

Chapitre 4

Matrices

Dans ce chapitre, p et n désignent deux entiers naturels non nuls et $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

I. Calcul matriciel

1. Opérations

Soient m et q des entiers naturels non nuls. Soient

$$A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in \mathcal{M}_{n,p}(\mathbb{K}), \quad B = (b_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq q}} \in \mathcal{M}_{m,q}(\mathbb{K})$$

et $\lambda \in \mathbb{K}$. On définit la matrice $\lambda \cdot A \in \mathcal{M}_{n,p}(\mathbb{K})$ (ou simplement λA) par :

$$\forall (i,j) \in \llbracket 1,n \rrbracket \times \llbracket 1,p \rrbracket, \quad (\lambda A)_{i,j} = \lambda a_{i,j}.$$

Si $n = m$ et $p = q$, on définit la matrice $A + B \in \mathcal{M}_{n,p}(\mathbb{K})$ par :

$$\forall (i,j) \in \llbracket 1,n \rrbracket \times \llbracket 1,p \rrbracket, \quad (A + B)_{i,j} = a_{i,j} + b_{i,j}.$$

Si $p = m$, on définit la matrice $AB \in \mathcal{M}_{n,q}(\mathbb{K})$ par :

$$\forall (i,j) \in \llbracket 1,n \rrbracket \times \llbracket 1,q \rrbracket, \quad (AB)_{i,j} = \sum_{k=1}^p a_{i,k} b_{k,j}.$$

Enfin, on définit la matrice transposée de A , notée tA ou A^T , par :

$${}^tA = (a_{j,i})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}} \in \mathcal{M}_{p,n}(\mathbb{K}).$$

On vérifie sans difficulté que, si $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$, alors

$${}^t(AB) = {}^tB {}^tA.$$

Notation – Si $A \in \mathcal{M}_{n,p}(\mathbb{K})$, les notations

$$A = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} \quad \text{et} \quad A = (C_1 \ \cdots \ C_p)$$

signifient que L_1, \dots, L_n sont les lignes de A , et que C_1, \dots, C_p sont les colonnes de A (dans cet ordre).

Remarques

• Si $A = (C_1 \ \cdots \ C_p) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $X = {}^t(x_1 \ \cdots \ x_p) \in \mathcal{M}_{p,1}(\mathbb{K})$, alors AX est une combinaison linéaire des colonnes de A :

$$AX = x_1 C_1 + \cdots + x_p C_p.$$

- Si $A = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (C_1 \ \cdots \ C_q) \in \mathcal{M}_{p,q}(\mathbb{K})$, alors :

- La j -ième colonne de AB est le produit AC_j de A par la j -ième colonne de B .
- La i -ième ligne de AB est le produit $L_i B$ de la i -ième ligne de A par B .

Attention !

- Le produit matriciel est associatif, mais non commutatif en général : si A et B sont deux éléments de $\mathcal{M}_n(\mathbb{K})$, on a en général $AB \neq BA$.
- Si $n \geq 2$, il existe des éléments tous deux non nuls A et B de $\mathcal{M}_n(\mathbb{K})$ tels que $AB = 0$.

Le résultat suivant est immédiat :

Propriété

- $(\mathcal{M}_{n,p}(\mathbb{K}), +, \cdot)$ est un \mathbb{K} -espace vectoriel.
- Pour tout $(i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$, on note $E_{i,j}$ la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ dont tous les coefficients sont nuls sauf celui en position (i, j) qui vaut 1. Alors la famille $(E_{i,j})_{1 \leq i \leq n, 1 \leq j \leq p}$ est une base de $\mathcal{M}_{n,p}(\mathbb{K})$, appelée base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$.
- $(\mathcal{M}_{n,p}(\mathbb{K}), +, \cdot)$ est de dimension finie égale à $n \times p$.

On montre également que la formule du binôme de Newton est valable pour deux matrices carrées de même taille **qui commutent**.

2. Polynômes de matrices

Si $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice **carrée**, on définit, de même qu'on l'a fait pour les endomorphismes, les polynômes de A , et les polynômes annulateurs de A .

Présentons une méthode très utile pour calculer les puissances d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$. Soit P un polynôme annulateur non nul de A . Pour $k \in \mathbb{N}$, effectuons la division euclidienne de X^k par P : il existe $Q_k \in \mathbb{K}[X]$ et $R_k \in \mathbb{K}[X]$ vérifiant $\deg(R_k) < \deg(P)$, tels que

$$X^k = P(X)Q_k(X) + R_k(X).$$

En évaluant cette relation en A , on obtient

$$A^k = P(A)Q_k(A) + R_k(A) = R_k(A),$$

car $P(A) = 0$ par définition. Le calcul de A^k se ramène donc à celui de R_k : il est d'autant plus simple que le degré de P est petit.

Par exemple, soit

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R}).$$

On vérifie facilement que $A^2 = A + 2I_3$, de sorte que le polynôme

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

est annulateur de A . Pour $k \in \mathbb{N}$, la division euclidienne de X^k par P peut s'écrire sous la forme

$$X^k = (X + 1)(X - 2)Q_k(X) + a_k X + b_k,$$

le reste R_k étant de degré au plus 1. On détermine a_k et b_k en évaluant la relation précédente en -1 et 2 (racines de P) :

$$\begin{cases} (-1)^k = -a_k + b_k \\ 2^k = 2a_k + b_k, \end{cases}$$

ce qui donne immédiatement

$$a_k = \frac{2^k + (-1)^{k+1}}{3}, \quad b_k = \frac{2^k + 2(-1)^k}{3}.$$

Finalement, on a montré que pour tout $k \in \mathbb{N}$,

$$A^k = \frac{2^k + (-1)^{k+1}}{3}A + \frac{2^k + 2(-1)^k}{3}I_3.$$

3. Matrices inversibles

Propriété/Définition

- Soit $A \in \mathcal{M}_n(\mathbb{K})$. On dit que A est **inversible** s'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = BA = I_n$.
Une telle matrice est alors unique, elle est notée A^{-1} et appelée **inverse** de A .
- L'ensemble des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ est noté $\mathcal{GL}_n(\mathbb{K})$, il est appelé **groupe linéaire** d'ordre n .

Démonstration de l'unicité

Si B et C vérifient les propriétés de la définition, alors

$$B = BI_n = B(AC) = (BA)C = I_n C = C. \quad \square$$

Propriété

Si A et B sont deux éléments de $\mathcal{GL}_n(\mathbb{K})$, alors $AB \in \mathcal{GL}_n(\mathbb{K})$ et

$$(AB)^{-1} = B^{-1}A^{-1}$$

Démonstration – On a

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_n A^{-1} = AA^{-1} = I_n,$$

et de même pour le produit $(B^{-1}A^{-1})(AB)$. D'où le résultat. \square

Remarque – Dans le chapitre précédent, on a donné, pour un endomorphisme u , une méthode pour prouver l'existence de u^{-1} à partir d'un polynôme annulateur de u de coefficient constant non nul. Cette méthode peut bien sûr être adaptée pour les matrices carrées.

II. Matrices, vecteurs et applications linéaires

Dans cette partie, sauf indication contraire, E et F désignent deux \mathbb{K} -espaces vectoriels de dimension finie non nulle. On note $p = \dim(E)$, $n = \dim(F)$, $\mathcal{B} = (e_1, \dots, e_p)$ une base de E et $\mathcal{C} = (f_1, \dots, f_n)$ une base de F . Enfin, u désigne une application linéaire entre E et F : $u \in \mathcal{L}(E, F)$.

1. Matrices d'une famille de vecteurs

Si x est un vecteur de F par exemple (ce qui suit s'adapte pour tout espace vectoriel de dimension finie), on peut décomposer x dans la base \mathcal{C} de F : on peut écrire $x = \sum_{i=1}^n a_i f_i$, où les a_i , éléments de \mathbb{K} , sont les coordonnées de x dans la base \mathcal{C} . Le vecteur

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

est appelé vecteur (ou matrice) colonne des coordonnées de x dans la base \mathcal{C} .

Plus généralement, si (x_1, \dots, x_k) est une famille de vecteurs de F , pour tout $j \in \llbracket 1, k \rrbracket$, on peut écrire

$$x_j = \sum_{i=1}^n a_{i,j} f_i,$$

où les $a_{i,j}$ sont des éléments de \mathbb{K} . Soit A la matrice

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,k} \\ \vdots & & \vdots & & \vdots \\ a_{i,1} & \dots & a_{i,j} & \dots & a_{i,k} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,j} & \dots & a_{n,k} \end{pmatrix} \in \mathcal{M}_{n,k}(\mathbb{K})$$

dont la j -ième colonne est, pour tout $j \in \llbracket 1, k \rrbracket$, le vecteur des coordonnées de x_j dans la base \mathcal{C} . La matrice A est appelée matrice de la famille (x_1, \dots, x_k) dans la base \mathcal{C} , notée $\text{Mat}_{\mathcal{C}}(x_1, \dots, x_k)$.

Exemple – Si $\mathcal{C} = (1, X, X^2)$ est la base canonique de $\mathbb{R}_2[X]$, la matrice de la famille

$$(2X^2 - X + 1, 3X^2 - 1)$$

dans la base \mathcal{C} est

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

2. Matrices d'une application linéaire

On sait que l'application linéaire u est entièrement déterminée par les p vecteurs $u(e_1), \dots, u(e_p)$, et donc, par leurs coordonnées dans la base \mathcal{C} . L'information concernant u est donc entièrement contenue dans la donnée de $n \times p$ scalaires.

Définition

La matrice de la famille $u(\mathcal{B}) = (u(e_1), \dots, u(e_p))$ dans la base \mathcal{C} , *i.e.*

$$\text{Mat}_{\mathcal{C}}(u(e_1), \dots, u(e_p)) \in \mathcal{M}_{n,p}(\mathbb{K})$$

est appelée **matrice de u dans les bases \mathcal{B} et \mathcal{C}** , et notée $\text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u)$.

Si $E = F$ et $\mathcal{B} = \mathcal{C}$, on note simplement $\text{Mat}_{\mathcal{B}}(u)$.

On retiendra notamment que pour tout $j \in \llbracket 1, p \rrbracket$, la j -ième colonne de $\text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u)$ est le vecteur des coordonnées de $u(e_j)$ dans la base \mathcal{C} : le fait que

$$\text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u) = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$$

est équivalent au fait que pour tout $j \in \llbracket 1, p \rrbracket$,

$$u(e_j) = \sum_{i=1}^n a_{i,j} f_i.$$

Exemple – Soit

$$\phi : \begin{cases} \mathbb{R}_3[X] & \rightarrow \mathbb{R}_2[X] \\ P & \mapsto P' \end{cases}$$

La matrice de ϕ dans les bases canoniques de $\mathbb{R}_3[X]$ et $\mathbb{R}_2[X]$ est

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

3. Isomorphisme entre $\mathcal{L}(E, F)$ et $\mathcal{M}_{n,p}(\mathbb{K})$

Ainsi, des bases \mathcal{B} de E et \mathcal{C} de F étant fixées, on peut associer à toute application linéaire $u \in \mathcal{L}(E, F)$ sa matrice dans les bases \mathcal{B} et \mathcal{C} , qui est un élément de $\mathcal{M}_{n,p}(\mathbb{K})$. On peut en fait en dire plus :

Théorème

L'application

$$\text{Mat}_{\mathcal{B}}^{\mathcal{C}} : \begin{cases} \mathcal{L}(E, F) & \rightarrow \mathcal{M}_{n,p}(\mathbb{K}) \\ u & \mapsto \text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u) \end{cases}$$

est un isomorphisme d'espaces vectoriels :

- Si $u, v \in \mathcal{L}(E, F)$ et $\lambda \in \mathbb{K}$,

$$\text{Mat}_{\mathcal{B}}^{\mathcal{C}}(\lambda u + v) = \lambda \text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u) + \text{Mat}_{\mathcal{B}}^{\mathcal{C}}(v).$$

- Pour tout $A \in \mathcal{M}_{n,p}(\mathbb{K})$, il existe une unique application linéaire $u \in \mathcal{L}(E, F)$ dont la matrice dans les bases \mathcal{B} et \mathcal{C} soit A .

Démonstration

- **Linéarité** : c'est immédiat par définition des opérations $+$ et \cdot sur les matrices.
- **Bijektivité** : Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$. On pose, pour tout $j \in \llbracket 1, p \rrbracket$, $y_j = \sum_{i=1}^n a_{i,j} f_i$. Une application linéaire $u \in \mathcal{L}(E, F)$ a pour matrice A dans les bases \mathcal{B} et \mathcal{C} si et seulement si pour tout $j \in \llbracket 1, p \rrbracket$, $u(e_j) = y_j$. Or, il existe une unique application linéaire u de E dans F satisfaisant ces conditions. Donc l'application $\text{Mat}_{\mathcal{B}}^{\mathcal{C}}$ est bijective. \square

Remarques

- Attention, on ne peut pas parler de « la » matrice de l'application linéaire u . Il est indispensable de préciser les bases au départ et à l'arrivée. Par exemple, l'identité de \mathbb{R}^n , dans les bases $2\mathcal{B}$ et \mathcal{B} (\mathcal{B} est la base canonique de \mathbb{R}^n) a pour matrice $2I_n$.
- L'unique endomorphisme u de E tel que $\text{Mat}_{\mathcal{B}}(u) = I_n$ est l'identité : $u = \text{Id}_E$ (remarquer la différence avec l'exemple précédent : ici on considère la même base au départ et à l'arrivée).
- L'unique endomorphisme u de E tel que $\text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u) = 0$ est l'application nulle.
- Cas particulier de $E = \mathcal{M}_{p,1}(\mathbb{K})$ et $F = \mathcal{M}_{n,1}(\mathbb{K})$. Pour k entier naturel non nul, l'espace $\mathcal{M}_{k,1}(\mathbb{K})$ admet pour base canonique la famille $\mathcal{B}_k = (V_1, \dots, V_k)$, où V_i est un vecteur-colonne à k composantes, toutes nulles sauf la i -ième qui vaut 1. D'après le théorème précédent, $\text{Mat}_{\mathcal{B}_p}^{\mathcal{B}_n}$ est un isomorphisme entre $\mathcal{L}(\mathcal{M}_{p,1}(\mathbb{K}), \mathcal{M}_{n,1}(\mathbb{K}))$ et $\mathcal{M}_{n,p}(\mathbb{K})$.

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. L'unique application linéaire de $\mathcal{M}_{p,1}(\mathbb{K})$ dans $\mathcal{M}_{n,1}(\mathbb{K})$ dont la matrice dans les bases \mathcal{B}_p et \mathcal{B}_n est A , est l'application linéaire canoniquement associée à A , *i.e.*

$$u_A : \begin{cases} \mathcal{M}_{p,1}(\mathbb{K}) & \rightarrow \mathcal{M}_{n,1}(\mathbb{K}) \\ X & \mapsto AX \end{cases}$$

- On fait souvent l'identification entre $\mathcal{M}_{n,1}(\mathbb{K})$ et \mathbb{K}^n (et de même, entre $\mathcal{M}_{p,1}(\mathbb{K})$ et \mathbb{K}^p) pour simplifier l'écriture. Il ne faut pas oublier ce que représente réellement chacun de ces espaces : $\mathcal{M}_{n,1}(\mathbb{K})$ est l'espace des vecteurs-colonnes à n coefficients, \mathbb{K}^n est l'espace des n -uplets d'éléments de \mathbb{K} (et donc, écrits en ligne, en séparant les composantes par des virgules).
- Les isomorphismes préservent la dimension, donc on retrouve le fait que

$$\dim(\mathcal{L}(E, F)) = p \times n = \dim(E) \times \dim(F).$$

4. Calcul de l'image d'un vecteur

Propriété

Soit $x \in E$ et $y = u(x)$. On note

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

les vecteurs colonne des coordonnées de x et y dans les bases \mathcal{B} et \mathcal{C} , respectivement. Soit $A = \text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u)$. Alors $Y = AX$.

Démonstration – Notons $A = (a_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$. On a $x = \sum_{j=1}^p x_j e_j$ et donc, par linéarité de u ,

$$u(x) = \sum_{j=1}^p x_j u(e_j) = \sum_{j=1}^p \left[x_j \left(\sum_{i=1}^n a_{i,j} f_i \right) \right] = \sum_{i=1}^n \left(\sum_{j=1}^p a_{i,j} x_j \right) f_i.$$

Par unicité des coordonnées dans la base \mathcal{C} , on en déduit

$$\forall i \in \llbracket 1, n \rrbracket, \quad y_i = \sum_{j=1}^p a_{i,j} x_j.$$

Par définition du produit matriciel, ces égalités signifient exactement que $Y = AX$. □

Remarque – Le produit matriciel a été défini pour que la propriété précédente soit vraie.

