Chapitre 6

Algèbre linéaire

6.1 Applications linéaires et espaces vectoriels

Rappels

Une *application linéaire f de $\mathbb{R}$ dans $\mathbb{R}$* est une fonction dont le graphe est une droite passant par l'origine. Par conséquent son expression fonctionnelle est la suivante.

$$f : \mathbb{R} \to \mathbb{R};\ v \mapsto av \quad \text{où } a \text{ est un nombre réel}$$

On remarque que cette fonction satisfait les propriétés suivantes.

$$\begin{aligned} &1) \quad f(v_1 + v_2) = f(v_1) + f(v_2) \text{ pour tout } v_1, v_2 \in \mathbb{R}. \\ &2) \quad f(\lambda v) = \lambda f(v) \text{ pour tout } \lambda, v \in \mathbb{R}. \end{aligned}$$

Ces propriétés sont intéressantes car elles concernent les deux opérations essentielles sur les nombres réels (l'addition et la multiplication).

En remplaçant λ par 0 dans la deuxième propriété, on montre qu'une application linéaire satisfait toujours la condition $f(0) = 0$.

Applications linéaires dans le plan

Il s'agit maintenant d'applications de $\mathbb{R}^2$ dans $\mathbb{R}^2$, qui associent un vecteur à un vecteur. Il y a deux opérations essentielles sur les vecteurs dans le plan : on peut additionner deux vecteurs ou multiplier un vecteur par un nombre réel. On va ainsi définir les applications linéaires dans le plan comme suit.

Une *application linéaire f de $\mathbb{R}^2$ dans $\mathbb{R}^2$* est une fonction qui satisfait :

$$\begin{aligned} &1) \quad \overrightarrow{f(\vec{v}_1 + \vec{v}_2)} = \overrightarrow{f(\vec{v}_1)} + \overrightarrow{f(\vec{v}_2)} \text{ pour tout } \vec{v}_1, \vec{v}_2 \in \mathbb{R}^2. \\ &2) \quad \overrightarrow{f(\lambda\vec{v})} = \lambda\overrightarrow{f(\vec{v})} \text{ pour tout } \lambda \in \mathbb{R} \text{ et } \vec{v} \in \mathbb{R}^2. \end{aligned}$$

On verra que l'expression fonctionnelle de f est la suivante.

$$f : \mathbb{R}^2 \to \mathbb{R}^2;\ \vec{v} \mapsto \overrightarrow{A\vec{v}} \quad \text{où } A \text{ est une matrice réelle de taille } 2 \times 2$$

En remplaçant λ par 0 dans la deuxième propriété, on montre qu'une application linéaire satisfait toujours la condition $f(0) = 0$.

Remarque sur la notation

On va alléger les notations en ne notant plus les petites flèches au dessus des vecteurs.

Applications linéaires dans l'espace

Il s'agit maintenant d'applications de $\mathbb{R}^3$ dans $\mathbb{R}^3$, qui associent un vecteur à un vecteur. Il y a deux opérations essentielles sur les vecteurs dans l'espace : on peut additionner deux vecteurs ou multiplier un vecteur par un nombre réel. On va ainsi définir les applications linéaires dans l'espace comme suit.

Une *application linéaire f de $\mathbb{R}^3$ dans $\mathbb{R}^3$* est une fonction qui satisfait :

$$1) \quad f(v_1 + v_2) = f(v_1) + f(v_2) \text{ pour tout } v_1, v_2 \in \mathbb{R}^3.$$

$$2) \quad f(\lambda v) = \lambda f(v) \text{ pour tout } \lambda \in \mathbb{R} \text{ et } v \in \mathbb{R}^3.$$

On verra que l'expression fonctionnelle de f est la suivante.

$$f : \mathbb{R}^3 \to \mathbb{R}^3; \; v \mapsto Av \quad \text{où } A \text{ est une matrice réelle de taille } 3 \times 3$$

En remplaçant λ par 0 dans la deuxième propriété, on montre qu'une application linéaire satisfait toujours la condition $f(0) = 0$.

Généralisation

Les propriétés des applications linéaires sont basées sur le fait qu'on peut additionner deux vecteurs entre-eux et qu'on peut multiplier un vecteur par un nombre réel. Pour généraliser ce principe, on va imiter ce qu'il se passe avec $\mathbb{R}$, $\mathbb{R}^2$ et $\mathbb{R}^3$.

Pour cela on a besoin d'ensembles (comme $\mathbb{R}$, $\mathbb{R}^2$ ou $\mathbb{R}^3$) où l'on peut additionner les éléments entre-eux (on dit que l'addition est une *opération interne*) et où l'on peut multiplier un élément par un nombre réel (on dit que la multiplication est une *opération externe*).

Un tel ensemble est appelé un *espace vectoriel* et est en général noté V (les lois régissant cette addition et cette multiplication se trouvent en page 70). Les éléments d'un espace vectoriel V sont appelés *vecteurs* et sont en général notés v (sans petite flèche).

Définition : applications linéaires

On peut maintenant définir ce qu'est une application linéaire dans ce cadre plus général. Soit V et W deux espaces vectoriels. On dit qu'une application $f : V \to W$ est une *application linéaire* si les propriétés suivantes sont vérifiées.

$$1) \quad f(v_1 + v_2) = f(v_1) + f(v_2) \text{ pour tout } v_1, v_2 \in V.$$

$$2) \quad f(\lambda v) = \lambda f(v) \text{ pour tout } \lambda \in \mathbb{R} \text{ et } v \in V.$$

On verra que l'expression fonctionnelle de f est la suivante.

$$f : V \to W; \; v \mapsto Av \quad \text{où } A \text{ est un opérateur linéaire}$$ [1]

En remplaçant λ par 0 dans la deuxième propriété, on montre qu'une application linéaire satisfait toujours la condition $f(0) = 0$.

1. Dans le cas où les dimensions de V et de W sont finies, de dimension respective m et n, l'opérateur linéaire d'une application linéaire $f : V \to W$ est une matrice réelle de taille $n \times m$

Exemples d'espace vectoriel

1. Les nombres réels forment un espace vectoriel réel, noté $\mathbb{R}$.
2. Les vecteurs du plan forment un espace vectoriel réel, noté $\mathbb{R}^2$. Dans cet espace, on va identifier les points P avec les vecteurs $\overrightarrow{OP}$.
3. Il en va de même pour $\mathbb{R}^3$, $\mathbb{R}^4$, ..., $\mathbb{R}^n$, ...
4. L'ensembles des suites de nombres réels indicées dans les nombres naturels forment un espace vectoriel réel, noté $\mathbb{R}^{\mathbb{N}}$.
5. Les polynômes à coefficients dans $\mathbb{R}$ de degré au plus n (avec le polynôme nul) forment un espace vectoriel réel, noté $\mathcal{P}_n(\mathbb{R})$.
6. Les polynômes à coefficients dans $\mathbb{R}$ forment un espace vectoriel réel, noté $\mathcal{P}(\mathbb{R})$.
7. Les fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ forment un espace vectoriel réel, noté $\mathcal{C}(\mathbb{R})$.

Exemples d'applications linéaires

1. La rotation du plan centrée à l'origine d'angle $\frac{\pi}{3}$.
$$r_{\frac{\pi}{3}} : \mathbb{R}^2 \to \mathbb{R}^2;\ \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \frac{1}{2}x - \frac{\sqrt{3}}{2}y \\ \frac{\sqrt{3}}{2}x + \frac{1}{2}y \end{pmatrix}$$
2. La dérivée est une application linéaire de $\mathcal{P}_n(\mathbb{R})$ dans $\mathcal{P}_{n-1}(\mathbb{R})$.
$$D : \mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_{n-1}(\mathbb{R});\ p(x) \mapsto D(p(x)) = p'(x)$$
3. L'intégrale est une fonction linéaire de $\mathcal{C}(\mathbb{R})$ dans $\mathcal{C}(\mathbb{R})$ (ou de $\mathcal{P}_{n-1}(\mathbb{R})$ dans $\mathcal{P}_n(\mathbb{R})$).
$$I : \mathcal{C}(\mathbb{R}) \to \mathcal{C}(\mathbb{R});\ f(x) \mapsto I(f(x)) = \int_0^x f(t)\,dt$$
4. Le décalage à gauche est une application linéaire de $\mathbb{R}^{\mathbb{N}}$ dans $\mathbb{R}^{\mathbb{N}}$.
$$\sigma_l : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}};\ (x_0; x_1; x_2; x_3; \ldots) \to \sigma_l((x_0; x_1; x_2; x_3; \ldots)) = (x_1; x_2; x_3; x_4; \ldots)$$
5. Le décalage à droite est aussi une application linéaire de $\mathbb{R}^{\mathbb{N}}$ dans $\mathbb{R}^{\mathbb{N}}$.
$$\sigma_r : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}};\ (x_0; x_1; x_2; x_3; \ldots) \to \sigma_r((x_0; x_1; x_2; x_3; \ldots)) = (0; x_0; x_1; x_2; x_3; \ldots)$$
6. Le produit scalaire dans $\mathbb{R}^2$ est une application bilinéaire. Cela signifie que le produit scalaire est linéaire en chacune des deux variables.
$$\bullet : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R};\ (\vec{v}_1; \vec{v}_2) \mapsto \vec{v}_1 \bullet \vec{v}_2$$
7. Le produit scalaire dans $\mathbb{R}^n$ est une application bilinéaire.
$$\bullet : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R};\ \left(\begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}; \begin{pmatrix} y_1 \\ \cdots \\ y_n \end{pmatrix} \right) \mapsto x_1y_1 + \cdots + x_ny_n$$
8. Le produit vectoriel dans $\mathbb{R}^3$ est une application bilinéaire.
$$\wedge : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3;\ (\vec{v}_1; \vec{v}_2) \mapsto \vec{v}_1 \wedge \vec{v}_2$$
9. Le déterminant dans $\mathbb{R}^n$ est une application multilinéaire. Cela signifie que le déterminant est linéaire en chaque variable.
$$\bullet : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ termes}} \to \mathbb{R};\ (\vec{v}_1; \ldots; \vec{v}_n) \mapsto \det(\vec{v}_1; \ldots; \vec{v}_n)$$

6.2 Bases et matrices

6.2.1 Combinaison linéaire de vecteurs

Soit $S = \{v_1, \ldots, v_n\}$ un ensemble de n vecteurs dans un espace vectoriel V.

1. Le vecteur v est une *combinaison linéaire* de ces n vecteurs lorsque

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \quad \text{avec } \lambda_i \in \mathbb{R}$$

   Les nombres λ_i sont les *coefficients* de cette combinaison linéaire.

2. Les vecteurs de S sont *linéairement indépendants* si la seule solution de l'équation

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0 \quad \text{d'inconnues } \lambda_i \in \mathbb{R}$$

   est la solution $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.
   Autrement dit si on ne peut pas exprimer un des vecteurs de S comme combinaison linéaire des autres.

6.2.2 Bases d'un espace vectoriel

Soit $S = \{v_1, \ldots, v_n\}$ un ensemble de n vecteurs dans un espace vectoriel V.

1. S est un *système générateur* si tout élément de V s'écrit comme combinaison linéaire des vecteurs de S.
2. S est un *système libre* si les vecteurs de S sont linéairement indépendants.

Remarque importante

Si S est un système libre et que v s'écrit comme combinaison linéaire des vecteurs de S, alors cette combinaison linéaire est unique.

En effet, supposons qu'on ait deux combinaisons linéaires de v.

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$$

En basculant tout du même côté, on obtient :

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_n - \beta_n)v_n = 0$$

Comme les vecteurs de S sont linéairement indépendants, on a :

$$\alpha_1 - \beta_1 = 0,\ \alpha_2 - \beta_2 = 0,\ \ldots,\ \alpha_n - \beta_n = 0$$

Autrement dit :

$$\alpha_1 = \beta_1,\ \alpha_2 = \beta_2,\ \ldots,\ \alpha_n = \beta_n$$

Ce qui montre que les combinaisons linéaires sont les mêmes.

Base d'un espace vectoriel

Le système S est une base de V si tout élément de V s'écrit comme unique combinaison linéaire des vecteurs de S. Grâce à la remarque importante ci-dessus, on a l'équivalence suivante.

$$S \text{ est une base de } V \iff S \text{ est un système libre et générateur}$$

Théorème (sans preuve)

Deux bases d'un même espace vectoriel ont toujours le même nombre de vecteurs. Si ce nombre est fini, on parle d'espace vectoriel de dimension finie ; sinon, on parle d'espace vectoriel de dimension infinie.

Définition

La *dimension* de l'espace vectoriel est le nombre de vecteurs de base de cet espace.

Convention. La dimension d'un point étant nulle, on dit que $\dim(\{0\}) = 0$.

Exemples de bases

1. Le système S formé des n vecteurs suivants est une base de $\mathbb{R}^n$. C'est la *base canonique.*
$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \cdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \cdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ \cdots \\ 0 \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \cdots \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \cdots \\ 0 \\ 1 \end{pmatrix} \right\}$$
Ainsi, $\mathbb{R}^n$ est un espace vectoriel de dimension n.

2. Le système S ci-dessous est une base de $\mathcal{P}_n(\mathbb{R})$.
$$S = \{1, x, x^2, x^3, \ldots, x^n\}$$
Ainsi, $\mathcal{P}_n(\mathbb{R})$ est un espace vectoriel de dimension $n + 1$.

3. L'espace vectoriel des polynômes à coefficients réels $\mathcal{P}(\mathbb{R})$ n'admet pas de bases avec un nombre fini d'éléments. C'est donc un espace vectoriel de dimension infinie.

