -3 & -1 \\ -4 & 1 & 1 \end{pmatrix}$$

Le calcul de son polynôme caractéristique donne $P_A(X) = -(X + 4)^2(X + 2)$.

Le sous-espace propre associé à la valeur propre -4 , $\ker(A + 4I)$ est une droite de base $u_1 = (1, -1, 1)$ donc A n'est pas diagonalisable; on calcule alors une base de $\ker(A + 4I)^2$ en complétant la base u_1 de $\ker(A + 4I)$: $\ker(A + 4I)^2$ est un plan d'équation $-x + y + 2z = 0$, donc, en posant $u_2 = (1, 1, 0)$, (u_1, u_2) est une base de $\ker(A + 4I)^2$.

D'autre part, le sous-espace propre associé à la valeur propre -2 , $\ker(A + 2I)$ est une droite de base $u_3 = (1, 1, 1)$; $\mathcal{B} = (u_1, u_2, u_3)$ est alors une base de \mathbb{R}^3 dans laquelle la matrice de f est triangulaire:

$$M_{\mathcal{B}}(f) = T = \begin{pmatrix} -4 & -3 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

4.5 Corollaire

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$ dont le polynôme caractéristique P_A est scindé sur K

$$P_A(X) = (-1)^n (X - \lambda_1)^{m_1} \cdots (X - \lambda_p)^{m_p}$$

avec $\lambda_1, \dots, \lambda_p$ distinctes deux à deux. Alors il existe une matrice inversible P et une matrice triangulaire supérieure T telles que $A = PTP^{-1}$; de plus T est de la forme

$$T = \text{diag}(T_1, \dots, T_p)$$

où chaque bloc diagonal $T_i = \lambda_i I_{m_i} + N_i$ avec N_i triangulaire supérieure et nilpotente d'indice $\leq m_i$.

Donc T s'écrit $T = D + N$ où $D = \text{diag}(\lambda_1 I_{m_1}, \dots, \lambda_p I_{m_p})$ est diagonale, $N = \text{diag}(N_1, \dots, N_p)$ est triangulaire supérieure et nilpotente et D et N commutent.

De plus, A est diagonalisable si et seulement si $N = 0$.

4.6 Calcul de la puissance k ième d'une matrice

Soit $A \in \mathcal{M}_n(\mathbb{C})$: on veut calculer A^k pour tout entier $k \geq 1$. Deux cas se présentent :

1er cas : A est diagonalisable. Alors on diagonalise A , c'est-à-dire on calcule une matrice P inversible telle que $A = PDP^{-1}$ où D est diagonale ; il est alors facile de montrer par récurrence que $A^k = PD^kP^{-1}$ et que si

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix} \text{ alors } D^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2^k & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1}^k & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n^k \end{pmatrix}$$

d'où le calcul de A^k .

2ème cas : A n'est pas diagonalisable, alors comme le corps de base est \mathbb{C} , A est trigonalisable ; de plus, si on emploie la méthode décrite en 4.4 pour trigonaliser A , on peut construire une matrice P inversible telle que $A = PTP^{-1}$ où T est une matrice formée de blocs diagonaux du type $T_i = \lambda_i I + N_i$ où N_i est nilpotente d'indice $\leq m_i$. Alors $A^k = PT^kP^{-1}$ et T^k est une matrice formée par les blocs diagonaux T_i^k : il ne reste plus qu'à calculer T_i^k . Comme $T_i = \lambda_i I + N_i$ et les deux matrices $\lambda_i I$ et N_i commutent, on peut appliquer la formule du binôme de Newton :

$$T_i^k = \sum_{l=0}^k C_k^l (\lambda_i I)^{k-l} N_i^l$$

or $(\lambda_i I)^{k-l} = \lambda_i^{k-l} I$ et $N_i^l = 0$ si $l \geq m_i$, d'où si $k \geq m_i$

$$T_i^k = \lambda_i^k I + \lambda_i^{k-1} C_k^1 N_i + \cdots + \lambda_i^{k-m_i+1} C_k^{m_i-1} N_i^{m_i-1}.$$

d'où le calcul de A^k .

Exemple

Reprenons l'exemple traité en 4.4

$$A = \begin{pmatrix} -8 & 1 & 5 \\ 2 & -3 & -1 \\ -4 & 1 & 1 \end{pmatrix}$$

$A = PTP^{-1}$ où

$$T = \begin{pmatrix} -4 & -3 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \text{ et } P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & -2 \\ -1 & 1 & 2 \end{pmatrix}$$

alors pour tout $k \geq 1$, $A^k = PT^kP^{-1}$; de plus

$$T = \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix}$$

avec

$$T_1 = \begin{pmatrix} -4 & -3 \\ 0 & -4 \end{pmatrix} \text{ et } T_2 = (-2)$$

donc

$$T^k = \begin{pmatrix} T_1^k & 0 \\ 0 & T_2^k \end{pmatrix}$$

Calculons T_1^k : $T_1 = -4I + N_1$ où N_1 vérifie $N_1^2 = 0$ donc, puisque $-4I$ et N_1 commutent, on a :

$$T_1^k = (-4I + N_1)^k = (-4I)^k + C_k^1(-4I)^{k-1}N_1 + 0$$

i.e

$$T_1^k = \begin{pmatrix} (-4)^k & -3k(-4)^{k-1} \\ 0 & (-4)^k \end{pmatrix}$$

d'où

$$T^k = \begin{pmatrix} (-4)^k & -3k(-4)^{k-1} & 0 \\ 0 & (-4)^k & 0 \\ 0 & 0 & (-2)^k \end{pmatrix}$$

il ne reste plus alors qu'à calculer $A^k = PT^kP^{-1}$:

$$A^k = \begin{pmatrix} (-4)^{k-1}(-3k-6) + (-2)^{k-1} & 2(-4)^{k-1} - (-2)^{k-1} & (-4)^{k-1}(3k+4) + (-2)^k \\ (-4)^{k-1}(3k-2) + (-2)^{k-1} & -2(-4)^{k-1} - (-2)^{k-1} & (-4)^{k-1}(-3k+4) + (-2)^k \\ (-4)^{k-1}(-3k-2) + (-2)^{k-1} & 2(-4)^{k-1} - (-2)^{k-1} & (-4)^{k-1}(3k) + (-2)^k \end{pmatrix}$$

III EXPONENTIELLES DE MATRICES-SYSTÈMES D'ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES

Dans tout le chapitre K désigne le corps \mathbb{R} ou \mathbb{C} .

1 Normes sur K^n et $M_n(K)$

1.1 Définition

Rappelons qu'une **norme** sur un K -espace vectoriel E est une application

$$\begin{aligned} E &\rightarrow \mathbb{R}^+ \\ x &\mapsto \|x\| \end{aligned}$$

vérifiant les propriétés suivantes :

- a) $\forall x \in K^n, \|x\| = 0 \iff x = 0$;
- b) $\forall x \in K^n, \forall \lambda \in K, \|\lambda x\| = |\lambda| \|x\|$;
- c) $\forall x, y \in K^n, \|x + y\| \leq \|x\| + \|y\|$.

1.2 Proposition

- a) Les applications de K^n dans \mathbb{R}^+ définies par :
pour tout $X = (x_1, x_2, \dots, x_n) \in K^n$

$$\begin{aligned} \|X\|_1 &= \sum_{i=1}^n |x_i| \\ \|X\|_\infty &= \max_{1 \leq i \leq n} |x_i| \end{aligned}$$

sont des normes sur K^n et on a les inégalités

$$\|X\|_\infty \leq \|X\|_1 \leq n \|X\|_\infty$$

- b) L'application de $M_n(K)$ dans \mathbb{R}^+ définie par

$$\text{pour toute matrice } A = (a_{ij}) \in M_n(K), \quad \|A\| = n \max_{1 \leq i, j \leq n} |a_{ij}|$$

est une norme sur $M_n(K)$ et on a les propriétés suivantes :

- * $\forall A, B \in M_n(K), \|AB\| \leq \|A\| \cdot \|B\|$
- * $\forall A \in M_n(K), \forall k \in \mathbb{N}^*, \|A^k\| \leq \|A\|^k$.

Dans tout le chapitre, $M_n(K)$ sera muni de cette norme $\|\cdot\|$.

Preuve: sans difficulté.

2 Exponentielles de matrices

2.1 Théorème et définition

Soit $A = (a_{ij}) \in M_n(K)$; pour tout entier N , posons

$$S_N = I + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{N!}A^N$$

et notons s_{Nij} le terme général de la matrice S_N . Alors pour tous $1 \leq i, j \leq n$, la suite $(s_{Nij})_{N \in \mathbb{N}}$ converge vers une limite $s_{ij} \in K$: on note alors e^A la matrice de terme général s_{ij} et on l'appelle **exponentielle** de la matrice A ; e^A est aussi la limite de la suite $(S_N)_{N \in \mathbb{N}}$ dans $M_n(K)$ muni de la norme $\|\cdot\|$ i.e $\lim_{N \rightarrow +\infty} \|S_N - e^A\| = 0$, ce que l'on écrit :

$$e^A = \sum_{k=0}^{+\infty} \frac{1}{k!} A^k.$$

Preuve:

Notons $a_{ij}^{(k)}$ le terme général de la matrice A^k pour tout entier k . Alors, pour tous $1 \leq i, j \leq n$, on a :

$$s_{Nij} = \delta_{ij} + a_{ij} + \cdots + \frac{1}{N!} a_{ij}^{(N)}$$

or, pour tout entier k et pour tous $1 \leq i, j \leq n$, on a d'après 1.2

$$\frac{1}{k!} |a_{ij}^{(k)}| \leq \frac{1}{n} \frac{1}{k!} \|A^k\| \leq \frac{1}{n} \frac{1}{k!} \|A\|^k.$$

Or la série de terme général $\frac{1}{k!} \|A\|^k$ converge dans \mathbb{R} vers $e^{\|A\|}$ donc la série de terme général $\frac{1}{k!} a_{ij}^{(k)}$ est absolument convergente dans K et ainsi la suite $(s_{Nij})_{N \in \mathbb{N}}$ converge vers une limite $s_{ij} \in K$. Notons e^A la matrice de terme général s_{ij} et montrons que e^A est aussi la limite de la suite $(S_N)_{N \in \mathbb{N}}$ dans $M_n(K)$ muni de la norme $\|\cdot\|$:

$$\text{pour tout } N \in \mathbb{N}, \|e^A - S_N\| = n \max_{1 \leq i, j \leq n} |s_{ij} - s_{Nij}|$$

or pour tous $1 \leq i, j \leq n$, on a $|s_{ij} - s_{Nij}| = r_{Nij} \rightarrow 0$ quand $N \rightarrow +\infty$, donc r_N défini par $r_N = \max_{1 \leq i, j \leq n} r_{Nij} \rightarrow 0$ quand $N \rightarrow +\infty$ et ainsi $S_N \rightarrow e^A$ pour la norme $\|\cdot\|$ quand $N \rightarrow \infty$.

□

2.2 Proposition

Soient A et $B \in M_n(K)$, alors on a :

a) si $AB = BA$, alors $e^{A+B} = e^A \cdot e^B = e^B \cdot e^A$;

b) e^A est inversible et $(e^A)^{-1} = e^{-A}$;

c) l'application

$$\begin{aligned} \mathbb{R} &\rightarrow M_n(K) \\ t &\mapsto e^{tA} \end{aligned}$$

est dérivable sur \mathbb{R} et $\frac{d}{dt}(e^{tA}) = Ae^{tA} = e^{tA}A$;

d) si A est diagonale :

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

alors e^A est aussi diagonale et

$$e^A = \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 & 0 \\ 0 & e^{\lambda_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_{n-1}} & 0 \\ 0 & 0 & \cdots & 0 & e^{\lambda_n} \end{pmatrix};$$

e) si A est nilpotente, il existe $p \geq 1$ tel que $A^p = 0$; alors

$$e^A = I + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{(p-1)!}A^{p-1};$$

f) s'il existe une matrice inversible P telle que $A = P^{-1}BP$, alors $e^A = P^{-1}e^BP$.

Preuve :

a) se démontre comme dans le cas des exponentielles de réels, sachant que la condition $AB = BA$ permet de développer

$$(A + B)^k = \sum_{j=0}^k C_k^j A^j B^{k-j}.$$

b) On utilise a) : $e^{A-A} = e^A e^{-A}$ car A commute avec $-A$; or $e^{A-A} = e^0 = I$ d'où le résultat.

c) Pour tout $t \in \mathbb{R}$, on a

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

donc

$$Ae^{tA} = A \sum_{k=0}^{+\infty} \frac{1}{k!} t^k A^k = \sum_{k=0}^{+\infty} \frac{1}{k!} t^k A^{k+1} = \left(\sum_{k=0}^{+\infty} \frac{1}{k!} t^k A^k \right) A = e^{tA}A$$

et ainsi A et e^{tA} commutent ; de plus $e^{tA} = (s_{ij}(t))_{ij}$ où

$$s_{ij}(t) = \sum_{k=0}^{+\infty} a_{ij}^{(k)} \frac{1}{k!} t^k.$$

Chaque fonction $s_{ij}(t)$ est donc la somme d'une série entière de rayon de convergence infini, par conséquent elle est dérivable et

$$\frac{d}{dt}s_{ij}(t) = \sum_{k=1}^{+\infty} a_{ij}^{(k)} \frac{1}{k!} k t^{k-1} = \sum_{k=0}^{+\infty} a_{ij}^{(k+1)} \frac{1}{k!} t^k.$$

Ainsi $\frac{d}{dt}s_{ij}(t)$ n'est autre que le terme général de la matrice $\sum_{k=0}^{+\infty} \frac{1}{k!} t^k A^{k+1} = Ae^{tA}$, d'où

$$\frac{d}{dt}(e^{tA}) = Ae^{tA}.$$

d) Si

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix} \text{ alors } A^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2^k & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1}^k & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n^k \end{pmatrix}$$

Comme $e^A = \sum_{k=0}^{+\infty} \frac{1}{k!} A^k$, on voit que e^A est diagonale et que son terme diagonal d'indice i est égal à $\sum_{k=0}^{+\infty} \frac{1}{k!} \lambda_i^k = e^{\lambda_i}$, d'où

$$e^A = \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 & 0 \\ 0 & e^{\lambda_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_{n-1}} & 0 \\ 0 & 0 & \cdots & 0 & e^{\lambda_n} \end{pmatrix}$$

e) clair puisque $A^k = 0$ pour tout $k \geq p$.

f) Si $A = P^{-1}BP$, alors $A^k = P^{-1}B^kP$ d'où

$$e^A = \sum_{k=0}^{+\infty} \frac{1}{k!} A^k = \sum_{k=0}^{+\infty} \frac{1}{k!} P^{-1}B^kP = P^{-1} \left(\sum_{k=0}^{+\infty} \frac{1}{k!} B^k \right) P = P^{-1}e^B P.$$

□

2.3 Calcul de l'exponentielle d'une matrice

Soit $A \in M_n(K)$ dont le polynôme caractéristique P_A est scindé sur K

$$P_A(X) = (-1)^n (X - \lambda_1)^{m_1} \dots (X - \lambda_p)^{m_p}$$

avec $\lambda_1, \dots, \lambda_p$ distinctes deux à deux.