5. Lien entre produit de matrices et composition d'applications

Propriété

Soient E, E', E'' trois \mathbb{K} -espaces vectoriels de dimension finie. Soit \mathcal{B} une base de E , \mathcal{B}' une base de E' et \mathcal{B}'' une base de E'' . Soit $u \in \mathcal{L}(E, E')$ et $v \in \mathcal{L}(E', E'')$. On sait que

$$v \circ u : \begin{cases} E & \xrightarrow{u} & E' & \xrightarrow{v} & E'' \\ x & \mapsto & u(x) & \mapsto & v(u(x)) \end{cases}$$

appartient à $\mathcal{L}(E, E'')$. Alors

$$\text{Mat}_{\mathcal{B}}^{\mathcal{B}''}(v \circ u) = \text{Mat}_{\mathcal{B}'}^{\mathcal{B}''}(v) \text{Mat}_{\mathcal{B}}^{\mathcal{B}'}(u).$$

Démonstration – Notons $M = \text{Mat}_{\mathcal{B}}^{\mathcal{B}''}(v \circ u)$, $A = \text{Mat}_{\mathcal{B}'}^{\mathcal{B}''}(v)$ et $B = \text{Mat}_{\mathcal{B}}^{\mathcal{B}'}(u)$. Soit $x \in E$ et X le vecteur colonne des coordonnées de x dans la base \mathcal{B} . On sait que MX est le vecteur colonne des coordonnées de $(v \circ u)(x)$ dans la base \mathcal{B}'' . Or le vecteur colonne des coordonnées de $u(x)$ dans la base \mathcal{B}' est $Y = AX$ et le vecteur colonne des coordonnées de $v(u(x))$ dans la base \mathcal{B}'' est $BY = BAX$. Donc

$$\forall X \in \mathcal{M}_{p,1}(\mathbb{K}), \quad MX = BAX,$$

où $p = \dim(E)$. On en déduit que $M = BA$ en choisissant pour X les vecteurs de la base canonique de $\mathcal{M}_{p,1}(\mathbb{K})$. □

Corollaire

Soient E et F deux \mathbb{K} -espaces vectoriels de dimension n . Soit \mathcal{B} une base de E et \mathcal{C} une base de F . Soit $u \in \mathcal{L}(E, F)$. Alors on a l'équivalence :

$$u \text{ est un isomorphisme} \Leftrightarrow \text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u) \in \mathcal{GL}_n(\mathbb{K}).$$

Dans ce cas,

$$(\text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u))^{-1} = \text{Mat}_{\mathcal{C}}^{\mathcal{B}}(u^{-1}).$$

Cas particulier : si $E = F$ et $u \in \mathcal{L}(E)$, on a l'équivalence

$$u \text{ est un automorphisme} \Leftrightarrow \text{Mat}_{\mathcal{B}}(u) \in \mathcal{GL}_n(\mathbb{K}).$$

Dans ce cas,

$$(\text{Mat}_{\mathcal{B}}(u))^{-1} = \text{Mat}_{\mathcal{B}}(u^{-1}).$$

Démonstration

\Rightarrow Si u est un isomorphisme, alors il existe une application linéaire $v = u^{-1}$ telle que $u \circ v = \text{Id}_F$ et $v \circ u = \text{Id}_E$. Alors d'après la propriété précédente, $\text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u) \text{Mat}_{\mathcal{C}}^{\mathcal{B}}(v) = \text{Mat}_{\mathcal{C}}^{\mathcal{B}}(v) \text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u) = I_n$, donc $\text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u)$ est inversible, d'inverse $\text{Mat}_{\mathcal{C}}^{\mathcal{B}}(v)$.

\Leftarrow Soit $A = \text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u)$. Si A est inversible, alors il existe une matrice $B = A^{-1}$ telle que $AB = BA = I_n$. Soit v l'unique application linéaire de F dans E telle que $\text{Mat}_{\mathcal{C}}^{\mathcal{B}}(v) = B$. Alors d'après la propriété précédente, $\text{Mat}_{\mathcal{C}}(u \circ v) = \text{Mat}_{\mathcal{B}}(v \circ u) = I_n$, donc $u \circ v = \text{Id}_F$ et $v \circ u = \text{Id}_E$. \square

6. Changements de bases

Définition – Matrice de passage

Soient $\mathcal{B} = (e_1, \dots, e_p)$ et $\mathcal{B}' = (e'_1, \dots, e'_p)$ deux bases de E . Pour tout $j \in \llbracket 1, p \rrbracket$, on peut écrire $e'_j = \sum_{i=1}^p p_{i,j} e_i$, c'est-à-dire exprimer e'_j dans la base \mathcal{B} .

La matrice $P = (p_{i,j}) \in \mathcal{M}_p(\mathbb{K})$ est appelée **matrice de passage** de la base \mathcal{B} à la base \mathcal{B}' .

Les colonnes de P sont les coordonnées des vecteurs de la « nouvelle » base dans « l'ancienne ».

Remarque – On a $P = \text{Mat}_{\mathcal{B}}(e'_1, \dots, e'_p) = \text{Mat}_{\mathcal{B}'}^{\mathcal{B}}(\text{Id})$. En particulier, P est inversible et P^{-1} est la matrice de passage de \mathcal{B}' à \mathcal{B} .

Exemple – Les familles $\mathcal{B} = (1, X, X^2)$ et $\mathcal{B}' = (1 - X + X^2, X + 2X^2, 2 - X + 2X^2)$ sont des bases de $\mathbb{R}_2[X]$: la première est la base canonique, la seconde comporte $3 = \dim(\mathbb{R}_2[X])$ vecteurs et on vérifie très facilement qu'elle est libre. La matrice de passage de \mathcal{B} à \mathcal{B}' est

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

Propriété – Formule de changement de bases pour les vecteurs

Soient x un vecteur de E , X la matrice colonne des coordonnées de x dans la base \mathcal{B} , X' la matrice colonne des coordonnées de x dans la base \mathcal{B}' et P la matrice de passage de \mathcal{B} à \mathcal{B}' .

Alors on a la relation $X = PX'$, c'est-à-dire que l'on obtient les *anciennes coordonnées* en fonction des *nouvelles*.

Démonstration – D’après la propriété sur le calcul matriciel de l’image d’un vecteur par une application linéaire, et d’après la remarque ci-dessus,

$$X = \text{Mat}_{\mathcal{B}'}^{\mathcal{B}}(\text{Id}) X' = P X'. \quad \square$$

Exemple – Dans le cas de l’exemple ci-dessus, un polynôme $a + bX + cX^2$ se décompose dans la base \mathcal{B}' sous la forme

$$\alpha(1 - X + X^2) + \beta(X + 2X^2) + \gamma(2 - X + 2X^2)$$

où

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = P \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}, \quad \text{soit} \quad \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = P^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

À titre d’illustration, en identifiant le coefficient constant dans les deux décompositions, on obtient $a = \alpha + 2\gamma$, où l’on reconnaît le premier coefficient du produit

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = P \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}.$$

Propriété – Formule de changement de bases pour les applications linéaires

Soient \mathcal{B} et \mathcal{B}' deux bases de E , \mathcal{C} et \mathcal{C}' deux bases de F . Soient P la matrice de passage de \mathcal{B} à \mathcal{B}' dans E , Q la matrice de passage de \mathcal{C} à \mathcal{C}' dans F . Soient $u \in \mathcal{L}(E, F)$, $A = \text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u)$ et $B = \text{Mat}_{\mathcal{B}'}^{\mathcal{C}'}(u)$. Alors $B = Q^{-1}AP$.

Démonstration – Avec les notations précédentes pour $x \in E$ et des notations analogues pour $y = u(x)$, on a $X = PX'$, $Y = QY'$, $Y = AX$ et $Y' = BX'$, et donc

$$BX' = Y' = Q^{-1}Y = Q^{-1}AX = Q^{-1}APX'.$$

Ceci est vrai pour tout $X' \in \mathcal{M}_{p,1}(\mathbb{K})$ avec $p = \dim(E)$. On en déduit le résultat en choisissant pour X' les vecteurs de la base canonique de $\mathcal{M}_{p,1}(\mathbb{K})$. \square

Exemple – Soit

$$u : \begin{cases} \mathbb{R}_2[X] & \rightarrow \mathbb{R}_2[X] \\ a + bX + cX^2 & \mapsto (-a - 2b + c) + (3a + 3b - 2c)X - 2bX^2 \end{cases}$$

Il est immédiat que $u \in \mathcal{L}(\mathbb{R}_2[X])$. Écrivons la matrice de u dans la base canonique \mathcal{B} de $\mathbb{R}_2[X]$; on a

$$u(1) = -1 + 3X, \quad u(X) = -2 + 3X - 2X^2, \quad u(X^2) = 1 - 2X,$$

donc

$$A = \text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} -1 & -2 & 1 \\ 3 & 3 & -2 \\ 0 & -2 & 0 \end{pmatrix}.$$

De même, écrivons la matrice de u dans la base \mathcal{B}' de $\mathbb{R}_2[X]$ des deux exemples précédents; on a

$$\begin{aligned} u(1 - X + X^2) &= 2 - 2X + 2X^2 = 2(1 - X + X^2), \\ u(X + 2X^2) &= -X - 2X^2 = -(X + 2X^2), \\ u(2 - X + 2X^2) &= 2 - X + 2X^2, \end{aligned}$$

donc

$$B = \text{Mat}_{\mathcal{B}'}(u) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Les matrices A et B sont reliées par la formule de changement de base $B = P^{-1}AP$, soit $A = PBP^{-1}$. Du fait de la simplicité de la matrice B , cette relation facilite, par exemple, le calcul des puissances de A : pour tout $k \in \mathbb{N}$, $A^k = PB^kP^{-1}$ où $B^k = \text{diag}(2^k, (-1)^k, 1)$.

Définition – Matrices semblables

Si $(A, B) \in (\mathcal{M}_n(\mathbb{K}))^2$, on dit que A et B sont **semblables** si

$$\exists P \in \mathcal{GL}_n(\mathbb{K}); B = P^{-1}AP.$$

Propriété

Deux matrices de $\mathcal{M}_n(\mathbb{K})$ sont semblables si et seulement si elles représentent le même endomorphisme d'un espace de dimension n , quitte à faire *le même* changement de base au départ et à l'arrivée.

Remarque – La relation de similitude entre matrices définit une relation d'équivalence sur $\mathcal{M}_n(\mathbb{K})$.

III. Image, noyau et rang d'une matrice

1. Définitions, propriétés du rang

Toutes les définitions et propriétés des applications linéaires se transposent aux matrices $A \in \mathcal{M}_{n,p}(\mathbb{K})$ par l'intermédiaire de l'application linéaire canoniquement associée

$$u_A : \begin{cases} \mathcal{M}_{p,1}(\mathbb{K}) & \rightarrow \mathcal{M}_{n,1}(\mathbb{K}) \\ X & \mapsto AX \end{cases}$$

En particulier, pour $A \in \mathcal{M}_{n,p}(\mathbb{K})$, on définit :

- le noyau de A comme le noyau de u_A , *i.e.*

$$\text{Ker}(A) = \{X \in \mathcal{M}_{p,1}(\mathbb{K}); AX = 0\}.$$

- l'image de A comme l'image de u_A , *i.e.*

$$\text{Im}(A) = \{Y \in \mathcal{M}_{n,1}(\mathbb{K}); \exists X \in \mathcal{M}_{p,1}(\mathbb{K}); Y = AX\}.$$

- le rang de A comme le rang de u_A .

Remarques

- Déterminer $\text{Ker}(A)$ revient à résoudre le système linéaire sans second membre de matrice A .
- Si $A = (C_1 \cdots C_p)$ et $X = {}^t(x_1 \cdots x_p)$, alors $AX = x_1 C_1 + \cdots + x_p C_p$. En particulier, les colonnes de A forment une famille liée si et seulement s'il existe un vecteur non nul dans $\text{Ker}(A)$, et un tel vecteur donne explicitement une relation de dépendance linéaire entre les colonnes de A .
- $\text{Im}(A)$ est engendré par les images par l'application u_A des vecteurs de la base canonique de $\mathcal{M}_{p,1}(\mathbb{K})$, c'est-à-dire par les colonnes de A .
- En particulier, $\text{rg}(A)$ est le rang de la famille des vecteurs colonnes de A .

On sait que le rang d'une application linéaire n'est pas modifié par composition (à droite ou à gauche) par un isomorphisme. Matriciellement, ce résultat se traduit ainsi :

Propriété

Le rang d'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ n'est pas modifié par multiplication (à droite ou à gauche) par une matrice inversible : si $P \in \mathcal{GL}_n(\mathbb{K})$ et $Q \in \mathcal{GL}_p(\mathbb{K})$, alors

$$\text{rg}(PAQ) = \text{rg}(A).$$

Propriété

Soient $u \in \mathcal{L}(E, F)$ et $A = \text{Mat}_{\mathcal{B}}^{\mathcal{C}}(u)$.

Alors $\text{rg}(u) = \text{rg}(A)$, *i.e.*, pour calculer le rang de u , il suffit de calculer le rang d'une de ses matrices.

Démonstration – Par définition, en notant $A = (a_{i,j})$, on a

$$\text{rg}(u) = \dim(\text{Vect}(u(e_1), \dots, u(e_p))) = \dim\left(\text{Vect}\left(\sum_{i=1}^n a_{i,1} f_i, \dots, \sum_{i=1}^n a_{i,p} f_i\right)\right).$$

Notons $(V_i)_{1 \leq i \leq n}$ la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$ et ϕ l'isomorphisme de F sur $\mathcal{M}_{n,1}(\mathbb{K})$ tel que

$$\forall i \in \llbracket 1, n \rrbracket, \quad \phi(f_i) = V_i.$$

Ainsi, ϕ est l'application qui à tout vecteur de F associe la matrice colonne de ses coordonnées dans la base \mathcal{C} . Les isomorphismes préservant la dimension, on a

$$\text{rg}(u) = \dim\left(\text{Vect}\left(\phi\left(\sum_{i=1}^n a_{i,1} f_i\right), \dots, \phi\left(\sum_{i=1}^n a_{i,p} f_i\right)\right)\right).$$

Ainsi

$$\text{rg}(u) = \dim\left(\text{Vect}\left(\sum_{i=1}^n a_{i,1} V_i, \dots, \sum_{i=1}^n a_{i,p} V_i\right)\right),$$

où l'on reconnaît le rang de la famille des colonnes de A , et donc $\text{rg}(A)$. \square

Corollaire

Soit (x_1, \dots, x_m) une famille de vecteurs de E et $A = \text{Mat}_{\mathcal{B}}(x_1, \dots, x_m)$. Alors

$$\text{rg}(x_1, \dots, x_m) = \text{rg}(A),$$

i.e., pour calculer le rang d'une famille de vecteurs, il suffit de calculer le rang d'une de ses matrices.

Démonstration – Soit $\mathcal{C} = (e_j)_{1 \leq j \leq m}$ la base canonique de \mathbb{K}^m et u l'unique application linéaire de \mathbb{K}^m dans E telle que pour tout $j \in \llbracket 1, m \rrbracket$, $u(e_j) = x_j$. Alors $\text{Mat}_{\mathcal{C}}^{\mathcal{B}}(u) = \text{Mat}_{\mathcal{B}}(x_1, \dots, x_m) = A$ et $\text{Im}(u) = \text{Vect}(x_1, \dots, x_m)$ donc

$$\text{rg}(x_1, \dots, x_m) = \text{rg}(u) = \text{rg}(A),$$

la dernière égalité provenant de la propriété précédente. \square

2. Caractérisation des matrices inversibles, théorème du rang

Théorème – Caractérisation des matrices inversibles

Soit $A \in \mathcal{M}_n(\mathbb{K})$. On a les équivalences :

$$A \text{ est inversible} \Leftrightarrow \text{Im}(A) = \mathcal{M}_{n,1}(\mathbb{K}) \Leftrightarrow \text{rg}(A) = n \Leftrightarrow \text{Ker}(A) = \{0_{\mathcal{M}_{n,1}(\mathbb{K})}\}.$$

Ceci permet de montrer que A est inversible sans calculer son inverse.

Démonstration – La matrice A est inversible si et seulement si u_A est un isomorphisme. La caractérisation des isomorphismes en dimension finie donne alors le résultat. \square

Corollaire

Soient (x_1, \dots, x_p) une famille de vecteurs de E (on rappelle que $p = \dim(E)$) et

$$A = \text{Mat}_{\mathcal{B}}(x_1, \dots, x_p).$$

Alors, pour que (x_1, \dots, x_p) soit une base de E , il faut et il suffit que A soit inversible.

Démonstration – La famille (x_1, \dots, x_p) est une famille de p vecteurs de E avec $p = \dim(E)$, donc c'est une base de E si et seulement si elle est génératrice de E , ce qui équivaut à $\text{rg}(x_1, \dots, x_p) = p$, *i.e.*, à $\text{rg}(A) = p$. D'après le théorème précédent, ceci équivaut à l'inversibilité de A . \square

En appliquant le théorème du rang à u_A avec $A \in \mathcal{M}_{n,p}(\mathbb{K})$ (l'espace de départ étant alors de dimension p), on obtient :

Théorème du rang pour les matrices

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors

$$\dim(\text{Ker}(A)) + \text{rg}(A) = p.$$

Propriété

Soient A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$.

Si $AB = I_n$ alors A et B sont inversibles et inverses l'une de l'autre.

Démonstration – Si $AB = I_n$ alors $u_A \circ u_B = \text{Id}$ donc u_A est surjective et u_B est injective. D'après la caractérisation des isomorphismes en dimension finie, on en déduit que u_A et u_B sont des isomorphismes, donc A et B sont inversibles. De plus, $(u_A)^{-1} = u_B$ et donc $A^{-1} = B$. \square

Remarque – Lorsque $AB = I_n$ pour deux matrices A et B de $\mathcal{M}_n(\mathbb{K})$, il est donc inutile de vérifier que $BA = I_n$, on peut directement conclure que A et B sont inversibles et inverses l'une de l'autre.

IV. La méthode de Gauss-Jordan

1. Opérations élémentaires

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. On appelle **opérations élémentaires** les manipulations suivantes (où λ désigne un scalaire) :

► Opérations élémentaires sur les colonnes de A :

- Multiplier la i -ième colonne de A par un scalaire λ **non nul** : $C_i \leftarrow \lambda C_i$.
- Permuter les colonnes i et j de A : $C_i \leftrightarrow C_j$,
- Ajouter à la i -ième colonne de A , λ fois la j -ième ($j \neq i$) : $C_i \leftarrow C_i + \lambda C_j$,

► Opérations élémentaires sur les lignes de A :

- Multiplier la i -ième ligne de A par un scalaire λ **non nul** : $L_i \leftarrow \lambda L_i$.
- Permuter les lignes i et j de A : $L_i \leftrightarrow L_j$,
- Ajouter à la i -ième ligne de A , λ fois la j -ième ($j \neq i$) : $L_i \leftarrow L_i + \lambda L_j$,

Définition – Matrices élémentaires

Dans les matrices qui suivent, les coefficients non précisés sont égaux à 0. Soit $m \in \mathbb{N}^*$.