4. Les espaces vectoriels $\mathbb{R}^{\mathbb{N}}$ et $\mathcal{C}(\mathbb{R})$ sont aussi de dimension infinie.

5. Les nombres complexes $\mathbb{C}$ forment un espace vectoriel réel de dimension 2 de base $\{1, i\}$ (ou un espace vectoriel complexe de dimension 1 de base $\{1\}$).

Remarque

L'algèbre linéaire traite principalement de l'étude des applications linéaires dans des espaces vectoriels de dimension finie, tandis que l'analyse fonctionnelle traite principalement de l'étude des applications linéaires dans des espaces vectoriels de dimension infinie.

Puisque ce cours est un cours d'algèbre linéaire, on se restreindra aux espaces vectoriels dont la dimension est finie.

Exemples de combinaison linéaire

On considère la base canonique $S = \{v_1, v_2\}$ du plan $\mathbb{R}^2$ où $v_1 = \binom{1}{0}$ et $v_2 = \binom{0}{1}$.

Voici quelques vecteurs de $\mathbb{R}^2$ exprimés comme combinaison linéaire des vecteurs de S.
$$\begin{pmatrix} 3 \\ 4 \end{pmatrix} = 3v_1 + 4v_2 \qquad \begin{pmatrix} -5 \\ 2 \end{pmatrix} = -5v_1 + 2v_2 \qquad \begin{pmatrix} 1 \\ -7 \end{pmatrix} = v_1 - 7v_2$$

6.2.3 La notation en vecteurs

Soit V un espace vectoriel de dimension finie m. Soit $S = \{v_1, v_2, \dots, v_m\}$ une base de V. Considérons deux vecteurs v et v' de l'espace vectoriel V. On peut décrire v et v' comme unique combinaison des vecteurs de S et ainsi se ramener à la notation habituelle des vecteurs

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m = \sum_{i=1}^{m} \alpha_i v_i \overset{\text{notation}}{=} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_m \end{pmatrix}$$

$$v' = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_m v_m = \sum_{i=1}^{m} \beta_i v_i \overset{\text{notation}}{=} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdots \\ \beta_m \end{pmatrix}$$

Important. L'écriture en vecteurs de v et de v' dépend du choix de la base S de V.

Exemple

Dans le plan $\mathbb{R}^2$, on considère la base canonique $S = \{v_1, v_2\}$ et la base $T = \{v_3, v_4\}$ où $v_1 = \binom{1}{0}$, $v_2 = \binom{0}{1}$, $v_3 = \binom{2}{1}$ et $v_4 = \binom{-3}{2}$. Le vecteur $\binom{-1}{3}$ s'écrit

$$\begin{pmatrix} -1 \\ 3 \end{pmatrix} = -v_1 + 3v_2 \overset{\text{notation}}{=} \begin{pmatrix} -1 \\ 3 \end{pmatrix}_S \quad \text{dans la base } S \text{ et}$$

$$\begin{pmatrix} -1 \\ 3 \end{pmatrix} = v_3 + v_4 \overset{\text{notation}}{=} \begin{pmatrix} 1 \\ 1 \end{pmatrix}_T \quad \text{dans la base } T.$$

Les deux propriétés de la notation en vecteurs

1. **Propriété d'addition**

   Regardons comment réagit la notation en vecteurs pour la somme de v et de v'. Grâce à une propriété du symbole $\sum$, on obtient

$$\begin{aligned} v + v' &= \sum_{i=1}^{m} \alpha_i v_i + \sum_{i=1}^{m} \beta_i v_i = \sum_{i=1}^{m} (\alpha_i + \beta_i) v_i \\ &\overset{\text{notation}}{\Longleftrightarrow} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_m \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdots \\ \beta_m \end{pmatrix} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \cdots \\ \alpha_m + \beta_m \end{pmatrix} \end{aligned}$$

2. **Propriété de multiplication par un nombre réel**

   Regardons comment réagit la notation en vecteurs pour le produit λv avec $\lambda \in \mathbb{R}$. Grâce à une propriété du symbole $\sum$, on obtient

$$\begin{aligned} \lambda v &= \lambda \sum_{i=1}^{m} \alpha_i v_i = \sum_{i=1}^{m} (\lambda \alpha_i) v_i \\ &\overset{\text{notation}}{\Longleftrightarrow} \lambda \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} \lambda\alpha_1 \\ \lambda\alpha_2 \\ \cdots \\ \lambda\alpha_m \end{pmatrix} \end{aligned}$$

On retrouve ainsi les principes de calculs qu'on a avec les vecteurs de $\mathbb{R}^2$ et de $\mathbb{R}^3$.

Important. Ces propriétés ne sont valables que si les écritures en vecteurs de v et de v' sont effectuées dans la même base. Le vecteur résultant est aussi écrit dans cette base.

6.2.4 Matrices associées à une application linéaire

Soit V et W deux espaces vectoriels de dimension finie respective m et n.

Soit $S = \{v_1, v_2, \ldots, v_m\}$ une base de V et $T = \{w_1, w_2, w_3, \ldots, w_n\}$ une base de W.

On peut décrire chaque vecteur v de V comme unique combinaison des vecteurs de S (on se retient d'utiliser la notation en vecteurs pour l'instant).

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m = \sum_{i=1}^{m} \alpha_i v_i$$

Comme l'application $f : V \to W$ est linéaire, on peut écrire

$$f(v) = f(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m) \overset{f \text{ linéaire}}{=} \alpha_1 f(v_1) + \alpha_2 f(v_2) + \cdots + \alpha_m f(v_m) = \sum_{i=1}^{m} \alpha_i f(v_i)$$

Ainsi, il suffit de connaître les images des vecteurs de base $f(v_1)$, $f(v_2)$, $\ldots$, $f(v_m)$ pour pouvoir connaître l'image de n'importe quel vecteur par la fonction f.

Or $f(v_1)$, $f(v_2)$, $\ldots$, $f(v_m)$ sont des vecteurs de W qui s'écrivent aussi dans la base $T = \{w_1, w_2, w_3, \ldots, w_n\}$ de W. Ainsi, il existe des nombres $a_{i,j}$ $(1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n)$ tels que

$$\begin{array}{cccc}
f(v_1) & f(v_2) & \cdots & f(v_m) \\
\parallel & \parallel & & \parallel \\
a_{1,1}w_1 & a_{1,2}w_1 & \cdots & a_{1,m}w_1 \\
+ & + & & + \\
a_{2,1}w_2 & a_{2,2}w_2 & \cdots & a_{2,m}w_2 \\
+ & + & & + \\
a_{3,1}w_3 & a_{3,2}w_3 & \cdots & a_{3,m}w_3 \\
+ & + & & + \\
\vdots & \vdots & \cdots & \vdots \\
+ & + & & + \\
a_{n,1}w_n & a_{n,2}w_n & \cdots & a_{n,m}w_n
\end{array}$$

Autrement dit, avec la notation en vecteurs

$$f(v_i) = \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ a_{3,i} \\ \cdots \\ a_{n,i} \end{pmatrix}$$

pour tout i de 1 à m

En utilisant les deux propriétés de la notation en vecteurs vues en page 40, on obtient

$$f(v) = \sum_{i=1}^{m} \alpha_i f(v_i) \overset{\text{notation}}{=} \sum_{i=1}^{m} \alpha_i \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ a_{3,i} \\ \cdots \\ a_{n,i} \end{pmatrix} = \sum_{i=1}^{m} \begin{pmatrix} a_{1,i}\alpha_i \\ a_{2,i}\alpha_i \\ a_{3,i}\alpha_i \\ \cdots \\ a_{n,i}\alpha_i \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{m} a_{1,i}\alpha_i \\ \sum_{i=1}^{m} a_{2,i}\alpha_i \\ \sum_{i=1}^{m} a_{3,i}\alpha_i \\ \cdots \\ \sum_{i=1}^{m} a_{n,i}\alpha_i \end{pmatrix}$$

Ainsi quelque soit le vecteur v, on voit que pour définir $f(v)$, on fait apparaître $n \times m$ nombres réels. Une idée géniale a été de les mettre dans un tableau appelé *matrice*.

$$f(v) \overset{\text{notation}}{=} \underbrace{\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ a_{3,1} & a_{3,2} & \cdots & a_{3,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix}}_{\text{matrice associée à } f\text{, notée } A} \underbrace{\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_m \end{pmatrix}}_{\text{vecteur } v} = \underbrace{\begin{pmatrix} a_{1,1}\alpha_1 + a_{1,2}\alpha_2 + \cdots + a_{1,m}\alpha_m \\ a_{2,1}\alpha_1 + a_{2,2}\alpha_2 + \cdots + a_{2,m}\alpha_m \\ a_{3,1}\alpha_1 + a_{3,2}\alpha_2 + \cdots + a_{3,m}\alpha_m \\ \cdots \\ a_{n,1}\alpha_1 + a_{n,2}\alpha_2 + \cdots + a_{n,m}\alpha_m \end{pmatrix}}_{\text{vecteur image } f(v)\text{, noté } Av}$$

La matrice A est une matrice de taille $n \times m$ (on indique toujours le nombre de lignes puis ensuite le nombre de colonnes).

Mieux encore, on remarque que les colonnes de la matrice A correspondent aux vecteurs $f(v_i)$. En effet, on a :

$$A = \begin{pmatrix} \boxed{\begin{matrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \\ \cdots \\ a_{n,1} \end{matrix}} & \boxed{\begin{matrix} a_{1,2} \\ a_{2,2} \\ a_{3,2} \\ \cdots \\ a_{n,2} \end{matrix}} & \boxed{\begin{matrix} \cdots \\ \cdots \\ \cdots \\ \cdots \\ \cdots \end{matrix}} & \boxed{\begin{matrix} a_{1,m} \\ a_{2,m} \\ a_{3,m} \\ \cdots \\ a_{n,m} \end{matrix}} \end{pmatrix} = \begin{pmatrix} \boxed{f(v_1)} & \boxed{f(v_2)} & \boxed{\cdots} & \boxed{f(v_m)} \end{pmatrix}$$

Cela nous permet d'énoncer la règle pour la construction de la matrice A associée à f :

LES COLONNES DE LA MATRICE SONT LES IMAGES DES VECTEURS DE BASE

Lorsqu'on calcule $f(v)$ à l'aide de la matrice A et de la notation vectorielle de v, on voit une façon géométrique pour calculer Av. Il s'agit de la multiplication matricielle entre une matrice et un vecteur, elle s'effectue ligne par colonne.

$$Av = \begin{pmatrix} \boxed{a_{1,1} \quad a_{1,2} \quad \cdots \quad a_{1,m}} \\ \boxed{a_{2,1} \quad a_{2,2} \quad \cdots \quad a_{2,m}} \\ \boxed{a_{3,1} \quad a_{3,2} \quad \cdots \quad a_{3,m}} \\ \boxed{\cdots} \\ \boxed{a_{n,1} \quad a_{n,2} \quad \cdots \quad a_{n,m}} \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_m \end{matrix}} \end{pmatrix} = \begin{pmatrix} \boxed{a_{1,1}\alpha_1 + a_{1,2}\alpha_2 + \cdots + a_{1,m}\alpha_m} \\ \boxed{a_{2,1}\alpha_1 + a_{2,2}\alpha_2 + \cdots + a_{2,m}\alpha_m} \\ \boxed{a_{3,1}\alpha_1 + a_{3,2}\alpha_2 + \cdots + a_{3,m}\alpha_m} \\ \boxed{\cdots} \\ \boxed{a_{n,1}\alpha_1 + a_{n,2}\alpha_2 + \cdots + a_{n,m}\alpha_m} \end{pmatrix}$$

Addition de transformations linéaires

Soit V et W deux espaces vectoriels de dimension finie respective m et n. On considère $S = \{v_1, v_2, \ldots, v_m\}$ une base de V et $T = \{w_1, w_2, w_3, \ldots, w_n\}$ une base de W. Soit $f, g : V \to W$ deux applications linéaires dont les matrices respectives sont A et B.

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ a_{3,1} & a_{3,2} & \cdots & a_{3,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,m} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,m} \\ b_{3,1} & b_{3,2} & \cdots & b_{3,m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,m} \end{pmatrix}$$

La matrice associée à la transformation $f+g$ est donnée par l'*addition des matrices* A et B (on additionne les matrices entrée par entrée) :

$$A + B = \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \cdots & a_{1,m} + b_{1,m} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \cdots & a_{2,m} + b_{2,m} \\ a_{3,1} + b_{3,1} & a_{3,2} + b_{3,2} & \cdots & a_{3,m} + b_{3,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} + b_{n,1} & a_{n,2} + b_{n,2} & \cdots & a_{n,m} + b_{n,m} \end{pmatrix}$$

Preuve

En effet, comme les colonnes de la matrice sont les images des vecteurs de base, il suffit de calculer l'image de v_i par l'application $f+g$ pour $i \in \{1, 2, \ldots, m\}$.

$$(f+g)(v_i) = f(v_i) + g(v_i) \overset{\text{notation}}{=} \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ a_{3,i} \\ \cdots \\ a_{n,i} \end{pmatrix} + \begin{pmatrix} b_{1,i} \\ b_{2,i} \\ b_{3,i} \\ \cdots \\ b_{n,i} \end{pmatrix} = \begin{pmatrix} a_{1,i} + b_{1,i} \\ a_{2,i} + b_{2,i} \\ a_{3,i} + b_{3,i} \\ \cdots \\ a_{n,i} + b_{n,i} \end{pmatrix}$$

On reconnaît ainsi les colonnes de $A + B$. $\square$

Composition de transformations linéaires

Soit U, V et W trois espaces vectoriels de dimension respective k, m et n.

