Algèbre linéaire 2

Julia Matos

Université d'Evry Val d'Essonne
L2 PCSPI

Année 2012/2013

Table des matières

1 Déterminant d'une matrice **2**
- 1.1 Matrices . . . 2
- 1.2 Définition du déterminant d'une matrice . . . 4
- 1.3 Propriétés du déterminant . . . 4

2 Rappels d'algèbre linéaire **7**
- 2.1 Espaces vectoriels. Sous-espaces vectoriels. . . . 7
- 2.2 Applications linéaires . . . 11
- 2.3 Matrice d'une application linéaire . . . 12
- 2.4 Changement de base . . . 14
- 2.5 Matrices semblables . . . 15

3 Diagonalisation **17**
- 3.1 Valeur propre. Vecteur propre . . . 17
- 3.2 Polynôme caractéristique . . . 18
- 3.3 Sous-espaces propres . . . 21
- 3.4 Endomorphisme et matrice diagonalisables . . . 21
- 3.5 Théorème de Cayley-Hamilton . . . 25
- 3.6 Endomorphisme et matrice triagonalisables . . . 26

4 Applications : systèmes de récurrence et systèmes différentiels **27**
- 4.1 Systèmes de récurrence linéaires . . . 27
- 4.2 Systèmes différentiels linéaires . . . 30

5 Orthonormalité **33**
- 5.1 Produit scalaire sur $\mathbb{R}^n$ . . . 33
- 5.2 Bases orthonormées . . . 35
- 5.3 Orthonormalisation de Gram-Schmidt . . . 36
- 5.4 Matrices orthogonales. . . . 37
- 5.5 Endomorphisme symétrique . . . 38
- 5.6 Isométrie . . . 40
- 5.7 Produit hermitien sur $\mathbb{C}^n$ . . . 41

1 Déterminant d'une matrice

L'ensemble $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$. Les éléments de $\mathbb{K}$ seront souvent appelés des *scalaires*.

1.1 Matrices

Soient n, p des entiers positifs. Une *matrice* $n \times p$ à coefficients dans $\mathbb{K}$ est un tableau d'éléments de $\mathbb{K}$ à n lignes et p colonnes, que l'on note

$$A = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix} \quad \text{ou en abrégé} \quad A = (a_{ij}).$$

Les a_{ij} s'appellent les *coefficients de la matrice*, où i désigne l'indice de la ligne et j indique l'indice de la colonne.

Deux matrices $A = (a_{ij})$ et $B = (b_{ij})$ sont telles que $A = B$ si et seulement si $a_{ij} = b_{ij}$ pour tous $1 \leq i \leq n$ et tous $1 \leq j \leq p$.

L'ensemble des matrices $n \times p$ à coefficients dans $\mathbb{K}$ se note $M_{n,p}(\mathbb{K})$ et plus simplement $M_n(\mathbb{K})$ si $n = p$.

Si $n = p$, on dit que la matrice est *carrée*. Elle a le même nombre de lignes que de colonnes. Dans ce cas, les nombres a_{ii}, pour $1 \leq i \leq n$, s'appellent les *coefficients diagonaux* de A.

On dit que la matrice A est :
- *diagonale* si $a_{ij} = 0$ pour tout $i \neq j$,
- *triangulaire supérieure* si $a_{ij} = 0$ pour tout $i > j$,
- *triangulaire inférieure* si $a_{ij} = 0$ pour tout $i < j$.

Exemples.

1. Soient $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$. Alors $A \neq B$.
2. Si $n \geq 1$, on note I_n la matrice identité de $M_n(\mathbb{R})$, la matrice diagonale dont tous les coefficients diagonaux sont égaux à 1.
3. La matrice $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 3 \\ 0 & 0 & -4 \end{pmatrix}$ de $M_3(\mathbb{R})$ est triangulaire supérieure.

Si $A = (a_{ij})$, la i-ième ligne de la matrice A est $\begin{pmatrix} a_{i1} & \dots & a_{ip} \end{pmatrix}$ et la j-ième colonne de A est

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}.$$

Opérations sur les matrices.

Soient $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices dans $M_{n,p}(\mathbb{K})$ et $\alpha \in \mathbb{K}$. On définit $A + B$ la matrice somme et αA le produit de la matrice A par le scalaire α par : $A + B = (a_{ij} + b_{ij})$ et $\alpha A = (\alpha a_{ij})$.

Si $A = \begin{pmatrix} a_1 & \dots & a_p \end{pmatrix}$ est une matrice ligne dans $M_{1,p}(\mathbb{K})$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}$ est une matrice colonne dans $M_{p,1}(\mathbb{K})$. Le produit de A par B, noté AB, est la matrice 1×1 dont le coefficient est $a_1b_1 + \ldots + a_pb_p$.

Soient $A \in M_{n,p}(\mathbb{K})$ et $B \in M_{p,1}(\mathbb{K})$. Alors la matrice produit AB est la matrice dans $M_{n,1}(\mathbb{K})$ dont la i-ième ligne est le produit de la i-ième ligne de A par la matrice colonne B.

Exemple. Soient $A \in M_{4,3}(\mathbb{R})$ et $B \in M_{3,1}(\mathbb{R})$ définies par

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 3 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 1 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Alors, $AB = \begin{pmatrix} 4 \\ -5 \\ 0 \\ 4 \end{pmatrix}$.

Soient $A \in M_{n,p}(\mathbb{K})$ et $B \in M_{p,q}(\mathbb{K})$. Alors la matrice produit AB est la matrice dans $M_{n,q}(\mathbb{K})$ dont la j-ième colonne est le produit de A par la j-ième colonne de B.

Remarques.

1. Le produit AB est possible uniquement lorsque le numéro de colonnes de A est égal au numéro de lignes de B.
2. En général, si $A, B \in M_n(\mathbb{R})$, $AB \neq BA$. Par exemple, si $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, alors
$$AB = \begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 2 \\ -1 & 0 \end{pmatrix}.$$

Définition 1.1 *On dit que $A \in M_n(\mathbb{K})$ est inversible s'il existe une matrice $B \in M_n(\mathbb{K})$ telle que $AB = BA = I_n$. La matrice B est unique et s'appelle l'inverse de A. On note $B = A^{-1}$.*

Exemples.

1. On a $I_n^{-1} = I_n$, pour tout $n \geq 1$.
2. Soit $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$. Alors, $A^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{pmatrix}$.
3. Soit $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Alors, $A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$.
4. Si A est inversible, alors A^{-1} est inversible et $(A^{-1})^{-1} = A$.

1.2 Définition du déterminant d'une matrice

Si $A \in M_n(\mathbb{K})$, le déterminant de A, noté $\det A$ ou $|A|$, est un élément de $\mathbb{K}$ définit par récurrence par :

si $n = 1$, $A = (a)$ et $\det A = a$,

si $n \geq 2$, $A = (a_{ij})$ alors

$$\det A = a_{11}\Delta_{11} - a_{21}\Delta_{21} + \ldots + (-1)^{n+1}a_{n1}\Delta_{n1},$$

où Δ_{i1} est le déterminant de la matrice de $M_{n-1}(\mathbb{R})$ obtenue à partir de A en rayant la première colonne et la i-ième ligne.

Si $n \geq 2$, le déterminant d'une matrice $n \times n$ est donc défini au moyen de n déterminants de matrices $(n-1) \times (n-1)$.

Exemples

1. Si $n = 2$ et $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, alors

$$\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

2. Soit $A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 3 & 1 \\ 2 & 4 & 1 \end{pmatrix}$. Alors,

$$\det A = \begin{vmatrix} 1 & 2 & 1 \\ -1 & 3 & 1 \\ 2 & 4 & 1 \end{vmatrix} = 1 \times \begin{vmatrix} 3 & 1 \\ 4 & 1 \end{vmatrix} + 1 \times \begin{vmatrix} 2 & 1 \\ 4 & 1 \end{vmatrix} + 2 \times \begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} = -1 - 2 - 2 = -5.$$

Noter que les parenthèses sont utilisées pour désigner une matrice et les barres verticales pour désigner le déterminant d'une matrice.

1.3 Propriétés du déterminant

Calculer un déterminant en utilisant la définition est en général labourieux. Nous allons étudier des propriétés qui rendent plus simple le calcul d'un déterminant.

Proposition 1.1 *Soit $A \in M_n(\mathbb{K})$. Si A est une matrice triangulaire, supérieure ou inférieure, alors le déterminant de A est égal au produit des coefficients diagonaux de A.*

Preuve. Par récurrence sur $n \geq 1$.

Ainsi, pour tout $n \geq 1$, $\det I_n = 1$.

Proposition 1.2 *Soit $a \in \mathbb{K}$. Si l'on multiplie par a l'une des lignes d'une matrice de $M_n(\mathbb{K})$, alors le déterminant est multiplié par a.*

En particulier, si une ligne de la matrice A est nulle, il vient $\det A = 0 \times \det A = 0$. D'autre part, si $A \in M_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$, alors $\det(\lambda A) = \lambda^n \det A$ (puisque chaque ligne de A est multiplié par λ).

Proposition 1.3 *Soient $A, A', A'' \in M_n(\mathbb{K})$. Si A, A' et A'' ne diffèrent que par leur i-ième ligne et si la i-ième ligne de A est la somme des i-ièmes lignes de A' et A'' alors*

$$\det A = \det A' + \det A''.$$

Proposition 1.4 *Supposons $n \geq 2$. Si l'on échange deux lignes d'une matrice de $M_n(\mathbb{K})$ alors le déterminant est multiplié par -1.*

Proposition 1.5 *Si $A \in M_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$, alors le déterminant de A ne change pas si l'on ajoute à une ligne de A le produit par λ d'une autre ligne de A.*

Soit $A \in M_n(\mathbb{K})$ et notons $L_1, \ldots, L_n$ les n lignes de A. Le résultat précedent s'écrit :

$$\det A = \begin{vmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_n \end{vmatrix} \underset{L_i+\lambda L_j \to L_i}{=} \begin{vmatrix} L_1 \\ \vdots \\ L_i + \lambda L_j \\ \vdots \\ L_n \end{vmatrix}.$$

Par conséquent, si deux lignes de A sont égales, $\det A = 0$.

Théorème 1.1 *Soient $A, B \in M_n(\mathbb{K})$. Alors,* $\det(AB) = (\det A)(\det B)$.

Théorème 1.2 *Une matrice carrée A est inversible si et seulement si son déterminant est non nul. De plus, si A est inversible,* $\det(A^{-1}) = \frac{1}{\det(A)}$.

Exemple. Soit $n = 2$ et $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Alors A est inversible si et seulement si $ad - bc \neq 0$.

Proposition 1.6 *Si $A \in M_n(\mathbb{K})$ alors* $\det({}^tA) = \det A$.

Grâce à cette propriété chacune des propriétés des déterminants relatives aux lignes de la matrice fournit, en utilisant la transposée de la matrice, une propriété relative aux colonnes :

- si l'on multiplie une colonne de A par $a \in \mathbb{K}$, alors le déterminant est multiplié par a,
- si l'on échange deux colonnes de la matrice, le déterminant est multiplié par -1,
- si deux colonnes de A sont égales, alors $\det A = 0$,
- le déterminant ne change pas lorsqu'on ajoute à une colonne λ fois une autre colonne.

Exemple. On a

$$\begin{vmatrix} 1 & \lambda x + x' & 0 \\ 2 & \lambda y + y' & 1 \\ 1 & \lambda z + z' & 1 \end{vmatrix} = \begin{vmatrix} 1 & \lambda x & 0 \\ 2 & \lambda y & 1 \\ 1 & \lambda z & 1 \end{vmatrix} + \begin{vmatrix} 1 & x' & 0 \\ 2 & y' & 1 \\ 1 & z' & 1 \end{vmatrix} = \lambda \begin{vmatrix} 1 & x & 0 \\ 2 & y & 1 \\ 1 & z & 1 \end{vmatrix} + \begin{vmatrix} 1 & x' & 0 \\ 2 & y' & 1 \\ 1 & z' & 1 \end{vmatrix}.$$

Dans la définition de déterminant, la première colonne de la matrice joue un rôle privilégié. Le résultat suivant montre qu'on peut faire jouer un rôle analogue à n'importe quelle colonne et même à n'importe quelle ligne de la matrice.

Soit $n \geq 2$ et $A \in M_n(\mathbb{K})$. Si Δ_{kl} est le déterminant de la matrice $M_{n-1}(\mathbb{K})$ obtenue à partir de A en rayant la k-ième ligne et la p-ième colonne alors, pour tous $1 \leq i, j \leq n$, on a

$$\det A = (-1)^{i+1}\left(a_{i1}\Delta_{i1} - a_{i2}\Delta_{i2} + \ldots + (-1)^{n+1}a_{in}\Delta_{in}\right). \tag{1}$$

et

$$\det A = (-1)^{j+1}\left(a_{1j}\Delta_{1j} - a_{2j}\Delta_{2j} + \ldots + (-1)^{n+1}a_{nj}\Delta_{nj}\right). \tag{2}$$

Si l'on choisit $j = 1$ dans la seconde formulation, on retrouve la définition du déterminant.

Calculer $\det A$ en utilisant la formule (1) s'appelle développer le déterminant de A selon la i-ième ligne. Alors que calculer $\det A$ en utilisant la formule (2) s'appelle développer le déterminant de A selon la j-ième colonne.

Pour calculer un déterminant, on doit :

- choisir une ligne (ou une colonne) qui a le plus de 0,
- si cela est possible, augmenter le nombre de zéros par opérations élémentaires sur les colonnes ou sur les lignes,
- développer le déterminant selon cette ligne (ou colonne), sans oublier l'alternance des signes.

Remarque. Pour savoir par quel signe il faut commencer quand on développe un déterminant selon une ligne ou une colonne, on peut remplir un tableau à n lignes et n colonnes avec des signes : les signes sont alternés, on commence en haut à gauche par un signe $+$. Par exemple, les tableaux pour $n = 3$, $n = 4$ et $n = 5$ sont, respectivement,

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}, \quad \begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}, \quad \begin{pmatrix} + & - & + & - & + \\ - & + & - & + & - \\ + & - & + & - & + \\ - & + & - & + & - \\ + & - & + & - & + \end{pmatrix}.$$

Exemple. Soit $A = \begin{pmatrix} x+1 & -2 & 2 & -2 \\ -1 & x & -1 & 1 \\ -2 & 2 & x-3 & 2 \\ -1 & 1 & -1 & x \end{pmatrix}$. Alors,

$$\begin{aligned}
\det A &= \begin{vmatrix} x+1 & -2 & 2 & -2 \\ -1 & x & -1 & 1 \\ -2 & 2 & x-3 & 2 \\ -1 & 1 & -1 & x \end{vmatrix} \underset{C_4+C_3\to C_4}{=} \begin{vmatrix} x+1 & -2 & 2 & 0 \\ -1 & x & -1 & 0 \\ -2 & 2 & x-3 & x-1 \\ -1 & 1 & -1 & x-1 \end{vmatrix} \underset{L_3-L_4\to L_3}{=} \\
&= \begin{vmatrix} x+1 & -2 & 2 & 0 \\ -1 & x & -1 & 0 \\ -1 & 1 & x-2 & 0 \\ -1 & 1 & -1 & x-1 \end{vmatrix} = (x-1)\begin{vmatrix} x+1 & -2 & 2 \\ -1 & x & -1 \\ -1 & 1 & x-2 \end{vmatrix} \underset{C_3+C_2\to C_3}{=} \\
&= (x-1)\begin{vmatrix} x+1 & -2 & 0 \\ -1 & x & x-1 \\ -1 & 1 & x-1 \end{vmatrix} \underset{L_2-L_3\to L_2}{=} (x-1)\begin{vmatrix} x+1 & -2 & 0 \\ 0 & x-1 & 0 \\ -1 & 1 & x-1 \end{vmatrix} \\
&= (x-1)^2\begin{vmatrix} x+1 & -2 \\ 0 & x-1 \end{vmatrix} = (x-1)^2(x+1)(x-1) \\
&= (x-1)^3(x+1).
\end{aligned}$$

2 Rappels d'algèbre linéaire

Comme dans le chapitre précédent $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$. Les éléments de $\mathbb{K}$ seront souvent appelés des *scalaires*.