- Pour tout $i \in \llbracket 1, m \rrbracket$ et $\lambda \in \mathbb{K}^*$, on définit la matrice de **dilatation**

$$D_i^m(\lambda) = \lambda E_{i,i} + \sum_{\substack{k \in \llbracket 1, m \rrbracket \\ k \neq i}} E_{k,k} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \lambda & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix} \begin{matrix} C_i \\ \\ \\ L_i \\ \\ \\ \\ \\ \\ \\ \end{matrix} \in \mathcal{M}_m(\mathbb{K})$$

- Pour tout $(i, j) \in \llbracket 1, m \rrbracket^2$ tel que $i \neq j$, on définit la matrice de **transposition**

$$\tau_{i,j}^m = E_{i,j} + E_{j,i} + \sum_{\substack{k \in \llbracket 1, m \rrbracket \\ k \neq i, k \neq j}} E_{k,k} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & \cdots & 1 & \\ & & \vdots & \ddots & \vdots & \\ & & 1 & \cdots & 0 & \\ & & & & & \ddots \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix} \begin{matrix} C_i & C_j \\ \\ \\ L_i \\ L_j \\ \\ \\ \\ \\ \end{matrix} \in \mathcal{M}_m(\mathbb{K})$$

Dans la matrice précédente, on a choisi $i < j$, ce qui n'est pas restrictif car pour tout $(i, j) \in \llbracket 1, m \rrbracket^2$ tel que $i \neq j$, $\tau_{i,j}^m = \tau_{j,i}^m$.

- Pour tout $(i, j) \in \llbracket 1, m \rrbracket^2$ tel que $i \neq j$, pour tout $\lambda \in \mathbb{K}$, on définit la matrice de **transvection**

$$T_{i,j}^m(\lambda) = I_m + \lambda E_{i,j} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \vdots & \ddots & & \\ & & \lambda & \cdots & 1 & \\ & & & & & \ddots \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix} \begin{matrix} C_j & C_i \\ \\ \\ L_j \\ L_i \\ \\ \\ \\ \\ \end{matrix} \in \mathcal{M}_m(\mathbb{K})$$

Dans la matrice précédente, on a choisi $i > j$, ce qui est restrictif : il est bien sûr possible de choisir $i < j$, auquel cas le coefficient λ sera placé au-dessus de la diagonale.

Propriété – Traduction matricielle des opérations élémentaires

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors :

1. Opérations élémentaires sur les colonnes de A :

- La matrice obtenue à partir de A par l'opération $C_i \leftarrow \lambda C_i$ est $AD_i^p(\lambda)$.
- La matrice obtenue à partir de A par l'opération $C_i \leftrightarrow C_j$ est $A\tau_{i,j}^p$.
- La matrice obtenue à partir de A par l'opération $C_i \leftarrow C_i + \lambda C_j$ est $AT_{j,i}^p(\lambda)$.

2. Opérations élémentaires sur les lignes de A :

- La matrice obtenue à partir de A par l'opération $L_i \leftarrow \lambda L_i$ est $D_i^n(\lambda)A$.
- La matrice obtenue à partir de A par l'opération $L_i \leftrightarrow L_j$ est $\tau_{i,j}^n A$.
- La matrice obtenue à partir de A par l'opération $L_i \leftarrow L_i + \lambda L_j$ est $T_{i,j}^n(\lambda)A$.

Démonstration

1. On rappelle qu'en général, la k -ième colonne d'un produit AB est le produit de A par la k -ième colonne de B . On notera C_1, \dots, C_p les colonnes de A .

• En notant T_1, \dots, T_p les colonnes de $D_i^p(\lambda)$, on a $AT_i = \lambda C_i$ et $AT_k = C_k$ si $k \neq i$. D'où le résultat du premier point.

• De même, en notant T_1, \dots, T_p les colonnes de $\tau_{i,j}^p$, on a $AT_i = C_j$, $AT_j = C_i$ et $AT_k = C_k$ si $k \neq i$ et $k \neq j$. D'où le résultat du deuxième point.

• Enfin, en notant T_1, \dots, T_p les colonnes de $T_{j,i}^p(\lambda)$, on a $AT_i = C_i + \lambda C_j$ et $AT_k = C_k$ si $k \neq i$. D'où le résultat du troisième point.

2. De même, la k -ième ligne d'un produit BA est le produit de la k -ième ligne de B par A . On notera L_1, \dots, L_n les lignes de A .

• En notant T_1, \dots, T_n les lignes de $D_i^n(\lambda)$, on a $T_i A = \lambda L_i$ et $T_k A = L_k$ si $k \neq i$. D'où le résultat du premier point.

• De même, en notant T_1, \dots, T_n les lignes de $\tau_{i,j}^n$, on a $T_i A = L_j$, $T_j A = L_i$ et $T_k A = L_k$ si $k \neq i$ et $k \neq j$. D'où le résultat du deuxième point.

• Enfin, en notant T_1, \dots, T_n les lignes de $T_{i,j}^n(\lambda)$, on a $T_i A = L_i + \lambda L_j$ et $T_k A = L_k$ si $k \neq i$. D'où le résultat du troisième point. \square

Propriété

Toutes les matrices élémentaires sont inversibles ; plus précisément : pour tout $m \in \mathbb{N}^*$, pour tout $i \in \llbracket 1, m \rrbracket$, tout $j \in \llbracket 1, m \rrbracket$ tel que $i \neq j$, pour tout $\lambda \in \mathbb{K}$,

$$(D_i^m(\lambda))^{-1} = D_i^m(1/\lambda) \text{ si } \lambda \neq 0, \quad (\tau_{i,j}^m)^{-1} = \tau_{i,j}^m, \quad (T_{i,j}^m(\lambda))^{-1} = T_{i,j}^m(-\lambda).$$

Démonstration – On raisonne à l'aide d'opérations sur les lignes : si $\lambda \neq 0$, $D_i^m(1/\lambda)D_i^m(\lambda)$ est, d'après la propriété précédente, la matrice obtenue à partir de $D_i^m(\lambda)$ par l'opération $L_i \leftarrow L_i/\lambda$, c'est-à-dire la matrice identité I_m . Donc $D_i^m(\lambda)$ est inversible d'inverse $D_i^m(1/\lambda)$.

De même, $\tau_{i,j}^m \tau_{i,j}^m$ est la matrice obtenue à partir de $\tau_{i,j}^m$ par l'opération $L_i \leftrightarrow L_j$, c'est-à-dire, la matrice I_m . Donc $\tau_{i,j}^m$ est inversible et égale à sa propre matrice inverse.

Enfin, $T_{i,j}^m(-\lambda)T_{i,j}^m(\lambda)$ est la matrice obtenue à partir de $T_{i,j}^m(\lambda)$ par l'opération $L_i \leftarrow L_i - \lambda L_j$, c'est-à-dire, ici encore, la matrice I_m . Donc $T_{i,j}^m(\lambda)$ est inversible d'inverse $T_{i,j}^m(-\lambda)$. \square

Remarque – On comprend bien cette propriété et sa démonstration en voyant les choses ainsi : les opérations élémentaires sont « reversibles », l'opération $L_i \leftarrow \lambda L_i$ pour $\lambda \neq 0$ est compensée par l'opération $L_i \leftarrow L_i/\lambda$, l'opération $L_i \leftrightarrow L_j$ est compensée par elle-même, et l'opération $L_i \leftarrow L_i + \lambda L_j$ est compensée par l'opération $L_i \leftarrow L_i - \lambda L_j$ (de même pour les colonnes).

Définition – Matrices équivalentes par lignes ou par colonnes

Deux matrices A et A' de $\mathcal{M}_{n,p}(\mathbb{K})$ sont dites **équivalentes par lignes** (resp., par colonnes) si elles se déduisent l'une de l'autre par une suite finie d'opérations élémentaires sur les lignes (resp., les colonnes). Ceci se note : $A \underset{L}{\sim} A'$ (resp. $A \underset{C}{\sim} A'$).

Les opérations élémentaires étant réversibles, il est équivalent d'écrire $A \underset{L}{\sim} A'$ et $A' \underset{L}{\sim} A$ (de même pour les colonnes).

De plus, grâce aux résultats précédents, $A \underset{L}{\sim} A'$ si et seulement s'il existe une matrice $E \in \mathcal{G}l_n(\mathbb{K})$, qui est un produit de matrices élémentaires, telle que $A = EA'$. De même, $A \underset{C}{\sim} A'$ si et seulement s'il existe une matrice $E \in \mathcal{G}l_p(\mathbb{K})$ produit de matrices élémentaires, telle que $A = A'E$.

2. Algorithme du pivot de Gauss-Jordan

On a montré plus haut que le rang n'est pas modifié par multiplication à gauche ou à droite par une matrice inversible. En fait, il y a d'autres invariants de ce type : soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$.

- Si $P \in \mathcal{G}l_n(\mathbb{K})$, alors pour tout vecteur $X \in \mathcal{M}_{p,1}(\mathbb{K})$, on a l'équivalence :

$$AX = 0 \Leftrightarrow PAX = 0,$$

car P est inversible. Ceci montre que $\text{Ker}(A) = \text{Ker}(PA)$: le noyau d'une matrice n'est donc pas modifié par multiplication à gauche par une matrice inversible.

- Si $P \in \mathcal{G}l_p(\mathbb{K})$, alors pour tout vecteur $Y \in \mathcal{M}_{n,1}(\mathbb{K})$, on a l'équivalence

$$\exists X \in \mathcal{M}_{p,1}(\mathbb{K}); Y = AX \Leftrightarrow \exists X \in \mathcal{M}_{p,1}(\mathbb{K}); Y = (AP)(P^{-1}X).$$

Sachant que $P^{-1}X$ décrit $\mathcal{M}_{p,1}(\mathbb{K})$ lorsque X décrit $\mathcal{M}_{p,1}(\mathbb{K})$ (car P est inversible), on en déduit que $\text{Im}(A) = \text{Im}(AP)$: l'image d'une matrice n'est donc pas modifiée par multiplication à droite par une matrice inversible.

Reprenons ces considérations dans le cas où P est une matrice élémentaire : le rang d'une matrice A n'est pas modifié par les opérations élémentaires, son noyau n'est pas modifié par les opérations élémentaires sur ses *lignes*, son image n'est pas modifiée par les opérations élémentaires sur ses *colonnes*. En d'autres termes :

Propriété

- Deux matrices équivalentes par lignes ou par colonnes ont le même rang.
- Deux matrices équivalentes par lignes ont le même noyau.
- Deux matrices équivalentes par colonnes ont la même image.

Il est donc naturel d'essayer, au moyen d'opérations élémentaires bien choisies, d'obtenir à partir de A une matrice A' sur laquelle il sera plus facile de lire les informations telles que son rang, son noyau ou son image, qui seront les mêmes que ceux de A . C'est l'objectif de l'algorithme de Gauss-Jordan.

Commençons par décrire la forme équivalente la plus simple à laquelle on souhaite aboutir :

Définition – Matrices échelonnées, échelonnées réduites

Soit $B \in \mathcal{M}_{n,p}(\mathbb{K})$.

- On dit que B est **échelonnée par lignes** si elle vérifie les propriétés suivantes :
 - (i) Si une ligne de B est nulle, alors toutes les lignes suivantes de B sont nulles.
 - (ii) Le cas échéant, dans chaque ligne non nulle à partir de la deuxième ligne, le premier coefficient non nul (à partir de la gauche) et situé strictement à droite du premier coefficient non nul de la ligne précédente.

Le premier coefficient non nul d'une ligne non nulle est appelé **pivot**.

- On dit que B est **échelonnée réduite par lignes** si elle est échelonnée par lignes et si tous ses pivots sont égaux à 1 et sont les seuls éléments non nuls de leur colonne.
- On dit que B est **échelonnée par colonnes** (resp. **échelonnée réduite par colonnes**) si tB est échelonnée par lignes (resp. échelonnée réduite par lignes).

Remarque – Une matrice échelonnée réduite par lignes non nulle a la forme suivante (les pivots sont notés en gras, le symbole * désigne un coefficient éventuellement non nul.) :

$$\begin{pmatrix} 0 & \dots & \mathbf{1} & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * \\ 0 & \dots & 0 & 0 & \dots & 0 & \mathbf{1} & * & \dots & * & 0 & * & \dots & * \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & * & \dots & * & 0 & * & \dots & * \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \ddots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & * & 0 & * & \dots & * \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \mathbf{1} & * & \dots & * \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Le trait de séparation représenté en partie dans la matrice ci-dessus permet de mettre en valeur sa structure de matrice échelonnée et ses pivots. On parle de schéma en escalier.

Chaque ligne et chaque colonne possédant au plus un pivot, le nombre r de pivots d'une matrice échelonnée $B \in \mathcal{M}_{n,p}(\mathbb{K})$ vérifie $r \leq n$ et $r \leq p$.

Exemple – La matrice

$$\begin{pmatrix} \mathbf{1} & 0 & 3 & 5 & 7 \\ 0 & \mathbf{2} & 4 & 6 & 8 \\ 0 & 0 & 0 & \mathbf{1} & 9 \end{pmatrix}$$

est échelonnée par lignes, mais pas échelonnée réduite par lignes. La matrice

$$\begin{pmatrix} \mathbf{1} & 0 & 3 & 0 & 7 \\ 0 & \mathbf{1} & 4 & 0 & 8 \\ 0 & 0 & 0 & \mathbf{1} & 9 \end{pmatrix}$$

est échelonnée réduite par lignes. Une telle matrice peut donc tout à fait posséder, en dehors des pivots, des coefficients non nuls.

La matrice précédente n'est pas échelonnée par colonnes. La matrice

$$\begin{pmatrix} \mathbf{1} & 0 & 0 \\ 2 & 0 & 0 \\ 0 & \mathbf{1} & 0 \end{pmatrix}$$

est échelonnée réduite par colonnes.

Théorème – Algorithme de Gauss-Jordan et traduction matricielle

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors :

- A est équivalente par lignes à une unique matrice échelonnée réduite par lignes.

De façon équivalente : il existe une matrice $E \in \mathcal{G}\ell_n(\mathbb{K})$, qui est un produit de matrices élémentaires, et une unique matrice $R \in \mathcal{M}_{n,p}(\mathbb{K})$ échelonnée réduite par lignes, telles que $A = ER$.

- A est équivalente par colonnes à une unique matrice échelonnée réduite par colonnes.

De façon équivalente : il existe une matrice $E' \in \mathcal{G}\ell_p(\mathbb{K})$, qui est un produit de matrices élémentaires, et une unique matrice $R' \in \mathcal{M}_{n,p}(\mathbb{K})$ échelonnée réduite par colonnes, telles que $A = R'E'$.

Démonstration de l'existence (la démonstration de l'unicité, non exigible, est admise)

Montrons tout d'abord que le premier point entraîne le second ; pour cela on applique le résultat du premier point à tA : il existe E , produit de matrices élémentaires, et R échelonnée réduite par lignes, telles que ${}^tA = ER$. Alors $A = {}^tR{}^tE$. Par définition, $R' = {}^tR$ est échelonnée par colonnes ; la transposée de toute matrice élémentaire étant une matrice élémentaire, $E' = {}^tE$ est un produit de matrices élémentaires. D'où l'existence dans le cas des opérations sur les colonnes.

Dans le cas des opérations sur les lignes, on procède par récurrence sur le nombre p de colonnes de $A = (a_{i,j})$.

Initialisation : si $p = 1$, A est une matrice colonne. Si $A = 0$, le résultat est vrai, sinon, il existe $i_0 \in \llbracket 1, n \rrbracket$ tel que $a_{i_0,1} \neq 0$. L'opération $L_{i_0} \leftrightarrow L_1$ (ce qui revient à multiplier A à gauche par τ_{1,i_0}^n) fournit une matrice A' équivalente par lignes à A dont le coefficient en position $(1,1)$ vaut $a_{i_0,1}$ (et dont le coefficient en position $(i_0,1)$ vaut $a_{1,1}$). Après l'opération $L_1 \leftrightarrow L_1/a_{i_0,1}$ (multiplication de A' à gauche par $D_1^n(1/a_{i_0,1})$) le coefficient en position $(1,1)$ vaut 1. On fait alors, si $n \geq 2$, les $n - 1$ opérations suivantes : $L_i \leftarrow L_i - a_{i,1}L_1$, pour $i \in \llbracket 2, n \rrbracket$ (multiplication à gauche par $T_{i,1}^n(-a_{i,1})$), ce qui prouve que A est équivalente par lignes à

$$R = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

et prouve le résultat dans ce cas, car R est évidemment échelonnée réduite par lignes.

Hérédité : supposons le résultat vrai au rang p , et soit $A \in \mathcal{M}_{n,p+1}(\mathbb{K})$.

Premier cas : la première colonne de A est nulle. On applique alors l'hypothèse de récurrence à la matrice B obtenue en extrayant les p dernières colonnes de A . Les opérations faites sur B peuvent alors être faites sur A ; elles ne modifient pas la première colonne de A car cette colonne est nulle. La matrice obtenue à partir de B est échelonnée réduite par lignes, il en est de même pour celle obtenue à partir de A .