Soit $R = \{u_1, \ldots, u_k\}$ une base de U, $S = \{v_1, v_2, \ldots, v_m\}$ une base de V et encore $T = \{w_1, w_2, w_3, \ldots, w_n\}$ une base de W.

Soit $f : V \to W$ et $g : U \to V$ deux applications linéaires. Il est important que l'espace vectoriel de départ de f soit le même que celui d'arrivée de g pour pouvoir parler de $f \circ g$.

Notons A et B les matrices associées respectivement à f et à g.

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ a_{3,1} & a_{3,2} & \cdots & a_{3,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,k} \\ b_{2,1} & \cdots & b_{2,k} \\ \cdots & \cdots & \cdots \\ b_{m,1} & \cdots & b_{m,k} \end{pmatrix}$$

La matrice associée à la transformation $f \circ g : U \to W$ est donnée par la *multiplication des matrices* A et B. Cette multiplication s'effectue ligne par colonne.

$$\begin{aligned} AB &= \begin{pmatrix} \boxed{a_{1,1} \quad a_{1,2} \quad \cdots \quad a_{1,m}} \\ \boxed{a_{2,1} \quad a_{2,2} \quad \cdots \quad a_{2,m}} \\ \boxed{a_{3,1} \quad a_{3,2} \quad \cdots \quad a_{3,m}} \\ \boxed{\cdots \quad \cdots \quad \cdots \quad \cdots} \\ \boxed{a_{n,1} \quad a_{n,2} \quad \cdots \quad a_{n,m}} \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} b_{1,1} \\ b_{2,1} \\ \cdots \\ b_{m,1} \end{matrix}} & \boxed{\begin{matrix} \cdots \\ \cdots \\ \cdots \\ \cdots \end{matrix}} & \boxed{\begin{matrix} b_{1,k} \\ b_{2,k} \\ \cdots \\ b_{m,k} \end{matrix}} \end{pmatrix} \\ &= \begin{pmatrix} \boxed{a_{1,1}b_{1,1} + a_{1,2}b_{2,1} + \cdots + a_{1,m}b_{m,1}} & \boxed{\cdots} & \boxed{a_{1,1}b_{1,k} + a_{1,2}b_{2,k} + \cdots + a_{1,m}b_{m,k}} \\ \boxed{a_{2,1}b_{1,1} + a_{2,2}b_{2,1} + \cdots + a_{2,m}b_{m,1}} & \boxed{\cdots} & \boxed{a_{2,1}b_{1,k} + a_{2,2}b_{2,k} + \cdots + a_{2,m}b_{m,k}} \\ \boxed{a_{3,1}b_{1,1} + a_{3,2}b_{2,1} + \cdots + a_{3,m}b_{m,1}} & \boxed{\cdots} & \boxed{a_{3,1}b_{1,k} + a_{3,2}b_{2,k} + \cdots + a_{3,m}b_{m,k}} \\ \boxed{\cdots} & \boxed{\cdots} & \boxed{\cdots} \\ \boxed{a_{n,1}b_{1,1} + a_{n,2}b_{2,1} + \cdots + a_{n,m}b_{m,1}} & \boxed{\cdots} & \boxed{a_{n,1}b_{1,k} + a_{n,2}b_{2,k} + \cdots + a_{n,m}b_{m,k}} \end{pmatrix} \end{aligned}$$

Preuve

En effet, comme les colonnes de la matrice sont les images des vecteurs de base, il suffit de calculer l'image de u_i par l'application $f \circ g$ pour $i \in \{1, \ldots, k\}$.

$$(f \circ g)(u_i) \overset{\text{définition}}{=} f(g(u_i)) \overset{\text{notation}}{=} A(Bu_i)$$

Grâce à la règle de construction des matrices, on constate que Bu_i est la i-ième colonne de B. Cela permet de continuer le calcul, puisqu'on sait multiplier une matrice et un vecteur.

$$A(Bu_i) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ a_{3,1} & a_{3,2} & \cdots & a_{3,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix} \begin{pmatrix} b_{1,i} \\ b_{2,i} \\ \cdots \\ b_{m,i} \end{pmatrix} = \begin{pmatrix} a_{1,1}b_{1,i} + a_{1,2}b_{2,i} + \cdots + a_{1,m}b_{m,i} \\ a_{2,1}b_{1,i} + a_{2,2}b_{2,i} + \cdots + a_{2,m}b_{m,i} \\ a_{3,1}b_{1,i} + a_{3,2}b_{2,i} + \cdots + a_{3,m}b_{m,i} \\ \cdots \\ a_{n,1}b_{1,i} + a_{n,2}b_{2,i} + \cdots + a_{n,m}b_{m,i} \end{pmatrix}$$

On reconnaît ainsi les colonnes de AB. □

Important. En général, les matrices AB et BA ne sont pas les mêmes !

Exemples de matrices

On utilise la règle pour la construction de la matrice A associée à l'application linéaire f désirée.

LES COLONNES DE LA MATRICE SONT LES IMAGES DES VECTEURS DE BASE

Lorsque rien n'est précisé, on prend la base canonique de l'espace vectoriel de départ et la base canonique de l'espace d'arrivée.

Matrice identité et matrice nulle

La matrice identité I est associée à l'applications linéaires $\mathrm{id} : \mathbb{R}^n \to \mathbb{R}^n;\ v \mapsto v$ et la matrice nulle $\mathbb{O}$ est associée à l'application linéaire $0 : \mathbb{R}^n \to \mathbb{R}^n;\ v \mapsto 0$.

$$I = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \qquad \mathbb{O} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

Ces deux matrices sont les mêmes quelque soit la base choisie pour $\mathbb{R}^n$ (néanmoins il faut utiliser la même base pour l'espace de départ que celle pour l'espace d'arrivée).

Matrice de rotation d'un quart de tour

Voici la matrice $R_{\frac{\pi}{2}}$ associée à la rotation d'angle $\frac{\pi}{2}$ dans le plan.

$$R_{\frac{\pi}{2}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Matrice associée à une homothétie de facteur 2

Voici la matrice H associée à une homothétie de facteur 2 dans l'espace.

$$H = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Matrices associées à la dérivée et à l'intégrale (de polynômes)

Voici les matrices associées à l'application dérivée de $\mathcal{P}_n(\mathbb{R})$ dans $\mathcal{P}_{n-1}(\mathbb{R})$ et à l'intégrale de $\mathcal{P}_{n-1}(\mathbb{R})$ dans $\mathcal{P}_n(\mathbb{R})$ dans les bases canoniques de ces espaces de polynômes.

$$D = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n-1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & n \end{pmatrix} \qquad I = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{n} \end{pmatrix}$$

La matrice D est de taille $n \times (n+1)$ et la matrice I est de taille $(n+1) \times n$.

6.3 Le noyau et l'image

6.3.1 Sous-espaces d'un espace vectoriel

Soit V un espace vectoriel et W un sous-ensemble non vide de V. On dit que W est un *sous-espace vectoriel* de V si les éléments de W sont stables par l'opération interne et l'opération externe de V. Autrement dit si

$$\begin{aligned} &1) \quad w_1 + w_2 \in W \text{ pour tout } w_1, w_2 \in W. \\ &2) \quad \lambda w \in W \text{ pour tout } \lambda \in \mathbb{R}, w \in W. \end{aligned}$$

En remplaçant λ par 0 dans la deuxième propriété, on montre qu'un sous-espace vectoriel contient toujours le vecteur 0.

Exemples

1. Considère une droite d dans l'espace $\mathbb{R}^n$ $(n \geqslant 2)$.
   La droite d est un sous-espace de $\mathbb{R}^n$ si et seulement si d passe par l'origine 0.
2. Considère un plan π dans l'espace $\mathbb{R}^3$.
   Le plan π est un sous-espace de $\mathbb{R}^3$ si et seulement si π passe par l'origine 0.
3. Voici une inclusion de sous-espaces vectoriels de $\mathcal{C}(\mathbb{R})$.

$$\mathcal{P}_0(\mathbb{R}) \subset \mathcal{P}_1(\mathbb{R}) \subset \mathcal{P}_2(\mathbb{R}) \subset \cdots \subset \mathcal{P}_n(\mathbb{R}) \subset \cdots \subset \mathcal{P}(\mathbb{R}) \subset \mathcal{C}(\mathbb{R})$$

6.3.2 Le noyau d'une application linéaire

Soit V et W deux espaces vectoriels et $f : V \to W$ une application linéaire.
Le *noyau*[2] de f, noté $\mathrm{Ker}(f)$, est un sous-espace vectoriel de V défini comme suit :

$$\mathrm{Ker}(f) = \{v \in V : f(v) = 0\}$$

Lorsque A est une matrice associée à l'application f, on pose

$$\mathrm{Ker}(A) = \{v \in V : Av = 0\}$$

Avec le langage des fonctions en première année, le noyau de f est l'ensemble des zéros de f.

6.3.3 L'image d'une application linéaire

Soit V et W deux espaces vectoriels et $f : V \to W$ une application linéaire.
L'*image* de f, noté $\mathrm{Im}(f)$, est un sous-espace vectoriel de W défini comme suit :

$$\mathrm{Im}(f) = \{w \in W : \text{ il existe } v \in V \text{ tel que } f(v) = w\}$$

Lorsque A est une matrice associée à l'application f, on pose

$$\mathrm{Im}(A) = \{w \in W : \text{ il existe } v \in V \text{ tel que } Av = w\}$$

Avec le langage des fonctions en première année, l'image de f est le domaine image de f.

2. La terminologie *noyau* provient de la théorie des groupes. $\mathrm{Ker}(f)$ est l'ensemble des éléments qui sont envoyés sur le neutre du groupe d'arrivée (correspondant ici au zéro de W).

6.3.4 Théorème de la base incomplète

Soit V_k un sous-espace vectoriel de dimension k d'un espace vectoriel V_n de dimension n. Soit $U_k = \{u_1, \dots, u_k\}$ une base de V_k et $S = \{v_1, v_2, \dots, v_n\}$ une base de V_n.

Alors, il existe une base de V_n qui contient le système U_k.

Preuve

On va donner un procédé permettant de construire une telle base V_n. Pour cela, on va ajouter $n-k$ vecteurs bien choisis au système U_k pour obtenir un système U_n qui sera une base de V_n (et ainsi il contiendra automatiquement le système U_k).

Ajoutons un vecteur à U_k pour obtenir un système U_{k+1} qui est une base d'un espace vectoriel de dimension $k+1$, noté V_{k+1}. Si[3] $k < n$ (autrement dit si V_k est strictement contenu dans V_n), il existe un vecteur parmi la base S qui ne s'écrit pas comme combinaison linéaire des vecteurs de U_k. On ajoute ce vecteur au système U_k pour obtenir le système U_{k+1}.

En continuant ainsi de suite, on construit un système U_n qui sera une base d'un espace vectoriel de dimension n qui sera forcément V_n car c'est le seul sous-espace vectoriel de dimension n dans V_n. □

6.3.5 Théorème de la dimension

Soit V et W deux espaces vectoriels de dimension finie respective m et n et $f : V \to W$ une application linéaire. Alors :

$$\dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) = \dim(V)$$

Preuve

On sait que le noyau est un sous-espace de V de dimension k avec $0 \leqslant k \leqslant m$. Prenons une base du noyau de f. Cette base aura k éléments, disons $S = \{v_1, \dots, v_k\}$.
Cette base peut être complétée en une base de V, notée $T = \{\underbrace{v_1, \dots, v_k}_{\text{base de Ker}(f)}, \underbrace{w_1, \dots, w_{m-k}}_{\text{vecteurs ajoutés}}\}$.

On va démontrer que les images des vecteurs ajoutés $\{f(w_1), \dots, f(w_{m-k})\}$ forment une base de $\mathrm{Im}(f)$ et ainsi que la dimension de $\mathrm{Im}(f)$ sera $m-k$.

1. Les vecteurs $f(w_1)$, ..., $f(w_{m-k})$ engendrent $\mathrm{Im}(f)$.

   Soit $w \in \mathrm{Im}(f)$, alors il existe $v \in V$ tel que $f(v) = w$. Le vecteur v s'écrit comme combinaison (unique) des vecteurs de la base T, donc :

   $$v = \alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 w_1 + \cdots + \beta_{m-k} w_{m-k}$$

   Ainsi, on a $w = f(v) = f(\alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 w_1 + \cdots + \beta_{m-k} w_{m-k})$
   $\overset{f \text{ linéaire}}{=} \alpha_1 f(v_1) + \cdots + \alpha_k f(v_k) + \beta_1 f(w_1) + \cdots + \beta_{m-k} f(w_{m-k})$

   Or, comme les vecteurs v_1, ..., v_k sont dans $\mathrm{Ker}(f)$, on a $f(v_1) = \cdots = f(v_k) = 0$ et donc :

   $$w = \beta_1 f(w_1) + \cdots + \beta_{m-k} f(w_{m-k})$$

   On a ainsi montré que $\{f(w_1), \dots, f(w_{m-k})\}$ est un système générateur de $\mathrm{Im}(f)$.

3. Dans le cas où $k = n$, on aurait rien à faire puisque U_k serait la base U_n cherchée.

2. Les vecteurs $f(w_1), \ldots, f(w_{m-k})$ sont linéairement indépendants.

   Soit $\beta_1, \ldots, \beta_{m-k}$ des nombres réels tels que

$$\beta_1 f(w_1) + \cdots + \beta_{m-k} f(w_{m-k}) = 0$$

   On doit montrer que $\beta_1 = \cdots = \beta_{m-k} = 0$.