2.1 Espaces vectoriels. Sous-espaces vectoriels.

Un *espace vectoriel sur* $\mathbb{K}$ ou un $\mathbb{K}$-*espace vectoriel*, est un ensemble non vide E dont les éléments s'appellent des *vecteurs*, et qui est muni de deux opérations, la somme de vecteurs et la multiplication d'un vecteur par un scalaire, vérifiant un certain nombre de propriétés.

La somme des vecteurs x et y de E est le vecteur de E noté $x+y$. De plus, si λ est un élément de $\mathbb{K}$, la multiplication du vecteur x par la scalaire λ est le vecteur de E noté λx.

Les règles de calcul sont les suivantes :

- pour tout $x, y \in E$, on a $x+y = y+x$,
- pour tous $x, y, z \in E$, on a $(x+y)+z = x+(y+z)$ et ce vecteur est noté $x+y+z$,
- il existe un vecteur $e \in E$ tel que $x+e = x$ pour tout $x \in E$; de plus, pour tout $y \in E$, il existe un vecteur $y' \in E$ tel que $y+y' = e$,
- pour tout $x \in E$, on a $1x = x$,
- pour tout $x \in E$ et tous $\lambda, \mu \in \mathbb{K}$, on a $(\lambda+\mu)x = \lambda x + \mu x$ et $(\lambda\mu)x = \lambda(\mu x)$, ce dernier vecteur étant noté $\lambda\mu x$,
- pour tous $x, y \in E$ et pour tout $\lambda \in \mathbb{K}$, on a $\lambda(x+y) = \lambda x + \lambda y$.

L'élément e défini ci-dessus est nécessairement unique. Il s'appelle le *vecteur nul* de E et se note 0_E. Alors, on en déduit que pour tout $y \in E$, il existe un unique $y' \in E$ tel que $y + y' = 0_E$ et $y' = -y$.

Exemples.

1. L'ensemble $\mathbb{K}$ lui-même est un espace vectoriel sur $\mathbb{K}$. L'addition et la multiplication par un scalaire.
2. L'ensemble des matrices $n \times p$ à coefficients dans $\mathbb{K}$ est un espace vectoriel sur $\mathbb{K}$. Les opérations somme et multiplication par un scalaire sont celles qui ont été définies dans le chapitre 1. On note cet espace $M_{n,p}(\mathbb{K})$.
3. Pour un entier $n \geq 1$, l'espace $\mathbb{K}^n$ est un espace vectoriel sur $\mathbb{K}$, muni des opérations usuelles. Pour tous $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{K}^n$, $\lambda \in \mathbb{K}$,

$$x + y = (x_1 + y_1, \ldots, x_n + y_n), \quad \lambda x = (\lambda x_1, \ldots, \lambda x_n).$$

   Le vecteur $0_{\mathbb{K}^n} = (0, \ldots, 0)$ est le vecteur nul.

Définition 2.1 *Soit F une partie de E. On dit que F est un sous-espace vectoriel de E si*

1. $0_E \in F$,
2. *pour tous $x, y \in F$, $x + y \in F$,*
3. *pour tout $x \in F$ et pour tout $\lambda \in \mathbb{K}$, $\lambda x \in F$.*

Remarque. Les conditions 2 et 3 peuvent s'écrire : pour tous $x, y \in F$ et pour tous $\lambda, \mu \in \mathbb{K}$, $\lambda x + \mu y \in F$, c'est-à-dire que toute combinaison linéaire d'éléments de F appartient encore à F.

Proposition 2.1 *Soient $n \geq 2$ et $a_1, \ldots, a_n \in \mathbb{K}$. Soit F l'ensemble des vecteurs $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ tels que $a_1x_1 + \ldots + a_nx_n = 0$. Alors F est un sous-espace vectoriel de $\mathbb{K}^n$.*

Proposition 2.2 *Soient $n, k \in \mathbb{N}$ et $v_1, \ldots v_k$ des vecteurs de $\mathbb{K}^n$. Soit F l'ensemble de vecteurs de $\mathbb{K}^n$ qui sont combinaison linéaire de $v_1, \ldots v_k$. Alors, F est un sous-espace vectoriel de $\mathbb{K}^n$ et s'appelle le sous-espace vectoriel engendré par les vecteurs $v_1, \ldots v_k$.*

On note $F =< v_1, \ldots v_k >$ ou $F = \mathrm{Vect}(v_1, \ldots, v_k)$.

Exemple. Soit $F = \{(x, y, z) \in \mathbb{R}^3 : x - 2y + z = 0\}$. Alors, si $(x, y, z) \in F$, on a $x = 2y - z$ et

$$(x, y, z) = (2y - z, y, z) = y(2, 1, 0) + z(-1, 0, 1).$$

On en déduit que F est engendré par les vecteurs $(2, 1, 0)$ et $(-1, 0, 1)$, c'est-à-dire

$$F =< (2, 1, 0), (-1, 0, 1) > .$$

Définition 2.2 *Soient F et G deux sous-espaces vectoriels de E. On dit que F et G sont supplémentaires si tout vecteur de E s'écrit comme la somme d'un vecteur de F et d'un vecteur de G. Cette propriété se note $E = F \oplus G$.*

Proposition 2.3 *Soient F et G deux sous-espaces vectoriels de E. Les sous-espaces F et G sont supplémentaires si et seulement si $E = F + G$ et $F \cap G = \{0_E\}$.*

Exemple. On a $\mathbb{R}^2 = F \oplus G$ où $F =< (1, 0) >$ et $G =< (0, 1) >$.

Définition 2.3 *Soient n un entier positif et $x_1, \ldots, x_n$ des vecteurs de E. Les vecteurs sont dits linéairement indépendants si pour tous $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, on a l'implication*

$$\lambda_1x_1 + \ldots + \lambda_nx_n = 0 \implies \lambda_1 = \ldots = \lambda_n = 0.$$

Nous supposons ensuite que $E \neq \{0_E\}$, c'est-à-dire que E contient des vecteurs autres que le vecteur nul.

Définition 2.4 *Soient n un entier positif et $e_1, \ldots, e_n$ des vecteurs de E. On dit que $(e_1, \ldots, e_n)$ est une base de E si les vecteurs $e_1, \ldots, e_n$ sont linéairement indépendants et si E est engendré par les vecteurs $e_1, \ldots, e_n$.*

Exemples.

1. Si $n \geq 2$ et $e_1, \ldots, e_n$ sont les vecteurs de $\mathbb{K}^n$ définis par

$$e_1 = (1, 0, \ldots, 0), \;\; e_2 = (0, 1, 0, \ldots, 0), \;\; \ldots, \;\; e_n = (0, \ldots, 0, 1),$$

alors $(e_1, \ldots, e_n)$ est une base de $\mathbb{K}^n$. Cette base s'appelle la base canonique de $\mathbb{K}^n$.

2. Considérons les matrices suivantes de $M_2(\mathbb{K})$:

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \;\; E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \;\; E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \;\; E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

Alors $(E_{11}, E_{12}, E_{21}, E_{22})$ est une base de $M_2(\mathbb{K})$.

Théorème 2.1 (Théorème de la base incomplète) *Soient $p \in \mathbb{N}$ et $f_1, \dots f_p$ des vecteurs de E linéairement indépendants. Soit $q \in \mathbb{N}$ et $g_1, \dots, g_q$ des vecteurs de E. On suppose que E est engendré par $g_1, \dots, g_q$. Alors il existe un entier n et une base $(e_1, \dots, e_n)$ de E tels que : $p \leq n \leq q$ et*

$$e_i = f_i \quad \textit{pour tout} \quad i \leq p \quad \textit{et} \quad e_i \quad \textit{est un des vecteurs} \quad g_j \quad \textit{pour tout} \quad p < i \leq n.$$

Comme conséquence de ce théorème, si E est engendré par un nombre fini de vecteurs, alors il existe une base de E.

Proposition 2.4 *Soit $(e_1, \dots, e_n)$ une base de E. Pour tout $x \in E$, il existe $x_1, \dots, x_n \in \mathbb{K}$ uniques tels que $x = x_1 e_1 + \dots + x_n e_n$. Les scalaires $x_1, \dots, x_n$ s'appellent les coordonnées de x dans la base $(e_1, \dots, e_n)$.*

Exemple. Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{K}^n$. Alors, pour tout $x = (x_1, \dots, x_n) \in \mathbb{K}^n$,

$$x = \sum_{i=1}^{n} x_i e_i.$$

Donc, les scalaires $x_1, \dots, x_n$ sont les coordonnées de x dans la base $(e_1, \dots, e_n)$.

Théorème 2.2 *Si E est engendré par un nombre fini de vecteurs, alors toutes les bases de E ont le même nombre de vecteurs. Ce nombre s'appelle la dimension de E et se note* $\dim E$.

Notons que si $E = \{0_E\}$ alors $\dim E = 0$.

Exemple. L'espace $\mathbb{K}^n$ a dimension n. Soient $v_1, \dots, v_p$ des vecteurs de $\mathbb{K}^n$. On a les propriétés suivantes :

1. Si $p > n$, $v_1, \dots, v_p$ ne sont pas linéairement indépendants.
2. Si $p = n$ et $v_1, \dots, v_p$ sont linéairement indépendants, alors $(v_1, \dots, v_p)$ est une base de E.
3. Toutes les bases de $\mathbb{K}^n$ ont n vecteurs.
4. Si $v_1, \dots, v_p$ sont linéairement indépendants, alors $p \leq n$.

Proposition 2.5 *Supposons E de dimension $n \geq 1$. Soient $(e_1, \dots, e_n)$ une base de E et $f_1, \dots, f_n$ des vecteurs de E. Notons A la matrice de $M_n(\mathbb{K})$ dont les coefficients de la j-ième colonne sont les coordonnées du vecteur f_j dans la base $(e_1, \dots, e_n)$. Alors, $(f_1, \dots, f_n)$ est une base de E si et seulement si la matrice A est inversible.*

Exemple. Soit m un nombre réel. On considère les vecteurs de $\mathbb{R}^4$: $f_1 = (1, m, 1, 0)$, $f_2 = (1, 0, 0, 1)$, $f_3 = (1, 1, m, 1)$, $f_4 = (1, 1, 1, 1)$. On considère (e_1, e_2, e_3, e_4) la base canonique de $\mathbb{R}^4$. D'après le résultat précédent, (f_1, f_2, f_3, f_4) est une base de $\mathbb{R}^4$ si et seulement si la matrice

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

est inversible, donc si et seulement si $\det A \neq 0$. On a

$$\det A = \begin{vmatrix} 1 & 1 & 1 & 1 \\ m & 0 & 1 & 1 \\ 1 & 0 & m & 1 \\ 0 & 1 & 1 & 1 \end{vmatrix} \underset{L_1 - L_4 \to L_1}{=} \begin{vmatrix} 1 & 0 & 0 & 0 \\ m & 0 & 1 & 1 \\ 1 & 0 & m & 1 \\ 0 & 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 0 & m & 1 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ m & 1 \end{vmatrix} = 1 - m.$$

On conclut que (f_1, f_2, f_3, f_4) est une base de $\mathbb{R}^4$ si et seulement si $m \neq 1$.

Supposons E de dimension finie. Si F est un sous-espace vectoriel de E, F est de dimension finie, inférieure ou égale à n. De plus, si $\dim F = n$, alors $F = E$. Si F est un sous-espace vectoriel de E de dimension p, alors il existe une base de E dont les p premiers vecteurs forment une base de F (conséquence du théorème de la base incomplète).

Proposition 2.6 *Si E est de dimension finie et si $E = F \oplus G$ alors* $\dim E = \dim F + \dim G$.

Remarque. Soit F un sous-espace vectoriel de $\mathbb{K}^n$. Il y a deux façons de décrire F :

1. se donner des équations en F, autrement dit considérer F comme l'ensemble des vecteurs $x = (x_1, \ldots, x_N) \in \mathbb{K}^n$ solutions d'un système d'équations linéaires,
2. se donner une base de F, ce qui permet de paramètrer F, c'est-à-dire trouver explicitement les vecteurs de F.

Chaque description a son utilité. Il est indispensable de savoir passer d'une description à l'autre.

Exemple. Considérons F et G les sous-espaces vectoriels de $\mathbb{R}^4$:

$$F = \{(x, y, z, w) \in \mathbb{R}^4 : x - y - 2w = 0 \quad \text{et} \quad x + w = 0\},$$

$$G = \{(x, y, z, w) \in \mathbb{R}^4 : x - z + w = 0 \quad \text{et} \quad y + z = 0\}.$$

On a

$$\begin{cases} x - y - 2w = 0 \\ x + w = 0 \end{cases} \iff \begin{cases} y = x - 2w = -3w \\ x = -w \end{cases} \iff (x, y, z, w) = z(0, 0, 1, 0) + w(-1, -3, 0, 1)$$

et

$$\begin{cases} x - z + w = 0 \\ y + z = 0 \end{cases} \iff \begin{cases} x = z - w \\ y = -z \end{cases} \iff (x, y, z, w) = z(1, -1, 1, 0) + w(-1, 0, 0, 1).$$

Puisqu'on a pratiqué la méthode de Gauss, les vecteurs $u_1 = (0, 0, 1, 0)$ et $u_2 = (-1, -3, 0, 1)$ forment une base de F. De même $u_3 = (1, -1, 1, 0)$ et $u_4 = (-1, 0, 0, 1)$ forment une base de G.

Le déterminant de la matrice dont les colonnes sont les coordonnées des vecteurs u_1, u_2, u_3 et u_4 dans la base canonique de $\mathbb{R}^4$ est

$$\begin{vmatrix} 0 & -1 & 1 & -1 \\ 0 & -3 & -1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} -1 & 1 & -1 \\ -3 & -1 & 0 \\ 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 \\ -3 & -1 & 0 \\ 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ -3 & -1 \end{vmatrix} = 3 \neq 0.$$

Donc, (u_1, u_2, u_3, u_4) est une base de $\mathbb{R}^4$. Les sous-espaces vectoriels F et G de $\mathbb{R}^4$ sont donc supplémentaires.

2.2 Applications linéaires

Dans ce paragraphe, E et F sont des espaces vectoriels sur $\mathbb{K}$.

Définition 2.5 *Une application linéaire de E dans F est une application $f : E \to F$ telle que*

1. *pour tous $x, y \in E$, $f(x+y) = f(x) + f(y)$,*
2. *pour tout $x \in E$, pour tout $\alpha \in \mathbb{K}$, $f(\alpha x) = \alpha f(x)$.*

Remarque. Les propriétés 1 et 2 sont équivalentes à la condition suivante :
– pour tous $x, y \in E$ et pour tous $\alpha, \beta \in \mathbb{K}$, $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$.

Exemple. L'application identité de E est une application linéaire de E dans E. On rapelle que cette application est notée id_E et que l'on a par définition $\mathrm{id}_E(x) = x$ pour tout $x \in E$.

Définition 2.6 *On appelle un endomorphisme de E une application linéaire de E dans E.*

Définition 2.7 *Soit $n \geq 2$. Si f est une application de E dans $\mathbb{K}^n$, les composantes de f sont les applications $f_1, f_2, \ldots, f_n$ de E dans $\mathbb{K}$ définies par $f(x) = (f_1(x), \ldots, f_n(x))$ pour tout $x \in E$.*

Il est assez facile de vérifier que $f : E \to \mathbb{K}^n$ est linéaire si et seulement si ses composantes $f_1, f_2, \ldots, f_n$ sont linéaires de E dans $\mathbb{K}$.

Définition 2.8 *Un isomorphisme de E dans F est une application linéaire de E dans F qui est bijective.*

On dit que E et F sont isomorphes s'il existe un isomorphisme de E dans F.

Alors, si E et F sont de dimension finie, E et F sont isomorphes si et seulement si $\dim E = \dim F$.

Rappelons que si G est une partie de E, alors l'image de G par f est la partie $f(G)$ de F définie par

$$f(G) = \{f(x) : x \in G\}.$$

Si G est un sous-espace vectoriel de E alors $f(G)$ est un sous-espace vectoriel de F. En effet, puisque $f(0_E) = 0_F$, $0_F \in f(G)$. Soient $y, y' \in f(G)$. Alors, il existe $x, x' \in G$ tels que $f(x) = y$ et $f(x') = y'$. Il vient, $f(x + x') = f(x) + f(x') = y + y'$, où $x + x' \in G$ car G est un sous-espace vectoriel de E. Donc, par définition, $y + y' \in f(G)$. De même, on montre que si $\alpha \in \mathbb{K}$ alors $\alpha y \in f(G)$. L'ensemble $f(G)$ est donc un sous-espace vectoriel de F.

