Cours d'algèbre du second semestre de L1

(écrit par Jean-Yves Ducloux)

Table des matières

Sources :

• Liret et Martinais, *Mathématiques pour le DEUG. Algèbre 1ère année.* Éd. Dunod. [512 LIR]

• E. Ramis, C. Deschamps, J. Odoux, Cours de mathématiques [51 L RAM] :
[512 RAM], [514 RAM], [515 RAM], [517 RAM], [517 RAM], niveau L1 seul dans [51 L1 RAM]

• R. L. Burden, J. D. Faires, *Numerical analysis*, ch. 6 [519.6 BUR].

• Marc Hindry, *Cours de Mathématiques, Première Année.* Université Paris 7.
(`http://www.imj-prg.fr/~marc.hindry/Cours-L1.pdf`)

Ch. 1. Polynômes

Dans tout ce chapitre, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$ (ou plus généralement un « corps commutatif »).

I. L'anneau $\mathbb{K}[X]$

1. Polynômes et fonctions polynomiales

Définition

(a) Un *polynôme à coefficients dans* $\mathbb{K}$ est une suite $P = (a_k)_{k\geq 0}$ d'éléments de $\mathbb{K}$ pour laquelle il existe $n \in \mathbb{N}$ tel que : $\forall k > n \quad a_k = 0$.

On le notera $P = a_n X^n + \cdots + a_1 X + a_0$ au lieu de $(a_k)_{k\geq 0}$.

Les nombres a_0, a_1, ..., a_n s'appellent les *coefficients de* P.

(b) À chaque polynôme $P = a_n X^n + \cdots + a_1 X + a_0$ à coefficients dans $\mathbb{K}$, on associe l'application notée ici $f_P \colon \mathbb{K} \to \mathbb{K}$ définie par :

$$f_P(x) = \underbrace{a_n x^n + \cdots + a_1 x + a_0}_{\text{noté } P(x)} \text{ pour tout } x \in \mathbb{K}.$$

Une *fonction polynomiale sur* $\mathbb{K}$ est une application f de $\mathbb{K}$ dans $\mathbb{K}$ telle qu'il existe un polynôme P à coefficients dans $\mathbb{K}$ pour lequel $f = f_P$.

(c) Le *degré* d'un polynôme $P = a_n X^n + \cdots + a_1 X + a_0$ non nul, noté $\deg P$, est le plus grand $d \in \mathbb{N}$ tel que $a_d \neq 0$. Par convention on pose : $\deg 0 := -\infty$.

Exemple

On suppose ici que $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$ et $P = \overline{1}X^2 + (-\overline{1})X \overset{\text{notation}}{=} X^2 - X$.

Ici : $a_0 = \overline{1}$, $a_1 = -\overline{1}$, $a_2 = \overline{1}$, et $a_n = \overline{0}$ quand $n \geq 3$, donc $P \neq \underbrace{0}_{\text{suite nulle}}$ et $\deg P = 2$.

Par contre : $P(\overline{0}) = \overline{0}$ et $P(\overline{1}) = \overline{0}$ donc $f_P = \underbrace{0}_{\text{application nulle}}$.

Remarque

Soit P un polynôme non-nul de degré d à coefficients dans $\mathbb{R}$:

$$P = a_d X^d + \cdots + a_1 X + a_0 \text{ avec } a_0, ..., a_d \in \mathbb{R} \text{ et } a_d \neq 0.$$

On a : $|P(x)| = |x|^d \left| a_d + \frac{a_{d-1}}{x} + \cdots + \frac{a_0}{x^d} \right| \underset{x\to+\infty}{\longrightarrow} +\infty$ si $d \geq 1$ et $\underbrace{P(0)}_{a_0} \neq 0$ si $d = 0$.

En particulier : $f_P \neq 0$. (On n'utilisera plus la notation « f_P ».)

Ainsi, lorsque $\mathbb{K} = \mathbb{R}$:

si une fonction polynomiale réelle $x \mapsto a_n x^n + \cdots + a_1 x + a_0$ est nulle, alors $a_0 = ... = a_n = 0$.

Notations

(a) On note $\mathbb{K}[X]$ l'ensemble des polynômes à coefficients dans $\mathbb{K}$.

Il contient $\mathbb{K}$ vu comme ensemble des $\underbrace{\text{polynômes } a_0 \text{ pour lesquels on peut choisir } n = 0}_{\text{« polynômes constants »}}$.

(b) Étant donné $N \in \mathbb{N}$, on note $\mathbb{K}_N[X]$ l'ensemble des polynômes à coefficients dans $\mathbb{K}$ qui sont non-nuls de degré inférieur ou égal à N, ou nuls.

(La convention $\deg 0 = -\infty$ permet d'appeler $\mathbb{K}_N[X]$ *l'ensemble des polynômes de degré* $\leq N$.)

Ainsi : $\mathbb{K}_N[X] = \left\{ a_N X^N + \cdots + a_1 X + a_0 \ ; \ a_0, a_1, \ldots a_N \in \mathbb{K} \right\}$.

Proposition (immédiate)

a) L'ensemble $\mathbb{K}[X]$ est un sous-espace vectoriel de $\mathbb{K}^{\mathbb{N}}$.

(b) Soit $N \in \mathbb{N}$. L'ensemble $\mathbb{K}_N[X]$ est le sous-espace vectoriel $\mathrm{Vect}(1, X, ..., X^N)$ de $\mathbb{K}[X]$. Il admet la base suivante, appelée *base canonique de* $\mathbb{K}_N[X]$: $\mathscr{B} := (1, X, ..., X^N)$.

En particulier : $\dim \mathbb{K}_N[X] = N + 1$.

(c) L'ensemble, qu'on notera ici $\mathrm{Pol}(\mathbb{K})$, des fonctions polynomiales sur $\mathbb{K}$ est un sous-espace vectoriel de $\mathbb{K}^{\mathbb{K}}$.

(d) L'application $\begin{array}{ccc} \mathbb{K}[X] & \to & \mathrm{Pol}(\mathbb{K}) \\ P & \mapsto & f_P \end{array}$ est linéaire et surjective.

Définition

Soient $P = a_pX^p + \cdots + a_1X + a_0$ et $Q = b_qX^q + \cdots + b_1X + b_0$ deux éléments de $\mathbb{K}[X]$. On prolonge les coefficients en posant : $a_k = 0$ pour $k > p$ et $b_l = 0$ pour $l > q$.

(a) On note : $P + Q := \sum\limits_{k=0}^{\max(p,q)} (a_k + b_k)\, X^k$. $\longleftarrow$[suite $(a_k + b_k)_{k\geq 0}$]

(b) On note : $P\,Q := \sum\limits_{k=0}^{p+q} \big(\sum\limits_{i+j=k} a_ib_j \big)\, X^k$. $\longleftarrow$[suite $(c_k)_{k\geq 0}$ avec $c_k = \sum\limits_{i=0}^{k} a_ib_{k-i}$]

Concrètement :
$$\begin{aligned} PQ &= \sum_{k=0}^{p+q} (a_0b_k + a_1b_{k-1} + a_2b_{k-2} + \cdots + a_kb_0)X^k \\ &= a_pb_qX^{p+q} + (a_pb_{q-1} + a_{p-1}b_q)X^{p+q-1} + \cdots + a_0b_0. \end{aligned}$$

(c) Soit $\alpha \in \mathbb{K}$. On note : $\alpha P = (\alpha a_p)X^p + \cdots + (\alpha a_1)X + (\alpha a_0)$. $\longleftarrow$[suite $(\alpha a_k)_{k\geq 0}$]
Il s'agit du cas particulier du produit du polynôme constant α par P.

(d) Les puissances de P sont définies par récurrence : $P^0 = 1$ et $P^n = P^{n-1}P$ pour $n \geq 1$.

Remarque

On *définit* « l'indéterminée » X par : $X = (x_k)_{k\geq 0}$ où $x_0 = 0$, $x_1 = 1$ et $x_k = 0$ si $k \geq 2$.
Quand $P = X$, on a par récurrence sur $n \geq 0$: $P^n = (\delta_n(k))_{k\geq 0}$ avec $\delta_n(k) = \begin{cases} 1 & \text{si } k = n \\ 0 & \text{sinon} \end{cases}$ ou.

Il en résulte qu'un polynôme $P = (a_k)_{k\geq 0}$ à coefficients dans $\mathbb{K}$ avec $a_k = 0$ pour $k > n$ s'écrit effectivement $P = a_nX^n + \cdots + a_1X + a_0$ où X^k représente maintenant la puissance k^{e} de X qu'on vient de calculer.

Proposition

Soient $P, Q \in \mathbb{K}[X]$, $\alpha, x \in \mathbb{K}$ et $n \in \mathbb{N}$. On a posé $\deg 0 := -\infty$.

(a) On a : $(P + Q)(x) = P(x) + Q(x)$, $(PQ)(x) = P(x)Q(x)$, et $(\alpha P)(x) = \alpha P(x)$.

(b) On a : $\deg(P + Q) \leq \max(\deg(P), \deg(Q))$ et $\deg(PQ) = \deg(P) + \deg(Q)$.

(c) Si $PQ = 0$ alors $P = 0$ ou $Q = 0$.

(d) On a : $(P + Q)^n = \sum\limits_{k=0}^{n} \binom{n}{k} P^{n-k}Q^k$ « formule du binôme de Newton ».

Démonstration

(a) Immédiat.

(b) On suppose d'abord que $P = 0$ ou $Q = 0$. On a : $P + Q \in \{P, Q\}$ et $PQ = 0$. Donc $\deg(P + Q) \in \{\deg(P), \deg(Q)\}$ et $\deg(PQ) = -\infty = \deg(P) + \deg(Q)$.

On suppose maintenant que $P \neq 0$ et $Q \neq 0$:
$$P = a_pX^p + \cdots + a_1X + a_0 \text{ avec } a_p \neq 0 \text{ et } Q = b_qX^q + \cdots + b_1X + b_0 \text{ avec } b_q \neq 0.$$
Donc $P + Q = (a_r + b_r)\, X^r + (a_{r-1} + b_{r-1})\, X^{r-1} + \cdots + (a_0 + b_0)$ en posant $r = \max(p, q)$ et $PQ = \underbrace{a_pb_q}_{\neq 0} X^{p+q} + (a_pb_{q-1} + a_{p-1}b_q)X^{p+q-1} + \cdots + a_0b_0$. Cela fournit le résultat.

(c) Si $PQ = 0$, alors $\deg(P) + \deg(Q) = -\infty$ d'après (b), puis $P = 0$ ou $Q = 0$.

(d) L'égalité se démontre par récurrence comme dans le cas de $(u + v)^n$ avec $u, v \in \mathbb{C}$. □

Définition

Soient $P = a_pX^p + \cdots + a_1X + a_0$ et Q deux éléments de $\mathbb{K}[X]$.

(a) On note : $P' = pa_pX^{p-1} + \cdots + 2a_2X + a_1 \in \mathbb{K}[X]$.

On définit les dérivées successives par récurrence : $P^{(0)} = P$ et $P^{(n)} = (P^{(n-1)})'$ pour $n \geq 1$.

(b) On note : $P(Q) = a_pQ^p + \cdots + a_1Q + a_0 \in \mathbb{K}[X]$.

Exemple

On prend $P = X^2 + X + 1$ et $Q = X^2$.
On a : $P' = 2X + 1$ et $P(X^2) = (X^2)^2 + (X^2) + 1 = X^4 + X^2 + 1$.
On a aussi : $P(X) = X^2 + X + 1 = P$.

Proposition

Soient $P, Q \in \mathbb{K}[X]$, $\alpha, \beta \in \mathbb{K}$ et $n \in \mathbb{N}$.

(a) On a : $(\alpha P + \beta Q)' = \alpha P' + \beta Q'$.

(b) On a : $(PQ)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(n-k)}Q^{(k)}$ « formule de Leibniz ».

Démonstration

Exercice. □

Définition-Proposition

(a) Une *loi de composition* sur un ensemble E est une application $\star\colon E \times E \to E$.
Suivant l'usage, on notera ici $x \star y$ l'image par $\star$ de (x, y).

(b) Un *anneau* est un triplet $(A, +, \times)$, où A est un ensemble, $+$ et $\times$ sont des lois de composition sur A, avec

(i) $(x + y) + z = x + (y + z)$ pour tous $x, y, z \in A$;
(ii) il existe $0_A \in A$ tel que pour tout $x \in A$, on a : $0_A + x = x + 0_A = x$;
(iii) pour tout $x \in A$ il existe $x' \in A$ tel que : $x + x' = x' + x = 0_A$;
(iv) $x + y = y + x$ pour tous $x, y \in A$:
(v) $(x \times y) \times z = x \times (y \times z)$ pour tous $x, y, z \in A$;
(vi) $x \times (y + z) = (x \times y) + (x \times z)$ et $(x + y) \times z = (x \times z) + (y \times z)$ pour $x, y, z \in A$;
(vii) il existe $1_A \in A$ tel que pour tout $x \in A$, on a : $1_A \times x = x \times 1_A = x$.

Dans ce cas, l'élément 0_A du (ii) est unique (immédiat) appelé *zéro de* $(A, +)$, pour chaque $x \in A$ l'élément x' de A du (iii) est unique (admis) appelé *opposé de* x et noté $-x$, l'élément 1_A du (vii) est unique (immédiat) et appelé *élément unité de* A.

(c) On dit qu'un anneau $(A, +, \times)$ est *commutatif* si pour tous $x, y \in A$, on a : $x \times y = y \times x$.

(d) Un *corps* est un anneau $(K, +, \times)$ qui vérifie :

$$K \neq \{0\} \text{ et pour tout } x \in K \setminus \{0\} \text{ il existe } x' \in K \text{ tel que } x \times x' = x' \times x = 1_K.$$

Dans ce cas, pour chaque $x \in K \setminus \{0\}$ l'élément x' de K est unique (admis) et noté x^{-1}.

Démonstration

Vérification des unicités annoncées : exercice. □

Proposition

(a) L'ensemble $\mathbb{K}[X]$ muni de l'addition et de la multiplication des polynôme est un anneau commutatif de zéro le polynôme 0 et d'élément unité le polynôme constant 1.

(b) Les ensembles $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ munis des lois $+$ et $\times$ usuelles sont des corps commutatifs.

Démonstration

Exercice. □

2. Division euclidienne

Définition

Soient $A, B \in \mathbb{K}[X]$.
On dit que B *divise* A (ou A *est multiple de* B), et note $B \mid A$, si :

$$\text{il existe } Q \in \mathbb{K}[X] \text{ tel que } A = BQ.$$

Quand $B \neq 0$ et B ne divise pas A, on va disposer du reste dans la division de A par B :

Théorème

Soient $A \in \mathbb{K}[X]$ et $B \in \mathbb{K}[X] \setminus \{0\}$.
Il existe un unique $(Q, R) \in \mathbb{K}[X]^2$ tel que :

$$A = BQ + R \ \text{ et } \ \deg R < \deg B \qquad \text{« division euclidienne de } A \text{ par } B \text{ ».}$$

Dans ce cas, R s'appelle *le reste* et Q s'appelle *le quotient*, de la division de A par B.

Démonstration

• On vérifie par récurrence sur $n \in \mathbb{N}$ que :

$$(H_n) \quad \text{existence de } (Q, R) \in \mathbb{K}[X]^2 \text{ quand } \deg A < n \text{ et } B \in \mathbb{K}[X] \setminus \{0\}.$$

(i) On suppose que $n = 0$. On a : $A = 0 = B\,0 + 0$.
Cela montre que (H_0) est vraie.

(ii) On suppose que (H_n) est vraie et $\deg A \leq n$.
Si $\deg A < \deg B$: l'égalité $A = B\,0 + A$ fournit un couple (Q, R).
Si $\deg A \geq \deg B$: $A = a_pX^p + \cdots + a_0$ et $Q = b_qX^q + \cdots + b_0$ avec $a_p \neq 0$, $b_q \neq 0$, $p \geq q$.
On pose : $A_0 = A - \frac{a_p}{b_q}X^{p-q}B$ donc $\deg A_0 < \deg A$ en examinant les coefficients de X^p.
Grâce à (H_n) il existe $(Q_0, R_0) \in \mathbb{K}[X]^2$ tel que $A_0 = BQ_0 + R_0$ et $\deg R_0 < \deg B$.
D'où : $A = \frac{a_p}{b_q}X^{p-q}B + BQ_0 + R_0 = B(\frac{a_p}{b_q}X^{p-q} + Q_0) + R_0$ ce qui donne un couple (Q, R).
On en déduit que (H_{n+1}) est vraie.

(iii) Ainsi l'existence est obtenue quand $A \in \mathbb{K}[X]$ et $B \in \mathbb{K}[X] \setminus \{0\}$.

• On montre maintenant l'unicité.
On se donne $(Q, R), (Q_1, R_1) \in \mathbb{K}[X]^2$ tels que :

$$A = BQ + R = BQ_1 + R_1 \ \text{ avec } \ \deg R < \deg B \ \text{ et } \ \deg R_1 < \deg B.$$

On a : $B(Q - Q_1) = R_1 - R$. On suppose par l'absurde que $Q_1 \neq Q$.
Donc : $\deg B \leq \deg(\underbrace{B(Q - Q_1)}_{R_1 - R}) < \deg B$ <u>contradiction</u>.
Par conséquent $Q_1 = Q$ puis $R_1 = R$. □

Algorithme

On suppose que $\deg(A) \geq \deg(B)$ (sinon $Q = 0$ et $R = A$) et note :

$$A = a_pX^p + \cdots + a_1X + a_0 \ \text{ et } \ B = b_qX^q + \cdots + b_1X + b_0, \text{ où } a_p \neq 0 \text{ et } b_q \neq 0.$$

On effectue la division euclidienne de A par B ainsi :

$$
\begin{array}{cccccc|l}
 & a_pX^p + & a_{p-1}X^{p-1} & + & \cdots & + \quad a_0 & b_qX^q + b_{q-1}X^{q-1} \ \cdots \ + b_0 \\ \cline{7-7}
\ominus & a_pX^p + & \frac{a_pb_{q-1}}{b_q}X^{p-1} & + & \cdots & & \underbrace{\frac{a_p}{b_q}X^{p-q} + \frac{a'_{p-1}}{b_q}X^{p-q-1} + \cdots}_{\text{quotient } Q} \\ \cline{2-6}
= & & a'_{p-1}X^{p-1} & + & \cdots & + \quad a'_0 & \\
 & \ominus & \cdots & & \cdots & & \\ \cline{3-6}
 & = & & & \cdots & & \\
 & & & & // \quad // & // & \\ \cline{5-6}
 & & & & \multicolumn{2}{c|}{\underbrace{a''_{q-1}X^{q-1} + \cdots + a''_0}_{\text{reste } R}} &
\end{array}
$$

Exemple

On prend : $A = X^5 + X^3 + X^2 + 1$ et $B = X^4 + X^3 + 2X^2 + X + 1$ avec $\mathbb{K} = \mathbb{R}$.

Division :

$$\begin{array}{rrrrrrr|l}
 & X^5 + & 0X^4 & + X^3 & + X^2 & + 0X & + 1 & X^4 + X^3 + 2X^2 + X + 1 \\ \cline{8-8}
\ominus & X^5 + & X^4 & + 2X^3 & + X^2 & + X & & X - 1 \\ \cline{2-7}
= & & -X^4 & - X^3 & + 0X^2 & - X & + 1 & \\
\ominus & & -X^4 & - X^3 & - 2X^2 & - X & - 1 & \\ \cline{3-7}
= & & & & 2X^2 & + 0X & + 2 &
\end{array}$$

Ainsi : $A = BQ + R$ avec $Q = X - 1$ et $R = 2X^2 + 2$, qui vérifient $\underbrace{\deg R}_{2} < \underbrace{\deg B}_{4}$.

Définition

(a) Un *polynôme unitaire* dans $\mathbb{K}[X]$ est un polynôme P de la forme suivante :

$$P = a_n X^n + \cdots + a_1 X + a_0 \quad \text{avec} \quad n \in \mathbb{N},\ a_0, \ldots, a_n \in \mathbb{K} \text{ et } a_n = 1.$$

(b) Soit $P \in \mathbb{K}[X]$. On note : $P\mathbb{K}[X] = \{PQ \,;\, Q \in \mathbb{K}[X]\}$.

(c) Soient $A, B \in \mathbb{K}[X]$. On note : $A\mathbb{K}[X] + B\mathbb{K}[X] = \{AU + BV \,;\, U, V \in \mathbb{K}[X]\}$.

Lemme

On appelle ici « idéal bilatère » d'un anneau $(A, +, \times)$ une partie $\mathscr{I}$ de A vérifiant :

(i) $0_A \in \mathscr{I}$;
(ii) pour tous $x, y \in \mathscr{I}$, on a : $x - y \in \mathscr{I}$;
(iii) pour tous $x \in \mathscr{I}$ et $a \in A$, on a : $x \times a \in \mathscr{I}$ et $a \times x \in \mathscr{I}$.

Les idéaux bilatères de $(\mathbb{K}[X], +, \times)$ sont les ensembles $P\mathbb{K}[X]$ avec $P \in K[X]$.
Ce sont aussi les ensembles distincts $P\mathbb{K}[X]$ avec $P \in \mathbb{K}[X]$ unitaire et $\underbrace{\{0\}}_{0\,\mathbb{K}[X]}$.

Démonstration

• Soit $P \in K[X]$. L'ensemble $P\mathbb{K}[X]$ est un idéal bilatère de $\mathbb{K}[X]$ car :
(i) $0 = P0 \in P\mathbb{K}[X]$;
(ii) pour tous $PQ, PR \in P\mathbb{K}[X]$, on a : $PQ - PR = P(Q - R) \in P\mathbb{K}[X]$;
(iii) pour tous $PQ \in P\mathbb{K}[X]$ et $a \in A$, on a : $\underbrace{PQA}_{APQ} \in P\mathbb{K}[X]$.

• Soit $\mathscr{I}$ un idéal bilatère de $(\mathbb{K}[X], +, \times)$.
On a : $0 \in \mathscr{I}$. Si $\mathscr{I} = \{0\}$: $\mathscr{I} = 0\,\mathbb{K}[X]$.
On suppose que $\mathscr{I} \neq \{0\}$. On fixe $P \in \mathscr{I} \setminus \{0\}$ de degré minimal.
– On a : $P \in \mathscr{I}$ donc $P\mathbb{K}[X] \subseteq \mathscr{I}$.
– Soit $A \in \mathscr{I}$. Division euclidienne : $A = BQ + R$ avec $\deg R < \deg P$.
On a : $R = A - PQ \in \mathscr{I}$ et $\deg R < \deg P$ donc $R = 0$ puis $A \in P\mathbb{K}[X]$.
– Ainsi, en fixant $k \in \mathbb{K} \setminus \{0\}$ tel que kP est unitaire, on a : $\mathscr{I} = P\mathbb{K}[X] = kP\mathbb{K}[X]$.

• Tout d'abord quand $P \in \mathbb{K}[X]$ est unitaire, on a : $P \in P\mathbb{K}[X]$ donc $P\mathbb{K}[X] \neq \{0\}$.
Soient $P, Q \in \mathbb{K}[X]$ unitaires tels que $P\mathbb{K}[X] = Q\mathbb{K}[X]$.
Il existe $C, D \in \mathbb{K}[X]$ tels que $P = CQ$ et $Q = DP$, On a : $\deg P = \deg C + \deg Q$ et $\deg Q = \deg D + \deg P$, donc $\deg C = \deg D = 0$ et $P = Q$. □

Définition-Proposition

Soient $A, B \in \mathbb{K}[X]$.

(a) Il existe un unique $D \in \mathbb{K}[X]$ unitaire ou nul qui divise A et B, et tel que tout diviseur dans $\mathbb{K}[X]$ de A et B divise D.
Il est caractérisé par : $A\mathbb{K}[X] + B\mathbb{K}[X] = D\mathbb{K}[X]$.
On le note : $D = \operatorname{pgcd}(A, B)$ car, quand $A \neq 0$ et $B \neq 0$, D est le diviseur unitaire de degré maximum dans $\mathbb{K}[X]$ commun à A et B.

(b) Il existe un unique $M \in \mathbb{K}[X]$ unitaire ou nul qui est multiple de A et B, et tel que tout multiple dans $\mathbb{K}[X]$ de A et B est multiple de M.

Il est caractérisé par : $A\mathbb{K}[X] \cap B\mathbb{K}[X] = M\mathbb{K}[X]$.

On le note : $M = \text{ppcm}(A,B)$ car, quand $A \neq 0$ et $B \neq 0$, M est le multiple unitaire de degré minimum dans $\mathbb{K}[X]$ commun à A et B.

(c) Il existe $\lambda \in \mathbb{K} \setminus \{0\}$ tel que : $\boxed{AB = \lambda MD}$.

Démonstration

On constate que $A\mathbb{K}[X]+B\mathbb{K}[X]$ et $A\mathbb{K}[X]\cap B\mathbb{K}[X]$ sont des idéaux bilatères de $(\mathbb{K}[X],+,\times)$.

D'après le lemme, il existe D, M unitaires ou nuls uniques tels que :

$$A\mathbb{K}[X] + B\mathbb{K}[X] = D\mathbb{K}[X] \quad \text{et} \quad A\mathbb{K}[X] \cap B\mathbb{K}[X] = M\mathbb{K}[X].$$

(a) On a : $A = A1 + B0 \in D\mathbb{K}[X]$ et $B = A0 + B1 \in D\mathbb{K}[X]$ donc $D \mid A$ et $D \mid B$.
Ill existe $U_0, V_0 \in \mathbb{K}[X]$ tels que $D = AU_0 + BV_0$, donc tout diviseur de A et B divise D.

Si D_1 et D_2 conviennent, alors $D_1 \mid D_2$ et $D_2 \mid D_1$ puis $D_1\mathbb{K}[X] = D_2\mathbb{K}[X]$ et $D_1 = D_2$.

(b) On a : $M \in A\mathbb{K}[X] \cap B\mathbb{K}[X]$ donc M est multiple de A et B.
Tout multiple de A et B est dans $A\mathbb{K}[X] \cap B\mathbb{K}[X]$ donc dans $M\mathbb{K}[X]$ donc est multiple de M.

Si M_1 et M_2 conviennent, alors $M_2 \mid M_1$ et $M_1 \mid M_2$ puis $M_1\mathbb{K}[X] = M_2\mathbb{K}[X]$ et $M_1 = M_2$.

(c) Il existe $A_0, B_0, U_0, V_0 \in \mathbb{K}[X]$ tels que : $A = A_0D$, $B = B_0D$, $D = AU_0 + BV_0$.
On a : A_0B_0D est multiple de A et B puis de M, donc $AB = (A_0B_0D)D$ est multiple de MD.
On a aussi : AB divise AM et BM, donc AB divise $MAU_0 + MBV_0$ qui vaut MD.
Vu que $AB \in MD\mathbb{K}[X]$ et $MD \in AB\mathbb{K}[X]$, on a $AB\mathbb{K}[X] = MD\mathbb{K}[X]$ donc il existe $\lambda \in \mathbb{K}$ tel que le polynôme unitaire $\frac{1}{\lambda}AB$ est égal à MD. □

Remarques

Soient $A, B \in \mathbb{K}[X] \setminus \{0\}$. On pose $D = \text{pgcd}(A,B)$, donc $D \neq 0$.

1. On se donne $P \in \mathbb{K}[X]$ unitaire. On a : $PA\mathbb{K}[X] + PB\mathbb{K}[X] = PD\mathbb{K}[X]$,
Donc : $\text{pgcd}(PA, PB) = P\,\text{pgcd}(A,B)$. De même : $\text{ppcm}(PA, PB) = P\,\text{ppcm}(A,B)$.

2. On introduit $A_0, B_0 \in \mathbb{K}[X]$ tels que $A = DA_0$ et $B = DB_0$.
On a d'après 1 : $D = \text{pgcd}(DA_0, DB_0) = D\,\text{pgcd}(A_0, B_0)$.
D'où : $\text{pgcd}(A_0, B_0) = 1$.

Définition

On dit que $A, B \in \mathbb{K}[X]$ sont *premiers entre eux* si : $\text{pgcd}(A,B) = 1$.

Théorème

Soient $A, B \in \mathbb{K}[X]$. Les polynômes A et B sont premiers entre eux si et seulement si il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$ « théorème de Bézout ».

Démonstration

On a, d'après la définition-proposition ci-dessus :

$$\begin{aligned}\text{pgcd}(A,B) = 1 &\iff \mathbb{K}[X] = A\mathbb{K}[X] + B\mathbb{K}[X] \iff \mathbb{K}[X] \subseteq A\mathbb{K}[X] + B\mathbb{K}[X] \\ &\iff 1 \in A\mathbb{K}[X] + B\mathbb{K}[X] \iff \exists U, V \in \mathbb{K}[X] \quad AU + BV = 1.\end{aligned}$$ □

Proposition

Soient $A, B \in \mathbb{K}[X] \setminus \{0\}$. On pose $R_0 = A$ et $R_1 = B$. ←[$\text{pgcd}(A,0) = |A|$, $\text{pgcd}(0,B) = |B|$, $\text{pgcd}(0,0) = 0$]

On effectue les divisions euclidiennes successives de R_k par R_{k+1} :

$A = BQ_2 + R_2$ avec $R_2 \neq 0$ et $\deg R_2 < \deg R_1$

$B = R_2Q_3 + R_3$ avec $R_3 \neq 0$ et $\deg R_3 < \deg R_2$

...

$R_{n-2} = R_{n-1}Q_n + R_n$ avec $R_n \neq 0$ et $\deg R_n < \deg R_{n-1}$

$R_{n-1} = R_nQ_{n+1} + 0$. ←[on atteint 0 car les degrés des restes décroissent strictement dans $\mathbb{N}$]

Dans ce cas, il existe $k \in \mathbb{K} \setminus \{0\}$ tel que : $\boxed{\text{pgcd}(A,B) = kR_n}$ « algorithme d'Euclide ».

Démonstration

On utilise la définition du pgcd en termes de diviseurs :

$$\mathrm{pgcd}(\underbrace{A}_{R_0},\underbrace{B}_{R_1})\underbrace{=}_{\substack{\text{mêmes}\\ \text{diviseurs communs}}}\mathrm{pgcd}(\underbrace{A-R_1Q_2}_{R_2},R_1)=\mathrm{pgcd}(R_1,R_2)=\cdots=\underbrace{\mathrm{pgcd}(R_n,0)}_{\in(\mathbb{K}\setminus\{0\})R_n}.$$ □

Remarque (« couple de Bézout »)

L'algorithme d'Euclide fournit aussi $(U,V)\in\mathbb{K}[X]^2$ tels que $AU+BV=\mathrm{pgcd}(A,B)$:

$$\begin{array}{rcll}
A-BQ_2 & = & \cancel{R_2} & \times(\ldots)\quad \text{pour équilibrer les } R_2 \text{ sur les 3 premières lignes}\\
B-\cancel{R_2Q_3} & = & \cancel{R_3} & \times(\ldots)\quad \text{pour équilibrer les } R_3 \text{ sur les 3 lignes suivantes}\\
\ldots & & & \\
\cancel{R_{n-3}}-\cancel{R_{n-2}Q_{n-1}} & = & \cancel{R_{n-1}} & \times(-Q_{n-1})\quad \text{pour équilibrer les } R_{n-1} \text{ sur les 2 dernières lignes}\\
\cancel{R_{n-2}}-\cancel{R_{n-1}Q_n} & = & R_n & \\
\hline
\frac{1}{k}(AU+BV) & = & R_n &
\end{array}$$

On pose : $R_0=A$ et $R_1=B$. Ainsi : $R_{i+2}=R_i-R_{i+1}\,Q_{i+2}$ pour $0\le i\le n-2$.
On construit une suite $((U_i,V_i))_{0\le i\le n}$ par :
(i) $(U_0,V_0)=(1,0)$ et $(U_1,V_1)=(0,1)$;
(ii) $U_{i+2}=U_i-Q_{i+2}\,U_{i+1}$ et $V_{i+2}=V_i-Q_{i+2}\,V_{i+1}$ pour $0\le i\le n-2$.
On a par récurrence : $R_i=AU_i+BV_i$ pour $0\le i\le n$, donc $A\,U_n+BV_n=\mathrm{pgcd}(A,B)$.

Exemple

On prend $A=X^5+X^3+X^2+1$ et $B=X^4-X^2-2$.

Calcul du pgcd de A et B ? Couple de Bézout associé ?

On a :

$$\begin{array}{rcl}
A & = & BX+R \qquad \text{où } R=2X^3+X^2+2X+1\\
B & = & R(\frac{1}{2}X-\frac{1}{4})-\frac{7}{4}X^2-\frac{7}{4}\\
R & = & (X^2+1)(2X+1)+0
\end{array}$$

donc $\mathrm{pgcd}(A,B)=X^2+1$.

Ensuite :

$$\begin{array}{rcl}
A-BX & = & \cancel{R} \qquad \times(-2X+1)\quad \text{équilibre les } R\\
4B-\cancel{(2X-1)R} & = & -7(X^2+1)\\
\hline
(-2X+1)A+(2X^2-X+4)B & = & -7(X^2+1)
\end{array}$$

II. Racines d'un polynôme

1. Racines avec leur multiplicité

Définition

Soient $P\in\mathbb{K}[X]$ et $r\in\mathbb{K}$.
On dit que r *est racine de* P (ou *zéro de la fonction* $x\mapsto P(x)$) si $P(r)=0$.

Exemple

• Quelles sont les racines « évidentes » (rationnelles) du polynôme réel $2X^3-5X^2-3X+9$?

• On considère plus généralement : $P=a_nX^n+\cdots+a_1X+a_0$ où $n\ge 1$ et $a_0,\ldots,a_n\in\mathbb{Z}$.
On suppose que $r=\frac{p}{q}\in\mathbb{Q}$ vérifie $P(r)=0$ avec p et q de seuls diviseurs communs -1 et 1.

On a : $a_np^n+a_{n-1}p^{n-1}q+\cdots+a_1pq^{n-1}+a_0q^n=0$

donc $\begin{cases} a_0q^n=p(-a_np^{n-1}-\cdots-a_1q^{n-1}) & \text{puis} \quad p\mid a_0q^n\\ a_np^n=q(-a_{n-1}p^{n-1}-\cdots-a_0q^{n-1}) & \text{puis} \quad q\mid a_np^n\end{cases}$

donc, par le lemme de Gauss : $p\mid a_0$ et $q\mid a_n$.

• Les racines rationnelles de $2X^3-5X^2-3X+9$ sont dans $\{\pm1,\pm3,\pm9,\pm\frac{1}{2},\pm\frac{3}{2},\pm\frac{9}{2}\}$. On constate que $\frac{3}{2}$ est effectivement une racine de $2X^3-5X^2-3X+9$.

Proposition

Soient $P \in \mathbb{K}[X]$ et $r \in \mathbb{K}$.
On a : r est racine de P si et seulement si $X - r$ divise P.

Démonstration

($\Leftarrow$) On suppose que $X - r$ divise P.
Il existe $Q \in \mathbb{K}[X]$ tel que $P = (X - r)Q$. Donc $P(r) = 0$.

($\Rightarrow$) On suppose que $P(r) = 0$.
Division euclidienne : $P = (X - r)Q + R$ avec $(Q, R) \in \mathbb{K}[X]^2$ et $\deg R < 1$.
Donc $P(r) = R(r)$ et R est constant. D'où $R = 0$, ce qui montre que $X - r$ divise P. □

Corollaire 1

Soient $P, Q \in \mathbb{K}[X]$. Les racines communes à P et Q sont les racines de $\mathrm{pgcd}(P, Q)$.

Corollaire 2

Un polynôme non-nul $P \in \mathbb{K}[X] \setminus \{0\}$ de degré n a au plus n racines dans $\mathbb{K}$.

(On verra plus tard que quand P a n racines comptées avec multiplicité $(x_1, \alpha_1), ..., (x_p, \alpha_p)$, où $p \in \mathbb{N}$ et $\alpha_1, ..., \alpha_p \in \mathbb{N} \setminus \{0\}$ vérifient $\alpha_1 + ... + \alpha_p = n$, il existe $\lambda \in \mathbb{K} \setminus \{0\}$ tel que :
$$P = \lambda(X - x_1)^{\alpha_1} ... (X - x_p)^{\alpha_p}.)$$

Démonstration

Récurrence sur $n \in \mathbb{N}$.

(i) Si $n = 0$: P est constant non-nul donc n'a pas de racine.

(ii) Soit $n \geq 0$ tel que tout polynôme non-nul de degré n a au plus n racines.
On se donne un polynôme non-nul P de degré $n + 1$. On suppose que P possède au moins une racine r (sinon il n'y a rien à faire). Il existe $Q \in \mathbb{K}[X]$ tel que $P = (X - r)Q$. On applique l'hypothèse de récurrence à Q et constate que P a au plus $n + 1$ racines.

(iii) D'où le résultat. □

Remarque

On suppose ici que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ $^{(*)}$.

Soit $P \in \mathbb{K}[X] \setminus \{0\}$. En notant $d = \deg P$, le polynôme P a au plus d racines. En particulier, comme on l'a déjà vu dans le cas réel, l'application $f_P : x \in \mathbb{K} \mapsto P(x) \in \mathbb{K}$ est non-nulle.

Par contre quand $\mathbb{K}$ est fini, le polynôme non-nul $P := \prod_{x \in \mathbb{K}} (X - x)$ vérifie $f_P = 0$.

Proposition (hors programme)

Soient $P = a_n X^n + \cdots + a_1 X + a_0 \in \mathbb{K}[X]$ avec $a_n \neq 0$ et $x_1, ..., x_n \in \mathbb{K}$.

Il existe $\lambda \in \mathbb{K} \setminus \{0\}$ tel que $P = \lambda(X - x_1)...(X - x_n)$ si et seulement si :
$x_1 + ... + x_n = -\frac{a_{n-1}}{a_n}$ et $\sum_{1 \leq i < j \leq n} x_i x_j = \frac{a_{n-2}}{a_n}$ et $\sum_{1 \leq i < j < k \leq n} x_i x_j x_k = -\frac{a_{n-3}}{a_n}$ et ... et $x_1 ... x_n = (-1)^n \frac{a_0}{a_n}$.