1er cas : A est diagonalisable, alors il existe une matrice inversible P et une matrice diagonale D telles que $A = PDP^{-1}$ et D est de la forme :

$$D = \text{diag}(\lambda_1 I_{m_1}, \dots, \lambda_p I_{m_p})$$

donc, d'après 2.2, $e^A = Pe^D P^{-1}$ et $e^D = \text{diag}(e^{\lambda_1} I_{m_1}, \dots, e^{\lambda_p} I_{m_p})$.

2ème cas : A n'est pas diagonalisable, alors A est trigonalisable puisque P_A est scindé ; d'après II 4.4, il existe alors une matrice inversible P et une matrice triangulaire supérieure T telles que $A = PTP^{-1}$. De plus T est de la forme

$$T = \text{diag}(T_1, \dots, T_p)$$

où chaque bloc $T_i = \lambda_i I_{m_i} + N_i$ avec N_i triangulaire supérieure et nilpotente d'indice m_i . Donc T s'écrit $T = D + N$ où $D = \text{diag}(\lambda_1 I_{m_1}, \dots, \lambda_p I_{m_p})$ est diagonale et $N = \text{diag}(N_1, \dots, N_p)$ est triangulaire supérieure et nilpotente : il existe $q \geq 1$ tel que $N^q = 0$; de plus D et N commutent, donc d'après 2.2, on a $e^T = e^D e^N$ avec

$$e^D = \text{diag}(e^{\lambda_1} I_{m_1}, \dots, e^{\lambda_p} I_{m_p})$$

$$e^N = I + N + \frac{1}{2!} N^2 + \dots + \frac{1}{(q-1)!} N^{q-1}$$

plus précisément

$$e^N = \text{diag}(e^{N_1}, \dots, e^{N_p})$$

avec

$$e^{N_i} = I + N_i + \frac{1}{2!} N_i^2 + \dots + \frac{1}{(m_i-1)!} N_i^{m_i-1}$$

et $e^A = Pe^T P^{-1}$.

Exemple Reprenons l'exemple traité en II 4.4

$$A = \begin{pmatrix} -8 & 1 & 5 \\ 2 & -3 & -1 \\ -4 & 1 & 1 \end{pmatrix}$$

$A = PTP^{-1}$ où

$$T = \begin{pmatrix} -4 & -3 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \quad \text{et} \quad P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & -2 \\ -1 & 1 & 2 \end{pmatrix}$$

donc $e^A = Pe^T P^{-1}$.

Comme

$$T = \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix}$$

avec

$$T_1 = \begin{pmatrix} -4 & -3 \\ 0 & -4 \end{pmatrix} \text{ et } T_2 = (2)$$

on a

$$e^T = \begin{pmatrix} e^{T_1} & 0 \\ 0 & e^{T_2} \end{pmatrix}$$

d'autre part $T_1 = -4I + N_1$ où N_1 vérifie $N_1^2 = 0$ et $-4I$ et N_1 commutent, donc

$$e^{T_1} = e^{-4I+N_1} = e^{-4I}e^{N_1} = e^{-4}I(I + N_1)$$

d'où

$$e^T = \begin{pmatrix} e^{-4} & -3e^{-4} & 0 \\ 0 & e^{-4} & 0 \\ 0 & 0 & e^{-2} \end{pmatrix}$$

on en déduit

$$e^A = \begin{pmatrix} -\frac{3}{2}e^{-4} - \frac{1}{2}e^{-2} & -\frac{1}{2}e^{-4} + \frac{1}{2}e^{-2} & 2e^{-4} + e^{-2} \\ \frac{7}{2}e^{-4} - \frac{1}{2}e^{-2} & \frac{1}{2}e^{-4} + \frac{1}{2}e^{-2} & -4e^{-4} + e^{-2} \\ -\frac{5}{2}e^{-4} - \frac{1}{2}e^{-2} & -\frac{1}{2}e^{-4} + \frac{1}{2}e^{-2} & 3e^{-4} + e^{-2} \end{pmatrix}$$

3 Systèmes d'équations différentielles linéaires

3.1 Définition

Soit $A \in M_n(K)$ et $B(t)$ un vecteur colonne formé de n fonctions coordonnées $b_i(t)$ définies sur un intervalle J de \mathbb{R} ; on considère n fonctions inconnues dérivables sur J

$$x_i(t) : J \rightarrow K$$

et on note $X(t)$ le vecteur colonne formé des n fonctions coordonnées $x_i(t)$: ainsi $X(t)$ est une fonction inconnue dérivable sur J à valeurs dans K^n :

$$X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} \text{ et } \frac{d}{dt}X(t) = \begin{pmatrix} x'_1(t) \\ x'_2(t) \\ \vdots \\ x'_n(t) \end{pmatrix}$$

On appelle système d'équations différentielles linéaires à coefficients constants toute équation différentielle du type:

$$(E) \quad : \quad \frac{d}{dt}X(t) = AX(t) + B(t).$$

On appelle système d'équations différentielles linéaires homogène associé à (E) le système suivant :

$$(H) \quad : \quad \frac{d}{dt}X(t) = AX(t)$$

3.2 Théorème

Avec les notations de 3.1 :

a) la solution générale du système homogène (H) est donnée par :

$$X(t) = e^{tA}C$$

où C est un vecteur colonne de K^n quelconque ;

b) les solutions de (E) s'obtiennent en ajoutant à la solution générale de (H) une solution particulière de (E) ;

c) pour calculer une solution particulière de (E), on applique la méthode dite de "variation de la constante" qui consiste à chercher une solution particulière sous la forme

$$X(t) = e^{tA}C(t)$$

où $C(t)$ est une fonction inconnue dérivable sur J à déterminer : on obtient alors le système différentiel

$$\frac{d}{dt}C(t) = e^{-tA}B(t)$$

et il ne reste plus qu'à calculer n primitives pour obtenir une solution.

Preuve :

a) pour tout $t \in J$, posons $F(t) = e^{-tA}X(t)$, alors

$$\frac{d}{dt}F(t) = \left(\frac{d}{dt}e^{-tA}\right)X(t) + e^{-tA}\frac{d}{dt}X(t) = e^{-tA}(-A)X(t) + e^{-tA}\frac{d}{dt}X(t) = e^{-tA}\left(-AX(t) + \frac{d}{dt}X(t)\right)$$

d'où

$$\frac{d}{dt}F(t) = e^{-tA} \times 0 = 0$$

donc $F(t)$ est un vecteur colonne constant C pour tout $t \in J$. Réciproquement, si C est un vecteur colonne constant, alors il est clair que $t \mapsto e^{tA}C$ est solution de (H).

b) clair.

c) on cherche une solution de (E) de la forme $X(t) = e^{tA}C(t)$; alors on a

$$\frac{d}{dt}X(t) = Ae^{tA}C(t) + e^{tA}\frac{d}{dt}C(t) = AX(t) + e^{tA}\frac{d}{dt}C(t)$$

mais aussi

$$\frac{d}{dt}X(t) = AX(t) + B(t)$$

d'où

$$e^{tA}\frac{d}{dt}C(t) = B(t)$$

ou encore

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

□

Exemple Considérons le système différentiel

$$(E) \quad \begin{cases} x'(t) &= -8x + y + 5z \\ y'(t) &= 2x - 3y - z \\ z'(t) &= -4x + y + z \end{cases}$$

Ce système s'écrit sous la forme

$$(E) \quad : \quad \frac{d}{dt}X(t) = AX(t)$$

où

$$A = \begin{pmatrix} -8 & 1 & 5 \\ 2 & -3 & -1 \\ -4 & 1 & 1 \end{pmatrix}$$

La solution générale de (E) s'écrit alors $X(t) = e^{tA}C$ où C est un vecteur constant quelconque. Or $A = PTP^{-1}$ où

$$T = \begin{pmatrix} -4 & -3 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \quad \text{et} \quad P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & -2 \\ -1 & 1 & 2 \end{pmatrix}$$

et ainsi $e^{tA} = Pe^{tT}P^{-1}$. Donc $X(t) = Pe^{tT}P^{-1}C = Pe^{tT}\Gamma$ si on pose $\Gamma = P^{-1}C$: $\Gamma = (\gamma_1, \gamma_2, \gamma_3)$ est un vecteur constant quelconque, tout comme C .

Le calcul de e^{tT} donne

$$e^{tT} = \begin{pmatrix} e^{-4t} & -3te^{-4t} & 0 \\ 0 & e^{-4t} & 0 \\ 0 & 0 & e^{-2t} \end{pmatrix}$$

d'où

$$X(t) = P \begin{pmatrix} \gamma_1 e^{-4t} - 3\gamma_2 t e^{-4t} \\ \gamma_2 e^{-4t} \\ \gamma_3 e^{-2t} \end{pmatrix}$$

et finalement

$$\begin{cases} x(t) &= (\gamma_1 + \gamma_2 - 3\gamma_2 t)e^{-4t} + \gamma_3 e^{-2t} \\ y(t) &= (-\gamma_1 + \gamma_2 + 3\gamma_2 t)e^{-4t} + \gamma_3 e^{-2t} \\ z(t) &= (\gamma_1 - 3\gamma_2 t)e^{-4t} + \gamma_3 e^{-2t} \end{cases}$$

4 Equations différentielles linéaires homogènes d'ordre n à coefficients constants

4.1 Définition et théorème

On appelle équation différentielle linéaire homogène d'ordre n à coefficients constants toute équation différentielle du type

$$(E) \quad : \quad y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0.$$

où a_0, a_1, \dots, a_{n-1} sont des nombres réels avec $a_0 \neq 0$ et la fonction inconnue $y(t)$ est supposée définie sur un intervalle J de \mathbb{R} à valeurs dans \mathbb{R} et n fois dérivable sur J .

Considérons le polynôme

$$P(X) = X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0$$

$P(X)$ est scindé sur \mathbb{C} et s'écrit

$$P(X) = (X - \lambda_1)^{m_1} \dots (X - \lambda_p)^{m_p}$$

où $\lambda_1, \dots, \lambda_p$ sont distincts deux à deux. Alors les solutions de E sont de la forme

$$y(t) = e^{\lambda_1 t} P_1(t) + \dots + e^{\lambda_p t} P_p(t)$$

où P_1, \dots, P_p sont des polynômes vérifiant $d^\circ(P_i) \leq m_i - 1$ pour tout $1 \leq i \leq p$.

Si une racine λ_j est complexe non réelle, alors elle s'écrit $\lambda_j = \alpha_j + i\beta_j$ et dans ce cas le polynôme P étant à coefficient réels, $\overline{\lambda_j} = \alpha_j - i\beta_j$ est aussi racine de P et est donc égale à λ_k pour un certain k et on écrit alors

$$e^{\lambda_j t} P_j(t) + e^{\overline{\lambda_j} t} P_k(t) = e^{\alpha_j t} \cos(\beta_j t) Q_j(t) + e^{\alpha_j t} \sin(\beta_j t) R_j(t)$$

où Q_j et R_j sont des polynômes (à coefficients réels) de degré $\leq m_j - 1$.

Preuve :

On peut ramener la résolution de l'équation (E) à la résolution d'un système d'équations différentielles linéaires de la façon suivante :

Posons

$$X(t) = \begin{pmatrix} y(t) \\ y'(t) \\ \vdots \\ y^{(n-1)}(t) \end{pmatrix} \text{ alors } \frac{d}{dt} X(t) = \begin{pmatrix} y'(t) \\ y''(t) \\ \vdots \\ y^{(n)}(t) \end{pmatrix}$$

or y est solution de (E) si et seulement si $y^{(n)} = -a_{n-1}y^{(n-1)} - \dots - a_1y' - a_0y$ donc si et seulement si $X(t)$ vérifie le système différentiel

$$(E') \quad : \quad \frac{d}{dt} X(t) = AX(t)$$

où A désigne la matrice

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 & 0 \\ 0 & 0 & \dots & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

alors les solutions du système (E') sont données par

$$X(t) = e^{tA} C$$

où C est un vecteur colonne constant quelconque. Or, on voit facilement que le polynôme caractéristique de A est égal à

$$P_A(X) = (-1)^n (X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0) = (-1)^n (X - \lambda_1)^{m_1} \dots (X - \lambda_p)^{m_p}$$

de plus, chaque sous espace propre $\ker(A - \lambda_j I)$ est de dimension 1, engendré par le vecteur

$$u_j = \begin{pmatrix} \lambda_j \\ \lambda_j^2 \\ \vdots \\ \lambda_j^{n-1} \end{pmatrix}$$

ce vecteur étant non nul puisque 0 n'est pas racine de P_A . Donc, sauf dans le cas où P_A possède n racines distinctes, A n'est pas diagonalisable: pour calculer e^{tA} on trigonalise alors A . Ainsi $A = PTP^{-1}$ où T est triangulaire supérieure (ou diagonale dans le cas diagonalisable), et alors $e^{tA} = Pe^{tT}P^{-1}$, d'où

$$X(t) = e^{tA}C = Pe^{tT}P^{-1}C = Pe^{tT}\Gamma$$

en posant $P^{-1}C = \Gamma$; or la fonction $y(t)$ est la première coordonnée du vecteur $X(t)$ et d'après le § 2.3 sur le calcul de l'exponentielle d'une matrice, il est facile de voir qu'en effectuant le calcul $Pe^{tT}\Gamma$, on obtient

$$y(t) = e^{\lambda_1 t} P_1(t) + \dots + e^{\lambda_p t} P_p(t)$$

où P_1, \dots, P_p sont des polynômes vérifiant $d^\circ(P_i) \leq m_i - 1$ pour tout $1 \leq i \leq p$. □

Exemples

a) Considérons l'équation différentielle suivante :

$$y^{(6)} - 6y^{(5)} + 12y^{(4)} - 6y^{(3)} - 9y'' + 12y' - 4y = 0$$

Le polynôme caractéristique associé est

$$P(X) = X^6 - 6X^5 + 12X^4 - 6X^3 - 9X^2 + 12X - 4 = (X - 1)^3(X - 2)^2(X + 1)$$

donc la solution générale de l'équation différentielle est donnée par

$$y(t) = (a_2 t^2 + a_1 t + a_0)e^t + (b_1 t + b_0)e^{2t} + c_0 e^{-t}$$

où $a_0, a_1, a_2, b_0, b_1, c_0$ sont des réels quelconques.

b) Considérons l'équation différentielle suivante :

$$y^{(7)} - y^{(6)} - 2y^{(4)} + 2y^{(3)} + y' - y = 0$$

Le polynôme caractéristique associé est

$$P(X) = X^7 - X^6 - 2X^4 + 2X^3 + X - 1 = (X - 1)^3(X^2 + X + 1)^2$$

donc la solution générale de l'équation différentielle est donnée par

$$y(t) = (a_2 t^2 + a_1 t + a_0)e^t + e^{-\frac{t}{2}} \cos\left(\frac{\sqrt{3}}{2}t\right) (b_1 t + b_0) + e^{-\frac{t}{2}} \sin\left(\frac{\sqrt{3}}{2}t\right) (c_1 t + c_0)$$

où $a_0, a_1, a_2, b_0, b_1, c_0, c_1$ sont des réels quelconques.

IV ESPACES EUCLIDIENS

Dans tout le chapitre E désigne un espace vectoriel sur \mathbb{R} .