Deuxième cas : la première colonne de A est non nulle. On fait tout d'abord, pour cette première colonne, exactement le même raisonnement que pour l'initialisation. On note A_1 la matrice équivalente par lignes à A à laquelle on aboutit alors, puis on applique, si $n \geq 2$, l'hypothèse de récurrence à la matrice B_1 obtenue en extrayant les $n - 1$ dernières lignes et les p dernières colonnes de A_1 . Les opérations faites sur B_1 peuvent alors être traduites en des opérations sur les $n - 1$ dernières lignes de A_1 ; on obtient ainsi une matrice notée A_2 . Ces opérations ne modifient pas la première colonne de A_1 car les coefficients de A_1 en position $(i,1)$ avec $i \geq 2$ sont nuls.

Pour résumer les notations, on a donc l'enchaînement suivant dans ce cas :

$$\begin{array}{ccc}
 A = \begin{pmatrix} a_{1,1} & \cdots & \cdots & \cdots & \cdots \\ \vdots & \ddots & & & \vdots \\ a_{i_0,1} & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ a_{n,1} & \cdots & \cdots & \cdots & \cdots \end{pmatrix} & \xrightarrow{\text{traitement de la première colonne}} & A_1 = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots \\ 0 & & & & \\ \vdots & & B_1 & & \\ \vdots & & & & \\ 0 & & & & \end{pmatrix} \\
 & & \xrightarrow{\text{échelonnement de } B_1} & & A_2 = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots \\ 0 & & & & \\ \vdots & & B_2 & & \\ \vdots & & & & \\ 0 & & & & \end{pmatrix},
 \end{array}$$

la matrice B_2 étant échelonnée réduite par lignes.

En particulier, il est clair que A_2 est échelonnée par lignes. En revanche, elle peut ne pas être échelonnée *réduite* car un pivot de B_2 peut ne pas être le seul coefficient non nul de sa colonne dans la matrice A_2 . Ceci n'est possible que si $B_2 \neq 0$. Dans ce cas, notons $(i_1, j_1), \dots, (i_r, j_r)$ les positions dans la matrice A_2 des pivots de B_2 , avec $i_1 < \dots < i_r$. Pour k de r à 1, on effectue sur A_2 l'opération $L_1 \leftarrow L_1 - a_{1, j_k} L_{i_k}$. Pour chacune de ces opérations, la première colonne de A_2 n'est pas modifiée, les coefficients a_{1, j_m} tels que $m > k$ ne sont pas modifiés ; à l'issue de ces opérations, les pivots de la matrice obtenue, notée R , sont donc les seuls coefficients non nuls de leur colonne. La matrice R est échelonnée réduite par lignes, ce qui prouve le résultat du point de vue « opérations élémentaires ».

Du point de vue matriciel, chaque opération revient à multiplier à gauche par une matrice élémentaire (comme indiqué dans l'étape d'initialisation). On en déduit qu'il existe une matrice D , produit de matrices élémentaires, telle que $DA = R$. Une matrice élémentaire étant inversible et son inverse étant une matrice élémentaire, D est inversible et D^{-1} est un produit de matrices élémentaires. En posant $E = D^{-1}$, on a bien $A = ER$ avec la forme voulue. \square

Remarques

- La démonstration ci-dessus décrit entièrement une méthode effective d'échelonnement par lignes ou colonnes. Elle est en particulier programmable pour un traitement par ordinateur.
- Dans le cas d'opérations sur les lignes, la première étape de l'algorithme est dite étape de descente, elle aboutit à une forme échelonnée par lignes. La deuxième étape, qui aboutit à la forme échelonnée réduite par lignes, est dite étape de remontée.
- Le théorème affirme l'unicité de R mais pas celle de E . Cela est lié au fait qu'il n'y a pas une unique suite d'opérations élémentaires qui permet de passer de A à R . En revanche, quelle que soit la suite d'opérations convenable, on aboutira à la même matrice échelonnée réduite R . Toute suite d'opérations élémentaires sur les lignes qui permet de déduire de A une matrice échelonnée réduite par lignes est donc acceptée (de même pour les colonnes).
- Notamment, même si la démonstration précédente présente l'annulation des coefficients situés au-dessus d'un pivot seulement en fin de procédure, de la droite vers la gauche, on vérifie facilement qu'il est possible de le faire au fur et à mesure, c'est-à-dire de traiter entièrement une colonne avant de passer à la suivante. On remarquera cependant que cela entraîne des calculs moins simples (report de coefficients non nuls) lors des opérations du type $L_i \leftarrow L_i + \lambda L_j$.
- Il faut être vigilant lorsque l'on fait plusieurs opérations à la suite, par exemple sur les lignes : si après la première opération, la ligne i est modifiée, et si l'opération suivante utilise L_i , il s'agit de la ligne *modifiée*. C'est ce qui se passe par exemple lors de la suite d'opérations $L_2 \leftarrow L_2 - L_1$, $L_3 \leftarrow L_3 - L_2$: la ligne L_2 utilisée pour la deuxième opération est celle qui est issue de la première opération !

Exemples

• Échelonnons par lignes la matrice M suivante. À gauche, on indique les différentes matrices équivalentes par lignes obtenues, jusqu'à la forme échelonnée réduite par lignes, et à droite, on indique l'opération qui permet de passer à l'étape suivante, et sa traduction matricielle (m. à g. signifie « multiplication à gauche »).

$$\begin{array}{l}
 M = \begin{pmatrix} 0 & 0 & -1 & 2 \\ 2 & 4 & 6 & -2 \\ 3 & 6 & 5 & 3 \end{pmatrix} \\
 \sim_L \begin{pmatrix} 2 & 4 & 6 & -2 \\ 0 & 0 & -1 & 2 \\ 3 & 6 & 5 & 3 \end{pmatrix} \\
 \sim_L \begin{pmatrix} 1 & 2 & 3 & -1 \\ 0 & 0 & -1 & 2 \\ 3 & 6 & 5 & 3 \end{pmatrix} \\
 \sim_L \begin{pmatrix} 1 & 2 & 3 & -1 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & -4 & 6 \end{pmatrix}
 \end{array}
 \quad
 \begin{array}{l}
 L_1 \leftrightarrow L_2, \\
 L_1 \leftarrow L_1/2, \\
 L_3 \leftarrow L_3 - 3L_1,
 \end{array}
 \quad
 \begin{array}{l}
 \text{m. à g. par } \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \text{m. à g. par } \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \text{m. à g. par } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} .
 \end{array}$$

On a alors traité la première colonne, on poursuit l'algorithme en raisonnant sur la matrice extraite d'ordre 2×3 en bas à droite. La première colonne de cette matrice étant nulle, on poursuit en raisonnant sur la matrice extraite d'ordre 2×2 en bas à droite :

$$\begin{array}{l}
 \sim_L \begin{pmatrix} 1 & 2 & 3 & -1 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & -4 & 6 \end{pmatrix} \\
 \sim_L \begin{pmatrix} 1 & 2 & 3 & -1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & -4 & 6 \end{pmatrix} \\
 \sim_L \begin{pmatrix} 1 & 2 & 3 & -1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & -2 \end{pmatrix} \\
 \sim_L \begin{pmatrix} 1 & 2 & 3 & -1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix} .
 \end{array}
 \quad
 \begin{array}{l}
 L_2 \leftarrow -L_2, \\
 L_3 \leftarrow L_3 + 4L_2, \\
 L_3 \leftarrow L_3/(-2),
 \end{array}
 \quad
 \begin{array}{l}
 \text{m. à g. par } \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \text{m. à g. par } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{pmatrix} \\
 \text{m. à g. par } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/2 \end{pmatrix}
 \end{array}$$

À ce stade, on a une matrice échelonnée par lignes mais pas échelonnée réduite. On annule donc, en partant de la droite, les coefficients situés au-dessus des pivots :

$$\begin{array}{l}
 \sim_L \begin{pmatrix} 1 & 2 & 3 & -1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
 \sim_L \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
 \sim_L \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} .
 \end{array}
 \quad
 \begin{array}{l}
 L_2 \leftarrow L_2 + 2L_3, \\
 L_1 \leftarrow L_1 + L_3, \\
 L_1 \leftarrow L_1 - 3L_2,
 \end{array}
 \quad
 \begin{array}{l}
 \text{m. à g. par } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \text{ puis } \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \text{m. à g. par } \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
 \end{array}$$

La matrice ci-dessus, notée R , est la matrice échelonnée réduite par lignes associée à M . On peut donner directement à partir des calculs précédents, une décomposition ER pour la matrice M .

- L'exemple suivant illustre, sur une même matrice A , les deux méthodes. On remarquera que l'échelonnement en ligne ou en colonne n'aboutit pas à la même matrice.

Échelonnement par lignes :

$$A = \begin{pmatrix} 1 & 2 & 8 \\ -1 & 0 & -2 \\ 1 & 1 & 5 \end{pmatrix} \quad \begin{array}{l} L_3 \leftarrow L_3 + L_2 \\ L_2 \leftarrow L_2 + L_1 \end{array}$$

$$\underset{L}{\sim} \begin{pmatrix} 1 & 2 & 8 \\ 0 & 2 & 6 \\ 0 & 1 & 3 \end{pmatrix} \quad L_2 \leftrightarrow L_3$$

$$\underset{L}{\sim} \begin{pmatrix} 1 & 2 & 8 \\ 0 & 1 & 3 \\ 0 & 2 & 6 \end{pmatrix} \quad L_3 \leftarrow L_3 - 2L_2$$

$$\underset{L}{\sim} \begin{pmatrix} 1 & 2 & 8 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \quad L_1 \leftarrow L_1 - 2L_2$$

$$\underset{L}{\sim} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

Échelonnement par colonnes :

$$A = \begin{pmatrix} 1 & 2 & 8 \\ -1 & 0 & -2 \\ 1 & 1 & 5 \end{pmatrix} \quad \begin{array}{l} C_2 \leftarrow C_2 - 2C_1 \\ C_3 \leftarrow C_3 - 8C_1 \end{array}$$

$$\underset{C}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 6 \\ 1 & -1 & -3 \end{pmatrix} \quad C_3 \leftarrow C_3 - 3C_2$$

$$\underset{C}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 1 & -1 & 0 \end{pmatrix} \quad C_2 \leftarrow C_2/2$$

$$\underset{C}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1/2 & 0 \end{pmatrix} \quad C_1 \leftarrow C_1 + C_2$$

$$\underset{C}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & -1/2 & 0 \end{pmatrix}$$

Rappelons que l'image de A n'est pas modifiée par l'algorithme de Gauss-Jordan sur les colonnes. On en déduit que

$$\text{Im}(A) = \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1/2 \end{pmatrix} \right),$$

et en particulier $\text{rg}(A) = 2$. Plus généralement, cette méthode permet de déterminer l'espace vectoriel engendré par une famille finie de vecteurs d'un espace de dimension finie.

De même, le noyau n'est pas modifié par l'algorithme de Gauss-Jordan sur les lignes, et donc :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \text{Ker}(A) \quad \Leftrightarrow \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \text{Ker} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Leftrightarrow \quad \begin{cases} x + 2z = 0 \\ y + 3z = 0 \end{cases}$$

On en déduit que

$$\text{Ker}(A) = \left\{ \begin{pmatrix} -2z \\ -3z \\ z \end{pmatrix}; z \in \mathbb{K} \right\} = \text{Vect} \left(\begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix} \right),$$

de dimension 1, en accord avec le théorème du rang.

3. Échelonnement, rang et matrices inversibles

Propriété

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors le rang de A est égal au nombre de pivots de sa matrice échelonnée réduite par lignes, et égal au nombre de pivots de sa matrice échelonnée réduite par colonnes.

Démonstration – Soit R la matrice échelonnée réduite par lignes associée à A . La matrice R peut avoir des coefficients non nuls : en position de pivot, à droite dans la ligne de chaque pivot (mais pas au-dessus d'autres pivots), et seulement à ces positions.

À l'aide d'opérations sur les colonnes de R , on obtient une matrice équivalente par colonnes à R , où les coefficients autres que les pivots ont été remplacés par des 0. Cette matrice est encore

échelonnée réduite par lignes et a les mêmes pivots que R (attention cependant, elle n'est en général *pas* équivalente par lignes à A).

Finalement, il existe une suite finie d'opérations élémentaires sur les lignes et les colonnes, qui permet de déduire de A une matrice échelonnée réduite par ligne, dont les seuls coefficients non nuls sont les pivots de R , un tel pivot étant le seul coefficient non nul de sa ligne et de sa colonne. Le rang d'une telle matrice est égal au nombre r des pivots, car la famille de ses colonnes non nulles est clairement libre, et constituée de r vecteurs.

Les opérations élémentaires ne modifient pas le rang, donc le rang de A est égal au nombre de pivots de sa matrice échelonnée réduite par lignes. En raisonnant de la même façon, mais en échelonnant d'abord par colonnes, on obtient que le rang de A est égal au nombre de pivots de sa matrice échelonnée réduite par colonnes. \square

Remarques

- Dans le cas des exemples ci-dessus, les transformations du raisonnement précédent sont les suivantes :

$$M \underset{L}{\sim} \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \underset{C}{\sim} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A \underset{L}{\sim} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \underset{C}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

- Lorsque l'on passe, par opérations élémentaires, d'une matrice échelonnée à une matrice échelonnée réduite (par lignes ou colonnes), le nombre et la position des pivots ne sont pas modifiés. On en déduit que le rang d'une matrice échelonnée (même si elle n'est pas échelonnée réduite) est égal au nombre de ses pivots. En particulier, le rang de $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est égal au nombre de pivots de toute matrice échelonnée équivalente par lignes ou par colonnes à A .

Propriété – Rang de la transposée

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors $\text{rg}({}^tA) = \text{rg}(A)$.

En particulier, le rang de A (qui est le rang de la famille des colonnes de A) est aussi égal au rang de la famille de ses lignes.

Démonstration – Le rang de tA est égal au nombre de pivots de sa matrice échelonnée réduite par lignes. Or, échelonner tA par lignes revient à échelonner A par colonnes, et à transposer le résultat obtenu. Le nombre de pivots de la matrice échelonnée réduite par lignes de tA est donc égal au nombre de pivots de la matrice échelonnée réduite par colonnes de A , qui est le rang de A . On a donc $\text{rg}({}^tA) = \text{rg}(A)$. On en déduit que le rang de A est le rang de la famille des colonnes de tA , c'est-à-dire le rang de la famille des lignes de A . \square

Propriété

Soit $A \in \mathcal{M}_n(\mathbb{K})$. La matrice A est inversible si et seulement si elle est équivalente (par lignes ou par colonnes) à la matrice I_n .

Démonstration – On raisonne dans le cas des lignes, celui des colonnes est analogue. Notons R la matrice échelonnée réduite par lignes de A . La matrice A est carrée, on sait qu'elle est inversible si et seulement si $\text{rg}(A) = n$, c'est-à-dire, si et seulement si R possède n pivots. D'après la définition d'une matrice échelonnée réduite, c'est le cas si et seulement si $R = I_n$.

Or, on remarque que la matrice I_n est échelonnée réduite par lignes, donc si $A \underset{L}{\sim} I_n$, alors par unicité, $R = I_n$. La réciproque est évidente car $A \underset{L}{\sim} R$ par définition de R . Finalement, A est inversible si et seulement si $A \underset{L}{\sim} I_n$. \square

Or, on sait exactement comment déterminer la matrice échelonnée réduite par lignes de A : c'est l'algorithme de Gauss-Jordan. Il en découle un moyen effectif de déterminer A^{-1} lorsque A est inversible :

Propriété – Calcul de l'inverse par l'algorithme de Gauss-Jordan

Soit $A \in \mathcal{G}\ell_n(\mathbb{K})$. On note (L) une suite finie d'opérations élémentaires sur les lignes de A à partir de laquelle on obtient sa matrice échelonnée réduite par lignes, I_n .

Alors la matrice déduite de I_n par la suite d'opérations (L) est A^{-1} . On peut donner le même résultat sur les colonnes.

Démonstration – La suite (L) correspond à une matrice $E \in \mathcal{G}\ell_n(\mathbb{K})$, produit de matrices élémentaires, telle que $EA = I_n$. On en déduit que $E = A^{-1}$, c'est-à-dire, $EI_n = A^{-1}$. En effectuant sur I_n la suite (L) d'opérations élémentaires, on obtient donc A^{-1} . \square

Remarque – L'algorithme de Gauss-Jordan permet aussi de *prouver* que A est inversible : l'algorithme aboutit à I_n si et seulement si A est inversible.

Exemple – Soit $C = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ -1 & 2 & 2 \end{pmatrix}$. On fait en parallèle les mêmes opérations sur les lignes de C et de I_3 :

$$\begin{array}{lll} \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ -1 & 2 & 2 \end{pmatrix} & \begin{array}{l} L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 + L_1 \end{array} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 3 \end{pmatrix} & L_3 \leftarrow L_3 - 2L_2 & \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} & \begin{array}{l} L_1 \leftarrow L_1 - L_3 \\ L_2 \leftarrow L_2 - L_3 \end{array} & \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -2 & 1 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & & \begin{pmatrix} -4 & 2 & -1 \\ -7 & 3 & -1 \\ 5 & -2 & 1 \end{pmatrix}. \end{array}$$

On en déduit que C est inversible et que $C^{-1} = \begin{pmatrix} -4 & 2 & -1 \\ -7 & 3 & -1 \\ 5 & -2 & 1 \end{pmatrix}$. \square

Soit (x_1, \dots, x_p) une famille d'un \mathbb{K} -espace vectoriel E de dimension n . On rappelle que :

- (x_1, \dots, x_p) est libre si et seulement si $\text{rg}(x_1, \dots, x_p) = p$.
- (x_1, \dots, x_p) engendre E si et seulement si $\text{rg}(x_1, \dots, x_p) = n$.
- (x_1, \dots, x_p) est une base de E si et seulement si $p = n$ et $\text{rg}(x_1, \dots, x_p) = n$.