Démonstration

Il suffit de traduire l'égalité $P = \lambda(X - x_1)...(X - x_n)$ en identifiant successivement les coefficients de X^n, X^{n-1}, X^{n-2}, ..., X^0. □

Définition-Proposition

Soient $P \in \mathbb{K}[X] \setminus \{0\}$ et $r \in \mathbb{K}$.

(a) On appelle *ordre de multiplicité de r dans P* l'unique $m \in \mathbb{N}$ tel que :
$$(X - r)^m \text{ divise } P \quad \text{et} \quad (X - r)^{m+1} \text{ ne divise pas } P \text{ (existe).}$$
$m \geq 1$: « r est racine de P » ;
$m = 1$: « r est racine simple de P » ;

$(*)$ (ou plus généralement que $\mathbb{K}$ est un corps commutatif infini)

$m=2$: « r est racine double de P » ;
$m \geq 2$: « r est racine multiple de P ».

(b) Lorsque $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$ $^{(**)}$l'ordre de multiplicité m de r dans P est caractérisé par :

$$P(r)=P'(r)=\ldots=P^{(m-1)}(r)=0 \quad \text{et} \quad P^{(m)}(r)\neq 0.$$

En particulier, on a : $m\geq 2 \iff P(r)=P'(r)=0$.

Par convention : l'ordre de multiplicité de r dans 0 est $+\infty$ (dérivées de 0 nulles en r).

Démonstration

(a) Le polynôme $(X-r)^0$ divise P et $(X-r)^k$ ne divise pas P quand $k>\deg P$. Ainsi il existe un plus grand $m\in\mathbb{N}$ tel que $(X-r)^m$ divise P.

(b) On suppose que $\underbrace{1+\ldots+1}_{n \text{ fois}}\neq 0$ pour tout $n\in\mathbb{N}\setminus\{0\}$ (par exemple $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$).

On montre $P(r)=P'(r)=\ldots=P^{(m-1)}(r)=0$ et $P^{(m)}(r)\neq 0$ (ce qui déterminera m).
Il existe $Q\in\mathbb{K}[X]$ tel que $P=(X-r)^mQ$. Donc $Q(r)\neq 0$ car $(X-r)^{m+1}$ ne divise pas P.
Formule de Leibniz : $((X-r)^mQ)^{(p)}=\sum_{k=0}^{p}\binom{p}{k}\underbrace{((X-r)^m)^{(k)}}_{m\cdots(m-k+1)(X-r)^{m-k}}Q^{(p-k)}$ pour $p\in\{0,\ldots,m\}$.

D'où : $P(r)=P'(r)=\ldots=P^{(m-1)}(r)=0$ et $P^{(m)}(r)=\underbrace{m!\,Q(r)}_{\neq 0}$. □

Remarque

Par contre, lorsque $\mathbb{K}=\mathbb{Z}/2\mathbb{Z}$, l'ordre de multiplicité de 0 dans X^2 est 2 bien que :

$$(X^2)(0)=(X^2)'(0)=(X^2)''(0)=0.$$

Exemple ($\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$)

On prend $P=X^3+pX+q$ avec $p,q\in\mathbb{K}$.

La division euclidienne de P par P' donne : $P=P'\times(\frac{1}{3}X)+\frac{2}{3}pX+q$. D'où, pour $r\in\mathbb{K}$:

$$\begin{cases}P(r)=0\\P'(r)=0\end{cases}\iff\begin{cases}\overbrace{P'(r)}^{3r^2+p}=0\\\frac{2}{3}pr+q=0\end{cases}\iff\begin{cases}pr=-\frac{3}{2}q\\3r^2+p=0\end{cases}\iff\left(\begin{cases}p\neq 0\\r=-\frac{3q}{2p}\\3(-\frac{3q}{2p})^2+p=0\end{cases}\text{ou}\quad\begin{cases}p=0\\q=0\\r=0\end{cases}\right).$$

Ainsi P a une racine multiple dans $\mathbb{K}$ si et seulement si : $4p^3+27q^2=0$.

(Quand $p\neq 0$ et $q\neq 0$, on aurait aussi pu calculer $\mathrm{pgcd}(P,P')$ par l'algorithme d'Euclide :
$\underbrace{P}_{R_0}=\underbrace{P'}_{R_1}\times(\frac{1}{3}X)+\underbrace{\frac{2p}{3}X+q}_{R_2}$, $R_1=R_2\times(\frac{9}{2p}X)+\underbrace{-\frac{9q}{2p}X+p}_{R_3}$, $R_2=R_3\times(-\frac{4p^2}{27q})+\underbrace{\frac{4p^3}{27q}+q}_{R_4}$.)

2. Polynômes irréductibles

Définition

Un *polynôme irréductible à coefficients dans* $\mathbb{K}$ est un élément P de $\mathbb{K}[X]$ non-constant tel que les seules égalités $P=AB$ avec $A,B\in\mathbb{K}[X]$ s'obtiennent quand A ou B est constant.

Remarques (immédiates)

1. Tout polynôme de degré 1 est irréductible.
(Si $\deg P=1$ et $P=AB$, on a : $\deg A+\deg B=1$ donc $\deg A=0$ ou $\deg B=0$.)

2. Un polynôme à coefficients dans $\mathbb{K}$ de degré 2 ou 3 est irréductible si et seulement si il n'a aucune racine dans $\mathbb{K}$.

$(**)$ (ou plus généralement quand $\mathbb{K}$ est un corps commutatif dans lequel $\underbrace{1+\ldots+1}_{n \text{ fois}}\neq 0$ pour tout $n\in\mathbb{N}\setminus\{0\}$)

$$\left(\begin{array}{l}\text{Si } \deg P \in \{2,3\} \text{ et } P = AB \text{ avec } A \text{ et } B \text{ non-constants :}\\ \underbrace{\deg A}_{\geq 1} + \underbrace{\deg B}_{\geq 1} \leq 3 \text{ donc } A \text{ ou } B \text{ est de degré 1 de la forme } \lambda(X-r) \text{ avec } \lambda, r \in \mathbb{K}.\\ \text{Si } \deg P \in \{2,3\} \text{ et } P = (X-r)\,C \text{ avec } r \in \mathbb{K} \text{ et } C \in \mathbb{K}[X] :\\ \deg(X-r) = 1 \text{ et } \deg C \in \{1,2\} \text{ donc } X-r \text{ et } C \text{ sont non-constants.}\end{array}\right)$$

Par exemple :

– X^2+1 n'est pas irréductible dans $\mathbb{C}[X]$;
– X^2+1 est irréductible dans $\mathbb{R}[X]$ et dans $\mathbb{Q}[X]$;
– X^2-2 n'est pas irréductible dans $\mathbb{R}[X]$;
– X^2-2 est irréductible dans $\mathbb{Q}[X]$ (exemple du début du paragraphe 1).

Proposition

(a) Rappel. Soient $A, B \in \mathbb{K}[X]$ sans diviseur commun non constant.
Il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$ « théorème de Bézout ».

(b) Soient $A, B \in \mathbb{K}[X]$ et $P \in \mathbb{K}[X]$ irréductible.
Si $P \mid AB$, alors $P \mid A$ ou $P \mid B$ « lemme d'Euclide ».

(c) Soient $A, B, C \in \mathbb{K}[X]$.
Si $A \mid BC$ et $\operatorname{pgcd}(A,B) = 1$, alors $A \mid C$ « lemme de Gauss » (généralise celui d'Euclide).

Démonstration

(a) Tout d'abord : $A \neq 0$ et $B \neq 0$. On raisonne par récurrence sur $n := \deg A + \deg B$.

(i) Si $n = 0$: $B = b_0$ puis on prend $U = 1$ et $V = \frac{1}{b_0}(1-A)$.

(ii) Soit $n \geq 1$ tel que le résultat est acquis au niveau de $n-1$.
Supposons par exemple que $\deg A \geq \deg B$.
On effectue la division euclidienne de A par B : $A = QB + R$.
Si $R = 0$: B est un diviseur commun de A et B donc B est constant, et on fait comme au (i).
Si $R \neq 0$: $\deg B > \deg R$ où R et B n'ont pas de diviseur commun non constant, ce qui donne $U_0R + BV_0 = 1$ puis $U_0A + B(V_0 - QU_0) = 1$ enfin on prend $U := U_0$ et $V := V_0 - QU_0$.

(b) On suppose que $P \mid AB$, et $P \nmid A$.
Comme P est irréductible, le seul diviseur unitaire de P et A est 1. Donc $\operatorname{pgcd}(A,P) = 1$.
Bézout : il existe $U, V \in \mathbb{K}[X]$ tels que $AU + PV = 1$, donc $\underbrace{BA}_{\text{divisible par } P}U + BPV = B$ et $P \mid B$.

$$\left(\begin{array}{l}\text{Variante sans utiliser le théorème de Bézout.}\\ \text{On suppose par l'absurde qu'il existe } P \in \mathbb{K}[X] \text{ irréductible et } A, B \in \mathbb{K}[X] \text{ non multiples}\\ \text{de } P \text{ tels que } AB \text{ est multiple de } P \text{ (a fortiori } A \text{ et } B \text{ sont non-constants).}\\ \text{On fixe } P \text{ et } A, \text{ et choisit } B \text{ comme ci-dessus de degré minimal.}\\ \text{Division euclidienne : } B = PQ_1 + B_1 \text{ avec } \deg B_1 < \deg P \text{ donc } B_1 \text{ est non multiple de } P \text{ et}\\ AB_1 \text{ est multiple de } P, \text{ donc } \deg B \leq \deg B_1 < \deg P \text{ par minimalité du degré de } B.\\ \text{Division euclidienne : } P = BQ_2 + B_2 \text{ avec } B_2 \neq 0 \text{ (} P \text{ irréductible) et } \deg B_2 < \deg B < \deg P.\\ \text{Ainsi } B_2 \text{ n'est pas multiple de } P \text{ et } AB_2 = AP - Q_2AB \text{ est multiple de } P.\\ \text{Cela contredit la minimalité du degré de } B.\end{array}\right)$$

(c) On suppose que $A \mid BC$ et $\operatorname{pgcd}(A,B) = 1$.
Bézout : il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$, donc $AUC + \overbrace{BC}^{\text{divisible par } A}V = B$ et $A \mid C$. □

Théorème

Soit $P \in \mathbb{K}[X] \setminus \{0\}$.

(a) Il existe $\lambda \in \mathbb{K} \setminus \{0\}$, $k \in \mathbb{N}$, des polynômes unitaires irréductibles distincts P_1, ..., P_k et $\alpha_1, ..., \alpha_k \in \mathbb{N} \setminus \{0\}$, tels que : $P = \lambda \underbrace{P_1^{\alpha_1} \cdots P_k^{\alpha_k}}_{1 \text{ quand } k=0}$.

(b) La décomposition du (a) est « unique à l'ordre près »$^{(\star)}$ au sens où pour toute autre décomposition $P = \mu\, Q_1^{\beta_1} \cdots Q_l^{\beta_l}$, on a : $\mu = \lambda$ et $\{(Q_1,\beta_1), ..., (Q_l,\beta_l)\} = \{(P_1,\alpha_1), ..., (P_k,\alpha_k)\}$.

(c) Quand P est décomposé comme au (a), les éléments de $\mathbb{K}[X]$ qui divisent P sont les polynômes de la forme $\mu\, P_1^{\beta_1} \cdots P_k^{\beta_k}$ avec $\mu \in \mathbb{K} \setminus \{0\}$ et $0 \leq \beta_1 \leq \alpha_1$, ..., $0 \leq \beta_k \leq \alpha_k$.

$(\star)$ (les autres décompositions s'obtiennent en « changeant l'ordre dans lequel on écrit (P_1,α_1), ..., (P_k,α_k) »)

Démonstration

(a) On montre l'existence d'une décomposition de P en polynômes irréductibles.
Récurrence sur $n \geq 0$: $\ (H_n)\ $ existence d'une décomposition en polynômes irréductibles pour tous les polynômes non-nuls de degré $\leq n$.

(i) Cas $n = 0$: $k = 0$ convient. Donc (H_0) est vraie.

(ii) On suppose que (H_n) est vraie et se donne $Q \in \mathbb{K}[X] \setminus \{0\}$ de degré $n+1$:
– soit Q est irréductible et se décompose en $\mu \widetilde{Q}^1$ où $Q = \mu \widetilde{Q}$ avec $\widetilde{Q}$ unitaire irréductible ;
– soit Q n'est pas irréductible et il existe $A, B \in \mathbb{K}[X]\setminus\{0\}$ tels que $Q = AB$, puis en appliquant (H_n) au niveau de A et de B on décompose Q en polynômes irréductibles.

On en déduit que (H_{n+1}) est vraie.

(iii) Ainsi, pour chaque $n \geq 1$, (H_n) est vraie et en particulier P possède une décomposition en polynômes irréductibles.

(b) On montre l'unicité de la décomposition en partant de $P = \lambda P_1^{\alpha_1} \cdots P_k^{\alpha_k} = \mu Q_1^{\beta_1} \cdots Q_l^{\beta_l}$ avec $\lambda, \mu \in \mathbb{K} \setminus \{0\}$, $\underbrace{P_1, ..., P_k}_{\text{distincts}}$, $\underbrace{Q_1, ..., Q_l}_{\text{distincts}}$ unitaires irréductibles et $\alpha_1, ..., \alpha_k, \beta_1, ..., \beta_l \in \mathbb{N} \setminus \{0\}$.

Le lemme d'Euclide montre que chaque P_j $(1 \leq j \leq k)$ divise un des polynômes $Q_1, ..., Q_l$, noté Q_{i_j}, puis $P_j = Q_{i_j}$ par irréductibilité de Q_{i_j}. Donc : $\{P_1, ..., P_k\} \subseteq \{Q_1, ..., Q_l\}$. De même : $\{Q_1, ..., Q_l\} \subseteq \{P_1, ..., P_k\}$. D'où $l = k$. Après simplification : $\lambda P_1^{\alpha_1} ... P_k^{\alpha_k - \beta_{i_k}} = \mu Q_{i_1}^{\beta_{i_1}} ... Q_{i_{k-1}}^{\beta_{i_{k-1}}}$ si $\alpha_k > \beta_{i_k}$, ou, $\lambda P_1^{\alpha_1} ... P_{k-1}^{\alpha_{k-1}} = \mu Q_{i_1}^{\beta_{i_1}} ... Q_{i_k}^{\beta_{i_k} - \alpha_k}$ si $\alpha_k < \beta_{i_k}$, chacun de ces deux cas menant à une contradiction comme ci-dessus. Ainsi $\alpha_k = \beta_{i_k}$. Ensuite $\alpha_{k-1} = \beta_{i_{k-1}}$, ..., $\alpha_1 = \beta_{i_1}$, $\mu = \nu$.

(c) On cherche maintenant les diviseurs de P. Ceux proposés conviennent.

Soit $B \in \mathbb{K}[X]$ qui divise P. Donc $B \neq 0$ et il existe $C \in \mathbb{K}[X] \setminus \{0\}$ tel que $P = BC$. Quitte à ajouter des Q_i^0, on peut écrire $B = \mu Q_1^{\beta_1} ... Q_l^{\beta_l}$ et $C = \nu Q_1^{\gamma_1} ... Q_l^{\gamma_l}$, avec $\lambda, \mu \in \mathbb{K} \setminus \{0\}$, $Q_1, ..., Q_l$ unitaires irréductibles distincts et $\beta_1, ..., \beta_l, \gamma_1, ..., \gamma_l \in \mathbb{N}$, $\beta_1 + \gamma_1 \neq 0$, ..., $\beta_l + \gamma_l \neq 0$. Donc $P = \mu\nu\, Q_1^{\beta_1+\gamma_1} ... Q_l^{\beta_l+\gamma_l}$ puis par unicité de la décomposition en polynômes irréductibles : $l = k$ et $Q_1 = P_{j_1}$ et ... et $Q_k = P_{j_k}$ et $\beta_1 + \gamma_1 = \alpha_{j_1}$ et ... et $\beta_k + \gamma_k = \alpha_{j_k}$. D'où le résultat. □

Corollaire

Soient $A, B \in \mathbb{K}[X] \setminus \{0\}$. On peut donc les écrire $A = \lambda P_1^{\alpha_1} ... P_k^{\alpha_k}$ et $B = \mu P_1^{\beta_1} ... P_l^{\beta_k}$ avec $\lambda, \mu \in \mathbb{K} \setminus \{0\}$, $P_1, ..., P_k \in \mathbb{K}[X]$ unitaires irréductibles et $\alpha_1, ..., \alpha_k, \beta_1, ..., \beta_k \in \mathbb{N}$.
On a : $\operatorname{pgcd}(A, B) = P_1^{\min(\alpha_1, \beta_1)} \cdots P_k^{\min(\alpha_k, \beta_k)}$ et $\operatorname{ppcm}(A, B) = P_1^{\max(\alpha_1, \beta_1)} \cdots P_k^{\max(\alpha_k, \beta_k)}$.

Démonstration

D'après le théorème, le diviseur commun de degré maximal est $P_1^{\min(\alpha_1, \beta_1)} \cdots P_k^{\min(\alpha_k, \beta_k)}$.
La formule pour le ppcm découle par exemple de : $AB = \lambda \operatorname{ppcm}(A, B) \operatorname{pgcd}(A, B)$. □

Remarque

Soient $P = \lambda P_1^{\alpha_1} \cdots P_k^{\alpha_k} \in \mathbb{K}[X] \setminus \{0\}$ comme ci-dessus et $r \in \mathbb{K}$.

On a : r est racine de P si et seulement si $X - r$ est l'un des P_i, $1 \leq i \leq k$.
Dans ce cas α_i est la multiplicité de la racine r de P.

(D'après le (c) du théorème, la multiplicité de r dans P est le plus grand $m \in \mathbb{N}$ tel que $(X - r)^m$ s'écrive $\mu P_1^{\beta_1} \cdots P_k^{\beta_k}$ avec $\mu \in \mathbb{K} \setminus \{0\}$ et $0 \leq \beta_1 \leq \alpha_1$, ..., $0 \leq \beta_k \leq \alpha_k$.
Or $X - r$ est unitaire irréductible. On utilise l'unicité de la décomposition.)

Théorème (« théorème de d'Alembert-Gauss »)

Tout polynôme non-constant à coefficients dans $\mathbb{C}$ a au moins une racine dans $\mathbb{C}$.

Démonstration

Admise (difficile). □

Corollaire (important)

(a) Les polynômes irréductibles dans $\mathbb{C}[X]$ sont les polynômes de degré 1.

En particulier, tout $P \in \mathbb{C}[X] \setminus \{0\}$ s'écrit de manière unique à l'ordre près sous la forme

$$P = \lambda\,(X - r_1)^{\alpha_1} ... (X - r_m)^{\alpha_m}$$

avec $\lambda \in \mathbb{C} \setminus \{0\}$, $\underbrace{\{r_1, ..., r_m\}}_{\text{distincts}} \subseteq \mathbb{C}$, et $\alpha_1, ..., \alpha_m \in \mathbb{N} \setminus \{0\}$.

(b) Les polynômes irréductibles dans $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 dont le discriminant est strictement négatif.

En particulier, tout $P \in \mathbb{R}[X] \setminus \{0\}$ s'écrit de manière unique à l'ordre près sous la forme

$$P = \lambda\,(X - r_1)^{\alpha_1} ... (X - r_m)^{\alpha_m} (X^2 + u_1 X + v_1)^{\beta_1} ... (X^2 + u_n X + v_n)^{\beta_n}$$

avec $\lambda \in \mathbb{R} \setminus \{0\}$, $\underbrace{\{r_1, ..., r_m\}}_{\text{distincts}} \subseteq \mathbb{R}$, $\underbrace{\{(u_1, v_1), ..., (u_n, v_n)\}}_{\text{couples distincts}} \subseteq \mathbb{R}^2$ tels que les $X^2 + u_k X + v_k$ $(1 \le k \le n)$ n'ont pas de racine réelle, et $\alpha_1, ..., \alpha_m, \beta_1, ..., \beta_n \in \mathbb{N} \setminus \{0\}$.

Démonstration

(a) Tout polynôme $P \in \mathbb{C}[X]$ de degré 1 est irréductible (déjà vu).

Soit $P \in \mathbb{C}[X]$ irréductible. D'après le théorème de d'Alembert-Gauss, P a un diviseur de la forme $X - r$ avec $r \in \mathbb{C}$, donc par irréductibilité on a $P = \lambda(X - r)$ pour un $\lambda \in \mathbb{C}$.

(b) Tout $P \in \mathbb{R}[X]$ de degré 1, ou de degré 2 sans racine réelle, est irréductible (déjà vu).

Soit $P \in \mathbb{R}[X]$ irréductible.

Si P a une racine $r \in \mathbb{R}$, on a comme ci-dessus $P = \lambda(X - r)$ pour un $\lambda \in \mathbb{R}$.

Sinon, d'après le théorème de d'Alembert-Gauss, P a une racine $r \in \mathbb{C} \setminus \mathbb{R}$.
Donc il existe $A \in \mathbb{C}[X]$ tel que : $P = (X - r)A$.
Comme P est à coefficients réels, on a : $P(\overline{r}) = \overline{P(r)} = 0$, puis $A(\overline{r}) = 0$.
Donc il existe $B \in \mathbb{C}[X]$ tel que : $A = (X - \overline{r})B$. On a : $P = (X^2 - 2\,\mathrm{Re}(r)X + |r|^2)\,B$.
De plus : $B \in \mathbb{R}[X]$ par existence de la division euclidienne dans $\mathbb{R}[X]$ et unicité dans $\mathbb{C}[X]$.
Ainsi $\underbrace{P}_{\text{irréductible}} = (X^2 - 2\,\mathrm{Re}(r)X + |r|^2)\,B$ puis $P = \lambda(X^2 - 2\,\mathrm{Re}(r)X + |r|^2)$ pour un $\lambda \in \mathbb{R}$. $\square$

Exemple

On obtient les décompositions en polynômes irréductibles unitaires suivantes :

– $\underbrace{X^4 + 1}_{\text{4 racines connues}} \underbrace{=}_{\substack{\text{4 facteurs} \\ \text{irréductibles connus}}} \underbrace{(X - e^{i\frac{\pi}{4}})(X - e^{-i\frac{\pi}{4}})}_{X^2 - \sqrt{2}X + 1} \underbrace{(X - e^{i\frac{3\pi}{4}})(X - e^{-i\frac{3\pi}{4}})}_{X^2 + \sqrt{2}X + 1}$ dans $\mathbb{C}[X]$;

– $X^4 + 1 = (X^2 - \sqrt{2}\,X + 1)(X^2 + \sqrt{2}\,X + 1)$ dans $\mathbb{R}[X]$.

Par contre $X^4 + 1$ est irréductible dans $\mathbb{Q}[X]$. En effet un facteur irréductible de $X^4 + 1$ dans $\mathbb{Q}[X]$ de degré < 4 devrait être un diviseur particulier de $X^4 + 1$ dans $\mathbb{R}[X]$ de la forme $\lambda\,(X^2 - \sqrt{2}\,X + 1)$ avec $\lambda \in \mathbb{R} \setminus \{0\}$ ou $\mu\,(X^2 + \sqrt{2}\,X + 1)$ avec $\mu \in \mathbb{R} \setminus \{0\}$, donc hors de $\mathbb{Q}[X]$.
[En effet, on a $\sqrt{2} \notin \mathbb{Q}$ car le polynôme $X^2 - 2$ est irréductible dans $\mathbb{Q}[X]$.]

Proposition

Soit $P \in \mathbb{K}[X] \setminus \{0\}$.

(a) Pour tout $A \in \mathbb{K}[X]$, on note ici $\underbrace{\dot{A}}_{\text{« classe de } A \text{ »}} := \{A + B\ ;\ B \in P\mathbb{K}[X]\}$.

On munit l'ensemble $\mathbb{K}[X]/(P\mathbb{K}[X]) := \{\dot{A}\ ;\ A \in \mathbb{K}[X]\}$ d'une structure d'anneau avec les lois $+$ et $\times$ définies par :

$$U + V := \widehat{A + B}^{\,\cdot} \quad \text{et} \quad U \times V := \dot{\widehat{AB}} \quad \text{pour } U, V \in \mathbb{K}[X]/(P\mathbb{K}[X])$$

indépendamment du choix de $A, B \in \mathbb{K}[X]$ tels que $U = \dot{A}$ et $V = \dot{B}$

(b) Soit $A \in \mathbb{K}[X] \setminus \{0\}$. L'élément $\dot{A}$ de l'anneau quotient $\mathbb{K}[X]/(P\mathbb{K}[X])$ est « inversible » (c-à-d il existe $B \in \mathbb{K}[X]$ tel que $\dot{A}\dot{B} = \dot{1}$) si et seulement si A et P sont premiers entre eux.

Démonstration

(a) Admis (facile).

(b) On a : $\dot{A}$ est inversible dans $\mathbb{K}[X]/(P\mathbb{K}[X])$
si et seulement si il existe $B \in \mathbb{K}[X]$ tel que $\dot{A}\dot{B} = \dot{1}$,
c'est-à-dire il existe $B, C \in \mathbb{K}[X]$ tel que $AB = 1 + CP$,
c'est-à-dire il existe $B, C \in \mathbb{K}[X]$ tel que $AB + (-C)P = 1$,
c'est-à-dire par le théorème de Bézout A et P sont premiers entre eux. □

Proposition

Soit $P \in \mathbb{K}[X] \backslash \{0\}$. On a : $\mathbb{K}[X]/(P\mathbb{K}[X])$ est un corps si et seulement si P est irréductible.

Démonstration

($\Rightarrow$) Si $\deg P = 0$: $\mathbb{K}[X]/(P\mathbb{K}[X]) = \{\dot{0}\}$ donc $\mathbb{K}[X]/(P\mathbb{K}[X])$ n'est pas un corps.

On suppose $\deg P \geq 1$ et P non irréductible : $P = AB$ pour certains $A, B \in \mathbb{K}[X] \backslash \mathbb{K}$.
On a : $\dot{1} \neq \dot{0}$, $\dot{A} \neq \dot{0}$, $\dot{B} \neq \dot{0}$ dans $\mathbb{K}[X]/(P\mathbb{K}[X])$, et, $\dot{A}\dot{B} = \dot{0}$ donc $\dot{A}$ n'est pas inversible.
Par conséquent $\mathbb{K}[X]/(P\mathbb{K}[X])$ n'est pas un corps.

($\Leftarrow$) On suppose que P est irréductible. Comme $\deg P \neq 0$, on a : $\mathbb{K}[X]/(P\mathbb{K}[X]) \neq \{\dot{0}\}$.

Soit $A \in \mathbb{K}[X] \backslash P\mathbb{K}[X]$. On a $\dot{A} = \dot{R}$ où R est le reste dans la division euclidienne de A par P. Les polynômes B et P n'ont pas de diviseur commun non constant.
Donc $\dot{B}$ est inversible dans $\mathbb{K}[X]/(P\mathbb{K}[X])$ d'après la dernière proposition.

Ainsi $\mathbb{K}[X]/(P\mathbb{K}[X])$ est un corps. □

Exemples

1. On prend $P = X^2 + 1$, donc P est irréductible dans $\mathbb{R}[X]$.

L'application $f\colon \begin{array}{ccc} \mathbb{R}[X]/(P\mathbb{R}[X]) & \longrightarrow & \mathbb{C} \\ a\dot{X} + b & \longmapsto & a + ib \end{array}$ où $a, b \in \mathbb{R}$, est bijective et vérifie :

$$f(\dot{A} + \dot{B}) = f(\dot{A}) + f(\dot{B}),\ f(\dot{1}) = 1 \text{ et } f(\dot{A}\dot{B}) = f(\dot{A})f(\dot{B}) \text{ pour } A, B \in \mathbb{R}[X].$$

2. On note $\mathbb{F}_2 := \mathbb{Z}/2\mathbb{Z}$. On prend $Q = X^2 + X + \bar{1}$, donc Q est irréductible dans $\mathbb{F}_2[X]$.

Le corps $\mathbb{F}_4 := \mathbb{F}_2[X]/(Q\mathbb{F}_2[X])$ a 4 éléments : $\mathbb{F}_4 = \{\dot{\bar{0}}, \dot{\bar{1}}, \dot{X}, \dot{X} + \dot{\bar{1}}\}$.

III. Fractions rationnelles

Comment dériver 1000 fois et comment intégrer une fonction rationnelle $f\colon [a,b] \to \mathbb{R}$?

Méthode : on écrit f comme somme de fonctions rationnelles f_1, ..., f_n « plus simples ».

(Pour dériver, on utilisera les éléments simples sur $\mathbb{C}$. Pour intégrer un élément simple f_m $(1 \leq m \leq n)$ de seconde espèce, on cherche $\alpha \in \mathbb{R}$, puis $\beta \in \mathbb{R}$ et $\gamma > 0$, tels que :

$$f_m(x) = \frac{\mu x + \nu}{(x^2 + u_l x + v_l)^j} = \alpha \frac{2x + u_l}{(x^2 + u_l x + v_l)^j} + \beta \frac{1}{((x + \frac{u_l}{2})^2 + \gamma^2)^j} \quad \text{pour tout } x \in \mathbb{R}.$$

Les primitives $\int \frac{2x+u_l}{(x^2+u_l x+v_l)^j}\,\mathrm{d}x$ se calculent à l'aide du CV : $t = x^2 + u_l x + v_l$.

Le calcul des primitives $\int \frac{\mathrm{d}x}{((x+\frac{u_l}{2})^2+\gamma^2)^j}$ se ramène à celui de $\int \frac{\mathrm{d}x}{((x+\frac{u_l}{2})^2+\gamma^2)^{j-1}}$ en intégrant par parties à partir de $\int \underbrace{1}_{\text{dérivée de } x \mapsto x + \frac{u_l}{2}} \times \frac{1}{((x+\frac{u_l}{2})^2+\gamma^2)^{j-1}}\,\mathrm{d}x$.)

1. Fractions rationnelles

Définition

(a) Une *fraction rationnelle à coefficients dans* $\mathbb{K}$ est un « quotient » $F = \frac{A}{B}$ (en un sens qu'on pourrait préciser) où $A \in \mathbb{K}[X]$ et $B \in \mathbb{K}[X] \setminus \{0\}$.

(b) À chaque fraction rationnelle $F = \dfrac{a_p X^p + \cdots + a_1 X + a_0}{b_q X^q + \cdots + b_1 X + b_0}$ à coefficients dans $\mathbb{K}$, on associe l'application notée ici abusivement f_F et définie par :

$f_F(x) = \underbrace{\dfrac{a_p x^p + \cdots + a_1 x + a_0}{b_q x^q + \cdots + b_1 x + b_0}}_{\text{noté } F(x)}$ pour (au moins) tout $x \in \mathbb{K}$ tel que $b_q x^q + \cdots + b_1 x + b_0 \neq 0$.

[L'application f_F est définie sur cet ensemble pour l'écriture $\frac{A}{B}$ irréductible et B unitaire.]

Une *application rationnelle sur* $\mathbb{K}$ est une application f d'une certaine partie de $\mathbb{K}$ dans $\mathbb{K}$ telle qu'il existe une fraction rationnelle F à coefficients dans $\mathbb{K}$ pour laquelle $f = f_F$.

(c) Le *degré* d'une fraction rationnelle $F = \frac{A}{B}$ avec $A \in \mathbb{K}[X]$ et $B \in \mathbb{K}[X] \setminus \{0\}$, noté $\deg F$, est défini par : $\deg F := \deg A - \deg B$.

Remarque (hors programme)

On appelle « quotient d'un élément A de $\mathbb{K}[X]$ par un élément B de $\mathbb{K}[X] \setminus \{0\}$ » l'ensemble :

$$\frac{A}{B} := \{(A_0, B_0) \in \mathbb{K}[X] \times (\mathbb{K}[X] \setminus \{0\}) \mid AB_0 = A_0 B\} \quad \text{(il contient } (A, B)\text{)}.$$

Par définition on a donc : $\frac{A}{B} = \frac{PA}{PB}$ pour tout $P \in \mathbb{K}[X] \setminus \{0\}$.

Chaque notion attachée à $\frac{A}{B}$ devra être indépendante du choix de (A, B) dans $\frac{A}{B}$.

Par exemple on peut noter : $\left(\frac{A}{B}\right)(x) := \frac{A(x)}{B(x)}$ pour certains $x \in \mathbb{K}$, et $\left(\frac{A}{B}\right)' := \frac{A'B - AB'}{B^2}$

car pour $(A_0, B_0) \in \frac{A}{B}$, on a : $\frac{A_0(x)}{B_0(x)} = \frac{A(x)}{B(x)}$ quand $B_0(x)B(x) \neq 0$, et $\frac{A_0' B_0 - A_0 B_0'}{B_0^2} = \frac{A'B - AB'}{B^2}$.

(En effet : $AB_0 = A_0 B$ donc $A_0' B_0 B^2 + AB' B_0^2 = (A_0' B + A_0 B') B_0 B = (A' B_0 + A B_0') B_0 B = A' B B_0^2 + A_0 B_0' B^2$.)

Notation

On note $\mathbb{K}(X)$ l'ensemble des fractions rationnelles à coefficients dans $\mathbb{K}$.
En « identifiant » A à $\frac{A}{1}$, on obtient : $\mathbb{K}[X] \subseteq \mathbb{K}(X)$ (choix de $B = 1$).

Définition-Proposition

Soient $F = \frac{A}{B}$ et $G = \frac{C}{D}$ deux éléments de $\mathbb{K}(X)$, $\alpha, x \in \mathbb{K}$ avec $B(x) \neq 0$ et $D(x) \neq 0$.

On note : $F + G := \frac{AD + BC}{BD}$, $FG := \frac{AC}{BD}$, $\alpha F = \frac{\alpha A}{B}$, et aussi $\frac{F}{G} := \frac{AD}{BC}$ quand $G \neq 0$.

Ainsi, on a : $(F + G)(x) = F(x) + G(x)$, $(FG)(x) = F(x)G(x)$, $(\alpha F)(x) = \alpha F(x)$.

Proposition

L'ensemble $\mathbb{K}(X)$ muni de l'addition et de la multiplication des fractions rationnelles est un corps commutatif de zéro le polynôme 0 et d'élément unité le polynôme constant 1.

Remarques

1. Soit $F = \frac{A}{B} \in \mathbb{K}(X) \setminus \{0\}$.

On décompose $A = \lambda P_1^{\alpha_1} \ldots P_k^{\alpha_k}$ et $B = \mu P_1^{\beta_1} \ldots P_k^{\beta_k}$ avec $\lambda, \mu \in \mathbb{K} \setminus \{0\}$, $P_1, \ldots, P_k \in \mathbb{K}[X]$ unitaires irréductibles et $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k \in \mathbb{N}$. On pose $D = P_1^{\min(\alpha_1, \beta_1)} \cdots P_k^{\min(\alpha_k, \beta_k)}$.

On obtient tout de suite $A_0, B_0 \in \mathbb{K}[X] \setminus \{0\}$ sans facteur unitaire irréductible commun tels que $A = A_0 D$ et $B = B_0 D$. On a : $F = \frac{A_0}{B_0}$ « écriture de F comme fraction irréductible ».

2. Soient $P = a_n X^n + \cdots + a_1 X + a_0 \in \mathbb{C}[X]$ et $Q \in \mathbb{C}[X]$, $F = \frac{A}{B} \in \mathbb{C}(X)$ et $G \in \mathbb{C}(X)$.

On pose : $\overline{P} := \overline{a_n} X^n + \cdots + \overline{a_1} X + \overline{a_0} \in \mathbb{C}[X]$, puis $\overline{F} := \frac{\overline{A}}{\overline{B}} \in \mathbb{C}(X)$ (cf. $\overline{PQ} = \overline{P}\,\overline{Q}$).

On obtient : $\overline{F + G} = \overline{F} + \overline{G}$, $\overline{FG} = \overline{F}\,\overline{G}$, et $\overline{F(z)} = \overline{F}(\overline{z})$ quand $z \in \mathbb{C}$ vérifie $B(z) \neq 0$.

2. Décomposition en éléments simples

Lemme

Soient $A, \widetilde{B} \in \mathbb{K}[X]$ et $r \in \mathbb{K}$ tels que $\widetilde{B}(r) \neq 0$.

Pour tout $\alpha \in \mathbb{N} \setminus \{0\}$ il existe $c_0, c_1, ..., c_{\alpha-1} \in \mathbb{K}$ et $R_{\alpha-1} \in \mathbb{K}[X]$ uniques tels que :

$$A = \widetilde{B}_{\times}\big(c_0 + c_1(X-r) + \cdots + c_{\alpha-1}(X-r)^{\alpha-1}\big) + (X-r)^{\alpha} R_{\alpha-1}$$

« division de A par $\widetilde{B}$ suivant les puissances croissantes de $X - r$ jusqu'à l'ordre $\alpha - 1$ ».

Démonstration

• Existence ?

On peut supposer que $A \neq 0$ (clair sinon). On note m la multiplicité de r dans A.

On vérifie par récurrence sur $n \in \mathbb{N}$ avec α fixé que :

$$(H_n) \quad \text{existence d'une décomposition quand } A \neq 0 \text{ et } \alpha - m \leq n.$$

(i) Si $n = 0$: on écrit $A = (X-r)^m C$ pour un certain $C \in \mathbb{K}[X]$ où $m \geq \alpha$, puis l'égalité $A = \widetilde{B}_{\times}0 + (X-r)^{\alpha}\big((X-r)^{m-\alpha}C\big)$ donne une décomposition.