1. Formes bilinéaires symétriques - Formes quadratiques

1.1 Définition

On appelle **forme bilinéaire symétrique** sur E une application φ de $E \times E$ dans \mathbb{R} vérifiant les propriétés suivantes :

a) φ est linéaire par rapport à chaque variable, i.e :

$$\forall x_1, x_2, y \in E, \forall \alpha, \beta \in \mathbb{R}, \varphi(\alpha x_1 + \beta x_2, y) = \alpha \varphi(x_1, y) + \beta \varphi(x_2, y)$$

$$\forall x, y_1, y_2 \in E, \forall \alpha, \beta \in \mathbb{R}, \varphi(x, \alpha y_1 + \beta y_2) = \alpha \varphi(x, y_1) + \beta \varphi(x, y_2).$$

b) φ est symétrique : $\forall x, y \in E \varphi(x, y) = \varphi(y, x)$.

1.2 Exemples

a) La forme bilinéaire canonique dans \mathbb{R}^n : $\forall X = (x_1, \dots, x_n), Y = (y_1, \dots, y_n) \in \mathbb{R}^n$

$$\varphi(X, Y) = \sum_{i=1}^n x_i y_i.$$

b) $\varphi(P, Q) = \int_a^b P(t)Q(t)dt$ définit une forme bilinéaire symétrique sur $\mathbb{R}[X]$.

1.3 Définition

Soit φ une forme bilinéaire symétrique sur E , on appelle **forme quadratique** associée à φ l'application q de E dans \mathbb{R} définie par :

$$\forall x \in E, q(x) = \varphi(x, x).$$

1.4 Identités de polarisation

$$\forall x, y \in E, \varphi(x, y) = \frac{1}{2}[q(x+y) - q(x) - q(y)] = \frac{1}{4}[q(x+y) - q(x-y)].$$

1.5 Expression dans une base

Soit E un espace de dimension finie n et soit φ une forme bilinéaire symétrique sur E . On appelle matrice de φ dans la base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E et on note $M_{\mathcal{B}}(\varphi)$ la matrice dont le terme général $a_{i,j} = \varphi(e_i, e_j)$. La matrice $M_{\mathcal{B}}(\varphi)$ est symétrique et on obtient φ en fonction de $M_{\mathcal{B}}(\varphi)$ de la manière suivante :

Pour tous x et y dans E , si X et Y désignent les vecteurs colonnes des coordonnées de x et y dans la base \mathcal{B} , alors :

$$\varphi(x, y) = {}^t X M_{\mathcal{B}}(\varphi) Y.$$

Réciproquement, si A est une matrice symétrique, en posant :

$$\varphi(x, y) = {}^t X A Y$$

on définit une forme bilinéaire symétrique φ dont la matrice est A dans la base \mathcal{B} . Si (x_1, \dots, x_n) (resp. (y_1, \dots, y_n)) désignent les coordonnées de x (resp. y) dans la base \mathcal{B} , on a ainsi :

$$\varphi(x, y) = \sum_{i=1}^n a_{ii} x_i y_i + \sum_{1 \leq i < j \leq n} a_{ij} (x_i y_j + x_j y_i)$$

$$q(x) = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j.$$

1.6 Changement de base

Soit E un espace de dimension finie n muni de deux bases $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ et $\mathcal{B}' = (e'_i)_{1 \leq i \leq n}$ et soit P la matrice de passage de \mathcal{B} à \mathcal{B}' . Alors pour toute forme bilinéaire symétrique φ on a :

$$M_{\mathcal{B}'}(\varphi) = {}^t P M_{\mathcal{B}}(\varphi) P.$$

Preuve :

Soient x et y dans E et soient X (resp. X') le vecteur colonne de x dans \mathcal{B} (resp. \mathcal{B}'), Y (resp. Y') le vecteur colonne de y dans \mathcal{B} (resp. \mathcal{B}') : on a $X = PX'$ et $Y = PY'$. Notons $A = M_{\mathcal{B}}(\varphi)$ et $A' = M_{\mathcal{B}'}(\varphi)$. Alors $\varphi(x, y) = {}^t X A Y = {}^t X' A' Y'$. On en déduit ${}^t (PX') A P Y' = {}^t X' A' Y'$, d'où ${}^t X {}^t P A P Y' = {}^t X' A' Y'$. Donc ${}^t P A P = A'$. □

1.7 Définition

On dit qu'une forme bilinéaire symétrique φ sur E (ou la forme quadratique q associée) est **positive** (resp. **définie positive**) si et seulement si :

$$\forall x \in E, q(x) \geq 0 \text{ (resp. } \forall x \in E - \{0\}, q(x) > 0).$$

1.8 Exemples :

a) La forme bilinéaire canonique dans \mathbb{R}^n est définie positive ; en effet $q(x) = \sum_{i=1}^n x_i^2$, donc $q(x) \geq 0$, et si $q(x) = 0$, alors $\sum_{i=1}^n x_i^2 = 0$ or $\forall i \in [1, n]$ on a $x_i^2 \geq 0$ donc la somme des x_i^2 ne peut être nulle que si tous les x_i^2 sont nuls donc $x_i = 0$ pour tout $i \in [1, n]$ et ainsi $x = 0$.

b) Si $a < b$, la forme bilinéaire $\varphi(P, Q) = \int_a^b P(t)Q(t)dt$ dans $\mathbb{R}[X]$ est définie positive. (cf. exercice)

2. Espaces euclidiens

2.1 Définition

On appelle **produit scalaire** sur E une forme bilinéaire symétrique φ définie positive : on le note souvent $\varphi(x, y) = \langle x | y \rangle$.

On appelle **espace euclidien** tout espace vectoriel E muni d'un produit scalaire $\langle | \rangle$, et pour tout $x \in E$ on note $\|x\| = \sqrt{\langle x | x \rangle}$: **norme euclidienne** de x .

On dit qu'un vecteur x d'un espace euclidien E est **unitaire** si $\|x\| = 1$.

2.2 Proposition

Soit E un espace euclidien. Les identités de polarisation s'expriment sous la forme :

$$\forall x, y \in E, \langle x|y \rangle = \frac{1}{2}[\|x + y\|^2 - \|x\|^2 - \|y\|^2] = \frac{1}{4}[\|x + y\|^2 - \|x - y\|^2].$$

2.3 Inégalité de Cauchy-Schwarz

Soit E un espace euclidien. On a :

$$\forall x, y \in E, |\langle x|y \rangle| \leq \|x\| \|y\|.$$

Et on a l'égalité si et seulement si x et y sont colinéaires.

Preuve: Soient x et y dans E . Considérons $P(\lambda) = \|\lambda x + y\|^2$ pour $\lambda \in \mathbb{R}$. On a :

$$P(\lambda) = \lambda^2 \|x\|^2 + 2\lambda \langle x|y \rangle + \|y\|^2.$$

Si $x \neq 0$, P est un polynôme en λ de degré 2 et on a : $\forall \lambda \in \mathbb{R}, P(\lambda) \geq 0$ donc le discriminant Δ de P est négatif ou nul : $|\langle x|y \rangle|^2 - \|x\|^2 \|y\|^2 \leq 0$ d'où l'inégalité cherchée. Si $\Delta = 0$, P possède une racine réelle λ_0 , donc $\|\lambda_0 x + y\|^2 = 0$ d'où $\lambda_0 x + y = 0$ et x et y sont colinéaires et réciproquement.

Si $x = 0$, pour tout $y \in E$, l'inégalité se réduit à l'égalité $0 = 0$ et 0 et y sont trivialement colinéaires.

□

2.4 Théorème

Soit E un espace euclidien : alors la norme euclidienne $\|\cdot\|$ est bien une norme, i.e elle vérifie les propriétés suivantes :

$$\forall x, y \in E, \forall \lambda \in \mathbb{R}$$

- a) $\|x\| \geq 0$ et $\|x\| = 0$ si et seulement si $x = 0$;
- b) $\|\lambda x\| = |\lambda| \cdot \|x\|$;
- c) $\|x + y\| \leq \|x\| + \|y\|$: inégalité de Minkowski.

De plus, on a :

- d) $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$: identité de la médiane ;
- e) $\|x\|^2 - \|y\|^2 = \langle x + y|x - y \rangle$.

Preuve:

- a) Le produit scalaire est une forme bilinéaire définie positive ;
- b) $\|\lambda x\| = \sqrt{\langle \lambda x|\lambda x \rangle} = \sqrt{\lambda^2 \langle x|x \rangle} = |\lambda| \cdot \|x\|$;
- c) Proviennent de l'inégalité de Cauchy-Schwarz :

$$\|x + y\|^2 = \langle x + y|x + y \rangle = \langle x|x \rangle + \langle y|y \rangle + 2\langle x|y \rangle$$

$$\text{Donc : } \|x + y\|^2 \leq \langle x|x \rangle + \langle y|y \rangle + 2\|x\| \|y\| = (\|x\| + \|y\|)^2.$$

- d) et e) proviennent de la bilinéarité du produit scalaire.

□

2.5 Proposition

Soient $A \in \mathcal{M}_{n,p}(\mathbb{R})$, $X \in \mathbb{R}^p$ et $Y \in \mathbb{R}^n$; alors on a

$$\langle AX|Y \rangle_{\mathbb{R}^n} = \langle X|^t AY \rangle_{\mathbb{R}^p}.$$

Preuve: En effet $\langle AX|Y \rangle_{\mathbb{R}^n} = {}^t (AX)Y = {}^t X^t AY = \langle X|^t AY \rangle_{\mathbb{R}^p}$.

□

3. Orthogonalité

3.1 Définitions

Soit E un espace euclidien.

a) Soient $x, y \in E$; on dit que x et y sont **orthogonaux** si et seulement si $\langle x|y \rangle = 0$ ou, ce qui est équivalent, si et seulement si $\langle y|x \rangle = 0$.

b) Soient E_1, E_2 des sous-espaces vectoriels de E : on dit que E_1 et E_2 sont **orthogonaux** si et seulement si:

$$\forall x_1 \in E_1, \forall x_2 \in E_2, \langle x_1|x_2 \rangle = 0.$$

c) Soit E_1 un sous-espace vectoriel de E : on appelle **orthogonal** de E_1 et on note E_1^\perp le sous-ensemble:

$$E_1^\perp = \{x \in E / \forall x_1 \in E_1, \langle x_1|x \rangle = 0\}.$$

3.2 Proposition

Soit E un espace euclidien. Alors

a) $E^\perp = \{0\}$;

b) soit E_1 un sous-espace vectoriel de E , alors E_1^\perp est un sous-espace vectoriel de E ;

c) soient E_1, E_2 des sous-espaces vectoriels de E , alors on a:

$$(E_1 + E_2)^\perp = E_1^\perp \cap E_2^\perp$$

$$E_1 \subset E_2 \implies E_1^\perp \supset E_2^\perp$$

$$(E_1 \cap E_2)^\perp \supset E_1^\perp + E_2^\perp$$

$$(E_1^\perp)^\perp \supset E_1.$$

Preuve:

a) Soit $x \in E^\perp$: $\forall y \in E, \langle x|y \rangle = 0$. En particulier $\langle x|x \rangle = 0$ i.e $\|x\| = 0$ d'où $x = 0$. Réciproquement il est clair que $0 \in E^\perp$.

b) $\forall x, y \in E_1^\perp, \forall \lambda, \mu \in \mathbb{R}, \forall z \in E_1, \langle \lambda x + \mu y|z \rangle = \lambda \langle x|z \rangle + \mu \langle y|z \rangle = \lambda 0 + \mu 0 = 0$ donc $\lambda x + \mu y \in E_1^\perp$.

c) cf. exercice.

□

3.3 Proposition

Soit E un espace euclidien.

a) Soient $x, y \in E$, alors :

x et y sont orthogonaux si et seulement si $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ (Pythagore)

b) soient $x_1, x_2, \dots, x_p \in E$ deux à deux orthogonaux, alors :

$$\|x_1 + x_2 + \dots + x_p\|^2 = \|x_1\|^2 + \|x_2\|^2 + \dots + \|x_p\|^2$$

(la réciproque est fautive si $p \geq 3$)

c) soient $x_1, x_2, \dots, x_p \in E$ non nuls deux à deux orthogonaux, alors (x_1, x_2, \dots, x_p) est un système libre ;

d) soient E_1, E_2, \dots, E_p des sous-espaces vectoriels de E deux à deux orthogonaux, alors ils sont en somme directe dans E : on parle de **somme directe orthogonale**. En particulier E_1 et E_1^\perp sont en somme directe orthogonale.

Preuve :

a) $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x|y \rangle$ d'où l'égalité si et seulement si $\langle x|y \rangle = 0$;

b) par récurrence sur p en utilisant a) ;

c) si $x_1, x_2, \dots, x_p \in E$ sont deux à deux orthogonaux :

Soient $\lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R}$ tels que $\lambda_1 x_1 + \dots + \lambda_p x_p = 0$. Alors pour tout $i \in [1, p]$ on a :

$\langle x_i | \lambda_1 x_1 + \dots + \lambda_p x_p \rangle = 0 \implies \lambda_1 \langle x_i | x_1 \rangle + \dots + \lambda_p \langle x_i | x_p \rangle = 0 \implies \lambda_i \|x_i\|^2 = 0$ car x_i est orthogonal à $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_p$. Donc $\lambda_i = 0$ puisque $\|x_i\| \neq 0$;

d) si E_1, E_2, \dots, E_p sont deux à deux orthogonaux, considérons $x_i \in E_i$ pour tout $i \in [1, p]$; si $x_1 + \dots + x_p = 0$, alors $\forall i \in [1, p]$, $\langle x_i | x_1 + \dots + x_p \rangle = 0 \implies \|x_i\|^2 = 0$ car x_i est orthogonal à $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_p$, d'où $x_i = 0$: la somme est directe. □

3.4 Procédé d'orthogonalisation de Gram-Schmidt

Soit E un espace euclidien et soit (x_1, x_2, \dots, x_p) un système libre de vecteurs de E . Alors il existe une famille orthogonale (y_1, y_2, \dots, y_p) de vecteurs de E telle que les sous-espaces vectoriels engendrés $Vect(x_1, x_2, \dots, x_p)$ et $Vect(y_1, y_2, \dots, y_p)$ sont égaux. On l'obtient par récurrence de la manière suivante :

- $x_1 = y_1$
- $x_2 = y_2 + \lambda_{2,1} y_1$: on calcule $\lambda_{2,1}$ de sorte que $\langle y_1 | y_2 \rangle = 0$: $\langle y_1 | x_2 \rangle = \langle y_1 | y_2 \rangle + \lambda_{2,1} \langle y_1 | y_1 \rangle$ donc $\lambda_{2,1} = \langle y_1 | x_2 \rangle / \|y_1\|^2$ car $y_1 = x_1$ est non nul. Alors $y_2 \neq 0$ sinon x_2 est colinéaire à x_1 , et on a $Vect(x_1, x_2) = Vect(y_1, y_2)$.
- $x_i = y_i + \lambda_{i,i-1} y_{i-1} + \dots + \lambda_{i,1} y_1$ en supposant construits y_1, \dots, y_{i-1} non nuls orthogonaux deux à deux tels que $Vect(x_1, \dots, x_{i-1}) = Vect(y_1, \dots, y_{i-1})$, et en imposant que y_i soit orthogonal à y_1, \dots, y_{i-1} . Alors pour tout $j \in [1, i-1]$, $\langle y_j | x_i \rangle = 0 + \dots + \lambda_{i,j} \|y_j\|^2 + \dots + 0$ d'où $\lambda_{i,j} = \langle y_j | x_i \rangle / \|y_j\|^2$ et $y_i \neq 0$ sinon x_i est combinaison linéaire de y_1, y_2, \dots, y_{i-1} i.e de x_1, x_2, \dots, x_{i-1} . On a ainsi une famille orthogonale (y_1, y_2, \dots, y_i) de vecteurs non nuls donc libre. Donc $Vect(y_1, y_2, \dots, y_i)$ est de dimension i tout comme $Vect(x_1, x_2, \dots, x_i)$. Or $x_i \in Vect(y_1, y_2, \dots, y_i)$ donc $Vect(x_1, x_2, \dots, x_i) = Vect(y_1, y_2, \dots, y_i)$.