Définition 2.9 *On appelle image de f le sous-espace vectoriel $f(E)$ de F et on le note* $\mathrm{Im} f$.

Si $(e_1, \ldots, e_n)$ est une base de E, $\mathrm{Im} f$ est engendré par les vecteurs $f(e_1), \ldots, f(e_n)$. Donc, $\dim \mathrm{Im} f \leq n$.

Proposition 2.7 *L'ensemble des vecteurs $x \in E$ tels que $f(x) = 0_F$ est un sous-espace vectoriel de E. Ce sous-espace vectoriel s'appelle le noyau de f et se note* $\mathrm{Ker} f$.

Preuve. On a $f(0_E) = 0_F$, alors $0_E \in \mathrm{Ker}f$. D'autre part, si $x, y \in \mathrm{Ker}f$, on a $f(x) = f(y) = 0_F$. Alors, pour tous $\alpha, \beta \in \mathbb{K}$,

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = 0_F,$$

c'est-à-dire $\alpha x + \beta y \in \mathrm{Ker}f$.

Proposition 2.8 *L'application linéaire f est injective si et seulement si* $\mathrm{Ker}f = \{0_E\}$.

Preuve. Supposons que f est injective. Si $x \in \mathrm{Ker}f$, alors $f(x) = 0_F = f(0_E)$. Par l'injectivité de f, on a $x = 0_E$ et donc $\mathrm{Ker}f = \{0_E\}$. Réciproquement, suppposons $\mathrm{Ker}f = \{0_E\}$. Soient $x, y \in E$ tels que $f(x) = f(y)$. Alors, $f(x - y) = f(x) - f(y) = 0_F$ et donc $x - y \in \mathrm{Ker}f$. Donc $x - y = 0_E$, c'est-à-dire $x = y$. On conclut que f est injective.

Théorème 2.3 (Théorème de la dimension) *Soit E espace vectoriel de dimension finie et f un endomorphisme de E. Alors,*

$$\dim E = \dim \mathrm{Ker}f + \dim \mathrm{Im}f.$$

Corollaire 2.1 *Soit f un endomorphisme de E espace vectoriel de dimension finie. Alors f est un isomorphisme (application linéaire bijective) si et seulement si elle est injective ou elle est surjective.*

2.3 Matrice d'une application linéaire

Dans ce paragraphe, E et E' sont des espaces vectoriels de dimension finie non nulle. Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E et $\mathcal{B}' = (e'_1, \ldots, e'_m)$ une base de E', $n, m \in \mathbb{N}^*$.

Définition 2.10 *Soit f une application linéaire de E dans E'. La matrice de f dans les bases $\mathcal{B}$ et $\mathcal{B}'$, notée $M(f; \mathcal{B}, \mathcal{B}')$, est la matrice appartenant à $M_{m,n}(\mathbb{K})$ dont les coefficients de la j-ième colonne sont les coordonnées du vecteur $f(e_j)$ dans la base $\mathcal{B}'$.*

Si $E = E'$ et $e'_i = e_i$, pour tout $1 \leq i \leq n$, alors la matrice de f dans les bases $\mathcal{B} = \mathcal{B}' = (e_1, \ldots, e_n)$ s'appelle plus simplement la matrice de f dans la base $\mathcal{B}$ et on note $M(f; \mathcal{B}, \mathcal{B})$.

Exemples.

1. La matrice de id_E, où E est un espace vectoriel de dimension n, dans n'importe quelle base de E est I_n.
2. Soit $f : \mathbb{R}^2 \to \mathbb{R}^3$ l'application linéaire définie par $f(x, y) = (x + y, -x + 2y, y)$ pour tous $x, y \in \mathbb{R}$. Si $\mathcal{B}$ et $\mathcal{B}'$ denotent, respectivement, les bases canoniques de $\mathbb{R}^2$ et de $\mathbb{R}^3$, on a
$$M(f; \mathcal{B}, \mathcal{B}') = \begin{pmatrix} 1 & 1 \\ -1 & 2 \\ 0 & 1 \end{pmatrix}.$$

La matrice d'une application linéaire permet de calculer les coordonnées de $f(x)$ lorsqu'on connait celles de x.

Proposition 2.9 *Soit f une application linéaire de E dans E' et soit $A = M(f;\mathcal{B},\mathcal{B}')$ dans les bases $\mathcal{B} = (e_1,\ldots,e_n)$ de E et $\mathcal{B}' = (e'_1,\ldots,e'_m)$ de E'. Si x est un vecteur de E et si l'on note*

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad et \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},$$

où $x_1,\ldots,x_n$ sont les coordonnées de x dans la base $\mathcal{B}$ et où $y_1,\ldots,y_m$ sont les coordonnées de $f(x)$ dans la base $\mathcal{B}'$, alors on a $Y = AX$.

Preuve. L'application f est linéaire, alors $f(x) = x_1 f(e_1) + \ldots + x_n f(e_n)$. Si A est la matrice (a_{ij}), alors par la définition de A, on a

$$\begin{cases} f(e_1) = a_{11}e'_1 + \ldots + a_{m1}e'_m \\ \qquad\qquad \vdots \\ f(e_n) = a_{1n}e'_1 + \ldots + a_{mn}e'_m \end{cases}$$

et donc $f(x) = (a_{11}x_1 + \ldots + a_{1n}x_n)e'_1 + \ldots + (a_{m1}x_1 + \ldots + a_{mn}x_n)e'_m$. Par unicité des coordonnées de $f(x)$ dans la base $\mathcal{B}'$, on a donc

$$\begin{cases} y_1 = a_{11}x_1 + \ldots + a_{1n}x_n \\ \qquad\qquad \vdots \\ y_m = a_{m1}x_1 + \ldots + a_{mn}x_n \end{cases}$$

ce qui se traduit matriciellement par $Y = AX$.

Exemple. Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'application linéaire dont la matrice dans $\mathcal{B}$ la base canonique de $\mathbb{R}^3$ est

$$A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 1 \\ 2 & 3 & 1 \end{pmatrix}.$$

Alors, pour tous nombres réels x, y, z, on a

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + z \\ -x + z \\ 2x + 3y + z \end{pmatrix}.$$

On a donc $f(x, y, z) = (x + 2y + z, -x + z, 2x + 3y + z)$ pour tous $x, y, z \in \mathbb{R}$.

Dans la proposition suivante, nous allons voir que le produit matriciel correspond à la composée des applications linéaires.

Proposition 2.10 *Soient $\mathcal{B}$ une base de E, $\mathcal{B}'$ une base de E' et $\mathcal{B}''$ une base de E''. Soit f une application linéaire de E dans E' et g une application linéaire de E' dans E''. Si $A = M(f;\mathcal{B},\mathcal{B}')$ et $B = M(g;\mathcal{B}',\mathcal{B}'')$, alors $M(g \circ f;\mathcal{B},\mathcal{B}'') = BA$.*

Proposition 2.11 *Supposons que les espaces E et E' ont la même dimension n. Soit f une application linéaire de E dans E' et soit $A = M(f;\mathcal{B},\mathcal{B}')$ où $\mathcal{B} = (e_1,\ldots,e_n)$ est une base de E et $\mathcal{B}' = (e'_1,\ldots,e'_n)$ est une base de E'. L'application f est un isomorphisme si et seulement si la matrice A est inversible. De plus, si f est un isomorphisme $M(f^{-1};\mathcal{B}',\mathcal{B}) = A^{-1}$.*

Preuve. Nous savons que f est un isomorphisme si et seulement si $(f(e_1), \dots, f(e_n))$ est une base de E'. La matrice A est celle des coordonnées de $(f(e_1), \dots, f(e_n))$ dans la base $\mathcal{B}'$. D'après la proposition 2.5, $(f(e_1), \dots, f(e_n))$ est une base de E' si et seulement si la matrice A est inversible.

Si f est un isomorphisme, $f^{-1} \circ f = \mathrm{id}_E$. Alors,

$$I_n = M(\mathrm{id}_E; \mathcal{B}, \mathcal{B}) = M(f^{-1} \circ f; \mathcal{B}, \mathcal{B}) = M(f^{-1}; \mathcal{B}', \mathcal{B}) M(f; \mathcal{B}, \mathcal{B}') = M(f^{-1}; \mathcal{B}', \mathcal{B}) A.$$

Donc, $M(f^{-1}; \mathcal{B}', \mathcal{B}) = A^{-1}$.

2.4 Changement de base

Supposons que $\mathcal{B} = (u_1, \dots, u_n)$ et $\mathcal{B}' = (u'_1, \dots, u'_n)$ sont des bases de E.

Définition 2.11 *La matrice P de $M_n(\mathbb{K})$ dont les coefficients de la j-ième colonne sont les coordonnées du vecteur u'_j dans la base $\mathcal{B}$ s'appelle la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$.*

Remarque. On a $P = M(\mathrm{Id}_E; \mathcal{B}', \mathcal{B})$.

D'après la proposition précédente, une matrice de passage est toujours inversible car id_E est un isomorphisme de E sur E.

Soit $x \in E$. Notons $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ les coordonnées de x dans la base $\mathcal{B}$ et $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ les coordonnées de x dans la base $\mathcal{B}'$. D'après la proposition 2.9, $X = PX'$, c'est-à-dire $X' = P^{-1}X$.

Étudions ensuite l'effet d'un changement de base sur la matrice d'un endomorphisme de E.

Théorème 2.4 (Formule de changement de base) *Supposons que $\mathcal{B} = (u_1, \dots, u_n)$ et $\mathcal{B}' = (u'_1, \dots, u'_n)$ sont des bases de E. Soient f un endomorphisme de E, $A = M(f; \mathcal{B}, \mathcal{B})$ et $A' = M(f; \mathcal{B}', \mathcal{B}')$. Si P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, alors on a $A' = P^{-1}AP$.*

Preuve. D'après la proposition 2.10, nous avons

$$M(f \circ \mathrm{id}_E; \mathcal{B}', \mathcal{B}) = M(f; \mathcal{B}, \mathcal{B}) M(\mathrm{id}_E; \mathcal{B}', \mathcal{B}) = AP$$

et

$$M(\mathrm{id}_E \circ f; \mathcal{B}', \mathcal{B}) = M(\mathrm{id}_E; \mathcal{B}', \mathcal{B}) M(f; \mathcal{B}', \mathcal{B}') = PA'.$$

Puisqu'on a $f \circ \mathrm{id}_E = \mathrm{id}_E \circ f = f$, il vient $AP = PA'$. On en déduit $A' = P^{-1}AP$ et ce qui est équivalent, $A = PA'P^{-1}$.

Exemple. Soit $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique de $\mathbb{R}^3$ et soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'application linéaire définie par $f(x, y, z) = (-x + y + z, -2x + y + z, -6x + 2y + 4z)$ pour tous $x, y, z, \in \mathbb{R}$. La matrice de f dans la base $\mathcal{B}$ est

$$A = M(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} -1 & 1 & 1 \\ -2 & 1 & 1 \\ -6 & 2 & 4 \end{pmatrix}.$$

Considérons les vecteurs de $\mathbb{R}^3$: $v_1 = (1, 0, 2)$, $v_2 = (1, 1, 2)$ et $v_3 = (2, 1, 5)$. Soit P la matrice de $M_3(\mathbb{R})$ dont les colonnes sont les coordonnées des vecteurs v_1, v_2 et v_3 dans la base $\mathcal{B}$. Il vient

$$P = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 2 & 2 & 5 \end{pmatrix}.$$

Puisque $\det P = 1$, P est inversible et par conséquent $\mathcal{B}' = (v_1, v_2, v_3)$ est une base de $\mathbb{R}^3$.

On a $f(v_1) = (1, 0, 2) = v_1$, $f(v_2) = (2, 1, 4)$ et $f(v_3) = (4, 2, 10) = 2v_3$. Les coordonnées de $f(v_2)$ dans la base (v_1, v_2, v_3) sont les nombres réels a, b, c tels que $f(v_2) = av_1 + bv_2 + cv_3$, c'est-à-dire les solutions du système d'équations linéaires suivant :

$$\left\{ \begin{aligned} a + b + 2c &= 2 \\ b + c &= 1 \\ 2a + 2b + 5c &= 4 \end{aligned} \right. \iff \left\{ \begin{aligned} a + b + 2c &= 2 \\ b + c &= 1 \\ c &= 0 \end{aligned} \right. \iff \left\{ \begin{aligned} a &= 1 \\ b &= 1 \\ c &= 0. \end{aligned} \right.$$

Il s'ensuit $f(v_2) = v_1 + v_2$. La matrice de f dans la base $\mathcal{B}' = (v_1, v_2, v_3)$ est donc

$$B = M(f; \mathcal{B}', \mathcal{B}') = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

D'après la formule de changement de base, nous savons que l'on a l'égalité matricielle $B = P^{-1}AP$. Notons que $P^{-1} = \begin{pmatrix} 3 & -1 & -1 \\ 2 & 1 & -1 \\ -2 & 0 & 1 \end{pmatrix}$.

2.5 Matrices semblables

Nous notons $\mathrm{GL}_n(\mathbb{K})$ l'ensemble des matrices inversibles de $M_n(\mathbb{K})$.

Définition 2.12 *Soit $A, B \in M_n(\mathbb{K})$. On dit que la matrice A est semblable à B, ou que les matrices A et B sont semblables, s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.*

Nous remarquons que $B = P^{-1}AP$ si et seulement si $A = PBP^{-1}$.

Exemple. Les matrices $A = \begin{pmatrix} 5 & 4 \\ 0 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 0 \\ 2 & 5 \end{pmatrix}$ et $C = \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}$ de $M_2(\mathbb{R})$ sont semblables.

Proposition 2.12 *Soient A et B deux matrices semblables de $M_n(\mathbb{K})$. Alors les matrices A et B ont le même determinant et la même trace.*

Preuve. S'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$, alors

$$\det(B) = \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P) = (\det(P))^{-1}\det(A)\det(P) = \det(A),$$

et

$$\mathrm{tr(B) = tr(P^{-1}AP) = tr(APP^{-1}) = tr(A)}.$$

Mais attention, **la réciproque est fausse**. Par exemple, les matrices $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ont le même determinant et la même trace mais elles ne sont pas semblables, car I_2 est la seule matrice semblable à I_2.

Proposition 2.13 *Supposons E de dimension n. Soient A et B des matrices de $M_n(\mathbb{K})$. Les matrice A et B sont semblables si et seulement si A et B sont les matrices d'un même endomorphisme sur E, dans des bases éventuellement différentes.*

3 Diagonalisation

Dans ce chapitre, E est un espace vectoriel de dimension finie non nulle et f est un endomorphisme de E. Nous allons donner une méthode pour trouver une base de E dans laquelle la matrice de f est "la plus simple possible", et plus précisemment, est une matrice diagonale.

3.1 Valeur propre. Vecteur propre

Définition 3.1 *Soit $\lambda \in \mathbb{K}$. On dit que λ est valeur propre de f s'il existe un vecteur non nul $x \in E$ tel que $f(x) = \lambda x$.*

Exemples.

1. Le scalaire $\lambda = 1$ est l'unique valeur propre de l'application identité id_E.
2. Soit $\lambda \in \mathbb{R}$. On définit $f : E \to E$ par $f(x) = \lambda x$. Cet endomorphisme s'appelle l'homothétie de E par rapport à λ. On a $f = \lambda \mathrm{id}_E$. Alors, λ est l'unique valeur propre de f.
3. Si $\mathrm{Ker} f \neq \{0_E\}$ alors 0 est valeur propre de f. En effet, il existe $x \in E \setminus \{0_E\}$ tel que $f(x) = 0 = 0 \times x$.

Remarque. Le scalaire λ est valeur propre de f si et seulement si $f - \lambda \mathrm{id}_E$ n'est pas bijectif.