(ii) Soit $n \geq 1$ tel que (H_{n-1}) est vraie. On suppose donc que $\alpha - m = n$:
$A = \tilde{a}_m(X-r)^m + \cdots + \tilde{a}_p(X-r)^p$ et $\widetilde{B} = \tilde{b}_0 + \tilde{b}_1(X-r) + \cdots + \tilde{b}_{\tilde{q}}(X-r)^{\tilde{q}}$ avec $\tilde{b}_0 \neq 0$.
On pose : $A_0 = A - \frac{\tilde{a}_m}{\tilde{b}_0}(X-r)^m\, \widetilde{B}$. Si $A_0 \neq 0$, la multiplicité de r dans A_0 est $\geq m+1$.
D'après (H_{n-1}), il existe $c_0, c_1, ..., c_{\alpha-1} \in \mathbb{K}$ et $R_{\alpha-1} \in \mathbb{K}[X]$ tels que :
$A_0 = \widetilde{B}_{\times}\big(c_0 + c_1(X-r) + \cdots + c_{\alpha-1}(X-r)^{\alpha-1}\big) + (X-r)^{\alpha}R_{\alpha-1}$.
Ainsi : $A = \widetilde{B}_{\times}\big(c_0 + c_1(X-r) + \cdots + (c_m + \frac{\tilde{a}_m}{\tilde{b}_0})(X-r)^m + \cdots + c_{\alpha-1}(X-r)^{\alpha-1}\big) + (X-r)^{\alpha}R_{\alpha-1}$.

(iii) D'où le résultat.

• Unicité ?

Soient $\widetilde{Q}_1, \widetilde{Q}_2 \in \mathbb{K}_{\alpha-1}[X]$ et $R_1, R_2 \in \mathbb{K}[X]$ tels que :

$$A = \widetilde{B}\widetilde{Q}_1 + (X-r)^{\alpha}R_1 = \widetilde{B}\widetilde{Q}_2 + (X-r)^{\alpha}R_2.$$

Par différence on a : $\widetilde{B}(\widetilde{Q}_1 - \widetilde{Q}_2) = (X-r)^{\alpha}(R_2 - R_1)$, où $\widetilde{Q}_1, \widetilde{Q}_2 \in \mathbb{K}_{\alpha-1}[X]$.

On suppose par l'absurde que $\widetilde{Q}_2 \neq \widetilde{Q}_1$, donc $R_2 \neq R_1$. La multiplicité de r dans $\widetilde{B}(\widetilde{Q}_1 - \widetilde{Q}_2)$ est $< \alpha$ tandis qu'elle est $\geq \alpha$ dans $(X-r)^{\alpha}(R_2 - R_1)$ contradictiopn.

Par conséquent, on a : $\widetilde{Q}_2 = \widetilde{Q}_1$, puis $R_2 = R_1$. □

Théorème (« *théorème de décomposition en éléments simples* » sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$)

Soient $A \in \mathbb{K}[X]$ et $B \in \mathbb{K}[X] \setminus \{0\}$. On pose : $F := \frac{A}{B}$.

(a) On suppose que $\mathbb{K} = \mathbb{C}$. On décompose B :

$$B = \gamma\,(X-r_1)^{\alpha_1} ... (X-r_m)^{\alpha_m}$$

avec $\gamma \in \mathbb{C} \setminus \{0\}$, $\underbrace{\{r_1, ..., r_m\}}_{\text{distincts}} \subseteq \mathbb{C}$, et $\alpha_1, ..., \alpha_m \in \mathbb{N} \setminus \{0\}$.

La fraction rationnelle F est somme de manière unique :

– d'un polynôme $E \in \mathbb{C}[X]$ appelé *partie entière de* F ;

– « d'éléments simples » $\frac{\lambda}{(X-r_k)^i}$ avec $\lambda \in \mathbb{C}$, $1 \leq k \leq m$, $1 \leq i \leq \alpha_k$.

$\Big[F = E + \sum_{k=1}^{m} \Big(\frac{\lambda_{k,1}}{X-r_k} + \cdots + \frac{\lambda_{k,\alpha_k}}{(X-r_k)^{\alpha_k}}\Big).\Big]$

Quand A n'est pas divisible par $X - r_k$ l'élément simple $\frac{\lambda}{(X-r_k)^{\alpha_k}}$ vérifie $\lambda \neq 0$.

(b) On suppose que $\mathbb{K} = \mathbb{R}$. On décompose B :

$$B = \gamma\,(X-r_1)^{\alpha_1} ... (X-r_m)^{\alpha_m}(X^2 + u_1X + v_1)^{\beta_1} ... (X^2 + u_nX + v_n)^{\beta_n}$$

avec $\gamma \in \mathbb{R} \setminus \{0\}$, $\underbrace{\{r_1, ..., r_m\}}_{\text{distincts}} \subseteq \mathbb{R}$, $\underbrace{\{(u_1, v_1), ..., (u_n, v_n)\}}_{\text{couples distincts}} \subseteq \mathbb{R}^2$ tels que les $X^2 + u_kX + v_k$ $(1 \leq k \leq n)$ n'ont pas de racine réelle, et $\alpha_1, ..., \alpha_m, \beta_1, ..., \beta_n \in \mathbb{N} \setminus \{0\}$.

La fraction rationnelle F est somme de manière unique :

– d'un polynôme $E \in \mathbb{R}[X]$ appelé *partie entière de* F ;

– « d'éléments simples de 1re espèce » $\frac{\lambda}{(X-r_k)^i}$ avec $\lambda \in \mathbb{R}$, $1 \leq k \leq m$, $1 \leq i \leq \alpha_k$;

– « d'éléments simples de 2e espèce » $\frac{\mu X + \nu}{(X^2+u_lX+v_l)^j}$ avec $\mu, \nu \in \mathbb{R}$, $1 \leq l \leq n$, $1 \leq j \leq \beta_l$.

$$\left[F = E + \sum_{k=1}^{m}\left(\frac{\lambda_{k,1}}{X-r_k} + \cdots + \frac{\lambda_{k,\alpha_k}}{(X-r_k)^{\alpha_k}}\right) + \sum_{l=1}^{n}\left(\frac{\mu_{l,1}X+\nu_{l,1}}{X^2+u_lX+v_l} + \cdots + \frac{\mu_{l,\beta_l}X+\nu_{l,\beta_l}}{(X^2+u_lX+v_l)^{\beta_l}}\right).\right]$$

Quand A n'est pas divisible par $X-r_k$ (resp. par $X^2+u_lX+v_l$) l'élément simple $\frac{\lambda}{(X-r_k)^{\alpha_k}}$ (resp. $\frac{\mu X+\nu}{(X^2+u_lX+v_l)^{\beta_l}}$) vérifie $\lambda \neq 0$ (resp. $(\mu,\nu) \neq (0,0)$).

(c) Le polynôme E du (a) ou du (b) est le quotient de la division euclidienne de A par B. Il est nul si $\deg(A) < \deg(B)$, et non-nul avec un degré égal à $\deg(A) - \deg(B)$ sinon.

Démonstration (idées)

(a) • On montre l'existence d'une décomposition sur $\mathbb{C}$ par récurrence sur m.

(i) On suppose que $m = 0$. On a : $F \in \mathbb{K}[X]$ et la décomposition est évidente.

(ii) On se donne $m \geq 1$ pour lequel la décomposition est obtenue quand B a $m-1$ racines.

On utilise le lemme avec $\widetilde{B} := \frac{B}{(X-r_m)^{\alpha_m}}$, $r = r_m$ et $\alpha = \alpha_m$:

$$A = \widetilde{B}\times\left(c_0 + c_1(X-r) + \cdots + c_{\alpha_m-1}(X-r)^{\alpha_m-1}\right) + (X-r)^{\alpha_m}R_{\alpha_m-1}.$$

Donc $F = \frac{c_0+c_1(X-r)+\cdots+c_{\alpha_m-1}(X-r)^{\alpha_m-1}}{(X-r_m)^{\alpha_m}} + \frac{R_{\alpha_m-1}}{\widetilde{B}}$, puis on applique l'hypothèse de récurrence.

On en déduit que la décomposition est obtenue quand B a m racines.

(iii) Ainsi l'existence est obtenue.

• Unicité sur $\mathbb{C}$? On se donne $E \in \mathbb{C}[X]$ et $\lambda_{k,i} \in \mathbb{C}$ tels que

$$E + \sum_{k=1}^{m}\left(\frac{\lambda_{k,1}}{X-r_k} + \cdots + \frac{\lambda_{k,\alpha_k}}{(X-r_k)^{\alpha_k}}\right) = 0.$$

On fixe un k. On multiplie par $(X-r_k)^{\alpha_k}$ puis prend la valeur en r_k. On obtient $\lambda_{k,\alpha_k} = 0$. On recommence en multipliant par $(X-r_k)^{\alpha_k-1}$ ce qui donne $\lambda_{k,\alpha_k-1} = 0$. Puis ... $\lambda_{k,1} = 0$.

Enfin $E = 0$.

• Si $\lambda = 0$: en multipliant F décomposé par B, on constate que A est divisible par $X - r_k$.

(b) • On montre l'existence d'une décomposition sur $\mathbb{R}$ en utilisant la décomposition sur $\mathbb{C}$.

On considère un élément simple $\frac{\lambda}{(X-r_k)^i}$. Si $r_k \in \mathbb{R}$, on a : $\overline{F} = F$ donc $\lambda \in \mathbb{R}$ par unicité.

On suppose que $r_k \notin \mathbb{R}$. Comme $\overline{F} = F$, on a : $\frac{\overline{\lambda}}{(X-\overline{r_k})^i}$ est un élément simple de F sur $\mathbb{C}$.

On décompose sur $\mathbb{R}$: $\frac{\lambda}{(X-r_k)^i} + \frac{\overline{\lambda}}{(X-\overline{r_k})^i}$ qui vaut $\frac{\lambda(X-\overline{r_k})^i+\overline{\lambda}(X-r_k)^i}{((X-r_k)(X-\overline{r_k}))^i}$, est somme d'éléments simples de 2^e^ espèce (diviser suffisamment de fois $\lambda(X-\overline{r_k})^i+\overline{\lambda}(X-r_k)^i$ par $(X-r_k)(X-\overline{r_k})$).

• Unicité sur $\mathbb{R}$? On se donne $E \in \mathbb{R}[X]$ et $\lambda_{k,i}, \mu_{l,j}, \nu_{l,j} \in \mathbb{R}$ tels que

$$E + \sum_{k=1}^{m}\left(\frac{\lambda_{k,1}}{X-r_k} + \cdots + \frac{\lambda_{k,\alpha_k}}{(X-r_k)^{\alpha_k}}\right) + \sum_{l=1}^{n}\left(\frac{\mu_{l,1}X+\nu_{l,1}}{X^2+u_lX+v_l} + \cdots + \frac{\mu_{l,\beta_l}X+\nu_{l,\beta_l}}{(X^2+u_lX+v_l)^{\beta_l}}\right) = 0.$$

Comme ci-dessus, on obtient $\lambda_{k,i} = 0$ pour tous les k, i.

On fixe un l. On introduit une racine s_l de $X^2 + u_lX + v_l$ (non-réelle). On multiplie par $(X^2+u_lX+v_l)^{\beta_l}$ et prend la valeur en s_l. Donc : $\mu_{l,\beta_l}s_l + \nu_{l,\beta_l} = 0$ puis $\mu_{l,\beta_l} = 0$ et $\nu_{l,\beta_l} = 0$. La multiplication par $(X^2+u_lX+v_l)^{\beta_l-1}$ donne $\mu_{l,\beta_l-1} = \nu_{l,\beta_l-1} = 0$. Puis ... $\mu_{l,1} = \nu_{l,1} = 0$.

Enfin $E = 0$.

• Si $\lambda = 0$: on constate comme ci-dessus que A est divisible par $X - r_k$.

Si $\mu = \nu = 0$: en multipliant F par B, on constate que A est divisible par $X^2 + u_lX + v_l$.

(c) À la fois dans (a) et dans (b), on obtient en multipliant F décomposé par B :

$$A = BE + R \ \text{ avec } \ \deg R \leq \deg B - 1.$$

Il s'agit du résultat de la division euclidienne de A par B. □

Algorithmes (l'algorithme du (b) n'est pas au programme mais peut être utilisé)

On reprend les notations du théorème. (Idée : calcul des DL de $x \mapsto F(x)$ en $+\infty$ et en r_k.)

a) On suppose que $\deg(A) \geq \deg(B)$ (sinon $E = 0$) et note :

$A = a_pX^p + \cdots + a_1X + a_0$ et $B = b_qX^q + \cdots + b_1X + b_0$, avec $a_p \neq 0$ et $b_q \neq 0$.

Division euclidienne de A par B : son quotient Q vérifie $\boxed{Q = E}$ (cf. le th. (c)).

Astuce : on utilise dans ce calcul autant de coefficients de A et de B, à partir de a_p et b_q, qu'on en attend dans E.

b) Soit $k \in \{1, ..., m\}$. On pose $X = r_k + Y$ et $\widetilde{B} := \frac{B}{Y^{\alpha_k}}$. On écrit A et $\widetilde{B}$ ainsi :

$$A = \tilde{a}_0 + \tilde{a}_1 Y + \cdots + \tilde{a}_p Y^p \quad \text{et} \quad \widetilde{B} = \tilde{b}_0 + \tilde{b}_1 Y + \cdots + \tilde{b}_{q-\alpha_k} Y^{q-\alpha_k} \quad \text{où} \quad \tilde{b}_0 \neq 0.$$

La fraction rationnelle F_k somme des éléments simples de 1^{re} espèce associés à r_k s'obtient par division suivant les puissances croissantes de A par $\widetilde{B}$, jusqu'à faire apparaître α_k coefficients dans le quotient (méthode incontournable quand $\alpha_k \geq 4$) :

$$\begin{array}{rl|l}
 & \tilde{a}_0 + \tilde{a}_1 Y + \cdots + \tilde{a}_p Y^p & \tilde{b}_0 + \tilde{b}_1 Y + \cdots + \tilde{b}_{q-\alpha_k} Y^{q-\alpha_k} \\
\ominus & \tilde{a}_0 + \frac{\tilde{a}_0 \tilde{b}_1}{\tilde{b}_0} Y + \cdots & \hline \underbrace{\frac{\tilde{a}_0}{\tilde{b}_0} + \frac{\tilde{a}'_1}{\tilde{b}_0} Y + \cdots} \\
\cline{2-2}
= & \tilde{a}'_1 Y + \cdots \quad \cdots + \tilde{a}'_{p'} Y^{p'} & \widetilde{Q}_k := c_0 + c_1 (X - r) + \cdots + c_{\alpha_k - 1} (X - r)^{\alpha_k - 1} \\
\ominus & \cdots \quad \cdots \quad \cdots & \\
\cline{2-2}
= & \cdots \quad \cdots \quad \cdots & \text{donc} \quad \boxed{F_k = \frac{\widetilde{Q}_k}{(X - r_k)^{\alpha_k}}} \quad \text{(cf. la dém. du th.)} \\
 & // \quad // \quad // \quad // \quad // & \\
\cline{2-2}
 & \tilde{a}''_{\alpha_k} Y^{\alpha_k} + \cdots + \tilde{a}''_{p''} Y^{p''} &
\end{array}$$

Astuce : on utilise dans ce calcul autant de coefficients de A et de $\widetilde{B}$, à partir de $\tilde{a}_0$ et $\tilde{b}_0$, qu'on en attend dans F_k.

Remarques

1. Après avoir utilisé les algorithmes (a), et (b) pour chaque $k \in \{1, ..., m\}$, on peut :

– appliquer la décomposition en un $z \in \mathbb{C}$, par exemple en une racine dans $\mathbb{C}$ d'un $X^2 + u_l X + v_l$;

– réduire au même dénominateur la décomposition en éléments simples de F obtenue pour l'instant et identifier les coefficients de son numérateur avec ceux de A ;

– réduire au même dénominateur $F - \left(E + \sum\limits_{k=1}^{m} F_k\right)$ puis le simplifier lorsque $\mathbb{K} = \mathbb{R}$ et $n = 1$.

2. On peut éviter de factoriser B sous la forme $(X - r_k)^{\alpha_k} C$ pour calculer le coefficient λ de l'élément simple $\frac{\lambda}{(X - r_k)^{\alpha_k}}$. En effet, il s'écrit : $\lambda = \frac{A(r_k)}{C(r_k)} = \alpha_k! \frac{A(r_k)}{B^{(\alpha_k)}(r_k)}$.

(On utilise la formule de Leibniz pour calculer $B^{(\alpha_k)}(r_k)$ à partir de l'égalité $B = (X - r_k)^{\alpha_k} C$.)

Exemple 1

On considère : $F = \dfrac{X^5 + X^4 + X^2 + 1}{X^3 - X^2 - X + 1} \in \mathbb{C}(X)$.

On a : $\underbrace{X^3 - X^2 - X + 1}_{\text{s'annule en 1}} = (X - 1)(X^2 + 0X - 1) = (X + 1)(X - 1)^2$.

On note $r_1 = -1$ et $r_2 = 1$. Il existe $E \in \mathbb{C}[X]$ de degré 2 et $a, b, c \in \mathbb{C}$ uniques tels que :

$$\underbrace{\frac{X^5 + X^4 + X^2 + 1}{(X + 1)(X - 1)^2}}_{F} = E + \underbrace{\frac{a}{X + 1}}_{F_1, \text{ cf. (b)}} + \underbrace{\frac{b}{(X - 1)^2} + \frac{c}{X - 1}}_{F_2, \text{ cf. (b)}}.$$

On commence par calculer les 3 coefficients de E :

$$\begin{array}{rr|l}
 & X^5 + X^4 + 0X^3 + \cdots & X^3 - X^2 - X + \cdots \\
\cline{3-3}
\ominus & X^5 - X^4 - X^3 + \cdots & X^2 + 2X + 3 \\
\cline{2-2}
= & 2X^4 + X^3 + \cdots & \\
\ominus & 2X^4 - 2X^3 + \cdots & \\
\cline{2-2}
= & 3X^3 + \cdots & \\
 & // \quad // \quad // &
\end{array}
\qquad \text{donc} \quad E = X^2 + 2X + 3.$$

On multiplie l'égalité plus haut par $X + 1$ et prend les valeurs en $x = -1$: $a = \frac{1}{2}$.

• Méthode 1

On continue en généralisant cette technique et choisissant une autre valeur particulière :

– on multiplie par $(X - 1)^2$ et prend les valeurs en $x = 1$: $b = 2$;

– valeurs en $x = 0$: $1 = 3 + a + b - c$, donc $c = \frac{9}{2}$.

• Méthode 2

On utilise l'algorithme (b) pour calculer b et c. On pose $X = 1+Y$ et cherche 2 coefficients :

$$F = \frac{(1+Y)^5+(1+Y)^4+(1+Y)^2+1}{(1+Y)^3-(1+Y)^2-(1+Y)+1} = \frac{4+11Y+\cdots}{Y^3(2+Y)}.$$

Division suivant les puissances de Y croissantes jusqu'à obtenir 2 coefficients :

$$\begin{array}{rl|l} & 4+11Y+\cdots & 2+Y \\ \hline - & 4+\ 2Y & 2+\frac{9}{2}Y \\ \hline = & 9Y\ +\cdots & \\ & // \ // & \end{array}$$

donc $F_2(X) = \frac{2}{Y^2}+\frac{\frac{9}{2}}{Y}$ puis $b=2$ et $c=\frac{9}{2}$.

• Finalement : $F \ = \ X^2+2X+3 \ + \ \frac{\frac{1}{2}}{X+1} \ + \ \frac{2}{(X-1)^2} \ + \ \frac{\frac{9}{2}}{X-1}$.

Exemple 2

On considère : $F = \dfrac{1}{X^3+1} \in \mathbb{R}(X)$.

On a : $\underbrace{X^3+1}_{\text{s'annule en } -1} = (X+1)(\underbrace{X^2+\ ?\ }_{X^2-X+1})$).

Il existe donc $\lambda, \mu, \nu \in \mathbb{R}$ uniques tels que :

$$\underbrace{\frac{1}{(X+1)(X^2-X+1)}}_{F} \ = \ \underbrace{0}_{E} \ + \ \frac{\lambda}{X+1} \ + \ \frac{\mu X+\nu}{X^2-X+1}.$$

On pourrait réduire le membre de droite au même dénominateur, puis identifier les coefficients des numérateurs (c'est tout à fait possible). D'habitude, on utilise les astuces suivantes :

– on multiplie par $X+1$ et prend les valeurs en $x=-1$: $\lambda = \frac{1}{3}$;

– valeurs en $x=0$: $1 = \lambda+\nu$, donc $\nu = \frac{2}{3}$;

– on multiplie par X, prend les valeurs en x et passe à $x \to +\infty$: $0 = \lambda+\mu$, donc $\mu = -\frac{1}{3}$.

(Au lieu d'utiliser les valeurs 0 et $+\infty$, on aurait aussi pu se placer dans $\mathbb{C}(X)$ et prendre une valeur complexe non-réelle, par exemple $\frac{1+i\sqrt{3}}{2}$ après avoir multiplié par X^2-X+1.)

En conclusion : $F \ = \ \frac{\frac{1}{3}}{X+1} \ + \ \frac{-\frac{1}{3}X+\frac{2}{3}}{X^2-X+1}$.

Ch. 2. Matrices sur $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$

Plan

I. Systèmes linéaires (rappels)

1. Introduction

Deux problèmes

1. Comment passer d'une $\overbrace{\text{équation cartésienne}}^{\text{(image réciproque d'un singleton par une application)}}$ d'un « sous-espace affine de $\mathbb{K}^p$ » à une équation paramétrique qui décrit les points de ce sous-espace affine ? Par exemple :

$$(E_{\text{cart}})\begin{cases}-4x+12y-5z=\ \ 1\\ \quad x-3y+2z=-1\\ \quad 2x-6y+z=\ \ 1\end{cases}\ \text{dans } \mathbb{K}^3 \text{ détermine } \mathscr{S}_{E_{\text{cart}}}:=\underbrace{\{(x,y,z)\in\mathbb{K}^3 \mid E_{\text{cart}}\}}_{\text{image réciproque de } \{(1,-1,1)\} \text{ par } \dots}.$$

2. Comment passer d'une $\overbrace{\text{équation paramétrique}}^{\text{(image directe d'une application)}}$ d'un « sous-espace affine de $\mathbb{K}^p$ » à une équation cartésienne de ce sous-espace affine ? Par exemple :

$$\mathscr{D}\colon \begin{cases}x=\ \ s-t+1\\ y=-s+t+2\end{cases},\ s,t\in\mathbb{K} \quad \text{détermine} \quad \mathscr{D}:=\underbrace{\{(s-t+1,-s+t+2)\ ;\ s,t\in\mathbb{K}\}}_{\text{image directe de } \mathbb{K}^2 \text{ par } \dots}\subseteq\mathbb{K}^2.$$

Une motivation

De nombreuses équations issues de la physique se résolvent numériquement (par « discrétisation ») en se ramenant à des « systèmes linéaires », cf. :
`http://math.nist.gov/MatrixMarket/` (cliquer sur « Search by application area »).

But

On désire « résoudre » des équations comme (E_{cart}) au sens où on écrira $\mathscr{S}_{E_{\text{cart}}}$ sous forme paramétrique, sous réserve d'avoir $\mathscr{S}_{E_{\text{cart}}}\neq 0$. On va utiliser une méthode qui permettra aussi de passer d'une forme paramétrique à une forme cartésienne. Plus précisément, on cherche :
– une forme paramétrique sans paramètre inutile (ce n'est pas le cas dans la définition de $\mathscr{D}$ où tout s'exprime avec $u:=s-t$) ;
– une forme cartésienne sans égalité inutile (ce n'est pas le cas dans l'écriture de (E_{cart}) où ligne $3=-$ligne $1-2$ ligne 2).

2. Transformation d'un système linéaire

Définition

(a) Un système *d'équations « linéaires » de n équations à p inconnues dans* $\mathbb{K}$ est du type :

$$(E)\quad\begin{cases}a_{11}x_1+\dots+a_{1p}x_p &= b_1\\ \qquad\dots & \\ a_{n1}x_1+\dots+a_{np}x_p &= b_n\end{cases}\quad \text{d'inconnue } (x_1,\dots,x_p)\in\mathbb{K}^p$$

où sont donnés $a_{11},\dots,a_{1p},a_{21},\dots,a_{2p},\dots,a_{n1},\dots,a_{np}\in\mathbb{K}$ et $b_1,\dots,b_n\in\mathbb{K}$.

Dans la suite du I on fixe un tel système (E) et note $\mathscr{S}_E$ l'ensemble de ses solutions dans $\mathbb{K}^p$.

(b) Le *système d'équations linéaires homogène associé à* (E) est le système suivant :

$$(H)\quad\begin{cases}a_{11}x_1+\dots+a_{1p}x_p &= 0\\ \qquad\dots & \\ a_{n1}x_1+\dots+a_{np}x_p &= 0\end{cases}\quad \text{d'inconnue } (x_1,\dots,x_p)\in\mathbb{K}^p.$$

Remarques (importantes)

1. (H) admet toujours comme solution $(0,\dots,0)$.

2. On suppose que $\mathscr{S}_E \neq \emptyset$ et fixe une solution « particulière » $(x_1^E,\dots,x_p^E)$ de (E).
Pour tout $(x_1,\dots,x_p)\in\mathbb{K}^p$, on a tout de suite :

$(x_1,\dots,x_p)$ vérifie (E) si et seulement si $(x_1-x_1^E,\dots,x_p-x_p^E)$ vérifie (H).

En résumé : $\boxed{\text{solutions de } (E) = \text{solutions de } (H) + \text{une solution particulière de } (E)}$.

Notations

(a) On utilisera les « matrices » suivantes :

$$A=\begin{pmatrix} a_{11} & \dots & a_{1p}\\ & \dots & \\ a_{n1} & \dots & a_{np}\end{pmatrix} \textit{ matrice de } (E),\quad X=\begin{pmatrix} x_1\\ \vdots \\ x_p\end{pmatrix}\underbrace{\textit{inconnue}}_{\text{identifiée à } (x_1,\dots,x_p)},\text{ et }\ B=\begin{pmatrix} b_1\\ \vdots\\ b_n\end{pmatrix} \textit{ second membre de } (E).$$

On écrira en abrégé $(E) : AX=B$.

(b) La *matrice augmentée* de (E) est la matrice $(A|B) := \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1p} & b_1\\ & \dots & & \vdots\\ a_{n1} & \dots & a_{np} & b_n\end{array}\right)$, où la barre verticale n'a aucune signification mathématique.

On résoudra (E) en travaillant sur les lignes de $(A|B)$ et sur les colonnes de A.

Exemples (cf. l'introduction de la partie I de ce chapitre)

1. $\underbrace{\text{L'équation}}_{\text{d'inconnue } (x,y,z)\in\mathbb{K}^3}$ $(E_{\text{cart}})\begin{cases} -4x+12y-5z = 1\\ x-3y+2z=-1\\ 2x-6y+z= 1\end{cases}$ a pour matrice augmentée $\left(\begin{array}{ccc|c} -4 & 12 & -5 & 1\\ 1 & -3 & 2 & -1\\ 2 & -6 & 1 & 1\end{array}\right)$.

2. On fixe $(x,y)\in\mathbb{K}^2$. On a : $(x,y)\in\mathscr{D} \iff \left(\exists (s,t)\in\mathbb{K}^2 \quad \begin{cases} x = s-t+1\\ y=-s+t+2\end{cases}\right)$.

$\underbrace{\text{L'équation}}_{\text{d'inconnue } (s,t)\in\mathbb{K}^2}$ $(E_{\text{param}})\begin{cases} s-t=x-1\\ -s+t=y-2\end{cases}$ a pour matrice augmentée $\left(\begin{array}{cc|c} 1 & -1 & x-1\\ -1 & 1 & y-2\end{array}\right)$.

Définition

Dans ce qui suit on va noter L_1, ..., L_n les lignes de $(A|B)$ et C_1, ..., C_p les colonnes de A. Étant donnés $k\in\{1,...,n\}$ et $\alpha_1,...,\alpha_n\in\mathbb{K}$, on écrira « $L_k \leftarrow \alpha_1L_1+...+\alpha_nL_n$ » pour exprimer qu'on remplace la ligne L_k par $\alpha_1L_1+...+\alpha_nL_n$.

(a) Une *opération élémentaire sur les lignes de* $(A|B)$ est une des transformations suivantes :

(i) $L_i \overset{\text{(échange)}}{\longleftrightarrow} L_j$ avec $i\neq j$ « permutation de deux lignes » ;

(ii) $L_i \leftarrow c\,L_i$ avec $c\neq 0$ « multiplication d'une ligne par un scalaire non-nul » ;

(iii) $L_i\leftarrow L_i+cL_j$ avec $i\neq j$ et $c\in\mathbb{K}$ « ajout à une ligne d'un multiple d'une autre ligne ».

(b) Une *permutation de deux colonnes* de A est une transformation $C_i \overset{\text{(échange)}}{\longleftrightarrow} C_j$ avec $i\neq j$.

Remarque

Pour toute opération élémentaire l sur les lignes des matrices à n lignes et p (resp. $p+1$) colonnes et toute permutation c de deux colonnes de ces matrices, on a : $c\circ l = l\circ c$.

Lemme

Soit $(A'|B')$ une matrice obtenue à partir de $(A|B)$ après avoir effectué un nombre fini d'étapes de l'un des deux types suivants :

$L_i \leftrightarrow L_j$ avec $i\neq j$;

$L_i \leftarrow c_1L_1+\dots+c_iL_i+\dots+c_nL_n$ avec $1\le i\le n$, $c_1,...,c_n\in\mathbb{K}$ et $\underline{\underline{c_i\neq 0}}$.

Pour tout $X=\begin{pmatrix} x_1\\ \vdots\\ x_p\end{pmatrix}\in\mathbb{K}^p$, le système $(E) : AX=B$ équivaut au système $(E') : A'X=B'$.

Démonstration (idée)

En appliquant à (E) les opérations sur les lignes qui font passer de $(A|B)$ à $(A'|B')$, on constate que : si $(x_1, \dots, x_p)$ vérifie (E), alors $(x_1, \dots, x_p)$ vérifie (E').

Chaque étape se décompose en une succession d'opérations élémentaires sur les lignes. Pour être sûr de pouvoir remonter de (E') à (E), il reste à constater que ces opérations élémentaires sont réversibles :
(i) $L_i \leftrightarrow L_j$ avec $i \neq j$ a pour réciproque $L_i \leftrightarrow L_j$;
(ii) $L_i \leftarrow c\,L_i$ avec $c \neq 0$ a pour réciproque $L_i \leftarrow \frac{1}{c}\,L_i$;
(iii) $L_i \leftarrow L_i + cL_j$ avec $i \neq j$ et $c \in \mathbb{K}$ a pour réciproque $L_i \leftarrow L_i - cL_j$. □

Exemple

On va écrire une suite de matrices augmentées associées à des systèmes linéaires équivalents.
On peut commencer la résolution de (E_{cart}) ainsi, en entourant le « pivot » :

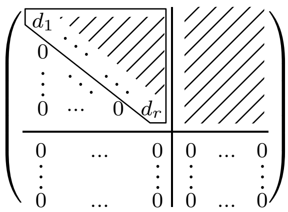
$$\left(\begin{array}{ccc|c} -4 & 12 & -5 & 1 \\ 1 & -3 & 2 & -1 \\ 2 & -6 & 1 & 1 \end{array}\right) \xrightarrow[\substack{\text{(pour utiliser le pivot 1, ce}\\ \text{qui n'est pas indispensable)}}]{L_1 \leftrightarrow L_2} \left(\begin{array}{ccc|c} 1 & -3 & 2 & -1 \\ -4 & 12 & -5 & 1 \\ 2 & -6 & 1 & 1 \end{array}\right) \xrightarrow[\substack{\text{puis } L_3 \leftarrow L_3 - 2L_1 \\ \text{(pour amener des 0 dans } C_1 \\ \text{strictement sous la diagonale)}}]{L_2 \leftarrow L_2 + 4L_1} \left(\begin{array}{ccc|c} 1 & -3 & 2 & -1 \\ \boxed{0} & 0 & 3 & -3 \\ \boxed{0} & 0 & -3 & 3 \end{array}\right).$$

3. Méthode du pivot de Gauss

On va travailler sur les colonnes de A dans $(A|B)$ en allant de la gauche vers la droite pour se ramener à un système triangulaire. À chaque étape on va chercher un scalaire non-nul (« pivot ») vers le bas à partir du terme diagonal, et éventuellement à droite ce qui nécessitera de faire un échange de colonnes. On amène ce pivot sur la diagonale puis des 0 sous ce pivot.

Proposition (*)

a) On peut passer par une suite finie d'opérations élémentaires sur les lignes et d'échanges de deux colonnes, de la matrice A à une matrice de la forme

$A' =$  avec $d_1 \neq 0$, ..., $d_r \neq 0$.

b) Le système (E) : $AX = B$ d'inconnue $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ équivaut au système (E') : $A'X' = B'$ d'inconnue $X' := \begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_p} \end{pmatrix}$, où les transformations du (a) envoient $(\overset{x_1 \dots x_p}{A} \,|\, B)$ sur $(\overset{x_{j_1} \dots x_{j_p}}{A'} \,|\, B')$.

[Noter que les lettres qui se trouvent au-dessus de A ou de A' représentent ici les noms des variables-coordonnées d'un vecteur de $\mathbb{K}^n$ et non pas les valeurs particulières qu'elles prennent pour le vecteur X. De plus, ce sont les éventuels échanges de colonnes qui feront passer de $x_1, \dots, x_p$ à $x_{j_1}, \dots, x_{j_p}$.]

c) On note $\mathscr{S}_E$ l'ensemble des solutions de (E) et $B' = \begin{pmatrix} b'_1 \\ \vdots \\ b'_n \end{pmatrix}$.

On a : $\quad \mathscr{S}_E \neq \emptyset \iff \underbrace{b'_{r+1} = \cdots = b'_n = 0}_{n-r \text{ conditions}}$.

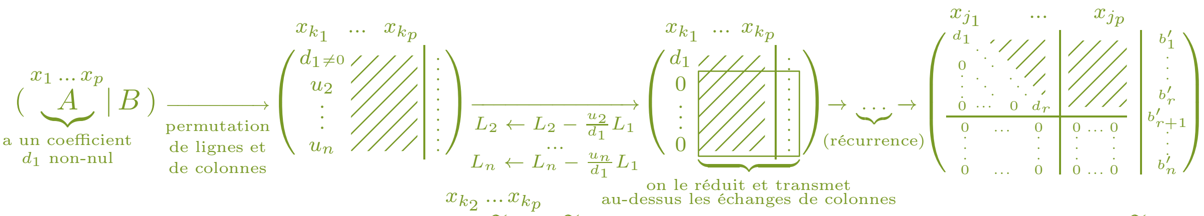
Dans ce cas, on obtient une équation paramétrique de $\mathscr{S}_E$ à partir de (E') en écrivant successivement x_{j_r}, ..., x_{j_1} en fonction des paramètres $\underbrace{t_1 := x_{j_{r+1}}, \dots, t_{p-r} := x_{j_p}}_{p-r \text{ paramètres}}$.

Ainsi, quand $r = n = p$ (« système de Cramer ») le système (E) a une unique solution.

(*) Pour un point de vue sans échanges de colonnes, voir la fin de ce paragraphe (« matrices échelonnées »).
Les échanges de deux colonnes peuvent permettre de minimiser les erreurs d'arrondis sur ordinateur : en travaillant sur la j^{e} colonne on choisirait comme pivot a_{i_0,j_0} avec $i_0, j_0 \geq j$ tels que $|a_{i_0,j_0}| = \max_{i',j' \geq j} |a_{i',j'}|$.

Démonstration

(a) On écarte le cas $A=\begin{pmatrix}0&\dots&0\\&\vdots&\\0&\dots&0\end{pmatrix}$ qui donnerait $r=0$ et $A'=\begin{pmatrix}0&\dots&0\\&\vdots&\\0&\dots&0\end{pmatrix}$. On a :



où on s'arrête quand le rectangle $(\overset{x_{k_2}\dots x_{k_p}}{\widetilde{A}}\,|\,\widetilde{B})$ qu'on réduit par récurrence sur n vérifie $\widetilde{A}=0$.

(b) (c) D'après le lemme du 2, et le fait qu'un échange de colonnes donne deux systèmes linéaires équivalents, (E) équivaut à :

$$(E')\left\{\begin{array}{rcl} d_1x_{j_1}+\quad \cdots\cdots \quad \cdots\;\cdots & = & b'_1\\ d_1x_{j_2}+\;\cdots\quad \cdots\;\cdots & = & b'_2\\ \ddots \qquad\qquad & & \\ d_1x_{j_r}+\;\cdots & = & b'_r\\ 0 & = & b'_{r+1}\\ \vdots & & \vdots\\ 0 & = & b'_n \end{array}\right. \qquad \text{c'est-à-dire à :}$$

$$b'_{r+1}=\cdots=b'_n=0 \quad \text{et} \quad \exists t_1,\dots,t_{p-r}\in\mathbb{K} \quad \left\{\begin{array}{l} x_{j_{r+1}}=t_1\\ \quad\vdots\\ x_{j_p}=t_{p-r}\\ \left.\begin{array}{l} x_{j_r}=\frac{1}{d_r}(\dots)\\ \quad\vdots\\ x_{j_1}=\frac{1}{d_1}(\dots)\end{array}\right\}\text{à exprimer avec } t_1,\dots,t_{p-r} \end{array}\right. .$$

□

Remarques

1. Quand (E) a strictement plus d'inconnues que d'équations (c'est-à-dire $p>n$) et $\mathscr{S}_E\neq\emptyset$ (par exemple (E) est homogène), on a : $\underbrace{p-r}_{\text{nombre de paramètres}}>n-r\geq 0$, donc $\mathscr{S}_E$ est infini.

> Tout système linéaire homogène dont le nombre d'inconnues est strictement supérieur au nombre d'équations admet une infinité de solutions.