Il suffit alors de poursuivre le processus jusqu'au rang p . On remarque que le procédé de Schmidt laisse les vecteurs x_1, \dots, x_p invariants si x_1, \dots, x_p sont déjà orthogonaux deux à deux.

□

3.5 Déterminant de Gram

Soit E un espace euclidien et soient $(x_1, x_2, \dots, x_p) \in E$.

On appelle **matrice de Gram** associée au système (x_1, x_2, \dots, x_p) et on note $\text{Gram}(x_1, x_2, \dots, x_p)$ la matrice de $M_p(\mathbb{R})$ dont le terme général est donné par $\langle x_i | x_j \rangle$. On appelle **déterminant de Gram** et on note $\text{gram}(x_1, x_2, \dots, x_p)$ le déterminant de la matrice de Gram.

3.6 Proposition

Un déterminant de Gram est invariant lorsqu'on ajoute à un vecteur une combinaison linéaire des autres.

Preuve :

Soient $(x_1, x_2, \dots, x_p) \in E$: il suffit de démontrer que $\text{gram}(x_1, x_2, \dots, x_p)$ reste inchangé lorsqu'on remplace x_i par $x_i + \lambda x_j$: par itération on aura la proposition.

$$\text{gram}(x_1, \dots, x_{i-1}, x_i + \lambda x_j, x_{i+1}, \dots, x_p) =$$

$$\begin{vmatrix} \langle x_1 | x_1 \rangle & \cdots & \langle x_1 | x_i + \lambda x_j \rangle & \cdots & \langle x_1 | x_p \rangle \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \langle x_i + \lambda x_j | x_1 \rangle & \cdots & \langle x_i + \lambda x_j | x_i + \lambda x_j \rangle & \cdots & \langle x_i + \lambda x_j | x_p \rangle \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \langle x_p | x_1 \rangle & \cdots & \langle x_p | x_i + \lambda x_j \rangle & \cdots & \langle x_p | x_p \rangle \end{vmatrix}$$

en développant le produit scalaire dans la ième colonne, on obtient alors :

$$\text{gram}(x_1, \dots, x_{i-1}, x_i + \lambda x_j, x_{i+1}, \dots, x_p) =$$

$$\begin{vmatrix} \langle x_1 | x_1 \rangle & \cdots & \langle x_1 | x_i \rangle + \lambda \langle x_1 | x_j \rangle & \cdots & \langle x_1 | x_p \rangle \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \langle x_i + \lambda x_j | x_1 \rangle & \cdots & \langle x_i + \lambda x_j | x_i \rangle + \lambda \langle x_i + \lambda x_j | x_j \rangle & \cdots & \langle x_i + \lambda x_j | x_p \rangle \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \langle x_p | x_1 \rangle & \cdots & \langle x_p | x_i \rangle + \lambda \langle x_p | x_j \rangle & \cdots & \langle x_p | x_p \rangle \end{vmatrix}$$

On trouve alors dans la i-ème colonne, la j-ème colonne multipliée par λ : on peut donc supprimer ce terme, d'où :

$$\text{gram}(x_1, \dots, x_{i-1}, x_i + \lambda x_j, x_{i+1}, \dots, x_p) =$$

$$\begin{vmatrix} \langle x_1|x_1 \rangle & \cdots & \langle x_1|x_i \rangle & \cdots & \langle x_1|x_p \rangle \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \langle x_i + \lambda x_j|x_1 \rangle & \cdots & \langle x_i + \lambda x_j|x_i \rangle & \cdots & \langle x_i + \lambda x_j|x_p \rangle \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \langle x_p|x_1 \rangle & \cdots & \langle x_p|x_i \rangle & \cdots & \langle x_p|x_p \rangle \end{vmatrix}$$

On fait la même opération sur la i ème ligne et on obtient le résultat voulu :

$$\text{gram}(x_1, \dots, x_{i-1}, x_i + \lambda x_j, x_{i+1}, \dots, x_p) = \text{gram}(x_1, \dots, x_p).$$

□

3.7 Théorème

Soit E un espace euclidien et soient $x_1, x_2, \dots, x_p \in E$. Alors

(x_1, \dots, x_p) est un système libre si et seulement si $\text{gram}(x_1, \dots, x_p) \neq 0$.

et dans ce cas $\text{gram}(x_1, \dots, x_p) > 0$.

Preuve :

Si (x_1, x_2, \dots, x_p) est lié, l'un des vecteurs est combinaison linéaire des autres donc :

$$\text{gram}(x_1, \dots, x_p) = \text{gram}(x_1, \dots, 0, \dots, x_p) = 0$$

d'après 3.6, d'où par contraposée, $\text{gram}(x_1, \dots, x_p) \neq 0 \implies (x_1, \dots, x_p)$ libre.

Réciproquement, si (x_1, \dots, x_p) est libre : construisons par le procédé de Schmidt la famille orthogonale (y_1, \dots, y_p) associée : comme $x_i = y_i + \lambda_{i,i-1}y_{i-1} + \dots + \lambda_{i,1}y_1$, il est clair d'après la proposition 3.6, que $\text{gram}(x_1, \dots, x_p) = \text{gram}(y_1, \dots, y_p)$. Or la famille (y_1, \dots, y_p) étant orthogonale, on a :

$$\text{gram}(y_1, \dots, y_p) = \begin{vmatrix} \|y_1\|^2 & 0 & \cdots & 0 \\ 0 & \|y_2\|^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \|y_p\|^2 \end{vmatrix} = \|y_1\|^2 \cdots \|y_p\|^2 > 0$$

car $y_i \neq 0, \forall i \in [1, p]$.

□

4. Bases orthogonales et orthonormées

4.1 Définitions

Soit E un espace euclidien. On dit qu'une base de E est **orthogonale** si ses vecteurs sont orthogonaux deux à deux. On dit qu'une base de E est **orthonormée** si elle est orthogonale et si tous ses vecteurs sont unitaires.

4.2 Théorème

Tout espace euclidien de dimension finie possède une base orthonormée.

Preuve :

Soit (e_1, \dots, e_n) une base de E : il suffit d'appliquer le procédé de Schmidt au système libre (e_1, \dots, e_n) pour obtenir une base orthogonale de E puis de diviser chaque vecteur obtenu par sa norme.

□

4.3 Expression du produit scalaire dans une base orthonormée

Soit (e_1, \dots, e_n) une base orthonormée de l'espace euclidien E et soient $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{i=1}^n y_i e_i$. Alors on a :

$$\langle x|y \rangle = \sum_{i=1}^n x_i y_i \text{ et } \|x\| = \sqrt{\sum_{i=1}^n x_i^2}.$$

De plus

$$x = \sum_{i=1}^n \langle x|e_i \rangle e_i.$$

4.4 Proposition

Soit E un espace euclidien de dimension finie n et soit E_1, E_2 des sous-espaces vectoriels de E . Alors :

- $E = E_1 \oplus E_1^\perp$: E_1 et son orthogonal sont supplémentaires ;
- $(E_1^\perp)^\perp = E_1$;
- $(E_1 \cap E_2)^\perp = E_1^\perp + E_2^\perp$.

Preuve :

- On sait déjà que E_1 et E_1^\perp sont en somme directe : il ne reste donc qu'à montrer $\dim(E_1^\perp) = n - \dim E_1$.

Soit (e_1, \dots, e_p) une base orthogonale de E_1 : on la complète en une base (e_1, \dots, e_n) de E et on applique le procédé de Schmidt à cette base. Les p premiers vecteurs sont inchangés puisqu'ils étaient déjà orthogonaux deux à deux, d'où une nouvelle base $(e_1, \dots, e_p, e'_{p+1}, \dots, e'_n)$ orthogonale pour E . Ainsi l'espace $E' = \text{Vect}(e'_{p+1}, \dots, e'_n)$ est contenu dans E_1^\perp donc $\dim E_1^\perp \geq \dim E' = n - p$. Or $E_1 \oplus E_1^\perp \subset E$ donc $\dim E_1^\perp \leq n - p$, d'où $\dim E_1^\perp = n - p$.

- On sait déjà que $E_1 \subset (E_1^\perp)^\perp$. Or $\dim E_1^\perp = n-p$ donc $\dim (E_1^\perp)^\perp = n-(n-p) = \dim E_1$, d'où $(E_1^\perp)^\perp = E_1$.
- On a $(E_1^\perp + E_2^\perp)^\perp = (E_1^\perp)^\perp \cap (E_2^\perp)^\perp = E_1 \cap E_2$ donc $(E_1^\perp + E_2^\perp)^{\perp\perp} = (E_1 \cap E_2)^\perp$ i.e $E_1^\perp + E_2^\perp = (E_1 \cap E_2)^\perp$.

□

5. Projection orthogonale - Symétrie orthogonale

5.1 Théorème et Définition

Soit E un espace euclidien de dimension quelconque et soit E' un sous-espace vectoriel de E de dimension finie p : considérons une base orthonormée (e_1, \dots, e_p) de E' et pour tout

$x \in E$, posons $p(x) = \sum_{i=1}^p \langle x|e_i \rangle e_i$. Alors on a :

- $p(x) \in E'$.
- $p(x) - x \in E'^\perp$.
- $\|p(x) - x\| = \min\{\|y - x\| / y \in E'\}$.
- $p(x)$ ne dépend pas de la base orthonormée (e_1, \dots, e_p) de E' choisie.

On dit que $p(x)$ est la **projection orthogonale** de x sur E' . On définit ainsi une application $p : E \rightarrow E$ appelée projection orthogonale sur E' qui est linéaire et qui vérifie :

$$\ker p = E'^\perp, \quad \text{Im } p = E' \quad \text{et} \quad p \circ p = p$$

Dans le cas où E est de dimension finie, si on décompose $x = x' + x''$ dans la somme directe orthogonale $E = E' \oplus E'^\perp$, alors $p(x) = x'$.

Preuve :

- Il est clair que $p(x) \in E'$. Pour tout $i \in [1, p]$, on a :

$$\langle e_i | p(x) - x \rangle = \langle e_i | \sum_{k=1}^p \langle x | e_k \rangle e_k \rangle - \langle e_i | x \rangle = \langle e_i | x \rangle - \langle e_i | x \rangle = 0 \text{ car } (e_1, \dots, e_p) \text{ est orthonormée.}$$

Donc $p(x) - x \in E'^\perp$.

- pour tout $y \in E'$, $\|y - x\|^2 = \|(y - p(x)) + (p(x) - x)\|^2 = \|y - p(x)\|^2 + \|p(x) - x\|^2$ par Pythagore puisque $p(x) - x \in E'^\perp$ et $y - p(x) \in E'$. On en déduit : $\|y - x\|^2 \geq \|p(x) - x\|^2$ d'où $\|p(x) - x\| = \min\{\|y - x\| / y \in E'\}$.
- $p(x)$ ne dépend pas de la base orthonormée (e_1, \dots, e_p) de E' choisie : prenons une autre

base orthonormée (e'_1, \dots, e'_p) de E' et posons $x' = \sum_{i=1}^p \langle x | e'_i \rangle e'_i$. Alors il est clair que $x' \in E'$,

d'où $p(x) - x' \in E'$, et que $x' - x \in E'^\perp$. D'autre part, $p(x) - x' = (p(x) - x) + (x - x')$ d'où $p(x) - x' \in E'^\perp$. Comme $E' \cap E'^\perp = \{0\}$, on en déduit aussitôt que $p(x) = x'$.

Montrons que $\ker p = E'^\perp$: si $x \in E'^\perp$, alors, pour tout $i \in [1, p]$, on a $\langle x | e_i \rangle = 0$,

d'où $p(x) = \sum_{i=1}^p \langle x | e_i \rangle e_i = 0$. Réciproquement, si $p(x) = \sum_{i=1}^p \langle x | e_i \rangle e_i = 0$, alors, pour tout

$i \in [1, p]$, on a $\langle x | e_i \rangle = 0$ et ainsi $x \in E'^\perp$.

Montrons que $\text{Im } p = E'$: on a vu que $\text{Im } p \subset E'$; réciproquement soit $y \in E'$, alors, (e_1, \dots, e_p) étant une base orthonormée de E' , on a, d'après 4.3, $y = \sum_{i=1}^p \langle y | e_i \rangle e_i$, d'où $y = p(y) \in \text{Im } p$ et $E' \subset \text{Im } p$. On en déduit également que $p \circ p = p$, puisque, pour tout $x \in E$, $p(x) \in E'$ d'où $p(x) = p(p(x))$.

Enfin si E est de dimension finie, on a $E = E' \oplus E'^{\perp}$ et tout $x \in E$ se décompose de manière unique sous la forme $x = x' + x''$ où $x' \in E'$ et $x'' \in E'^{\perp}$. Complétons la base orthonormée (e_1, \dots, e_p) de E' en une base orthonormée (e_1, \dots, e_n) de E par le procédé de Gram-Schmidt ; alors, par construction, (e_{p+1}, \dots, e_n) est une base orthonormée de E'^{\perp} et x s'écrit dans la base orthonormée (e_1, \dots, e_n) sous la forme

$$x = \sum_{i=1}^n \langle x | e_i \rangle e_i = \sum_{i=1}^p \langle x | e_i \rangle e_i + \sum_{i=p+1}^n \langle x | e_i \rangle e_i = p(x) + y$$

où $y = \sum_{i=p+1}^n \langle x | e_i \rangle e_i \in E'^{\perp}$. On en déduit que $p(x) = x'$ et $y = x''$ par unicité d'écriture. □

5.2 Définition

Soit E un espace euclidien de dimension finie n et soit E' un sous-espace vectoriel de E . Pour tout $x \in E$, décomposons $x = x' + x''$ dans la somme directe orthogonale $E = E' \oplus E'^{\perp}$ et posons $s(x) = x' - x''$. L'application $s : E \rightarrow E$ ainsi définie est appelée **symétrie orthogonale** par rapport à E' .

5.3 Théorème

Soit E un espace euclidien de dimension finie n et soit E' un sous-espace vectoriel de E . On considère s la symétrie orthogonale par rapport à E' et p la projection orthogonale sur E' . Alors on a :

- $s = 2p - Id$ donc s est linéaire ;
- $s^2 = Id$ donc s est un automorphisme et $s^{-1} = s$;
- $\ker(s - Id) = E'$ et $\ker(s + Id) = E'^{\perp}$.

Réciproquement, si s est un endomorphisme de E vérifiant les deux conditions suivantes :

- $s^2 = Id$;
- $\ker(s - Id) = \ker(s + Id)^{\perp}$.

alors s est la symétrie orthogonale par rapport à $\ker(s - Id)$.

Preuve : La première partie du théorème est immédiate à vérifier.

Réciproquement, soit s vérifiant $s^2 = Id$ et $\ker(s + Id) = \ker(s - Id)^{\perp}$.