D'après le théorème de la dimension,

$$n = \dim \mathrm{Ker}(f - \lambda \mathrm{id}_E) + \dim \mathrm{Im}(f - \lambda \mathrm{id}_E).$$

Dans le chapitre précédent, nous avons vu que d'après cette égalité, $f - \lambda \mathrm{id}_E$ est bijectif si et seulement s'il est injectif, donc si et seulement si $\mathrm{Ker}(f - \lambda \mathrm{id}_E) = \{0_E\}$. Mais $x \in \mathrm{Ker}(f - \lambda \mathrm{id}_E)$ si et seulement si $f(x) = \lambda x$. Ainsi, $f - \lambda \mathrm{id}_E$ n'est pas bijectif si et seulement s'il existe un vecteur x non nul tel que $f(x) = \lambda x$, c'est-à-dire, λ est valeur propre de f.

En particulier, f et bijectif si et seulement si 0 n'est pas valeur propre de f.

Définition 3.2 *Soit $x \in E$. On dit que x est vecteur propre de f pour la valeur propre λ si $x \neq 0_E$ et si $f(x) = \lambda x$.*

Remarque. Un vecteur propre x est associé à une unique valeur propre. En effet, si $f(x) = \lambda x$ et $f(x) = \mu x$, alors $(\lambda - \mu)x = 0$. Or le vecteur x est non nul, donc $\lambda - \mu = 0$, c'est-à-dire $\lambda = \mu$.

Exemple. Soit f l'endomorphisme de $\mathbb{R}^2$ dont la matrice dans la base canonique est

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

Pour tout vecteur $x = (a, b) \in \mathbb{R}^2$ et tout $\lambda \in \mathbb{R}$, on a

$$f(x) = \lambda x \iff (A - \lambda I_2) \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \left\{ \begin{array}{l} (1-\lambda)a + b = 0 \\ (2-\lambda)b = 0 \end{array} \right. .$$

Alors,

$$f(x) = 2x \iff b - a = 0 \iff b = a \iff x = (a, a) = a(1, 1).$$

Donc, 2 est valeur propre de f et les vecteurs propres de f pour la valeur propre 2 sont les vecteurs non nuls colinéaires à $(1, 1)$. D'autre part,

$$f(x) = x \iff b = 0 \iff x = (a, 0) = a(1, 0).$$

Donc, 1 est valeur propre de f et les vecteurs propres de f pour la valeur propre 1 sont les vecteurs non nuls colinéaires à $(1, 0)$. Enfin, si λ est différent de 1 et 2, on a

$$\left\{ \begin{array}{l} (1-\lambda)a + b = 0 \\ (2-\lambda)b = 0 \end{array} \right. \iff \left\{ \begin{array}{l} (1-\lambda)a = 0 \\ b = 0 \end{array} \right. \iff \left\{ \begin{array}{l} a = 0 \\ b = 0 \end{array} \right.$$

On en déduit que si $\lambda \neq 1, 2$, l'égalité $f(x) = \lambda x$ entraîne $x = 0$, donc λ n'est pas valeur propre de f. Les seules valeurs propres de f sont par conséquent 1 et 2.

Théorème 3.1 *Soient $\lambda_1, \ldots, \lambda_k$ les valeurs propres distinctes de f. Si x_i est vecteur propre de f pour la valeur propre λ_i, alors les vecteurs $x_1, \ldots, x_k$ sont linéairement indépendants.*

Preuve. Puisqu'un vecteur propre n'est pas nul, le résultat est vrai si $k = 1$.

On considère le cas $k = 2$. Supposons par contradiction que x_1, x_2 ne sont pas linéairement indépendants. Alors, il existe $\alpha_1, \alpha_2 \in \mathbb{K}$ tels que $\alpha_1 x_1 + \alpha_2 x_2 = 0$ avec $\alpha_1 \neq 0$ ou $\alpha_2 \neq 0$. Supposons $\alpha_1 \neq 0$. Alors, $x_1 = \frac{\alpha_2}{\alpha_1} x_2$ et

$$\lambda_1 x_1 = f(x_1) = \frac{\alpha_2}{\alpha_1} f(x_2) = \frac{\alpha_2}{\alpha_1} \lambda_2 x_2 = \lambda_2 x_1,$$

donc $(\lambda_1 - \lambda_2)x_1 = 0$. Puisque x_1 est un vecteur propre, $x_1 \neq 0$ et donc $\lambda_1 = \lambda_2$ ce qui est une contradiction. On conclut que x_1, x_2 sont linéairement indépendants.

3.2 Polynôme caractéristique

Nous allons introduire un polynôme dont les racines dans $\mathbb{K}$ sont les valeurs propres de f. Ce polynôme sera le déterminant d'une matrice et donc nous pourrons l'expliciter en utilisant les techniques habituelles de calcul d'un déterminant.

Soit $A \in M_n(\mathbb{K})$ et soit I_n la matrice identité dans $M_n(\mathbb{K})$. Posons $A = (a_{i,j})_{1 \leq i,j \leq n}$, $B = A - \lambda I_n$ et $B = (b_{i,j})_{1 \leq i,j \leq n}$ avec $b_{i,i} = a_{i,i} - \lambda$ et $b_{i,j} = a_{i,j}$ si $i \neq j$. Par exemple, si $n = 2$,

$$B = \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix}$$

et

$$\det B = (a_{11} - \lambda)(a_{22} - \lambda) - a_{21}a_{12} = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{21}a_{12}.$$

Observons que $\det B$ est un polynôme en λ à valeurs dans $\mathbb{K}$.

Proposition 3.1 *Soit $A \in M_n(\mathbb{K})$. Le déterminant de la matrice $A - \lambda I_n$ est un polynôme à coefficients dans $\mathbb{K}$ de degré n. De plus, on a*

$$\det(A - \lambda I_n) = (-1)^n \lambda^n + (-1)^{n-1}(\mathrm{tr} A)\lambda^{n-1} + \ldots + \det A.$$

Définition 3.3 *Soit $A \in M_n(\mathbb{K})$. Le polynôme $\det(A - \lambda I_n)$ de $\mathbb{K}[\lambda]$ s'appelle le polynôme caractéristique de la matrice A. Nous le noterons $Q_A(\lambda)$ ou $Q(\lambda)$.*

Exemple. Soit $A \in M_4(\mathbb{R})$ définie par $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \\ -1 & 2 & 0 & 3 \\ 2 & 3 & 0 & 0 \end{pmatrix}$. Son polynôme caractéristique est :

$$\begin{aligned}
Q_A(\lambda) &= \begin{vmatrix} 1-\lambda & 0 & 1 & 0 \\ 0 & 1-\lambda & 0 & 2 \\ -1 & 2 & -\lambda & 3 \\ 2 & 3 & 0 & -\lambda \end{vmatrix} \underset{L3+\lambda L1 \to L3}{=} \begin{vmatrix} 1-\lambda & 0 & 1 & 0 \\ 0 & 1-\lambda & 0 & 2 \\ -\lambda^2+\lambda-1 & 2 & 0 & 3 \\ 2 & 3 & 0 & -\lambda \end{vmatrix} \\
&= \begin{vmatrix} 0 & 1-\lambda & 2 \\ -\lambda^2+\lambda-1 & 2 & 3 \\ 2 & 3 & -\lambda \end{vmatrix} = (\lambda^2-\lambda+1) \begin{vmatrix} 1-\lambda & 2 \\ 3 & -\lambda \end{vmatrix} + 2 \begin{vmatrix} 1-\lambda & 2 \\ 2 & 3 \end{vmatrix} \\
&= (\lambda^2-\lambda+1)(\lambda^2-\lambda-6) - 2(3\lambda+1) = \lambda^4 - 2\lambda^3 - 4\lambda^2 - \lambda - 8.
\end{aligned}$$

Proposition 3.2 *Soient A et B deux matrices appartenant à $M_n(\mathbb{K})$. S'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$ (c'est-à-dire, A et B sont des matrices semblables), alors $Q_A(\lambda) = Q_B(\lambda)$.*

Preuve. Nous avons $B - \lambda I_n = P^{-1}AP - \lambda P^{-1}I_nP = P^{-1}(A - \lambda I_n)P$. Puisqu'on a $\det(P^{-1}MP) = \det M$ si $M \in M_n(\mathbb{K})$, alors

$$Q_A(\lambda) = \det(A - \lambda I_n) = \det(B - \lambda I_n) = Q_B(\lambda).$$

Si A et B sont les matrices d'un même endomorphisme de E dans des bases éventuellements différentes, il existe une matrice inversible P telle que $B = P^{-1}AP$, d'après la formule de changement de base. La proposition précédente permet donc de définir le polynôme caractéristique d'un endomorphisme.

Définition 3.4 *On appelle polynôme caractéristique de f, le polynôme caractéristique de la matrice de f dans une base de E. Le polynôme caractéristique de f sera noté Q_f.*

Par définition du déterminant d'un endomorphisme, on a donc $Q_f(\lambda) = \det(f - \lambda \mathrm{id}_E)$, pour tout $\lambda \in \mathbb{K}$.

Proposition 3.3 *Soit $\lambda \in \mathbb{K}$. Le scalaire λ est valeur propre de f si et seulement si λ est racine du polynôme caractéristique de f.*

Preuve. On sait que λ est valeur propre de f si et seulement si l'endomorphisme $f - \lambda\mathrm{id}_E$ n'est pas bijectif, donc si et seulement si $Q_f(\lambda) = \det(f - \lambda\mathrm{id}_E) = 0$. Autrement dit, λ est valeur propre de f si et seulement si λ est racine du polynôme Q_f.

Si E est de dimension n, f a au plus n valeurs propres.

Exemple 1. Soit f l'endomorphisme de $\mathbb{R}^2$ dont la matrice dans la base canonique est

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

Alors,

$$Q(\lambda) = \begin{vmatrix} 1-\lambda & -1 \\ 2 & 4-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda) + 2 = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3).$$

Les valeurs propres de f sont $\lambda_1 = 2$ et $\lambda_2 = 3$.

Exemple 2. Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique est

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

Alors,

$$\begin{aligned} Q(\lambda) &= \begin{vmatrix} 1-\lambda & 0 & 1 \\ -1 & 2-\lambda & 1 \\ 1 & -1 & 1-\lambda \end{vmatrix} \underset{C_1+C_2\to C_1}{=} \begin{vmatrix} 1-\lambda & 0 & 1 \\ 1-\lambda & 2-\lambda & 1 \\ 0 & -1 & 1-\lambda \end{vmatrix} \\ &\underset{L_2-L_1\to L_2}{=} \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 0 & -1 & 1-\lambda \end{vmatrix} = (1-\lambda) \begin{vmatrix} 2-\lambda & 0 \\ -1 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)(2-\lambda)(1-\lambda) = (1-\lambda)^2(2-\lambda). \end{aligned}$$

Donc, f admet deux valeurs propres distinctes : $\lambda_1 = 1$ de multiplicité 2 et $\lambda_2 = 2$ de multiplicité 1.

Proposition 3.4 *Soit A la matrice de f dans une base de E. Si A est triangulaire supérieure ou inférieure, alors les valeurs propres de f sont les coefficients diagonaux de A.*

Exemple. Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique est

$$A = \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 5 \\ 0 & 0 & 1 \end{pmatrix}.$$

Les valeurs propres de f sont : 1, 2 et 3. Soient u_1 vecteur propre de f pour la valeur propre 1, u_2 vecteur propre de f pour la valeur propre 2 et u_3 vecteur propre de f pour la valeur propre 3. D'après le théorème 3.1, les vecteurs u_1, u_2, u_3 sont linéairement indépendents, donc ils forment une base de $\mathbb{R}^3$. La matrice de f dans la base $\mathcal{B}' = (u_1, u_2, u_3)$ est la matrice diagonale

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Si l'on note P la matrice de passage de la base canonique $\mathcal{B}$ de $\mathbb{R}^3$ à la base $\mathcal{B}'$, d'après la formule de changement de base, on a $A = PDP^{-1}$.

Définition 3.5 *Soit A une matrice de $M_n(\mathbb{K})$. On dit que $\lambda \in \mathbb{K}$ est une valeur propre de A si λ est une racine de $Q_A(\lambda)$. Soit λ une valeur propre de A. On dit que le vecteur colonne V est un vecteur propre de A pour la valeur propre λ si $V \neq 0$ et $AV = \lambda V$.*

3.3 Sous-espaces propres

Si λ est une valeur propre de f, alors les vecteurs propres de f pour λ sont les vecteurs non nuls appartenant à $\mathrm{Ker}(f - \lambda \mathrm{id}_E)$.

Définition 3.6 *Soit λ une valeur propre de f. Le sous-espace vectoriel $\mathrm{Ker}(f - \lambda \mathrm{id}_E)$ de E s'appelle le sous-espace propre de f pour la valeur propre λ et se note E_λ.*

Exemple. Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique est

$$A = \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 5 \\ 0 & 0 & 1 \end{pmatrix}.$$

Les valeurs propres de f sont : 1, 2 et 3. Les sous-espaces propres de f sont alors

$$E_1 = \mathrm{Ker}(A - I_n) = \{(x, y, z) \in \mathbb{R}^3 : 2x = z, \ \ y = -5z\} =< (1, -10, 2) >,$$

$$E_2 = \mathrm{Ker}(A - 2I_n) = \{(x, y, z) \in \mathbb{R}^3 : x + y = 0, \ \ z = 0\} =< (1, -1, 0) >,$$

$$E_3 = \mathrm{Ker}(A - 3I_n) = \{(x, y, z) \in \mathbb{R}^3 : y = 0, \ \ z = 0\} =< (1, 0, 0) > .$$

Soient $u_1 = (1, -10, 2)$, $u_2 = (1, -1, 0)$ et $u_3 = (1, 0, 0)$. Alors $\mathcal{B}' = (u_1, u_2, u_3)$ est une base de $\mathbb{R}^3$ de vecteurs propres de f.

Théorème 3.2 *Soient $\lambda_1, \ldots, \lambda_k$ les valeurs propres distinctes de f. Alors les sous-espaces propres $E_{\lambda_1}, \ldots, E_{\lambda_k}$ sont en somme directe.*

Soit λ_0 une valeur propre de f, c'est-à-dire une racine du polynôme caractéristique $Q = Q_f$. Notons $m(\lambda_0)$ la multiplicité de λ_0 dans le polynôme Q, c'est-à-dire le plus grand entier r tel que $(\lambda - \lambda_0)^r$ divise $Q(\lambda)$. Puisque λ_0 est racine de Q, on a $m(\lambda_0) \geq 1$.

Théorème 3.3 *Soit λ une valeur propre de f. Alors, on a $1 \leq \dim E_\lambda \leq m(\lambda)$.*

3.4 Endomorphisme et matrice diagonalisables

Soit $(e_1, \ldots, e_n)$ base de E. Pour que la matrice de f dans la base $(e_1, \ldots, e_n)$ soit la matrice diagonale $D = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix}$ il faut et il suffit que l'on ait $f(e_i) = \lambda e_i$ pour tout $1 \leq i \leq n$. On a donc la proposition suivante.

Proposition 3.5 *La matrice de f dans la base $\mathcal{B} = (e_1, \ldots, e_n)$ est diagonale si et seulement si les vecteurs $e_1, \ldots, e_n$ sont des vecteurs propres de f.*

Définition 3.7 *On dit que f est diagonalisable s'il existe une base de E formée de vecteurs propres de f.*

D'après la proposition, l'endomorphisme f est diagonalisable si et seulement il existe une base de E dans laquelle la matrice de f est diagonale.

Exemple.

1. Une homothétie de E est un endomorphisme diagonalisable. En effet, dans n'importe quelle base de E, la matrice d'une homothétie est de la forme λI_n.
2. Soit $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'endomorphisme dont la matrice dans la base canonique de $\mathbb{R}^3$ est
$$A = \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 5 \\ 0 & 0 & 1 \end{pmatrix}.$$
Nous avons vu que $u_1 = (1, -10, 2)$, $u_2 = (1, -1, 0)$ et $u_3 = (1, 0, 0)$ sont des vecteurs propres de f, pour les valeurs propres 1, 2 et 3, respectivement. Puisque (u_1, u_2, u_3) est une base de $\mathbb{R}^3$ de vecteurs propres de f, f est diagonalisable.
3. Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'endomorphisme dont la matrice dans la base canonique $\mathcal{B}$ de $\mathbb{R}^2$ est
$$A = \begin{pmatrix} -1 & 2 \\ -4 & 5 \end{pmatrix}.$$
Puisque $Q(\lambda) = (-1-\lambda)(5-\lambda) + 8 = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3)$, les valeurs propres de f sont 1 et 3. On vérifie facilement que $E_1 =< (1, 1) >$ et $E_3 =< (1, 2) >$. Alors, $\mathcal{B}' = ((1, 1), (1, 2))$ est une base de $\mathbb{R}^2$ de vecteurs propres de f. On conclut que f est diagonalisable. De plus, $A = PDP^{-1}$ où D est la matrice diagonale de coefficients diagonaux 1 et 3 et P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.