   On utilise la linéarité de f pour dire que :

$$\beta_1 f(w_1) + \cdots + \beta_{m-k} f(w_{m-k}) = f(\beta_1 w_1 + \cdots + \beta_{m-k} w_{m-k}) = 0$$

   Ainsi, $\beta_1 w_1 + \cdots + \beta_{m-k} w_{m-k} \in \mathrm{Ker}(f)$, donc s'écrit dans la base du $\mathrm{Ker}(f)$. Ainsi :

$$\beta_1 w_1 + \cdots + \beta_{m-k} w_{m-k} = \alpha_1 v_1 + \cdots + \alpha_k v_k$$

   En passant tout du même côté, on obtient :

$$-\alpha_1 v_1 - \cdots - \alpha_k v_k + \beta_1 w_1 + \cdots + \beta_{m-k} w_{m-k} = 0$$

   Comme les vecteurs de T forment une base de V, on obtient

$$-\alpha_1 = \cdots = -\alpha_k = \beta_1 = \cdots = \beta_{m-k} = 0$$

On a ainsi démontré que $\mathrm{Im}(f)$ est un espace de dimension $m - k$. En se rappelant que $\mathrm{Ker}(f)$ est de dimension k, on a

$$\dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) = k + m - k = m = \dim(V)$$

□

6.3.6 Premier théorème sur les applications linéaires injectives

Soit V et W deux espaces vectoriels (même pas forcément de dimension finie). Soit $f : V \to W$ une application linéaire. On a

$$f \text{ est injective} \iff \mathrm{Ker}(f) = \{0\}$$

Preuve

Rappelons qu'une fonction $f : V \to W$ est injective si pour tout $x, y \in V$ tels que $f(x) = f(y)$, on a $x = y$ (deux points ayant la même image sont forcément les mêmes ou de manière équivalente, deux points différents ont des images différentes (c'est la contraposée)).

"$\Rightarrow$" Si la fonction est injective, on doit montrer que $\mathrm{Ker}(f) = \{0\}$. Autrement dit, il faut vérifier que le seul zéro de f est 0. Pour cela, on prend un zéro de f, disons $v \in V$ tel que $f(v) = 0$ et on montre que $v = 0$.

Comme f est linéaire, on a $f(0) = 0$, ainsi on a $f(v) = f(0)$. Puisque f est injective, on a $v = 0$.

"$\Leftarrow$" Si $\mathrm{Ker}(f) = \{0\}$, on doit montrer que la fonction est injective. Ainsi, on prend x et y dans V tels que $f(x) = f(y)$. Cette équation est équivalente à $f(x) - f(y) = 0$. Comme f est linéaire, on a $f(x) - f(y) = f(x - y) = 0$.

Par conséquent, $x - y$ est un zéro de f (autrement dit $x - y \in \mathrm{Ker}(f)$). Comme le seul zéro est 0 (car $\mathrm{Ker}(f) = \{0\}$), on a $x - y = 0$, c'est-à-dire $x = y$. □

6.3.7 Deuxième théorème sur les applications linéaires injectives

Soit V et W deux espaces vectoriels de dimension finie n (les dimensions de V et de W sont les mêmes). Soit $f : V \to W$ une application linéaire. Sont équivalents.

$i)$ f est injective. $ii)$ f est bijective. $iii)$ f est surjective.

Preuve

On va montrer $i) \Rightarrow ii) \Rightarrow iii) \Rightarrow i)$.

"$i) \Rightarrow ii)$" Si la fonction est injective, alors $\mathrm{Ker}(f) = \{0\}$. Or, on a :

$$\dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) = \dim(V)$$

Puisque $\mathrm{Ker}(f) = \{0\}$, on a $\dim(\mathrm{Ker}(f)) = 0$ et ainsi $\dim(\mathrm{Im}(f)) = \dim(V) = n$. Donc f est surjective puisque $\mathrm{Im}(f) = W$ (car le seul sous-espace de dimension n de W est W). Comme la fonction est injective et surjective, elle est bijective.

"$ii) \Rightarrow iii)$" Évident, puisqu'une fonction bijective est surjective.

"$iii) \Rightarrow i)$" Si la fonction est surjective, alors $\mathrm{Im}(f) = W$ et donc $\dim(\mathrm{Im}(f)) = n = \dim(V)$. Or, on a :

$$\dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) = \dim(V)$$

On en déduit aisément que $\dim(\mathrm{Ker}(f)) = 0$ et donc que $\mathrm{Ker}(f) = \{0\}$ (c'est le seul sous-espace de dimension 0). Par conséquent, f est injective.

$\square$

6.4 Systèmes d'équations et opérations de lignes

En mathématiques, on peut être confronté à résoudre un système de n équations linéaires à m inconnues $x_1, x_2, \ldots, x_m$.

$$(S) : \left\{ \begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,m}x_m & = & y_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,m}x_m & = & y_2 \\ a_{3,1}x_1 & + & a_{3,2}x_2 & + & \cdots & + & a_{3,m}x_m & = & y_3 \\ & & & & \cdots & & & & \\ a_{n,1}x_1 & + & a_{n,2}x_2 & + & \cdots & + & a_{n,m}x_m & = & y_n \end{array} \right.$$

Pour résoudre un tel système, on a vu deux méthodes : la substitution ou la combinaison (on multiplie et on additionne).

En algèbre linéaire, on privilégie toujours la méthode utilisant des combinaisons (appelées ici opérations de lignes). On distingue trois types d'opérations de lignes.

- multiplication d'une ligne par un nombre non nul.
- permutation de deux lignes.
- addition d'un multiple d'une ligne à une autre.

À chaque étape, on n'effectue qu'une seule opération de lignes (cette opération sera évidemment réversible, ainsi un système obtenu après une telle opération sera forcément équivalent au système précédent). De plus, on conserve toujours les lignes que l'on n'a pas touchées.

Exemple de résolution d'un système d'équations par opération de lignes

Voici un problème de géométrie spatiale : on cherche la droite d'intersection des plans d'équations cartésiennes $3x + 4y + 5z = 6$ et $6x + 9y + 3z = 3$.

La droite est l'ensemble des points $(x; y; z)$ qui satisfont les deux équations cartésiennes. Pour trouver x, y et z, il faut donc résoudre le système suivant.

$$\left\{\begin{array}{rcrcrcr} 3x & + & 4y & + & 5z & = & 6 \\ 6x & + & 9y & + & 3z & = & 3 \end{array}\right.$$

Effectuons des opérations de lignes afin de rendre ce système triangulaire !

$$\left\{\begin{array}{rcrcrcr} 3x & + & 4y & + & 5z & = & 6 \\ 6x & + & 9y & + & 3z & = & 3 \end{array}\right. \overset{l_2-2l_1}{\Longleftrightarrow} \left\{\begin{array}{rcrcrcr} 3x & + & 4y & + & 5z & = & 6 \\ & & y & - & 7z & = & -9 \end{array}\right.$$

$$\overset{l_1-4l_2}{\Longleftrightarrow} \left\{\begin{array}{rcrcrcr} 3x & & & + & 33z & = & 42 \\ & & y & - & 7z & = & -9 \end{array}\right. \overset{l_1:3}{\Longleftrightarrow} \left\{\begin{array}{rcrcrcr} x & & & + & 11z & = & 14 \\ & & y & - & 7z & = & -9 \end{array}\right.$$

Le terme triangulaire se comprend maintenant de manière évidente. On remarque que les variables x et y (ce sont les *variables principales*) sont déterminées par z (c'est la *variable secondaire* ou le *paramètre*). En posant $z = \lambda$, le système de départ est donc équivalent à une représentation paramétrique de la droite cherchée.

$$\left\{\begin{array}{rcrcr} x & = & 14 & - & 11\lambda \\ y & = & -9 & + & 7\lambda \\ z & = & & & \lambda \end{array}\right. , \qquad \lambda \in \mathbb{R}$$

6.4.1 Opérations de lignes sur les matrices

Le système (S) de n équations linéaires à m inconnues $x_1, x_2, \ldots, x_m$ de la page précédente peut s'écrire à l'aide la notation matricielle de la façon suivante.

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ a_{3,1} & a_{3,2} & \cdots & a_{3,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_m \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \cdots \\ y_n \end{pmatrix} \quad \Longleftrightarrow \quad Ax = y$$

Or, les opérations de lignes modifient les lignes de A et de y. Ainsi, lorsqu'on résout un tel système par opérations de lignes, on travaille sur la matrice étendue suivante, sauf si y est le vecteur nul auquel cas on peut se passer d'étendre la matrice.

$$\left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,m} & y_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} & y_2 \\ a_{3,1} & a_{3,2} & \cdots & a_{3,m} & y_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} & y_n \end{array}\right)$$

Exemple de résolution d'un système d'équations par opération de lignes

Reprenons l'exemple précédent, mais cette fois, en utilisant la notation avec la matrice étendue. On cherchait la droite d'intersection des plans d'équations cartésiennes $3x + 4y + 5z = 6$ et $6x + 9y + 3z = 3$.

La droite est l'ensemble des points $(x; y; z)$ qui satisfont les deux équations cartésiennes. Pour trouver x, y et z, il faut donc résoudre le système suivant.

$$\left\{\begin{array}{rcrcrcr} 3x & + & 4y & + & 5z & = & 6 \\ 6x & + & 9y & + & 3z & = & 3 \end{array}\right.$$

On prend la matrice étendue associée à ce système et on effectue des opérations de lignes.

$$\left(\begin{array}{ccc|c} 3 & 4 & 5 & 6 \\ 6 & 9 & 3 & 3 \end{array}\right) \overset{l_2-2l_1}{\sim} \left(\begin{array}{ccc|c} 3 & 4 & 5 & 6 \\ 0 & 1 & -7 & -9 \end{array}\right) \overset{l_1-4l_2}{\sim} \left(\begin{array}{ccc|c} 3 & 0 & 33 & 42 \\ 0 & 1 & -7 & -9 \end{array}\right) \overset{l_2:3}{\sim} \left(\begin{array}{ccc|c} 1 & 0 & 11 & 14 \\ 0 & 1 & -7 & -9 \end{array}\right)$$

Ainsi, on a montré l'équivalence suivante.

$$\left\{\begin{array}{rcrcrcr} x & & & + & 11z & = & 14 \\ & & y & - & 7z & = & -9 \end{array}\right. \overset{z=\lambda}{\Longleftrightarrow} \left\{\begin{array}{rcrcr} x & = & 14 & - & 11\lambda \\ y & = & -9 & + & 7\lambda \\ z & = & & & \lambda \end{array}\right. , \quad \lambda \in \mathbb{R}$$

Exemple de résolution d'un système de 3 équations à 3 inconnues

Prenons à nouveau un problème de géométrie spatiale : cherchons l'intersection de trois plans donnés par leur équation cartésienne.

$$\left\{\begin{array}{rcrcrcr} & & 3y & - & 4z & = & -12 \\ 2x & + & 3y & + & 5z & = & -3 \\ x & + & 5y & - & z & = & 2 \end{array}\right.$$

On prend la matrice étendue associée à ce système et on effectue des opérations de lignes.

$$\left(\begin{array}{ccc|c} 0 & 3 & -4 & -12 \\ 2 & 3 & 5 & -3 \\ 1 & 5 & -1 & 2 \end{array}\right) \overset{l_1\leftrightarrow l_3}{\sim} \left(\begin{array}{ccc|c} 1 & 5 & -1 & 2 \\ 2 & 3 & 5 & -3 \\ 0 & 3 & -4 & -12 \end{array}\right) \overset{l_2-2l_1}{\sim} \left(\begin{array}{ccc|c} 1 & 5 & -1 & 2 \\ 0 & -7 & 7 & -7 \\ 0 & 3 & -4 & -12 \end{array}\right)$$

$$\overset{l_2:(-7)}{\sim} \left(\begin{array}{ccc|c} 1 & 5 & -1 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & 3 & -4 & -12 \end{array}\right) \overset{l_3-3l_2}{\sim} \left(\begin{array}{ccc|c} 1 & 5 & -1 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -1 & -15 \end{array}\right) \overset{l_3:(-1)}{\sim} \left(\begin{array}{ccc|c} 1 & 5 & -1 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 15 \end{array}\right)$$

$$\overset{l_2+l_3}{\sim} \left(\begin{array}{ccc|c} 1 & 5 & -1 & 2 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 15 \end{array}\right) \overset{l_1+l_3}{\sim} \left(\begin{array}{ccc|c} 1 & 5 & 0 & 17 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 15 \end{array}\right) \overset{l_1-5l_2}{\sim} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -63 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 15 \end{array}\right)$$

Ainsi, on arrive à placer la matrice identité I là où au départ on avait A. Puis en revenant au système d'équations, on obtient directement la solution !

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & -63 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 15 \end{array}\right) \Longleftrightarrow \left\{\begin{array}{rcr} x & = & -63 \\ y & = & 16 \\ z & = & 15 \end{array}\right.$$

Ces trois plans se coupent tous en un seul point $I(-63; 16; 15)$.

6.5 Inversion des applications linéaires bijectives

Soit V un espace vectoriel de dimension finie n. Considérons une application linéaire $f : V \to V$ bijective dont la matrice associée est notée A.