On a $(s - Id)(s + Id) = 0$, d'où la somme directe $E = \ker(s - Id) \oplus \ker(s + Id)$ d'après le théorème de décomposition des noyaux II3.6, donc, comme $\ker(s + Id) = \ker(s - Id)^{\perp}$, en décomposant $x = x' + x''$ dans la somme directe orthogonale $E = \ker(s - Id) \oplus \ker(s + Id)$, on a bien $s(x) = x' - x''$, d'où le résultat. □

6 Problèmes de moindres carrés

Dans ce paragraphe on munit \mathbb{R}^p du produit scalaire canonique et de la norme euclidienne $\|\cdot\|$ associée.

6.1 Approximation de fonctions

Soit $y(t)$ une fonction qui dépend linéairement de n paramètres x_j :

$$y(t) = \sum_{j=1}^n x_j f_j(t)$$

où les fonctions f_1, f_2, \dots, f_n sont choisies linéairement indépendantes.

On connaît par des mesures la valeur de $y(t)$ en p points t_1, \dots, t_p : $y(t_i) = y_i$ pour $1 \leq i \leq p$ avec $p > n$ et on cherche à calculer les paramètres x_j pour connaître $y(t)$; comme on a plus d'équations que d'inconnues, il n'y a pas toujours de solution. On cherche alors à calculer x_1, \dots, x_n pour que la quantité

$$E(x_1, \dots, x_n) = \sum_{i=1}^p \left(\sum_{j=1}^n x_j f_j(t_i) - y_i \right)^2$$

soit minimale, à défaut d'être nulle.

Si on note A la matrice de $\mathcal{M}_{p,n}(\mathbb{R})$ de terme général $f_j(t_i)$ et

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ et } y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}$$

alors

$$E(x_1, \dots, x_n) = \|Ax - y\|^2.$$

Un vecteur x qui réalise le minimum de E est appelé une **solution du système au sens des moindres carrés**.

6.2 Régression linéaire

On a mesuré sur p individus n variables Y, X_1, \dots, X_{n-1} : on note \underline{x}_{ij} la mesure de la variable X_j (resp. \underline{y}_i la mesure de la variable Y) sur l'individu i pour $1 \leq j \leq n-1$ et $1 \leq i \leq p$.

Considérons \underline{y} le vecteur de \mathbb{R}^p défini par $\underline{y} = \begin{pmatrix} \underline{y}_1 \\ \vdots \\ \underline{y}_p \end{pmatrix}$. On cherche à calculer Y à l'aide des prédicteurs X_j , supposés linéairement indépendants, sous la forme

$$Y = u_0 + u_1 X_1 + \dots + u_{n-1} X_{n-1}$$

où u_0, \dots, u_{n-1} sont des réels à déterminer.

Considérons y le vecteur de \mathbb{R}^p défini par $y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$ où

$$y_i = u_0 + u_1 \underline{x}_{i1} + \cdots + u_{n-1} \underline{x}_{in-1}$$

On veut que y soit le plus proche de \underline{y} en ce sens que $\|y - \underline{y}\|^2$ soit minimal ; en notant

$$A = \begin{pmatrix} 1 & \underline{x}_{11} & \cdots & \underline{x}_{1n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \underline{x}_{p1} & \cdots & \underline{x}_{pn-1} \end{pmatrix} \text{ et } u = \begin{pmatrix} u_0 \\ \vdots \\ u_{n-1} \end{pmatrix}$$

cela revient à chercher u tel que $\|Au - \underline{y}\|^2$ soit minimal : on est ramené à un problème de moindres carrés comme en 6.1.

6.3 Théorème

Soit $A \in \mathcal{M}_{p,n}(\mathbb{R})$ avec $p > n$ et soit $y \in \mathbb{R}^p$; on note q la projection orthogonale de \mathbb{R}^p sur $\text{Im}A$.

Alors la fonction $E(x) = \|Ax - y\|^2$ atteint son minimum en tout vecteur $\lambda \in \mathbb{R}^n$ vérifiant

$$A\lambda = q(y)$$

De plus λ vérifie

$${}^tAA\lambda = {}^tAy.$$

Si la matrice tAA est inversible, i.e si $rg(A) = n$, on a alors une unique solution donnée par $\lambda = ({}^tAA)^{-1}{}^tAy$.

Preuve :

On a $\min_{x \in \mathbb{R}^n} E(x) = \min_{x \in \mathbb{R}^n} \|Ax - y\|^2 = \min_{v \in \text{Im}A} \|v - y\|^2$, donc d'après 5.1, ce minimum est atteint pour $v = q(y)$ la projection orthogonale de y sur $\text{Im}A$; ainsi il existe $\lambda \in \mathbb{R}^n$ tel que $v = A\lambda$ et λ réalise le minimum de E .

De plus, si (e_1, \dots, e_n) désigne la base canonique de \mathbb{R}^n et si $\lambda = \sum_{i=1}^n \lambda_i e_i$, $v = A\lambda$ s'écrit

$v = \sum_{i=1}^n \lambda_i Ae_i$ et ainsi y s'écrit dans la somme directe orthogonale $\mathbb{R}^p = \text{Im}A \oplus (\text{Im}A)^\perp$ sous la forme

$$y = v + z = \sum_{i=1}^n \lambda_i Ae_i + z$$

où $z \in (\text{Im}A)^\perp$.

Or d'après 2.5 on a pour tout $1 \leq j \leq n$

$$\langle {}^tAy | e_j \rangle = \langle y | Ae_j \rangle = \sum_{i=1}^n \lambda_i \langle Ae_i | Ae_j \rangle + \langle z | Ae_j \rangle = \sum_{i=1}^n \lambda_i \langle e_i | {}^tAAe_j \rangle + 0 = \langle \lambda | {}^tAAe_j \rangle$$

$$\text{donc } \langle {}^tAy | e_j \rangle = \langle {}^t({}^tAA)\lambda | e_j \rangle = \langle ({}^tAA)\lambda | e_j \rangle$$

d'où ${}^tAy = {}^tAA\lambda$.

Il reste à montrer que la matrice carrée d'ordre n tAA est inversible si et seulement si $rg(A) = n$; or $rg(A) = n$ si et seulement si les vecteurs Ae_1, \dots, Ae_n sont linéairement indépendants, ou encore d'après 3.7, si et seulement si la matrice de Gram associée au système (Ae_1, \dots, Ae_n) est inversible; or il est facile de voir que

$$\text{Gram}(Ae_1, \dots, Ae_n) = {}^tAA$$

puisque la base canonique (e_1, \dots, e_n) est orthonormée, d'où le résultat. □

Exemple

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$; on cherche à approximer f par un polynôme $P(t)$ de degré $\leq n - 1$ connaissant les valeurs de f en p points distincts t_1, t_2, \dots, t_p avec $p > n$. Posons

$$P(t) = a_{n-1}t^{n-1} + \dots + a_1t + a_0$$

et considérons

$$E(a_{n-1}, \dots, a_1, a_0) = \sum_{i=1}^p [P(t_i) - f(t_i)]^2 = \|Ax - y\|^2$$

où

$$y = \begin{pmatrix} f(t_1) \\ f(t_2) \\ \vdots \\ f(t_p) \end{pmatrix}, \quad x = \begin{pmatrix} a_{n-1} \\ a_{n-2} \\ \vdots \\ a_0 \end{pmatrix} \quad \text{et} \quad Ax = \begin{pmatrix} P(t_1) \\ P(t_2) \\ \vdots \\ P(t_p) \end{pmatrix}$$

ainsi

$$A = \begin{pmatrix} t_1^{n-1} & t_1^{n-2} & \dots & t_1 & 1 \\ t_2^{n-1} & t_2^{n-2} & \dots & t_2 & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ t_p^{n-1} & t_p^{n-2} & \dots & t_p & 1 \end{pmatrix}$$

La matrice A est de rang n car t_1, t_2, \dots, t_p sont distincts deux à deux, donc le problème de moindres carrés considéré possède une unique solution $\lambda \in \mathbb{R}^n$ vérifiant $\lambda = ({}^tAA)^{-1}{}^tAy$.

V ESPACES HERMITIENS

On va exposer dans ce chapitre une théorie analogue à celle des espaces euclidiens mais cette fois pour des espaces vectoriels complexes. Dans tout le chapitre, E désigne un espace vectoriel sur \mathbb{C} .

1. Formes sesquilineaires hermitiennes - Formes quadratiques

1.1 Définition

On appelle **forme sesquilineaire hermitienne** sur E une application φ de $E \times E$ dans \mathbb{C} vérifiant les propriétés suivantes :

a) φ est sesquilineaire, i.e :

$$\begin{aligned} \forall x_1, x_2, y \in E, \forall \alpha \in \mathbb{C}, \varphi(x_1 + x_2, y) &= \varphi(x_1, y) + \varphi(x_2, y) \text{ et } \varphi(\alpha x_1, y) = \alpha \varphi(x_1, y) \\ \forall x, y_1, y_2 \in E, \forall \alpha \in \mathbb{C}, \varphi(x, y_1 + y_2) &= \varphi(x, y_1) + \varphi(x, y_2) \text{ et } \varphi(x, \alpha y_1) = \bar{\alpha} \varphi(x, y_1). \end{aligned}$$

b) φ est hermitienne : $\forall x, y \in E \varphi(x, y) = \overline{\varphi(y, x)}$.

1.2 Exemples

a) La forme hermitienne canonique dans \mathbb{C}^n : $\forall X = (x_1, \dots, x_n), Y = (y_1, \dots, y_n) \in \mathbb{C}^n$

$$\varphi(X, Y) = \sum_{i=1}^n x_i \bar{y}_i.$$

b) $\varphi(P, Q) = \int_a^b P(t) \overline{Q(t)} dt$ définit une forme sesquilineaire hermitienne sur $\mathbb{C}[X]$.

1.3 Définition

Soit φ une forme sesquilineaire hermitienne sur E : on appelle **forme quadratique hermitienne** associée à φ l'application q de E dans \mathbb{R} définie par :

$$\forall x \in E, q(x) = \varphi(x, x).$$

Remarque : la forme quadratique hermitienne q est bien à valeurs réelles puisque :

$$\forall x \in E, \varphi(x, x) = \overline{\varphi(x, x)}.$$

1.4 Identités de polarisation

Elles sont beaucoup moins "sympathiques" que dans le cas réel : $\forall x, y \in E$

$$\begin{aligned} \varphi(x, y) &= \frac{1}{2} [q(x + y) + iq(x + iy) - (1 + i)(q(x) + q(y))] \\ &= \frac{1}{4} [q(x + y) - q(x - y) + i(q(x + iy) - q(x - iy))] \end{aligned}$$

1.5 Définitions

Pour toute matrice $A = (a_{i,j})_{1 \leq i,j \leq n}$ à coefficients complexes, on appelle **matrice conjuguée** de A et on note \bar{A} la matrice dont le terme général est donné par $\overline{a_{i,j}}$.

On dit que A est **hermitienne** si et seulement si $A = {}^t \bar{A}$.

1.6 Expression dans une base

Soit E un espace de dimension finie n et soit φ une forme sesquilinéaire hermitienne sur E . On appelle matrice de φ dans la base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E et on note $M_{\mathcal{B}}(\varphi)$ la matrice dont le terme général $a_{i,j} = \varphi(e_i, e_j)$. La matrice $M_{\mathcal{B}}(\varphi)$ est hermitienne et on obtient φ en fonction de $M_{\mathcal{B}}(\varphi)$ de la manière suivante :

Pour tous x et y dans E , si X et Y désignent les vecteurs colonnes des coordonnées de x et y dans la base \mathcal{B} , alors :

$$\varphi(x,y) = {}^t X M_{\mathcal{B}}(\varphi) \bar{Y}.$$

Réciproquement, si A est une matrice hermitienne, en posant :

$$\varphi(x,y) = {}^t X A \bar{Y}$$

on définit une forme sesquilinéaire hermitienne φ dont la matrice est A dans la base \mathcal{B} .

On a ainsi, si $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{i=1}^n y_i e_i$

$$\varphi(x,y) = \sum_{i=1}^n a_{ii} x_i \bar{y}_i + \sum_{1 \leq i < j \leq n} (a_{ij} x_i \bar{y}_j + \overline{a_{ij}} x_j \bar{y}_i)$$

$$q(x) = \sum_{i=1}^n a_{ii} |x_i|^2 + 2 \sum_{1 \leq i < j \leq n} \operatorname{Re}(a_{ij} x_i \bar{x}_j)$$

1.7 Changement de base

Soit E un espace de dimension finie n muni de deux bases $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ et $\mathcal{B}' = (e'_i)_{1 \leq i \leq n}$ et soit P la matrice de passage de \mathcal{B} à \mathcal{B}' . Alors, pour toute forme sesquilinéaire hermitienne φ sur E , on a :

$$M_{\mathcal{B}'}(\varphi) = {}^t P M_{\mathcal{B}}(\varphi) \bar{P}.$$

Preuve : analogue à celle faite dans le cas réel.

1.8 Définition

On dit qu'une forme sesquilinéaire hermitienne φ sur E (ou la forme quadratique q associée) est **positive** (resp. **définie positive**) si et seulement si :

$$\forall x \in E, q(x) \geq 0 \text{ (resp. } \forall x \in E - \{0\}, q(x) > 0 \text{)}.$$

1.9 Exemples : cf. Exemples 1.2.

2. Espaces hermitiens

2.1 Définition

On appelle **produit scalaire hermitien** sur E une forme sesquilinéaire hermitienne φ définie positive : on la note comme dans le cas réel $\varphi(x,y) = \langle x|y \rangle$.

On appelle **espace hermitien** tout espace vectoriel E muni d'un produit scalaire hermitien $\langle | \rangle$, et pour tout $x \in E$ on note $\|x\| = \sqrt{\langle x|x \rangle}$: **norme hermitienne** de x .

On dit qu'un vecteur x d'un espace hermitien E est unitaire si $\|x\| = 1$.

2.2 Inégalité de Cauchy-Schwarz

Soit E un espace hermitien. On a :

$$\forall x,y \in E, |\langle x|y \rangle| \leq \|x\| \|y\|.$$

Et on a l'égalité si et seulement si x et y sont colinéaires.

Preuve: Remarquons d'abord qu'ici $| \cdot |$ désigne le module d'un complexe. Soient x et y dans E et notons $\langle x|y \rangle = \rho e^{i\theta}$ avec $\rho = |\langle x|y \rangle|$. Considérons $P(\lambda) = \|\lambda e^{-i\theta} x + y\|^2$ pour $\lambda \in \mathbb{R}$. On a :

$$P(\lambda) = \lambda^2 \|x\|^2 + 2\lambda \operatorname{Re}(e^{-i\theta} \langle x|y \rangle) + \|y\|^2 = \lambda^2 \|x\|^2 + 2\lambda |\langle x|y \rangle| + \|y\|^2.$$

Si $x \neq 0$, P est un polynôme en λ à coefficients réels de degré 2 et on a $\forall \lambda \in \mathbb{R}, P(\lambda) \geq 0$ donc le discriminant Δ de P est négatif ou nul : $|\langle x|y \rangle|^2 - \|x\|^2 \|y\|^2 \leq 0$ d'où l'inégalité cherchée.

Si $\Delta = 0$, P possède une racine réelle λ_0 , donc $\|\lambda_0 e^{-i\theta} x + y\|^2 = 0$ d'où $\lambda_0 e^{-i\theta} x + y = 0$ et x et y sont colinéaires et réciproquement.