2. La matrice $(A'|B')$ dépend des opérations sur les lignes et les colonnes qu'on a utilisées.

Cependant, on verra dans la suite du cours que :

– r ne dépend que de A, et sera appelé « le rang de A » ;

– $\mathscr{S}_E$ n'a pas d'équation paramétrique avec strictement moins que $p-r$ paramètres ;

(lorsque $\mathscr{S}_E\colon AX=B$ et $\mathscr{S}_E=\{UT+V\,;\,T\in\mathbb{K}^k\}$, on a : $k\geq\dim\overrightarrow{\mathscr{S}_E}=\dim\operatorname{Ker}A=p-r$)

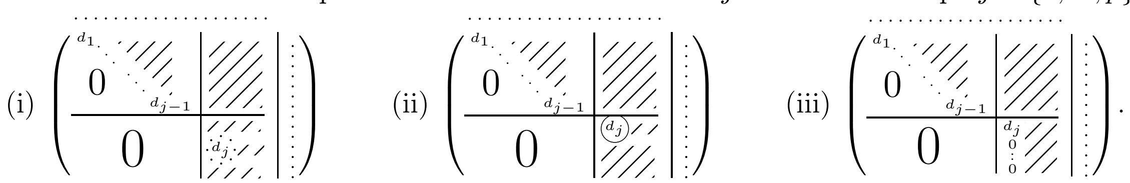
– une équation cartésienne obtenue par la méthode de Gauss à partir d'une équation paramétrique a un nombre d'égalités qui est minimal.

(lorsque $\mathscr{E}=\{AT+C\,;\,T\in\mathbb{K}^p\}$ et $\mathscr{E}\colon UX=V$, on a : $\operatorname{rg}U=n-\dim\overrightarrow{\mathscr{E}}=n-r$)

Algorithme 1 (« pivot total » au sens où on cherche d_j dans n'importe quelle colonne)

On détaille la méthode de la démonstration pour construire $(\overset{x_{j_1}\dots x_{j_p}}{A'}\,|\,B')$ comme au (b).

On utilise une construction par récurrence. On construit la j^{e} colonne à l'étape $j\in\{1,\dots,p\}$:



(i) On choisit une colonne $C_{j'}$ avec $j\leq j'\leq p$ qui a, à partir de la j^{e} ligne, au moins un coefficient $d_j\neq 0$ (s'il n'y a pas une telle colonne on arrête), puis on échange C_j et $C_{j'}$.

(ii) On échange la j^{e} ligne avec la ligne contenant d_j.

(iii) On amène des 0 strictement au-dessous de d_j par des opérations sur les lignes (composées de deux opérations élémentaires) de la forme $L_i\leftarrow c_iL_i+c_jL_j$ avec $c_i\neq 0$ quand $i>j$.

Après avoir travaillé sur toutes les colonnes en reportant les opérations sur les lignes au niveau de la colonne à droite de la barre verticale, on regarde si les membres de droite des égalités dont le membre de gauche est 0 vaut lui-même 0 (sinon le système n'a pas de solution) auquel cas on résout le système en remontant de la dernière égalité à la première égalité.

Remarques

En pratique, on exploitera la proposition précédente en allégeant la présentation :
– lorsque $B = 0$ on ne fera pas apparaître la dernière colonne (seconds membres nuls) ;
– on n'introduira les noms des variables au dessus des p premières colonnes qu'à partir du moment où un échange de colonnes aura été introduit (ce qui est rarement indispensable) ;
– on résoudra « de tête » le système triangulaire correspondant à la dernière étape de la méthode de Gauss, en partant de la dernière égalité et remontant vers la première.

Exemples (cf. l'introduction de la partie I de ce chapitre)

1. Fin de la résolution de (E_{cart}) :

$$\left(\begin{array}{ccc|c} -4 & 12 & -5 & 1 \\ 1 & -3 & 2 & -1 \\ 2 & -6 & 1 & 1 \end{array}\right) \to \underbrace{\ldots}_{\text{déjà vu}} \to \begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \left(\begin{array}{ccc|c} 1 & -3 & 2 & -1 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & -3 & 3 \end{array}\right) \end{array} \xrightarrow[C_2 \leftrightarrow C_3]{} \begin{array}{c} \begin{array}{ccc} x & z & y \end{array} \\ \left(\begin{array}{ccc|c} 1 & 2 & -3 & -1 \\ 0 & 3 & 0 & -3 \\ 0 & -3 & 0 & 3 \end{array}\right) \end{array} \xrightarrow[L_3 \leftarrow L_3 + L_2]{} \begin{array}{c} \begin{array}{ccc} x & z & y \end{array} \\ \left(\begin{array}{ccc|c} 1 & 2 & -3 & -1 \\ 0 & 3 & 0 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right) \end{array}.$$

En renommant dans la conclusion la variable y en t, et calculant d'abord z puis ensuite x, on obtient l'équation paramétrique suivante de $\mathscr{S}_{E_{\text{cart}}}$:

$$\mathscr{S}_{E_{\text{cart}}} : \begin{cases} x = \frac{1}{1}(-2(-1) + 3t - 1) = 3t + 1 \\ y = t \\ z = \frac{1}{3}(-3) = -1 \end{cases}, \qquad t \in \mathbb{K}.$$

2. Soit $(x, y) \in \mathbb{K}^2$. Existence de $(s, t) \in \mathbb{K}^2$ vérifiant $(E_{\text{param}}) \begin{cases} s - t = x - 1 \\ -s + t = y - 2 \end{cases}$?

Calcul par la méthode de Gauss : $\left(\begin{array}{cc|c} 1 & -1 & x-1 \\ -1 & 1 & y-2 \end{array}\right) \xrightarrow[L_2 \leftarrow L_2 + L_1]{} \left(\begin{array}{cc|c} 1 & -1 & /// \\ 0 & 0 & \underbrace{x + y - 3}_{\text{est-ce égal à 0 ?}} \end{array}\right)$.

Ainsi $\mathscr{D} : \begin{cases} x = s - t + 1 \\ y = -s + t + 2 \end{cases}$, $s, t \in \mathbb{K}$ a pour équation cartésienne : $\mathscr{D} : x + y - 3 = 0$.

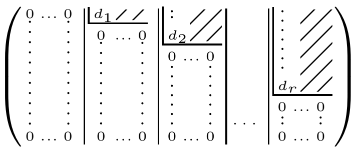
(Inutile de calculer explicitement les valeurs des paramètres s et t en fonction de x et y.)

3. Le système (E_C) $\begin{cases} 3x + 2y = 5 \\ 2x - y = 4 \end{cases}$ se résout ainsi : $\left(\begin{array}{cc|c} 3 & 2 & 5 \\ 2 & -1 & 4 \end{array}\right) \xrightarrow[L_2 \leftarrow 3L_2 - 2L_1]{} \left(\begin{array}{cc|c} 3 & 2 & 5 \\ 0 & -7 & 2 \end{array}\right)$.

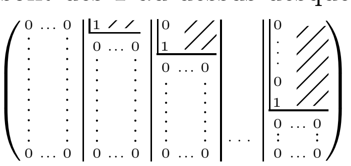
C'est un système de Cramer d'unique solution donnée par : $\begin{cases} x = \frac{1}{3}\big(-2(-\frac{2}{7}) + 5\big) = \frac{13}{5} \\ y = -\frac{2}{7} \end{cases}$.

Définition (« méthode de Gauss sans échange de colonnes »)

(a) Une *matrice échelonnée* — suivant les lignes — est une matrice telle que le nombre de termes nuls au début de chaque ligne augmente lorsqu'on passe d'une ligne à la suivante, et ce nombre augmente même strictement lorsqu'on passe d'une ligne non nulle à la suivante.

Il s'agit d'une matrice de la forme  avec $d_1 \neq 0$, ..., $d_r \neq 0$.

(b) Une *matrice échelonnée réduite* est une matrice échelonnée dans laquelle les premiers termes non-nuls sur chaque ligne sont des 1 au-dessus desquels se trouvent des 0. Elle s'écrit :

.

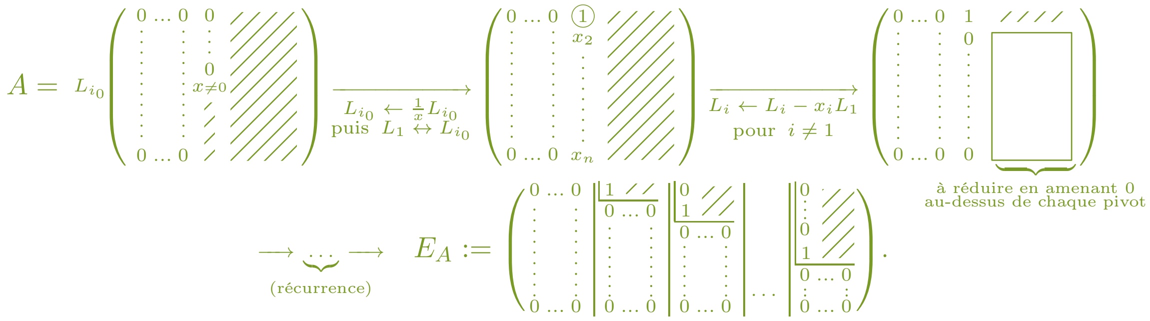
Proposition

Il existe une <u>unique</u> matrice échelonnée réduite E_A qui se déduit de A par des opérations élémentaires sur les lignes.

Démonstration (idées)

• Existence ?

Elle s'obtient par exemple ainsi :



• Unicité ?

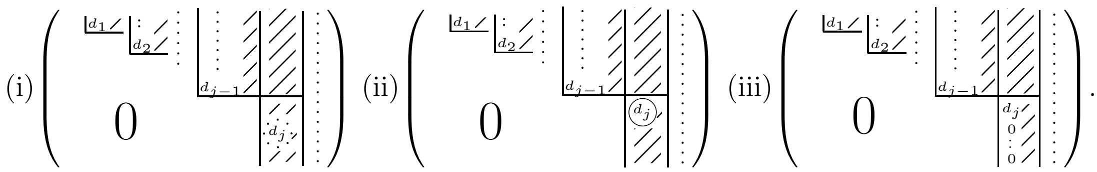
On vérifie par récurrence sur p qu'une matrice échelonnée réduite A' avec r lignes non-nulles déduite de A par des opérations élémentaires sur les lignes est déterminée par A.
On remarque tout d'abord que les relations $\alpha_1 C_1 + \cdots + \alpha_p C_p = 0$ sont vérifiées pour les mêmes $(\alpha_1, ..., \alpha_p) \in \mathbb{K}^p$ au niveau de A et de A' (il s'agit des solutions du système $(H) : AX = 0$).

Lorsque $p = 0$ la seule matrice échelonnée réduite est la matrice vide. En supposant que la proposition est vraie avec $p-1$ à la place de p, les colonnes $C_1, ..., C_{p-1}$ de A' sont déterminées par A. Il reste à constater que la colonne C_p de A' est nulle quand $A = 0$, et $\underbrace{\text{sinon}}_{\text{avec } r \neq 0}$ contient :

– un seul terme non-nul égal à 1 au niveau de la r^{e} ligne si $C_p \notin \operatorname{Vect}(C_1, ..., C_{p-1})$;
– les coordonnées de C_p suivant les r colonnes des pivots si $C_p \in \operatorname{Vect}(C_1, ..., C_{p-1})$. □

Algorithme 2 (« pivot partiel » au sens où on cherche d_j dans la 1^{re} colonne possible)

On détaille la méthode de la démonstration pour construire $(A'|B')$ avec A' échelonnée.
On utilise une construction par récurrence. On s'intéresse à la j^{e} ligne à l'étape $j \in \{1, ..., p\}$:



(i) On se place dans la colonne la plus à gauche qui a, à partir de la j^{e} ligne, au moins un coefficient $d_j \neq 0$ (s'il n'y a pas une telle colonne on arrête).

(ii) On échange la j^{e} ligne avec la ligne contenant d_j.

(iii) On amène des 0 strictement au-dessous de d_j par des opérations sur les lignes (composées de deux opérations élémentaires) de la forme $L_i \leftarrow c_i L_i + c_j L_j$ avec $c_i \neq 0$ quand $i > j$.

Pour obtenir une matrice échelonnée réduite, on amènerait aussi à la j^{e} étape dans (iii) des 0 strictement au-dessus du pivot d_j par des opérations sur les lignes de la forme $L_i \leftarrow c_i L_i + c_j L_j$ avec $c_i \neq 0$ quand $i < j$, puis en fin de calcul on diviserait chacune des lignes non-nulles par son premier coefficient non-nul (noté d_j à la j^{e} étape).

Après avoir travaillé sur toutes les lignes en reportant les opérations sur les lignes au niveau de la colonne qui se trouve à droite de la barre verticale, on regarde si les membres de droite des égalités dont le membre de gauche est 0 vaut lui-même 0 (sinon le système n'a pas de solution), auquel cas on résout le système en remontant de la dernière égalité à la première égalité.

On peut adapter cet algorithme au cas des matrices à valeurs dans $\mathbb{Z}$.

Pour cela on remplace au (iii) de la j^{e} étape les opérations sur les lignes par des opérations correspondant à la multiplication à gauche par un certain $P \in \mathfrak{M}(n, \mathbb{Z})$ inversible dans $\mathfrak{M}(n, \mathbb{Z})$.
On note a_1, ..., a_n les coefficients dans la j^{e} colonne. On fixe $i > j$ tel que $a_i \neq 0$.
On pose : $d_{i,j} = \operatorname{pgcd}(a_i, a_j) \in \mathbb{Z} \setminus \{0\}$, $b_i = \frac{a_i}{d_{i,j}}$ et $b_j = \frac{a_j}{d_{i,j}}$.
D'après le théorème de Bézout, il existe $u_i, u_j \in \mathbb{Z}$ tels que $u_i b_i + u_j b_j = 1$.
L'opération $(L_i, L_j) \leftarrow (u_i L_i + u_j L_j, -b_j L_i + b_i L_j)$ amène $d_{i,j}$ sur la j^{e} ligne et 0 sur la i^{e} ligne.

Remarques (lien entre les deux présentations de l'algorithme du pivot de Gauss)

1. En appliquant à A les opérations élémentaires sur les lignes — mais pas sur les colonnes — qui ont permis de passer de A à une matrice A' en suivant l'algorithme 1 avec à chaque étape le pivot dans la colonne la plus à gauche, on obtient une matrice échelonnée.

2. Réciproquement, à partir d'une matrice échelonnée déduite de A par des opérations élémentaires sur les lignes, en déplaçant en premières positions les colonnes contenant les termes non-nuls au début de chaque ligne, on obtient une matrice A' associée à l'algorithme 1.

II. Calcul matriciel

Dans cette partie, on fixe $n, p, q, r \in \mathbb{N} \setminus \{0\}$.

1. Opérations usuelles

On considère un système d'équations linéaires de n équations à p inconnues dans $\mathbb{K}$:

$$(E) \quad \begin{cases} a_{11}x_1 + \cdots + a_{1p}x_p &= b_1 \\ \qquad \cdots & \\ a_{n1}x_1 + \cdots + a_{np}x_p &= b_n \end{cases} \qquad \text{d'inconnue } (x_1, ..., x_p) \in \mathbb{K}^p$$

où sont donnés $a_{11},..., a_{1p}, a_{21},..., a_{2p}, \ldots, a_{n1},..., a_{np} \in \mathbb{K}$ et $b_1,..., b_n \in \mathbb{K}$.

On lui associe : $A := \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ & \cdots & \\ a_{n1} & \cdots & a_{np} \end{pmatrix}$ *matrice de* (E), $X := \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ *inconnue de* (E) et $B := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ *second membre de* (E). On écrit en abrégé $(E) : AX = B$.

Cette notation $(E) : AX = B$ va être un guide pour définir le produit des matrices.

Définition

(a) Une *matrice* $\underbrace{n}_{\text{nb lignes}} \times \underbrace{p}_{\text{nb colonnes}}$ *à coefficients dans* $\mathbb{K}$ est un tableau(**) de la forme :

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ & \cdots & \\ a_{n1} & \cdots & a_{np} \end{pmatrix} \quad \text{avec} \quad a_{1,1}, ..., a_{1p}; \ldots ; a_{n1}, ..., a_{np} \in \mathbb{K}.$$

On écrit aussi $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ et appelle *coefficient de A d'indice (i, j)* le nombre a_{ij}.
Cela sous-entend que « a_{ij} est placé à l'intersection de la i^{e} ligne et de la j^{e} colonne ».

(b) On note $\mathfrak{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices $n \times p$ à coefficients dans $\mathbb{K}$.

On pose ensuite $\mathfrak{M}(n, \mathbb{K}) := \mathfrak{M}_{n,n}(\mathbb{K})$ *matrices carrées d'ordre n à coefficients dans* $\mathbb{K}$ et $\mathfrak{M}_{0,p}(\mathbb{K}) = \mathfrak{M}_{n,0}(\mathbb{K}) = \mathfrak{M}(0, \mathbb{K}) := \{0\} \subseteq \mathbb{K}$ où à la matrice vide () est identifiée à 0.

Exemples

Matrice *nulle* $0 := \underbrace{\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}}_{p} \Big\updownarrow n$ et matrice (carrée) *identité* $I_n := \underbrace{\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}}_{n} \Big\updownarrow n$.

Matrice *scalaire* : (a_{11}), s'identifie à a_{11}.
Matrice *ligne* : $(a_{11} \ \cdots \ a_{1p})$, à ne pas confondre avec $(a_{11}, \ldots, a_{1p}) \in \mathbb{K}^p$.
Matrice *colonne* : $\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}$, sera par convention identifiée avec $(a_{11}, \ldots, a_{n1}) \in \mathbb{K}^n$.
Matrice (carrée) *diagonale* : $D = \begin{pmatrix} \ddots & & 0 \\ & \ddots & \\ 0 & & \ddots \end{pmatrix}$.
Matrice (carrée) *triangulaire supérieure* : $U = \begin{pmatrix} \ddots & & * \\ & \ddots & \\ 0 & & \ddots \end{pmatrix}$.

(**) Du point de vue mathématique, un « tableau » est un élément de $\mathbb{K}^{\{1,...,n\} \times \{1,...,p\}}$.

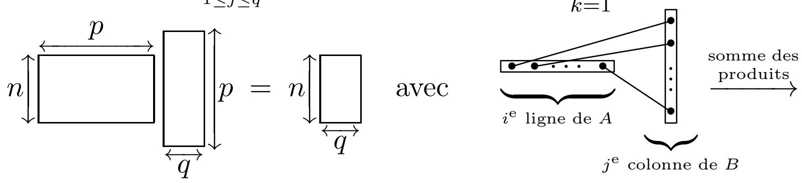
Définition

(a) Soient $\lambda \in \mathbb{K}$ et $\quad A = (a_{ij})_{\substack{1\le i\le n\\1\le j\le p}}, \ B = (b_{ij})_{\substack{1\le i\le n\\1\le j\le p}} \in \mathfrak{M}_{n,p}(\mathbb{K})$.

On pose : $\left| \begin{array}{l} \lambda A := (\lambda a_{ij})_{\substack{1\le i\le n\\1\le j\le p}} \ \text{ et } \ -A := (-a_{ij})_{\substack{1\le i\le n\\1\le j\le p}} \\ A+B := (a_{ij}+b_{ij})_{\substack{1\le i\le n\\1\le j\le p}} \ \text{ et } \ A-B := (a_{ij}-b_{ij})_{\substack{1\le i\le n\\1\le j\le p}} \end{array} \right.$.

(b) Soient $\ A = (a_{ij})_{\substack{1\le i\le n\\1\le j\le p}} \in \mathfrak{M}_{n,p}(\mathbb{K})$ et $\ B = (b_{ij})_{\substack{1\le i\le p\\1\le j\le q}} \in \mathfrak{M}_{p,q}(\mathbb{K})$.

On pose : $\ AB := (c_{ij})_{\substack{1\le i\le n\\1\le j\le q}} \in \mathfrak{M}_{n,q}(\mathbb{K})$ avec $\ c_{ij} = \sum\limits_{k=1}^{p} a_{ik}b_{kj} = a_{i1}b_{1j} + \cdots + a_{ip}b_{pj}$.

Schéma : 

coefficient de AB d'indice (i,j) .

Exemples

On considère : $U = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, $V = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$, $W = \begin{pmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Les sommes et produits possibles de 2 matrices différentes parmi U, V, W, X sont :

$U + V = V + U = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}$;

$UV = \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}$ et $VU = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, donc $UV \neq VU$;

$UW = \begin{pmatrix} -3 & -5 & 8 \\ 3 & 5 & -8 \end{pmatrix}$ et $VW = \begin{pmatrix} 3 & 3 & -4 \\ 6 & 6 & -8 \end{pmatrix}$;

$WX = \begin{pmatrix} -y+2z \\ 3x+4y-6z \end{pmatrix}$, ce qui est compatible avec la notation « $(E) : AX = B$ ».

Remarque

Dans l'exemple ci-dessus, on pourrait calculer U^n et V^n pour $n \in \mathbb{N}$ car U et V sont carrées.

Dans le cas favorable où des matrices carrées A et B commutent on peut déduire des A^n et B^n les puissances de $A + B$. En effet, la formule du binôme se généralise facilement ainsi :

si $\ A, B \in \mathfrak{M}(p, \mathbb{K})$ vérifient $AB = BA$, alors $\ (A+B)^n = \sum\limits_{k=0}^{n} \binom{n}{k} A^{n-k}B^k\ $ pour tout $n \in \mathbb{N}$.

Proposition

(a) Soient $\ A, B, C \in \mathfrak{M}_{n,p}(\mathbb{K})$, et $\ \lambda, \mu \in \mathbb{K}$.

On a : $\left| \begin{array}{l} 1A = A, \ \ A + 0 = A \ \text{ et } A + (-A) = 0 \\ (A+B)+C = A+(B+C) \ \text{ et } \ A+B = B+A \\ \lambda(A+B) = (\lambda A) + (\lambda B) \ \text{ et } \ (\lambda+\mu)A = (\lambda A) + (\mu A) \\ \lambda(\mu A) = (\lambda\mu)A \end{array} \right.$

(Cela s'exprimera en disant que « $(\mathfrak{M}_{n,p}(\mathbb{K}), +, .)$ est un $\mathbb{K}$-espace vectoriel ».)

(b) Soient $\ A, A_1, A_2 \in \mathfrak{M}_{n,p}(\mathbb{K})$, $\ B, B_1, B_2 \in \mathfrak{M}_{p,q}(\mathbb{K})$, $\ C \in \mathfrak{M}_{q,r}(\mathbb{K})$, et $\ \lambda \in \mathbb{K}$.

On a : $\left| \begin{array}{l} I_nA = A = AI_p \\ (AB)C = A(BC) \\ A(B_1+B_2) = (AB_1) + (AB_2) \ \text{ et } \ (A_1+A_2)B = (A_1B) + (A_2B) \\ \lambda(AB) = (\lambda A)B = A(\lambda B) \end{array} \right.$

En particulier : $\ (\mathfrak{M}(n, \mathbb{K}), +, .)$ est un anneau. ←[non-commutatif quand $n \geq 2$]

Démonstration

Exercice. □

Définition-Proposition

(a) Une *décomposition en blocs de taille* $\overbrace{(n_1,\dots,n_N)}^{\in\,\mathbb{N}^N}\times\overbrace{(p_1,\dots,p_P)}^{\in\,\mathbb{N}^P}$ d'une matrice $A\in\mathfrak{M}_{n,p}(\mathbb{K})$ est la donnée de matrices $A_{i,j}\in\mathfrak{M}_{n_i,p_j}(\mathbb{K})$, $i\in\{1,\dots,N\}$ et $j\in\{1,\dots,P\}$, telles que :

$$A=\begin{array}{c}\begin{array}{ccc}\overset{p_1}{\longleftrightarrow} & & \overset{p_P}{\longleftrightarrow}\end{array}\\ \begin{pmatrix} A_{1,1} & \cdots & A_{1,P}\\ \vdots & & \vdots\\ A_{N,1} & \cdots & A_{N,P}\end{pmatrix}\end{array}\begin{array}{l}\updownarrow n_1\\ \vdots\\ \updownarrow n_N\end{array}$$ où les parenthèses des matrices $A_{i,j}$ ont été enlevées.

(b) Soient $\lambda\in\mathbb{K}$ et deux matrices $A,B\in\mathfrak{M}_{n,p}(\mathbb{K})$ décomposées en blocs $\big(A_{i,j}\big)_{\substack{1\le i\le N\\1\le j\le P}}$ et $\big(B_{i,j}\big)_{\substack{1\le i\le N\\1\le j\le P}}$ de même taille. On a : $\lambda A=\big(\lambda A_{i,j}\big)_{\substack{1\le i\le N\\1\le j\le P}}$ et $A+B=\big(A_{i,j}+B_{i,j}\big)_{\substack{1\le i\le N\\1\le j\le P}}$.

(c) Soient $A\in\mathfrak{M}_{n,p}(\mathbb{K})$ et $B\in\mathfrak{M}_{p,q}(\mathbb{K})$ décomposées en blocs $A=\big(A_{i,j}\big)_{\substack{1\le i\le N\\1\le j\le P}}$ de taille $(n_1,\dots,n_N)\times(p_1,\dots,p_P)$ et $B=\big(B_{j,k}\big)_{\substack{1\le j\le P\\1\le k\le Q}}$ de taille $(p_1,\dots,p_P)\times(q_1,\dots,q_Q)$.

On a la décomposition en blocs suivante : $\quad AB=\big(C_{i,j}\big)_{\substack{1\le i\le N\\1\le j\le Q}}$ avec $C_{i,j}=\displaystyle\sum_{k=1}^{P}A_{i,k}\,B_{k,j}$.

Démonstration

(b) Calcul immédiat.

(c) Soit $(u,v)\in\{1,\dots,n\}\times\{1,\dots,q\}$. On peut écrire u et v sous la forme :

$u=n_1+\dots+n_{i-1}+r$ avec $1\le r\le n_i$ et $v=q_1+\dots+q_{j-1}+s$ avec $1\le s\le q_j$.

En notant $M(i,j)$ le coefficient d'indice (i,j) d'une matrice M, on a :

$$(AB)(u,v)=\sum_{w=1}^{p_1}A(u,w)B(w,v)+\sum_{w=p_1+1}^{p_1+p_2}A(u,w)B(w,v)+\cdots+\sum_{w=p_1+\dots+p_{P-1}+1}^{p_1+\dots+p_P}A(u,w)B(w,v)$$

$$=\sum_{t=1}^{p_1}A(u,t)B(t,v)+\sum_{t=1}^{p_2}A(u,p_1+t)B(p_1+t,v)+\cdots+\sum_{t=1}^{p_P}A(u,p_1+\dots+p_{P-1}+t)B(p_1+\dots+p_{P-1}+t,v)$$

$$=\sum_{t=1}^{p_1}A_{i,1}(r,t)B_{1,j}(t,s)+\sum_{t=1}^{p_1}A_{i,2}(r,t)B_{2,j}(t,s)+\cdots+\sum_{t=1}^{p_P}A_{i,P}(r,t)B_{1,j}(t,s)$$

$$=\sum_{k=1}^{P}\Big(\sum_{t=1}^{p_k}A_{i,k}(r,t)B_{k,j}(t,s)\Big)=\sum_{k=1}^{P}(A_{i,k}B_{k,j})(r,s)=C_{i,j}(r,s).$$

D'où le résultat. □

Exemples

On va délimiter des blocs par des traits horizontaux et verticaux.

(1) Exemple de calcul pour illustrer le point (c) :

$$\begin{pmatrix}0&1&1\\1&0&1\\2&1&-1\end{pmatrix}\begin{pmatrix}1&3\\2&4\\5&6\end{pmatrix}=\left(\begin{array}{cc|c}0&1&1\\1&0&1\\\hline 2&1&-1\end{array}\right)\left(\begin{array}{cc}1&3\\2&4\\\hline 5&6\end{array}\right)=\left(\begin{array}{c}\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}1&3\\2&4\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}5&6\end{pmatrix}\\\hline \begin{pmatrix}2&1\end{pmatrix}\begin{pmatrix}1&3\\2&4\end{pmatrix}+(-1)\begin{pmatrix}5&6\end{pmatrix}\end{array}\right)=\left(\begin{array}{c}\begin{pmatrix}2&4\\1&3\end{pmatrix}+\begin{pmatrix}5&6\\5&6\end{pmatrix}\\\hline \begin{pmatrix}4&10\end{pmatrix}+\begin{pmatrix}-5&-6\end{pmatrix}\end{array}\right)=\begin{pmatrix}7&10\\6&9\\-1&4\end{pmatrix}.$$

(2) Cas des matrices augmentées :

$P(A|B)=(PA|PB)$ quand $P\in\mathfrak{M}(n,\mathbb{K})$, $A\in\mathfrak{M}_{n,p}(\mathbb{K})$ et $B\in\mathbb{K}^n$.

2. Transposée d'une matrice, trace d'une matrice carrée

Définition (matrices rectangulaires)

Soit $A=\begin{pmatrix}a_{11}&\cdots&a_{1p}\\&\cdots&\\a_{n1}&\cdots&a_{np}\end{pmatrix}\in\mathfrak{M}_{n,p}(\mathbb{K})$.

On appelle *matrice transposée de* A la matrice ${}^tA:=\begin{pmatrix}a_{11}&\cdots&a_{n1}\\&\cdots&\\a_{1p}&\cdots&a_{np}\end{pmatrix}\in\mathfrak{M}_{p,n}(\mathbb{K})$.

Cela signifie que : ${}^tA=(a'_{i\,j})_{\substack{1\le i\le p\\1\le j\le n}}$ où $a'_{i\,j}:=a_{j\,i}$ quand $1\le i\le p$ et $1\le j\le n$.

Proposition

Soient $A\in\mathfrak{M}_{n,p}(\mathbb{K})$ et $B\in\mathfrak{M}_{p,q}(\mathbb{K})$. On a : ${}^tB\,{}^tA$ est définie et ${}^t(AB)={}^tB\,{}^tA$.

Démonstration

On pose $A=(a_{ij})_{\substack{1\le i\le n\\1\le j\le p}}$ et $B=(b_{ij})_{\substack{1\le i\le p\\1\le j\le q}}$.

Donc : ${}^tA=(a'_{i\,j})_{\substack{1\le i\le p\\1\le j\le n}}$ et ${}^tB=(b'_{i\,j})_{\substack{1\le i\le q\\1\le j\le p}}$, avec $a'_{i\,j}:=a_{j\,i}$ et $b'_{i\,j}:=b_{j\,i}$.

D'une part : $AB=(c_{ij})_{\substack{1\le i\le n\\1\le j\le q}}$ avec $c_{ij}=\sum\limits_{k=1}^{p}a_{ik}b_{kj}$, puis ${}^t(AB)=(c'_{i\,j})_{\substack{1\le i\le q\\1\le j\le n}}$ avec $c'_{i\,j}:=c_{j\,i}$.

D'autre part : ${}^tB{}^tA=(d_{ij})_{\substack{1\le i\le q\\1\le j\le n}}$ avec $d_{ij}=\sum\limits_{k=1}^{p}b'_{ik}a'_{kj}$.

Finalement : $c'_{i\,j}=\sum\limits_{k=1}^{p}a_{jk}b_{ki}=\sum\limits_{k=1}^{p}a'_{kj}b'_{ik}=d_{ij}$, ce qui s'écrit ${}^t(AB)={}^tB\,{}^tA$.

Variante

${}^t(AB)(i,j)=(AB)(j,i)=\sum\limits_{k=1}^{p}A(j,k)B(k,i)=\sum\limits_{k=1}^{p}({}^tB)(i,k)({}^tA)(k,j)=({}^tB\,{}^tA)(i,j)$. □

Définition (matrices carrées)

Soit $A=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\&\cdots&\\a_{n1}&\cdots&a_{nn}\end{pmatrix}\in\mathfrak{M}(n,\mathbb{K})$.

(a) On dit que A *est symétrique* si ${}^tA=A$ (c'est-à-dire $a_{j\,i}=a_{i\,j}$ quand $1\le i<j<\le n$).

(b) On appelle *trace de* A, l'élément de $\mathbb{K}$ noté $\operatorname{tr}A$, défini par :

$$\operatorname{tr}A=\sum_{i=1}^{n}a_{i,i}=a_{1,1}+\cdots+a_{n,n}.$$

Proposition

Soient $A\in\mathfrak{M}_{n,p}(\mathbb{K})$ et $B\in\mathfrak{M}_{p,n}(\mathbb{K})$. On a : $\operatorname{tr}(AB)=\operatorname{tr}(BA)$.

Démonstration

On pose $A=(a_{ij})_{\substack{1\le i\le n\\1\le j\le p}}$ et $B=(b_{ij})_{\substack{1\le i\le p\\1\le j\le n}}$.

On a : $AB=(c_{ij})_{\substack{1\le i\le n\\1\le j\le q}}$ avec $c_{ij}=\sum\limits_{k=1}^{p}a_{ik}b_{kj}$, donc $\operatorname{tr}(AB)=\sum\limits_{i=1}^{n}c_{i,i}=\sum\limits_{i=1}^{n}\Big(\sum\limits_{k=1}^{n}a_{ik}b_{ki}\Big)$.

On conclut en intervertissant $\sum\limits_{i=1}^{n}$ et $\sum\limits_{k=1}^{n}$. □

3. Matrices carrées inversibles

Définition-Proposition

(a) Soit $A\in\mathfrak{M}(n,\mathbb{K})$. On dit que A est *inversible* s'il existe $B\in\mathfrak{M}(n,\mathbb{K})$ tel que :

$$AB=I_n \quad\text{et}\quad BA=I_n.$$

Dans ce cas B est unique, notée A^{-1}, et appelée *inverse de* A.

(b) On note $GL(n,\mathbb{K})$ l'ensemble des éléments inversibles de $\mathfrak{M}(n,\mathbb{K})$.

Démonstration

On se place dans la situation du (a) avec deux « inverses » potentiels : si $B,C\in\mathfrak{M}(n,\mathbb{K})$ vérifient $AB=I_n$ et $CA=I_n$, alors $C=CI_n=C(AB)=(CA)B=I_nB=B$. □

Remarque (importante)

On suppose que $A\in\mathfrak{M}(n,\mathbb{K})$ est inversible et $B\in\mathbb{K}^n$.

Le système $(E):AX=B$ d'inconnue $X\in\mathbb{K}^n$ équivaut à l'équation $A^{-1}(AX)=A^{-1}B$.
Il a donc pour unique solution $X:=A^{-1}B$. $\longleftarrow$[dans ce cas le nombre de paramètres $n-r$ pour les solutions est nul]

Proposition

(a) Soit $A\in\mathfrak{M}(n,\mathbb{K})$ inversible. La matrice A^{-1} est inversible et $\boxed{\left(A^{-1}\right)^{-1}=A}$.

(b) Soit $A\in\mathfrak{M}(n,\mathbb{K})$ inversible. La matrice ${}^tA$ est inversible et $\boxed{({}^tA)^{-1}={}^t(A^{-1})}$.

(c) Soient $A,B\in\mathfrak{M}(n,\mathbb{K})$ inversibles. La matrice AB est inversible et $\boxed{(AB)^{-1}=B^{-1}A^{-1}}$.

Démonstration

(a) Il suffit de remarquer qu'en posant $B := A^{-1}$), on a : $A^{-1}\,B = I_n$ et $B\,A^{-1} = I_n$.

(b) D'après la fin du 2, on a : ${}^t\!A\,{}^t(A^{-1}) = {}^t(\underbrace{A^{-1}A}_{I_n}) = I_n$ et ${}^t(A^{-1}){}^t\!A = {}^t(\underbrace{AA^{-1}}_{I_n}) = I_n$. Donc ${}^t\!A$ est inversible d'inverse ${}^t(A^{-1})$.

(c) Il suffit encore de tester l'inverse proposé :
$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$
et $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$. □

Lemme

On se donne une suite finie d'opérations élémentaires sur les lignes des matrices à n lignes. Il existe une matrice inversible $P \in \mathfrak{M}(n,\mathbb{K})$ telle que cette suite d'opérations élémentaires envoie toute matrice $M \in \mathfrak{M}_{n,p}(\mathbb{K})$ sur PM. ⟵[il en résulte d'ailleurs que P est l'image de I_n]

Démonstration (au programme)

On est ramené à vérifier le lemme pour une seule opération élémentaire :

$$L_i \leftrightarrow L_j \text{ avec } i \neq j \text{ est associé à } \quad P \overset{\substack{\text{penser au}\\ \text{cas } M=I_n}}{:=} \underbrace{\begin{pmatrix} 1 &&&&&&&&\\ &\ddots&&&&&&&\\ &&1&&&&&&\\ &&&0&\cdots&0&1&&\\ &&&\vdots&1&&0&&\\ &&&\vdots&&\ddots&\vdots&&\\ &&&0&&1&\vdots&&\\ &&&1&0&\cdots&0&&\\ &&&&&&&1&\\ &&&&&&&&\ddots\\ &&&&&&&&&1 \end{pmatrix}}_{0 \text{ ailleurs}} \begin{array}{l} \\ \\ \\ \leftarrow \text{ligne } i \\ \\ \\ \\ \leftarrow \text{ligne } j \\ \\ \\ \\ \end{array} ;$$

$$L_i \leftarrow c\,L_i \text{ avec } c \neq 0 \text{ est associé à } \quad P \overset{\substack{\text{penser au}\\ \text{cas } M=I_n}}{:=} \underbrace{\begin{pmatrix} 1 &&&&&\\ &\ddots&&&&\\ &&1&&&\\ &&&c&&\\ &&&&1&\\ &&&&&\ddots\\ &&&&&&1 \end{pmatrix}}_{0 \text{ ailleurs}} \begin{array}{l} \\ \\ \\ \leftarrow \text{ligne } i \\ \\ \\ \\ \end{array} ;$$

$$L_i \leftarrow L_i + cL_j \text{ avec } i \neq j \text{ et } c \in \mathbb{K} \text{ est associé à } \quad P \overset{\substack{\text{penser au}\\ \text{cas } M=I_n}}{:=} \begin{array}{c} \text{colonne } j \\ \downarrow \\ \underbrace{\begin{pmatrix} 1 &&\\ &\ddots & c\\ &&\ddots\\ &&&1 \end{pmatrix}}_{0 \text{ ailleurs}} \end{array} \begin{array}{l} \\ \\ \leftarrow \text{ligne } i \\ \\ \\ \end{array} .$$

Chacune des trois matrices P ci-dessus est inversible car en appliquant l'opération élémentaire en question et sa réciproque de la forme $M \mapsto QM$ à I_n, on retrouve I_n. □

Définition (peu utilisée)

Les matrices associées à chaque opération élémentaire sur les lignes qui apparaissent dans la démonstration précédente s'appellent les *matrices d'opérations élémentaires*.