Comme f est bijective, elle admet une fonction réciproque. On peut montrer sans trop de difficulté que cette fonction réciproque, notée f^{-1}, est linéaire. On a :

$$f(x) = y \iff x = f^{-1}(y)$$

Avec la notation matricielle cela donne :

$$\boxed{Ax = y \iff x = A^{-1}y}(\bigstar)$$

Cela signifie que si on effectue des opérations de lignes sur la matrice étendue $(A|y)$, on ne trouverait qu'un seul x. L'exemple précédent nous permet d'affirmer que ce n'est possible que si on peut transformer A en la matrice identité I par des opérations de lignes.

$$(A|y) \sim \cdots \sim (I|x)$$

Or, les opérations de lignes peuvent être remplacées par des multiplications matricielles à gauche (on va décrire précisément ces matrices dans les exercices). Notons E_i la matrice correspondante à la i-ième opération de lignes. Disons qu'après n opérations de lignes, on a transformé la matrice A en la matrice identité I. On a ainsi

$$I = E_nE_{n-1}\cdots E_2E_1A \iff I = \big(E_nE_{n-1}\cdots E_2E_1\big)A$$

Ainsi, $\big(E_nE_{n-1}\cdots E_2E_1\big)$ est une matrice qui, multipliée par A, donne l'identité. On a donc

$$\begin{aligned} Ax = y &\iff \big(E_nE_{n-1}\cdots E_2E_1\big)Ax = \big(E_nE_{n-1}\cdots E_2E_1\big)y \\ &\iff Ix = \big(E_nE_{n-1}\cdots E_2E_1\big)y \\ &\iff x = \big(E_nE_{n-1}\cdots E_2E_1\big)y \end{aligned}$$

En comparant avec l'équivalence $(\bigstar)$, cela signifie que $A^{-1} = E_nE_{n-1}\cdots E_2E_1$ est la matrice associée à l'application f^{-1}.

Les propriétés des fonctions réciproques nous permettent d'affirmer que

$$A^{-1}A = I \quad \text{et} \quad AA^{-1} = I$$

En effet, il suffit de contempler le tableau suivant.

notation avec les fonctions	notation avec les matrices	relation matricielle
$(f^{-1} \circ f)(x) = x$	$A^{-1}Ax = x$	$A^{-1}A = I$
$(f \circ f^{-1})(x) = x$	$AA^{-1}x = x$	$AA^{-1} = I$

Première méthode pour calculer l'inverse d'une matrice

On vient de voir que $A^{-1} = E_n E_{n-1} \cdots E_2 E_1$ où E_i est la matrice correspondante à la i-ième opération de lignes qui permet de transformer la matrice A en la matrice I.

De plus, on a évidemment l'équivalence suivante.

$$A^{-1} = E_n E_{n-1} \cdots E_2 E_1 \iff A^{-1} = (E_n E_{n-1} \cdots E_2 E_1) I$$

Cela signifie que si on effectue exactement les mêmes opérations à la matrice I que celles effectuées à la matrice A, alors on transforme la matrice I en la matrice A^{-1}.

Par conséquent, si on effectue des opérations de lignes sur la matrice étendue $(A|I)$ afin de placer la matrice identité à gauche $(I|A^{-1})$, alors à droite on va transformer la matrice identité en la matrice A^{-1}. On peut ainsi éviter d'écrire et de multiplier les matrices d'opérations de lignes.

Exemple

Calculons l'inverse de la matrice de l'exemple précédent, qui est

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

En effectuant exactement les mêmes opérations de lignes, on obtient

$$\left(\begin{array}{rrr|rrr} 0 & 3 & -4 & 1 & 0 & 0 \\ 2 & 3 & 5 & 0 & 1 & 0 \\ 1 & 5 & -1 & 0 & 0 & 1 \end{array}\right) \sim \cdots \sim \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 4 & \frac{17}{7} & -\frac{27}{7} \\ 0 & 1 & 0 & -1 & -\frac{4}{7} & \frac{8}{7} \\ 0 & 0 & 1 & -1 & -\frac{3}{7} & \frac{6}{7} \end{array}\right)$$

Ainsi

$$A^{-1} = \frac{1}{7}\begin{pmatrix} 28 & 17 & -27 \\ -7 & -4 & 8 \\ -7 & -3 & 6 \end{pmatrix}$$

Application

À l'aide de la matrice A^{-1}, on peut aussi résoudre le système de l'exemple précédent.

$$\left\{\begin{array}{rcrcrcr} & & 3y & - & 4z & = & -12 \\ 2x & + & 3y & + & 5z & = & -3 \\ x & + & 5y & - & z & = & 2 \end{array}\right. \iff \begin{pmatrix} 0 & 3 & -4 \\ 2 & 3 & 5 \\ 1 & 5 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -12 \\ -3 \\ 2 \end{pmatrix}$$

$$\iff A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -12 \\ -3 \\ 2 \end{pmatrix} \iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1}\begin{pmatrix} -12 \\ -3 \\ 2 \end{pmatrix}$$

$$\iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 28 & 17 & -27 \\ -7 & -4 & 8 \\ -7 & -3 & 6 \end{pmatrix}\begin{pmatrix} -12 \\ -3 \\ 2 \end{pmatrix} = \begin{pmatrix} -63 \\ 16 \\ 15 \end{pmatrix} \iff \left\{\begin{array}{rcr} x & = & -63 \\ y & = & 16 \\ z & = & 15 \end{array}\right.$$

Cette méthode nous fait gagner du temps si on connaît déjà l'inverse de A ou si on doit résoudre plusieurs fois un tel système en changeant uniquement le vecteur de droite.

6.6 Changement de base

Parfois la règle de construction des matrices nous donne envie de définir une matrice dans une autre base que la base canonique. Par exemple, il est plus facile de donner la matrice correspondant à une rotation autour d'un axe dans l'espace par rapport à la base formée d'un vecteur directeur de l'axe et de deux vecteurs perpendiculaires à cet axe bien choisis, que de décrire directement la matrice dans la base canonique de $\mathbb{R}^3$.

C'est dans ce but que l'on va développer une méthode générale qui permet de décrire la matrice dans une base bien choisie pour ensuite la décrire dans la base canonique.

Soit V un espace vectoriel de dimension finie n pour lequel on considère la base canonique $C = \{e_1, e_2, \ldots, e_n\}$ et une (autre) base $T = \{w_1, w_2, \ldots, w_n\}$.

En général, on connaît l'expression des vecteurs w_i dans la base canonique. Disons

$$w_i = a_{1,i}e_1 + a_{2,i}e_2 + \cdots + a_{n,i}e_n \overset{\text{notation}}{=} \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ \cdots \\ a_{n,i} \end{pmatrix}$$

Or, dans la base T les vecteurs $w_1, w_2, \ldots, w_n$ s'écrivent comme suit.

$$w_1 = \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}_T, w_2 = \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}_T, \ldots, w_n = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}_T$$

La notation sous forme de vecteurs est pratique, mais il ne faut pas oublier que l'écriture d'un vecteur dépend de la base dans laquelle on travaille (voir la section 6.2.3).
Ainsi, on peut décrire l'application qui passe des vecteurs de V exprimés dans la base T aux vecteurs de V exprimés dans la base canonique comme suit.

$$\begin{array}{ccccccccccc}
w_1 & \mapsto & w_1 & & w_2 & \mapsto & w_2 & \cdots & w_n & \mapsto & w_n \\
\| & & \| & & \| & & \| & & \| & & \| \\
\begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}_T & \mapsto & \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \cdots \\ a_{n,1} \end{pmatrix} & & \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}_T & \mapsto & \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \cdots \\ a_{n,2} \end{pmatrix} & \cdots & \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}_T & \mapsto & \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \cdots \\ a_{n,n} \end{pmatrix}
\end{array}$$

Puisque, les colonnes de la matrice sont les images des vecteurs de base, on a la matrice de changement de base

$$P = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

La matrice de changement de base qui va de $(V;C)$ dans $(V;T)$ est la matrice inverse, notée P^{-1}. Ainsi, si on connaît la matrice d'une application f dans la base T, notée A_T, alors on peut trouver la matrice de f dans la base canonique C, notée A, en composant les matrices de changement de base avec la matrice A_T selon le schéma suivant.

$$\begin{tikzcd}
(V;T) \arrow[r, "A_T"] & (V;T) \arrow[d, "P"] \\
(V;C) \arrow[u, "P^{-1}"] \arrow[r] \arrow[ur, phantom, "\boxed{A = PA_TP^{-1}}"] & (V;C)
\end{tikzcd}$$

Exemple

Dans l'espace vectoriel $\mathbb{R}^3$, on considère la rotation d'angle π autour de la droite passant par l'origine (c'est obligé, sinon l'application ne serait pas linéaire) et dont le vecteur directeur est le suivant.

$$v = \begin{pmatrix} 3 \\ 0 \\ -4 \end{pmatrix}$$

Afin de trouver la matrice associée à une telle application linéaire, on va travailler avec une base de $\mathbb{R}^3$ pour laquelle la rotation peut aisément se décrire. On va prendre v comme vecteur de base (puisqu'il est immobile par la rotation). Comme $\mathbb{R}^3$ est un espace vectoriel de dimension 3, il nous faut encore trouver deux vecteurs pour former une base.

Puisqu'on sait donner la matrice de rotation dans le plan, on va construire deux vecteurs orthogonaux de même longueur [4] qui sont tous les deux perpendiculaires à v.

Pour cela, prenons n'importe quel vecteur perpendiculaire à v. Par exemple

$$v = \begin{pmatrix} 3 \\ 0 \\ -4 \end{pmatrix} \quad \perp \quad \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} \parallel \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = v_2$$

Il nous reste à trouver un vecteur perpendiculaire à v et à v_2. On utilise un produit vectoriel.

$$v_3 = v \wedge v_2 = \begin{pmatrix} 3 \\ 0 \\ -4 \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix} = v_3$$

Il faut maintenant faire en sorte que ces deux vecteurs aient la même longueur[4]. Puisque v_3 est de longueur 5 et que v_2 est de longueur 1, multiplions la longueur de v_2 par 5. Ainsi la base dans laquelle on va décrire l'application est

$$T = \left\{ w_1 = \begin{pmatrix} 3 \\ 0 \\ -4 \end{pmatrix}, w_2 = \begin{pmatrix} 0 \\ 5 \\ 0 \end{pmatrix}, w_3 = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix} \right\}$$

Grâce au slogan «les colonnes de la matrice sont les images des vecteurs de base», on trouve la matrice A_T. Par la rotation, le vecteur w_1 est fixe, le vecteur w_2 est envoyé sur son opposé, tout comme le vecteur w_3.

$$w_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}_T \mapsto \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}_T \qquad w_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}_T \mapsto \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}_T \qquad w_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}_T \mapsto \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}_T$$

Quant à la matrice de changement de base P, il suffit de regarder la base T. Pour P^{-1}, on calcule l'inverse de P. Ainsi, on a :

$$P = \begin{pmatrix} 3 & 0 & 4 \\ 0 & 5 & 0 \\ -4 & 0 & 3 \end{pmatrix} \quad A_T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad P^{-1} = \frac{1}{25} \begin{pmatrix} 3 & 0 & -4 \\ 0 & 5 & 0 \\ 4 & 0 & 3 \end{pmatrix}$$

On trouve la matrice exprimée dans la base canonique en calculant le produit matriciel suivant.

$$A = PA_TP^{-1} = \frac{1}{25} \begin{pmatrix} -7 & 0 & -24 \\ 0 & -25 & 0 \\ -24 & 0 & 7 \end{pmatrix}$$

4. Ce n'est pas nécessaire si l'angle de rotation est de π. Mais c'est capital pour un angle qui n'est pas un multiple de π comme, par exemple, $\frac{\pi}{4}$.

6.7 Les propriétés du déterminant

En première année, on a vu comment calculer un déterminant de n vecteurs de $\mathbb{R}^n$. On sait ainsi calculer le déterminant d'une matrice carrée.

Le déterminant a une fabuleuse propriété que l'on ne peut pas démontrer dans le cadre d'un tel cours (puisqu'on n'a pas vu de définition formelle de ce qu'est un déterminant). Cette propriété est la suivante :

6.7.1 Propriété fondamentale du déterminant

Soit A et B deux matrices carrées de même taille. On a

$$\det(AB) = \det(A)\det(B)$$

6.7.2 Conséquence

Si A est une matrice carrée inversible (cela signifie que A est associée à une fonction bijective), alors on a

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

Preuve

En effet, l'inverse de A, noté A^{-1}, satisfait la propriété $AA^{-1} = I$. On a ainsi, grâce à la propriété fondamentale du déterminant

$$1 = \det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1})$$

D'où la formule annoncée. $\square$

6.7.3 Invariance par changement de base

Le déterminant est invariant par changement de base (c'est-à-dire que le déterminant d'une matrice est le même quelle que soit la base qu'on a utilisée pour décrire la matrice).

Preuve

On considère la matrice A_T d'une application $f : V \to V$ (où V est un espace vectoriel de dimension finie n) exprimée dans une base T (pour les vecteurs de base et d'arrivée) et la matrice A de la même application f exprimée dans la base canonique. Notons P la matrice de changement de base (qui va de la base T dans la base canonique). On a vu dans le chapitre sur les changements de base que

$$A = PA_TP^{-1}$$

Grâce à la propriété fondamentale du déterminant, on a

$$\begin{aligned} \det(A) &= \det(PA_TP^{-1}) = \det(P)\det(A_T)\det(P^{-1}) = \det(P)\det(P^{-1})\det(A_T) \\ &= \det(PP^{-1})\det(A_T) = \det(I)\det(A_T) = \det(A_T) \end{aligned}$$

$\square$

6.7.4 Théorème du déterminant

Soit A une matrice carrée. On a :

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

Preuve

"$\Rightarrow$" Grâce à la propriété fondamentale du déterminant et par le fait que A admet une matrice inverse A^{-1}, on a

$$\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(I) = 1$$

Par conséquent, on a $\det(A) \neq 0$.