Si $x = 0$, l'inégalité se réduit à l'égalité $0 = 0$ et 0 et y sont trivialement colinéaires.

□

2.3 Théorème

Soit E un espace hermitien ; alors la norme hermitienne $\|\cdot\|$ est bien une norme, i.e elle vérifie les propriétés suivantes :

$$\forall x, y \in E, \forall \lambda \in \mathbb{C}$$

a) $\|x\| \geq 0$ et $\|x\| = 0$ si et seulement si $x = 0$;

b) $\|\lambda x\| = |\lambda| \cdot \|x\|$;

c) $\|x + y\| \leq \|x\| + \|y\|$: inégalité de Minkowski.

De plus, on a :

d) $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$: identité de la médiane.

Preuve: quasiment identique à celle faite dans le cas euclidien.

3. Orthogonalité

3.1 Définitions

Soit E un espace hermitien.

• Soient $x, y \in E$; on dit que x est **orthogonal** à y si et seulement si $\langle x|y \rangle = 0$ ou, ce qui est équivalent, si et seulement si $\langle y|x \rangle = 0$: la condition est donc symétrique en x et y .

• Soient E_1, E_2 des sous-espaces vectoriels de E ; on dit que E_1 et E_2 sont **orthogonaux** si et seulement si :

$$\forall x_1 \in E_1, \forall x_2 \in E_2, \langle x_1|x_2 \rangle = 0.$$

• Soit E_1 un sous-espace vectoriel de E ; on appelle **orthogonal** de E_1 et on note E_1^\perp le sous-ensemble :

$$E_1^\perp = \{x \in E / \forall x_1 \in E_1, \langle x_1|x \rangle = 0\}.$$

3.2 Proposition

Soit E un espace hermitien.

a) $E^\perp = \{0\}$;

b) soit E_1 un sous-espace vectoriel de E , alors E_1^\perp est un sous-espace vectoriel de E ;

c) soient E_1, E_2 des sous-espaces vectoriels de E , alors on a :

$$(E_1 + E_2)^\perp = E_1^\perp \cap E_2^\perp$$

$$E_1 \subset E_2 \implies E_1^\perp \supset E_2^\perp$$

$$(E_1 \cap E_2)^\perp \supset E_1^\perp + E_2^\perp$$

$$(E_1^\perp)^\perp \supset E_1.$$

Preuve: quasiment identique à celle faite dans le cas euclidien.

3.3 Proposition

Soit E un espace hermitien.

a) Soient $x, y \in E$, alors, si x et y sont orthogonaux, on a $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ (Pythagore).

La réciproque est fautive: si $\|x + y\|^2 = \|x\|^2 + \|y\|^2$, on a seulement $\operatorname{Re}(\langle x|y \rangle) = 0$.

b) soient $x_1, x_2, \dots, x_p \in E$ deux à deux orthogonaux, alors :

$$\|x_1 + x_2 + \dots + x_p\|^2 = \|x_1\|^2 + \|x_2\|^2 + \dots + \|x_p\|^2$$

c) soient $x_1, x_2, \dots, x_p \in E$ non nuls deux à deux orthogonaux, alors (x_1, x_2, \dots, x_p) est un système libre;

d) soient E_1, E_2, \dots, E_p des sous-espaces vectoriels de E deux à deux orthogonaux, alors ils sont en somme directe dans E : on parle de somme directe orthogonale. En particulier E_1 et E_1^\perp sont en somme directe orthogonale.

Preuve : quasiment identique à celle faite dans le cas euclidien.

3.4 Procédé d'orthogonalisation de Gram-Schmidt

Tout ce qui a été dit dans le cas euclidien est encore valable dans le cas hermitien.

4. Bases orthogonales et orthonormées

4.1 Définitions

Soit E un espace hermitien. On dit qu'une base de E est **orthogonale** si ses vecteurs sont orthogonaux deux à deux. On dit qu'une base de E est **orthonormée** si elle est orthogonale et si tous ses vecteurs sont unitaires.

4.2 Théorème

Tout espace hermitien de dimension finie possède une base orthonormée.

Preuve : quasiment identique à celle faite dans le cas euclidien.

4.3 Expression du produit scalaire dans une base orthonormée

Soit (e_1, \dots, e_n) une base orthonormée de l'espace hermitien E et soient $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{i=1}^n y_i e_i$. Alors, on a :

$$\langle x|y \rangle = \sum_{i=1}^n x_i \bar{y}_i \text{ et } \|x\| = \sqrt{\sum_{i=1}^n |x_i|^2}.$$

4.4 Proposition

Soit E un espace hermitien de dimension finie n et soient E_1, E_2 des sous-espaces vectoriels de E . Alors, on a :

- $E = E_1 \oplus E_1^\perp$: E_1 et son orthogonal sont supplémentaires ;
- $(E_1^\perp)^\perp = E_1$;
- $(E_1 \cap E_2)^\perp = E_1^\perp + E_2^\perp$.

Preuve : quasiment identique à celle faite dans le cas euclidien.

VI RÉDUCTION DES MATRICES NORMALES

Dans tout le chapitre, on considère le produit scalaire hermitien canonique $\langle \cdot | \cdot \rangle$ sur \mathbb{C}^n .

1. Matrice adjointe

1.1 Définition

Soit A une matrice de $\mathcal{M}_n(\mathbb{C})$; on appelle matrice **adjointe** de A et on note A^* la matrice de $\mathcal{M}_n(\mathbb{C})$ définie par

$$A^* = {}^t \bar{A}.$$

1.2 Propriétés

a) Soit $A \in \mathcal{M}_n(\mathbb{C})$; pour tous X et $Y \in \mathbb{C}^n$, on a

$$\langle AX | Y \rangle = \langle X | A^* Y \rangle \text{ et } \langle A^* X | Y \rangle = \langle X | AY \rangle ;$$

b) soient A et $B \in \mathcal{M}_n(\mathbb{C})$ et $\lambda \in \mathbb{C}$, alors on a

$$* (A + B)^* = A^* + B^*$$

$$* (AB)^* = B^* A^*$$

$$* (A^*)^* = A$$

$$* (\lambda A)^* = \bar{\lambda} A^* ;$$

c) les valeurs propres de A^* sont les conjuguées des valeurs propres de A .

Preuve :

a) On a $\langle AX | Y \rangle = {}^t (AX) \bar{Y} = {}^t X {}^t A \bar{Y}$ et $\langle X | A^* Y \rangle = {}^t X \overline{A^* Y} = {}^t X \overline{{}^t A Y} = {}^t X {}^t A \bar{Y}$ d'où $\langle AX | Y \rangle = \langle X | A^* Y \rangle$. De même $\langle A^* X | Y \rangle = \langle X | AY \rangle$.

b) On a $(AB)^* = {}^t (\overline{AB}) = {}^t (\bar{A} \bar{B}) = {}^t (\bar{B}) {}^t (\bar{A}) = B^* A^*$; les autres égalités s'obtiennent de manière analogue.

c) Soit $\lambda \in \mathbb{C}$; on a $\det(A^* - \bar{\lambda}I) = \det({}^t \bar{A} - \bar{\lambda}I) = \overline{\det({}^t A - \lambda I)} = \overline{\det(A - \lambda I)}$, d'où λ valeur propre de $A \iff \det(A - \lambda I) = 0 \iff \det(A^* - \bar{\lambda}I) = 0 \iff \bar{\lambda}$ valeur propre de A^* .

□

1.3 Définitions et proposition

Soit $A \in \mathcal{M}_n(\mathbb{C})$. On dit que

* A est **hermitienne** si $A^* = A$;

* A est **unitaire** si $A^* A = I$: alors A est inversible et $A^{-1} = A^*$, d'où $AA^* = I$ également; de plus les vecteurs colonnes (resp. lignes) de A forment une base orthonormée de \mathbb{C}^n et pour tout $X \in \mathbb{C}^n$, on a $\|AX\| = \|X\|$.

* La matrice de passage entre deux bases orthonormées d'un espace hermitien de dimension finie est une matrice unitaire.

Preuve :

Si u_1, u_2, \dots, u_n sont les vecteurs colonnes d'une matrice unitaire A , alors $AA^* = I$ signifie que pour tous $i, j \in [1, n]$, $\langle u_i | u_j \rangle = \delta_{ij}$ le symbole de Kronecker, autrement dit (u_1, u_2, \dots, u_n) est un système orthonormé de n vecteurs, donc une base orthonormée de \mathbb{C}^n . De même pour les vecteurs lignes.

D'autre part, soit $X \in \mathbb{C}^n$, alors $\|AX\|^2 = \langle AX | AX \rangle = \langle X | A^*AX \rangle = \langle X | X \rangle = \|X\|^2$.

Soit P la matrice de passage entre deux bases orthonormées \mathcal{B} et \mathcal{B}' d'un espace hermitien E de dimension finie n ; si on note φ le produit scalaire hermitien sur E , on a d'après V1.7,

$$M_{\mathcal{B}'}(\varphi) = {}^t P M_{\mathcal{B}}(\varphi) \overline{P}$$

or \mathcal{B} et \mathcal{B}' sont orthonormées donc $M_{\mathcal{B}}(\varphi) = M_{\mathcal{B}'}(\varphi) = I$, on en déduit alors que ${}^t P \overline{P} = I$ d'où, par passage à la transposée, ${}^t \overline{P} P = I$ i.e P est unitaire. □

1.4 Proposition

Soit f un endomorphisme de \mathbb{C}^n dont la matrice A dans une base orthonormée \mathcal{B} de \mathbb{C}^n est unitaire; alors si F est un sous-espace vectoriel de \mathbb{C}^n stable par f , F^\perp est également stable par f :

$$f(F) \subset F \implies f(F^\perp) \subset F^\perp$$

et la matrice de f dans toute base orthonormée de \mathbb{C}^n est unitaire.

Preuve :

La matrice A étant unitaire est inversible, donc f est bijectif et par conséquent $\dim f(F) = \dim F$; or $f(F) \subset F$ par hypothèse, d'où $f(F) = F$ puis $f^{-1}(f(F)) = f^{-1}(F)$, autrement dit $F = f^{-1}(F)$.

Considérons maintenant $x \in F^\perp$ et montrons que $f(x) \in F^\perp$;

soit $y \in F$, alors $\langle f(x) | y \rangle = \langle AX | Y \rangle$ si X et Y sont les vecteurs colonnes coordonnées respectivement de x et y dans la base orthonormée considérée. Ainsi

$$\langle f(x) | y \rangle = \langle AX | Y \rangle = \langle X | A^*Y \rangle = \langle X | A^{-1}Y \rangle = \langle x | f^{-1}(y) \rangle.$$

Or $y \in F$, donc, d'après ce qui précède, $f^{-1}(y) \in F$ d'où $\langle x | f^{-1}(y) \rangle = 0$ d'où $\langle f(x) | y \rangle = 0$: $f(x) \in F^\perp$.

Soit \mathcal{B}' une base orthonormée de \mathbb{C}^n et soit B la matrice de f dans cette base \mathcal{B}' ; si on note P la matrice de passage de la base \mathcal{B} à la base \mathcal{B}' , alors on a $B = P^{-1}AP$, or d'après 1.3, P est une matrice unitaire donc $B = P^*AP$ d'où $B^* = P^*A^*P$ et ainsi $BB^* = P^*APP^*A^*P = P^*AA^*P = P^*P = I$, i.e B est unitaire. □

1.5 Théorème de Schur

Soit $A \in \mathcal{M}_n(\mathbb{C})$; alors il existe une matrice unitaire U et une matrice triangulaire supérieure T telles que

$$A = UTU^* = UTU^{-1}.$$

Ainsi A est trigonalisable dans une base orthonormée.

Preuve :

On fait une démonstration par récurrence sur la dimension n .

Le théorème est vrai pour $n = 1$: $A = (a) = (1)(a)(1)$.

On suppose le théorème vrai pour les matrices de $\mathcal{M}_{n-1}(\mathbb{C})$:

soit $A \in \mathcal{M}_n(\mathbb{C})$ et soit f l'endomorphisme de \mathbb{C}^n dont A est la matrice dans la base canonique; A possédant au moins une valeur propre λ , considérons x un vecteur propre associé à la valeur propre λ : alors $u_1 = \frac{x}{\|x\|}$ est un vecteur propre unitaire, et par le procédé de Gram-Schmidt on peut construire une base orthonormée $\mathcal{B} = (u_1, u_2, \dots, u_n)$ de \mathbb{C}^n dont u_1 est le premier vecteur. Notons P la matrice de passage de la base canonique à la base \mathcal{B} , alors, d'après la formule de changement de base (V 1.7) on a $I = {}^t P I \bar{P} = {}^t P \bar{P}$ et ainsi P est unitaire.

Posons $B = P^{-1}AP = P^*AP$: B est la matrice de f dans la base \mathcal{B} , donc B est de la forme

$$B = \begin{pmatrix} \lambda & \cdot & \cdot & \cdot \\ 0 & & & \\ \vdots & B_1 & & \\ 0 & & & \end{pmatrix}$$

où B_1 est une matrice de $\mathcal{M}_{n-1}(\mathbb{C})$. Appliquons l'hypothèse de récurrence à B_1 : il existe une matrice unitaire U_1 et une matrice triangulaire supérieure T_1 dans $\mathcal{M}_{n-1}(\mathbb{C})$ telles que $B_1 = U_1 T_1 U_1^*$ ou encore $T_1 = U_1^* B_1 U_1$. Posons alors

$$\tilde{U} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & U_1 & & \\ 0 & & & \end{pmatrix}$$

ainsi

$$\tilde{U}^* = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & U_1^* & & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & U_1^{-1} & & \\ 0 & & & \end{pmatrix}$$

d'où $\tilde{U}^* \tilde{U} = I$ et ainsi \tilde{U} est unitaire. De plus, on a

$$\tilde{U}^* B \tilde{U} = \begin{pmatrix} \lambda & \cdot & \cdot & \cdot \\ 0 & & & \\ \vdots & U_1^* B_1 U_1 & & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \lambda & \cdot & \cdot & \cdot \\ 0 & & & \\ \vdots & T_1 & & \\ 0 & & & \end{pmatrix} = T$$

où T est triangulaire supérieure, ou encore $B = \tilde{U}T\tilde{U}^*$. On en déduit

$$A = PBP^* = P\tilde{U}T\tilde{U}^*P^* = (P\tilde{U})T(P\tilde{U})^* = UTU^*$$

en posant $U = P\tilde{U}$; or $U^* = (P\tilde{U})^* = \tilde{U}^*P^* = \tilde{U}^{-1}P^{-1} = (P\tilde{U})^{-1} = U^{-1}$ donc U est unitaire et ainsi les vecteurs colonnes de U forment une base orthonormée dans laquelle la matrice de f est triangulaire.

□

2 Matrices normales

2.1 Définition

Soit $A \in \mathcal{M}_n(\mathbb{C})$; on dit que A est **normale** si elle commute avec sa matrice adjointe :

$$AA^* = A^*A.$$

2.2 Exemples

- a) Toute matrice hermitienne est normale.
- b) Toute matrice unitaire est normale.

2.3 Proposition

Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice normale; alors on a

- a) pour tout $X \in \mathbb{C}^n$, $\|AX\| = \|A^*X\|$;
- b) pour tout $\lambda \in \mathbb{C}$, $\ker(A - \lambda I) = \ker(A^* - \bar{\lambda}I)$;
- c) les sous-espaces propres de A sont orthogonaux deux à deux.