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ la matrice de (x_1, \dots, x_p) dans une base quelconque de E . Le rang de (x_1, \dots, x_p) est égal au rang de la matrice A , qui lui-même, est égal au nombre r de pivots de toute matrice échelonnée équivalente par lignes ou colonnes à la matrice A . On en déduit que :

- (x_1, \dots, x_p) est libre si et seulement si $r = p$.
- (x_1, \dots, x_p) engendre E si et seulement si $r = n$.
- (x_1, \dots, x_p) est une base de E si et seulement si $r = p = n$.

Exemples

• Avec la matrice A des exemples précédents, montrons que $\mathcal{M}_{3,1}(\mathbb{R}) = \text{Ker}(A) \oplus \text{Im}(A)$: pour cela on montre que la famille

$$\left(\begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1/2 \end{pmatrix} \right)$$

est une base de $\mathcal{M}_{3,1}(\mathbb{R})$ grâce à l'algorithme de Gauss-Jordan sur sa matrice D dans la base canonique :

$$\begin{aligned}
 D &= \begin{pmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \\ 1 & 1/2 & -1/2 \end{pmatrix} & C_1 \leftrightarrow C_2 \\
 &\underset{C}{\sim} \begin{pmatrix} 1 & -2 & 0 \\ 0 & -3 & 1 \\ 1/2 & 1 & -1/2 \end{pmatrix} & C_2 \leftarrow C_2 + 2C_1 \\
 &\underset{C}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 1 \\ 1/2 & 2 & -1/2 \end{pmatrix} & C_2 \leftrightarrow C_3 \\
 &\underset{C}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 1/2 & -1/2 & 2 \end{pmatrix} & C_3 \leftarrow C_3 + 3C_2 \\
 &\underset{C}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & -1/2 & 1/2 \end{pmatrix}
 \end{aligned}$$

La matrice précédente est échelonnée par colonnes. Avec les notations précédentes, on a dans ce cas $r = p = n = 3$, d'où le résultat. D'après un résultat du chapitre précédent (fractionnement d'une base), on a donc

$$\mathcal{M}_{3,1}(\mathbb{R}) = \text{Vect} \left(\begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix} \right) \oplus \text{Vect} \left(\begin{pmatrix} 1 \\ 0 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1/2 \end{pmatrix} \right) = \text{Ker}(A) \oplus \text{Im}(A).$$

• Dans $E = \mathbb{R}_1[X]$, on considère la famille

$$\mathcal{F} = (X + 1, X + 2, X + 3).$$

On sait que cette famille est liée car elle est constituée de 3 vecteurs en dimension 2 ; on cherche une relation de dépendance linéaire entre ses éléments. On met en œuvre l'algorithme de Gauss-Jordan sur les lignes de N , matrice de la famille \mathcal{F} dans la base $(X, 1)$ de $\mathbb{R}_1[X]$:

$$\begin{aligned}
 N &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} & L_2 \leftrightarrow L_2 - L_1 \\
 &\underset{L}{\sim} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} & L_1 \leftarrow L_1 - L_2 \\
 &\underset{L}{\sim} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}
 \end{aligned}$$

On a $r = n = 2$ et $p = 3$. On retrouve le fait que la famille \mathcal{F} est liée, mais on sait aussi qu'elle engendre $\mathbb{R}_1[X]$. De plus, déterminer les relations de dépendance linéaire entre les éléments de \mathcal{F} revient à déterminer les éléments non nuls de $\text{Ker}(N)$, qui d'après le calcul précédent, est caractérisé par le système

$$\begin{cases} x - z = 0 \\ y + 2z = 0 \end{cases}$$

On en déduit que $\text{Ker}(N) = \text{Vect}({}^t(1 \ -2 \ 1))$ et notamment,

$$(X + 1) - 2(X + 2) + (X + 3) = 0.$$

4. Résolution de systèmes linéaires

On s'intéresse dans ce paragraphe à la résolution des systèmes linéaires par l'algorithme de Gauss-Jordan. On rappelle que la forme générale d'un tel système est

$$(S) : \begin{cases} a_{1,1}x_1 + \cdots + a_{1,p}x_p = b_1 \\ a_{2,1}x_1 + \cdots + a_{2,p}x_p = b_2 \\ \vdots \\ a_{n,1}x_1 + \cdots + a_{n,p}x_p = b_n \end{cases}$$

et qu'en notant $A = (a_{i,j})_{1 \leq i \leq n, 1 \leq j \leq p} \in \mathcal{M}_{n,p}(\mathbb{K})$, $X \in \mathcal{M}_{p,1}(\mathbb{K})$ la matrice colonne de coefficients x_1, \dots, x_p et $B \in \mathcal{M}_{n,1}(\mathbb{K})$ la matrice colonne de coefficients b_1, \dots, b_n , ce système se met sous la forme matricielle $(E) : AX = B$, c'est-à-dire que (x_1, \dots, x_p) est solution de (S) si et seulement si X est solution de (E) .

Définition

Avec les notations précédentes :

- On dit que A est la **matrice du système linéaire** (S) .
- On appelle **seconds membres** du système (S) les scalaires b_1, \dots, b_n ; on appelle **colonne des seconds membres** de (S) la matrice colonne B .
- On appelle **système homogène** (ou sans second membre) associé à (S) le système obtenu à partir de (S) en remplaçant tous les b_i par 0. Ce système s'écrit matriciellement $AX = 0$.
- On appelle **matrice augmentée** associée à (S) la matrice $(A|B)$ obtenue en mettant côte à côte A et B (dans cet ordre) dans une même matrice, *i.e.*

$$\forall (i,j) \in \llbracket 1,n \rrbracket \times \llbracket 1,p+1 \rrbracket, \quad (A|B)_{i,j} = \begin{cases} a_{i,j} & \text{si } j \leq p \\ b_i & \text{si } j = p+1. \end{cases}$$

Comme on l'a expliqué dans le chapitre précédent, le système (S) possède au moins une solution si et seulement si l'équation $AX = B$ possède au moins une solution, ce qui équivaut au fait que $B \in \text{Im}(A)$.

Dans ce cas, l'ensemble des solutions de l'équation $AX = B$ est $\{X_0 + Y; Y \in \text{Ker}(A)\}$, où X_0 désigne une solution particulière de l'équation. En d'autres termes, l'ensemble des solutions de (S) est $\{x_0 + y; y \in \mathcal{S}_h\}$, où x_0 est une solution particulière de (S) et \mathcal{S}_h désigne l'ensemble des solutions du système homogène (S_h) associé à (S) .

Définition – Système compatible/incompatible

- On dit que le système (S) est **compatible** s'il possède au moins une solution (c'est-à-dire, avec les notations précédentes, si $B \in \text{Im}(A)$).
- On dit que (S) est **incompatible** dans le cas contraire.

Remarque – Un vecteur $(x_1, \dots, x_p) \in \mathbb{K}^p$ est solution de (S) si et seulement si

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,p}x_p - b_1 = 0 \\ a_{2,1}x_1 + \cdots + a_{2,p}x_p - b_2 = 0 \\ \vdots \\ a_{n,1}x_1 + \cdots + a_{n,p}x_p - b_n = 0, \end{cases} \quad \text{ce qui équivaut à : } \begin{pmatrix} x_1 \\ \vdots \\ x_p \\ -1 \end{pmatrix} \in \text{Ker}((A|B)).$$

Nous allons maintenant expliquer comment résoudre en pratique les systèmes linéaires.

Définition – Opérations sur les lignes d'un système linéaire

- On définit les mêmes opérations élémentaires *sur les lignes* d'un système linéaire que sur les matrices (en tenant compte des seconds membres).
- On dit que deux systèmes linéaires sont **équivalents** si on peut passer de l'un à l'autre par une suite finie d'opérations élémentaires sur les lignes.

Remarques

- Les opérations élémentaires étant réversibles, il n'est pas ambigu de dire que deux systèmes sont équivalents.
- Soient (S) et (S') deux systèmes linéaires de matrices respectives A et A' de même taille, et de colonnes des seconds membres respectives B et B' . Alors, pour que (S) et (S') soient équivalents, il faut et il suffit que $(A|B)$ et $(A'|B')$ soient équivalentes par lignes. Plus précisément, si (L) désigne une suite finie d'opérations sur les lignes, alors on peut passer de (S) à (S') par la suite (L) si et seulement si on peut passer de $(A|B)$ à $(A'|B')$ par la suite (L) .

Ceci justifie la présentation matricielle des systèmes linéaires : pour passer d'un système linéaire (S) à un système (S') qui lui soit équivalent, on peut former la matrice augmentée $(A|B)$ associée à (S) , effectuer des opérations élémentaires sur les lignes de $(A|B)$, ce qui fournit une matrice de la forme $(A'|B')$ à partir de laquelle on obtient (S') .

L'intérêt des opérations élémentaires sur les lignes d'un système linéaire vient notamment de la propriété suivante :

Propriété

Deux systèmes linéaires équivalents ont le même ensemble de solutions.

Démonstration – Avec les notations précédentes, si (S) et (S') sont équivalents, $(A|B)$ et $(A'|B')$ sont équivalentes par lignes. Comme on l'a déjà montré, elles ont donc le même noyau. Ainsi, pour $(x_1, \dots, x_p) \in \mathbb{K}^p$, on a l'équivalence

$$\begin{pmatrix} x_1 \\ \vdots \\ x_p \\ -1 \end{pmatrix} \in \text{Ker}((A|B)) \Leftrightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_p \\ -1 \end{pmatrix} \in \text{Ker}((A'|B')),$$

et donc, d'après une remarque faite plus haut, (x_1, \dots, x_p) est solution de (S) si et seulement si (x_1, \dots, x_p) est solution de (S') . \square

Résolution pratique d'un système linéaire

Un système linéaire (S) de forme matricielle $AX = B$, dont la matrice A est échelonnée réduite par lignes, est particulièrement facile à résoudre : en gardant à l'esprit la forme générale d'une matrice échelonnée réduite par lignes donnée page 77, notons $(i_1, j_1), \dots, (i_r, j_r)$ les positions des pivots de A (on suppose A non nulle) ; en particulier, on a :

- $j_1 < \dots < j_r$,
- les r premières lignes de A sont non nulles et, le cas échéant, les $n - r$ dernières sont nulles,
- pour tout $k \in \llbracket 1, r \rrbracket$, $a_{i_k, j_k} = 1$, et $a_{i_k, j} = 0$ si $j < j_k$, $a_{i, j_k} = 0$ si $i \neq i_k$.

Commençons par examiner, le cas échéant, les $n - r$ dernières lignes de A . Elles correspondent aux équations $0 = b_i$, pour $i \in \llbracket r + 1, n \rrbracket$. Si l'un des b_i , pour $i \in \llbracket r + 1, n \rrbracket$, est non nul, alors (S) ne possède aucune solution : il est incompatible.

Sinon, la r -ième équation s'écrit

$$x_{j_r} + a_{r, j_r+1} x_{j_r+1} + \dots + a_{r, p} x_p = b_r,$$

elle donne directement x_{j_r} en fonction de b_r et x_{j_r+1}, \dots, x_p . On remonte alors dans le système, jusqu'à la première équation,

$$x_{j_1} + a_{1,j_1+1} x_{j_1+1} + \dots + a_{1,p} x_p = b_1,$$

ce qui donne directement x_{j_1} en fonction de b_1 et x_{j_1+1}, \dots, x_p , mais l'expression ne fait pas intervenir x_{j_2}, \dots, x_{j_r} car la matrice A est échelonnée *réduite*. Finalement, dans ce cas, (S) possède des solutions, et tout choix de valeurs pour les x_j tels que $j \notin \{j_1, \dots, j_r\}$ donne explicitement une solution de (S) .

Dans le cas général, (A quelconque), il existe une matrice E , produit de matrices élémentaires, et R , échelonnée réduite par lignes, telles que $EA = R$. L'équation $AX = B$ équivaut à l'équation $RX = EB$: on retrouve la situation précédente. On remarque que EB est la matrice colonne obtenue en effectuant sur B les opérations faites pour passer de A à sa forme échelonnée réduite par lignes R .

En pratique, pour résoudre l'équation $AX = B$, on forme la matrice augmentée $(A|B)$, sur laquelle on met en œuvre l'algorithme de Gauss-Jordan sur les lignes :

- À l'issue de la phase de descente, on peut déjà déterminer si le système est compatible ou incompatible : il est compatible si et seulement si la dernière colonne (correspondant au second membre) ne contient aucun pivot. Les opérations élémentaires que l'on aurait faites en traitant uniquement A suffisent à faire cette vérification.
- Si le système est compatible, la phase de remontée fera intervenir les mêmes opérations élémentaires que si l'on échelonnait uniquement A , car le dernier pivot ne se situe pas dans la dernière colonne correspondant au second membre. On obtient donc bien la forme équivalente $RX = EB$.

Exemple – Résolvons le système linéaire

$$\begin{cases} x & +2y & +8z & = 7 \\ -x & & -2z & = -3 \\ x & +y & +5z & = 5 \end{cases}$$

La matrice de ce système est la matrice A d'un exemple traité page 81. La colonne des seconds membres est $B = {}^t(7 \ -3 \ 5)$. On met en œuvre l'algorithme de Gauss-Jordan sur les lignes de la matrice augmentée $(A|B)$:

$$\begin{aligned} (A|B) &= \begin{pmatrix} 1 & 2 & 8 & 7 \\ -1 & 0 & -2 & -3 \\ 1 & 1 & 5 & 5 \end{pmatrix} & \begin{array}{l} L_3 \leftarrow L_3 + L_2 \\ L_2 \leftarrow L_2 + L_1 \end{array} \\ \sim_L & \begin{pmatrix} 1 & 2 & 8 & 7 \\ 0 & 2 & 6 & 4 \\ 0 & 1 & 3 & 2 \end{pmatrix} & L_2 \leftrightarrow L_3 \\ \sim_L & \begin{pmatrix} 1 & 2 & 8 & 7 \\ 0 & 1 & 3 & 2 \\ 0 & 2 & 6 & 4 \end{pmatrix} & L_3 \leftarrow L_3 - 2L_2 \\ \sim_L & \begin{pmatrix} 1 & 2 & 8 & 7 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Le système est compatible. L'opération $L_1 \leftarrow L_1 - 2L_2$ montre finalement que

$$(A|B) \sim_L \begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Le système (S) est donc équivalent à

$$\begin{cases} x + 2z = 3 \\ y + 3z = 2 \end{cases} \quad i.e., \quad \text{à} \quad \begin{cases} x = 3 - 2z \\ y = 2 - 3z \end{cases}$$

L'inconnue z n'est liée par aucune équation, on la choisit comme paramètre, que l'on peut renommer λ , c'est-à-dire que l'ensemble des solutions de (S) est

$$\{(3 - 2\lambda, 2 - 3\lambda, \lambda); \lambda \in \mathbb{K}\} = \{(3, 2, 0) + \lambda(-2, -3, 1); \lambda \in \mathbb{K}\}.$$

On obtient une représentation paramétrique de l'ensemble des solutions, et on retrouve bien, pour les solutions, la forme générale $x_0 + y$ où $x_0 = (3, 2, 0)$ est une solution particulière (obtenue pour $\lambda = 0$), et $y \in \text{Vect}(-2, -3, 1)$ avec

$$\text{Vect} \left(\begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix} \right) = \text{Ker}(A).$$

Définition

Soit (S) un système linéaire de matrice A non nulle.

- Les inconnues x_{j_1}, \dots, x_{j_r} dont les indices sont ceux des colonnes des pivots de la matrice échelonnée réduite par lignes associée à A , sont appelées **inconnues principales** de (S) .
- Les autres inconnues sont appelées **inconnues secondaires**, ou **paramètres**.
- On appelle **rang** du système (S) le nombre r , c'est-à-dire le nombre de pivots de la matrice échelonnée réduite par lignes associée à A .
- Le nombre de paramètres est donc égal à $p - r$, c'est-à-dire, à la différence du nombre d'inconnues et du rang de (S) .

Remarques

- Comme on l'a montré dans le paragraphe 3, r est aussi le rang de la matrice A , ce qui montre la cohérence de l'appellation.
- Un système sans second membre est toujours compatible, car le p -uplet $(0, \dots, 0)$ en est solution.
- Dans l'exemple traité ci-dessus, les inconnues principales sont x et y , le paramètre est z . Le rang du système est 2.

De l'étude précédente, on déduit que trois cas se présentent quant à l'ensemble \mathcal{S} des solutions d'un système linéaire (S) de rang r et de matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$:

- Si le système est incompatible, alors $\mathcal{S} = \emptyset$.
- Si le système est compatible et si $r = p$, alors le système n'a que des inconnues principales, et donc, possède une unique solution : \mathcal{S} est réduit à un point.
- Si le système est compatible et si $r < p$, alors le système a $p - r$ paramètres et \mathcal{S} est infini.

Le théorème du rang montre aussi que

$$p = \dim(\text{Ker}(A)) + \text{rg}(A), \quad \text{c'est-à-dire que } p - r = \dim(\text{Ker}(A)).$$

Le nombre de paramètres d'un système compatible de matrice A est donc égal à $\dim(\text{Ker}(A))$. Ceci est bien sûr cohérent avec la description de l'ensemble des solutions de (S) .