On rappelle que $\mathrm{GL}_n(\mathbb{K})$ denote l'espace des matrices de $M_n(\mathbb{K})$ inversibles.

Définition 3.8 *Soit $A \in M_n(\mathbb{K})$. On dit que la matrice A est diagonalisabe sur $\mathbb{K}$ s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}AP$ est une matrice diagonale.*

Les définitions de matrice diagonalisable et d'endomorphisme diagonalisable sont liées.

Proposition 3.6 *L'endomorphisme f est diagonalisable si et seulement si la matrice de f dans une base de E est diagonalisable sur $\mathbb{K}$.*

Preuve. Soient $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E et A la matrice de f dans la base $\mathcal{B}$. Supposons que f est diagonalisable. Il existe donc une base $\mathcal{B}'$ de E dans laquelle la matrice de f est une matrice diagonale D. Notons P la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$. Par la formule du changement de base, $D = P^{-1}AP$ et donc A est diagonalisable sur $\mathbb{K}$. Réciproquement, supposons qu'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ et $D \in M_n(\mathbb{K})$ telles que $D = P^{-1}AP$. Considérons les vecteurs $u_1, \ldots, u_n$ de E dont les coordonnées dans la base $\mathcal{B}$ sont les colonnes de P. Puisque la matrice P est inversible, $\mathcal{B}' = (u_1, \ldots, u_n)$ est une base de E et P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. La matrice de f dans la base $\mathcal{B}'$ est la matrice diagonale $D = P^{-1}AP$, donc l'endomorphisme f est diagonalisable.

Proposition 3.7 *Supposons E de dimension n. Si f a n vecteurs propres distinctes, alors l'endomorphisme f est diagonalisable.*

Preuve. Supposons que f a n valeurs propres distinctes $\lambda_1, \ldots, \lambda_n$. Pour tout entier $1 \leq i \leq n$, soit u_i un vecteur propre de f pour la valeur propre λ_i. D'après le théorème 3.1, les vecteurs $u_1, \dots, u_n$ sont linéairement indépendents et donc ils forment une base de E. D'où $(u_1, \dots, u_n)$ est une base de E formée de vecteurs propres de f.

Notons n la dimension de E. Le polynôme caractéristique de f est de degré n, donc a au plus n racines. D'après la proposition, s'il a exactement n racines dans $\mathbb{K}$, alors f est diagonalisable sur $\mathbb{K}$. Mais cette condition pour que l'endomorphisme soit diagonalisable n'est pas nécessaire. Par exemple, l'application identité est diagonalisable mais n'a qu'une seule valeur propre : 1.

Proposition 3.8 *Supposons que l'endomorphisme possède au moins une valeur propre et notons $\lambda_1, \dots, \lambda_r$ les valeurs propres distinctes de f. Alors, l'endomorphisme f est diagonalisable si et seulement si $E = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_r}$.*

Définition 3.9 *Soit Q un polynôme non constant de $\mathbb{K}[X]$. On dit que Q a toutes ses racines dans $\mathbb{K}$ si Q est le produit dans $\mathbb{K}[X]$ de polynômes de degré 1.*

Exemples.

1. Un polynôme $Q(X) = aX^2 + bX + c$ de degré deux appartenant à $\mathbb{R}[X]$ a toutes ses racines dans $\mathbb{R}$ si et seulement si son discriminant $b^2 - 4ac$ est positif ou nul.
2. Tout polynôme non constant dans $\mathbb{C}[X]$ a une racine complexe (théorème d'Alembert). Il est assez simple d'en déduire que tout polynôme non constant dans $\mathbb{C}[X]$ a toutes ses racines dans $\mathbb{C}$.

Théorème 3.4 *L'endomorphisme f est diagonalisable si et seulement si le polynôme $Q(\lambda)$ a toutes ses racines dans $\mathbb{K}$ et si pour chaque valeur propre λ, on a $\dim E_\lambda = m(\lambda)$.*

Remarque 1. Pour toute valeur propre λ, on a $1 \leq \dim E_\lambda \leq m(\lambda)$. Donc, la deuxième condition du théorème est automatiquement vérifiée lorsque λ est racine simple de $Q(\lambda)$, c'est-à-dire lorsque $m(\lambda) = 1$. Pour étudier si f est diagonalisable, il suffit donc de calculer la dimension des sous-espaces propres pour les valeurs propres de multiplicité supérieure ou égale à 2.

Remarque 2. Soit $A \in M_n(\mathbb{K})$. Si λ est valeur propre de A de multiplicité n, alors A est diagonalisable si et seulement si $A = \lambda I_n$.

Exemple 1. Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique est

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

Nous avons montré que f a deux valeurs propres : 1 de multiplicité 2 et 2 de multiplicité 1 ($Q(\lambda) = (1-\lambda)^2(2-\lambda)$). Ensuite, on détermine le sous-espace propre E_1. Soit $(x, y, z) \in \mathbb{R}^3$. On a

$$(A - I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\iff \begin{cases} z=0 \\ -x+y+z=0 \\ x-y=0 \end{cases} \iff \begin{cases} x=y \\ z=0 \end{cases}$$

Alors, $E_1 =< (1,-1,0) >$ et $\dim E_1 = 1 < 2 = m(1)$. Donc, f n'est pas diagonalisable.

Exemple 2. Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique est

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

Alors, le polynôme caractéristique de f est

$$Q(\lambda) = \begin{vmatrix} 1-\lambda & 0 & 1 \\ -1 & 2-\lambda & 1 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda) \begin{vmatrix} 1-\lambda & 0 \\ -1 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda)^2.$$

Donc, f a deux valeurs propres : $\lambda_1 = 1$ de multiplicité 1 et $\lambda_2 = 2$ de multiplicité 2.

Pour $(x,y,z) \in \mathbb{R}^3$, on a

$$(A-2I_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff x = z.$$

Alors, $E_2 =< (0,1,0), (1,0,1), >$ et $\dim E_2 = 2 = m(2)$. On peut déjà conclure que f est diagonalisable. Pour $(x,y,z) \in \mathbb{R}^3$, on a

$$(A-I_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\iff \begin{cases} z=0 \\ -x+y+z=0 \\ z=0 \end{cases} \iff \begin{cases} x=y \\ z=0 \end{cases}$$

Alors, $E_1 =< (1,1,0) >$ et $\dim E_1 = 1 = m(1)$. De plus, $\mathcal{B}' = ((1,1,0),(0,1,0),(1,0,1))$ est une base de $\mathbb{R}^3$ de vecteurs propres de f et $M(f;\mathcal{B}',\mathcal{B}') = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Calcul des puissances d'une matrice diagonalisable.

Soit $D = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$ une matrice diagonale. Il est facile de montrer par récurrence sur $k \in \mathbb{N}$ que, pour tout $k \in \mathbb{N}$,

$$D^k = \begin{pmatrix} \lambda_1^k & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n^k \end{pmatrix}.$$

Soit $A \in M_n(\mathbb{K})$ une matrice diagonalisable. Alors, il existe P matrice inversible et D matrice diagonale telles que $A = PDP^{-1}$. On a $A^0 = D^0 = I_n$ et donc $A^0 = PD^0P^{-1}$. Pour tout $k \in \mathbb{N}$, on a

$$A^k = PD^kP^{-1},$$

ce qui se démontre par récurrence.

3.5 Théorème de Cayley-Hamilton

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$. Le polynôme caractéristique de A est

$$Q(\lambda) = \lambda^2 - (a+d)\lambda + ad - bc.$$

Alors, on a $Q(A) = A^2 - (a+d)A + (ad-bc)I_2$, donc il vient

$$Q(A) = \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} - \begin{pmatrix} a^2+ad & ab+bd \\ ac+cd & ad+d^2 \end{pmatrix} + \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

On obtient ainsi $Q(A) = 0$. Ce résultat est général.

Théorème 3.5 (Cayley-Hamilton) *Si $Q(\lambda)$ est le polynôme caractéristique d'une matrice $A \in M_n(\mathbb{R})$ alors $Q(A) = 0$.*

Ce théorème permet de calculer les puissances d'une matrice plus simplement que par un calcul direct. On peut également utiliser la relation polynomiale $Q(A) = 0$ pour montrer que A est inversible et calculer son inverse.

Exemples.

1. On considère la matrice carrée d'ordre 3,

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

   On a $Q(\lambda) = (1-\lambda)^3$. Alors, par le théorème de Cayley-Hamilton, $(I_3 - A)^3 = 0$.

2. Soit $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. On a $Q(\lambda) = (1-\lambda)(4-\lambda) - 6 = \lambda^2 - 5\lambda - 2$. Le théorème de Cayley-Hamilton affirme que $A^2 - 5A - 2I_2 = 0$. Alors,

$$A(A - 5I_2) = 2I_2 \text{ et } (A - 5I_2)A = 2I_2.$$

   On conclut que A est inversible et

$$A^{-1} = \frac{1}{2}(A - 5I_2) = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}.$$

   D'autre part,

$$A^2 = 5A + 2I_2, \quad A^3 = 5A^2 + 2A = 27A + 10I_2,$$
$$A^4 = A^3A = 27A^2 + 10A = 145A + 54I_2.$$

Un corollaire important du Théorème de Cayley-Hamilton affirme que le polynôme minimal d'une matrice donnée est un diviseur du polynôme caractéristique.

3.6 Endomorphisme et matrice triagonalisables

Nous avons vu au précédent paragraphe qu'il existe des endomorphismes non diagonalisables. Il est néanmoins parfois possible, pour certains d'entre eux, de trouver une base pour laquelle la matrice est assez simple, par exemple triangulaire.

Définition 3.10 *Soit $A \in M_n(\mathbb{K})$. On dit que A est trigonalisable sur $\mathbb{K}$ s'il existe $P \in GL_n(\mathbb{K})$ telle que la matrice $P^{-1}AP$ est triangulaire supérieure.*

Théorème 3.6 *Soit $A \in M_n(\mathbb{K})$. Alors A est trigonalisable si et seulement si le polynôme caractéristique de A a toutes les racines dans $\mathbb{K}$.*

Exemple. Soit $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. On a $P_A(\lambda) = \lambda^2 + 1$, donc A n'est pas trigonalisable sur $\mathbb{R}$, car le polynôme P n'a aucune racine réelle. Par contre A a deux racines propres complexes distinctes : i et $-i$, donc A est diagonalisable sur $\mathbb{C}$.

Définition 3.11 *On dit que l'endomorphisme f est trigonalisable s'il existe une base de E dans laquelle la matrice de f est triangulaire supérieure.*

Grâce à la formule de changement de base, l'endomorphisme f est trigonalisable si et seulement si la matrice de f dans une base de E est trigonalisable sur $\mathbb{K}$.

Tout polynôme non constant à coefficients dans $\mathbb{C}$ a toutes ses racines dans $\mathbb{C}$ d'après le théorème de d'Alembert-Gauss, donc toute matrice dans $M_n(\mathbb{C})$ est trigonalisable sur $\mathbb{C}$.

4 Applications : systèmes de récurrence et systèmes différentiels

Dans ce chapitre, nous illustrons les méthodes de réduction de matrices à travers la résolution de systèmes de récurrence et systèmes différentiels.

4.1 Systèmes de récurrence linéaires

Un système de récurrence linéaire (à coefficients constants homogène) d'ordre 1 est un système du type

$$U_{n+1} = AU_n, \quad n \geq 0. \tag{3}$$

où $U_n \in \mathbb{R}^p$ est le vecteur inconnu de p états à déterminer pour tout n, $A = (a_{ij})_{1 \leq i,j \leq p}$ est un matrice carrée à coefficients réels d'ordre p. On écrit (3) sous forme développée par :

$$\begin{cases} u_{1,n+1} = a_{11}u_{1,n} + a_{12}u_{2,n} + \ldots + a_{1p}u_{p,n} \\ u_{2,n+1} = a_{21}u_{1,n} + a_{22}u_{2,n} + \ldots + a_{2p}u_{p,n} \\ \vdots \\ u_{p,n+1} = a_{p1}u_{1,n-1} + a_{p2}u_{2,n} + \ldots + a_{pp}u_{p,n} \end{cases} \quad \text{et} \quad U_n = \begin{pmatrix} u_{1,n} \\ u_{2,n} \\ \vdots \\ u_{p,n} \end{pmatrix}.$$

Résoudre (3) consiste à trouver toutes les solutions réelles de (3), c'est-à-dire la solution générale. Les solutions de (3) forment un sous-espace vectoriel de dimension p de l'espace des suites de vecteurs de $\mathbb{R}^p$.

Si $p = 1$, le système (3) s'écrit

$$u_{n+1} = au_n, \quad n \geq 0,$$

où $a \in \mathbb{R}$ et, est appelé équation de récurrence d'ordre 1. Alors, $(u_n)_{n\in\mathbb{N}}$ s'agit d'une suite géométrique de raison a et premier terme u_0. La solution générale de cette équation est $u_n = \alpha a^n$, avec $\alpha = u_0 \in \mathbb{R}$.

Remarque 1. On rappelle qu'une suite numérique est une application de $\mathbb{N}$ dans un espace vectoriel F, notée $(u_n)_{n\in\mathbb{N}}$ où $u_n = f(n)$ est appelé le *terme général* de la suite. Si $F = \mathbb{R}^k$ ou $F = \mathbb{C}^k$, on dit que la suite est réelle ou complexe, respectivement. On rappelle que l'espace des suites numériques $(u_n)_{n\in\mathbb{N}}$ muni de l'addition : $(u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}} = (u_n + v_n)_{n\in\mathbb{N}}$, et de la multiplication par un scalaire λ : $\lambda(u_n)_{n\in\mathbb{N}} = (\lambda u_n)_{n\in\mathbb{N}}$, est un espace vectoriel.

La solution générale de (3) est la suite réelle $(U_n)_{n\in\mathbb{N}}$ donnée par $U_n = A^n C$, avec $C \in \mathbb{R}^p$ vecteur arbitraire. Déterminer la solution générale de (3) se ramène donc à déterminer la puissance n-ième de A qui se fait à l'aide de la diagonalisation ou de trigonalisation de la matrice A.

On remarque également que si $\lambda \in \mathbb{R}$ est une valeur propre de A et $v \in \mathbb{R}^p$ est un vecteur propre pour λ, alors $(U_n)_{n\in\mathbb{N}}$, avec $U_n = \lambda^n V$, est solution de (3), où V désigne le vecteur colonne associé à v.

A est diagonalisable sur $\mathbb{R}$.

Il existe une base $\mathcal{B}' = (v_1, \ldots, v_p)$ de $\mathbb{R}^p$ formée de vecteurs propres de A, de valeurs propres correspondants : $\lambda_1, \ldots, \lambda_p$. Alors, $A = PDP^{-1}$ où P est la matrice inversible dont la j-ième colonne est V_j (c'est-à-dire la matrice de passage de la base canonique de $\mathbb{R}^p$ à $\mathcal{B}'$) et

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots \\ 0 & \lambda_2 & 0 & \dots \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & \lambda_p \end{pmatrix}.$$

La solution générale de (3) est alors donnée par la combinaison linéaire des p solutions linéairement indépendentes de (3) :

$$(\lambda_1^n V_1)_{n\in\mathbb{N}}, \dots, (\lambda_p^n V_p)_{n\in\mathbb{N}},$$

c'est-à-dire

$$U_n = \alpha_1 \lambda_1^n V_1 + \dots + \alpha_p \lambda_p^n V_p, \quad \alpha_1, \dots, \alpha_p \in \mathbb{R}. \tag{4}$$

En effet, on a

$$A^n = PD^nP^{-1} \quad \text{avec} \quad D^n = \begin{pmatrix} \lambda_1^n & 0 & 0 & \dots \\ 0 & \lambda_2^n & 0 & \dots \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & \lambda_p^n \end{pmatrix}.$$

La solution générale du système est $(U_n)_{n\in\mathbb{N}}$ de terme général

$$U_n = A^n C = PD^nP^{-1}C$$

On note $\tilde{C} = P^{-1}C \in \mathbb{R}^p$. Alors,

$$U_n = PD^n\tilde{C} = \sum_{j=1}^{p} \tilde{C}_j \lambda_j^n V_j.$$

La solution générale de (3) est alors donnée par (4).