Proposition

Soit $A \in \mathfrak{M}(n,\mathbb{K})$. Les propriétés suivantes sont équivalentes :

(i) A est inversible ;
(ii) il existe $B \in \mathfrak{M}(n,\mathbb{K})$ tel que $AB = I_n$;
(iii) il existe $C \in \mathfrak{M}(n,\mathbb{K})$ tel que $CA = I_n$;
(iv) il existe une suite finie d'opérations élémentaires sur les lignes qui envoie A sur I_n.

Dans les cas (ii) à (iv), on a respectivement : $A^{-1} = B$, $A^{-1} = C$, et A^{-1} est l'image de I_n par la suite finie d'opérations élémentaires.

Démonstration (idée)

On vérifie que (i) ⇒ (ii) ⇒ (iv) ⇒ (i) et (i) ⇔ (iii).

• Par définition de l'inversibilité, on a : (i) ⇒ (ii) et (i) ⇒ (iii).

• On suppose (iv) vrai, ce qui fournit une suite finie Σ d'opérations élémentaires sur les lignes envoyant A sur I_n qui d'après le lemme est de la forme $M \mapsto PM$ pour un certain $P \in \mathfrak{M}(n,\mathbb{K})$

inversible. En choisissant M tout d'abord égal à A, et ensuite égal à I_n, on obtient : $PA = I_n$ puis $A = P^{-1}$ (par produit à gauche par P^{-1}) et l'image de I_n par Σ est P.

Ainsi (i) est vrai et A^{-1} qui est égal à P est bien l'image de I_n par la suite Σ.

• On montre que (ii) $\Rightarrow$ (iv). On sait que pour toute $M_0 \in \mathfrak{M}_{n,p}(\mathbb{K})$ il existe une suite finie d'opérations élémentaires sur les lignes qui transforme M_0 en une matrice échelonnée réduite :

Par le lemme, cette suite d'opérations s'écrit $M \mapsto PM$ pour un certain $P \in GL(n, \mathbb{K})$

On suppose maintenant (ii) vrai et choisit $M_0 = A$. Si la dernière ligne de la matrice PA à droite est nulle, alors celle de $(PA)(BP^{-1})$ aussi. Cela est faux, car $(PA)(BP^{-1}) = PP^{-1} = I_n$. La matrice PA a donc exactement n blocs, puis comme elle a n colonnes, on a : $PA = I_n$.

Ainsi (iv) est vérifié. Donc (i) est vrai car « (iv) $\Rightarrow$ (i) », puis $A^{-1} = A^{-1}(AB) = B$.

• On suppose (iii) vrai. La matrice C vérifie l'analogue du (ii) avec le rôle de A. On vient de voir que « (ii) $\Rightarrow$ (i) ». On en déduit que C est inversible, puis $A = C^{-1}(CA) = C^{-1}$.

Ainsi (i) est vrai et $A^{-1} = C$. □

Remarques

1. La condition (iv) de la proposition signifie que I_n est l'unique matrice échelonnée réduite qui se déduit de A par des opérations élémentaires sur les lignes.

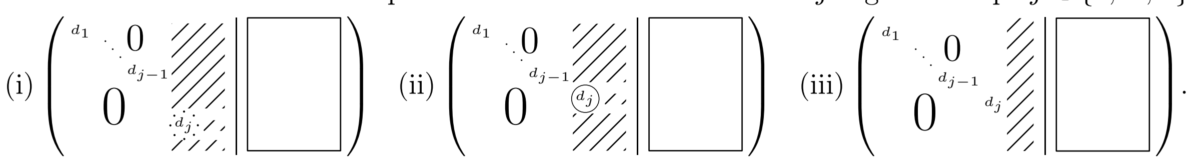
2. L'ensemble des transformations obtenues à l'issue d'une suite finie d'opérations élémentaires sur les lignes des matrices à n lignes est égal à celui des transformations $M \mapsto PM$ avec $P \in \mathfrak{M}(n, \mathbb{K})$ inversible. En effet :
($\subseteq$) toute suite finie d'opérations élémentaires sur les lignes des matrices à n lignes s'écrit d'après le lemme $M \mapsto PM$ pour un certain $P \in \mathfrak{M}(n, \mathbb{K})$ inversible ;
($\supseteq$) étant donné $P \in \mathfrak{M}(n, \mathbb{K})$ inversible, il existe d'après la proposition une suite finie Σ d'opérations élémentaires sur les lignes qui envoie P^{-1} sur I_n, et comme Σ est d'après le lemme de la forme $M \mapsto QM$ avec $Q \in \mathfrak{M}(n, \mathbb{K})$ inversible, l'égalité $\underbrace{QP^{-1} = I_n}_{\text{signifie que } Q = P}$ obtenue en choisissant $M = P^{-1}$ montre aussi que Σ s'écrit $M \mapsto PM$.

Algorithme (calcul de l'inverse d'une matrice par la « méthode de Gauss-Jordan »)

On suppose que A est inversible. On va construire A^{-1} par des opérations sur les lignes de A.

On applique la méthode de Gauss sans échange de colonnes sur A dans $(A|I)$ en travaillant colonne par colonne de la gauche vers la droite pour faire apparaître des zéros en dehors de la diagonale. Une dernière étape permet de faire apparaître I à gauche de la barre verticale.

On utilise une construction par récurrence. On s'intéresse à la j^{e} ligne à l'étape $j \in \{1, ..., n\}$:



(i) On choisit dans la j^{e} colonne à gauche de la barre verticale un coefficient $d_j \neq 0$ (existe).

(ii) On échange la j^{e} ligne de la matrice augmentée avec la ligne contenant d_j.

(iii) On amène des 0 strictement au-dessus et au-dessous de d_j par des opérations sur les lignes de la matrice augmentée de la forme $L_i \leftarrow c_i L_i + c_j L_j$ avec $c_i \neq 0$ quand $i \neq j$.

En fin de calcul on divise chacune des lignes de la matrice augmentée par son coefficient d_j. La matrice A^{-1} est la matrice qui se trouve à droite de la barre verticale.

Exemple

$A := \begin{pmatrix} 5 & 4 \\ 6 & 5 \end{pmatrix}$ est-elle inversible ? Si oui, que vaut A^{-1} ?

On a : $(A|I)=\left(\begin{array}{cc|cc} 5 & 4 & 1 & 0 \\ 6 & 5 & 0 & 1 \end{array}\right) \xrightarrow[L_2 \leftarrow 5L_2 - 6L_1]{} \left(\begin{array}{cc|cc} 5 & 4 & 1 & 0 \\ 0 & 1 & -6 & 5 \end{array}\right) \xrightarrow[L_1 \leftarrow L_1 - 4L_2]{} \left(\begin{array}{cc|cc} 5 & 0 & 25 & -20 \\ 0 & 1 & -6 & 5 \end{array}\right) \xrightarrow[L_1 \leftarrow \frac{1}{5}L_1]{} \left(\begin{array}{cc|cc} 1 & 0 & 5 & -4 \\ 0 & 1 & -6 & 5 \end{array}\right).$

(On verra plus bas que le calcul de rang à mi-chemin prouve à cette étape que A est inversible.)

D'après la proposition précédente : A est inversible et $A^{-1} = \begin{pmatrix} 5 & -4 \\ -6 & 5 \end{pmatrix}$.

[On disposera aussi plus tard (en L2) d'une formule simple pour A^{-1} quand $A \in GL(2, \mathbb{K})$.]

4. Rang

La définition qui suit va utiliser la notion de rang d'une famille $(v_1, ..., v_p)$ avec $v_1, ..., v_p \in \mathbb{K}^n$, qui ne sera introduite que dans le chapitre suivant. En première lecture on admettra donc que le nombre r qui apparaitra dans l'algorithme qui va suivre la définition est indépendant de la matrice A', et considèrera que c'est pour l'instant ce nombre qui sera appelé *rang de A*.

Définition

Soit $A = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ & \cdots & \\ a_{n1} & \cdots & a_{np} \end{pmatrix} \in \mathfrak{M}_{n,p}(\mathbb{K})$.

On note : $\operatorname{rg} A := \operatorname{rg}\left(\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, ..., \begin{pmatrix} a_{1p} \\ \vdots \\ a_{np} \end{pmatrix}\right) \in \mathbb{N}$.

Algorithme

On transforme $\overset{x_1 \dots x_p}{A} := \begin{array}{c} \begin{array}{ccc} x_1 & \dots & x_p \end{array} \\ \begin{pmatrix} a_{11} & \dots & a_{1p} \\ & \cdots & \\ a_{n1} & \dots & a_{np} \end{pmatrix} \end{array}$ en $\overset{x_{j_1} \dots x_{j_p}}{A'} = \begin{array}{c} \begin{array}{cccccc} x_{j_1} & \dots & x_{j_r} & x_{j_{r+1}} & \dots & x_{j_p} \end{array} \\ \left(\begin{array}{cccc|ccc} d_1 & & & & & & \\ 0 & \ddots & & & & & \\ \vdots & \ddots & \ddots & & & & \\ 0 & \cdots & 0 & d_r & & & \\ \hline 0 & \cdots & & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & & 0 & 0 & \cdots & 0 \end{array}\right) \end{array}$

par la méthode de Gauss, avec $d_1 \neq 0$, ..., $d_r \neq 0$.

On a : $\boxed{\operatorname{rg} A = r}$ « calcul du rang par la méthode de Gauss ».

(Le nombre r est donc indépendant de la manière dont on applique la méthode de Gauss.)

Démonstration

On note v_1, ..., v_P les colonnes de A et montre que $(v_{j_1}, ..., v_{j_r})$ est une base de $\operatorname{Vect}(v_1, ..., v_p)$.

• On vérifie que la famille $(v_{j_1}, ..., v_{j_r})$ est libre. On montre que $(\star)$ $\alpha_1 v_{j_1} + ... + \alpha_r v_{j_r} = 0$ d'inconnue $(\alpha_1, ..., \alpha_r) \in \mathbb{K}^r$ a pour seule solution $(0, ..., 0)$.

On reprend le calcul permettant de passer de A à A' en ne prenant en compte que les colonnes qui sont sous les symboles $x_{j_1}, ..., x_{j_r}$. On constate à la fin du calcul que l'équation $(\star)$ a une unique solution : $(\alpha_1, ..., \alpha_r) = (0, ..., 0)$.

Variante

On a : $(\star) \iff \alpha_1 v_{j_1} + ... + \alpha_r v_{j_r} + 0 v_{j_{r+1}} + ... + 0 v_{j_p} = 0 \iff A' \begin{pmatrix} \alpha_1 \\ \cdots \\ \alpha_r \\ 0 \\ \cdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

En revenant à un système linéaire, on constate que l'équation $(\star)$ a pour seule solution $(0, ..., 0)$.

• On vérifie maintenant que $(v_{j_1}, ..., v_{j_r})$ engendre $\operatorname{Vect}(v_1, ..., v_p)$. Soit $k \in \{r+1, ..., p\}$. On montre que l'équation $(\star\star)$ $\alpha_1 v_{j_1} + ... + \alpha_r v_{j_r} = v_{j_k}$ d'inconnue $(\alpha_1, ..., \alpha_r) \in \mathbb{K}^r$ a au moins une solution. Le résultat en découlera car on aura ensuite $\underbrace{\operatorname{Vect}(v_1, ..., v_p)}_{\text{plus petit sous-espace vectoriel de } \mathbb{K}^n \text{ contenant } v_1, \, ..., \, v_p} \subseteq \operatorname{Vect}(v_{j_1}, ..., v_{j_r})$.

On reprend le calcul permettant de passer de A à A' en ne prenant en compte que les colonnes qui sont sous les symboles $x_{j_1}, ..., x_{j_r}$ et x_{j_k}, et place une barrer verticale juste avant la dernière colonne. On constate que $(\star\star)$ a une solution (des 0 à droite au bons endroits).

Variante

On a : $(\star\star) \iff \alpha_1 v_{j_1} + ... + \alpha_r v_{j_r} + 0 v_{j_{r+1}} + ... + (-1) v_{j_k} + ... + 0 v_{j_p} = 0 \iff A' \begin{array}{r} \\ \\ \\ \\ \text{ligne } k \rightarrow \\ \\ \\ \end{array}\!\begin{pmatrix} \alpha_1 \\ \cdots \\ \alpha_r \\ 0 \\ \cdots \\ -1 \\ \cdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

Une remontée triangulaire fournira une solution de $(\star\star)$. $\square$

Remarque

Lorqu'on adopte le point de vue de la méthode de Gauss sans permutation de colonnes, qui fournit une matrice échelonnée A' pas forcément réduite, le rang de A reste égal au nombre de pivots non nuls qui apparaissent dans la matrice A'.

Proposition

Soit $A \in \mathfrak{M}(n, \mathbb{K})$.
On a : A est inversible si et seulement si $\operatorname{rg} A = n$.

Démonstration

($\Rightarrow$) On suppose que A est inversible. Elle vérifie la condition (iv) de la dernière proposition, et on a donc $\operatorname{rg} A = n$ en calculant $\operatorname{rg} A$ par la méthode de Gauss.

($\Leftarrow$) On suppose que $\operatorname{rg} A = n$. On reproduit la démonstration de l'implication (ii) $\Rightarrow$ (iv) de la dernière proposition en remplaçant l'hypothèse (ii) par l'hypothèse « $\operatorname{rg} A = n$ ».
La matrice PA fournit $\operatorname{rg} A$. Elle a donc n blocs $\lfloor 1 \diagup \diagup$, puis (iv) est vérifié et A est inversible. □

Proposition

Soit $A \in \mathfrak{M}_{n,p}(\mathbb{K})$. On a : $\operatorname{rg}\left({}^tA\right) = \operatorname{rg}(A)$.

Démonstration (une variante sera donnée plus tard)

On calcule le rang r de A par la méthode de Gauss, ce qui donne A' avec r lignes non-nulles.

On pose : $A' = \begin{pmatrix} \overset{\overset{r}{\longleftrightarrow}}{U} & \overset{\overset{p-r}{\longleftrightarrow}}{V} \\ 0 & 0 \end{pmatrix} \updownarrow r$ avec U triangulaire supérieure de termes diagonaux non-nuls.

On passe de A à A' à l'aide d'opérations élémentaires sur les lignes de matrices associées P_1, ..., P_k dans $GL(n, \mathbb{K})$, et aussi à l'aide d'échanges $C_i \leftrightarrow C_j$ de colonnes pour lesquels on note R_1, ..., R_L les matrices associées aux $L_i \leftrightarrow L_j$ dans $GL(p, \mathbb{K})$. On a : $P_k ... P_1 A R_1 ... R_l = A'$.

D'où : $P_k ... P_1 A R_1 ... R_l = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U & V \\ 0 & I_{p-r} \end{pmatrix}$ avec $\begin{pmatrix} U & V \\ 0 & I_{p-r} \end{pmatrix} \in GL(p, \mathbb{K})$ vu son rang.

Cela fournit $P \in GL(n, \mathbb{K})$ et $Q \in GL(p, \mathbb{K})$ tels que : $A = P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q$.

On en déduit que : ${}^tA = {}^tQ \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} {}^tP$ avec ${}^tP \in GL(n, \mathbb{K})$ et ${}^tQ \in GL(p, \mathbb{K})$.

On peut donc passer de ${}^tA$ à $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} {}^tP$ par des opérations élémentaires sur les lignes.
La dernière matrice dont les $p - r$ dernières lignes sont nulles a un rang $\leq r$.

Ainsi $\operatorname{rg}\left({}^tA\right) \leq \operatorname{rg}(A)$. De même $\operatorname{rg} A = \operatorname{rg}\left({}^t({}^tA)\right) \leq \operatorname{rg}\left({}^tA\right)$. Cela donne le résultat. □

III. Déterminant d'ordre 2

Dans cette partie, on fixe $n \in \mathbb{N} \setminus \{0\}$.

Seul le cas $n = 2$ est au programme du semestre.

1. Caractérisation du déterminant

La décomposition d'une matrice carrée inversible en produit de matrices d'opérations élémentaire sur les lignes permet de fournir une construction concrète du déterminant.

Définition-Proposition

(a) Il existe une unique application $\det : \mathfrak{M}(n, \mathbb{K}) \to \mathbb{K}$, appelée *déterminant*, telle que :

(i) $\det A' = -\det A$ quand A' se déduit de A par $L_i \leftrightarrow L_j$ avec $i \neq j$;
(ii) $\det A' = c \det A$ quand A' se déduit de A par $L_i \leftarrow c\, L_i$ avec $c \in \mathbb{K}$ (par exemple 0) ;
(iii) $\det A' = \det A$ quand A' se déduit de A par $L_i \leftarrow L_i + cL_j$ avec $i \neq j$ et $c \in \mathbb{K}$;
(iv) $\det I_n = 1$.

(b) Cette application det, appelée *déterminant*, vérifie aussi :

(v) $\det A = \det A' + \det A''$ quand $A = \begin{pmatrix} L_1 \\ \vdots \\ L'_{i_0}+L''_{i_0} \\ \vdots \\ L_n \end{pmatrix}$ avec $A' = \begin{pmatrix} L_1 \\ \vdots \\ L'_{i_0} \\ \vdots \\ L_n \end{pmatrix}$ et $A'' = \begin{pmatrix} L_1 \\ \vdots \\ L''_{i_0} \\ \vdots \\ L_n \end{pmatrix}$;

(vi) $\det A = d_1 \cdots d_n$ quand $A = \begin{pmatrix} d_1 & & \\ 0 & \ddots & \\ \vdots & \ddots & \ddots \\ 0 & \cdots & 0 & d_n \end{pmatrix}$.

Démonstration (idées)

(a) • On montre d'abord l'unicité de « det ».

Gauss : $A \to \overbrace{\cdots}^{\substack{\text{opérations élémentaires} \\ \text{sur les lignes}}} \to \underset{\text{échelonnée}}{A'} = \left(\begin{array}{c|c|c|c|c} 0 \ldots 0 & 1 \\ \vdots \quad \vdots & 0 \ldots 0 & \begin{array}{c} 0 \\ 1 \end{array} & & \begin{array}{c} 0 \\ \vdots \\ 0 \\ 1 \end{array} \\ & \vdots \quad \vdots & 0 \ldots 0 & \ldots & 0 \ldots 0 \\ 0 \ldots 0 & 0 \ldots 0 & 0 \ldots 0 & & 0 \ldots 0 \end{array}\right)$.

– si la dernière ligne de A' est nulle :

$A' \mapsto A'' = A'$ par $L_n \leftarrow 2L_n$, donc $\det A' = 2 \det A'$ puis $\det A = \det A' = 0$;

– sinon : $A' = I$ donc $\det A' = 1$ puis $\det A$ est connu.

• On construit le déterminant des $A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}} \in \mathfrak{M}(n, \mathbb{K})$ par récurrence sur n :

– si $n = 1$: $\det(a_{11}) := a_{11}$;

– si $n \geq 2$: $\det A := a_{11}\Delta_{11} - a_{21}\Delta_{21} + \cdots + (-1)^{n-1} a_{n1}\Delta_{n1}$,

où Δ_{i1} est le déterminant de la matrice carrée d'ordre $n-1$ obtenue à partir de A en rayant la 1^re^ colonne et la i^e^ ligne.

(i) Récurrence sur n. Cas $n = 1$: rien à montrer.

L'opération élémentaire $L_i \leftrightarrow L_j$ avec $i < j$ se décompose en $L_i \leftrightarrow L_{i+1}$, puis ..., puis $L_{j-1} \leftrightarrow L_j$ (qui aboutit à $\begin{pmatrix} \vdots \\ L_{i-1} \\ L_{i+1} \\ \vdots \\ L_j \\ L_i \\ L_{j+1} \\ \vdots \end{pmatrix}$) suivi de $L_{j-1} \leftrightarrow L_{j-2}$, puis ..., puis $L_{i+1} \leftrightarrow L_i$.

On peut donc supposer que $j = i+1$. Dans la formule « $\det A = a_{11}\Delta_{11} - a_{21}\Delta_{21} + \cdots + (-1)^{n-1}a_{n1}\Delta_{n1}$ », l'échange de L_i avec L_{i+1} laisse invariant $(a_{1,1}, ..., a_{i-1,1}, a_{i+2,1}, ..., a_{n,1})$, remplace $(\Delta_{1,1}, ..., \Delta_{i-1,1}, \Delta_{i+2,1}, ..., \Delta_{n,1})$ par son opposé grâce à l'hypothèse de récurrence, et échange $(a_{i,1}, \Delta_{i,1})$ et $(a_{i+1,1}, \Delta_{i+1,1})$. D'où la formule (i).

(ii) Immédiat par récurrence.

(iii) On suppose que $A \mapsto A'$ par $L_i \leftarrow L_i + cL_j$ avec $i \neq j$ et $c \in \mathbb{K}$.

On va voir ci-dessous que la fonction déterminant qu'on vient de proposer réalise la première égalité du (b). En choisissant $L'_i = L_i$ et $L''_i = cL_j$, on en déduit d'après (ii) que :

$\det A' = \det \begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_n \end{pmatrix} + c \det \begin{pmatrix} L_1 \\ \vdots \\ \widetilde{L}_i \\ \vdots \\ L_n \end{pmatrix}$ où $A = \begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_n \end{pmatrix}$ et la matrice $\begin{pmatrix} L_1 \\ \vdots \\ \widetilde{L}_i \\ \vdots \\ L_n \end{pmatrix}$ avec $\widetilde{L}_i = L_j$ qui est invariante par $L_i \leftrightarrow L_j$ a d'après (i) un déterminant nul.

(iv) Immédiat par récurrence.

(b) L'égalité (v) découle par récurrence de la formule pour $\det A$ donnée dans la démonstration du (a), en distinguant pour $a_{i1}\Delta_{i1}$ les cas $i = i_0$ et $i \neq i_0$.

L'égalité (vi) découle par récurrence de cette même formule. □

Remarque

Soient $A \in \mathfrak{M}(n, \mathbb{K})$ et $\lambda \in \mathbb{K}$. En factorisant successivement chaque ligne de A par λ, on obtient grâce à l'égalité (ii) de la définition-proposition : $\det(\lambda A) = \lambda^n \det A$.

Notations

Soit $A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}} \in \mathfrak{M}(n, \mathbb{K})$. On note $v_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}$, ..., $v_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}$ les colonnes de A.

On pose : $\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} := \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ et $\det(v_1, ..., v_n) := \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$.

Corollaire

Soient $A, B \in \mathfrak{M}(n, \mathbb{K})$ et $v_1, ..., v_n \in \mathbb{K}^n$.

(a) $\boxed{\text{La matrice } A \text{ est inversible si et seulement si } \det A \neq 0}$.

(b) La famille $(v_1, ..., v_n)$ est libre si et seulement si $\det(v_1, ..., v_n) \neq 0$.

(c) On a : $\det(AB) = (\det A)(\det B)$.

(d) On a : $\det({}^tA) = \det A$.

DÉMONSTRATION

(a) On sait que A est inversible si et seulement si $\operatorname{rg} A = n$.
On reprend la démonstration de l'existence du déterminant.

Gauss : $$A \overbrace{\longrightarrow \dots \longrightarrow}^{\substack{\text{opérations élémentaires} \\ \text{sur les lignes}}} \underset{\text{échelonnée}}{A'} = \left(\begin{array}{c|c|c|c|c} 0 \dots 0 & 1 \ \ \ast & 0 \ \ \ast & & 0 \ \ \ast \\ \vdots \quad \vdots & 0 \dots 0 & 1 \ \ \ast & & \vdots \\ \vdots \quad \vdots & \vdots \quad \vdots & 0 \dots 0 & & 0 \\ \vdots \quad \vdots & \vdots \quad \vdots & \vdots \quad \vdots & \dots & 1 \ \ \ast \\ \vdots \quad \vdots & \vdots \quad \vdots & \vdots \quad \vdots & & 0 \dots 0 \\ 0 \dots 0 & 0 \dots 0 & 0 \dots 0 & & 0 \dots 0 \end{array}\right).$$

– si la dernière ligne de A' est nulle : $\det A = 0$ et $\operatorname{rg} A \neq n$ (calcul du rang par Gauss) ;
– sinon : $A' = I$ donc $\det A' = 1$ puis $\det A \neq 0$ et $\operatorname{rg} A = n$ (calcul du rang par Gauss).

(b) La famille $(v_1, ..., v_n)$ est libre si et seulement si elle est génératrice de $\mathbb{K}^n$, c'est-à-dire $\operatorname{rg}(v_1, ..., v_n) = n$, c'est-à-dire $\det(v_1, ..., v_n) \neq 0$ compte tenu de ce qui précède.

(c) • Si B n'est pas inversible : AB n'est pas inversible car sinon $\big((AB)^{-1}A\big)B = I_n$.
D'après (a) : $\det(AB) = 0 = (\det A) \times 0 = (\det A)(\det B)$.

• Si B est inversible : l'application $M \mapsto (\det B)^{-1} \det(MB)$ vérifie les conditions (i) à (iv) qui caractérisent le déterminant, car $MB = \begin{pmatrix} L_1 B \\ \vdots \\ L_n B \end{pmatrix}$ lorsque $M = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix}$.
Ainsi : $\det A = (\det B)^{-1} \det(AB)$.

(d) On peut supposer que A est inversible car sinon, vu les rangs, la formule s'écrit $0 = 0$.
Dans ce cas, il existe une suite finie d'opérations élémentaires sur les lignes, de la forme $M \mapsto P_k \cdots P_1 M$ pour une certaine matrice $P_1, ..., P_k$ inversibles associées à chacune de ces opérations élémentaires, qui envoie A sur I_n. En particulier $P_k \cdots P_1 A = I_n$.
On en déduit que : $A = P_1^{-1} \cdots P_k^{-1}$ et ${}^tA = {}^t(P_k^{-1}) \cdots {}^t(P_1^{-1})$.
Il reste à constater que la formule est vraie pour chacune des matrices P_1^{-1}, ..., P_k^{-1} d'opérations élémentaires sur les lignes. □

Remarques

On note $(e_1, ..., e_n)$ la base canonique de $\mathbb{K}^n$. Soient $v_1 = \sum_{i=1}^{n} a_{i,1} e_i, ..., v_n = \sum_{i=1}^{n} a_{i,n} e_i \in \mathbb{K}^n$.
On note $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ la matrice qui a pour colonnes v_1, ..., v_n.

1. Compte tenu de la définition-proposition du début du 1, on déduit du (d) que :

– $\boxed{\det(v_1, ..., v_n) = \alpha' \det(v_1, ..., v_{i_0-1}, v'_{i_0}, v_{i_0-1} v_n) + \alpha'' \det(v_1, ..., v_{i_0-1}, v''_{i_0}, v_{i_0-1} v_n)}$

quand $v_{i_0} = \alpha' v'_{i_0} + \alpha'' v''_{i_0}$ avec $i_0 \in \{1, ..., n\}$, $v'_{i_0}, v''_{i_0} \in \mathbb{K}^n$ et $\alpha', \alpha'' \in \mathbb{K}$;
– $\det(v_1, ..., v_n)$ change de signe lorsqu'on échange v_i et v_j avec $i \neq j$;
– $\det(v_1, ..., v_n)$ ne change pas lorsqu'on remplace v_i par $v_i + c v_j$ avec $i \neq j$ et $c \in \mathbb{K}$.

2. On note S_n l'ensemble des bijections σ de $\{1, ..., n\}$ sur $\{1, ..., n\}$.

On a : $\det A = \det(v_1, ..., v_n) = \det\Big(\sum_{i_1=1}^{n} a_{i_1,1} e_{i_1}, ..., \sum_{i_n=1}^{n} a_{i_n,n} e_{i_n} \Big)$

donc : $\det A = \sum_{\substack{1 \leq i_1, ..., i_n \leq n \\ \text{distincts}}} \det(e_{i_1}, ..., e_{i_n})\, a_{i_1,1} \cdots a_{i_n,n}$.

On pose $\sigma(1) = i_1$ et ... et $\sigma(n) = i_n$ et obtient : $\det A = \sum_{\sigma \in S_n} \underbrace{\det(e_{\sigma(1)}, ..., e_{\sigma(n)})}_{\text{noté } \varepsilon(\sigma)} a_{\sigma(1),1} ... a_{\sigma(n),n}$.

3. D'après le théorème de la base incomplète, le rang de $(v_1, ..., v_n)$ est le plus grand $r \in \mathbb{N}$ tel que $\mathbb{K}^n$ a une base de la forme $(e_{i_1}, ..., e_{i_{n-r}}, v_{j_1}, ..., v_{j_r})$. En développant le déterminant de

$(e_{i_1}, ..., e_{i_{n-r}}, v_{j_1}, ..., v_{j_r})$ suivant les premières colonnes, on constate que le rang de A est l'ordre maximal des déterminants extraits non-nuls de A par suppression de lignes (ici d'indices i_1, ..., i_{n-r}) et de colonnes (ici d'indices autres que j_1, ..., j_r).

2. Calcul et utilisation du déterminant

Proposition (hors programme)

Soit $A = (a_{ij})_{\substack{1\le i\le n\\1\le j\le n}} \in \mathfrak{M}(n, \mathbb{K})$.

À chaque $(i, j) \in \{1, ..., n\}^2$, on associe le signe $(-1)^{i+j}$ extrait de $\left(\begin{smallmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{smallmatrix}\right)$ et la matrice A_{ij} obtenue en supprimant dans A la i^{e} ligne et la j^{e} colonne.

(a) On a : $\det A = (-1)^{1+j} a_{1j} \det A_{1j} + \cdots + (-1)^{n+j} a_{nj} \det A_{nj}$ pour tout $j \in \{1, ..., n\}$ « développement suivant la j^{e} colonne ».

(b) On a : $\det A = (-1)^{i+1} a_{i1} \det A_{i1} + \cdots + (-1)^{i+n} a_{in} \det A_{in}$ pour tout $i \in \{1, ..., n\}$ « développement suivant la i^{e} ligne ».

Démonstration

(a) On note B la matrice déduite de A par $C_1 \leftrightarrow C_j$ puis $C_j \leftrightarrow C_{j-1}$, ..., $C_3 \leftrightarrow C_2$ (la colonne C_j est mise en première position et les autres sont dans l'ordre de leur indice).

On sait que $\det A = \det({}^tA)$. En appliquant les opérations élémentaires $L_1 \leftrightarrow L_j$ puis $L_j \leftrightarrow L_{j-1}$, ..., $L_3 \leftrightarrow L_2$ à ${}^tA$, on en déduit que : $\det A = (-1)^{j-1} \det B$ où $\det B$ se calcule comme dans la démonstration de la définition-proposition du début du paragraphe précédent.

(b) La formule « $\det A = \det({}^tA)$ » permet d'obtenir ce résultat en appliquant (a) à ${}^tA$. □

Exemples

Soient $a, b, c, d \in \mathbb{K}$.

• On a : $\boxed{\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc}$ en développant par rapport à C_1.

De plus : $\boxed{\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}$ quand $ad - bc \neq 0$ (vérification immédiate).

• On a : $\begin{vmatrix} 3 & 4 & 6 \\ 8 & 1 & -1 \\ -2 & 0 & 1 \end{vmatrix} \overset{\text{dév. / } L_3}{\overbrace{=}} +(-2) \overbrace{\begin{vmatrix} 4 & 6 \\ 1 & -1 \end{vmatrix}}^{-10} - 0 + 1 \overbrace{\begin{vmatrix} 3 & 4 \\ 8 & 1 \end{vmatrix}}^{-29} = -9.$

• $\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} \overset{\substack{L_2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow L_3 - L_1}}{\overbrace{=}} \begin{vmatrix} 1 & a & a^2 \\ 0 & b-a & (b-a)(b+a) \\ 0 & c-a & (c-a)(c+a) \end{vmatrix} \overset{\substack{\text{factoriser } L_2,\, L_3 \\ \text{puis dév. / } C_1}}{\overbrace{=}} (b-a)(c-a) \begin{vmatrix} 1 & b+a \\ 1 & c+a \end{vmatrix} = (a-b)(b-c)(c-a).$

Notation

Soit $A = (a_{ij})_{\substack{1\le i\le n\\1\le j\le n}} \in \mathfrak{M}(n, \mathbb{K})$. On lui associe sa *comatrice* :

$$\operatorname{Com} A := \left((-1)^{i+j} \det A_{ij}\right)_{\substack{1\le i\le n\\1\le j\le n}}$$

où A_{ij} est obtenue en supprimant dans A la i^{e} ligne et la j^{e} colonne.

Proposition

Soit $A \in \mathfrak{M}(n, \mathbb{K})$.

On a : ${}^t(\operatorname{Com} A)A = A\,{}^t(\operatorname{Com} A) = (\det A) I_n$.

En particulier, lorsque A est inversible on a : $A^{-1} = \frac{1}{\det A}\,{}^t(\operatorname{Com} A)$.

Démonstration (idée)

On note L_1, ..., L_n les lignes de A et C_1, ..., C_n les colonnes de A.

Le coefficient d'indice (i, j) de ${}^t(\operatorname{Com} A)A$ est, en développant suivant la i^{e} colonne, le déterminant $(-1)^{1+i} a_{1j} \det A_{1i} + \cdots + (-1)^{n+i} a_{nj} \det A_{ni}$ de la matrice obtenue en remplaçant C_i par C_j dans A. Il vaut $\det A$ si $i = j$ et 0 sinon.

Le coefficient d'indice (i, j) de $A\,{}^t(\operatorname{Com} A)$ est, en développant suivant la j^{e} ligne, le déterminant $(-1)^{j+1} a_{i1} \det A_{j1} + \cdots + (-1)^{j+n} a_{in} \det A_{jn}$ de la matrice obtenue en remplaçant L_j par L_i dans A. Il vaut $\det A$ si $i = j$ et 0 sinon. □

Exemple (déjà vu)

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathfrak{M}(2, \mathbb{K})$. On suppose A inversible.

On a : $\mathrm{Com}\, A = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$, donc $A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Proposition

Soient $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathfrak{M}(n, \mathbb{K})$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{K}^n$.

Le système $(E) \begin{cases} a_{11}x_1 + \ldots + a_{1p}x_p = b_1 \\ \ldots \\ a_{n1}x_1 + \ldots + a_{np}x_p = b_n \end{cases}$ a une unique solution si et seulement si $\det A \neq 0$.

Dans ce cas, cette solution $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ est donnée par :

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ & & \cdots & \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}} \quad \text{et} \ \ldots \ \text{et} \quad x_n = \frac{\begin{vmatrix} a_{11} & \cdots & a_{1,n-1} & b_1 \\ & \cdots & & \\ a_{n1} & \cdots & a_{n,n-1} & b_n \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}}$$ « formules de Cramer ».

Démonstration

Le début vient de l'équivalence entre l'unicité d'une solution et $\mathrm{rg}\, A = n$.

On note C_1, ..., C_n les colonnes de A. Le scalaire $\det(C_1, \ldots, C_{i-1}, B, C_{i+1}, \ldots, C_n)$ calculé par développement suivant la i^{e} colonne, est la i^{e} coordonnée de ${}^t(\mathrm{Com}\, A)B$.
La i^{e} coordonnée de $A^{-1}B$ s'écrit donc : $x_i = \frac{1}{\det A} {}^t(\mathrm{Com}\, A)B = \frac{\det(C_1, \ldots, C_{i-1}, B, C_{i+1}, \ldots, C_n)}{\det A}$. □

Ch. 3. Espaces vectoriels sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$

Plan
I. Règles de calcul
II. Constructions
III. Bases et dimension

I. Règles de calcul

1. Généralités

Définition-Proposition

(a) On appelle *espace vectoriel sur* $\mathbb{K}$ (ou $\mathbb{K}$-*ev*) un ensemble E muni de deux applications

$$\begin{array}{ccc} E\times E & \longrightarrow & E \\ (v,w) & \longmapsto & v+w \end{array} \quad \text{et} \quad \begin{array}{ccc} \mathbb{K}\times E & \longrightarrow & E \\ (\alpha,v) & \longmapsto & \underbrace{\alpha\times v}_{\text{noté par la suite } \alpha\, v} \end{array} \quad \text{telles que :}$$

(i) $(u+v)+w = u+(v+w)$ et $u+v = v+u$ pour tous $u,v,w\in E$;
(ii) il existe un élément $\underbrace{0_E}_{\text{noté par la suite } 0}$ de E tel que $v+0_E = v$ pour tout $v\in E$
et dans ce cas, compte tenu de (i) cet élément 0_E est unique ;
(iii) pour tout $v\in E$, il existe un élément $-v$ de E tel que $v+(-v) = 0_E$
et dans ce cas, compte tenu de (i) et (ii) cet élément $-v$ est unique ;
(iv) $\alpha\,(v+w) = (\alpha\, v)+(\alpha\, w)$ et $(\alpha+\beta)\,v = (\alpha\, v)+(\beta\, v)$
pour tous $\alpha,\beta\in\mathbb{K}$ et $v,w\in E$;
(v) $1\,v = v$ et $\alpha\,(\beta\, v) = (\alpha\beta)\,v$ pour tous $\alpha,\beta\in\mathbb{K}$ et $v\in E$.