De plus, on rappelle que le système est compatible si et seulement si $B \in \text{Im}(A)$. Par exemple, si $r = n$, alors $\text{rg}(A) = \dim(\mathcal{M}_{n,1}(\mathbb{R}))$, et donc $\text{Im}(A) = \mathcal{M}_{n,1}(\mathbb{R})$: le système est donc compatible quel que soit le choix de B . Si $r < n$, il existe des choix de B pour lesquels le système est incompatible. C'est par exemple le cas si $p < n$, car dans ce cas $r \leq p < n$.

La situation suivante est également intéressante : si $r = n = p$, alors quel que soit le choix de B , le système (S) est compatible et possède une unique solution (on dit dans ce cas que (S) est un système de Cramer). On retrouve ce résultat en remarquant que dans ce cas, A est une matrice carrée inversible ; pour tout B , on a l'équivalence $AX = B \Leftrightarrow X = A^{-1}B$.

On a en fait la caractérisation suivante :

Propriété – Matrices inversibles et résolution de systèmes

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Les propriétés suivantes sont équivalentes :

1. A est inversible.
2. Le système $AX = 0$ admet pour unique solution le vecteur nul.
3. Pour tout $B \in \mathcal{M}_{n,1}(\mathbb{K})$, le système $AX = B$ possède une unique solution.
4. Pour tout $B \in \mathcal{M}_{n,1}(\mathbb{K})$, le système $AX = B$ possède au moins une solution.

Démonstration – La matrice A est inversible si et seulement si l'application u_A canoniquement associée à A est un isomorphisme, ce qui équivaut au point 3. D'après la caractérisation des isomorphismes en dimension finie, ceci équivaut au fait que $\text{Ker}(A) = \{0_{\mathcal{M}_{n,1}(\mathbb{K})}\}$, *i.e.* au point 2, et également au fait que $\text{Im}(A) = \mathcal{M}_{n,1}(\mathbb{K})$ (*i.e.* au point 4). \square

Dans ce cas, la résolution du système $AX = B$ où B est un second membre quelconque permet même d'explicitier A^{-1} : si l'on résout le système $AX = B$, c'est-à-dire si l'on détermine $M \in \mathcal{M}_n(\mathbb{K})$ telle que $AX = B$ équivaut à $X = MB$, alors quel que soit le choix de B , on a $A^{-1}B = MB$, d'où l'on déduit que $A^{-1} = M$ en choisissant pour B les vecteurs de la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$. Par exemple, on vérifie par opérations sur les lignes que pour $(x_1, x_2, x_3) \in \mathbb{K}^3$ et $(b_1, b_2, b_3) \in \mathbb{K}^3$,

$$\begin{cases} 2x_1 + x_2 + x_3 = b_1 \\ x_1 + 2x_2 + x_3 = b_2 \\ x_1 + x_2 + 2x_3 = b_3 \end{cases} \Leftrightarrow \begin{cases} x_1 = \frac{1}{4}(3b_1 - b_2 - b_3) \\ x_2 = \frac{1}{4}(-b_1 + 3b_2 - b_3) \\ x_3 = \frac{1}{4}(-b_1 - b_2 + 3b_3) \end{cases}$$

d'où l'on déduit que

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \text{ est inversible avec } A^{-1} = \frac{1}{4} \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}.$$

V. Trace d'une matrice et d'un endomorphisme

1. Trace d'une matrice carrée

Définition – Trace d'une matrice

Soit $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. On appelle **trace** de A le scalaire

$$\text{Tr}(A) = \sum_{i=1}^n a_{i,i},$$

c'est-à-dire la somme des coefficients diagonaux de A .

Propriété

L'application Tr est une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$.

Démonstration – C'est immédiat, Tr est une somme de formes linéaires sur $\mathcal{M}_n(\mathbb{K})$. \square

Propriété

Soient A et B deux matrices carrées d'ordre n . Alors $\text{Tr}(AB) = \text{Tr}(BA)$.

Démonstration – Notons $a_{i,j}$ et $b_{i,j}$ les coefficients de A et B . Alors pour tout $(i,j) \in \llbracket 1,n \rrbracket^2$,

$$(AB)_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j},$$

de sorte que $\text{Tr}(AB) = \sum_{i=1}^n (AB)_{i,i} = \sum_{i=1}^n \sum_{k=1}^n a_{i,k} b_{k,i}$. En changeant d'indice, on peut écrire

$$\text{Tr}(AB) = \sum_{i,j=1}^n a_{i,j} b_{j,i}.$$

En échangeant les rôles de A et B , on a de même

$$\text{Tr}(BA) = \sum_{i,j=1}^n b_{i,j} a_{j,i}.$$

Le changement d'indice $i \leftrightarrow j$ montre alors que $\text{Tr}(BA) = \text{Tr}(AB)$. \square

Propriété

Deux matrices semblables ont la même trace.

Démonstration – Si A et B sont semblables, il existe $P \in \mathcal{GL}_n(\mathbb{K})$ tel que $B = P^{-1}AP$. Alors d'après la propriété précédente,

$$\text{Tr}(B) = \text{Tr}(P^{-1}AP) = \text{Tr}(APP^{-1}) = \text{Tr}(A). \quad \square$$

2. Trace d'un endomorphisme

Propriété

Soient E un \mathbb{K} -espace vectoriel de dimension finie et $u \in \mathcal{L}(E)$. Toutes les matrices représentant l'endomorphisme u ont la même trace : si \mathcal{B} et \mathcal{B}' sont deux bases de E , si $A = \text{Mat}_{\mathcal{B}}(u)$ et $B = \text{Mat}_{\mathcal{B}'}(u)$, alors $\text{Tr}(A) = \text{Tr}(B)$.

Démonstration – Si $A = \text{Mat}_{\mathcal{B}}(u)$ et $B = \text{Mat}_{\mathcal{B}'}(u)$, alors A et B sont semblables d'après les formules de changement de bases. Le résultat provient alors de la propriété précédente. \square

Définition – Trace d'un endomorphisme

Soient E un \mathbb{K} -espace vectoriel de dimension finie et $u \in \mathcal{L}(E)$. On définit la **trace** de u comme la trace d'une quelconque de ses matrices.

Exemple – Soit

$$u : \begin{cases} \mathcal{M}_2(\mathbb{R}) & \rightarrow \mathcal{M}_2(\mathbb{R}) \\ M & \mapsto {}^tM + 2M \end{cases}$$

Pour calculer $\text{Tr}(u)$, écrivons la matrice de u dans la base de $\mathcal{M}_2(\mathbb{R})$

$$\mathcal{B} = (e_1, e_2, e_3, e_4) = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right);$$

on a

$$u(e_1) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 3e_1$$

$$u(e_2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 2 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = 2e_2 + e_3$$

$$u(e_3) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = e_2 + 2e_3$$

$$u(e_4) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + 2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 3e_4,$$

et donc

$$\text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix},$$

d'où : $\text{Tr}(u) = \text{Tr}(\text{Mat}_{\mathcal{B}}(u)) = 10$.

VI. Sous-espaces stables

1. Matrices définies par blocs

Soient m et q deux entiers naturels non nuls et $(A_{i,j})_{1 \leq i \leq m, 1 \leq j \leq q}$ une famille de matrices à coefficients dans \mathbb{K} . On suppose que pour tout $j \in \llbracket 1, q \rrbracket$, toutes les matrices $A_{i,j}$ pour $i \in \llbracket 1, m \rrbracket$ ont le même nombre p_j de colonnes. De même, on suppose que pour tout $i \in \llbracket 1, m \rrbracket$, toutes les matrices $A_{i,j}$ pour $j \in \llbracket 1, q \rrbracket$ ont le même nombre n_i de lignes.

On définit alors la matrice

$$B = \begin{pmatrix} A_{1,1} & \dots & A_{1,j} & \dots & A_{1,q} \\ \vdots & & \vdots & & \vdots \\ A_{i,1} & \dots & A_{i,j} & \dots & A_{i,q} \\ \vdots & & \vdots & & \vdots \\ A_{m,1} & \dots & A_{m,j} & \dots & A_{m,q} \end{pmatrix}$$

obtenue en écrivant « côte à côte » le contenu des matrices $A_{i,j}$. La matrice B possède $\sum_{i=1}^m n_i$ lignes et $\sum_{j=1}^q p_j$ colonnes.

Propriété – Calculs par blocs

Sous réserve de compatibilité des tailles des blocs, on a :

- $\begin{pmatrix} A & B \\ C & D \end{pmatrix} + \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} A + A' & B + B' \\ C + C' & D + D' \end{pmatrix},$
- $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} AA' + BC' & AB' + BD' \\ CA' + DC' & CB' + DD' \end{pmatrix}.$

Remarque – On généralisera sans difficulté cette propriété au cas d'un nombre quelconque de blocs.

Attention !

- Le sens des opérations est important. En effet par exemple, en général, $AB' \neq B'A$.
- Il est nécessaire que les dimensions des blocs soient compatibles. Pour le premier point, les matrices A et A' doivent être de même taille, de même pour B et B' , etc... Pour le second point, le nombre de colonnes de A doit être égal au nombre de lignes de A' , etc...

2. Sous-espaces stables

Dans ce paragraphe uniquement, E est un \mathbb{K} -espace vectoriel qui n'est pas supposé de dimension finie.

Si $E = V \oplus W$, un endomorphisme $u \in \mathcal{L}(E)$ est entièrement défini par ses restrictions à V et W , qui peuvent être plus simples si V et W sont bien choisis.

Définition – Sous-espace stable par un endomorphisme

Soit V un sous-espace vectoriel de E , et $u \in \mathcal{L}(E)$.

On dit que V est **stable** par u si $u(V) \subset V$, c'est-à-dire : $\forall x \in V, u(x) \in V$.

Exemple – Soit $E = \mathbb{K}[X]$, $V = \{P \in E; P(1) = 0\}$ et u l'application linéaire qui à $P \in E$ associe le polynôme $XP(X)$. Alors V est stable par u : si $P(1) = 0$, on a aussi $(u(P))(1) = 0$.

Propriété/Définition – Endomorphisme induit

Si V est un sous-espace vectoriel de E stable par u , alors l'application

$$u|_V : \begin{cases} V & \rightarrow V \\ x & \mapsto u(x) \end{cases}$$

est un endomorphisme de V , appelé **endomorphisme de V induit** par u .

Attention ! Il ne s'agit pas d'une simple restriction de u : l'espace d'arrivée est aussi restreint.

Propriété

Soient u et v deux endomorphismes de E qui commutent (*i.e.* $u \circ v = v \circ u$).

Alors $\text{Ker}(u)$ et $\text{Im}(u)$ sont stables par v .

Démonstration – Soit $x \in \text{Ker}(u)$. Montrons que $v(x) \in \text{Ker}(u)$: sachant que u et v commutent, on a $u(v(x)) = v(u(x)) = v(0) = 0$, *i.e.* $v(x) \in \text{Ker}(u)$, d'où le résultat.

De même si $y = u(x) \in \text{Im}(u)$, avec $x \in E$, alors $v(y) = v(u(x)) = u(v(x)) \in \text{Im}(u)$, donc $\text{Im}(u)$ est stable par v . \square

Remarque – En particulier, si $u \in \mathcal{L}(E)$, $\text{Ker}(u)$ et $\text{Im}(u)$ sont stables par u (en effet, u commute avec lui-même).

3. Traduction matricielle

Dans ce paragraphe, E est un \mathbb{K} -espace vectoriel de dimension finie $n \geq 2$.

Propriété

Soit V un sous-espace vectoriel de E de dimension $r \in \llbracket 1, n-1 \rrbracket$. Soit $\mathcal{B} = \mathcal{B}_V \sqcup \mathcal{F}$ une base de E adaptée à V (avec \mathcal{B}_V une base de V) et soit $u \in \mathcal{L}(E)$.

Alors V est stable par u si et seulement si $\text{Mat}_{\mathcal{B}}(u)$ est de la forme

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

où A est d'ordre r et 0 désigne un bloc nul. Dans ce cas, $A = \text{Mat}_{\mathcal{B}_V}(u|_V)$.

Démonstration – Notons $\mathcal{B} = (e_1, \dots, e_n)$, de sorte que $\mathcal{B}_V = (e_1, \dots, e_r)$. Le sous-espace V est stable par u si et seulement si pour tout vecteur x de V , $u(x) \in \text{Vect}(e_1, \dots, e_r)$. En raisonnant avec des combinaisons linéaires, il est immédiat que ceci équivaut à : pour tout $i \in \llbracket 1, r \rrbracket$, $u(e_i) \in \text{Vect}(e_1, \dots, e_r)$.

Ainsi, V est stable par u si et seulement si les coordonnées de $u(e_1), \dots, u(e_r)$ selon e_{r+1}, \dots, e_n sont nulles. Or les r premières colonnes de $\text{Mat}_{\mathcal{B}}(u)$ sont les matrices colonnes des coordonnées de $u(e_1), \dots, u(e_r)$ dans la base \mathcal{B} . On a donc l'équivalence souhaitée.

Lorsque V est stable par u , les colonnes de la matrice A sont les matrices colonnes des coordonnées de $u(e_1), \dots, u(e_r)$ dans la base (e_1, \dots, e_r) de V , donc $A = \text{Mat}_{\mathcal{B}_V}(u|_V)$. \square

En raisonnant de façon analogue, on obtient :

Propriété

Soient $m \geq 2$, $u \in \mathcal{L}(E)$ et $\mathcal{B} = \mathcal{B}_1 \sqcup \dots \sqcup \mathcal{B}_m$ une base de E , où, pour tout $i \in \llbracket 1, m \rrbracket$, \mathcal{B}_i est composée de n_i vecteurs.

Les propriétés suivantes sont équivalentes :

- $\text{Mat}_{\mathcal{B}}(u)$ est de la forme

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_m \end{pmatrix}$$

avec, pour tout $i \in \llbracket 1, m \rrbracket$, A_i d'ordre n_i .

- Pour tout $i \in \llbracket 1, m \rrbracket$, $E_i = \text{Vect}(\mathcal{B}_i)$ est stable par u .

Dans ce cas, on a $A_i = \text{Mat}_{\mathcal{B}_i}(u|_{E_i})$ pour tout $i \in \llbracket 1, m \rrbracket$.

On a alors une décomposition $E = E_1 \oplus \dots \oplus E_m$ en somme directe de sous-espaces stables par u .

Définition

On appelle matrice **diagonale par blocs** une matrice carrée de la forme précédente.

Cas particulier – Une matrice diagonale est une matrice diagonale par blocs telle que, avec les notations précédentes, pour tout $i \in \llbracket 1, m \rrbracket$, A_i n'a qu'un coefficient (on a alors $m = n$).

D'après la propriété précédente, si $\mathcal{B} = (e_1, \dots, e_n)$, ceci équivaut au fait que pour tout $i \in \llbracket 1, n \rrbracket$,

$$u(e_i) \in \text{Vect}(e_i).$$

c'est-à-dire, au fait que $u(e_i)$ soit un multiple de e_i .

Définition

On appelle matrice **triangulaire supérieure par blocs** une matrice carrée, définie par blocs, de la forme

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,m} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & A_{m-1,m} \\ 0 & \cdots & 0 & A_{m,m} \end{pmatrix},$$

où $m \geq 2$ et pour tout $i \in \llbracket 1, m \rrbracket$, $A_{i,i}$ est une matrice carrée.

Propriété

Soient $m \geq 2$, $u \in \mathcal{L}(E)$ et $\mathcal{B} = \mathcal{B}_1 \sqcup \dots \sqcup \mathcal{B}_m$ une base de E , où, pour tout $i \in \llbracket 1, m \rrbracket$, \mathcal{B}_i est composée de n_i vecteurs. On note, pour tout $i \in \llbracket 1, m \rrbracket$, $E_i = \text{Vect}(\mathcal{B}_i)$.

Les propriétés suivantes sont équivalentes :

- $\text{Mat}_{\mathcal{B}}(u)$ est de la forme précédente avec, pour tout $i \in \llbracket 1, m \rrbracket$, $A_{i,i}$ d'ordre n_i ,
- Pour tout $i \in \llbracket 1, m \rrbracket$, $u(E_i) \subset E_1 \oplus \dots \oplus E_i$.

Remarque – Dans ce cas, E_1 est stable par u , mais en général, pas E_2, \dots, E_m .

Cas particulier – Une matrice triangulaire supérieure est une matrice triangulaire par blocs telle que, avec les notations précédentes, pour tout $i \in \llbracket 1, m \rrbracket$, $A_{i,i}$ n'a qu'un coefficient (on a alors $m = n$).

D'après la propriété précédente, si $\mathcal{B} = (e_1, \dots, e_n)$, ceci équivaut au fait que pour tout $i \in \llbracket 1, n \rrbracket$,

$$u(e_i) \in \text{Vect}(e_1, \dots, e_i).$$

Exemple – Les matrices

$$A = \begin{pmatrix} 1 & 2 & 5 & 3 & 1 \\ 3 & 4 & 3 & 2 & 4 \\ 0 & 0 & 2 & 3 & 0 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 3 & 2 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 3 & 2 \end{pmatrix}$$

sont respectivement triangulaire par blocs et diagonale par blocs. Si A et B sont les matrices respectives de deux endomorphismes u et v de E dans une base $(e_1, e_2, e_3, e_4, e_5)$, alors en notant $E_1 = \text{Vect}(e_1, e_2)$, $E_2 = \text{Vect}(e_3)$, $E_3 = \text{Vect}(e_4, e_5)$, on a $E = E_1 \oplus E_2 \oplus E_3$, avec E_1 stable par u et v , E_2 et E_3 stables par v , et $u(E_2) \subset E_1 \oplus E_2$.

Remarque – Un objectif fondamental de l'algèbre linéaire consiste à construire des sous-espaces stables par u ou à en prouver l'existence, voire à construire des décompositions de l'espace en somme directe de sous-espaces stables par u . Dans le cas idéal, l'endomorphisme induit par u sur chacun de ces sous-espaces est une homothétie ; la matrice de u dans une base adaptée est alors diagonale, ce qui simplifie tous les calculs. C'est l'objectif de la réduction des endomorphismes, voir les chapitres **Réduction** et **Endomorphismes remarquables des espaces euclidiens**.