A n'est pas diagonalisable sur $\mathbb{R}$.

Si A est diagonalisable sur $\mathbb{C}$, il existe $(V_1, \dots, V_p)$ base de $\mathbb{C}^p$ formé de vecteurs propres de A. On peut supposer, sans perte de généralité, que $V_1, \dots, V_k \in \mathbb{R}^p$ de valeur propres correspondantes $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ et $V_{k+1}, \dots, V_p \in \mathbb{R}^p$ de valeur propres correspondantes $\mu_1, \bar{\mu}_1 \dots, \mu_m, \bar{\mu}_m \in \mathbb{C}$, avec $p = k + 2m$. La solution générale de (3) est alors donnée par

$$\begin{aligned} U_n &= \alpha_1 \lambda_1^n V_1 + \dots + \alpha_k \lambda_k^n V_k + \\ &\quad + \beta_1 \mathrm{Re}(\mu_1^n V_{k+1}) + \beta_2 \mathrm{Im}(\mu_1^n V_{k+1}) + \dots + \beta_{2m-1} \mathrm{Re}(\mu_m^n V_{p-1}) + \beta_{2m} \mathrm{Im}(\mu_m^n V_{p-1}). \end{aligned}$$

Si A n'est pas diagonalisable ni sur $\mathbb{R}$ ni sur $\mathbb{C}$, on peut toujours trouver une matrice de Jourdan J carrée et triangulaire supérieure d'ordre p et une matrice inversible P telles que $A = PJP^{-1}$. Donc, la solution de (3) est de la forme $U_n = A^n C = PJ^nP^{-1}C = PJ^n\tilde{C}$ avec $\tilde{C} = P^{-1}C \in \mathbb{R}^p$. Enfin, en résolvant le système

$$\tilde{U}_n = P^{-1}U_n = J^n\tilde{C}$$

à partir de la dernière équation (car la matrice J^n est triangulaire supérieure), on obtient $\tilde{U}_n$ et ensuite $U_n = P\tilde{U}_n$.

Dans le cas d'un système de deux équations ($p = 2$), si A n'est pas diagonalisable (ni dans $\mathbb{R}$ ni dans $\mathbb{C}$) alors A admet une seule valeur propre λ_0 réelle telle que $\ker(A - \lambda_0 I)$ est un sous-espace propre de dimension 1.

On peut trouver une matrice inversible P telle que $A = PJP^{-1}$ où $J = \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}$. Les colonnes P_1 et P_2 de la matrice P sont respectivement un vecteur propre associé à λ_0 et un vecteur tel que

$$(A - \lambda_0 I)P_2 = P_1.$$

En effet,

$$A = PJP^{-1} \iff AP = PJ \iff \begin{cases} AP_1 = \lambda_0 P_1 \\ AP_2 = P_1 + \lambda_0 P_2 \end{cases}$$

Il est facile de vérifier (par récurrence sur n) que

$$J^n = \begin{pmatrix} \lambda_0^n & n\lambda_0^n \\ 0 & \lambda_0^n \end{pmatrix}.$$

La solution générale de (3) est alors

$$U_n = A^n C = PJ^n\tilde{C} = (\tilde{c}_1\lambda_0^n + \tilde{c}_2 n\lambda_0^{n-1})P_1 + \tilde{c}_2\lambda_0^n P_2,$$

avec $\tilde{C} = (\tilde{c}_1, \tilde{c}_2)^t = P^{-1}C \in \mathbb{R}^2$.

Exemple 1. On considère deux populations u_n et v_n dont la dynamique est décrite par le système suivant

$$\begin{cases} u_{n+1} = \frac{1}{3}u_n + \frac{1}{2}v_n \\ \\ v_{n+1} = \frac{2}{3}u_n + \frac{1}{2}v_n \end{cases}$$

On veut étudier l'évolution des ces populations lorsque n tend vers $+\infty$.

On pose $U_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix}$. Alors, on a $U_{n+1} = AU_n$, pour tout $n \geq 0$, avec

$$A = \begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \\ \frac{2}{3} & \frac{1}{2} \end{pmatrix}.$$

Les valeurs propres de A sont $\lambda_1 = -\frac{1}{6}$ et $\lambda_2 = 1$. Alors, A est diagonalisable dans $\mathbb{R}$. On a $\ker(A - \lambda_1) = \mathbb{R}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et $\ker(A - \lambda_2) = \mathbb{R}\begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Donc, la solution générale du système de récurrence est

$$U_n = \alpha\left(-\frac{1}{6}\right)^n \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \beta\begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} \alpha\left(-\frac{1}{6}\right)^n + 3\beta \\ -\alpha\left(-\frac{1}{6}\right)^n + 4\beta \end{pmatrix}, \quad \alpha, \beta \in \mathbb{R}.$$

Finalement, on conclut que $(u_n, v_n) \to (3\beta, 4\beta)$ lorsque $n \to +\infty$.

Exemple 2. On considère le système de récurrence $U_{n+1} = AU_n$ avec $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Les valeurs propres de A sont i et $-i$. Un vecteur propre associé à i est $P = (1, -i) = W + iZ$, avec $W = (1, 0)$ et $Z = (0, -1)$. Alors, la solution générale est

$$U_n = \alpha(\cos(n\frac{\pi}{2})W - \sin(n\frac{\pi}{2})Z) + \beta(\cos(n\frac{\pi}{2})Z) + \sin(n\frac{\pi}{2})W) = \alpha \begin{pmatrix} \cos(n\frac{\pi}{2}) \\ \sin(n\frac{\pi}{2}) \end{pmatrix} + \beta \begin{pmatrix} \sin(n\frac{\pi}{2}) \\ -\cos(n\frac{\pi}{2}) \end{pmatrix},$$

$\alpha, \beta \in \mathbb{R}$.

Exemple 3. On considère le système de récurrence

$$\begin{cases} u_{n+1} = 4u_n + v_n \\ v_{n+1} = -u_n + 2v_n \end{cases}$$

Sous forme matricielle, on écrit $U_{n+1} = AU_n$ avec $U_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix}$ et $A = \begin{pmatrix} 4 & 1 \\ -1 & 2 \end{pmatrix}$. On a $\det(A - \lambda I) = (\lambda - 3)^2$. Alors, 3 est l'unique valeur propre de A, d'ordre de multiplicité 2. De plus, $\ker(A - 3I) = \mathbb{R} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et donc A n'est pas diagonalisable. On a $A = PJP^{-1}$ avec

$$P = (P_1, P_2), \quad P_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad J = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$$

et $P_2 = \begin{pmatrix} x \\ y \end{pmatrix}$ telle que

$$AP_2 = P_1 + 3P_2 \iff x + y = 1.$$

Soit $P_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Alors, la solution générale du système de récurrence est

$$U_n = (\alpha 3^n + \beta n 3^n)P_1 + \beta 3^n P_2 = \alpha 3^n \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \beta 3^{n-1} \begin{pmatrix} n \\ 3 - n \end{pmatrix},$$

avec $\alpha, \beta \in \mathbb{R}$.

4.2 Systèmes différentiels linéaires

Le paragraphe précédent montre comment la réduction d'une matrice permet de calculer les puissances d'une matrice et, par conséquent, d'en déduire l'expression des suites récurrentes linéaires associées à cette matrice. Une autre utilisation en analyse provient des systèmes différentielles linéaires, dont on peut exprimer les solutions à l'aide d'une base de vecteurs propres de la matrice associée à cette équation.

On appelle système différentiel linéaire du premier ordre (homogène à coefficients constants) dans $\mathbb{R}^n$ un système de la forme

$$X'(t) = AX(t) \tag{5}$$

où $A \in M_n(\mathbb{R})$ et $X(t) = (x_i(t))_{1 \leq i \leq n}$ est le vecteur des n fonctions inconnues. Sous forme développée, (5) s'écrit :

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

c'est-à-dire

$$x_i'(t) = \sum_{j=1}^{n} a_{ij}x_j(t), \quad \forall 1 \leq i \leq n.$$

Résoudre ou intégrer le système ci-dessus consiste à déterminer la *solution générale* $X : \mathbb{R} \longrightarrow \mathbb{R}^n$, c'est-à-dire toutes les solutions. Le problème de déterminer la solution de (5) qui satisfait la donné initiale $X(t_0) = X_0$, avec $(t_0, X_0) \in J \times \mathbb{R}^n$ donné, s'appelle le problème de Cauchy associé à (t_0, X_0).

L'ensemble des solutions du système (5) est un espace vectoriel de dimension n.

On dit qu'un ensemble de n fonctions vectorielles réelles $(X_1(t), \ldots, X_n(t))$ est une *solution fondamentale* du système (5) si c'est un ensemble de solutions linéairement indépendantes de (5). La solution générale de (5) s'écrit alors

$$X(t) = \alpha_1 X_1(t) + \ldots + \alpha_n X_n(t), \;\; \alpha_j \in \mathbb{R}, 1 \leq j \leq n.$$

On peut montrer que si $\lambda \in \mathbb{R}$ est une valeur propre de A et V est un vecteur propre pour λ, alors $X(t) = e^{\lambda t}V$ est une solution de (5).

Cas simple : A est diagonalisable sur $\mathbb{R}$.

Il existe alors une base $(V_1, \ldots, V_n)$ de $\mathbb{R}^n$ constituée de vecteurs propres de A, de valeurs propres correspondantes $\lambda_1, \ldots, \lambda_n$. On obtient donc n solutions linéairement indépendantes

$$X_j(t) = e^{\lambda_j t}V_j, \quad 1 \leq j \leq n.$$

La solution générale est alors donnée par

$$X(t) = \alpha_1 e^{\lambda_1 t}V_1 + \ldots + \alpha_n e^{\lambda_n t}V_n, \quad \alpha_j \in \mathbb{R}.$$

Cas de la dimension 2 : $n = 2$

On résout le système $X' = AX$ de la façon suivante :

1. Si la matrice A admet deux valeurs propres rélles distinctes λ_1 et λ_2 de vecteurs propres correspondants V_1 et V_2, alors A est diagonalisable sur $\mathbb{R}$ et la solution générale de (5) est
$$X(t) = \alpha_1 e^{\lambda_1 t}V_1 + \alpha_2 e^{\lambda_2 t}V_2, \quad \alpha_1, \alpha_2 \in \mathbb{R}.$$
2. Si la matrice A admet une valeur propre réelle double λ et $\dim\ker(A - \lambda I) = 2$, alors A est diagonalisable sur $\mathbb{R}$ et $A = \lambda I$. La solution générale de (5) est donnée par
$$X(t) = \alpha_1 e^{\lambda t}V_1 + \alpha_2 e^{\lambda t}V_2, \;\; \alpha_1, \alpha_2 \in \mathbb{R},$$
où V_1 et V_2 sont les vecteurs colonnes de la base canonique de $\mathbb{R}^2$.

3. Si la matrice A admet une valeur propre réelle double λ et $\dim\ker(A-\lambda I)=1$, alors A n'est pas diagonalisable sur $\mathbb{R}$. On prend $V_1 \in \ker(A-\lambda I)$ (c'est-à-dire, V_1 est un vecteur propre correspondant à λ) et on prend $V_2 \in \ker((A-\lambda I)^2)$ tel que (V_1, V_2) est une base de $\ker((A-\lambda I)^2)$. Alors, la solution générale de (5) est

$$X(t) = \alpha_1 e^{\lambda t} V_1 + \alpha_2 e^{\lambda t}(I + t(A-\lambda I))V_2, \quad \alpha_1, \alpha_2 \in \mathbb{R}.$$

4. Si la matrice A admet deux valeurs propres complexes conjuguées $\lambda = \alpha + i\beta$ et $\bar{\lambda} = \alpha - i\beta$, A est diagonalisable sur $\mathbb{C}$. Soit $V \in \mathbb{C}^2$ un vecteur propre correspondant à λ. Alors, la solution générale de (5) est

$$X(t) = \alpha_1 \mathrm{Re}(e^{\lambda t} V) + \alpha_2 \mathrm{Im}(e^{\lambda t} V), \quad \alpha_1, \alpha_2 \in \mathbb{R}.$$

Exemples.

1. On considère le système

$$\begin{cases} x'(t) = 5x(t) \\ y'(t) = 3y(t) \end{cases}$$

Il peut s'écrire sous la forme : $X' = AX$ avec

$$A = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix},$$

matrice diagonale.On peut alors résoudre séparément chacune des équations du système et on obtient $x(t) = \alpha e^{5t}$ et $y(t) = \beta e^{3t}$ avec $\alpha, \beta \in \mathbb{R}$.

2. On considère le système

$$\begin{cases} x'(t) = 5x(t) - 3y(t) \\ y'(t) = 2x(t) - 2y(t) \end{cases}$$

Les valeurs propres de la matrice A associée au système sont les racines de $\lambda^2 - 3\lambda - 4 = 0$, soit 4 et -1. La matrice A est diagonalisable. Un vecteur propre associé à 4 est $(3,1)$ et un vecteur propre associé à -1 est $(1,2)$. La solution générale du système est

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \alpha e^{4t} \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \beta e^{-t} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3\alpha e^{4t} + \beta e^{-t} \\ \alpha e^{4t} + 2\beta e^{-t} \end{pmatrix},$$

avec $\alpha, \beta \in \mathbb{R}$.

3. On considère le système

$$\begin{cases} x'(t) = y(t) \\ y'(t) = -x(t) \end{cases}$$

Les valeurs propres de la matrice A associée au système sont les racines de $\lambda^2 + 1 = 0$, donc $\lambda = \pm i$. Un vecteur propre de A associé à i est $(1, i)$. La solution générale est

$$X() = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \alpha \mathrm{Re}\left(e^{it} \begin{pmatrix} 1 \\ i \end{pmatrix}\right) + \beta \mathrm{Im}\left(e^{it} \begin{pmatrix} 1 \\ i \end{pmatrix}\right) = \begin{pmatrix} \alpha\cos t + \beta\sin t \\ -\alpha\sin t + \beta\cos t \end{pmatrix},$$

avec $\alpha, \beta \in \mathbb{R}$.

5 Orthonormalité

Nous nous proposons de définir la notion générale de produit scalaire dans un espace vectoriel réel et d'en étudier les propriétés.

5.1 Produit scalaire sur $\mathbb{R}^n$

Soit E espace vectoriel sur $\mathbb{R}$ (c'est-à-dire $\mathbb{K} = \mathbb{R}$).

Définition 5.1 *Soit $B : E \times E \to \mathbb{R}$ une application. On dit que B est une forme bilinéaire si*

1. *pour tous $x, y, u \in E$ et tout $\lambda \in \mathbb{R}$, $B(u, x+y) = B(u, x) + B(u, y)$ et $B(u, \lambda x) = \lambda B(u, x)$.*
2. *pour tous $x, y, v \in E$ et tout $\lambda \in \mathbb{R}$, $B(x+y, v) = B(x, v) + B(y, v)$ et $B(\lambda x, v) = \lambda B(x, v)$.*

Exemple. Soient $f : E \to \mathbb{R}$ et $g : E \to \mathbb{R}$ des applications linéaires. Alors, l'application $B : E \times E \to \mathbb{R}$ définie par $B(x, y) = f(x)g(y)$ est une forme bilinéaire.

Définition 5.2 *Soit $B : E \times E \to \mathbb{R}$ une forme bilinéaire. On dit que B est une forme bilinéaire symétrique si l'on a $B(x, y) = B(y, x)$ pour tous $x, y \in E$.*

Exemple 1. Soit $B : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ l'application définie par

$$B(x, y) = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1, \quad x = (x_1, x_2), \quad y = (y_1, y_2).$$

L'application B est bilinéaire mais elle n'est pas symétrique.

Exemple 2. Soit $B : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ l'application définie par

$$B(x, y) = 2x_1 y_1 - 3x_2 y_2, \quad x = (x_1, x_2), \quad y = (y_1, y_2).$$

L'application B est bilinéaire symétrique.