"$\Leftarrow$" On démontre la contraposée : si A n'est pas inversible, alors $\det(A) = 0$.

La matrice A peut être associée à une application $f : V \to W$ où V et W sont des espaces vectoriels de même dimension finie n.

Comme A n'est pas inversible, l'application associée f n'est pas bijective. Par le deuxième théorème sur les applications linéaires injectives, f n'est pas injective. Par le premier théorème sur les applications linéaires injectives, $\mathrm{Ker}(f) \neq \{0\}$. Ainsi, $\mathrm{Ker}(f)$ est un espace vectoriel de dimension k avec $k \geqslant 1$. Donc il existe une base du noyau que l'on note $\{v_1, \ldots, v_k\}$. Par le théorème de la base incomplète, il existe une base de V dont les k premiers vecteurs sont les vecteurs de base du noyau. Notons $T = \{v_1, \ldots, v_k, w_1, \ldots, w_{n-k}\}$ cette base.

Dans la base T, la matrice correspondante à f, notée A_T, a ses k premières colonnes qui sont constituées de zéros (puisque les colonnes de la matrice sont les images des vecteurs de la base T et que les k premiers vecteurs de la base T vont sur 0). Ainsi $\det(A_T) = 0$ (car la première colonne de A_T est nulle).

Puisque le déterminant est invariant par changement de base, on a

$$\det(A) = \det(A_T) = 0$$

□

6.7.5 Formule d'inversion des applications linéaires bijectives

Voici une formule pour inverser les matrices qui utilise des déterminants (sans preuve) :

$$A^{-1} = \tfrac{1}{\det(A)}\,\widetilde{A}$$

La matrice $\widetilde{A}$ est construite de la façon suivante : le coefficient de la i-ième ligne et de la j-ième colonne de $\widetilde{A}$ est égal à $(-1)^{i+j}$ fois le déterminant de la matrice A à laquelle on enlève la j-ième ligne et la i-ième colonne.

Remarque

Cette méthode peut être relativement rapide pour des matrices inversibles de taille 2×2 ou de taille 3×3. Mais est fortement déconseillée pour des matrices plus grandes, car on trouvera l'inverse beaucoup plus rapidement à l'aide des opérations de lignes.

6.8 Valeurs propres et vecteurs propres

Certaines applications linéaires envoient des vecteurs sur un multiple de ces derniers.

Exemples

1. Une homothétie de facteur λ, envoie tout vecteur v sur λv.
2. La symétrie centrale du plan centrée en 0 envoie tout vecteur v sur son opposé $-v$.
3. La rotation du plan centrée en 0 d'angle α (qui n'est pas un multiple de π) dans le plan n'envoie aucun vecteur non nul sur un multiple de lui-même.
4. Une rotation dans l'espace fixe les vecteurs qui sont sur l'axe de rotation (ils sont envoyés sur une fois eux-mêmes).
5. Une projection orthogonale sur une droite d passant par 0 fixe les vecteurs de la droite d et envoie les vecteurs perpendiculaires à d sur 0 fois eux-mêmes.

Définitions

Soit V un espace vectoriel de dimension finie. Soit f une application linéaire de V dans V dont la matrice associée s'appelle A.

1. On dit que $\lambda \in \mathbb{R}$ est une *valeur propre* de f, s'il existe $v \in V$, $v \neq 0$ tel que
$$f(v) = \lambda v$$
On dit que le vecteur v est un *vecteur propre associé à la valeur propre* λ.
2. On note $\sigma_{\mathbb{R}}(A)$ l'ensemble des valeurs propres réelles de l'application f à laquelle la matrice A est associée. Cet ensemble est appelé le *spectre réel de* A. On a
$$\sigma_{\mathbb{R}}(A) = \{\lambda \in \mathbb{R} : \text{ il existe } v \in V \setminus \{0\} \text{ tel que } Av = \lambda v\} = \{\text{valeurs propres de } A\}$$
3. Soit $\lambda \in \mathbb{R}$. L'espace vectoriel $\{v \in E : Av = \lambda v\}$ est noté E_λ.

   Si λ est une valeur propre de f, alors $E_\lambda \setminus \{0\}$ est l'ensemble des vecteurs propres associés à la valeur propre λ et E_λ est appelé *l'espace propre associé à* λ.

   Si λ n'est pas une valeur propre de f, alors $E_\lambda = \{0\}$.

   En résumé, on a
$$E_\lambda = \{0\} \cup \{\text{vecteurs propres de } f \text{ associés à } \lambda\}$$
4. Le polynôme $p_A(\lambda) = \det(A - \lambda I)$ est appelé *polynôme caractéristique* de la matrice A. Si V est de dimension n, alors $p_A(\lambda)$ est un polynôme de degré n.

Interprétations géométriques

Soit λ une valeur propre.

1. Si $\lambda > 0$. Alors l'application f se comporte comme une homothétie de facteur λ sur l'espace propre E_λ.
2. Si $\lambda = 0$. Alors tout vecteur de E_λ est envoyé sur 0 par l'application f.
3. Si $\lambda < 0$. Alors l'application f se comporte comme une homothétie de facteur λ, suivie d'une symétrie centrale centrée en 0 sur l'espace propre E_λ.

6.8.1 Méthode de calcul des valeurs et des vecteurs propres

Soit A la matrice associée à une application linéaire f. Soit $\lambda \in \mathbb{R}$. On a les équivalences suivantes.

$$\begin{aligned}
&\text{il existe } v \neq 0 \text{ tel que } f(v) = \lambda v \\
\iff \quad & \text{il existe } v \neq 0 \text{ tel que } Av = \lambda v \\
\iff \quad & \text{il existe } v \neq 0 \text{ tel que } Av - \lambda v = 0 \\
\iff \quad & \text{il existe } v \neq 0 \text{ tel que } (A - \lambda I)v = 0 \\
\iff \quad & \operatorname{Ker}(A - \lambda I) \neq \{0\} \ (\diamondsuit) \\
\overset{\text{théorèmes 1 et 2}}{\iff} \quad & A - \lambda I \text{ est une matrice non inversible} \\
\iff \quad & \det(A - \lambda I) = 0 \ (\heartsuit)
\end{aligned}$$

Ainsi, λ est une valeur propre si et seulement si λ satisfait $(\heartsuit)$. Autrement dit, les valeurs propres sont les zéros du polynôme caractéristique. Par conséquent

$$\boxed{\sigma_{\mathbb{R}}(A) = \{\lambda \in \mathbb{R} : \det(A - \lambda I) = 0\}}$$

Si λ est une valeur propre, alors $(\diamondsuit)$ montre que l'espace propre E_λ est égal à $\operatorname{Ker}(A-\lambda I)$. Autrement dit, l'espace propre associé à la valeur propre λ est en fait le noyau de la matrice $A - \lambda I$.

$$\boxed{E_\lambda = \operatorname{Ker}(A - \lambda I)}$$

Cela reste vrai même si λ n'est pas une valeur propre, car dans ce cas $E_\lambda = \{0\}$.

Remarques

1. Si V est de dimension n, l'application linéaire $f : V \to V$ admet au plus n valeurs propres. Ceci provient du fait qu'un polynôme de degré n admet au plus n zéros.
2. On a $E_0 = \operatorname{Ker}(A - 0I) = \operatorname{Ker}(A)$. Ainsi si A n'est pas inversible, alors 0 est une valeur propre de A et le noyau de A, $\operatorname{Ker}(A)$, est l'espace propre associé à 0.
3. L'ensemble des points fixes de l'application associée à la matrice A est l'espace propre associé à la valeur propre 1, qui est $E_1 = \{v \in V : Av = v\} = \operatorname{Ker}(A - I)$.

6.8.2 Invariance par changement de base

Le polynôme caractéristique et donc le spectre sont invariants par changement de base.

Preuve

Puisque le spectre est l'ensemble des zéros du polynôme caractéristique, il suffit de montrer que ce dernier est invariant par changement de base. Dans le contexte d'un changement de base, on a la relation suivante.

$$A = PA_TP^{-1}$$

Donc $p_A(\lambda) = p_{PA_TP^{-1}}(\lambda) = \det(PA_TP^{-1} - \lambda I) = \det(PA_TP^{-1} - P\lambda P^{-1})$

$$\begin{aligned}
&= \det(P(A_T - \lambda I)P^{-1}) = \det(P)\det(A_T - \lambda I)\det(P^{-1}) = \det(A_T - \lambda I) \\
&= p_{A_T}(\lambda) \qquad \square
\end{aligned}$$

6.8.3 Le théorème de Hamilton-Cayley (sans preuve)

Soit A une matrice carrée et $p_A(\lambda)$ son polynôme caractéristique. Alors, on a $p_A(A) = 0$.

Une autre formule pour inverser les matrices

Soit A une matrice carrée inversible. Par le théorème du déterminant, on a $\det(A) \neq 0$. Ainsi, $p_A(0) = \det(A - 0I) = \det(A) \neq 0$.

Donc $p_A(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_2\lambda^2 + a_1\lambda + a_0$ avec $a_0 \neq 0$. Par Hamilton-Cayley, on a

$$p_A(A) = a_nA^n + a_{n-1}A^{n-1} + \cdots + a_2A^2 + a_1A + a_0I = 0$$

Ainsi, en multipliant par A^{-1}, on a

$$a_nA^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I + a_0A^{-1} = 0$$

Par conséquent, comme $a_0 \neq 0$, on obtient

$$A^{-1} = \frac{1}{a_0}\left(-a_nA^{n-1} - a_{n-1}A^{n-2} - \cdots - a_2A - a_1I\right)$$

6.8.4 Diagonalisation

Définition

Soit V un espace vectoriel de dimension finie n. Une application linéaire $f : V \to V$ est dite *diagonalisable* s'il existe une base de V telle que la matrice de f écrite dans cette base soit une matrice diagonale.

Critère de diagonalisation

Une application $f : V \to V$ est diagonalisable si et seulement s'il existe une base de V formée de vecteurs propres.

Preuve

Si f est diagonalisable, il existe une base de V dans laquelle la matrice associée est diagonale. Comme les colonnes de la matrices sont les images des vecteurs de base, alors on constate que les vecteurs de base vont sur un multiple d'eux-mêmes. Par conséquent, chaque vecteur de base est un vecteur propre.

Pour la réciproque, on suppose qu'il existe une base de vecteurs propres. Dans cette base, la matrice associée à f est diagonale puisque les colonnes de la matrices sont les images des vecteurs de base et que ces derniers vont sur un multiple d'eux-mêmes. $\square$

Une utilité de la diagonalisation

Si A est une matrice diagonalisable, alors on peut calculer A^n pour tout $n \in \mathbb{N}$. De plus, si A est inversible, alors on peut calculer A^n pour tout $n \in \mathbb{Z}$.

En effet, par un changement de base, on a $A = PDP^{-1}$ où D est une matrice diagonale. Ainsi, on a

$$A^n = PD^nP^{-1} \quad \text{et si } A \text{ est inversible} \quad A^{-n} = PD^{-n}P^{-1} \quad \text{pour tout } n \in \mathbb{N}$$

6.9 Algèbre linéaire dans les espaces euclidiens

Un espace vectoriel V est dit *euclidien* s'il est de dimension finie et est muni d'un produit scalaire. On va se restreindre aux espaces euclidiens $\mathbb{R}^n$ muni du produit scalaire usuel.

Définition

La *transposée* d'une matrice A de taille $n \times m$, notée ${}^tA$, est la matrice de taille $m \times n$ obtenue en effectuant une symétrie par rapport à la diagonale de la matrice A. Ainsi les lignes de ${}^tA$ sont les colonnes de A et les colonnes de ${}^tA$ sont les lignes de A.

$$ {}^tA = {}^t\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ a_{3,1} & a_{3,2} & \cdots & a_{3,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{2,1} & a_{3,1} & \cdots & a_{n,1} \\ a_{1,2} & a_{2,2} & a_{3,2} & \cdots & a_{n,2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{1,m} & a_{2,m} & a_{3,m} & \cdots & a_{n,m} \end{pmatrix} $$

On peut aussi parler de la *transposée* d'un vecteur puisqu'un vecteur est une matrice qui ne possède qu'une colonne.

Le produit scalaire est un produit matriciel

Remarquons que le produit scalaire peut être écrit comme un produit matriciel

$$ \boxed{v_1 \bullet v_2 = {}^tv_1 v_2} $$

En effet, ${}^tv_1 v_2 = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = v_1 \bullet v_2.$

6.9.1 Propriétés de la transposée

Première propriété

Soit A et B deux matrices (non forcément carrées). Alors ${}^t(AB) = {}^tB\,{}^tA$.