Exemple – Le cas particulier des projecteurs et des symétries

Si p est un projecteur (c'est-à-dire, si $p \circ p = p$), alors

$$E = \text{Ker}(p - \text{Id}) \oplus \text{Ker}(p)$$

avec $\text{Ker}(p - \text{Id}) = \text{Im}(p)$. De plus $\text{Ker}(p - \text{Id})$ et $\text{Ker}(p)$ sont stables par p (car p commute avec lui-même) et $p|_{\text{Ker}(p - \text{Id})} = \text{Id}$, $p|_{\text{Ker}(p)} = 0$. On a donc, dans une base adaptée \mathcal{B} :

$$\text{Mat}_{\mathcal{B}}(p) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

où $r = \dim(\text{Im}(p)) = \text{rg}(p)$. On remarque en particulier que $\text{Tr}(p) = \text{rg}(p)$: le rang d'un projecteur est égal à sa trace.

De même, si s est une symétrie (c'est-à-dire, si $s \circ s = \text{Id}$), alors

$$E = \text{Ker}(s - \text{Id}) \oplus \text{Ker}(s + \text{Id}).$$

De plus $\text{Ker}(s - \text{Id})$ et $\text{Ker}(s + \text{Id})$ sont stables par s , avec $s|_{\text{Ker}(s - \text{Id})} = \text{Id}$ et $s|_{\text{Ker}(s + \text{Id})} = -\text{Id}$. On a donc dans une base adaptée \mathcal{B} :

$$\text{Mat}_{\mathcal{B}}(s) = \begin{pmatrix} I_m & 0 \\ 0 & -I_q \end{pmatrix}$$

avec $m = \dim(\text{Ker}(s - \text{Id}))$ et $q = \dim(\text{Ker}(s + \text{Id}))$.

VII. Déterminant

Notation – Si f est une application de $\mathcal{M}_n(\mathbb{K})$ dans \mathbb{K} , si $A = (C_1 \cdots C_n) \in \mathcal{M}_n(\mathbb{K})$ et $U \in \mathcal{M}_{n,1}(\mathbb{K})$, on sera amené à utiliser la notation

$$f(C_1 \cdots C_{i-1} U C_{i+1} \cdots C_n) \quad \text{ou simplement} \quad f(C_1 \cdots U \cdots C_n)$$

pour $i \in \llbracket 1, n \rrbracket$. Bien sûr, cette notation n'a pas toujours de sens, par exemple pour $i = 1$, ou $i = n$. Dans ces cas, on sous-entend respectivement

$$f(U \cdots C_n) \quad \text{et} \quad f(C_1 \cdots U)$$

c'est-à-dire que dans tous les cas, on remplace la colonne C_i par U dans l'expression $f(C_1 \cdots C_n)$.

De plus, pour favoriser la lisibilité dans certains cas, on utilisera un trait de séparation vertical entre les colonnes, c'est-à-dire que la matrice $(C_1 \cdots C_n)$ sera parfois notée $(C_1 \mid \cdots \mid C_n)$.

1. Déterminant d'une matrice carrée

Théorème/Définition : Déterminant d'une matrice carrée

Il existe une unique application $f : \mathcal{M}_n(\mathbb{K}) \rightarrow \mathbb{K}$ vérifiant les propriétés suivantes :

(i) f est linéaire par rapport à chacune des colonnes de sa variable :

$$\forall i \in \llbracket 1, n \rrbracket, \forall (C_1 \cdots C_n) \in \mathcal{M}_n(\mathbb{K}), \forall (U, V) \in (\mathcal{M}_{n,1}(\mathbb{K}))^2, \forall \lambda \in \mathbb{K} :$$

$$\begin{aligned} & f(C_1 \mid \cdots \mid C_{i-1} \mid \lambda U + V \mid C_{i+1} \mid \cdots \mid C_n) \\ &= \lambda f(C_1 \mid \cdots \mid C_{i-1} \mid U \mid C_{i+1} \mid \cdots \mid C_n) + f(C_1 \mid \cdots \mid C_{i-1} \mid V \mid C_{i+1} \mid \cdots \mid C_n). \end{aligned}$$

(ii) f est antisymétrique par rapport aux colonnes de sa variable :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2; i \neq j, \forall (C_1 \cdots C_n) \in \mathcal{M}_n(\mathbb{K}),$$

$$f(C_1 \cdots \underbrace{C_i}_{\text{position } i} \cdots \underbrace{C_j}_{\text{position } j} \cdots C_n) = -f(C_1 \cdots \underbrace{C_j}_{\text{position } i} \cdots \underbrace{C_i}_{\text{position } j} \cdots C_n).$$

(iii) $f(I_n) = 1$.

Cette application est appelée **déterminant** et notée \det .

Propriété

Une application $f : \mathcal{M}_n(\mathbb{K}) \rightarrow \mathbb{K}$ qui vérifie la propriété (ii) vérifie aussi la propriété suivante : si $A \in \mathcal{M}_n(\mathbb{K})$ a deux colonnes égales, alors $f(A) = 0$.

Démonstration – En effet, si les colonnes d'indices i et j de A sont égales, avec $i \neq j$, on a par antisymétrie

$$f(A) = f(C_1 \cdots C_i \cdots C_i \cdots C_n) = -f(C_1 \cdots C_i \cdots C_i \cdots C_n) = -f(A)$$

et donc $f(A) = 0$. □

Démonstration de l'existence et de l'unicité du déterminant

Démontrons cette propriété dans le cas où $n = 3$; la démonstration est plus facile dans les cas $n = 1$ et $n = 2$, elle est hors programme pour $n \geq 4$.

Unicité : Soit f une application vérifiant les trois propriétés ci-dessus et $A = (a_{i,j}) \in \mathcal{M}_3(\mathbb{K})$. En notant (e_1, e_2, e_3) la base canonique de $\mathcal{M}_{3,1}(\mathbb{K})$, on a donc

$$f(A) = f(a_{1,1}e_1 + a_{2,1}e_2 + a_{3,1}e_3 \mid a_{1,2}e_1 + a_{2,2}e_2 + a_{3,2}e_3 \mid a_{1,3}e_1 + a_{2,3}e_2 + a_{3,3}e_3).$$

Par linéarité de f par rapport à chacune des colonnes de sa variable, on peut développer l'expression ci-dessus. De plus, d'après la propriété précédente, tous les termes correspondant à des matrices ayant deux colonnes égales sont nuls. Ainsi

$$\begin{aligned} f(A) &= a_{1,1}a_{2,2}a_{3,3} f(e_1|e_2|e_3) + a_{1,1}a_{3,2}a_{2,3} f(e_1|e_3|e_2) \\ &\quad + a_{2,1}a_{1,2}a_{3,3} f(e_2|e_1|e_3) + a_{2,1}a_{3,2}a_{1,3} f(e_2|e_3|e_1) \\ &\quad + a_{3,1}a_{1,2}a_{2,3} f(e_3|e_1|e_2) + a_{3,1}a_{2,2}a_{1,3} f(e_3|e_2|e_1) \end{aligned}$$

De plus, d'après la propriété (iii), $f(I_3) = 1$, et par antisymétrie,

$$\begin{aligned} f(e_1|e_3|e_2) &= f(e_3|e_2|e_1) = f(e_2|e_1|e_3) = -f(e_1|e_2|e_3) = -1 \\ f(e_2|e_3|e_1) &= -f(e_1|e_3|e_2) = f(e_1|e_2|e_3) = 1 \\ f(e_3|e_1|e_2) &= -f(e_1|e_3|e_2) = f(e_1|e_2|e_3) = 1. \end{aligned}$$

Finalement,

$$\begin{aligned} f(A) &= a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} \\ &\quad - a_{3,1}a_{2,2}a_{1,3} - a_{2,1}a_{1,2}a_{3,3} - a_{1,1}a_{3,2}a_{2,3}. \end{aligned}$$

Pour tout $A \in \mathcal{M}_n(\mathbb{K})$, le scalaire $f(A)$ est donc entièrement déterminé par une même formule sur les coefficients de A . En particulier, il existe au plus une application f vérifiant les trois propriétés du théorème.

Existence : On définit f par la formule obtenue ci-dessus. Il est alors immédiat que $f(I_3) = 1$ car dans ce cas, seul le terme $a_{1,1}a_{2,2}a_{3,3}$ est non nul, et il vaut 1. Donc f vérifie la propriété (iii). De plus, échanger deux colonnes de A a pour effet d'échanger les indices de colonnes correspondants sur les $a_{i,j}$, à l'intérieur de chacun des termes de la somme. On remarque alors que chaque terme affecté d'un signe positif est échangé avec un terme affecté d'un signe négatif. L'image par f de la matrice obtenue est donc $-f(A)$, ce qui prouve que f vérifie la propriété (ii). Enfin, si la colonne j de la matrice A est de la forme $\lambda U + V$ avec $(U, V) \in \mathcal{M}_{3,1}(\mathbb{K})^2$ (de coefficients respectifs notés u_1, u_2, u_3 et v_1, v_2, v_3), alors pour tout $i \in \llbracket 1, 3 \rrbracket$, $a_{i,j} = \lambda u_i + v_i$. En reportant cette expression dans la somme donnant $f(A)$, en développant le résultat et en regroupant les termes, on obtient la linéarité de f par rapport à la j -ième colonne de sa variable, et ce pour tout $j \in \llbracket 1, n \rrbracket$. \square

Remarques

- Pour $n = 1$, si $A = (a)$ avec $a \in \mathbb{K}$, on a $\det(A) = a$. Pour $n = 2$, on obtient, pour tout $(a, b, c, d) \in \mathbb{K}^4$,

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

Pour $n = 2$ et $n = 3$, les formules démontrées sont appelées règle de Sarrus. Elle n'ont pas d'équivalent lorsque $n \geq 4$.

- On remarquera que pour chacun des termes de la somme donnant $\det(A)$, on choisit un coefficient dans la première colonne, puis un dans la seconde, jusqu'à la n -ième, en choisissant des indices de lignes deux à deux distincts. On fait ensuite la somme pour toutes les façons possibles de faire un tel choix, en affectant à chaque terme un signe (dépendant en fait de l'ordre dans lequel on a choisi les lignes). Cette structure apparaît nettement dans la démonstration d'existence ci-dessus.

Propriété – Effet des opérations élémentaires

Soit $A = (C_1 \cdots C_n) \in \mathcal{M}_n(\mathbb{K})$.

- Si B est obtenue à partir de A par l'opération $C_i \leftrightarrow C_j$ ($i \neq j$), alors

$$\det(B) = -\det(A).$$

- Si B est obtenue à partir de A par l'opération $C_i \leftarrow C_i + \lambda C_j$ ($i \neq j$), alors on a :

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

- Si B est obtenue à partir de A par l'opération $C_i \leftarrow \lambda C_i$ ($\lambda \in \mathbb{K}$), alors on a :

$$\det(B) = \lambda \det(A).$$

- Pour tout $\lambda \in \mathbb{K}$, $\det(\lambda A) = \lambda^n \det(A)$.

Démonstration

- C'est une réécriture de la propriété d'antisymétrie par rapport aux colonnes.
- Par linéarité du déterminant par rapport à la i -ième colonne de sa variable,

$$\det(B) = \det(C_1 \cdots C_{i-1} C_i C_{i+1} \cdots C_n) + \lambda f(C_1 \cdots C_{i-1} C_j C_{i+1} \cdots C_n).$$

Dans le dernier terme, la colonne C_j apparaît deux fois, car $i \neq j$. Ce terme est donc nul d'après une propriété du déterminant. On en déduit que

$$\det(B) = \det(C_1 \cdots C_{i-1} C_i C_{i+1} \cdots C_n) = \det(A).$$

- Il suffit d'utiliser la linéarité du déterminant par rapport à la i -ème colonne de sa variable.
- On applique successivement le point précédent à chacune des n colonnes de A . \square

Remarques

- En particulier, on remarquera que les opérations élémentaires sur les colonnes conservent le déterminant ou le multiplient par un scalaire *non nul*.
- D'après le troisième point, le déterminant d'une matrice de dilatation $D_i^n(\lambda)$ vérifie

$$\det(D_i^n(\lambda)) = \lambda \det(I_n) = \lambda.$$

Une matrice de transposition $\tau_{i,j}^n$ est obtenue à partir de I_n par l'opération $C_i \leftrightarrow C_j$, donc par antisymétrie,

$$\det(\tau_{i,j}^n) = -1.$$

Une matrice de transvection $T_{i,j}^n(\lambda)$ est obtenue à partir de I_n par l'opération $C_j \leftarrow C_j + \lambda C_i$ qui ne modifie pas le déterminant, donc

$$\det(T_{i,j}^n) = 1.$$

Corollaire – Matrices inversibles et déterminant

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Pour que A soit inversible, il faut et il suffit que $\det(A) \neq 0$.

Démonstration

$\boxed{\Rightarrow}$ Si A est inversible, alors $A \underset{C}{\sim} I_n$, donc on peut passer de I_n à A par une suite finie d'opérations élémentaires sur les colonnes ; d'après ce qui précède, il existe $\alpha \in \mathbb{K}^*$ tel que $\det(A) = \alpha \det(I_n) = \alpha$. En particulier, $\det(A) \neq 0$.

⇐ On raisonne par contraposition : si A n'est pas inversible, l'une de ses colonnes, disons C_i , est combinaison linéaire des autres : on peut écrire

$$C_i = \sum_{j \neq i} \lambda_j C_j$$

où les λ_j sont des scalaires. Alors, par linéarité du déterminant par rapport à la i -ième colonne de sa variable,

$$\det(A) = \sum_{j \neq i} \det(C_1 \cdots \underbrace{C_j}_{\text{position } i} \cdots C_n) = 0$$

car dans chacun des termes de cette somme, deux des colonnes sont égales. □

Propriété

Soient A et B deux éléments de $\mathcal{M}_n(\mathbb{K})$. Alors $\det(AB) = \det(A) \det(B)$.

Démonstration – Si AB est inversible, B l'est également : en effet, si $X \in \mathcal{M}_{n,1}(\mathbb{K})$ vérifie $BX = 0$, alors $ABX = 0$ et, AB étant inversible, $X = 0$, ce qui prouve que B est inversible. Par contraposition, si B n'est pas inversible, AB ne l'est pas non plus. Dans ce cas, la formule est vraie car $\det(B) = \det(AB) = 0$.

Si B est inversible, elle est équivalente par colonnes à I_n et en particulier, B est un produit de matrices élémentaires. Notons m le nombre de matrices de transpositions, et p le nombre de matrices de dilatations, figurant dans ce produit. Notons enfin $\lambda_1, \dots, \lambda_p$ les coefficients de ces matrices de dilatations (on peut toujours supposer que $p \geq 1$, quitte à ajouter la dilatation I_n dans le produit). D'après la propriété sur l'effet des opérations élémentaires sur le déterminant,

$$\det(AB) = \det(A) \times (-1)^m \times \prod_{i=1}^p \lambda_i.$$

Mais on a également $B = I_n B$, et donc

$$\det(B) = \det(I_n) \times (-1)^m \times \prod_{i=1}^p \lambda_i = (-1)^m \times \prod_{i=1}^p \lambda_i.$$

On en déduit que $\det(AB) = \det(A) \det(B)$. □

Attention ! Il n'y a pas de propriété analogue pour la somme si $n \geq 2$: $\det(I_n + I_n) = 2^n \geq 4$ tandis que $\det(I_n) + \det(I_n) = 2$.

Propriété

Si A est inversible, $\det(A^{-1}) = \frac{1}{\det(A)}$.

Démonstration – En effet, $\det(A) \det(A^{-1}) = \det(AA^{-1}) = \det(I_n) = 1$. □

Propriété

Deux matrices semblables ont le même déterminant.

Démonstration – Si A et B sont semblables, il existe $P \in \mathcal{GL}_n(\mathbb{K})$ tel que $B = P^{-1}AP$. Alors d'après ce qui précède,

$$\det(B) = \det(P^{-1}AP) = \det(P^{-1}) \det(AP) = \det(AP) \det(P^{-1}) = \det(APP^{-1}) = \det(A). \quad \square$$

Propriété

Si A est une matrice carrée, on a $\det(A) = \det({}^tA)$.

Démonstration – Si A n'est pas inversible, tA ne l'est pas non plus car $\text{rg}({}^tA) = \text{rg}(A)$. Dans ce cas, la formule est vraie car $\det(A) = \det({}^tA) = 0$.

Si A est inversible, on raisonne comme dans la démonstration de la formule du produit : A est un produit de matrices élémentaires. On peut alors écrire tA comme un produit de matrices élémentaires, avec les mêmes matrices de dilatations et de transpositions que A (car de telles matrices sont symétriques). Les déterminant de A et tA étant entièrement calculable à partir du nombre de matrices de transpositions, et des coefficients des matrices de dilatations apparaissant dans ces produits, on en déduit le résultat. \square

Corollaire

Toutes les propriétés du déterminant par rapport aux colonnes sont également vraies par rapport aux lignes.

2. Déterminant d'une famille de vecteurs

Soient E un \mathbb{K} -espace vectoriel de dimension finie n et \mathcal{B} une base de E .

Définition – Déterminant d'une famille de vecteurs dans une base

Soit $\mathcal{F} = (u_1, \dots, u_n)$ une famille de n vecteurs de E . On appelle **déterminant** de \mathcal{F} dans la base \mathcal{B} , le déterminant de la matrice de \mathcal{F} dans la base \mathcal{B} .