(b) Étant donné un $\mathbb{K}$-espace vectoriel $(E,+,\times)$ comme au (a), on appellera *scalaires* les éléments de $\mathbb{K}$ et *vecteurs* les éléments de E.
On notera : $v-w = v+(-w)$ quand $v,w\in E$.

Démonstration

Il s'agit de prouver les deux cas d'unicité signalés au (a).
(ii) Pour tout candidat $\tilde{0}$ au rôle de 0_E, on a d'après (i) :

$$\tilde{0} = \tilde{0}+0_E = 0_E+\tilde{0} = 0_E.$$

(iii) Pour tout candidat $\tilde{v}$ au rôle de $-v$, on a aussi d'après (i) et (ii) :
$\tilde{v} = \tilde{v}+0_E = \tilde{v}+(v+(-v)) = (\tilde{v}+v)+(-v) = (v+\tilde{v})+(-v) = 0_E+(-v) = -v.$ □

Remarques

On se donne un $\mathbb{K}$-espace vectoriel E.

(1) Pour $\alpha\in\mathbb{K}$ et $v\in E$, on a :

$0\,v = 0\,v+(1\,v-v) = 0_E$;
$\alpha\,0_E = \alpha(0_E+0_E)-\alpha\,0_E = 0_E$;
$(-\alpha)v = (-\alpha+\alpha)v-(\alpha\,v) = -(\alpha\,v)$.

(2) Règles de simplification. Pour $\alpha,\beta\in\mathbb{K}$ et $u,v,w\in E$, on a :

$u+v = u+w \implies v = w$; (ajouter $-u$ à chaque membre)
$\alpha\,v = 0_E \implies \alpha = 0$ ou $v = 0_E$; (si $\alpha\neq 0$, multiplier chaque membre par $\frac{1}{\alpha}$)
$\alpha\,v = \alpha\,w$ et $\alpha\neq 0 \implies v = w$; (tout faire passer au premier membre)
$\alpha\,v = \beta\,v$ et $v\neq 0_E \implies \alpha = \beta$. (tout faire passer au premier membre)

2. Exemples

Exemple 1

Soit $n \in \mathbb{N}$. L'ensemble $\mathbb{K}^n$ des n-uplets $(x_1, ..., x_n)$ d'éléments $x_1, ..., x_n$ de $\mathbb{K}$, muni de

$$(x_1, ..., x_n) + (y_1, ..., y_n) := (x_1 + y_1, ..., x_n + y_n) \quad \text{et} \quad \alpha{\times}(x_1, ..., x_n) := (\alpha x_1, ..., \alpha x_n)$$

est un $\mathbb{K}$-espace vectoriel (clair). Ses éléments sont les applications $i \mapsto x_i$ de $\{1, ..., n\}$ dans $\mathbb{K}$.

Par convention, on écrira : $\mathbb{K}^0 := \{0\} \subseteq \mathbb{K}$ en « identifiant » la suite vide () et 0.

Cas particulier

Le « plan » $\mathbb{R}^2$ est un $\mathbb{R}$-espace vectoriel et la « droite » $\mathbb{C}^1$ est un $\mathbb{C}$-espace vectoriel.

Ainsi $\mathbb{C}$ possède à la fois une structure de $\mathbb{C}$-espace vectoriel et de $\mathbb{R}$-espace vectoriel (d'ensemble $\mathbb{R}^2$ en définissant $x + iy$ comme le couple (x, y) quand $x, y \in \mathbb{R}$).

Exemple 2

Soient A un ensemble et E un $\mathbb{K}$-espace vectoriel.

L'ensemble, noté E^A, des applications de A dans E, muni des « lois » déterminées par

$$\underset{E^A}{\underset{\Cap}{f}} + \underset{E^A}{\underset{\Cap}{g}} : \begin{array}{l} A \to E \\ x \mapsto f(x) + g(x) \end{array} \quad \text{et} \quad \underset{\mathbb{K}}{\underset{\Cap}{\alpha}} \times \underset{E^A}{\underset{\Cap}{f}} : \begin{array}{l} A \to E \\ x \mapsto \alpha\,(f(x)) \end{array}$$

est un $\mathbb{K}$-espace vectoriel, avec $0_{E^A} : \begin{array}{l} A \to E. \\ x \mapsto 0 \end{array}$ ⟵[exercice]

Exemple 3

En particulier, l'ensemble $\mathbb{K}^{\mathbb{N}}$ des suites d'éléments de $\mathbb{K}$, est un $\mathbb{K}$-espace vectoriel.

Suivant l'usage, on notera $(u_n)_{n\in\mathbb{N}}$ au lieu de $u\colon \begin{array}{l} \mathbb{N} \to \mathbb{K} \\ n \mapsto u_n \end{array}$.

Ainsi, on a par définition : $(u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}} = (u_n + v_n)_{n\in\mathbb{N}}$ et $\alpha{\times}(u_n)_{n\in\mathbb{N}} = (\alpha\, u_n)_{n\in\mathbb{N}}$.

Exemple 4

Soient $n, p \in \mathbb{N}$. D'après l'exemple 2, l'ensemble $\mathfrak{M}_{n,p}(\mathbb{K})$ des matrices $n \times p$ à coefficients dans $\mathbb{K}$, qui est égal à $\mathbb{K}^{\{1,...,n\}\times\{1,...,p\}}$, est aussi un $\mathbb{K}$-espace vectoriel, avec la somme et le produit d'un scalaire par une matrice qui ont été définis dans le chapitre 2.

3. Combinaisons linéaires

Définition-Proposition

Soient E un $\mathbb{K}$-espace vectoriel, $p \in \mathbb{N}$ et $v_1, \ldots, v_p, v, w \in E$.

(a) Une *combinaison linéaire* (ou *cl*) des vecteurs $v_1, \ldots, v_p$ est un vecteur u de E de la forme $u = \underbrace{\alpha_1\, v_1 + \cdots + \alpha_n\, v_p}_{0_E \text{ quand } p=0}$ avec $\alpha_1, \ldots, \alpha_p \in \mathbb{K}$.

(b) On dit que v et w sont *colinéaire* (« sur une même droite ») s'il existe $\alpha \in \mathbb{K}$ tel que $v = \alpha\, w$ ou $w = \alpha\, v$. Cela équivaut à : $v = 0_E$ ou il existe $\alpha \in \mathbb{K}$ tel que $w = \alpha\, v$.

Démonstration

Il s'agit de prouver l'équivalence du (b).

($\Rightarrow$) Dans le cas où $v = \alpha\, w$, on a : soit $\alpha = 0$ et par suite $v = 0_E$, soit $\alpha \neq 0$ et $w = \frac{1}{\alpha}\, v$.

($\Leftarrow$) Lorsque $v = 0_E$, on a : $v = 0\, w$. □

Exemple

Dans $\mathbb{R}^3$, on choisit $v_1 = (0, 1, -1)$, $v_2 = (1, 0, -1)$, $v_3 = (1, -1, 0)$, et $v = (5, -2, -3)$.

Le vecteur v est combinaison linéaire de v_1, v_2, v_3 car : $v = v_1 + 2v_2 + 3v_3$.

Remarque

La notation « $\alpha_1\, v_1 + \cdots + \alpha_n\, v_p$ » avec des points de suspension est usuelle mais abusive. On pourrait la remplacer par « $\sum\limits_{1\leq k\leq p} \alpha_k\, v_k$ » avec la définition par récurrence suivante : $\sum\limits_{1\leq k\leq 0} u_k := 0_E$, et, $\sum\limits_{1\leq k\leq p+1} u_k = \Big(\sum\limits_{1\leq k\leq p} u_k\Big) + u_{p+1}$ pour tous $p \in \mathbb{N}$ et $(u_i)_{1\leq i\leq p+1} \in E^{p+1}$.

II. Constructions

Dans toute cette partie, on fixe un $\mathbb{K}$-espace vectoriel E.

1. Sous-espaces vectoriels

Définition

On dit qu'une partie F de E est un *sous-espace vectoriel de* E (ou *sous-ev de* E) si :

(i) $0_E \in F$;

(ii) pour tous $\alpha, \beta \in \mathbb{K}$ et $v, w \in F$, on a $\alpha\, v + \beta\, w \in F$.

Remarques

(1) Il est immédiat que : $\{0_E\}$ et E sont des sous-espaces vectoriels de E.

(2) Soient F un sous-espace vectoriel de E, $p \in \mathbb{N}$ et $v_1, \ldots, v_p \in F$.
Par récurrence sur p, on constate que toute combinaison linéaire de $v_1, \ldots, v_p$ appartient à F.

Proposition

Tout sous-espace vectoriel de E est un $\mathbb{K}$-espace vectoriel pour l'addition et la multiplication par un scalaire dans E. (*)

Démonstration

Soit F un sous-espace vectoriel de E. Par restriction de l'addition et la multiplication par un scalaire dans E, on définit les applications $\begin{array}{ccc} F \times F & \longrightarrow & F \\ (v,w) & \longmapsto & 1v + 1w \end{array}$ et $\begin{array}{ccc} \mathbb{K} \times F & \longrightarrow & F \\ (\alpha, v) & \longmapsto & \alpha v + 0 0_E \end{array}$.

Comme E est un $\mathbb{K}$-espace vectoriel, F satisfait clairement les conditions pour être un $\mathbb{K}$-espace vectoriel, en prenant $0_F := 0_E$ et l'opposé de $v \in F$ égal à $\underbrace{-v}_{\text{notation relative à } E}$. $\square$

Exemples

1. On considère $E = \{(x,y,z) \in \mathbb{R}^3 \mid x + y + z = 0\}$.

(i) On a : $0 + 0 + 0 = 0$ donc $0 \in E$.

(ii) Soient $\alpha, \beta \in \mathbb{R}$ et $v = (x', y', z'), w = (x'', y'', z'') \in E$.
On a : $\alpha\, v + \beta\, w = (x,y,z)$ avec $x := \alpha x' + \beta x''$, $y := \alpha y' + \beta y''$, $z := \alpha z' + \beta z''$.
Or : $x + y + z = \alpha(\underbrace{x' + y' + z'}_{0 \text{ car } v \in E}) + \beta(\underbrace{x'' + y'' + z''}_{0 \text{ car } w \in E})$. D'où : $\alpha\, v + \beta\, w \in E$.

En conclusion : E est un sous-espace vectoriel de $\mathbb{R}^3$.

2. Soit $A \in \mathfrak{M}_{n,p}(\mathbb{K})$. L'ensemble $\mathscr{S}_H$ des solutions du système d'équations linéaires homogène $(H) : AX = 0$ d'inconnue $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathbb{K}^p$ est un sous-ev de $\mathbb{K}^p$. (**) , car :
(i) on a : $A\, 0_{\mathbb{K}^p} = 0_{\mathbb{K}^n}$ donc $0_{\mathbb{K}^p} \in \mathscr{S}_H$;
(ii) pour $\alpha, \beta \in \mathbb{K}$ et $X, Y \in \mathscr{S}_H$, on a $A(\alpha X + \beta Y) = \alpha \underbrace{AX}_{0} + \beta \underbrace{AY}_{0} = 0$, donc $\alpha X + \beta Y \in \mathscr{S}_H$.

3. Par définition, un polynôme à coefficients dans $\mathbb{K}$ est une suite $P = (a_n)_{n \geq 0}$ d'éléments de $\mathbb{K}$ pour laquelle il existe $p \in \mathbb{N}$ tel que : $a_n = 0$ pour tout $n > p$. $\longleftarrow [P = a_p X^p + \ldots + a_1 X + a_0]$

L'ensemble $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$ est un sous-espace vectoriel de $\mathbb{K}^{\mathbb{N}}$ car :
(i) $0_{\mathbb{K}^{\mathbb{N}}} \in \mathbb{K}[X]$ en prenant par exemple $p = 0$;
(ii) pour tous $\alpha, \beta \in \mathbb{K}$ et $P = (a_n)_{n \geq 0}$, $Q = (b_n)_{n \geq 0} \in \mathbb{K}[X]$, en fixant $p, q \in \mathbb{N}$ tels que $a_n = 0$ quand $n > p$ et $b_n = 0$ quand $n > q$ (il en existe), on constate que $\alpha a_n + \beta b_n = 0$ quand $n > \max(p,q)$, ce qui montre que $\alpha P + \beta Q \in \mathbb{K}[X]$.

(*) Cette proposition fournira de nombreux espaces vectoriels formés de fonctions réelles de la variable réelle (applications de $\mathbb{R}$ dans $\mathbb{R}$) ou formés de suites de nombres réels (applications de $\mathbb{N}$ dans $\mathbb{R}$).
(**) On dira qu'il s'agit d'un sous-espace vectoriel de $\mathbb{K}^p$ donné par « équation cartésienne ».

Définition-Proposition

(a) Soit P une partie de E. On appelle *sous-espace vectoriel de E engendré par P*, et note $\operatorname{Vect} P$ (ou $\operatorname{Vect}_{\mathbb{K}} P$), l'ensemble des combinaisons linéaires d'un nombre fini d'éléments de P.

On a : $\operatorname{Vect} P$ est le plus petit sous-espace vectoriel de E contenant P (unique sous-espace vectoriel de E contenant P qui est inclus dans tout sous-espace vectoriel F de E contenant P).

(b) En particulier, quand $p \in \mathbb{N}$ et $v_1, ..., v_p \in E$, on pose :
$\operatorname{Vect}(v_1, ..., v_p) := \operatorname{Vect}\big(\{v_1, ..., v_p\}\big) \overset{\text{cf. (a)}}{=} \big\{\alpha_1\, v_1 + \cdots + \alpha_p\, v_p\ ;\ \alpha_1, \ldots, \alpha_p \in \mathbb{K}\big\}$.(* * *)

Démonstration

Il s'agit de démontrer la caractérisation de $\operatorname{Vect} P$ donnée à la fin du (a).

Tout d'abord, on vérifie que $\operatorname{Vect} P$ est un sous-espace vectoriel de E.

(i) On a $0_E \in \operatorname{Vect} P$ (seule combinaison linéaire avec 0 vecteur).

(ii) Soient $\alpha, \alpha' \in \mathbb{K}$ et $v = \underset{\mathbb{K}}{\alpha_1}\, \underset{P}{v_1} + \cdots + \underset{\mathbb{K}}{\alpha_n}\, \underset{P}{v_n}$, $v' = \underset{\mathbb{K}}{\alpha'_1}\, \underset{P}{v'_1} + \cdots + \underset{\mathbb{K}}{\alpha'_{n'}}\, \underset{P}{v'_{n'}} \in \operatorname{Vect} P$.

On a : $\alpha\, v + \alpha'\, v' = (\alpha\alpha_1)v_1 + \cdots + (\alpha\alpha_n)v_n + (\alpha'\alpha'_1)v'_1 + \cdots + (\alpha'\alpha'_{n'})v'_{n'}$ donc $\alpha\, v + \alpha'\, v' \in \operatorname{Vect} P$.

Cela montre que $\operatorname{Vect} P$ est un sous-espace vectoriel de E.

On a aussi : $\operatorname{Vect} P$ contient P car tout $v \in P$ s'écrit $1\, v$ où évidemment $1 \in \mathbb{K}$ et $v \in P$.

Pour terminer, on constate que la remarque 2 ci-dessus montre que $\operatorname{Vect} P$ est inclus dans tout sous-espace vectoriel F de E qui contient P. □

Définition

On appelle :

– *droite vectorielle de E* un sous-ev de E de la forme $\operatorname{Vect}(v)$ avec $v \in E \setminus \{0\}$;

– *plan vectoriel de E* un sous-ev de E de la forme $\operatorname{Vect}(v, w)$ avec $v, w \in E$ non-colinéaires.

Exemple

Soient $v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ 0 \\ \lambda \end{pmatrix} \in \mathbb{R}^3$ (identifiés à des triplets) avec $\lambda \in \mathbb{R}$ fixé.

Il est « clair » que v_1 et v_2 sont non-colinéaires (car $v_2 \neq 0$ et – au vu des 1res coordonnées – une égalité $v_1 = \alpha v_2$ impliquerait $1 = 0$).

(a) *A-t-on : $v_3 \in \operatorname{Vect}(v_1, v_2)$?*

On étudie par la méthode de Gauss l'existence de $\alpha_1, \alpha_2 \in \mathbb{R}$ tels que $\overbrace{\underbrace{\alpha_1 v_1 + \alpha_2 v_2}_{\text{inconnues } \alpha_1, \alpha_2} = v_3}^{\text{expression « paramétrique » des éléments de } \operatorname{Vect}(v_1, v_2)}$:

$$\left(\begin{array}{cc|c} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & \lambda \end{array}\right) \xrightarrow[L_2 \leftarrow L_2 + L_1]{} \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & \lambda \end{array}\right) \xrightarrow[L_3 \leftarrow L_3 + L_2]{} \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & \lambda+1 \end{array}\right).$$

Donc : $v_3 \in \operatorname{Vect}(v_1, v_2) \iff \lambda = -1$.

(b) On suppose que $\lambda \neq -1$. *A-t-on : $\operatorname{Vect}(v_1, v_2, v_3) = \mathbb{R}^3$?*

Soit $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$. On étudie par la méthode de Gauss l'existence de $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ tels que $\underbrace{\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = v}_{\text{inconnues } \alpha_1, \alpha_2, \alpha_3}$:

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & x \\ -1 & 1 & 0 & y \\ 0 & -1 & \lambda & z \end{array}\right) \xrightarrow[\cdots]{} \text{(étapes ci-dessus)} \xrightarrow[\cdots]{} \left(\begin{array}{ccc|c} 1 & 0 & 1 & /// \\ 0 & 1 & 1 & /// \\ 0 & 0 & \lambda+1 & /// \end{array}\right).$$

Donc : $v \in \operatorname{Vect}(v_1, v_2, v_3)$.

On peut en conclure – l'inclusion $\subseteq$ étant claire – que : $\operatorname{Vect}(v_1, v_2, v_3) = \mathbb{R}^3$.

2. Opérations sur les sous-espaces vectoriels

(* * *) On dira qu'il s'agit d'un sous-espace vectoriel de E donné par « équation paramétrique ».
On note ici $\{f(t)\,;\, t \in T\}$ l'ensemble image $\{v \in E \mid \exists\, t \in T \quad v = f(t)\}$ d'une application $f : \underbrace{T}_{\text{ensemble des « paramètres »}} \to E$.

Proposition

Soient F et G deux sous-espaces vectoriels de E.
On a : $F \cap G$ est un sous-espace vectoriel de E.

Démonstration

(i) On a : $0_E \in F$ et $0_E \in G$ donc $0_E \in F \cap G$.
(ii) Soient $\alpha, \beta \in \mathbb{K}$ et $v, w \in F \cap G$.
On a : $\alpha \underset{F}{\underset{\Cap}{v}} + \beta \underset{F}{\underset{\Cap}{w}} \in F$ et $\alpha \underset{G}{\underset{\Cap}{v}} + \beta \underset{G}{\underset{\Cap}{w}} \in G$, donc $\alpha v + \beta w \in F \cap G$.
Il en résulte que $F \cap G$ est un sous-espace vectoriel de E. □

Remarque

On prend ici $F = (Ox)$ et $G = (Oy)$ dans $\mathbb{R}^2$. Donc $F = \mathrm{Vect}((1,0))$ et $G = \mathrm{Vect}((0,1))$ sont des sous-espaces vectoriels de $\mathbb{R}^2$.

On a $F \cap G = \{(0,0)\}$, où $\{(0,0)\}$ est un sous–espace vectoriel bien connu de $\mathbb{R}^2$.

Par contre $F \cup G$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$ car : $1\underbrace{(1,0)}_{\in F} + 1\underbrace{(0,1)}_{\in G} = \underbrace{(1,1)}_{\notin F \cup G}$.

Exemple

Idée : des équations cartésiennes de F et G fournissent une équation cartésienne de $F \cap G$.

On considère $\Pi_1 : x + y + z = 0$ et $\Pi_2 : x + 2y = 0$ dans $\mathbb{R}^3$. Comme ensembles de solutions de systèmes d'équations homogènes, Π_1 et Π_2 sont des sous-espaces vectoriels de $\mathbb{R}^3$ (on verra dans le prochain exemple que ce sont des plans vectoriels).
On cherche à préciser la nature du sous-espace vectoriel $\Pi_1 \cap \Pi_2$ de $\mathbb{R}^3$.

On regroupe les équations cartésiennes de Π_1 et de Π_2 : $\Pi_1 \cap \Pi_2 : \begin{cases} x + y + z = 0 \\ \quad x + 2y = 0 \end{cases}$.
On va exhiber une équation paramétrique de $\Pi_1 \cap \Pi_2$ en résolvant par la méthode de Gauss :
$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \end{pmatrix} \xrightarrow[L_2 \leftarrow L_2 - L_1]{} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$. On effectue ensuite de tête la « remontée triangulaire ».

Ainsi : $\Pi_1 \cap \Pi_2 : \begin{cases} x = -2t \\ y = \quad t \\ z = \quad t \end{cases}$, $t \in \mathbb{R}$ puis $\Pi_1 \cap \Pi_2 = \mathrm{Vect}\left(\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}\right)$ est une droite vectorielle.

Définition-Proposition

Soient F, G et $F_1, \dots, F_k$ des sous-espaces vectoriels de E.

(a) On note : $F + G := \{v + w \ ; \ v \in F \text{ et } w \in G\}$.
On a : $F + G$ est un sous-espace vectoriel de E, appelé *somme de F et de G*.

On note plus généralement : $F_1 + \dots + F_k := \{v_1 + \dots + v_k \ ; \ v_1 \in F_1, \dots v_k \in F_k\}$.
On a : $F_1 + \dots + F_k$ est un sous-espace vectoriel de E, appelé *somme de $F_1, \dots, F_k$*.

(b) On a : $F + G = \mathrm{Vect}(P \cup Q)$ quand $F = \mathrm{Vect}\, P$ et $G = \mathrm{Vect}\, Q$ avec $P, Q \subseteq E$.
(On en déduit que $F + G = \mathrm{Vect}(F \cup G)$.)

On a même : $F_1 + \dots + F_k = \mathrm{Vect}(P_1 \cup \dots \cup P_k)$ quand $F_i = \mathrm{Vect}\, \underbrace{P_i}_{\subseteq E}$ pour $1 \leq i \leq k$.

Démonstration

Afin de ne pas alourdir la rédaction, on se limite ici à la démonstration du cas de $F + G$.

(a) (i) On a : $0_E = 0_E + 0_E \in F + G$.
(ii) Soient $\alpha, \alpha' \in \mathbb{K}$ et $u = \underset{F}{\underset{\Cap}{v}} + \underset{G}{\underset{\Cap}{w}}$, $u' = \underset{F}{\underset{\Cap}{v'}} + \underset{G}{\underset{\Cap}{w'}} \in F + G$.
On a : $\alpha u + \alpha' u' = (\underbrace{\alpha v + \alpha' v'}_{\in F}) + (\underbrace{\alpha w + \alpha' w'}_{\in G})$, donc $\alpha u + \alpha' u' \in F + G$.

Il en résulte que $F + G$ est un sous-espace vectoriel de E.

(b) On a : $P \cup Q \subseteq F + G$ donc $\underbrace{\mathrm{Vect}(P \cup Q)}_{\text{plus petit sous-ev de } E \text{ contenant } P \cup Q} \subseteq F + G$.

Soit $u = \underset{F}{\underset{\Cap}{v}} + \underset{G}{\underset{\Cap}{w}} \in F + G$. Comme $v \in \mathrm{Vect}\, P$ et $w \in \mathrm{Vect}\, Q$, on a $v, w \in \mathrm{Vect}(P \cup Q)$ puis $u = 1v + 1w \in \mathrm{Vect}(P \cup Q)$. Ainsi : $F + G \subseteq \mathrm{Vect}(P \cup Q)$.

En conclusion : $F + G = \mathrm{Vect}(P \cup Q)$. □

Exemple

Idée : des équations paramétriques de F et G donnent une équation paramétrique de $F+G$.

On reprend $\Pi_1 : x+y+z=0$ et $\Pi_2 : x+2y=0$ dans $\mathbb{R}^3$. On cherche à préciser $\Pi_1+\Pi_2$.

• Méthode de Gauss pour aboutir à $\Pi_1 = \mathrm{Vect}(v_1, ..., v_p)$ et $\Pi_2 = \mathrm{Vect}(w_1, ..., w_q)$:

$\begin{array}{ccc} x & y & z \\ \end{array}$ $(\boxed{1}\ \ 1\ \ 1)$ donne $\Pi_1 : \begin{cases} x = -s-t \\ y = \quad s \\ z = \quad t \end{cases}$ $s, t \in \mathbb{R}$, et $\begin{array}{ccc} x & y & z \\ \end{array}$ $(\boxed{1}\ \ 2\ \ 0)$ donne $\Pi_2 : \begin{cases} x = -2u \\ y = \quad u \\ z = \quad v \end{cases}$ $u, v \in \mathbb{R}$.

En particulier $\Pi_1 = \mathrm{Vect}\left(\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right)$ et $\Pi_2 = \mathrm{Vect}\left(\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$ sont des plans vectoriels.

On a ensuite : $\Pi_1+\Pi_2 : \begin{cases} x = -s-t-2u \\ y = \quad s+u \\ z = \quad t+v \end{cases}$ $s, t, u, v \in \mathbb{R}$. On va voir que : $\Pi_1+\Pi_2 = \mathbb{R}^3$.

• Soit $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$. On étudie l'existence de $s, t, u, v \in \mathbb{R}$ permettant de s'assurer que v vérifie l'équation paramétrique précédente de $\Pi_1+\Pi_2$, par la méthode de Gauss :

$$\left(\begin{array}{cccc|c} -1 & -1 & -2 & 0 & x \\ 1 & 0 & 1 & 0 & y \\ 0 & 1 & 0 & 1 & z \end{array}\right) \xrightarrow[L_2 \leftarrow L_2+L_1]{} \left(\begin{array}{cccc|c} -1 & -1 & -2 & 0 & x \\ 0 & -1 & -1 & 0 & x+y \\ 0 & 1 & 0 & 1 & z \end{array}\right) \xrightarrow[L_3 \leftarrow L_3+L_2]{} \left(\begin{array}{cccc|c} -1 & -1 & -2 & 0 & /// \\ 0 & -1 & -1 & 0 & /// \\ 0 & 0 & -1 & 1 & /// \end{array}\right).$$

Donc : $v \in \Pi_1+\Pi_2$.

En conclusion : $\Pi_1+\Pi_2 = \mathbb{R}^3$. ⟵[variante, plus tard : extraction de base d'une famille génératrice]

Proposition

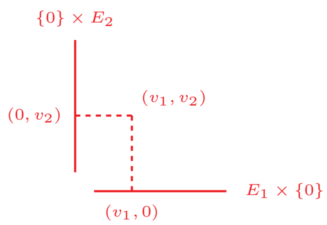
Soient E_1 et E_2 deux $\mathbb{K}$-espaces vectoriels.

L'ensemble $E_1 \times E_2$ muni des « lois » déterminées par $(v_1, v_2)+(w_1, w_2) := (v_1+w_1, v_2+w_2)$ et $\alpha \times (v_1, v_2) := (\alpha v_1, \alpha v_2)$ est un $\mathbb{K}$-espace vectoriel, appelé *produit de E_1 et E_2*.

Démonstration

On choisit $0_{E_1 \times E_2} := (0_{E_1}, 0_{E_2})$ et $-(v_1, v_2) := (-v_1, -v_2)$ quand $(v_1, v_2) \in E_1 \times E_2$. Le reste de cette démonstration, qui est facile, est omis. □

Remarque



Tout $(v_1, v_2) \in E_1 \times E_2$ se décompose de manière unique comme somme d'un élément de $\underbrace{E_1 \times \{0\}}_{\text{s'identifie à } E_1}$ et de $\underbrace{\{0\} \times E_2}_{\text{s'identifie à } E_2}$.

Problème : quelles décompositions de E jouent le même rôle ?

Proposition

Soient F et G deux sous-espaces vectoriels de E. Les propriétés suivantes sont équivalentes :
(i) pour tout $u \in E$, il existe $v \in F$ et $w \in G$ uniques$^{(***)}$ tels que $u = v+w$;
(ii) $E = F+G$ et $F \cap G = \{0\}$.

Démonstration

($\Rightarrow$) On suppose (i) vrai. En particulier, on a : $E = F+G$.
Soit $u \in F \cap G$. On a : $\underset{\in F}{u} + \underset{\in G}{0_E} = \underset{\in F}{0_E} + \underset{\in G}{u}$ donc, par unicité, $u = 0_E$. Ainsi : $F \cap G = \{0\}$.

($\Leftarrow$) On suppose (ii) vrai et se donne $u \in E$. En particulier, il existe au moins une décomposition $u = v+w$ avec $v \in F$ et $w \in G$.
Si $u = \underset{\in F}{v} + \underset{\in G}{w} = \underset{\in F}{v'} + \underset{\in G}{w'}$ alors $\underbrace{v-v'}_{\in F} = \underbrace{w'-w}_{\in G}$ puis $v-v', w'-w \in F \cap G$ ce qui donne $v = v'$ et $w = w'$. On en déduit l'unicité de la décomposition de u. □

$(***)$ Il faut interpréter l'expression « il existe $v \in F$ et $w \in G$ uniques » par « il existe $(v, w) \in F \times G$ unique ».

Définition-Proposition

Soient F, G et $F_1, \dots, F_k$ des sous-espaces vectoriels de E.

(a) On dit que *la somme de F et G est directe* si $F \cap G = \{0\}$.
On traduit cela en utilisant abusivement la notation $F \oplus G$ pour désigner $F + G$.

Plus généralement (c'est cohérent si $k = 2$), on dit que l*a somme de $F_1, \dots, F_k$ est directe* si, pour tous vecteurs $v_1 \in F_1, \dots, v_k \in F_k$ on a : $\quad v_1 + \dots + v_k = 0 \implies v_1 = \dots = v_k = 0.$
On traduit cela en utilisant abusivement la notation $F_1 \oplus \dots \oplus F_k$ pour désigner $F_1 + \dots + F_k$.

(b) On dit que *F et G sont supplémentaires dans E* (ou que *G est supplémentaires de F dans E*), ce qui s'écrit $E = F \oplus G$, s'ils vérifient les (i) et (ii) de la proposition précédente.

Plus généralement, on dit que $F_1, \dots, F_k$ *sont supplémentaires dans E*, ce qui s'écrit $E = F_1 \oplus \dots \oplus F_k$, s'ils vérifient l'une des deux propriétés équivalentes suivantes :
(i) pour tout $u \in E$, il existe $v_1 \in F_1, \dots, v_k \in F_k$ uniques tels que $u = v_1 + \dots + v_k$;
(ii) $E = F_1 + \dots + F_k$ et la somme de $F_1, \dots, F_k$ est directe.

Démonstration

(a) Le cas général avec $k = 2$ est bien compatible avec le cas de la somme de F et G car :
– si $F \cap G = \{0\}$ et $(v_1, v_2) \in F \times G$ vérifie $v_1 + v_2 = 0$, alors $v_1 = -v_2 \in F \cap G = \{0\}$ puis $v_1 = v_2 = 0$;
– si $u \in F \cap G$ et pour tout $(v_1, v_2) \in F \times G$ l'égalité $v_1 + v_2 = 0$ implique $v_1 = v_2 = 0$, alors en choisissant $v_1 = u$ et $v_2 = -u$ on constate que $u = 0$.

(b) Les idées de la proposition précédentes permettraient de démontrer encore ici l'équivalence entre (i) et (ii) (on ne détaille pas). □

Exemples

(1) Soient $v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ 0 \\ \lambda \end{pmatrix}$ avec $\lambda \in \mathbb{R}$ tel que $\lambda \neq -1$.
On constate en utilisant la méthode de Gauss que : $v_3 \notin \mathrm{Vect}(v_1, v_2)$ et $\mathrm{Vect}(v_1, v_2, v_3) = \mathbb{R}^3$.
On s'aperçoit donc que $\mathrm{Vect}(v_1, v_2)$ a pour supplémentaire $\mathrm{Vect}(v_3)$ dans $\mathbb{R}^3$.
En prenant différents λ, on constate que $\mathrm{Vect}(v_1, v_2)$ a une infinité de supplémentaires dans $\mathbb{R}^3$.

(2) Soient $\Pi_1 : x + y + z = 0$ et $\Pi_2 : x + 2y = 0$ dans $\mathbb{R}^3$.
On trouve par la méthode de Gauss que : $\Pi_1 \cap \Pi_2 \neq \{(0,0,0)\}$ $\big($et $\Pi_1 + \Pi_2 = \mathbb{R}^3\big)$.
Donc la somme de Π_1 et Π_2 n'est pas directe.

(3) Soient $D_1 = (Ox)$, $D_2 = (Oy)$, et $D_3 : x = y$ dans $\mathbb{R}^2$.
Les droites vectorielles D_1, D_2, D_3 vérifient $D_1 \cap D_2 = D_2 \cap D_3 = D_3 \cap D_1 = \{(0,0)\}$ et $D_1 + D_2 + D_3 = \mathbb{R}^2$ (facile).
Mais la somme de D_1, D_2, et D_3 n'est pas directe, car : $\underbrace{(1,0)}_{\in D_1} + \underbrace{(0,1)}_{\in D_2} + \underbrace{(-1,-1)}_{\in D_3} = (0,0)$.

3. Espace vectoriel quotient

Définition-Proposition

Soit F un sous-espace vectoriel de E.
On pose $\underbrace{\dot{x}}_{\text{« classe de } x \text{ »}} := \{x + y \,;\, y \in F\}$ quand $x \in E$, puis $E/F := \big\{\dot{x} \,;\, x \in E\big\}$.

L'ensemble E/F muni des « lois » déterminées par les égalités (où $\alpha \in \mathbb{K}$ et $u, v \in E/F$)

$u + v := \dot{\widehat{x+y}}$ indépendamment du choix d'éléments x, y de E tels que $u = \dot{x}$ et $v = \dot{y}$
$\alpha \times u := \dot{\widehat{\alpha x}}$ indépendamment du choix d'un élément x de E tel que $u = \dot{x}$

est un $\mathbb{K}$-espace vectoriel, appelé $\mathbb{K}$-*espace vectoriel quotient de E par F*.

Démonstration

Laissée en exercice : vérifier l'indépendance par rapport aux choix de représentants et montrer que le triplet $(E/F, +, \times)$ est un espace vectoriel, avec $0_{E/F} := \dot{\widehat{0_E}}$ et $-u := \dot{\widehat{-x}}$ indépendamment du choix d'un élément x de E tel que $u = \dot{x}$ quand $u \in E$. □

III. Bases et dimension

Dans cette partie, on fixe à nouveau un $\mathbb{K}$-espace vectoriel E.

1. Familles libres. Familles génératrices. Bases

Définition

(a) *Une famille numérotée de p vecteurs de E* où $p \in \mathbb{N}$, est un élément $(v_1, \dots, v_p)$ de E^p.

Plus généralement, *une famille de vecteurs de E* est une application $v\colon \underset{\text{ensemble}}{I} \to E,\ i \mapsto v_i$, qu'on notera plutôt $(v_i)_{i\in I}$.

Dans la suite de cette définition, on se donne suivant le cas : soit p vecteurs $v_1, \dots, v_p$ de E où $p \in \mathbb{N}$, soit une famille $(v_i)_{i\in I}$ indexée par I de vecteurs de E où I est un ensemble fixé.

(b) On dit que $\underbrace{\textit{la famille } (v_1,\dots,v_p) \textit{ est libre}}_{\text{ou « les vecteurs } v_1,\dots,v_p \text{ sont linéairement indépendants »}}$ si pour tous $\alpha_1, \dots, \alpha_p \in \mathbb{K}$, on a :

$$\alpha_1 v_1 + \cdots + \alpha_p v_p = 0 \quad \Longrightarrow \quad \alpha_1 = \cdots \alpha_p = 0\ .$$

Plus généralement, on dit que *la famille $(v_i)_{i\in I}$ est libre* si pour tous $k \in \mathbb{N}$ et $\underset{\text{distincts}}{i_1, \dots, i_k} \in I$, les seuls $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ tels que $\alpha_1 v_{i_1} + \cdots + \alpha_k v_{i_k} = 0$ sont $\alpha_1 = \cdots \alpha_k = 0$.

On exprimera qu'une famille de vecteurs de E n'est pas libre en disant qu'*elle est liée.*

(c) On dit que $\underbrace{\textit{la famille } (v_1,\dots,v_p) \textit{ est génératrice de } E}_{\text{ou « les vecteurs } v_1,\dots,v_p \text{ engendrent } E \text{ »}}$ si pour tout $v \in E$, il existe $\alpha_1, \dots, \alpha_p \in \mathbb{K}$ tels que : $v = \alpha_1 v_1 + \cdots + \alpha_p v_p$.

Plus généralement, on dit que *la famille $(v_i)_{i\in I}$ est génératrice de E* si pour tout $v \in E$, il existe $k \in \mathbb{N}$, $\underset{\text{distincts}}{i_1, \dots, i_k} \in I$, et $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ tels que : $v = \alpha_1 v_{i_1} + \cdots + \alpha_k v_{i_k}$.

Ainsi, la famille $(v_i)_{i\in I}$ est génératrice de E si et seulement si $\operatorname{Vect}\big(\{v_i\,;\, i \in I\}\big) = E$.

Exemples

(1) Les vecteurs 1 et i du $\mathbb{R}$-espace vectoriel $\mathbb{C}$ sont linéairement indépendants (sur $\mathbb{R}$), car : si $\alpha\,1 + \beta\,\mathrm{i} = 0$ avec $\alpha, \beta \in \mathbb{R}$ alors $\alpha = \beta = 0$.

Par contre, les vecteurs 1 et i du $\mathbb{C}$-espace vectoriel $\mathbb{C}$ sont linéairement dépendants (sur $\mathbb{C}$), car : $\alpha\,1 + \beta\,\mathrm{i} = 0$ quand $\alpha = \mathrm{i}$ et $\beta = -1$.