Preuve

Soit A une matrice de taille $n \times m$ et B une matrice de taille $m \times k$ (afin que le produit matriciel soit réalisable). La matrice $C = AB$ est de taille $n \times k$.
Posons $A = (a_{i,j})$, $B = (b_{j,k})$, $C = (c_{i,k})$. On sait que le produit de deux matrices s'effectue ligne par colonne. Ainsi, on a

$$ c_{i,k} = \sum_{j=1}^{m} a_{i,j} b_{j,k} $$

Si $c_{i,k}$ est le coefficient de la i-ième de ligne et de la k-ième colonne de C, alors le coefficient de la k-ième ligne et de la i-ième colonne de ${}^tC$ est noté $\tilde{c}_{k,i}$ et on a $\tilde{c}_{k,i} = c_{i,k}$. En utilisant la même notation pour les coefficients de ${}^tA$ et ${}^tB$, on a

$$ c_{i,k} = \sum_{j=1}^{m} a_{i,j} b_{j,k} \iff \tilde{c}_{k,i} = \sum_{j=1}^{m} \tilde{a}_{j,i} \tilde{b}_{k,j} \iff \tilde{c}_{k,i} = \sum_{j=1}^{m} \tilde{b}_{k,j} \tilde{a}_{j,i} $$

Ainsi, on reconnaît le produit matriciel ${}^tC = {}^tB\,{}^tA$. $\square$

Deuxième propriété

Soit v et w deux vecteurs dans $\mathbb{R}^n$ et A une matrice de taille $n \times n$. Alors

$$Av \bullet w = v \bullet {}^tAw$$

Preuve

Par définition, on a $Av \bullet w = {}^t(Av)w$. On applique la première propriété et on trouve

$$Av \bullet w = {}^t(Av)w = ({}^tv\,{}^tA)w = {}^tv({}^tAw) = v \bullet {}^tAw$$

$\square$

Troisième propriété

Soit A une matrice carrée. Alors $\det(A) = \det({}^tA)$.

Preuve

Par opérations de lignes, on transforme A en une matrice triangulaire T (avec des zéros sur ou sous la diagonale). Les opérations de lignes peuvent s'exécuter par des multiplications matricielles. donc

$$T = E_n E_{n-1} \cdots E_2 E_1 A$$

Ainsi

$${}^tT = {}^tA\,{}^tE_1\,{}^tE_2 \cdots {}^tE_{n-1}\,{}^tE_n$$

En utilisant la propriété fondamentale du déterminant, on a

$$\det(T) = \det(E_n)\det(E_{n-1})\cdots\det(E_2)\det(E_1)\det(A)$$

$$\det({}^tT) = \det({}^tA)\det({}^tE_1)\det({}^tE_2)\cdots\det({}^tE_{n-1})\det({}^tE_n)$$

Comme les matrices T et ${}^tT$ sont triangulaires et ont les mêmes éléments sur la diagonale, leur déterminant sont égaux. Quant aux déterminants des matrices d'opérations de lignes, les opérations de lignes qui ne sont pas des permutations de lignes ont aussi leur matrice sous forme triangulaire et les permutations de lignes ont une matrice qui est égale à leur transposée (elles sont symétriques par rapport à la diagonale). Donc

$$\det({}^tT) = \det(T) \quad \text{et} \quad \det({}^tE_i) = \det(E_i)$$

Par conséquent, on a

$$\begin{array}{ccl} \det(T) & = & \det(E_n)\det(E_{n-1})\cdots\det(E_2)\det(E_1)\det(A) \\ \parallel & & \\ \det(T) & = & \det({}^tA)\det(E_1)\det(E_2)\cdots\det(E_{n-1})\det(E_n) \end{array}$$

Donc, en simplifiant, on a $\det({}^tA) = \det(A)$. On peut simplifier car les matrices d'opérations de lignes sont toutes inversibles et ainsi leur déterminant est non nul. $\square$

Conséquence

On peut donc aussi développer les déterminants selon la première ligne (qui est la première colonne de la transposée).

6.9.2 Bases orthonormées

Soit E un espace euclidien de dimension n (penser à $\mathbb{R}^n$ et à son produit scalaire) et $S = \{e_1, e_2, \ldots, e_n\}$ une base de E. On dit que S est une *base orthonormée* si

$$e_i \bullet e_j = \delta_{i,j}$$

où $\delta_{i,j}$ est le *symbole de Kronecker* :

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

On définit la *norme* d'un vecteur v à l'aide du produit scalaire comme suit :

$$\|v\| = \sqrt{v \bullet v}$$

On définit aussi l'*angle* entre les deux vecteurs v et w par la formule suivante.

$$\cos(\angle(v, w)) = \frac{v \bullet w}{\|v\| \cdot \|w\|}$$

Ainsi une base est dite orthonormée si la norme de chaque vecteur est égale à 1 (car $e_i \bullet e_i = 1$) et si deux vecteurs différents sont perpendiculaires (car $e_i \bullet e_j = 0$).

6.9.3 Les isométries vectorielles

Définition

Soit E un espace euclidien de dimension finie n. Une application linéaire $f : E \to E$ qui conserve le produit scalaire est appelée une *isométrie vectorielle*. La conservation du produit scalaire signifie que

$$f(v) \bullet f(w) = v \bullet w \text{ pour tout } v, w \in E$$

Les isométries vectorielles préservent aussi les normes

Par définition, une isométrie vectorielle conserve les produits scalaires. Par conséquent, elle conserve aussi les normes. En effet

$$\|Av\| = \sqrt{Av \bullet Av} \overset{\text{iso.}}{=} \sqrt{v \bullet v} = \|v\|$$

Lemme

Si les vecteurs sont exprimés dans une base orthonormée et si $v \bullet Aw = v \bullet w$ pour tout v, $w \in E$, alors $A = I$.

Preuve

Notons e_i les n vecteurs de la base orthonormée. On a

$$\delta_{i,j} \overset{\text{H}}{=} e_i \bullet e_j \overset{\text{H}}{=} e_i \bullet (Ae_j) = e_i \bullet \left(\sum_{k=1}^{n} a_{k,j} e_k\right) = \sum_{k=1}^{n} a_{k,j} (e_i \bullet e_k) = \sum_{k=1}^{n} a_{k,j} \delta_{i,k} \overset{\text{H}}{=} a_{i,j}$$

Donc $a_{i,i} = 1$ et $a_{i,j} = 0$ si $i \neq j$. Ce qui montre bien que $A = I$. □

Théorème

Soit E un espace euclidien de dimension finie n et une application linéaire $f : E \to E$. Alors f est une isométrie vectorielle si et seulement si la matrice A associée à f écrite dans une base orthonormée de E satisfait les trois conditions équivalentes

$$i)\ \ {}^tAA = I \qquad ii)\ \ A^{-1} = {}^tA \qquad iii)\ \ A\,{}^tA = I$$

Une telle matrice est appelée *matrice orthogonale.*

Preuve

Si f est une isométrie vectorielle, alors en utilisant les propriétés de la transposée, on a $v \bullet w = Av \bullet Aw = v \bullet {}^tAAw$ pour tout v, $w \in E$. Donc, par le lemme (et le fait qu'on travaille dans une base orthonormée), on a ${}^tAA = I$. Cela montre que A est inversible et on a ainsi $A^{-1} = {}^tA$.

Pour la réciproque, puisque ${}^tAA = I$, on a $Av \bullet Aw = v \bullet {}^tAAw = v \bullet w$. Ainsi A est la matrice d'une isométrie vectorielle. $\square$

Proposition 1

Soit f une isométrie vectorielle de matrice A, alors $\det(A) = \pm 1$ et $\sigma_{\mathbb{R}}(A) \subset \{\pm 1\}$.

La dernière expression signifie qu'il ne peux pas y avoir d'autres valeurs propres que ± 1 ; ceci n'exclu pas que $\sigma_{\mathbb{R}}(A) = \emptyset$.

Preuve

1. Si A est la matrice associée à l'isométrie vectorielle f écrite dans une base orthonormée, on a $I = {}^tAA$. Par conséquent

$$1 = \det({}^tAA) = \det({}^tA)\det(A) = \big(\det(A)\big)^2 \iff \det(A) = \pm 1$$

2. Soit λ une valeur propre, alors il existe $v \neq 0$ tel que $Av = \lambda v$. Donc

$$\|v\| \overset{\text{iso.}}{=} \|Av\| = \|\lambda v\| = |\lambda| \cdot \|v\| \overset{v\neq 0}{\iff} |\lambda| = 1 \iff \lambda = \pm 1$$

$\square$

Proposition 2

Soit A une matrice orthogonale telle que $\sigma_{\mathbb{R}}(A) = \{\pm 1\}$. Soit $v_1 \in E_1$ et $v_{-1} \in E_{-1}$. Alors $v_1 \perp v_{-1}$.

Preuve

On a $v_1 \bullet v_{-1} \overset{\text{iso.}}{=} Av_1 \bullet Av_{-1} = v_1 \bullet (-v_{-1}) \iff v_1 \bullet v_{-1} = -(v_1 \bullet v_{-1})$

$$\iff 2(v_1 \bullet v_{-1}) = 0 \iff v_1 \bullet v_{-1} = 0 \iff v_1 \perp v_{-1} \qquad \square$$

Théorème de classification des isométries vectorielles du plan

Soit f une isométrie vectorielle du plan, alors :

1. Son déterminant vaut 1 si et seulement si f est une rotation du plan. De plus, il existe une base orthonormée dans laquelle la matrice de f est
$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$
2. Son déterminant vaut -1 si et seulement si f est une symétrie orthogonale. De plus, il existe une base orthonormée dans laquelle la matrice de f est
$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Preuve

Grâce à la proposition 1, on sait que soit $\det(A) = 1$, soit $\det(A) = -1$. Le théorème est donc bien un théorème de classification.

1. «$\Longrightarrow$»

   Si A est la matrice associée à l'application linéaire f écrite dans une base orthonormée, on a $A^{-1} = {}^tA$.

   Puisqu'on est en dimension 2, on peut poser $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

   On a ainsi
$${}^tA = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \qquad A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$
   La formule pour l'inverse nécessite le fait que $\det(A) = 1$. Le lecteur pourra la vérifier en la multipliant par A ou en appliquant la formule de la page 56.

   Comme $A^{-1} = {}^tA$, on peut affirmer que $a = d$ et $c = -b$. Ainsi
$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$
   Donc $\det(A) = a^2 + b^2 = 1$. Cela montre que le point $(a; b)$ est sur le cercle trigonométrique. Par conséquent, il existe un angle $\alpha \in [0, 2\pi[$ pour lequel on a $\cos(\alpha) = a$ et $\sin(\alpha) = b$. La matrice A s'écrit donc de la manière suivante.
$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$
   Il s'agit de la matrice de rotation α centrée en 0.

   «$\Longleftarrow$»

   Si f est une rotation du plan d'angle α (centrée en 0, car f est linéaire), alors, comme les colonnes de A sont les images des vecteurs de base, dans n'importe quelle base orthonormée, la matrice est
$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$
   On a donc $\det(A) = \cos^2(\alpha) + \sin^2(\alpha) = 1$.

2. «$\Longrightarrow$»

Si A est la matrice associée à l'application linéaire f écrite dans une base orthonormée, on a $A^{-1} = {}^tA$.

Puisqu'on est en dimension 2, on peut poser $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

On a ainsi

$$^tA = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \qquad A^{-1} = \begin{pmatrix} -d & b \\ c & -a \end{pmatrix}$$

La formule pour l'inverse nécessite le fait que $\det(A) = -1$. Le lecteur pourra la vérifier en la multipliant par A ou en appliquant la formule de la page 56.

Comme $A^{-1} = {}^tA$, on peut affirmer que $a = -d$ et $c = b$. Ainsi

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

Donc $\det(A) = -a^2 - b^2 = -1$. Par conséquent $a^2 + b^2 = 1$ ($\bigstar$).

Cherchons les valeurs propres en cherchant les zéros du polynômes caractéristique.

$$\begin{aligned} \det(A - \lambda I) = 0 &\iff \begin{vmatrix} a-\lambda & b \\ b & -a-\lambda \end{vmatrix} = 0 \\ &\iff -(a-\lambda)(a+\lambda) - b^2 = 0 \\ &\iff \lambda^2 - a^2 - b^2 = 0 \\ &\iff \lambda^2 = a^2 + b^2 \overset{(\bigstar)}{=} 1 \\ &\iff \lambda = \pm 1 \end{aligned}$$

On a donc deux valeurs propres $\lambda = 1$ et $\lambda = -1$. Prenons v_1 un vecteur propre de longueur 1 associé à la valeur propre $\lambda = 1$ et v_{-1} un vecteur propre de longueur 1 associé à la valeur propre $\lambda = -1$. Par la proposition 2, $\{v_{-1}, v_1\}$ est une base orthonormée. Dans cette base A s'écrit

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

«$\Longleftarrow$»

Si f est une symétrie orthogonale, on se place dans la base $\{v_1, v_2\}$ où v_1 est un vecteur de longueur 1 orthogonal à l'axe de symétrie et v_2 est un vecteur de longueur 1 parallèle à l'axe de symétrie, alors, comme les colonnes de A sont les images des vecteurs de base, dans cette base orthonormée, la matrice est

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

On a donc $\det(A) = -1$. $\square$

Théorème de classification des isométries vectorielles de l'espace

Soit f une isométrie vectorielle de l'espace, alors :

1. Son déterminant vaut 1 si et seulement si f est une rotation de l'espace autour d'un axe de rotation. De plus, il existe une base orthonormée dans laquelle la matrice de f est
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$
Dans ce cas, l'axe de rotation est égal à l'espace propre E_1 (associé à la valeur propre 1) si l'angle n'est pas un multiple *pair* de π. Si l'angle est un multiple pair de π, alors $E_1 = \mathbb{R}^3$ et f est en fait l'identité.

2. Son déterminant vaut -1 si et seulement si f est la composition d'une rotation autour d'un axe de rotation et d'une symétrie orthogonale dont le plan de symétrie est perpendiculaire à l'axe de la rotation. De plus, il existe une base orthonormée dans laquelle la matrice de f est
$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$
Dans ce cas, l'axe de rotation est égal à l'espace propre E_{-1} (associé à la valeur propre -1) si l'angle n'est pas un multiple *impair* de π. Si l'angle est un multiple impair de π, alors $E_{-1} = \mathbb{R}^3$ et f est en fait une symétrie centrale.