Il est noté $\det_{\mathcal{B}}(u_1, \dots, u_n)$.

Théorème – Caractérisation des bases

Une famille (u_1, \dots, u_n) de vecteurs de E est une base de E si et seulement si

$$\det_{\mathcal{B}}(u_1, \dots, u_n) \neq 0.$$

Démonstration – La famille (u_1, \dots, u_n) est une base de E si et seulement si sa matrice dans la base \mathcal{B} est inversible, *i.e.*, si et seulement si $\det_{\mathcal{B}}(u_1, \dots, u_n) \neq 0$. \square

3. Déterminant d'un endomorphisme

Soit E un \mathbb{K} -espace vectoriel de dimension finie n .

Propriété

Soit $u \in \mathcal{L}(E)$. Toutes les matrices représentant l'endomorphisme u ont le même déterminant : si \mathcal{B} et \mathcal{B}' sont deux bases de E , si $A = \text{Mat}_{\mathcal{B}}(u)$ et $B = \text{Mat}_{\mathcal{B}'}(u)$, alors $\det(A) = \det(B)$.

Démonstration – Si $A = \text{Mat}_{\mathcal{B}}(u)$ et $B = \text{Mat}_{\mathcal{B}'}(u)$, alors A et B sont semblables d'après les formules de changement de bases. Le résultat provient alors d'une propriété donnée ci-dessus. \square

Définition – Déterminant d'une application linéaire

On définit le **déterminant** de $u \in \mathcal{L}(E)$ comme le déterminant d'une quelconque de ses matrices.

Remarque – Si $A \in \mathcal{M}_n(\mathbb{K})$, l'application linéaire u_A canoniquement associée à A a pour matrice A dans la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$, donc $\det(u_A) = \det(A)$.

Propriété

Soient u et v deux endomorphismes de E .

- Pour tout $\lambda \in \mathbb{K}$, $\det(\lambda u) = \lambda^n \det(u)$.
- $\det(u \circ v) = \det(u) \times \det(v)$.
- u est un isomorphisme si et seulement si $\det(u) \neq 0$. Dans ce cas

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

Démonstration – C'est une conséquence immédiate de la définition et des propriétés analogues sur les matrices. \square

4. Matrices triangulaires

Propriété – Déterminant d'une matrice triangulaire

Soit $(a_{i,j})_{1 \leq i \leq j \leq n}$ une famille de scalaires. Alors

$$\begin{vmatrix} a_{1,1} & \cdots & \cdots & a_{1,n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,n} \end{vmatrix} = \prod_{i=1}^n a_{i,i}$$

(de même pour une matrice triangulaire inférieure).

Démonstration – Notons A la matrice dont on cherche à calculer le déterminant. Si $a_{1,1} = 0$, le résultat est vrai car A a une colonne nulle, elle n'est donc pas inversible, et $\det(A) = 0 = \prod_{i=1}^n a_{i,i}$. Sinon, on effectue successivement les opérations

$$C_2 \leftarrow C_2 - \frac{a_{1,2}}{a_{1,1}} C_1, \dots, \quad C_n \leftarrow C_n - \frac{a_{1,n}}{a_{1,1}} C_1$$

ce qui ne modifie pas la valeur de $\det(A)$. On a donc

$$\det(A) = \begin{vmatrix} a_{1,1} & 0 & \cdots & \cdots & 0 \\ 0 & a_{2,2} & \cdots & \cdots & a_{2,n} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n,n} \end{vmatrix}$$

On reproduit le raisonnement jusqu'à aboutir à

$$\det(A) = \begin{vmatrix} a_{1,1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{n,n} \end{vmatrix}.$$

Par linéarité du déterminant par rapport à chaque colonne, on a donc

$$\det(A) = \left(\prod_{i=1}^n a_{i,i} \right) \det(I_n) = \prod_{i=1}^n a_{i,i}. \quad \square$$

5. Calculs de déterminants par blocs

Lemme

On suppose $n \geq 2$. Soit $B \in \mathcal{M}_{n-1}(\mathbb{K})$, $L \in \mathcal{M}_{1,n-1}(\mathbb{K})$ et $C \in \mathcal{M}_{n-1,1}(\mathbb{K})$. Alors les matrices définies par blocs

$$A = \begin{pmatrix} 1 & L \\ 0 & B \end{pmatrix} \quad \text{et} \quad A' = \begin{pmatrix} B & C \\ 0 & 1 \end{pmatrix}$$

ont pour déterminant $\det(B)$.

Démonstration – On fait la démonstration dans le cas de A , l'autre cas est similaire. Si B n'est pas inversible, ses lignes forment une famille liée, donc celles de A également, et A n'est pas inversible. La formule est donc vraie dans ce cas. Sinon, lorsque l'on effectue l'algorithme de Gauss-Jordan sur les colonnes de A , il est évident que l'on aboutit à la matrice

$$\begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}$$

où R est la matrice échelonnée réduite par colonnes associée à B . De plus, les opérations effectuées sur A pour aboutir à ce résultat sont du type $C_i \leftarrow C_i - \lambda C_1$ pour $i \geq 2$ (remplacement de L par une ligne de 0), elles ne changent pas le déterminant, puis ce sont les mêmes que celles effectuées sur B . Le déterminant étant entièrement calculable à partir du nombre d'échanges de colonnes, et des coefficients des dilatations effectuées, on en déduit que $\det(A) = \det(B)$. On procède de même pour A' . \square

Propriété

Soit A une matrice carrée de la forme $A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$, avec B et D des matrices carrées.

Alors $\det(A) = \det(B) \times \det(D)$.

Démonstration – Soit r l'ordre de la matrice B . On remarque que

$$\begin{pmatrix} B & C \\ 0 & D \end{pmatrix} = \begin{pmatrix} I_r & C \\ 0 & D \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & I_{n-r} \end{pmatrix}.$$

De plus, en utilisant plusieurs fois le lemme précédent, on a

$$\det \begin{pmatrix} I_r & C \\ 0 & D \end{pmatrix} = \det(D)$$

et

$$\det \begin{pmatrix} B & 0 \\ 0 & I_{n-r} \end{pmatrix} = \det(B).$$

Le résultat suit car le déterminant d'un produit de matrices est le produit des déterminants. \square

Propriété – Matrice triangulaire par blocs ou diagonale par blocs

- Soit

$$A = \begin{pmatrix} A_1 & A_{1,2} & \cdots & A_{1,m} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & A_{m-1,m} \\ 0 & \cdots & 0 & A_m \end{pmatrix} \quad \text{ou} \quad A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_m \end{pmatrix}$$

une matrice triangulaire par blocs ou diagonale par blocs. Alors

$$\det(A) = \det(A_1) \times \cdots \times \det(A_m) = \prod_{i=1}^m \det(A_i).$$

- Soient E un \mathbb{K} -espace vectoriel de dimension finie, $u \in \mathcal{L}(E)$ et E_1, \dots, E_m des sous-espaces vectoriels de E stables par u tels que $E = E_1 \oplus \cdots \oplus E_m$. Alors

$$\det(u) = \det(u|_{E_1}) \times \cdots \times \det(u|_{E_m}) = \prod_{i=1}^m \det(u|_{E_i}).$$

Démonstration

- Elle se fait par une récurrence immédiate à partir de la propriété précédente.
- Soit $\mathcal{B} = \mathcal{B}_1 \sqcup \cdots \sqcup \mathcal{B}_m$ une base de E adaptée à cette décomposition en somme directe. On sait que $\text{Mat}_{\mathcal{B}}(u)$ est de la forme

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_m \end{pmatrix}$$

où, pour tout $i \in \llbracket 1, m \rrbracket$, A_i est d'ordre $\dim(E_i)$, et $A_i = \text{Mat}_{\mathcal{B}_i}(u|_{E_i})$. Le résultat vient alors du point précédent, et du fait que $\det(u) = \det(A)$ et $\det(u|_{E_i}) = \det(A_i)$ pour tout i . \square

6. Développement d'un déterminant par rapport aux lignes et colonnes

Théorème – Développement par rapport à une ligne ou une colonne

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, soit $A_{i,j} \in \mathcal{M}_{n-1}(\mathbb{K})$ la matrice obtenue en supprimant la i -ème ligne et la j -ème colonne de A . Alors :

- Développement par rapport à la j -ième colonne :

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

- Développement par rapport à la i -ième ligne :

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

Démonstration (non exigible)

On fait la démonstration de la formule de développement par rapport aux colonnes, celle sur les lignes est analogue. Pour $i \in \llbracket 1, n \rrbracket$, on note E_i le i -ème vecteur de la base canonique

de $\mathcal{M}_{n,1}(\mathbb{K})$. Notons $A = (a_{i,j})_{1 \leq i,j \leq n} = (C_1 \cdots C_n)$. On a donc, pour tout $j \in \llbracket 1, n \rrbracket$, $C_j = \sum_{i=1}^n a_{i,j} E_i$. Alors, par linéarité du déterminant par rapport à la j -ième colonne de sa variable,

$$\begin{aligned} \det(A) &= \det \left(C_1 \mid \cdots \mid C_{j-1} \mid \sum_{i=1}^n a_{i,j} E_i \mid C_{j+1} \mid \cdots \mid C_n \right) \\ &= \sum_{i=1}^n a_{i,j} \det(C_1 \cdots C_{j-1} E_i C_{j+1} \cdots C_n). \end{aligned}$$

Notons $M_{i,j} = (C_1 \cdots C_{j-1} E_i C_{j+1} \cdots C_n)$. En échangeant la ligne $i-1$ et la ligne i , puis la ligne $i-2$ et la ligne $i-1$, jusqu'à échanger la ligne 1 et la ligne 2, on se ramène à une matrice obtenue en plaçant la i -ème ligne de $M_{i,j}$ à la place de la première ligne. Au cours de chacune de ces $i-1$ opérations, le déterminant de A est multiplié par -1 . On procède de même avec les colonnes, ce qui amène à multiplier le déterminant par -1 , pour chacune des $j-1$ opérations. On obtient ainsi une matrice

$$B = \begin{pmatrix} 1 & * \\ 0 & A_{i,j} \end{pmatrix}$$

avec $\det(M_{i,j}) = (-1)^{i+j-2} \det(B) = (-1)^{i+j} \det(B)$. Or, le lemme ci-dessus montre que l'on a $\det(B) = \det(A_{i,j})$. On en déduit que

$$\det(C_1 \cdots C_{j-1} E_i C_{j+1} \cdots C_n) = (-1)^{i+j} \det(A_{i,j})$$

et le résultat. □

Remarques

- Ces formules sont très utiles, par exemple :
 - Lorsqu'une ligne ou colonne de A a un nombre important de coefficients nuls.
 - Pour calculer des déterminants par récurrence, lorsque la structure du déterminant s'y prête (par exemple, les déterminants tridiagonaux).
- En dimension 3, on retrouve des formules déjà connues, par exemple

$$\begin{aligned} & \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} \\ &= a_{1,1}(a_{2,2}a_{3,3} - a_{3,2}a_{2,3}) - a_{2,1}(a_{1,2}a_{3,3} - a_{3,2}a_{1,3}) + a_{3,1}(a_{1,2}a_{2,3} - a_{2,2}a_{1,3}). \end{aligned}$$

En développant le membre de gauche, on retrouve bien sûr la formule du déterminant et la règle de Sarrus.

7. Déterminant de Vandermonde

Définition

Soit $(a_1, \dots, a_n) \in \mathbb{K}^n$. On pose

$$M_n(a_1, \dots, a_n) = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix} = (a_i^{j-1})_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{K})$$

et $V_n(a_1, \dots, a_n) = \det(M_n(a_1, \dots, a_n))$.

Ce déterminant (ou celui de sa transposée) est appelé **déterminant de Vandermonde** associé aux scalaires a_1, \dots, a_n .

Il est non nul si et seulement si les a_i sont deux à deux distincts, ce que l'on peut prouver sans calculer le déterminant : si deux des a_i sont égaux, alors $M_n(a_1, \dots, a_n)$ a deux lignes égales, donc son déterminant est nul. Si les a_i sont deux à deux distincts, et si ${}^t(\lambda_0 \ \dots \ \lambda_{n-1}) \in \mathbb{K}^n$ appartient au noyau de $M_n(a_1, \dots, a_n)$, alors pour tout $i \in \llbracket 1, n \rrbracket$,

$$\sum_{j=0}^{n-1} \lambda_j a_i^j = 0.$$

Le polynôme $P(X) = \sum_{j=0}^{n-1} \lambda_j X^j$, de degré au plus $n-1$, a donc n racines deux à deux distinctes, ce qui montre qu'il est nul, et donc que tous les λ_j sont nuls. Donc la matrice carrée $M_n(a_1, \dots, a_n)$ est inversible et son déterminant est non nul.

Ce déterminant et la matrice associée ont d'importantes applications. Par exemple, soient (a_0, \dots, a_n) et (b_0, \dots, b_n) dans \mathbb{K}^{n+1} . On cherche une fonction polynomiale P telle que

$$P(a_0) = b_0, \dots, P(a_n) = b_n;$$

autrement dit, connaissant les valeurs prises par une fonction polynomiale en certains points, on recherche les coefficients du polynôme associé.

Cherchons P sous la forme $x_0 + x_1 X + \dots + x_n X^n$. Les conditions ci-dessus s'écrivent

$$\begin{pmatrix} 1 & a_0 & a_0^2 & \dots & a_0^n \\ 1 & a_1 & a_1^2 & \dots & a_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^n \end{pmatrix} \begin{pmatrix} x_0 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_0 \\ \vdots \\ \vdots \\ b_n \end{pmatrix}$$

c'est-à-dire comme un système linéaire dont la matrice est la matrice de Vandermonde $M_{n+1}(a_0, \dots, a_n)$. Si les a_i sont deux à deux distincts, cette matrice est inversible, et il existe un unique polynôme de $\mathbb{K}_n[X]$ qui est solution du problème. Ces problématiques interviennent notamment en théorie du signal.

On peut en fait calculer explicitement $V_n(a_1, \dots, a_n)$:

Propriété – Déterminant de Van der Monde

Avec les notations précédentes,

$$V_n(a_1, \dots, a_n) = \prod_{i < j} (a_j - a_i).$$

Démonstration

Première méthode : si $n \geq 2$, alors, pour tout j de n à 2 , on fait l'opération élémentaire $C_j \leftarrow C_j - a_1 C_{j-1}$, ce qui ne change pas la valeur du déterminant. On obtient

$$V_n(a_1, \dots, a_n) = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & a_2 - a_1 & a_2^2 - a_1 a_2 & \dots & a_2^{n-1} - a_1 a_2^{n-2} \\ 1 & a_3 - a_1 & a_3^2 - a_1 a_3 & \dots & a_3^{n-1} - a_1 a_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n - a_1 & a_n^2 - a_1 a_n & \dots & a_n^{n-1} - a_1 a_n^{n-2} \end{vmatrix}$$

c'est-à-dire

$$V_n(a_1, \dots, a_n) = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & a_2 - a_1 & (a_2 - a_1)a_2 & \dots & (a_2 - a_1)a_2^{n-2} \\ 1 & a_3 - a_1 & (a_3 - a_1)a_3 & \dots & (a_3 - a_1)a_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n - a_1 & (a_n - a_1)a_n & \dots & (a_n - a_1)a_n^{n-2} \end{vmatrix}$$

En développant par rapport à la première ligne, on a donc

$$V_n(a_1, \dots, a_n) = \begin{vmatrix} a_2 - a_1 & (a_2 - a_1)a_2 & \dots & (a_2 - a_1)a_2^{n-2} \\ a_3 - a_1 & (a_3 - a_1)a_3 & \dots & (a_3 - a_1)a_3^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_n - a_1 & (a_n - a_1)a_n & \dots & (a_n - a_1)a_n^{n-2} \end{vmatrix}$$

(déterminant d'ordre $n - 1$). Chaque ligne L_i étant multiple de $a_{i+1} - a_1$, on obtient

$$\begin{aligned} V_n(a_1, \dots, a_n) &= (a_2 - a_1)(a_3 - a_1) \cdots (a_n - a_1) \begin{vmatrix} 1 & a_2 & a_2^2 & \dots & a_2^{n-2} \\ 1 & a_3 & a_3^2 & \dots & a_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-2} \end{vmatrix} \\ &= (a_2 - a_1)(a_3 - a_1) \cdots (a_n - a_1) V_{n-1}(a_2, \dots, a_n). \end{aligned}$$

Une récurrence immédiate, avec le fait que $V_1(a_n) = 1$, montre alors le résultat.

Deuxième méthode : si $n \geq 2$, soit $P(X) = \prod_{i=1}^{n-1} (X - a_i) = X^{n-1} + \sum_{k=0}^{n-2} \lambda_k X^k$ où les λ_k sont des scalaires.

L'opération $C_n \leftarrow C_n + \sum_{k=0}^{n-2} \lambda_k C_{k+1}$ montre que

$$V_n(a_1, \dots, a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & P(a_1) \\ 1 & a_2 & a_2^2 & \dots & P(a_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \dots & P(a_{n-1}) \\ 1 & a_n & a_n^2 & \dots & P(a_n) \end{vmatrix} = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & 0 \\ 1 & a_2 & a_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \dots & 0 \\ 1 & a_n & a_n^2 & \dots & P(a_n) \end{vmatrix},$$

et donc, en développant par rapport à la dernière colonne,

$$V_n(a_1, \dots, a_n) = P(a_n) V_{n-1}(a_1, \dots, a_{n-1}) = \prod_{i=1}^{n-1} (a_n - a_i) V_{n-1}(a_1, \dots, a_{n-1}),$$

ce qui permet de conclure par récurrence comme ci-dessus (on a $V_1(a_1) = 1$). □