Définition 5.3 *On dit que l'application $B : E \times E \to \mathbb{R}$ est un produit scalaire si B est une forme bilinéaire symétrique et pour tout $x \in E$ tel que $x \neq 0$, on a $B(x, x) > 0$.*

Exemple 1. L'application $B : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ définie par

$$B((x_1, x_2), (y_1, y_2)) = x_1 y_1 + x_2 y_2,$$

est un produit scalaire. En effet, B est bilinéaire symétrique et, pour tout $x = (x_1, x_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, on a $B(x, x) = x_1^2 + x_2^2 > 0$.

Plus généralement, si $n \geq 2$, on considère l'application $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ définie par

$$B(x, y) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n \quad \text{si} \quad x = (x_1, \ldots, x_n), \quad y = (y_1, \ldots, y_n).$$

Alors B est un produit scalaire, appelé le *produit scalaire usuel* de $\mathbb{R}^n$.

Exemple 2. L'application $B : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ définie par

$$B((x_1, x_2), (y_1, y_2)) = 2x_1 y_1 + 3x_2 y_2 + 2x_1 y_2 + 2x_2 y_1,$$

est un produit scalaire.

Définition 5.4 *Un espace vectoriel réel muni d'un produit scalaire s'appelle un espace euclidien.*

Si E est un espace euclidien, nous notons en général $<x|y>$ le produit scalaire des vecteurs x et y.

Définition 5.5 *Soient $x, y \in E$ et F sous-espace vectoriel de E. On dit que x et y sont des vecteurs orthogonaux si $< x|y >= 0$. On dit que x est orthorgonal à F si l'on a $< x|y >= 0$, pour tout vecteur $y \in F$.*

Définition 5.6 *Soit F un sous-espace vectoriel de E. On définit l'espace des vecteurs orthogonaux à F, et on l'appelle l'espace orthogonal de F, par*

$$F^{\perp} = \{x \in E :< x|y >= 0, \ \forall\, y \in F\}.$$

Proposition 5.1 *Soient $u_1, \ldots, u_p$ des vecteurs de E et soit F le sous-espace vectoriel de E engendré par $u_1, \ldots, u_p$. Alors $x \in E$ soit orthogonal à F si et seulement si x est orthogonal à u_i, pour tout $1 \leq i \leq p$.*

Exemples.

1. Dans un espace euclidien, tout vecteur est orthogonal au vecteur nul.
2. Soit $\mathbb{R}^2$ muni du produit scalaire usuel. Le vecteur $(x, y) \in \mathbb{R}^2$ est orthogonal à $(1, 2)$ si et seulement si $x + 2y = 0$. Les vecteurs orthogonaux à $(1, 2)$ sont donc les vecteurs colinéaires à $(-2, 1)$.

Définition 5.7 *On dit que $N : E \to \mathbb{R}$ est une norme sur E si*

1. *Pour tout $x \in E$, $N(x) \geq 0$. De plus, si $N(x) = 0$ alors $x = 0$.*
2. *Pour tous $x, y \in E$, $N(x + y) \leq N(x) + N(y)$.*
3. *Pour tout $x \in E$ et pour tout $\lambda \in \mathbb{R}$, $N(\lambda x) = |\lambda| N(x)$.*

Proposition 5.2 *Soit E un espace euclidien et soit $N : E \to \mathbb{R}$ l'application qui à tout $x \in E$ associe $\sqrt{<x|x>}$. Alors, N est une norme sur E appelée norme associée au produit scalaire.*

Si E est un espace euclidien, nous notons $\|x\|$ la norme du vecteur x. Autrement dit, $\|x\| = \sqrt{<x|x>}$. Nous disons que le vecteur x est *unitaire* s'il est de norme 1, c'est-à-dire $\|x\| = 1$.

Si l'on muni $\mathbb{R}^n$ du produit scalaire usuel, la norme associée est définie par

$$\|x\| = \sqrt{<x|x>} = \sqrt{x_1^2 + \ldots + x_n^2}, \ \ x = (x_1, \ldots, x_n) \in \mathbb{R}^n,$$

appelée la norme euclidienne sur $\mathbb{R}^n$.

5.2 Bases orthonormées

Dans ce paragraphe, E est un espace euclidien de dimension finie non nulle.

Proposition 5.3 *Soient $u_1, \ldots, u_p$ des vecteurs de E non nuls et deux à deux orthogonaux. Alors les vecteurs $u_1, \ldots, u_p$ sont linéairement indépendents.*

Preuve. Par hypothèse, on a $<u_i|u_j>= 0$ si $i \neq j$ et $\|u_i\| \neq 0$ pour tout i. Soient $\lambda_1, \ldots, \lambda_p$ des scalaires tels que $\lambda_1 u_1 + \ldots + \lambda_p u_p = 0$. Alors,

$$0 =< \lambda_1 u_1 + \ldots + \lambda_p u_p | u_i >= \lambda_1 < u_1|u_i > + \ldots + \lambda_p < u_p|u_i >= \lambda_i < u_i|u_i >= \lambda_i \|u_i\|^2.$$

On en déduit $\lambda_i = 0$, pour tout $1 \leq i \leq p$. Donc $u_1, \ldots, u_p$ sont linéairement indépendents.

Corollaire 5.1 *Supposons E de dimension n. Si $u_1, \ldots, u_n$ sont des vecteurs de E non nuls et deux à deux orthogonaux, alors $(u_1, \ldots, u_n)$ est une base de E.*

Parmi les bases dont les vecteurs sont deux à deux orthogonaux, celles dont les vecteurs sont unitaires vont jouer un rôle privilégié.

Définition 5.8 *Soit $(e_1, \ldots, e_n)$ une base de E. On dit que $(e_1, \ldots, e_n)$ est une base orthonormée de E si les vecteurs $e_1, \ldots, e_n$ sont deux à deux orthogonaux et si $\|e_i\| = 1$, pour tout $1 \leq i \leq n$.*

Supposons E de dimension n et considérons des vecteurs $u_1, \ldots, u_n$ de E non nuls et deux à deux orthogonaux. Par hypothèse, on a $\|u_i\| \neq 0$ quel que soit i. Posons alors $e_i = \frac{u_i}{\|u_i\|}$ pour tout $1 \leq i \leq n$. Les vecteurs u_i sont unitaires et deux à deux orthogonaux, donc $(e_1, \ldots, e_n)$ est une base orthonormée de E, d'après le corollaire.

Exemple. Soit $\mathbb{R}^n$ muni du produit scalaire usuel. Alors la base canonique est une base orthonormée. Mais, il y en a d'autres. Par exemple, les vecteurs $\frac{1}{5}(3, 4)$ et $\frac{1}{5}(4, -3)$ forment aussi une base orthonormée de $\mathbb{R}^2$.

Théorème 5.1 *Soit $(e_1, \ldots, e_n)$ une base orthornormée de E. Alors*

1. *Pour tout vecteur $x \in E$, on a*

$$x = \sum_{i=1}^{n} < x|e_i > e_i \quad et \quad \|x\|^2 = \sum_{i=1}^{n} < x|e_i >^2 .$$

2. *Pour tous vecteurs $x, y \in E$, on a $<x|y>= \sum_{i=1}^{n} <x|e_i><y|e_i>$.*

Preuve. Soit $x \in E$ et soient $x_1, \ldots, x_n$ les coordonnées de x dans la base $(e_1, \ldots, e_n)$. Il vient

$$< x|e_j >=< \sum_{i=1}^{n} x_i e_i | e_j >= \sum_{i=1}^{n} x_i < e_i|e_j > .$$

La base $(e_1, \ldots, e_n)$ étant orthonormée, on a $<e_i|e_j>= 0$ si $i \neq j$ et $<e_j|e_j>= \|e_j\|^2 = 1$. Il s'ensuit $<x|e_j>= x_j$, pour tout j. Pour tout vecteur $y \in E$, on a donc

$$< x|y >=< \sum_{i=1}^{n} x_i e_i | y >= \sum_{i=1}^{n} x_i < e_i|y >= \sum_{i=1}^{n} < x|e_i >< y|e_i > .$$

Remarque. Soit $(e_1, \ldots, e_n)$ une base orthornormée de E. Alors, $<e_i|e_j>= \delta_{ij}$, symbole de Kronecker, où

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}.$$

5.3 Orthonormalisation de Gram-Schmidt

Le procédé d'orthonormalisation de Gram-Schmidt est un algorithme qui permet, à partir d'une base quelconque $(u_1 \ldots, u_n)$ de E, de trouver une base orthonormée. D'abord nous construisons une base $(v_1, \ldots, v_n)$ de E formée de vecteurs propres deux à deux orthogonaux. Ensuite, nous posons $e_i = \frac{v_i}{\|v_i\|}$, pour tout $1 \leq i \leq n$. Alors $(e_1, \ldots, e_n)$ est une base orthonormée de E.

Soit $(u_1 \ldots, u_n)$ une base de E. Pour tout entier $p \in \{1, \ldots, n\}$, notons F_p le sous-espace vectoriel de E engendré par $u_1 \ldots, u_p$. Ce sous-espace vectoriel est de dimension p.

1. On pose $v_1 = u_1$ et l'on détermine λ tel que $v_2 = u_2 + \lambda v_1$ soit orthogonal à F_1. On a

$$<v_2|v_1>=<u_2|v_1> +\lambda <v_1|v_1>= 0 \iff \lambda = -\frac{<u_2|v_1>}{\|v_1\|^2}.$$

Alors,

$$v_2 = u_2 - \frac{<u_2|v_1>}{\|v_1\|^2} v_1.$$

Pour le vecteur v_2 ainsi déterminé, (v_1, v_2) est une base de F_2 et les vecteurs v_1 et v_2 sont orthogonaux.

2. On cherche ensuite un vecteur $x \in F_2$ de la forme $x = av_1 + bv_2$, pour que le vecteur $v_3 = u_3 + x$ soit orthogonal à F_2, c'est-à-dire $<v_3|v_1>=<v_3|v_2>= 0$. Puisque $<v_1|v_2>= 0$, on a

$$\begin{cases} <u_3|v_1> +a\|v_1\|^2 = 0 \\ <u_3|v_2> +b\|v_2\|^2 = 0 \end{cases} \iff \begin{cases} a = -\frac{<u_3|v_1>}{\|v_1\|^2} \\ b = -\frac{<u_3|v_2>}{\|v_2\|^2} \end{cases}$$

Alors,

$$v_3 = u_3 - \frac{<u_3|v_1>}{\|v_1\|^2} v_1 - \frac{<u_3|v_2>}{\|v_2\|^2} v_2.$$

Les vecteurs non nuls v_1, v_2, v_3 étant deux à deux orthogonaux, ils sont linéairement indépendants et forment donc une base de F_3.

3. Supposons que p est un entier entre 3 et $n-1$ et que l'on a trouvé une base $(v_1, \ldots, v_p)$ de F_p formé de vecteurs deux à deux orthogonaux. On pose

$$v_{p+1} = u_{p+1} - \frac{<u_{p+1}|v_1>}{\|v_1\|^2} v_1 - \cdots - \frac{<u_{p+1}|v_p>}{\|v_p\|^2} v_p.$$

Ce vecteur est non nul, car u_{p+1} n'appartient pas à F_p. On montre facilement que $< v_{p+1}|v_j >= 0$ pour tout $1 \leq j \leq p$ et donc les vecteurs non nuls $v_1, \ldots, v_{p+1}$ sont orthogonaux deux à deux. Par suite $(v_1, \ldots, v_{p+1})$ est une base de F_{p+1}.

Cette méthode permet, à partir de la base $(u_1 \ldots, u_n)$, de calculer la base $(v_1 \ldots, v_n)$ formée de vecteurs deux à deux orthogonaux ayant la propriété suivante :

pour tout entier $p \in \{1, \dots, n\}$, le sous-espace vectoriel de E engendré par les vecteurs $v_1, \dots, v_p$ est égal au sous-espace vectoriel engendré par $u_1, \dots, u_p$.

Nous posons ensuite $e_i = \frac{v_i}{\|v_i\|}$, pour tout $1 \leq i \leq n$. On dit que la base orthonormée $(e_1, \dots, e_n)$ est obtenue à partir de la base $(u_1, \dots, u_n)$ à partir du procédé de Gram-Schmidt.

Exemple. On muni $\mathbb{R}^3$ du produit scalaire usuel. Soient u_1, u_2, u_3 les vecteurs de $\mathbb{R}^3$ définis par $u_1 = (1,1,1)$, $u_2 = (1,0,1)$ et $u_3 = (1,2,0)$. Le déterminant de ces vecteurs dans la base canonique de $\mathbb{R}^3$ est

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} - \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 2 - (2-1) = 1 \neq 0,$$

donc (u_1, u_2, u_3) est une base de $\mathbb{R}^3$. Pour orthonormaliser cette base selon le procédé de Gram-Schmidt, on pose $v_1 = u_1 = (1,1,1)$,

$$v_2 = u_2 - \frac{<u_2|v_1>}{\|v_1\|^2} v_1 = (1,0,1) - \frac{2}{3}(1,1,1) = (\frac{1}{3}, -\frac{2}{3}, \frac{1}{3})$$

et

$$v_3 = u_3 - \frac{<u_3|v_1>}{\|v_1\|^2} v_1 - \frac{<u_3|v_2>}{\|v_2\|^2} v_2 = (1,2,0) - \frac{3}{3}(1,1,1) - \frac{-1}{\frac{2}{3}}(\frac{1}{3}, -\frac{2}{3}, \frac{1}{3}) = (\frac{1}{2}, 0, -\frac{1}{2}).$$

On pose $e_i = \frac{v_i}{\|v_i\|}$, pour $i = 1,2,3$. Alors,

$$e_1 = \frac{\sqrt{3}}{3}(1,1,1), \quad e_2 = \frac{\sqrt{6}}{6}(1,-2,1) \text{ et } e_3 = \frac{\sqrt{2}}{2}(1,0,-1).$$

5.4 Matrices orthogonales.

Soit $(e_1, \dots, e_n)$ une base orthonormée de E. Soient $u_1, \dots, u_n \in E$ et $M \in M_n(\mathbb{R})$ la matrice dont les coefficients de la j-ième colonne sont les coordonnées de u_j dans la base $(e_1, \dots, e_n)$. Dans la matrice ${}^tMM$ le coefficient dans la i-ième ligne et j-ième colonne est égal à $< u_i|u_j >$.

On remarque que la matrice carrée ${}^tMM$ est symétrique, car

$${}^t({}^tMM) = {}^tM{}^t({}^tM) = {}^tMM.$$

Corollaire 5.2 *Si P est la matrice de passage de la base orthonormée $(e_1, \dots, e_n)$ à la base $(u_1, \dots, u_n)$ de E, alors $(u_1, \dots, u_n)$ est orthonormée si et seulement si ${}^tPP = I_n$.*

Définition 5.9 *On dit que $A \in M_n(\mathbb{R})$ est une matrice orthogonale si l'on a ${}^tAA = I_n$.*

Proposition 5.4 *Soit $A \in M_n(\mathbb{R})$ une matrice orthogonale. Alors A est inversible, $A^{-1} = {}^tA$ et le déterminant de A est égal à 1 ou -1.*

Preuve. La matrice A est orthogonale alors ${}^tAA = I_n$. Donc, A est inversible et $A^{-1} = {}^tA$. De plus,

$$1 = \det(I_n) = \det({}^tAA) = \det({}^tA)\det(A) = \det(A)^2.$$

Par conséquent, $\det A = 1$ ou $\det A = -1$.

On remarque si l'on écrit les vecteurs de $\mathbb{R}^n$ en colonnes, alors le produit

$$\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1y_1 + \dots + x_ny_n$$

est le produit scalaire usuel de $\mathbb{R}^n$. Il est donc utile de considérer les vecteurs de $\mathbb{R}^n$ comme des vecteurs-colonnes.

Proposition 5.5 *Une matrice $A \in M_n(\mathbb{R})$ est orthogonale si et seulement si les vecteurs-colonnes de A forment une base orthonormée de l'espace euclidien usuel $\mathbb{R}^n$.*

La matrice I_n est orthogonale et de déterminant 1. Par ailleurs, il existe de matrices orthogonales de déterminant -1, par exemple :

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \frac{\sqrt{2}}{2}\begin{pmatrix} -1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

5.5 Endomorphisme symétrique

Soit E un espace euclidien de dimension finie non nulle. Nous allons étudier les endomorphismes de E qui ont des “bonnes relations” avec le produit scalaire et la norme associée.