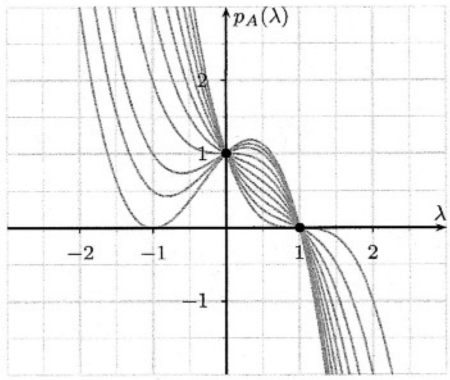
Preuve

Grâce à la proposition 1, on sait que soit $\det(A) = 1$, soit $\det(A) = -1$. Le théorème est donc bien un théorème de classification.

1. «$\Longrightarrow$»
Si A est la matrice associée à l'application linéaire f écrite dans une base orthonormée, on a $A^{-1} = {}^tA$.
Puisqu'on est en dimension 3, on peut poser $A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$.
Les valeurs propres sont les zéros du polynômes caractéristique. On a
$$p_A(\lambda) = \det(A - \lambda I) = 0 \iff \begin{vmatrix} a_{1,1} - \lambda & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} - \lambda & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} - \lambda \end{vmatrix} = 0$$
Sans faire de calculs, on remarque qu'il s'agit d'un polynôme de degré 3, de coefficient dominant -1. Disons
$$p_A(\lambda) = -\lambda^3 + p_2\lambda^2 + p_1\lambda + p_0$$
En remplaçant λ par 0, on trouve que $p_0 = p_A(0) = \det(A) = 1$.

Ainsi le polynôme caractéristique est

$$p_A(\lambda) = -\lambda^3 + p_2\lambda^2 + p_1\lambda + 1$$



On sait par la proposition 1, que ce polynôme ne peut s'annuler qu'en $\lambda = -1$ ou en $\lambda = 1$.

Or, 1) p_A est une fonction continue

2) $p_A(0) = 1$

3) $\lim_{\lambda \to +\infty} p_A(\lambda) = -\infty$

Donc $p_A(\lambda)$ doit s'annuler en $\lambda = 1$.

Ainsi, on sait que $\lambda = 1$ est une valeur propre. Soit v un vecteur propre de longueur 1 associé à cette valeur propre. On peut facilement construire un vecteur w_1 de longueur 1 orthogonal à v. Ainsi, en posant $w_2 = v \wedge w_1$, on obtient une base orthogonale $\{v, w_1, w_2\}$ de $\mathbb{R}^n$. Dans cette base, la matrice A s'écrit

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}$$

En effet, comme v est un vecteur propre associé à la valeur propre 1, son image est lui-même (ce qui explique la première colonne). Tandis que, comme f est une isométrie, les vecteurs w_1 et w_2 étant orthogonal à v, leur image est encore orthogonale à v (ce qui explique les zéros sur la première ligne).

Comme $\det(A) = 1$, la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ correspond à une isométrie du plan de déterminant 1.

On conclut en utilisant le théorème de classification des isométries du plan.

«$\Longleftarrow$»

Si f est une rotation de l'espace d'angle α autour d'un axe (qui passe par 0, car f est linéaire), on se place dans la base $\{v_1, v_2, v_3\}$ où v_1 est un vecteur de longueur 1 parallèle à l'axe de rotation, v_2 est un vecteur de longueur 1 perpendiculaire à v_1 et v_3 est donné par $v_3 = v_1 \wedge v_2$ afin d'être un vecteur de longueur 1 perpendiculaire à v_1 et v_2.

Ainsi, comme les colonnes de A sont les images des vecteurs de base, dans cette base orthonormée, la matrice est

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

On a donc $\det(A) = \cos^2(\alpha) + \sin^2(\alpha) = 1$.

2. «$\Longrightarrow$»

Si A est la matrice associée à l'application linéaire f écrite dans une base orthonormée, on a $A^{-1} = {}^tA$.

Puisqu'on est en dimension 3, on peut poser $A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$.

Les valeurs propres sont les zéros du polynômes caractéristique. On a

$$p_A(\lambda) = \det(A - \lambda I) = 0 \iff \begin{vmatrix} a_{1,1} - \lambda & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} - \lambda & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} - \lambda \end{vmatrix} = 0$$

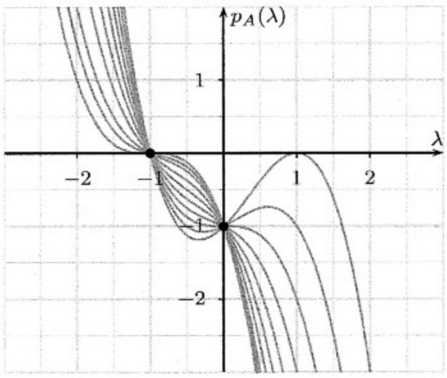
Sans faire de calculs, on remarque qu'il s'agit d'un polynôme de degré 3, de coefficient dominant -1. Disons

$$p_A(\lambda) = -\lambda^3 + p_2\lambda^2 + p_1\lambda + p_0$$

En remplaçant λ par 0, on trouve que $p_0 = p_A(0) = \det(A) = -1$.

Ainsi le polynôme caractéristique est

$$p_A(\lambda) = -\lambda^3 + p_2\lambda^2 + p_1\lambda - 1$$



On sait par la proposition 1, que ce polynôme ne peut s'annuler qu'en $\lambda = -1$ ou en $\lambda = 1$.

Or, 1) p_A est une fonction continue

2) $p_A(0) = -1$

3) $\lim\limits_{\lambda \to -\infty} p_A(\lambda) = +\infty$

Donc $p_A(\lambda)$ doit s'annuler en $\lambda = -1$.

Ainsi, on sait que $\lambda = -1$ est une valeur propre. Soit v un vecteur propre de longueur 1 associé à cette valeur propre. On peut facilement construire un vecteur w_1 de longueur 1 orthogonal à v. Ainsi, en posant $w_2 = v \wedge w_1$, on obtient une base orthogonale $\{v, w_1, w_2\}$ de $\mathbb{R}^n$. Dans cette base, la matrice A s'écrit

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}$$

En effet, comme v est un vecteur propre associé à la valeur propre -1, son image est $-v$ (ce qui explique la première colonne). Tandis que, comme f est une isométrie, les vecteurs w_1 et w_2 étant orthogonal à v, leur image est encore orthogonale à v (ce qui explique les zéros sur la première ligne).

Comme $\det(A) = -1$, la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ correspond à une isométrie du plan de déterminant 1.

On conclut en utilisant le théorème de classification des isométries du plan.

«⟸»

Si f est la composition d'une rotation autour d'un axe (qui passe par 0, car f est linéaire) et d'une symétrie orthogonale dont le plan de symétrie est perpendiculaire à l'axe de la rotation (et qui passe par 0, car f est linéaire), on se place dans la base $\{v_1, v_2, v_3\}$ où v_1 est un vecteur de longueur 1 parallèle à l'axe de rotation, v_2 est un vecteur de longueur 1 perpendiculaire à v_1 et v_3 est donné par $v_3 = v_1 \wedge v_2$ afin d'être un vecteur de longueur 1 perpendiculaire à v_1 et v_2

Ainsi, comme les colonnes de A sont les images des vecteurs de base, dans cette base orthonormée, la matrice est

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

On a donc $\det(A) = -\big(\cos^2(\alpha) + \sin^2(\alpha)\big) = -1$.

□

Proposition 3

Une matrice A de taille $n \times n$ est orthogonale si et seulement si ses colonnes sont orthonormées.

Preuve

On a

$$^tA = \begin{pmatrix} \boxed{\;{}^tv_1\;} \\ \boxed{\;{}^tv_2\;} \\ \boxed{\;\cdots\;} \\ \boxed{\;{}^tv_n\;} \end{pmatrix} \qquad \text{et} \qquad A = \begin{pmatrix} \boxed{v_1} & \boxed{v_2} & \boxed{\cdots} & \boxed{v_n} \end{pmatrix}$$

Ainsi

$$^tAA = \begin{pmatrix} v_1 \bullet v_1 & v_1 \bullet v_2 & \cdots & v_1 \bullet v_n \\ v_2 \bullet v_1 & v_2 \bullet v_2 & \cdots & v_2 \bullet v_n \\ \cdots & \cdots & \cdots & \cdots \\ v_n \bullet v_1 & v_n \bullet v_2 & \cdots & v_n \bullet v_n \end{pmatrix}$$

Par conséquent, on a

$$\begin{aligned} A \text{ est orthogonale} \quad &\Longleftrightarrow \quad {}^tAA = I \\ &\Longleftrightarrow \quad v_i \bullet v_j = \delta_{i,j} \quad \Longleftrightarrow \quad \text{les vecteurs } v_i \text{ sont orthonormés} \end{aligned}$$

□

6.10 Les axiomes d'un espace vectoriel

Un ensemble V est un *espace vectoriel réel*[5] s'il est muni d'une *opération interne* et d'une *opération externe* par des nombres réels.

L'opération interne, souvent notée $+$, satisfait les propriétés suivantes. En termes mathématiques, on dit que l'ensemble $(V, +)$ est un groupe commutatif ou groupe abélien.

1. L'addition est une opération interne, c'est une application qui transforme deux éléments de V en un élément de V.
$$\begin{array}{rrcl} + : & V \times V & \to & V \\ & (v_1, v_2) & \mapsto & v_1 + v_2 \end{array}$$
2. L'addition est associative. Cela signifie que pour tout v_1, v_2 et v_3 dans V, on a
$$(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$$
Cette règle signifie que l'on peut écrire une somme de trois vecteurs sans mettre de parenthèses.
3. Il existe un élément dans V qui joue le rôle du zéro, noté 0 et appelé l'*élément neutre* tel que pour tout $v \in V$ on a
$$0 + v = v$$
4. Pour chaque v dans V, il existe un élément de V, noté $-v$ et appelé l'*opposé* de v, tel que
$$v + (-v) = 0$$
5. L'addition est commutative. Cela signifie que pour tout v_1, v_2 dans V, on a
$$v_1 + v_2 = v_2 + v_1$$

L'opération externe, souvent notée $\cdot$, satisfait les propriétés suivantes. En termes mathématiques, il s'agit d'une action du groupe des nombres réels sur V.

1. La multiplication est une opération externe, c'est une application qui fait agir un nombre réel λ sur un élément v dans V pour fournir un élément de V.
$$\begin{array}{rrcl} \cdot : & \mathbb{R} \times V & \to & V \\ & (\lambda, v) & \mapsto & \lambda \cdot v \end{array}$$
2. Pour tout v dans V, on a
$$1 \cdot v = v$$
3. Pour tout λ_1, λ_2 dans $\mathbb{R}$ et tout v dans V, on a
$$(\lambda_1 \lambda_2) \cdot v = \lambda_1 \cdot (\lambda_2 \cdot v)$$
Cette règle signifie que l'on peut écrire $\lambda_1 \lambda_2 v$ sans parenthèses (et sans le symbole $\cdot$).

Il y a encore deux règles de compatibilité entre l'opération interne et externe. Ce sont les règles de mise en évidence.

$$(\lambda_1 + \lambda_2) \cdot v = \lambda_1 \cdot v + \lambda_2 \cdot v$$
$$\lambda \cdot (v_1 + v_2) = \lambda \cdot v_1 + \lambda \cdot v_2$$

5. On peut aussi définir des espaces vectoriels complexes en remplaçant $\mathbb{R}$ par $\mathbb{C}$ dans les règles.

Des règles ci-dessus, on peut en DÉDUIRE les trois règles suivantes

$$0 \cdot v = 0 \qquad (-1) \cdot v = -v \qquad \lambda \cdot 0 = 0$$

Montrons d'abord que pour chaque vecteur, il n'y a qu'un opposé possible (les règles disent a priori qu'il y a en au moins un).
Pour cela, supposons qu'on ait deux opposés d'un vecteur v donné, appelés v_1 et v_2. Il faut montrer que $v_1 = v_2$. Or on a

$$v_1 + v = 0 = v_2 + v$$

En faisant $-v$ de chaque côté, on trouve $v_1 = v_2$ (on a le droit car l'application $+$ est bien définie et tout vecteur possède un opposé).

Déduisons maintenant les trois règles énoncées ci-dessus.

1. Soit $v \in V$, on a

$$0 \cdot v + v = 0 \cdot v + 1 \cdot v = (0 + 1) \cdot v = 1 \cdot v = v$$

   En faisant $-v$ de chaque côté, on trouve $0 \cdot v = 0$ (on a le droit car l'application $+$ est bien définie et tout vecteur possède un opposé).

2. Soit $v \in V$, on a

$$(-1) \cdot v + v = (-1) \cdot v + 1 \cdot v = ((-1) + 1) \cdot v = 0 \cdot v = 0$$

   Donc, par unicité de l'opposé, on a $(-1) \cdot v = -v$ pour tout $v \in V$.

3. Soit $\lambda \in \mathbb{R}$ et $v \in V$, on a

$$0 = \lambda \cdot v - \lambda \cdot v = \lambda \cdot v + (-1)\lambda \cdot v = \lambda \cdot v + \lambda(-1) \cdot v = \lambda \cdot v + \lambda \cdot (-v) = \lambda \cdot (v - v) = \lambda \cdot 0$$

   Donc, on a $0 = \lambda \cdot 0$ pour tout $\lambda \in \mathbb{R}$.