(2) Les vecteurs $f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \cos x$ et $g\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \sin x$ du $\mathbb{R}$-espace vectoriel $\mathbb{R}^{\mathbb{R}}$ sont linéairement indépendants, car : si $\alpha f + \beta g = 0$ avec $\alpha, \beta \in \mathbb{R}$ alors $\begin{cases} \alpha = (\alpha f + \beta g)(0) = 0 \\ \beta = (\alpha f + \beta g)(\frac{\pi}{2}) = 0 \end{cases}$.

(3) Soient $u_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}$, $u_2 = \begin{pmatrix}1\\2\\0\end{pmatrix}$, $u_3 = \begin{pmatrix}2\\1\\2\end{pmatrix}$ dans $\mathbb{R}^3$.

On étudie $(\star)\colon \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 = 0$ et $(\star\star)\colon \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 = \begin{pmatrix}x\\y\\z\end{pmatrix}$ à $\begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3$ fixé avec comme inconnue $(\alpha_1, \alpha_2, \alpha_3)$, et obtient par la méthode de Gauss que :

– la famille (u_1, u_2, u_3) est libre, cf.
$$\begin{array}{c}\begin{array}{ccc}\alpha_1&\alpha_2&\alpha_3\end{array}\\ \begin{pmatrix}1&1&2\\0&2&1\\1&0&2\end{pmatrix}\end{array} \xrightarrow[L_3 \leftarrow L_3 - L_1]{} \begin{array}{c}\begin{array}{ccc}\alpha_1&\alpha_2&\alpha_3\end{array}\\ \begin{pmatrix}1&1&2\\0&2&1\\0&-1&0\end{pmatrix}\end{array} \xrightarrow[C_2 \leftrightarrow C_3]{} \underbrace{\begin{array}{c}\begin{array}{ccc}\alpha_1&\alpha_3&\alpha_2\end{array}\\ \begin{pmatrix}1&2&1\\0&1&2\\0&0&-1\end{pmatrix}\end{array}}_{\text{unique solution } (0,0,0)}\ ;$$

– la famille (u_1, u_2, u_3) est génératrice de $\mathbb{R}^3$, cf.
$$\begin{array}{c}\begin{array}{ccc}\alpha_1&\alpha_2&\alpha_3\end{array}\\ \left(\begin{array}{ccc|c}1&1&2&x\\0&2&1&y\\1&0&2&z\end{array}\right)\end{array} \xrightarrow[\cdots]{} \cdots \xrightarrow[\cdots]{} \underbrace{\begin{array}{c}\begin{array}{ccc}\alpha_1&\alpha_3&\alpha_2\end{array}\\ \left(\begin{array}{ccc|c}1&2&1&///\\0&1&2&///\\0&0&-1&///\end{array}\right)\end{array}}_{\text{existence d'une solution}}.$$

Remarques (exercices)

(1) Dans le cas de 0, 1, ou 2 vecteurs, on constate que :
– la famille vide () est libre (vu la convention $\sum\limits_{1\leq k\leq p} \alpha_k\, v_k = 0$ quand $p=0$) ;
– une famille (v_1) avec $v_1 \in E$ est libre si et seulement si v_1 est non-nul ;
– deux vecteur de E sont linéairement indépendants si et seulement si ils sont non-colinéaires.

(2) Une famille $(v_i)_{i\in I}$ de vecteurs de E est liée si et seulement s'il existe un vecteur v_i, $i\in I$, qui est combinaison linéaire d'un nombre fini de vecteurs de la forme v_j avec $j\neq i$.
En particulier, toute famille de vecteurs de E dans laquelle se trouve le vecteur nul ou deux vecteurs égaux, est liée.

(3) On dira qu'une famille de vecteurs de E est « extraite » d'une famille $(v_i)_{i\in I}$ de vecteurs de E si elle est de la forme $(v_j)_{j\in J}$ avec $J \subseteq I$.

Tout famille extraite d'une famille libre de vecteurs de E, est libre. Tout famille de vecteurs de E dont une famille extraite est génératrice de E, est elle-même génératrice de E.

Notation

Soit I un ensemble. On note $E^{(I)}$ l'ensemble des familles $(v_i)_{i\in I}$ d'éléments de E indexées par I telles que la partie $J := \{i \in I \mid v_i \neq 0\}$ de I est finie. $\longleftarrow$[par exemple : $\mathbb{K}^{(\mathbb{N})} = \mathbb{K}[X]$]
Pour un tel $(v_i)_{i\in I} \in E^{(I)}$, on pose : $\sum\limits_{i\in I} v_i \underset{\text{(indépendant de l'ordre sur } J)}{:=} v_{i_1} + \cdots + v_{i_k}$ où $\underset{\text{distincts}}{\{i_1,\ldots,i_k\}} := J$.

On peut donc écrire $P = \sum\limits_{n\in\mathbb{N}} a_n X^n$ pour tout $P\in\mathbb{K}[X]$ de coefficients a_n, $n\in\mathbb{N}$.

Le choix de $E=\mathbb{K}$ fournit en particulier la notation $\mathbb{K}^{(I)}$.

Proposition

Une famille finie $(v_1,\ldots,v_p)$ de vecteurs de E est libre et génératrice de E, si et seulement si, pour tout $v\in E$ il existe $(\alpha_1,\ldots,\alpha_p)\in\mathbb{K}^p$ unique tel que $v=\alpha_1\, v_1+\cdots+\alpha_p\, v_p$.

Plus généralement, une famille $(v_i)_{i\in I}$ de vecteurs de E est libre et génératrice de E, si et seulement si, pour tout $v\in E$ il existe $(\alpha_i)_{i\in I}\in\mathbb{K}^{(I)}$ unique tel que $v=\sum\limits_{i\in I}\underbrace{\alpha_i\, v_i}_{\text{détermine un élément de } E^{(I)}}$.

Démonstration

On traite seulement le cas général.

$(\Rightarrow)$ On suppose que $(v_i)_{i\in I}$ est libre et génératrice de E. Soit $v\in E$.

Comme $(v_i)_{i\in I}$ est génératrice de E, il existe $(\alpha_i)_{i\in I}\in\mathbb{K}^{(I)}$ tel que $v=\sum\limits_{i\in I}\alpha_i\, v_i$.

On se donne une autre décomposition $v=\sum\limits_{i\in I}\beta_i\, v_i$ avec $(\beta_i)_{i\in I}\in\mathbb{K}^{(I)}$. On pose :
$\underset{\text{distincts}}{\{i_1,\ldots,i_k\}} := \{i\in I\,|\,\alpha_i\neq 0\}$, $\underset{\text{distincts}}{\{j_1,\ldots,j_l\}} := \{i\in I\,|\,\beta_i\neq 0\}$, et $\underset{\text{distincts}}{\{t_1,\ldots,t_m\}} := \{i_1,\ldots,i_k,j_1,\ldots,j_l\}$.

On a : $\begin{cases} v &= \alpha_{i_1}v_{i_1}+\cdots+\alpha_{i_k}v_{i_k} &= \alpha_{t_1}v_{t_1}+\cdots+\alpha_{t_m}v_{t_m} \\ v &= \beta_{j_1}v_{j_1}+\cdots+\beta_{j_l}v_{j_l} &= \beta_{t_1}v_{t_1}+\cdots+\beta_{t_m}v_{t_m}\end{cases}$ donc, par soustraction, $(\alpha_{t_1}-\beta_{t_1})v_{t_1}+\cdots+(\alpha_{t_m}-\beta_{t_m})v_{t_m}=0$ puis, la famille $(v_i)_{i\in I}$ étant libre, $\beta_{t_1}=\alpha_{t_1}$ et … et $\beta_{t_m}=\alpha_{t_m}$. Cela donne l'unicité de la famille $(\alpha_i)_{i\in I}$.

$(\Leftarrow)$ On suppose que : $\forall v\in E\quad \underbrace{\exists!}_{\text{« il existe un unique »}}(\alpha_i)_{i\in I}\in\mathbb{K}^{(I)}\quad v=\sum\limits_{i\in I}\alpha_i\, v_i$.

L'existence d'un $(\alpha_i)_{i\in I}\in\mathbb{K}^{(I)}$ associé à chaque v montre que $(v_i)_{i\in I}$ est génératrice de E.
L'unicité de $(\alpha_i)_{i\in I}\in\mathbb{K}^{(I)}$ quand $v=0$ montre que $(v_i)_{i\in I}$ est libre. $\square$

Définition

(a) On dit qu'une famille $(v_i)_{i\in I}$ de vecteurs de E est *une base de E* si $(v_i)_{i\in I}$ est libre et génératrice de E.

(b) On suppose que E a une base finie $\mathscr{B}=(v_1,\ldots,v_n)$.

Pour tout $v\in E$, les uniques scalaires $\alpha_1,\ldots,\alpha_n\in\mathbb{K}$ tels que $v=\alpha_1\, v_1+\cdots+\alpha_n\, v_n$ s'appellent *les coordonnées de v suivant $\mathscr{B}$*. On le notera : $\underset{/\mathscr{B}}{v}\left|\begin{matrix}\alpha_1\\ \vdots\\ \alpha_n\end{matrix}\right.$.

Exemples

(1) Le $\mathbb{K}$-espace vectoriel $\mathbb{K}^n$ admet la base suivante, appelée *base canonique de* $\mathbb{K}^n$:

$$\mathscr{B} := (e_1, \ldots, e_n) \quad \text{avec} \quad e_1 := (1,0,...,0),\ e_2 := (0,1,0,...,0),\ \ldots,\ e_n := (0,...,0,1).$$

(2) Soit $(H) : AX = 0$ avec $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathbb{K}^p$ un système d'équations linéaires homogène. On sait que l'ensemble $\mathscr{S}_H$ de ses solutions est un sous-espace vectoriel de $\mathbb{K}^p$.

La méthode de Gauss fournit des vecteurs w_1, ..., w_{p-r} de $\mathbb{K}^p$ tels que :

– $\mathscr{S}_H = \{t_1 w_1 + \cdots + t_{p-r} w_{p-r} \ ;\ t_1, \ldots, t_{p-r} \in \mathbb{K}\}$; $\quad \longleftarrow \left[\text{cf. } \mathscr{S}_H : \begin{cases} x_{j_1} = \ldots \\ \vdots \\ x_{j_r} = \ldots \\ x_{j_{r+1}} = t_1 \\ \vdots \\ x_{j_p} = t_{p-r} \end{cases},\quad t_1, t_{p-r} \in \mathbb{R}. \right]$

– si $t_1 w_1 + \cdots + t_{p-r} w_{p-r} = 0$, alors $\begin{cases} \text{coordonnée } j_{r+1} : \ t_1 + 0 + \cdots + 0 = 0 \\ \vdots \\ \text{coordonnée } j_p \ : 0 + \cdots + 0 + t_{p-r} = 0 \end{cases}$, alors $t_1 = \cdots = t_{p-r} = 0$.

Ainsi $(w_1, ..., w_{p-r})$ est une base de $\mathscr{S}_H$.

(3) La famille $(X^n)_{n \geq 0}$ est une base du $\mathbb{K}$-espace vectoriel $\mathbb{K}[X]$.

Soit $N \in \mathbb{N}$. L'ensemble $\mathbb{K}_N[X]$ est le sous-espace vectoriel $\mathrm{Vect}(1, X, ..., X^N)$ de $\mathbb{K}[X]$. La famille $(1, X, ..., X^N)$ est une base de $\mathbb{K}_N[X]$, appelée *base canonique de* $\mathbb{K}_N[X]$.

(4) Le $\mathbb{K}$-espace vectoriel $\mathfrak{M}_{n,p}(\mathbb{K})$ admet la base suivante :

$$(E_{1,1}, ..., E_{1,p}; E_{2,1}, ..., E_{2,p}; \ldots; E_{n,1}, ..., E_{n,p}) \quad \text{avec} \quad E_{i,j} := \underbrace{\begin{array}{c} \text{colonne } j \\ \downarrow \\ \begin{pmatrix} 0 & & 0 \\ & 1 & \\ 0 & & 0 \end{pmatrix} \end{array}}_{0 \text{ ailleurs qu'en l'emplacement } (i,j)} \leftarrow \text{ligne } i$$

En effet, pour tous $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in \mathfrak{M}_{n,p}(\mathbb{K})$ et $(x_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in \mathbb{K}^{\{1,...,n\} \times \{1,...,p\}}$, on a :

$$A = \sum_{i,j} x_{ij} E_{ij} \iff \big(\forall (i,j) \in \{1, ..., n\} \times \{1, ..., p\} \quad x_{ij} = a_{ij}\big).$$

2. Dimension. Rang

Théorème (« théorème de la base incomplète »)

On suppose que E a une famille génératrice finie $(g_1, ..., g_k)$.

Toute famille libre $(h_1, ..., h_l)$ de E se complète à l'aide de certains vecteurs parmi $g_1, ..., g_k$ en une base $(e_1, ..., e_d)$ de E.

Démonstration

Rappel : toute partie non-vide de $\{0, 1, ..., N\}$, $N \in \mathbb{N}$, a un plus grand élément (par récurrence).

On note $\mathcal{F}$ l'ensemble des éléments n de $\{l, l+1, ..., k+l\}$ pour lesquels il existe une famille libre $(e_1, ..., e_n)$ de E vérifiant : $e_1 = h_1$, ..., $e_l = h_l$, et e_{l+1}, ..., e_n sont parmi g_1, ..., g_k.

L'ensemble $\mathcal{F}$ est non-vide, car il contient l qui correspond au cas $(e_1, ..., e_n) = (h_1, ..., h_l)$. Soit d le plus grand élément de $\mathcal{F}$ et $(e_1, ..., e_d)$ une famille associée. Pour terminer la démonstration, il reste à vérifier que cette famille libre $(e_1, ..., e_d)$ est aussi génératrice de E.

Soit $i \in \{1, ..., k\}$. Comme d est maximal, la famille $(e_1, ..., e_d, g_i)$ est liée. On dispose d'une égalité $\underbrace{\alpha_1 e_1 + ... + \alpha_d e_d + \alpha g_i}_{\alpha_1, ..., \alpha_d, \alpha \in \mathbb{K} \text{ non-tous nuls}} = 0$ donc $\alpha \neq 0$ et $g_i = -\frac{1}{\alpha}(\alpha_1 e_1 + ... + \alpha_d e_d) \in \mathrm{Vect}(e_1, ..., e_d)$.

D'où : $E = \underbrace{\mathrm{Vect}(g_1, ..., g_k)}_{\text{plus petit sous-ev de } E \text{ contenant } g_1, ..., g_k} \subseteq \mathrm{Vect}(e_1, ..., e_d)$, ce qui montre que $e_1, ..., e_n$ engendrent E. □

Corollaire 1 (« théorème de la dimension »)

On suppose que E a une famille génératrice finie.

On a : E a une base finie $(e_1, ..., e_d)$.

Les autres bases de E ont aussi d vecteurs.

DÉMONSTRATION

D'après le théorème de la base incomplète, on obtient une base finie $\mathscr{B}=(e_1,...,e_d)$ de E en complétant la famille () à l'aide de vecteurs pris dans une famille génératrice finie fixée de E.

On va vérifier que :
(i) toute base finie $\mathscr{B}'=(e'_1,...,e'_{d'})$ de E a d vecteurs, et pour cela quitte à échanger les rôles de $\mathscr{B}$ et $\mathscr{B}'$, il suffit de prouver que $d'\le d$:
(ii) E n'a pas de base infinie $\mathscr{B}''$.

Comme $\mathscr{B}'$ et $\mathscr{B}''$ sont libres, il reste à démontrer que toute famille libre $\mathscr{L}$ dans E a au plus d vecteurs. Une famille libre infinie dans E aurait une famille extraite formée de $d+1$ vecteurs, elle-même libre. On peut donc supposer que $\mathscr{L}$ est finie, de la forme $(h_1,...,h_l)$.

On considère l'équation $x_1h_1+...+x_lh_l=0$ d'inconnue $(x_1,...,x_l)\in\mathbb{K}^l$.
En notant $\underset{/\mathscr{B}}{h_1}\left|\begin{matrix}a_{11}\\ \vdots\\ a_{d1}\end{matrix}\right.$, ..., $\underset{/\mathscr{B}}{h_l}\left|\begin{matrix}a_{1l}\\ \vdots\\ a_{dl}\end{matrix}\right.$ elle équivaut à : $(H)\begin{cases}a_{11}\,x_1+\cdots+a_{1l}\,x_l=0\\ \quad\cdots\\ a_{d1}\,x_1+\cdots+a_{dl}\,x_l=0\end{cases}$.
Si $l>d$, alors comme le nombre d'équations est strictement supérieur au nombre d'inconnues, (H) a une solution non-nulle ; cela contredit la condition « $(h_1,...,h_l)$ est libre ». Ainsi : $l\le d$.

Variante

On a : $h_1,...,h_l\in\mathrm{Vect}(e_1,...,e_d)$. Si $l>d$, on aurait donc : $(\star)$ $h_1,...,h_{l+1}\in\mathrm{Vect}(e_1,...,e_d)$.

Pour montrer que $(\star)$ est fausse il reste à voir que pour tout $k\ge 0$, on a :
(H_k) pour tous $v_1,...,v_k\in E$ et $w_1,...,w_{k+1}\in\mathrm{Vect}(v_1,...,v_k)$, la famille $(w_1,...,w_{k+1})$ est liée.

(i) L'assertion (H_0), qui signifie que pour tout $w_1\in\{0\}$ la famille (w_1) est liée, est vraie.

(ii) Soit $k\ge 1$ tel que (H_{k-1}) est vraie. On se donne $v_1,...,v_k\in E$ et $w_1=\lambda_{1,1}v_1+...+\lambda_{1,k}v_k$, $\ldots$, $w_{k+1}=\lambda_{k+1,1}v_1+...+\lambda_{k+1,k}v_k\in\mathrm{Vect}(v_1,...,v_k)$. On vérifie que $(w_1,...,w_{k+1})$ est liée.
On peut supposer que $w_1\neq 0$, puis – quitte à renuméroter les v_i – que $\lambda_{1,1}\neq 0$.
En appliquant (H_{k-1}) aux vecteurs $u_2:=w_2-\frac{\lambda_{2,1}}{\lambda_{1,1}}w_1$, $\ldots$, $u_{k+1}:=w_{k+1}-\frac{\lambda_{k+1,1}}{\lambda_{1,1}}w_1$ de $\mathrm{Vect}(v_2,...,v_k)$, on obtient l'existence de scalaires $\alpha_2,...,\alpha_{k+1}\in\mathbb{K}$ non-tous nuls, tels que $\alpha_2u_2+...+\alpha_{k+1}u_{k+1}=0$. D'où l'égalité $\alpha_2w_2+...+\alpha_{k+1}w_{k+1}=\big(\alpha_2\frac{\lambda_{2,1}}{\lambda_{1,1}}+...+\alpha_{k+1}\frac{\lambda_{k+1,1}}{\lambda_{1,1}}\big)w_1$ qui montre que $(w_1,...,w_{k+1})$ est liée. □

Définition

(a) On dit que le $\mathbb{K}$-espace vectoriel E est *de dimension finie* s'il a une base finie $(e_1,...,e_d)$. Dans ce cas, on appelle *dimension de* E le nombre d, noté $\underbrace{\dim E}_{\text{ou }\dim_{\mathbb{K}}E}$, de vecteurs de ses bases.
(D'après le théorème de la dimension, ce d est bien indépendant du choix de la base.)

(b) Soient $v_1,...,v_p\in E$. On appelle *rang de la famille* $(v_1,...,v_p)$ le nombre
$$\mathrm{rg}\,(v_1,...,v_p):=\dim\mathrm{Vect}(v_1,...,v_p).$$
(Cela a un sens car d'après le théorème de la dimension, $\mathrm{Vect}(v_1,...,v_p)$ est de dimension finie.)

Exemples

(1) On reprend certaines définitions du paragraphe II.1. :
– $E=\{0\}$ si et seulement si $\dim E=0$;
– E est une droite vectorielle si et seulement si $\dim E=1$;
– E est un plan vectoriel si et seulement si $\dim E=2$.

(2) Tout $z\in\mathbb{C}$ s'écrit de manière unique $z=x\,1+y\,\mathrm{i}$ avec $x,y\in\mathbb{R}$.
Donc $(1,\mathrm{i})$ est une base du $\mathbb{R}$-espace vectoriel $\mathbb{C}$. En particulier : $\dim_{\mathbb{R}}\mathbb{C}=2$.
Par contre, comme $\mathbb{C}=\mathrm{Vect}_{\mathbb{C}}(1)$ avec $1\neq 0$, l'exemple 1 montre que : $\dim_{\mathbb{C}}\mathbb{C}=1$.

(3) Compte tenu des bases exhibées à la fin du paragraphe 1, on a :
$\dim\mathbb{K}^n=n$, $\dim\mathbb{K}_N[X]=N+1$ et $\dim\mathfrak{M}_{n,p}(\mathbb{K})=np$.

Corollaire 2

On suppose que E est de dimension finie. On note d la dimension de E.

(a) Toute famille libre de E a au plus d vecteurs ;
quand elle a d vecteurs, c'est une base de E.

(b) Toute famille génératrice de E a au moins d vecteurs ;
quand elle a d vecteurs, c'est une base de E.

Démonstration

(a) Une famille libre infinie dans E aurait une famille extraite formée de $d+1$ vecteurs, qui serait elle-même encore libre. On se contente donc de traiter le cas des familles libres finies.

Une famille libre finie de l vecteurs de E se complète d'après le théorème de la base incomplète en une base de E qui a d vecteurs d'après le théorème de la dimension, donc $l \leq d$.

(b) Une famille génératrice de E infinie a au moins d vecteurs. On se contente donc de traiter le cas des familles génératrices finies.

Soit $(g_1, ..., g_k)$ une famille génératrice finie de vecteurs de E. D'après le théorème de la base incomplète, la famille $()$ se complète en une base $\mathscr{B}'$ de E extraite de $(g_1, ..., g_k)$.
D'après le théorème de la dimension, la famille $\mathscr{B}'$ a d vecteurs. D'où le résultat. □

Remarques

(1) On reprend l'exemple 3 du début du paragraphe 1 : $u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, $u_3 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$.

On a vu à l'aide d'un premier calcul que (u_1, u_2, u_3) est libre. On remarque que la base canonique de $\mathbb{R}^3$ est formée de 3 vecteurs. On peut donc maintenant utiliser le a) du corollaire 2 pour en conclure, sans autre calcul, que (u_1, u_2, u_3) est génératrice de $\mathbb{R}^3$.

(2) On considère les applications $\begin{array}{rl} f_n \colon & \mathbb{R} \to \mathbb{R} \\ & x \mapsto x^n \end{array}$, $n \in \mathbb{N}$.

Elles forment une famille $(f_n)_{n\in\mathbb{N}}$ qui est libre dans $\mathbb{R}^{\mathbb{R}}$ car :
si des éléments $n_1 < ... < n_k$ de $\mathbb{N}$ et $\alpha_1, ..., \alpha_k \in \mathbb{R}$ vérifient $\alpha_1 f_{n_1} + ... + \alpha_k f_{n_k} = 0$,
alors $\alpha_1 = \frac{1}{n_1!}(\alpha_1 f_{n_1} + ... + \alpha_k f_{n_k})^{(n_1)}(0) = 0$ et ... et $\alpha_k = \frac{1}{n_k!}(\alpha_1 f_{n_1} + ... + \alpha_k f_{n_k})^{(n_k)}(0) = 0$.

Ainsi le a) du corollaire 2, appliqué à la famille $(f_n)_{n\in\mathbb{N}}$ qui est libre et indexée par un ensemble infini, montre que $\mathbb{R}^{\mathbb{R}}$ n'est pas un espace vectoriel de dimension finie.

D'autre part, la famille $(f_n)_{n\in\mathbb{N}}$ n'est pas génératrice de $\mathbb{R}^{\mathbb{R}}$ car $\mathrm{Vect}(\{f_n \,;\, n \in \mathbb{N}\})$ qui est formée de fonctions dérivables ne contient donc pas l'application valeur absolue.

Proposition

On suppose que E a une base finie $\mathscr{B} = (e_1, ..., e_n)$.

Soient $\underset{/\mathscr{B}}{v_1}\left|\begin{matrix} a_{11} \\ \vdots \\ a_{n1} \end{matrix}\right., \ldots, \underset{/\mathscr{B}}{v_p}\left|\begin{matrix} a_{1p} \\ \vdots \\ a_{np} \end{matrix}\right. \in E$.

(a) On transforme $\overset{x_1 \ldots x_p}{A} := \begin{array}{c} \begin{matrix} x_1 & \ldots & x_p \end{matrix} \\ \begin{pmatrix} a_{11} & \ldots & a_{1p} \\ & \ldots & \\ a_{n1} & \ldots & a_{np} \end{pmatrix} \end{array}$ en $\overset{x_{j_1} \ldots x_{j_p}}{A'} = \begin{array}{c} \begin{matrix} x_{j_1} & \ldots & x_{j_r} & x_{j_{r+1}} \ldots & x_{j_p} \end{matrix} \\ \left(\begin{array}{cccc|ccc} d_1 & & & & & & \\ 0 & \ddots & & & & & \\ \vdots & \ddots & \ddots & & & & \\ 0 & \cdots & 0 & d_r & & & \\ \hline 0 & & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & & \cdots & 0 & 0 & \cdots & 0 \end{array}\right) \end{array}$

par la méthode de Gauss, avec $d_1 \neq 0$, ..., $d_r \neq 0$.

On a : $(v_{j_1}, ..., v_{j_r})$ est une base de $\mathrm{Vect}(v_1, ..., v_p)$.

(b) En particulier : $\mathrm{rg}\,(v_1, ..., v_p) = r$ « calcul du rang par la méthode de Gauss ».
(Le nombre r est donc indépendant de la manière dont on applique la méthode de Gauss.)

Démonstration

Mêmes arguments que pour le calcul du rang d'une matrice par la méthode de Gauss. □

Remarque (importante au niveau de la rédaction au cours des partiels et examens)

La partie a) de cette proposition n'est pas classique.

Dans chaque exercice on la déduira des calculs simultanés de $\mathrm{rg}\,(v_1, ..., v_r)$ et $\mathrm{rg}\,(v_{j_1}, ..., v_{j_r})$:
– on a $\mathrm{rg}\,(v_1, ..., v_r) = r$ par calcul du rang par la méthode de Gauss ;
– on a aussi $\mathrm{rg}\,(v_{j_1}, ..., v_{j_r}) = r$ en reprenant le calcul précédent et rayant les colonnes qui ne sont pas sous les symboles $x_{j_1}, ..., x_{j_r}$, ce qui montrera cf. plus bas, que $(v_{j_1}, ..., v_{j_r})$ est libre ;
– $\mathrm{Vect}(v_{j_1}, ..., v_{j_r})$ est un sous-espace vectoriel de dimension r de $\mathrm{Vect}(v_1, ..., v_r)$ qui est aussi de dimension r, ce qui montrera cf. plus bas, que $(v_{j_1}, ..., v_{j_r})$ engendre $\mathrm{Vect}(v_1, ..., v_p)$.

3. Dimension des sous-espaces vectoriels

Proposition

On suppose E de dimension finie.

Tout sous-espace vectoriel F de E est de dimension finie avec $\dim F \leq \dim E$.

De plus, si $\dim F = \dim E$, alors $F = E$.

Démonstration

Toute famille libre $(v_1, ..., v_p)$ de F est libre dans E, donc vérifie : $p \leq \dim E$.
Il en existe, par exemple la famille vide. On choisit $(v_1, ..., v_p)$ de façon à avoir p maximal.

Soit $v \in F$. Comme p est maximal, la famille $(v_1, ..., v_p, v)$ est liée. On dispose d'une égalité $\underbrace{\alpha_1 v_1 + ... + \alpha_p v_p + \alpha v}_{\alpha_1, ..., \alpha_p, \alpha \in \mathbb{K} \text{ non-tous nuls}} = 0$ donc $\alpha \neq 0$ et $v = -\frac{1}{\alpha}(\alpha_1 v_1 + ... + \alpha_p v_p) \in \mathrm{Vect}(v_1, ..., v_p)$.

Par conséquent, $(v_1, ..., v_p)$ est génératrice de F et finalement $(v_1, ..., v_p)$ est une base de F.

En particulier : F est de dimension finie et $\dim F = p \leq \dim E$.

De plus : si $\dim F = \dim E$, alors la famille libre $(v_1, ..., v_p)$ de E qui est formée de $\dim E$ vecteurs est une base de E, puis $F = \mathrm{Vect}(v_1, ..., v_p) = E$. $\square$

Remarque

D'après cette proposition, les sous-espaces vectoriels de $\mathbb{R}^3$ ont une dimension, égale à 0 ou 1 ou 2 ou 3, le cas de la dimension 3 n'étant atteint que par $\mathbb{R}^3$ lui-même. Les sous-espaces vectoriels de $\mathbb{R}^3$ sont donc $\{0\}$, les droites vectorielles de $\mathbb{R}^3$, les plans vectoriels de $\mathbb{R}^3$, et $\mathbb{R}^3$.

Proposition (utile)

Soient $v_1, ..., v_p \in E$. On pose : $r = \mathrm{rg}\,(v_1, ..., v_p)$.

(a) La famille $(v_1, ..., v_p)$ est libre si et seulement si $r = p$.

(b) On suppose E de dimension finie, notée n.
La famille $(v_1, ..., v_p)$ est génératrice de E si et seulement si $r = n$.

Démonstration

(a) La famille $(v_1, ..., v_p)$, de p vecteurs, engendre $\mathrm{Vect}(v_1, ..., v_p)$ qui est de dimension r.
Elle est libre si et seulement si elle est base de $\mathrm{Vect}(v_1, ..., v_p)$, ce qui équivaut (compte tenu du corollaire 2 (b) du théorème de la base incomplète) à $r = p$.

(b) La famille $(v_1, ..., v_p)$ est génératrice de E si et seulement si $\mathrm{Vect}(v_1, ..., v_p) = E$, c'est à dire (vu la proposition précédente) $\dim \mathrm{Vect}(v_1, ..., v_p) = \dim E$, c'est à dire $r = n$. $\square$

Remarque (importante)

On revient sur l'exemple 3 du début du paragraphe 1 : $u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, $u_3 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \in \mathbb{R}^3$.

La rédaction la plus rapide pour démontrer que (u_1, u_2, u_3) est une base de $\mathbb{R}^3$ consiste à prouver par la méthode de Gauss que $\mathrm{rg}\,(u_1, u_2, u_3) = 3$.

Proposition

Soient E_1 et E_2 deux $\mathbb{K}$-espaces vectoriels de dimension finie.
L'espace vectoriel $E_1 \times E_2$ est de dimension finie et $\dim(E_1 \times E_2) = \dim E_1 + \dim E_2$.
(Cela sera généralisé plus bas, cf. $E_1 \times E_2 = (E_1 \times \{0\}) \oplus (\{0\} \times E_2)$.)

Démonstration

On se donne une base $(v_1, ..., v_p)$ de E_1 et une base $(w_1, ..., w_q)$ de E_2
Soit $(v, w) \in E_1 \times E_2$. Pour tous $x_1, ..., x_p, y_1, ..., y_q \in \mathbb{K}$, on a :

$$\begin{aligned} & (v, w) = x_1(v_1, 0) + ... + x_p(v_p, 0) + y_1(0, w_1) + ... + y_q(0, w_q) \\ \Longleftrightarrow \quad & v = x_1 v_1 + ... + x_p v_p \ \text{ et } \ w = y_1 w_1 + ... + y_q w_q \end{aligned}$$

L'unicité de tels scalaires $x_1, ..., x_p, y_1, ..., y_q$ montre que $\big((v_1, 0), ..., (v_p, 0), (0, w_1), ..., (0, w_q)\big)$ est une base de $E_1 \times E_2$, ce qui donne le résultat. □

Proposition

On suppose que E est de dimension finie. Soient F et G deux sous-espaces vectoriels de E, $(v_1, ..., v_p)$ une base de F et $(w_1, ..., w_q)$ une base de G.

On a : $\underbrace{E = F \oplus G}_{\text{« } F \text{ et } G \text{ sont supplémentaires dans } E \text{ »}}$ si et seulement si $(v_1, ..., v_p, w_1, ..., w_q)$ est une base de E.

Démonstration

On a $F + G = \mathrm{Vect}(v_1, ..., v_p) + \mathrm{Vect}(w_1, ..., w_q) = \mathrm{Vect}(v_1, ..., v_p, w_1, ..., w_q)$.
Donc : $F + G = E$ si et seulement si $(v_1, ..., v_p, w_1, ..., w_q)$ engendre E.
Il reste à vérifier que : $F \cap G = \{0\}$ si et seulement si $(v_1, ..., v_p, w_1, ..., w_q)$ est libre.

• On suppose que $F \cap G = \{0\}$.
Soient $\alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q \in \mathbb{K}$ tels que $\alpha_1 v_1 + ... + \alpha_p v_p + \beta_1 w_1 + ... + \beta_q w_q = 0$.
On a $\alpha_1 v_1 + ... + \alpha_p v_p = -\beta_1 w_1 - ... - \beta_q w_q \in F \cap G$. Puisque $F \cap G = \{0\}$, on en déduit que $\alpha_1 v_1 + ... + \alpha_p v_p = 0$ et $\beta_1 w_1 + ... + \beta_q w_q = 0$. Or $(v_1, ..., v_p)$ et $(w_1, ..., w_q)$ sont libres.
D'où $\alpha_1 = ... = \alpha_p = \beta_1 = ... = \beta_q = 0$.
Cela montre que la famille $(v_1, ..., v_p, w_1, ..., w_q)$ est libre.

• On suppose que la famille $(v_1, ..., v_p, w_1, ..., w_q)$ est libre.
Soit $u \in F \cap G$. Vu que $(v_1, ..., v_p)$ est une base de F et $(w_1, ..., w_q)$ est une base de G, il existe $x_1, ..., x_p, y_1, ..., y_q \in \mathbb{K}$ tels que $u = x_1 v_1 + ... + x_p v_p$ et $u = y_1 w_1 + ... + y_q w_q$. On a donc $x_1 v_1 + ... + x_p v_p - y_1 w_1 - ... - y_q w_q = 0$. Puisque $(v_1, ..., v_p, w_1, ..., w_q)$ est libre, il en résulte que $x_1 = ... = x_p = y_1 = ... = y_q = 0$, puis $u = 0v_1 + ... + 0v_p = 0$.
Cela montre que $F \cap G = \{0\}$. □

Remarque (importante)

On suppose que E est de dimension finie. Soit F un sous-espace vectoriel de E. On peut compléter une base $(v_1, ..., v_p)$ de F (qui est en particulier libre) à l'aide de certains vecteurs de E (par exemple les vecteurs d'une base donnée de E) en une base $(v_1, ..., v_p, w_1, ..., w_q)$ de E.

D'après la proposition, $G := \mathrm{Vect}(w_1, ..., w_q)$ est un supplémentaire de F dans E.

Corollaire (de la remarque)

On suppose que E est de dimension finie.
Tout sous-espace vectoriel F de E admet un supplémentaire G dans E.

Proposition

On suppose que E est de dimension finie.
Soient F et G deux sous-espaces vectoriels de E.

(a) On a : si $E = F \oplus G$, alors $\dim E = \dim F + \dim G$.

(b) Plus généralement : $\dim(F + G) + \dim(F \cap G) = \dim F + \dim G$.

Démonstration

(a) Conséquence de la proposition précédente.

(Cf. la forme vectorielle du théorème du rang pour la projection canonique de F sur $(F+G)/G$.)

(b) • On se donne un sous-espace vectoriel F' de F tel que : $F = F' \oplus (F \cap G)$ (existe par la remarque précédente). D'après (a), on a : $\dim F \overset{(*)}{=} \dim(F') + \dim(F \cap G)$.

• On a : $F + G = F' + \big(\underbrace{(F \cap G)}_{\subseteq G} + G = F' + G$ et $\underbrace{F'}_{\subseteq F} \cap G \subseteq F' \cap (F \cap G) = \{0\}$.

D'où : $F + G = F' \oplus G$. D'après (a), on a : $\dim(F+G) \overset{(**)}{=} \dim(F') + \dim G$.

• On obtient le résultat en soustrayant $(*)$ à $(**)$.

<u>Variante</u>

On complète une base $(u_1, ..., u_m)$ de $F \cap G$ d'une part en une base $(v_1, ..., v_p)$ de F et d'autre part en une base $(w_1, ..., w_q)$ de G, avec $(u_1, ..., u_m) = (v_1, ..., v_m) = (w_1, ..., w_m)$.

On constate que $(u_1, ..., u_m, v_{m+1}, ..., v_p, w_{m+1}, ..., w_q)$ est une base de $F+G$, car :
si $\alpha_1 u_1 + \cdots + \alpha_m u_m + \beta_{m+1} v_{m+1} + \cdots + \beta_p v_p + \gamma_{m+1} w_{m+1} + \cdots + \gamma_q w_w = 0$,
alors $\alpha_1 u_1 + ... + \alpha_m u_m + \beta_{m+1} v_{m+1} + ... + \beta_p v_p, \alpha_1 u_1 + ... + \alpha_m u_m + \gamma_{m+1} w_{m+1} + ... + \gamma_q w_w \in F \cap G$
puis $\alpha_1 = ... = \alpha_m = \beta_{m+1} = ... = \beta_p v_p = \gamma_{m+1} = ... = \gamma_q = 0$. □

Définition-Proposition

On suppose que E est de dimension finie, notée n.

On dit qu'un sous-espace vectoriel H de E est un *hyperplan* si $\dim H = n-1$, ce qui équivaut au fait que H a un supplémentaire dans E qui est une droite.

Démonstration

$(\Rightarrow)$ La dernière remarque montre que H a au moins un supplémentaire D dans E. Vu le b) de la proposition, ce supplémentaire D a pour dimension 1 et est donc une droite.