Définition 5.10 *On dit que l'endomorphisme f de E est symétrique si*

$$\forall x, y \in E, \quad <f(x)|y> = <x|f(y)> .$$

Exemple 1. On considère $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x, y) = (2x, 3x + y)$. Pour tous, $(x, y), (u, v) \in \mathbb{R}^2$,

$$<f(x,y)|(u,v)> = 2xu + 3xv + yv \quad \text{et} \quad <(x,y)|f(u,v)> = 2xu + 3yu + yv.$$

Donc, f n'est pas symétrique.

Exemple 2. On considère $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x, y) = (2x + 3y, 3x + y)$. Pour tous, $(x, y), (u, v) \in \mathbb{R}^2$,

$$<f(x,y)|(u,v)> = 2xu + 3yu + 3xv + yv = x(2u + 3v) + y(3u + v) = <(x,y)|f(u,v)> .$$

Donc, f est symétrique.

Si l'endomorphisme f est symétrique, alors la forme bilinéaire $B : E \times E \to \mathbb{R}$ définie par

$$B(x, y) = <x|f(y)>, \quad x, y \in E,$$

est symétrique. Si f est différent de id_E, cette forme bilinéaire B est différente du produit scalaire de départ. Par exemple, pour l'exemple 2 ci-dessous, on a

$$\begin{aligned} B((x,y),(u,v)) &= \,<(x,y)|f(u,v)> \\ &= \,<(x,y)|(2u+3v,3u+v)> \\ &= 2xu+3xv+3yu+yv, \end{aligned}$$

pour $(x,y),(u,v)\in\mathbb{R}^2$.

Si f et g sont endomorphismes symétriques et $\lambda\in\mathbb{R}$, alors λf, $f+g$ et $f\circ g$ le sont aussi.

On reconnaît facilement un endomorphisme symétrique par sa matrice dans une base orthonormée. On dit que $A\in M_n(\mathbb{R})$ est symétrique si ${}^tA=A$.

Proposition 5.6 *Soit $\mathcal{B}$ une base orthonormée de E et soit f un endomorphisme de E. Alors f est symétrique si et seulement si la matrice f dans la base $\mathcal{B}$ est symétrique.*

Exemple. On considère $f:\mathbb{R}^3\to\mathbb{R}^3$ la fonction définie par

$$f(x,y,z)=(x+2z,-y+3z,2x+3y+4z).$$

La matrice de f dans la base canonique de $\mathbb{R}^3$ est

$$A=\begin{pmatrix} 1 & 0 & 2\\ 0 & -1 & 3\\ 2 & 3 & 4\end{pmatrix}.$$

La matrice A est symétrique et la base canonique est une base orthonormée de $\mathbb{R}^3$ muni du produit scalaire usuel. Donc, f est un endomorphisme symétrique. En effet, pour tous $(x,y,z),(u,v,w)\in\mathbb{R}^3$,

$$\begin{aligned} <f(x,y,z)|(u,v,w)> &= (x+2z)u+(-y+3z)v+(2x+3y+4z)w \\ &= x(u+2w)+y(-v+3w)+z(2u+3v+4w) \\ &= \,<(x,y,z)|f(u,v,w)>. \end{aligned}$$

Proposition 5.7 *Soit $A\in M_n(\mathbb{R})$. Si A est une matrice symétrique alors les valeurs propres de A sont réelles.*

Remarque. Les valeurs propres de A sont les racines du polynôme caractéristique $Q(\lambda)=\det(A-\lambda I_n)$. Alors, les valeurs propres de A appartiennent a priori à $\mathbb{C}$ (l'ensemble des nombres complexes).

Théorème 5.2 *Soit f un endomorphisme de E. Si l'endomorphisme f est symétrique alors il existe une base orthonormée de E formée de vecteurs propres.*

Le résultat suivant est très important mais théorique.

Corollaire 5.3 *Soit $A\in M_n(\mathbb{R})$. Si A est symétrique alors il existe une matrice orthogonale P telle que ${}^tPAP$ est diagonale.*

Preuve. On muni $\mathbb{R}^n$ du produit scalaire usuel. Soit f l'endomorphisme de $\mathbb{R}^n$ dont la matrice dans la base canonique $\mathcal{B}$ de $\mathbb{R}^n$ est A. Alors, f est un endomorphisme symétrique. D'après le théorème, il existe une base orthonormée $\mathcal{B}'=(v_1,\dots,v_n)$ de $\mathbb{R}^n$ formée de vecteurs propres de f. Soit P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Puisque $\mathcal{B}'$ est formée de vecteurs propres de f, $P^{-1}AP=D$ est une matrice diagonale. De plus, la matrice P est orthogonale car ses colonnes forment une base orthonormée de $\mathbb{R}^n$ (voir proposition 5.5). Donc, $P^{-1}={}^tP$ et $D={}^tPAP$.

5.6 Isométrie

Nous allons nous intéresser aux endomorphismes de E qui ne modifient pas le produit scalaire de deux vecteurs quelconques de E.

Définition 5.11 *Soit f un endomorphisme de E. On dit que f est une isométrie si*

$$\forall x, y \in E, \quad < f(x)|f(y) >=< x|y > .$$

Théorème 5.3 *Soit f un endomorphisme de E. Alors, les propriétés suivantes sont équivalentes :*
(i) f est une isométrie.
(ii) Pour tout $x \in E$, $\|f(x)\| = \|x\|$.
(iii) L'application f est bijective et

$$\forall x, y \in E, \quad < f^{-1}(x)|y >=< x|f(y) > .$$

Preuve. Si f est une isométrie, alors on a $\|f(x)\|^2 =< f(x)|f(x) >=< x|x >= \|x\|^2$, pour tout $x \in E$ et donc (ii) est vérifiée. Réciproquement, supposons (ii). Alors, pour tout $x, y \in E$,

$$\begin{aligned} 2 < f(x)|f(y) > &= \|f(x) + f(y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2 \\ &= \|f(x + y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2 \\ &= \|x + y\|^2 - \|x\|^2 - \|y\|^2 = 2 < x|y > \quad \text{car} \quad \|f(u)\| = \|u\|. \end{aligned}$$

Les propriétés (i) et (ii) sont donc équivalentes.

Supposons (i). Soit $x \in E$, tel que $f(x) = 0$. Alors, $\|x\| = \|f(x)\| = 0$ par (ii). D'où, $x = 0$. L'application linéaire f est donc injective. Puisque f est un endomorphisme d'un espace vectoriel de dimension finie, f est un isomorphisme. De plus, pour tout $x, y \in E$,

$$< f^{-1}(x)|y >=< f(f^{-1}(x))|f(y) >=< x|f(y) > .$$

Réciproquement, si f vérifie la propriété (iii), alors on a, pour tout $y \in E$,

$$\|y\|^2 =< y|y >=< f^{-1}(f(y))|y >=< f(y)|f(y) >= \|f(y)\|^2.$$

Donc, f vérifie la propriété (ii) (et aussi (i)).

Exemples 1. Les applications id_E et $-\mathrm{id}_E$ sont des isométries.

Exemples 2. Soit $\mathbb{R}^4$ muni du produit scalaire usuel et soit $f : \mathbb{R}^4 \to \mathbb{R}^4$ l'application définie par $f(x, y, z, w) = (y, x, w, z)$. Alors, pour tout $(x, y, z, w) \in R^4$, on a

$$\|f(x, y, z, w)\|^2 = y^2 + x^2 + w^2 + z^2 = x^2 + y^2 + z^2 + w^2 = \|(x, y, z, w)\|^2.$$

On conclut que f est une isométrie de $\mathbb{R}^4$.

Le résultat suivant établi le lien entre les isométries et les matrices orthogonales.

Proposition 5.8 *Soit $\mathcal{B}$ une base orthonormée de E. L'endomorphisme f de E est une isométrie si et seulement si la matrice de f dans la base $\mathcal{B}$ est une matrice orthogonale.*

Preuve. Soit $A = M(f; \mathcal{B}, \mathcal{B})$. On sait que ${}^tAA = (b_{ij})$ où $b_{ij} =< f(e_i)|f(e_j) >$. Donc, f est une isométrie si et seulement si ${}^tAA = I_n$, c'est-à-dire A est orthogonale.

Soit $(e_1, \ldots, e_n)$ une base orthonormée de E. Nous déduisons du résultat précédent que la matrice de f dans la base $(e_1, \ldots, e_n)$ est orthogonale si et seulement si $(f(e_1), \ldots, f(e_n))$ est une base orthonormée de E.

Remarque. Si f est une isométrie de E, alors l'image de f d'une base orthonormée de E est une base orthonormée de E.

Exemple. Soit $\mathbb{R}^4$ muni du produit scalaire usuel et soit f l'endomorphisme de $\mathbb{R}^4$ dont la matrice dans la base canonique est

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

On veut montrer que f est une isométrie. Puisque la base canonique de $\mathbb{R}^4$ est orthonormée pour le produit scalaire usuel, f est une isométrie si et seulement si la matrice A est orthogonale. On a

$${}^tAA = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} = I_4,$$

donc A est orthogonale. On conclut que f est une isométrie.

Une isométrie n'a pas nécessairement de valeur propre. Par exemple, soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x, y) = (-y, x)$. Le polynôme caractéristique de f est $Q(\lambda) = \lambda^2 + 1$ qui n'a pas de racines dans $\mathbb{R}$. Donc, f n'a pas de valeur propre. Mais si une isométrie a une valeur propre, le résultat suivant montre que cette valeur n'est pas quelconque.

Proposition 5.9 *Si λ est une valeur propre d'une isométrie f de E, alors $\lambda = 1$ ou $\lambda = -1$.*

Preuve. Soit f une isométrie de E. Supposons que λ est une valeur propre de f et $x \in E$ est tel que $x \neq 0$ et $f(x) = \lambda x$. Alors, $|\lambda|\|x\| = \|f(x)\| = \|x\|$. On en déduit $|\lambda| = 1$, car $\|x\| \neq 0$. Donc, $\lambda = 1$ ou $\lambda = -1$.

5.7 Produit hermitien sur $\mathbb{C}^n$

La notion de produit scalaire sur $\mathbb{R}^n$ peut-être généralisée, en certain sens, à $\mathbb{C}^n$. Soit E espace vectoriel sur $\mathbb{C}$ (c'est-à-dire $\mathbb{K} = \mathbb{C}$).

Définition 5.12 *Soit $B : E \times E \to \mathbb{C}$ une application. On dit que B est une forme sesquilinéaire si*

1. *pour tous $x, y, u \in E$ et tout $\lambda \in \mathbb{C}$, $B(u, x + y) = B(u, x) + B(u, y)$ et $B(u, \lambda x) = \overline{\lambda} B(u, x)$.*

2. *pour tous $x, y, v \in E$ et tout $\lambda \in \mathbb{C}$, $B(x+y, v) = B(x, v) + B(y, v)$ et $B(\lambda x, v) = \lambda B(x, v)$.*

Exemple. Soient $f : E \to \mathbb{C}$ et $g : E \to \mathbb{C}$ des applications linéaires. Alors, l'application $B : E \times E \to \mathbb{C}$ définie par $B(x, y) = f(x)\bar{g}(y)$ est une forme bilinéaire.

Définition 5.13 *Soit $B : E \times E \to \mathbb{C}$ une forme sesquilinéaire. On dit que B est une forme hermitienne si l'on a $B(x, y) = \overline{B(y, x)}$ pour tous $x, y \in E$.*

Exemple 1. Soit $B : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ l'application définie par

$$B(x, y) = x_1\bar{y}_2 - x_2\bar{y}_1, \;\; x = (x_1, x_2), \;\; y = (y_1, y_2).$$

L'application B est sesquilinéaire mais elle n'est pas hermitienne.

Exemple 2. Soit $B : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ l'application définie par

$$B(x, y) = 2x_1\bar{y}_1 - 3x_2\bar{y}_2, \;\; x = (x_1, x_2), \;\; y = (y_1, y_2).$$

L'application B est hermitenne.

Définition 5.14 *On dit que l'application $B : E \times E \to \mathbb{C}$ est un produit hermitien si B est une forme hermitienne et pour tout $x \in E$ tel que $x \neq 0$, on a $B(x, x) > 0$.*

Exemple 1. L'application $B : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ définie par

$$B((x_1, x_2), (y_1, y_2)) = x_1\bar{y}_1 + x_2\bar{y}_2,$$

est un produit scalaire. En effet, B est hermitienne et, pour tout $x = (x_1, x_2) \in \mathbb{C}^2 \setminus \{(0, 0)\}$, on a $B(x, x) = |x_1|^2 + |x_2|^2 > 0$.

Plus généralement, si $n \geq 2$, on considère l'application $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ définie par

$$B(x, y) = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n \quad \text{si} \quad x = (x_1, \ldots, x_n), \;\; y = (y_1, \ldots, y_n).$$

Alors B est un produit hermitien, appelé le *produit hermitien usuel* de $\mathbb{C}^n$.

Exemple 2. L'application $B : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ définie par

$$B((x_1, x_2), (y_1, y_2)) = 2x_1\bar{y}_1 + 3x_2\bar{y}_2 + 2x_1\bar{y}_2 + 2x_2\bar{y}_1,$$

est un produit hermitien.

Définition 5.15 *Un espace vectoriel complexe muni d'un produit hermitien s'appelle un espace hermitien.*

Si E est un espace hermitien, nous notons en général $< x|y >$ le produit hermitien des vecteurs x et y.

Définition 5.16 *On dit que $N : E \to \mathbb{R}$ est une norme sur E si*

1. *Pour tout $x \in E$, $N(x) \geq 0$. De plus, si $N(x) = 0$ alors $x = 0$.*
2. *Pour tous $x, y \in E$, $N(x+y) \leq N(x) + N(y)$.*

3. Pour tout $x \in E$ et pour tout $\lambda \in \mathbb{C}$, $N(\lambda x) = |\lambda| N(x)$.

Proposition 5.10 *Soit E un espace euclidien et soit $N : E \to \mathbb{R}$ l'application qui à tout $x \in E$ associe $\sqrt{<x|x>}$. Alors, N est une norme sur E appelée norme associée au produit hermitien.*

Si E est un espace euclidien, nous notons $\|x\|$ la norme du vecteur x. Autrement dit, $\|x\| = \sqrt{<x|x>}$. Nous disons que le vecteur x est *unitaire* s'il est de norme 1, c'est-à-dire $\|x\| = 1$.

Si l'on muni $\mathbb{C}^n$ du produit hermitien usuel, la norme associée est définie par

$$\|x\| = \sqrt{<x|x>} = \sqrt{|x_1|^2 + \ldots + |x_n|^2}, \;\; x = (x_1, \ldots, x_n) \in \mathbb{C}^n.$$

Définition 5.17 *Soit E un espace hermitien et f un endomorphisme sur E. On dit que f est hermitien si*

$$\forall x, y \in E, \quad < f(x)|y >=< x|f(y) > .$$

Les opérateurs hermitiens jouent un rôle importanten mécanique quantique, car ils répresentent les grandeurs physiques. Les valeurs propres (réelles) répresentent les valeurs possibles de la grandeur et les fonctions propres (ou vecteurs) les états associés.

Soit $A \in M_n(\mathbb{C})$. On dit que la matrice A est *hermitienne* (ou *auto-adjointe*) si

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

Proposition 5.11 *Soit A la matrice d'un endomorphisme f dans une base orthonormée de E. Alors, f est un opérateur hermitien si et seulement si A est une matrice hermitienne.*

Exemple. La matrice $A = \begin{pmatrix} 3 & i & -5i \\ -i & -2 & 5 \\ 5i & 5 & 10 \end{pmatrix}$, dans $M_3(\mathbb{C})$, est hermitenne.

Références

[1] François Liret, Dominique Martinais, *Algèbre et Géométrie, 1er Année,* Dunod 2003.

[2] François Liret, Dominique Martinais, *Algèbre et Géométrie, 2ème Année,* Dunod 2002.

[3] Jean-Pierre Marco, Philippe Thieullen, Jacques-Arthur Weil, *Mathématiques 2,* Cursus LMD, Pearson Education 2007.