Preuve :

- a) pour tout $X \in \mathbb{C}^n$ on a :

$$\|AX\|^2 = \langle AX|AX \rangle = \langle X|A^*AX \rangle = \langle X|AA^*X \rangle = \langle A^*X|A^*X \rangle = \|A^*X\|^2 ;$$

- b) pour tout $\lambda \in \mathbb{C}$, $A - \lambda I$ est normale, en effet :

$(A - \lambda I)(A - \lambda I)^* = (A - \lambda I)(A^* - \bar{\lambda}I) = AA^* - \bar{\lambda}A - \lambda A^* + \lambda\bar{\lambda}I = A^*A - \bar{\lambda}A - \lambda A^* + \lambda\bar{\lambda}I = (A^* - \bar{\lambda}I)(A - \lambda I) = (A - \lambda I)^*(A - \lambda I)$; donc, d'après a) on a pour tout $X \in \mathbb{C}^n$, $\|(A - \lambda I)X\| = \|(A - \lambda I)^*X\| = \|(A^* - \bar{\lambda}I)X\|$ et ainsi $(A - \lambda I)X = 0 \iff (A - \lambda I)^*X = 0$ d'où $\ker(A - \lambda I) = \ker(A^* - \bar{\lambda}I)$.

- c) soient λ et μ deux valeurs propres distinctes de A et soient $X \in \ker(A - \lambda I)$ et $Y \in \ker(A - \mu I)$, alors on a : $\langle AX|Y \rangle = \langle \lambda X|Y \rangle = \lambda \langle X|Y \rangle$ mais aussi $\langle AX|Y \rangle = \langle X|A^*Y \rangle = \langle X|\bar{\mu}Y \rangle = \bar{\mu} \langle X|Y \rangle$. On en déduit $(\lambda - \mu)\langle X|Y \rangle = 0$, d'où $\langle X|Y \rangle = 0$ puisque $\lambda \neq \mu$; donc X et Y sont orthogonaux.

□

2.4 Théorème

Soit $A \in \mathcal{M}_n(\mathbb{C})$; alors on a

A est normale \iff Il existe U unitaire et D diagonale telles que $A = UDU^*$.

Ainsi une matrice normale est diagonalisable dans une base orthonormée.

Preuve :

S'il existe U unitaire et D diagonale telles que $A = UDU^*$ alors $A^* = UD^*U^*$ et ainsi $AA^* = UDU^*UD^*U^* = UDD^*U^*$ et $A^*A = UD^*U^*UDU^* = UD^*DU^*$; or D et D^* sont diagonales donc commutent entre elles d'où $AA^* = A^*A$ et A est normale.

Réciproquement, considérons une matrice normale A ; d'après le théorème de Schur, il existe une matrice unitaire U et une matrice triangulaire supérieure T telles que $A = UTU^*$, d'où $A^* = UT^*U^*$. On a alors

$$AA^* = A^*A \iff UTU^*UT^*U^* = UT^*U^*UTU^* \iff U(TT^*)U^* = U(T^*T)U^* \iff TT^* = T^*T$$

ainsi T est aussi une matrice normale. On va montrer par récurrence qu'alors T est diagonale.

Notons t_{ij} le terme général de T , b_{ij} le terme général de TT^* et c_{ij} le terme général de T^*T : on a ainsi pour tous $i, j \in \{1, 2, \dots, n\}$

$$b_{ij} = \sum_{k=1}^n t_{ik} \overline{t_{jk}} = c_{ij} = \sum_{l=1}^n \overline{t_{li}} t_{lj}.$$

En particulier, pour $i = j$

$$\sum_{k=1}^n t_{ik} \overline{t_{ik}} = \sum_{l=1}^n \overline{t_{li}} t_{li} \quad \text{ou encore} \quad \sum_{k=1}^n |t_{ik}|^2 = \sum_{l=1}^n |t_{li}|^2.$$

Or T est triangulaire supérieure, donc $t_{ik} = 0$ pour $i > k$, d'où, pour tout $i \in \{1, 2, \dots, n\}$

$$\sum_{k=i}^n |t_{ik}|^2 = \sum_{l=1}^i |t_{li}|^2 \quad (*)$$

Pour $i = 1$, on obtient donc $\sum_{k=1}^n |t_{1k}|^2 = |t_{11}|^2$, d'où $\sum_{k=2}^n |t_{1k}|^2 = 0$ et par conséquent $t_{1k} = 0$ pour $2 \leq k \leq n$.

Hypothèse de récurrence: dans les $p-1$ premières lignes de T , tous les termes, sauf éventuellement le terme diagonal, sont nuls:

$$t_{ik} = 0 \text{ pour } 1 \leq i \leq p-1, 1 \leq k \leq n \text{ et } i \neq k.$$

L'égalité (*) pour $i = p$ donne

$$\sum_{k=p}^n |t_{pk}|^2 = \sum_{l=1}^p |t_{lp}|^2 = |t_{pp}|^2$$

puisque $t_{lp} = 0$ pour $1 \leq l \leq p-1$ par hypothèse de récurrence. On en déduit que

$$\sum_{k=p+1}^n |t_{pk}|^2 = 0, \text{ d'où } t_{pk} = 0 \text{ pour } p+1 \leq k \leq n; \text{ or } t_{pk} = 0 \text{ pour } 1 \leq k \leq p-1$$

puisque T est triangulaire supérieure. Donc tous les termes de la p ème ligne sont nuls sauf éventuellement le terme diagonal. On en déduit par récurrence que dans la matrice T , tous les termes en dehors de la diagonale sont nuls : T est diagonale.

On a ainsi $A = UTU^* = UTU^{-1}$ avec T diagonale ; de plus, U étant unitaire, ses vecteurs colonnes forment une base orthonormée de \mathbb{C}^n et ainsi A est diagonalisable dans une base orthonormée.

□

2.5 Proposition

Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice **normale** ; alors on a :

- a) A est hermitienne \iff les valeurs propres de A sont réelles ;
- b) A est unitaire \iff les valeurs propres de A sont de module égal à 1.

Preuve :

A est une matrice normale donc il existe U unitaire et D diagonale telles que $A = UDU^*$.

a) Si A est hermitienne, alors $A = A^*$ i.e $UDU^* = UD^*U^*$, d'où $D = D^*$; or D est diagonale

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \text{ donc } D^* = \begin{pmatrix} \overline{\lambda_1} & 0 & \cdots & 0 \\ 0 & \overline{\lambda_2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \overline{\lambda_n} \end{pmatrix}$$

on en déduit $\lambda_i = \overline{\lambda_i}$ pour tout $1 \leq i \leq n$, i.e λ_i est réel.

Réciproquement, si toutes les valeurs propres de A sont réelles, alors $D = D^*$ donc $UDU^* = UD^*U^*$, c'est-à-dire $A = A^*$: A est hermitienne.

b) Si A est unitaire, alors par un raisonnement analogue, on a $D^* = D^{-1}$ donc $\overline{\lambda_i} = (\lambda_i)^{-1}$ pour tout $1 \leq i \leq n$, d'où $|\lambda_i| = \sqrt{\lambda_i \overline{\lambda_i}} = 1$.

Réciproquement, si toutes les valeurs propres λ_i de A sont de module égal à 1, alors $\overline{\lambda_i} = (\lambda_i)^{-1}$ pour tout $1 \leq i \leq n$ et ainsi $D^* = D^{-1}$, d'où $A^* = A^{-1}$: A est unitaire.

□

2.6 Corollaire

a) Soit A une matrice hermitienne de $\mathcal{M}_n(\mathbb{C})$; alors ses valeurs propres sont réelles et A est diagonalisable dans une base orthonormée de \mathbb{C}^n .

a) Soit A une matrice unitaire de $\mathcal{M}_n(\mathbb{C})$; alors ses valeurs propres sont de module égal à 1 et A est diagonalisable dans une base orthonormée de \mathbb{C}^n .

3 Cas des matrices réelles

3.1 Définition et proposition

Soit $A \in \mathcal{M}_n(\mathbb{R})$; on dit que A est **orthogonale** si et seulement si ${}^tAA = I$. Alors A est inversible et $A^{-1} = {}^tA$, d'où $A{}^tA = I$ également. De plus les vecteurs colonnes (resp. lignes) de A forment une base orthonormée de \mathbb{R}^n .

Remarque : Une matrice orthogonale n'est donc rien d'autre qu'une matrice unitaire à coefficients réels; de la même façon, une matrice symétrique à coefficients réels est hermitienne.

3.2 Proposition

a) Soient A et $B \in \mathcal{M}_n(\mathbb{R})$; s'il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{C})$ telle que $A = PBP^{-1}$, alors il existe une matrice inversible $Q \in \mathcal{M}_n(\mathbb{R})$ telle que $A = QBQ^{-1}$. Autrement dit si deux matrices à coefficients réels sont semblables dans $\mathcal{M}_n(\mathbb{C})$, alors elles sont semblables dans $\mathcal{M}_n(\mathbb{R})$.

b) Soit $A \in \mathcal{M}_n(\mathbb{R})$; si on considère A comme une matrice à coefficients complexes et si λ est une valeur propre de A complexe non réelle, alors $\bar{\lambda}$ est aussi une valeur propre de A .

Preuve :

a) S'il existe une matrice inversible $P \in \mathcal{M}_n(\mathbb{C})$ telle que $A = PBP^{-1}$, écrivons $P = P_1 + iP_2$ où P_1 et P_2 sont à coefficients réels; alors $A = PBP^{-1}$ donne $AP = PB$ ou encore $A(P_1 + iP_2) = (P_1 + iP_2)B$, on en déduit, en séparant partie réelle et partie imaginaire, que $AP_1 = P_1B$ et $AP_2 = P_2B$.

Considérons l'application f de \mathbb{C} dans \mathbb{C} définie par $f(z) = \det(P_1 + zP_2)$; alors f est une fonction polynomiale de z , donc f s'écrit sous la forme

$$f(z) = \alpha_p z^p + \dots + \alpha_1 z + \alpha_0$$

où $\alpha_0, \alpha_1, \dots, \alpha_p \in \mathbb{C}$. De plus, $f(i) = \det(P_1 + iP_2) = \det P \neq 0$ donc f n'est pas l'application nulle et par conséquent les coefficients $\alpha_0, \alpha_1, \dots, \alpha_p$ ne sont pas tous nuls. On en déduit alors qu'il existe au moins un réel a tel que $f(a) \neq 0$: en effet si on avait $f(x) = 0$ pour tout réel x , le polynôme $\alpha_p X^p + \dots + \alpha_1 X + \alpha_0$ aurait une infinité de racines dans \mathbb{R} , et par conséquent serait le polynôme nul, ce qui n'est pas le cas puisque les coefficients $\alpha_0, \alpha_1, \dots, \alpha_p$ ne sont pas tous nuls. Posons alors $Q = P_1 + aP_2$: Q est une matrice à coefficients réels et $\det Q = f(a) \neq 0$, i.e Q est inversible; de plus $AQ = A(P_1 + aP_2) = AP_1 + aAP_2 = P_1B + aP_2B = (P_1 + aP_2)B = QB$ d'où $A = QBQ^{-1}$. \square

3.3 Théorème

Soit A une matrice symétrique de $\mathcal{M}_n(\mathbb{R})$; alors il existe une matrice orthogonale P de $\mathcal{M}_n(\mathbb{R})$ et une matrice diagonale $D \in \mathcal{M}_n(\mathbb{R})$ telles que

$$A = PD{}^tP = PDP^{-1}.$$

Autrement dit, A est diagonalisable dans une base orthonormée de \mathbb{R}^n .

Preuve :

La matrice A vue comme matrice de $\mathcal{M}_n(\mathbb{C})$ est hermitienne, donc ses valeurs propres sont réelles et il existe une matrice unitaire $U \in \mathcal{M}_n(\mathbb{C})$ et une matrice diagonale $D \in \mathcal{M}_n(\mathbb{C})$ telles que $A = UDU^* = UDU^{-1}$. La matrice diagonale D est en fait à coefficients réels puisque ses termes diagonaux sont les valeurs propres de A ; donc d'après 3.2, il existe une matrice inversible $Q \in \mathcal{M}_n(\mathbb{R})$ telle que $A = QDQ^{-1}$, autrement dit les vecteurs colonnes de Q forment une base de \mathbb{R}^n formée de vecteurs propres de A . Or, d'après 2.3, les sous-espaces propres de A sont orthogonaux deux à deux; en appliquant alors le procédé de Gram-Schmidt dans chaque sous-espace propre $\ker(A - \lambda_i I)$, on construit une base orthonormée de $\ker(A - \lambda_i I)$ (formée de vecteurs à coordonnées réelles), et en réunissant toutes ces bases, on obtient une base orthonormée $\mathcal{B} = (e_1, e_2, \dots, e_n)$ de \mathbb{R}^n formée de vecteurs propres de A . Soit P la matrice obtenue en prenant (e_1, e_2, \dots, e_n) comme vecteurs colonnes: alors P est une matrice orthogonale, puisque (e_1, e_2, \dots, e_n) est une base orthonormée, et $A = PDP^{-1}$.

□

3.4 Théorème

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice orthogonale; alors il existe des entiers $p, q, k \in \mathbb{N}$ vérifiant $p + q + 2k = n$, une matrice orthogonale P de $\mathcal{M}_n(\mathbb{R})$ et une matrice A' de $\mathcal{M}_n(\mathbb{R})$ telles que $A = PA'P^{-1}$ où A' est de la forme:

$$A' = \begin{pmatrix} I_p & 0 & \cdots & \cdots & 0 \\ 0 & -I_q & \ddots & & \vdots \\ \vdots & \ddots & A_1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & A_k \end{pmatrix}$$

avec

$$A_i = \begin{pmatrix} \cos \theta_i & -\sin \theta_i \\ \sin \theta_i & \cos \theta_i \end{pmatrix} \text{ où } \theta_i \in]-\pi, \pi[-\{0\}$$

Preuve :

La matrice A vue comme matrice de $\mathcal{M}_n(\mathbb{C})$ est unitaire, donc ses valeurs propres sont de module égal à 1, A est diagonalisable dans $\mathcal{M}_n(\mathbb{C})$ et les sous-espaces propres de A sont orthogonaux deux à deux. Notons f l'endomorphisme de \mathbb{C}^n dont A est la matrice dans la base canonique.

Considérons une valeur propre λ de A , on a donc $|\lambda| = 1$; si $\lambda \in \mathbb{R}$, alors $\lambda = 1$ ou $\lambda = -1$ et si $\lambda \notin \mathbb{R}$, alors $\lambda = e^{i\theta}$ où $\theta \in]-\pi, \pi[-\{0\}$ et $\bar{\lambda} = e^{-i\theta}$ est aussi valeur propre de A : si e est vecteur propre de A pour λ , \bar{e} est vecteur propre de A pour $\bar{\lambda}$ et réciproquement, en effet on a:

$$Ae = \lambda e \iff \overline{Ae} = \overline{\lambda e} \iff \overline{A} \bar{e} = \bar{\lambda} \bar{e} \iff A\bar{e} = \bar{\lambda} \bar{e}$$

puisque A est à coefficients réels: on en déduit que $\dim \ker(A - \lambda I) = \dim \ker(A - \bar{\lambda} I)$; ainsi on peut construire une base \mathcal{B} de \mathbb{C}^n telle que

$$M_{\mathcal{B}}(f) = \text{diag}(I_p, -I_q, \lambda_1 I_{m_1}, \bar{\lambda}_1 I_{m_1}, \dots, \lambda_k I_{m_k}, \bar{\lambda}_k I_{m_k})$$

en prenant à chaque fois pour base du sous-espace propre $\ker(A - \bar{\lambda}_i I)$ la base conjuguée de celle du sous-espace propre $\ker(A - \lambda_i I)$. On va construire maintenant à partir de \mathcal{B} , une base orthonormée \mathcal{B}' de \mathbb{R}^n telle que $M_{\mathcal{B}'}(f) = A'$:

Si $\lambda = 1$ ou $\lambda = -1$, alors si e est un vecteur propre de A pour λ , \bar{e} est aussi un vecteur propre de A pour $\lambda = \bar{\lambda}$. Alors $u = \frac{e + \bar{e}}{2}$ et $v = \frac{e - \bar{e}}{2i}$ sont des vecteurs propres de A pour λ à coefficients réels et linéairement indépendants sur \mathbb{R} : ainsi, on peut construire une base du sous-espace $\ker(A - \lambda I)$ formée de vecteurs de \mathbb{R}^n , que l'on peut rendre orthonormée par le procédé de Gram-Schmidt.

Si $\lambda \notin \mathbb{R}$, alors $\lambda = e^{i\theta}$ et si e est vecteur propre de A pour λ , \bar{e} est vecteur propre de A pour $\bar{\lambda}$ et e et \bar{e} sont orthogonaux ; ainsi en posant $u = \frac{e + \bar{e}}{2}$ et $v = \frac{e - \bar{e}}{2i}$, u et v sont à coefficients réels, toujours orthogonaux, et on a

$$A \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

Ainsi, en remplaçant à chaque fois les vecteurs e et \bar{e} de la base \mathcal{B} par les vecteurs u et v correspondants et en réordonnant les vecteurs, on peut construire une base \mathcal{B}' de \mathbb{R}^n , qui est orthonormée, et telle que $M_{\mathcal{B}'}(f)$ soit de la forme A' , d'où le théorème. □

3.5 Etude pratique d'une matrice orthogonale en dimension 2

Soit f un endomorphisme de \mathbb{R}^2 dont la matrice dans la base canonique est orthogonale, alors $\det(A) = \pm 1$:

a) si $\det(A) = 1$, alors

$$A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \quad \text{où } \varphi \text{ est défini } \text{mod}(2\pi).$$

On dit que f est la rotation vectorielle de mesure $\varphi \text{ mod}(2\pi)$ dans le plan vectoriel \mathbb{R}^2 orienté par la base canonique.

De plus f n'est pas diagonalisable sauf si $f = Id$ ou $f = -Id$.

b) si $\det(A) = -1$, alors

$$A = \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix} \quad \text{où } \varphi \text{ est défini } \text{mod}(2\pi).$$

L'endomorphisme f n'est autre que la symétrie orthogonale par rapport à la droite $\ker(A - Id)$ et f est diagonalisable.

Preuve :

La matrice A est de la forme

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

a) Si $\det(A) = 1$, alors $ad - bc = 1$, de plus ${}^t A A = I$ donc $a^2 + c^2 = b^2 + d^2 = 1$, $ab + cd = 0$. On peut donc poser $a = \cos \varphi$ et $c = \sin \varphi$; de plus $ab + cd = 0$ donne $abd + cd^2 = 0$, or

$ad = bc + 1$ donc $b(bc + 1) + cd^2 = 0$ i.e $c(b^2 + d^2) + b = 0 = c + b$ d'où $b = -c = -\sin \varphi$. De même on prouve que $d = a = \cos \varphi$. Donc

$$A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

Si A possède une valeur propre, ce ne peut être que 1 ou -1 d'après 2.5. Or on voit facilement que si $A \neq \pm I$, alors $\det(A - I) \neq 0$ et $\det(A + I) \neq 0$, donc A ne possède pas de valeur propre, donc n'est pas diagonalisable.

b) Si $\det(A) = -1$, par un calcul analogue avec cette fois-ci $\det(A) = ad - bc = -1$, on obtient $a = -d = \cos \varphi$ et $b = c = \sin \varphi$.

Le calcul du polynôme caractéristique donne $P_A(X) = X^2 - 1$: $P_A(X)$ a deux racines réelles distinctes 1 et -1 , donc A est diagonalisable et $f^2 = Id$ d'après le théorème de Cayley-Hamilton.

Un calcul simple montre que le sous-espace propre $\ker(A - Id)$ est la droite de vecteur directeur $e_1 = (\cos(\varphi/2), \sin(\varphi/2))$ et que le sous-espace propre $\ker(A + Id)$ est la droite de vecteur directeur $e_2 = (-\sin(\varphi/2), \cos(\varphi/2))$, donc $\ker(A + Id) = \ker(A - Id)^\perp$. On en déduit que f est la symétrie orthogonale par rapport à la droite $\ker(A - Id)$.

□

3.6 Etude pratique d'une matrice orthogonale en dimension 3

Soit f un endomorphisme de \mathbb{R}^3 dont la matrice A dans la base canonique est orthogonale : son polynôme caractéristique $P_A(X)$ est alors un polynôme de $\mathbb{R}[X]$ de degré 3 donc possède au moins une racine réelle qui ne peut être que 1 ou -1 . On étudie alors $\ker(A - I)$:

a) Si $\dim \ker(A - I) = 3$, alors $A = I$ et $f = Id$.

b) Si $\dim \ker(A - I) = 2$, alors f est la symétrie orthogonale par rapport au plan $\ker(A - I)$: on dit que f est la réflexion vectorielle de plan $\ker(A - I)$.

c) Si $\dim \ker(A - I) = 1$, alors, en prenant e_1 un vecteur unitaire directeur de la droite $\ker(A - I)$, e_2 un vecteur unitaire orthogonal à e_1 , et $e_3 = e_1 \wedge e_2$, la matrice de f dans la base orthonormée (e_1, e_2, e_3) de \mathbb{R}^3 est de la forme :

$$A' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} \quad \text{où } \varphi \neq 0 \text{ mod}(2\pi).$$

On dit que f est la rotation vectorielle d'axe $\Delta = \ker(A - I)$ dirigé par e_1 , de mesure φ dans le plan Δ^\perp orienté par la base (e_2, e_3) .

Calcul de φ : $\cos \varphi = \frac{1}{2}(tr(A) - 1)$ et $\sin \varphi = \langle f(e_2) | e_3 \rangle$.

d) Si $\dim \ker(A - I) = 0$, alors on calcule $\ker(A + I)$ (et dans ce cas seulement !) qui est nécessairement de dimension 1 sauf dans le cas particulier où $A = -I$.

Si $A \neq -I$, en prenant e_1 un vecteur unitaire directeur de la droite $\Delta = \ker(A + I)$, e_2 un vecteur unitaire orthogonal à e_1 , et $e_3 = e_1 \wedge e_2$, la matrice de f dans la base orthonormée (e_1, e_2, e_3) de \mathbb{R}^3 est de la forme :

$$A' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} \quad \text{où } \varphi \neq 0 \text{ mod}(2\pi).$$

Alors $f = s \circ r = r \circ s$ où s est la réflexion de plan $P = \Delta^\perp$ et où r est la rotation d'axe Δ dirigé par e_1 et de mesure φ dans le plan P orienté par la base (e_2, e_3) . On dit que f est la réflexion-rotation d'axe $\Delta = \ker(A + I)$ dirigé par e_1 , de plan $P = \Delta^\perp$ et de mesure φ dans le plan P orienté par la base (e_2, e_3) .

Calcul de φ : $\cos \varphi = \frac{1}{2}(\text{tr}(A) + 1)$ et $\sin \varphi = \langle f(e_2) | e_3 \rangle$.

Preuve :

a) Il est clair que si $\dim \ker(A - I) = 3$, $\ker(A - I) = \mathbb{R}^3$ et $A = I$.

b) Si $\dim \ker(A - I) = 2$, alors 1 est racine de P_A de multiplicité au moins égale à 2 et ainsi P_A possède une troisième racine réelle qui ne peut être que 1 ou -1 . Si -1 n'est pas racine de P_A , alors 1 est racine de multiplicité 3 de P_A , donc $P_A(X) = -(X - 1)^3$. Or, comme -1 n'est pas valeur propre de f , d'après 3.4 il existe une base orthonormée de \mathbb{R}^3 dans laquelle la matrice A' de f est de la forme :

$$A' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} \quad \text{où } \varphi \neq \pi \text{ mod}(2\pi).$$

Alors $P_{A'}(X) = -(X - 1)(X^2 - 2X \cos \varphi + 1) = P_A(X)$ donc $\cos \varphi = 1$ et ainsi $A' = I$ d'où $f = Id$, ce qui est impossible puisque $\dim \ker(A - Id) = 2$. Donc -1 est racine de P_A et ainsi A est diagonalisable avec -1 comme valeur propre de multiplicité 1, et 1 comme valeur propre de multiplicité 2 donc A est semblable à la matrice

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

on en déduit que $D^2 = I$ et ainsi $f^2 = Id$. De plus, d'après 2.3, les sous-espaces propres $\ker(A - I)$ et $\ker(A + I)$ sont orthogonaux donc $\ker(A + I) = \ker(A - I)^\perp$ puisque $\dim \ker(A - I) = 2$ et $\dim \ker(A + I) = 1$. On en déduit que f est la symétrie orthogonale par rapport au plan $\ker(A - I)$ d'après IV 5.3.

c) Si $\dim \ker(A - I) = 1$, alors $\ker(A - I)$ est une droite Δ et la restriction de f à Δ^\perp ne possède plus 1 comme valeur propre. Donc, d'après 3.4, en prenant e_1 un vecteur unitaire directeur de la droite Δ , e_2 un vecteur unitaire orthogonal à e_1 , et $e_3 = e_1 \wedge e_2$, la matrice de f dans la base orthonormée (e_1, e_2, e_3) de \mathbb{R}^3 est de la forme :

$$A' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} \quad \text{où } \varphi \neq 0 \text{ mod}(2\pi).$$

d) Si $\dim \ker(A - I) = 0$, alors 1 n'est pas valeur propre de f et nécessairement -1 l'est. On en déduit, par un raisonnement analogue à celui fait dans b), que $\dim \ker(A + I) = 1$ ou $\dim \ker(A + I) = 3$, ce dernier cas signifiant que $A = -I$ et ainsi $f = -Id$.

Si $\dim \ker(A + I) = 1$, alors $\ker(A + I)$ est une droite Δ , et en prenant e_1 un vecteur unitaire directeur de la droite Δ , e_2 un vecteur unitaire orthogonal à e_1 , et $e_3 = e_1 \wedge e_2$, la matrice de f dans la base orthonormée (e_1, e_2, e_3) de \mathbb{R}^3 est de la forme :

$$A' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} \quad \text{où } \varphi \neq 0 \text{ mod}(2\pi).$$

Alors $f = s \circ r = r \circ s$ où s est la réflexion de plan $P = \Delta^\perp$ et où r est la rotation d'axe Δ dirigé par e_1 et de mesure φ dans le plan P orienté par la base (e_2, e_3) . En effet $A' = A_1 A_2 = A_2 A_1$ où :

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} \quad \text{et } A_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

□

3.6 Etude d'une matrice orthogonale en dimension 4

Soit f un endomorphisme de \mathbb{R}^4 dont la matrice A dans la base canonique est orthogonale ; alors, d'après le théorème 3.4, si $p = \dim \ker(A - I)$ et $q = \dim \ker(A + I)$, $p + q$ est pair et il y a 4 possibilités pour la forme réduite de la matrice :

1er cas : $A' = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$ avec $p + q = 4$ (éventuellement p ou q nul) ;

2ème cas : $A' = \begin{pmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \cos \theta & -\sin \theta \\ 0 & 0 & \sin \theta & \cos \theta \end{pmatrix}$ où $\theta \in]-\pi, \pi[- \{0\}$ (cas $p + q = 2$).

3ème cas : $A' = \begin{pmatrix} \cos \theta_1 & -\sin \theta_1 & 0 & 0 \\ \sin \theta_1 & \cos \theta_1 & 0 & 0 \\ 0 & 0 & \cos \theta_2 & -\sin \theta_2 \\ 0 & 0 & \sin \theta_2 & \cos \theta_2 \end{pmatrix}$ où θ_1 et $\theta_2 \in]-\pi, \pi[- \{0\}$ (cas $p = q = 0$).

Comme A est orthogonale, si A possède des valeurs propres réelles, ce ne peut être que 1 ou -1 ; on commence donc par calculer $p = \dim \ker(A - I)$ et $q = \dim \ker(A + I)$:

a) si $p + q = 4$, alors A est diagonalisable et on est dans le premier cas ;

b) si $0 < p + q < 4$, alors nécessairement on a $p + q = 2$ (éventuellement p ou $q = 0$) et on est dans le deuxième cas ; on calcule une base orthonormée (u_1, u_2) du plan $P = \ker(A - I) \oplus \ker(A + I)$ et, d'après 1.4, P^\perp est stable par f puisque P l'est ; on calcule alors une base orthonormée (u_3, u_4) du plan P^\perp et la matrice de f dans la base orthonormée (u_1, u_2, u_3, u_4) a la forme réduite ;

c) si $p = q = 0$, alors on est dans le 3ème cas : A ne possède pas de valeurs propres réelles, et on doit travailler dans \mathbb{C} ; on suit alors la démarche expliquée dans le théorème 3.4 : on calcule le polynôme caractéristique de A et on cherche ses racines complexes qui sont nécessairement de module 1 et conjuguées deux à deux. Soit $e^{i\theta_1}$ l'une d'entre elles : on calcule un vecteur propre $e \in \mathbb{C}^4$ pour $e^{i\theta_1}$ et ainsi \bar{e} est vecteur propre pour $e^{-i\theta_1}$, puis on considère les deux vecteurs de \mathbb{R}^4 $u = \frac{e + \bar{e}}{2}$ et $v = \frac{e - \bar{e}}{2i}$ qui sont orthogonaux et on pose $u_1 = u/\|u\|$ et $u_2 = v/\|v\|$; alors le plan P de base orthonormée (u_1, u_2) est stable par f et la matrice de la restriction de f à P est exactement

$$\begin{pmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{pmatrix}$$

Ensuite on détermine P^\perp , qui est donc stable par f ; il ne reste plus qu'à calculer une base orthonormée (u_3, u_4) de P^\perp et la matrice de la restriction de f à P^\perp est du type

$$\begin{pmatrix} \cos \theta_2 & -\sin \theta_2 \\ \sin \theta_2 & \cos \theta_2 \end{pmatrix}$$

où $e^{i\theta_2}$ et $e^{-i\theta_2}$ sont les deux autres valeurs propres de A . La matrice de f dans la base orthonormée (u_1, u_2, u_3, u_4) est alors la matrice réduite

$$A' = \begin{pmatrix} \cos \theta_1 & -\sin \theta_1 & 0 & 0 \\ \sin \theta_1 & \cos \theta_1 & 0 & 0 \\ 0 & 0 & \cos \theta_2 & -\sin \theta_2 \\ 0 & 0 & \sin \theta_2 & \cos \theta_2 \end{pmatrix}.$$