$(\Leftarrow)$ Cela découle aussi du b) de la proposition. □

Exemple

Soit $(a_1, ..., a_n) \in \mathbb{R}^n \setminus \{0\}$. La partie $F\colon a_1 x_1 + \cdots + a_n x_n = 0$ de $\mathbb{R}^n$ est un hyperplan vectoriel de $\mathbb{R}^n$ car on constate que la méthode de Gauss fournit une base $(w_1, ..., w_{n-1})$ de F.

4. Dimension d'un espace vectoriel quotient

Proposition

On suppose que E est de dimension finie. Soit F un sous-espace vectoriel de E.

L'espace vectoriel E/F est de dimension finie, et : $\dim E/F = \dim E - \dim F$.

Démonstration

Soit $(v_1, ..., v_p)$ une base de F. On la complète en une base $(v_1, ..., v_p, w_1, ..., w_q)$ de E.
On vérifie que $(\dot{w_1}, ..., \dot{w_q})$ est une base de E/F.

Tout d'abord, on a : $E/F = \mathrm{Vect}(\underbrace{\dot{v_1}}_{\dot{0}}, ..., \underbrace{\dot{v_p}}_{\dot{0}}, \dot{w_1}, ..., \dot{w_q}) = \mathrm{Vect}(\dot{w_1}, ..., \dot{w_q})$.

La famille $(\dot{w_1}, ..., \dot{w_q})$ est donc génératrice de E/F.

De plus, pour tous $\beta_1, ..., \beta_q \in \mathbb{K}$ vérifiant $\underbrace{\beta_1 \dot{w_1} + ... + \beta_q \dot{w_q}}_{\text{classe de } \beta_1 w_1 + ... + \beta_q w_q} = \dot{0}$, on a :

$\beta_1 w_1 + ... + \beta_q w_q \in F$, donc il existe $\alpha_1, ..., \alpha_p \in \mathbb{K}$ tels que $\beta_1 w_1 + ... + \beta_q w_q = \alpha_1 v_1 + ... + \alpha_p v_p$, puis l'égalité $\underbrace{\alpha_1 v_1 + ... + \alpha_p v_p + (-\beta_1) w_1 + ... + (-\beta_q) w_q}_{\text{où } v_1, ..., v_p, w_1, ..., w_q \text{ sont linéairement indépendants}} = 0$ montre que $\beta_1 = ... = \beta_q = 0$.

La famille $(\dot{w_1}, ..., \dot{w_q})$ est donc libre, et finalement base de E/F ce qui donne le résultat. □

Remarque

La condition « $\dim H = \dim E - 1$ » pour qu'un sous-espace vectoriel H d'un espace vectoriel de dimension finie E soit un hyperplan est donc équivalente à : $\dim E/H = 1$.

Ch. 4. Applications linéaires sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$

I. Applications linéaires

Dans toute cette partie, on fixe trois $\mathbb{K}$-espace vectoriel E, F et G.

1. Généralités

Définition

(a) On dit qu'une application $f : E \to F$ est *linéaire* si :

$$f(\alpha\, v + \beta\, w) = \alpha\, f(v) + \beta\, f(w) \ \text{ pour tous } \alpha, \beta \in \mathbb{K} \text{ et } v, w \in E.$$

Dans ce cas : $\left|\begin{array}{l} f(0_E) = 0_F \text{ en choisissant } \alpha = \beta = 0 \text{ et } v = w = 0_E\,; \\ f\Big(\sum\limits_{k=1}^{p} \alpha_k v_k\Big) \overset{\text{récurrence sur } p}{=} \sum\limits_{k=1}^{p} \alpha_k f(v_k) \ \text{ pour } \alpha_1, ..., \alpha_p \in \mathbb{K} \text{ et } v_1, ..., v_p \in E. \end{array}\right.$

(b) On appelle plus précisément :
– *forme linéaire sur* E une application linéaire de E dans $\mathbb{K}$;
– *endomorphisme de (l'espace vectoriel)* E une application linéaire de E dans E ;
– *isomorphisme (d'espaces vectoriels) de* E *sur* F une bijection linéaire de E sur F ;
– *automorphisme de (l'espace vectoriel)* E une bijection linéaire de E sur E.

(c) On dit que les espaces vectoriels E et F sont isomorphes s'il existe un isomorphisme d'espaces vectoriels de E sur F.

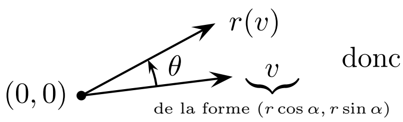
Exemples

(1) Soit $A \in \mathfrak{M}_{n,p}(\mathbb{K})$. On considère $f : \begin{array}{ccc} \mathbb{K}^p & \longrightarrow & \mathbb{K}^n \\ X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} & \longmapsto & Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \end{array}$ où $Y := AX$.

L'application f est linéaire, car pour tous $\alpha, \beta \in \mathbb{K}$ et $V, W \in \mathbb{K}^p$ on a :

$$f(\alpha V + \beta W) = A(\alpha V + \beta W) = \alpha AV + \beta AW = \alpha f(V) + \beta f(W).$$

(2) Soit $\theta \in \mathbb{R}$. L'application $r : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & \begin{pmatrix} x' \\ y' \end{pmatrix} \end{array}$ avec $\begin{cases} x' = (\cos\theta)\, x - (\sin\theta)\, y \\ y' = (\sin\theta)\, x + (\cos\theta)\, y \end{cases}$ est $\underbrace{\text{linéaire}}_{\text{cf. l'exemple (1) ci-dessus}}$.

Géométriquement : 

donc r *est la rotation (vectorielle) d'angle* θ.

(3) La *dérivation* $\delta : \begin{array}{ccc} \mathbb{K}[X] & \to & \mathbb{K}[X] \\ P & \mapsto & P' \end{array}$ est un endomorphisme de l'espace vectoriel $\mathbb{K}[X]$.

En effet, étant donnés $\alpha, \beta \in \mathbb{K}$ et $P = a_p X^p + \cdots + a_1 X + a_0$, $Q = b_q X^q + \cdots + b_1 X + b_0 \in \mathbb{K}[X]$, en posant $n = \max(p, q)$, $a_i = 0$ quand $i > p$ et $b_j = 0$ quand $j > q$, on a :

$$\begin{aligned} \delta(\alpha P + \beta Q) &= \delta\big((\alpha a_n + \beta b_n)X^n + \cdots + (\alpha a_1 + \beta b_1)X + (\alpha a_0 + \beta b_0)\big) \\ &= n(\alpha a_n + \beta b_n)X^{n-1} + \cdots + 2(\alpha a_2 + \beta b_2)X + (\alpha a_1 + \beta b_1) \\ &= \alpha(n a_n X^{n-1} + \cdots + 2a_2 X + a_1) + \beta(n b_n X^{n-1} + \cdots + 2b_2 X + b_1) \\ &= \alpha\delta(P) + \beta\delta(Q) \end{aligned}$$

(4) La transposition $t : \begin{array}{ccc} \mathfrak{M}_{n,p}(\mathbb{K}) & \to & \mathfrak{M}_{p,n}(\mathbb{K}) \\ A & \mapsto & {}^t\!A \end{array}$ est linéaire (exercice !).

Proposition

(a) L'ensemble, noté $\mathscr{L}(E,F)$, des applications linéaires de E dans F est un sous-espace vectoriel de F^E. On pose $\mathscr{L}(E) := \mathscr{L}(E,E)$ et $E^* := \mathscr{L}(E,\mathbb{K})$.

(b) Soit $f \in \mathscr{L}(E,F)$ bijective. On a : $f^{-1} \in \mathscr{L}(F,E)$.

(c) Soient $f \in \mathscr{L}(E,F)$ et $g \in \mathscr{L}(F,G)$. On a : $g \circ f \in \mathscr{L}(E,G)$.

(d) Soient $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{K}$, $f_1, f_2 \in \mathscr{L}(E,F)$ et $g_1, g_2 \in \mathscr{L}(F,G)$.
On a : $(\beta_1 g_1 + \beta_2 g_2) \circ (\alpha_1 f_1 + \alpha_2 f_2) = \beta_1\alpha_1(g_1 \circ f_1) + \beta_1\alpha_2(g_1 \circ f_2) + \beta_2\alpha_1(g_2 \circ f_1) + \beta_2\alpha_2(g_2 \circ f_2)$.

Démonstration

Les points (a), (c), (d) sont « immédiats ».

(b) Soient $\alpha, \beta \in \mathbb{K}$ et $v, w \in F$. On a, après avoir tenté un raisonnement par équivalences :

$$\begin{aligned} f^{-1}(\alpha\, v + \beta\, w) &= f^{-1}\left(\alpha\, f\big(f^{-1}(v)\big) + \beta\, f\big(f^{-1}(w)\big)\right) \\ &= f^{-1}\left(f\big(\alpha\, f^{-1}(v) + \beta\, f^{-1}(w)\big)\right) && \text{(car } f \text{ est linéaire)} \\ &= \alpha\, f^{-1}(v) + \beta\, f^{-1}(w) \end{aligned}$$

L'application f^{-1} est donc linéaire. □

Proposition

(a) On suppose que E est de dimension finie, de base $(v_1, ..., v_n)$.
Pour tous $w_1, ..., w_n \in F$, il existe $f \in \mathscr{L}(E,F)$ unique tel que :

$$f(v_1) = w_1 \ \text{ et ... et } \ f(v_n) = w_n.$$

Dans ce cas : f est bijective si et seulement si $(f(v_1), ..., f(v_n))$ est une base de F.

(b) On suppose que E ou F est de dimension finie.
Les espaces vectoriels E et F sont isomorphes si et seulement si E et F sont tous deux de dimension finie et $\dim E = \dim F$.
Ainsi, pour chaque $n \in \mathbb{N}$, les $\mathbb{K}$-espaces vectoriels de dimension n sont isomorphes à $\mathbb{K}^n$.

Démonstration

(a) Soient $w_1, ..., w_n \in F$. La seule application qui puisse convenir est

$$f : \begin{array}{ccc} E & \longrightarrow & F \\ \underset{\in\mathbb{K}}{x_1} v_1 + \cdots + \underset{\in\mathbb{K}}{x_n} v_n & \longmapsto & x_1 w_1 + \cdots + x_n w_n \end{array}.$$

On constate facilement que cette application f est linéaire.

On vérifie maintenant que f est bijective si et seulement si $(w_1, ..., w_n)$ est une base de F.

($\Rightarrow$) On suppose f bijective.
Pour tous $w \in F$ et $y_1, ..., y_n \in \mathbb{K}$, on a par linéarité de f^{-1} :
$w = y_1 w_1 + ... + y_n w_n \Leftrightarrow f^{-1}(w) = y_1 f^{-1}(w_1) + ... + y_n f^{-1}(w_n) \Leftrightarrow f^{-1}(w) = y_1 v_1 + ... + y_n v_n$.
L'unicité des composantes de w suivant $(w_1, ..., w_n)$ montre que $(w_1, ..., w_n)$ est une base de F.

($\Leftarrow$) On suppose que $(w_1, ..., w_n)$ est une base de F.
On introduit l'unique $g \in \mathscr{L}(F,E)$ unique tel que : $g(w_1) = v_1$ et ... et $g(w_n) = v_n$.
Donc au vu de la description de f donnée au début de la démonstration :
$g \circ f : \underset{\in\mathbb{K}}{x_1} v_1 + \cdots + \underset{\in\mathbb{K}}{x_n} v_n \mapsto x_1 v_1 + \cdots + x_n v_n$ et $f \circ g : \underset{\in\mathbb{K}}{y_1} w_1 + \cdots + \underset{\in\mathbb{K}}{y_n} w_n \mapsto y_1 w_1 + \cdots + y_n w_n$
puis $g \circ f = \mathrm{id}_E$ et $f \circ g = \mathrm{id}_F$, et par conséquent f est bijective.

(b) ($\Rightarrow$) On suppose qu'il existe un isomorphisme $f : E \to F$. On se place dans le cas où E est de dimension finie (le cas où F est de dimension finie s'en déduira en utilisant f^{-1} et échangeant les rôles de E et F). Soit $(v_1, ..., v_n)$ une base de E. D'après (a), la famille $(f(v_1), ..., f(v_n))$ est une base de F. D'où : F est de dimension finie et $\dim F = n = \dim E$.

($\Leftarrow$) On suppose que E et F sont tous deux de dimension finie avec une même dimension, notée n. Soient $(v_1, ..., v_n)$ une base de E et $(w_1, ..., w_n)$ une base de F. D'après (a), l'unique $f : E \to F$ linéaire tel que $f(v_1) = w_1$ et ... et $f(v_n) = w_n$ est un isomorphisme de E sur F.

La dernière phrase de l'énoncé découle de la dernière implication ci-dessus. □

Remarque

On suppose que E est de dimension finie, de base $\mathscr{B} = (v_1, ..., v_n)$, et note $(e_1, ..., e_n)$ la base canonique de $\mathbb{K}^n$. D'après la proposition (a), l'unique $f \in \mathscr{L}(E, \mathbb{K}^n)$ tel que $f(v_1) = e_1$ et ... et $f(v_n) = e_n$ est un isomorphisme de E sur $\mathbb{K}^n$. Par linéarité, f envoie $\underset{/\mathscr{B}}{v}\Big| X$ sur X.

Exemple (« changement de base »)

Les vecteurs $v_1 := \left(\begin{smallmatrix}3\\2\end{smallmatrix}\right)$ et $v_2 := \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ de $\mathbb{R}^2$ sont clairement non-colinéaires. Donc (v_1, v_2) est une base de $\mathbb{R}^2$. Soit $f \in \mathscr{L}(\mathbb{R}^2)$ déterminé par : $f(v_1) = 2v_2$ et $f(v_2) = v_1 + v_2$.
On cherche à calculer $f(x, y)$ pour $x, y \in \mathbb{R}$.

On exprime les vecteurs de la base canonique (e_1, e_2) de $\mathbb{R}^2$ à l'aide de (v_1, v_2).

On a : $\begin{cases} 3e_1 + 2e_2 = v_1 \\ \quad e_1 + e_2 = v_2 \end{cases}$. Présentation « à la Gauss », qui sera utile en L2 dans le cas triangulaire :

$$\begin{array}{cc} e_1 & e_2 \end{array} \qquad\qquad\qquad\qquad \begin{array}{cc} e_1 & e_2 \end{array}$$
$$\left(\begin{array}{cc|c} 3 & 2 & v_1 \\ 1 & 1 & v_2 \end{array}\right) \xrightarrow[L_2 \leftarrow 3L_2 - L_1]{} \left(\begin{array}{cc|c} 3 & 2 & v_1 \\ 0 & 1 & -v_1 + 3v_2 \end{array}\right) \quad \text{donc} \quad \uparrow \begin{cases} e_1 = \frac{1}{3}\big(-2(-v_1 + 3v_2) + v_1\big) = v_1 - 2v_2 \\ e_2 = -v_1 + 3v_2 \end{cases}$$

puis $\begin{cases} f(e_1) = \quad f(v_1) - 2f(v_2) \quad = \quad -2v_1 \quad = \left(\begin{smallmatrix}-6\\-4\end{smallmatrix}\right) \\ f(e_2) = -f(v_1) + 3f(v_2) = 3v_1 + v_2 = \left(\begin{smallmatrix}10\\7\end{smallmatrix}\right) \end{cases}$.

D'où $f\big(\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)\big) = f(xe_1 + ye_2) = xf(e_1) + yf(e_2) = \begin{pmatrix} -6x+10y \\ -4x+7y \end{pmatrix}$.

2. Noyau et image

Lemme

Soit $f \in \mathscr{L}(E, F)$.

(a) L'image réciproque $f^{-1}(W) := \{v \in E \mid f(v) \in W\}$ d'un sous-espace vectoriel W de F par f, est un sous-espace vectoriel de E.

(b) L'image directe $f(V) := \underbrace{\{f(v)\ ;\ v \in V\}}_{\text{notation pour } \{w \in F \mid \exists v \in V \;\; w = f(v)\}}$ d'un sous-espace vectoriel V de E par f, est un sous-espace vectoriel de F.

Démonstration

(a) Soit W un sous-espace vectoriel de F.

(i) On a : $f(0_E) = 0_F$ (car f est linéaire) avec $0_F \in W$, donc $0_E \in f^{-1}(W)$.
(ii) Soient $\alpha, \alpha' \in \mathbb{K}$ et $v, v' \in f^{-1}(W)$. On a donc : $f(v), f(v') \in W$.
Par linéarité de f on a ensuite : $f(\alpha v + \alpha' v') = \alpha f(v) + \alpha' f(v')$ avec $\alpha f(v) + \alpha' f(v') \in W$. D'où : $\alpha v + \alpha' v' \in f^{-1}(W)$.

Ainsi $f^{-1}(W)$ est un sous-espace vectoriel de E.

(b) Soit V un sous-espace vectoriel de E.

(i) On a : $0_F = f(0_E)$ avec $0_E \in V$, donc $0_F \in f(V)$.
(ii) Soient $\alpha, \alpha' \in \mathbb{K}$ et $w, w' \in f(V)$. Il existe $v, v' \in V$ tels que $w = f(v)$ et $w' = f(v')$.
Par linéarité de f on a : $\alpha w + \alpha' w' = \alpha f(v) + \alpha' f(v') = f(\alpha v + \alpha' v')$ avec $\alpha v + \alpha' v' \in V$. D'où : $\alpha w + \alpha' w' \in f(V)$.

Par conséquent $f(V)$ est un sous-espace vectoriel de F. □

Définition-Proposition

Soit $f \in \mathscr{L}(E, F)$.

(a) *Le noyau de* f, noté $\operatorname{Ker} f$, est le sous-espace vectoriel $f^{-1}(\{0\})$ de E :

$$\operatorname{Ker} f := \{v \in E \mid f(v) = 0\}. \qquad \longleftarrow \text{[équation cartésienne]}$$

(b) On a : f est injective si et seulement si $\operatorname{Ker} f = \{0\}$.

(c) *L'image de l'application* f, notée $\operatorname{Im} f$, est le sous-espace vectoriel $f(E)$ de F :

$$\operatorname{Im} f := \{w \in F \mid \exists v \in E \;\; f(v) = w\}. \qquad \longleftarrow \text{[équation paramétrique]}$$

(d) On a : f est surjective si et seulement si $\operatorname{Im} f = F$.

(e) Si $E = \operatorname{Vect}(v_1, ..., v_p)$ avec $v_1, ..., v_p \in E$, alors $\operatorname{Im} f = \operatorname{Vect}\big(f(v_1), ..., f(v_p)\big)$.

Démonstration

(a) D'après le lemme (a), $\operatorname{Ker} f$ est un sous-espace vectoriel de E.

(b) ($\Rightarrow$) On suppose que f est injective. Si $v \in \operatorname{Ker} f$, alors $f(v) = 0_F = f(0_E)$ alors, par injectivité de f, $v = 0_E$. Ainsi $\operatorname{Ker} f \subseteq \{0\}$. Comme $\{0\} \subseteq \operatorname{Ker} f$ d'après (a) : $\operatorname{Ker} f = \{0\}$.

($\Leftarrow$) On suppose que $\operatorname{Ker} f = \{0\}$. Si $v, v' \in E$ vérifient $f(v) = f(v')$, alors $f(v - v') = f(v) - f(v') = 0$, alors $v - v' \in \operatorname{Ker} f$, alors vu l'hypothèse $v = v'$. Ainsi f est injective.

(c) D'après le lemme (b), $\operatorname{Im} f$ est un sous-espace vectoriel de F.

(d) Découle des définitions des notions de « surjectivité » et « image d'une application ».

(e) On suppose que $v_1, ..., v_p \in E$ vérifient : $E = \operatorname{Vect}(v_1, ..., v_p)$. On a :

$$\begin{aligned} f(E) &= f\big(\{\alpha_1 v_1 + \cdots + \alpha_p v_p \ ; \ \alpha_1, ..., \alpha_p \in \mathbb{K}\}\big) = \{f\big(\alpha_1 v_1 + \cdots + \alpha_p v_p\big) \ ; \ \alpha_1, ..., \alpha_p \in \mathbb{K}\} \\ &= \big\{\alpha_1 f(v_1) + \cdots + \alpha_p f(v_p) \ ; \ \alpha_1, ..., \alpha_p \in \mathbb{K}\big\} = \operatorname{Vect}\big(f(v_1), ..., f(v_p)\big). \end{aligned}$$

D'où le résultat. □

Exemple (en « dimension infinie »)

On note ici $\mathcal{P}ol(\mathbb{K}, \mathbb{K})$ le sous-espace vectoriel $\operatorname{Vect}\big(\{f_n : x \mapsto x^n \ ; \ n \in \mathbb{N}\}\big)$ de $\mathbb{K}^{\mathbb{K}}$ formé des $p : \mathbb{K} \to \mathbb{K}$ de la forme $p : x \mapsto a_n x^n + \cdots + a_1 x + a_0$ pour certains $n \in \mathbb{N}$ et $a_0, ..., a_n \in \mathbb{K}$.

On considère $f : \begin{array}{ccc} \mathbb{K}[X] & \longrightarrow & \mathcal{P}ol(\mathbb{K}, \mathbb{K}) \\ P = a_n X^n + \cdots + a_1 X + a_0 & \longmapsto & (p : x \mapsto a_n x^n + \cdots + a_1 x + a_0) \end{array}$.

L'application f est linéaire (facile), et surjective par définition de $\mathcal{P}ol(\mathbb{K}, \mathbb{K})$.
On utilise le (b) de la définition-proposition pour montrer que f est injective.
Soit $P = a_n X^n + \cdots + a_1 X + a_0 \in \mathbb{K}[X]$ tel que l'application $p := f(P)$ est nulle.
Comme $\mathbb{R} \subseteq \mathbb{K}$, on a en particulier : $\forall x \in \mathbb{R} \quad a_n x^n + \cdots + a_1 x + a_0 = 0$,
donc $a_0 = p_{|\mathbb{R}}(0) = 0$, $a_1 = p_{|\mathbb{R}}{}'(0) = 0$, ..., $n!\, a_n = p_{|\mathbb{R}}{}^{n}(0) = 0$ et a fortiori $P = 0$.
Par conséquent $\operatorname{Ker} f = \{0\}$, ce qui prouve que f est injective.
Ainsi f est un isomorphisme de $\mathbb{K}$-espaces vectoriels.

Définition-Proposition

On suppose que E ou F est de dimension finie. Soit $f \in \mathscr{L}(E, F)$.

(a) L'espace vectoriel $\operatorname{Im} f$ est ici de dimension finie.
On note : $\operatorname{rg} f := \dim(\operatorname{Im} f)$ « rang de f ».

(b) Si $E = \operatorname{Vect}(v_1, ..., v_p)$ avec $v_1, ..., v_p \in E$, alors $\operatorname{rg} f = \operatorname{rg}\big(f(v_1), ..., f(v_p)\big)$.

Démonstration

(a) Si E est de dimension finie, on va voir dans la démonstration du (b) que $\operatorname{Im} f$ a une famille génératrice finie dont on pourra extraire une base finie .
Si F est de dimension finie, l'espace vectoriel $\operatorname{Im} f$ qui en est un sous-espace vectoriel est aussi de dimension finie.

(b) D'après le (e) de la définition-proposition qui précède : $\operatorname{Im} f = \operatorname{Vect}(f(v_1), ..., f(v_p))$.
D'où : $\operatorname{rg} f = \dim(\operatorname{Im} f) = \dim(\operatorname{Vect}(f(v_1), ..., f(v_p))) = \operatorname{rg}(f(v_1), ..., f(v_p))$. □

Exemple (exercice type)

Soit $f : \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} & \longmapsto & \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \end{array}$ définie par $\begin{cases} x' = -4x + 12y - 5z \\ y' = x - 3y + 2z \\ z' = 2x - 6y + z \end{cases}$.

(a) Équation cartésienne de $\operatorname{Im} f$ et base de $\operatorname{Im} f$?

Un vecteur $w = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3$ appartient à $\operatorname{Im} f$ si et seulement si l'équation $(\star)$ $f(v) = w$ d'inconnue $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ a au moins une solution. On étudie $(\star)$ par la méthode de Gauss :

$$\begin{array}{ccc|c} x & y & z & \\ \end{array}
\left(\begin{array}{ccc|c} -4 & 12 & -5 & a \\ 1 & -3 & 2 & b \\ 2 & -6 & 1 & c \end{array}\right)
\xrightarrow[\substack{L_2 \leftarrow 4L_2+L_1 \\ L_3 \leftarrow 2L_3+L_1}]{}
\begin{array}{ccc|c} x & y & z & \\ \end{array}
\left(\begin{array}{ccc|c} -4 & 12 & -5 & a \\ 0 & 0 & 3 & a+4b \\ 0 & 0 & -3 & a+2c \end{array}\right)$$

$$\xrightarrow[C_2 \leftrightarrow C_3]{}
\begin{array}{ccc|c} x & z & y & \\ \end{array}
\left(\begin{array}{ccc|c} -4 & -5 & 12 & a \\ 0 & 3 & 0 & a+4b \\ 0 & -3 & 0 & a+2c \end{array}\right)
\xrightarrow[L_3 \leftarrow L_3+L_2]{}
\begin{array}{ccc|c} x & z & y & \\ \end{array}
\left(\begin{array}{ccc|c} -4 & -5 & 12 & a \\ 0 & 3 & 0 & a+4b \\ 0 & 0 & 0 & 2a+4b+2c \end{array}\right)$$

donc : $w \in \operatorname{Im} f \iff 2a+4b+2c=0 \iff a+2b+c=0$.
Ainsi, dans les coordonnées habituelles : $\operatorname{Im} f$ a pour équation cartésienne $x+2y+z=0$.

En ne considérant que les colonnes sous x, y, z d'une part et sous x, z d'autre part, le calcul précédent donne aussi : $\dim(\operatorname{Im} f) = \operatorname{rg}(f(e_1), f(e_2), f(e_3)) = 2$ et $\operatorname{rg}(f(e_1), f(e_3)) = 2$, donc $(f(e_1), f(e_3))$ engendre $\operatorname{Im} f$ qui est de dimension 2, puis : $\operatorname{Im} f$ a pour base $\left(\begin{pmatrix}-4\\1\\2\end{pmatrix}, \begin{pmatrix}-5\\2\\1\end{pmatrix}\right)$.

(b) Équation cartésienne de $\operatorname{Ker} f$ avec deux égalités (entre réels) et base de $\operatorname{Ker} f$?

Un vecteur $v = \begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3$ appartient à $\operatorname{Ker} f$ si et seulement si $(\star\star)$ $f(v) = 0$.
La méthode de Gauss pour résoudre $(\star\star)$ est le calcul du (a) avec un second membre nul.

Donc : $\operatorname{Ker} f \ : \ \begin{cases} -4x - 5z + 12y = 0 \\ 3z = 0 \end{cases}$.

Par remontée triangulaire, on obtient :

$\operatorname{Ker} f \begin{cases} x = 3t \\ y = t \\ z = 0 \end{cases}$, $t \in \mathbb{R}$. Donc $\operatorname{Ker} f = \operatorname{Vect}\Big(\underbrace{\begin{pmatrix}3\\1\\0\end{pmatrix}}_{\text{vecteur non-nul}}\Big)$ puis $\operatorname{Ker} f$ a pour base $\left(\begin{pmatrix}3\\1\\0\end{pmatrix}\right)$. (∗)

Proposition

Soit $f \in \mathscr{L}(E, F)$. On suppose donné un supplémentaire S de $\operatorname{Ker} f$ dans E (on sait qu'il en existe quand E est de dimension finie).

L'application $\tilde{f} : \begin{array}{l} S \to \operatorname{Im} f \\ x \mapsto f(x) \end{array}$ est linéaire bijective.

(*Variante.* On peut définir une application $\tilde{f} : \begin{array}{l} E/\operatorname{Ker} f \to \operatorname{Im} f \\ \dot{x} \quad\ \mapsto f(x) \end{array}$, qui est linéaire bijective.)

Démonstration

Il est clair que $f(x) \in \operatorname{Im} f$ pour tout $x \in S$, et ensuite que $\tilde{f}$ est linéaire.

On a : $\operatorname{Im} \tilde{f} = \tilde{f}(S) = f(S) = f(S + \operatorname{Ker} f) = f(E) = \operatorname{Im} f$
et $\operatorname{Ker} \tilde{f} = \{x \in S \mid f(x) = 0\} = S \cap \operatorname{Ker} f = \{0\}$. Donc $\tilde{f}$ est bijective.

(*Variante.* Si $x, y \in E$ et $\dot{x} = \dot{y}$ alors $x - y \in \operatorname{Ker} f$ puis $f(x) = \underbrace{f(x-y)}_{0} + f(y) = f(y)$.
Cela permet de définir $\tilde{f}$. Elle est clairement linéaire et surjective.
De plus : $\operatorname{Ker} \tilde{f} = \{\dot{x} \in E/\operatorname{Ker} f \mid x \in \operatorname{Ker} f\} = \{\dot{0}\}$. Ainsi $\tilde{f}$ est bijective.) □

Théorème (« *théorème du rang* »)

Soit $f \in \mathscr{L}(E, F)$. On suppose que E est de dimension finie.
On a : $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim \underbrace{E}_{\text{espace de départ}}$ où $\operatorname{Ker} f$ et $\operatorname{Im} f$ sont de dimension finie.

(∗) Voici un point de vue sans systèmes linéaires pour traiter cet exercice type, relatif à $A \in \mathfrak{M}_{n,p}(\mathbb{K})$.

(i) Il existe $P \in GL(n, \mathbb{K})$ tel que $P(A \ \underbrace{I_n}_{\text{ou } Y}) = \begin{pmatrix} D & P_1 \\ 0 & P_2 \end{pmatrix}$, où $D = \begin{pmatrix} d_1 & * & \cdots & * \\ & d_2 & \cdots & * \\ 0 & & \ddots & \vdots \\ & & & d_r \end{pmatrix}$ avec $d_1 \neq 0$, ..., $d_r \neq 0$.
Par conséquent : $\operatorname{Ker} A$ a pour équation cartésienne $DX = 0$, où le nombre de lignes de D est minimal (car ces lignes engendrent $(\operatorname{Ker} A)^\perp$ qui vaut $\operatorname{Im} {}^tD$ puis en forment une base). Par surjectivité de D, on obtient : $\operatorname{Im} A$ a pour équation cartésienne $P_2 Y = 0$, où le nombre de lignes de P_2 est minimal (car ces lignes engendrent $(\operatorname{Im} A)^\perp$ qui vaut $\operatorname{Im} {}^tP_2$ puis en forment une base). Les colonnes des pivots de D fourniraient une base de $\operatorname{Im} A$.

(ii) Il existe aussi $Q \in GL(p, \mathbb{K})$ tel que ${}^tQ({}^tA\, I_p) = \begin{pmatrix} {}^t\widetilde{D} & {}^tQ_1 \\ 0 & {}^tQ_2 \end{pmatrix}$ c-à-d $\begin{pmatrix} A \\ I_p \end{pmatrix} Q = \begin{pmatrix} \widetilde{D} & 0 \\ Q_1 & Q_2 \end{pmatrix}$ où $\widetilde{D}$ a la forme de ${}^tD$.
D'après (i) : $\operatorname{Ker} {}^tA = \operatorname{Ker} {}^t\widetilde{D}$ et $\operatorname{Im} {}^tA = \operatorname{Ker} {}^tQ_2$. Leurs orthogonaux $\operatorname{Im} A$ et $\operatorname{Ker} A$ ont pour bases les colonnes de $\widetilde{D}$ et celles de Q_2. En pratique, on pourra utiliser le (ii) pour obtenir une base de $\operatorname{Ker} A$ qui manquait au (i), en travaillant sur les colonnes de $\begin{pmatrix} D \\ I_p \end{pmatrix}$ car D même noyau que A (cf. la décomposition $PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$).

Démonstration

Il existe – et on fixe – un supplémentaire S de $\operatorname{Ker} f$ dans E.

On a d'une part $\dim E = \dim \operatorname{Ker} f + \dim S$ car $E = \operatorname{Ker} f \oplus S$, et d'autre part $\operatorname{Im} f$ est de dimension finie avec $\dim S = \dim \operatorname{Im} f$ d'après la proposition. D'où le résultat.

(*Variante.* D'après la proposition, on a : $\dim \operatorname{Im} f = \dim(E/\operatorname{Ker} f) = \dim E - \dim \operatorname{Ker} f$.) □

Exemple

L'application linéaire $f : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x,y) & \mapsto & (0,x) \end{array}$ vérifie : $\operatorname{Ker} f = \operatorname{Im} f = (Oy)$.

On a bien : $\underbrace{\dim \operatorname{Ker} f}_{1} + \underbrace{\dim \operatorname{Im} f}_{1} = \underbrace{\dim \mathbb{R}^2}_{2}$.

Mais $\operatorname{Ker} f$ et $\operatorname{Im} f$ ne sont pas supplémentaires dans $\mathbb{R}^2$!

Corollaire (important)

Soit $f \in \mathscr{L}(E)$. On suppose que E est de dimension finie.

Les propriétés suivantes sont équivalentes :

(i) f est injective ; (ii) f est surjective ; (iii) f est bijective.

Démonstration

On vient de voir que : $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim E$.

$$\begin{array}{lll} \text{Donc : (i)} & \iff \operatorname{Ker} f = \{0\} & \iff \dim \operatorname{Ker} f = 0 \\ & \iff \dim \operatorname{Im} f = \dim E & \iff \operatorname{Im} f = E \\ & \iff \text{(ii)}. & \end{array}$$

On en déduit que : (i) $\iff$ ((i) et (ii)) $\iff$ (iii). □

Remarque

Le corollaire devient faux en enlevant l'hypothèse « E est de dimension finie ».

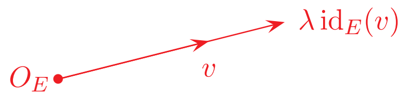
Par exemple, on peut montrer que la dérivation $\delta : \mathbb{K}[X] \to \mathbb{K}[X]$ définie dans le paragraphe 1, est surjective mais n'est pas injective.

3. Endomorphismes particuliers

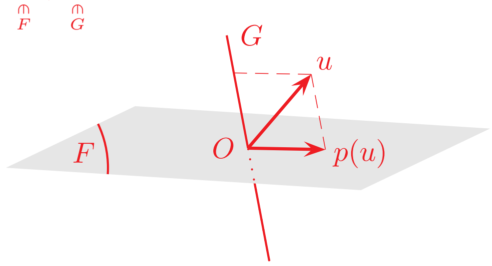
Définition

Soient $\lambda \in \mathbb{K}$, et, F et G deux sous-espaces vectoriels supplémentaires dans E.

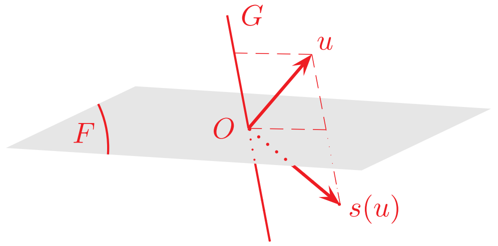
(a) L'application $\lambda\,\mathrm{id}_E : \begin{array}{ccc} E & \to & E \\ v & \mapsto & \lambda v \end{array}$ s'appelle *l'homothétie (vectorielle) de E de rapport λ.*



(b) L'application $p : \begin{array}{ccc} E = F \oplus G & \to & E \\ u = \underset{F}{\underset{\cap}{v}} + \underset{G}{\underset{\cap}{w}} & \mapsto & v \end{array}$ s'appelle *la projection de E sur F parallèlement à G.*



(c) L'application $\begin{array}{rccl} s: & E = F \oplus G & \longrightarrow & E \\ & u = \underset{F}{\underset{\cap}{v}} + \underset{G}{\underset{\cap}{w}} & \longmapsto & v - w \end{array}$ s'appelle *la symétrie de E par rapport à F parallèlement à G.*



Proposition

On utilise les notations de la définition précédente.

(a) L'homothétie $\lambda\,\mathrm{id}_E$ est un endomorphisme de E.
Quand $\lambda \neq 0$, elle est bijective de réciproque $\lambda^{-1}\,\mathrm{id}_E$.

(b) La projection p est un endomorphisme de E vérifiant $p \circ p = p$, et tel que :

$$F = \operatorname{Im} p \quad \text{et} \quad G = \operatorname{Ker} p.$$

Réciproquement, pour tout $f \in \mathscr{L}(E)$ vérifiant $f \circ f = f$, on a : $E = \operatorname{Im} f \oplus \operatorname{Ker} f$ et f est la projection de E sur $\operatorname{Im} f$ parallèlement à $\operatorname{Ker} f$.

(c) La symétrie s est un automorphisme de E vérifiant $s \circ s = \mathrm{id}_E$, et tel que :

$$F = \{u \in E \mid s(u) = u\} \quad \text{et} \quad G = \{u \in E \mid s(u) = -u\}.$$

Réciproquement, pour tout $f \in \mathscr{L}(E)$ vérifiant $f \circ f = \mathrm{id}_E$, on a : $E = \operatorname{Ker}(f - \mathrm{id}_E) \oplus \operatorname{Ker}(f + \mathrm{id}_E)$ et f est la symétrie de E par rapport à $\operatorname{Ker}(f - \mathrm{id}_E)$ parallèlement à $\operatorname{Ker}(f + \mathrm{id}_E)$.

Démonstration

Exercice. □

Exemple

On reprend les vecteurs $v_1 = \binom{3}{2}$ et $v_2 = \binom{1}{1}$ de $\mathbb{R}^2$ utilisés à la fin du paragraphe 1.
On choisit $E := \mathbb{R}^2$, $F := \operatorname{Vect} v_1$ et $G := \operatorname{Vect} v_2$ dans la définition précédente.
On a bien $E = F \oplus G$ car (v_1, v_2) est une base de $\mathbb{R}^2$.

Par définition de p et de s, on a : $\begin{cases} p(v_1) = v_1 \\ p(v_2) = 0 \end{cases}$ et $\begin{cases} s(v_1) = v_1 \\ s(v_2) = -v_2 \end{cases}$.

Par ailleurs, on a vu que : $\begin{cases} e_1 = v_1 - 2v_2 \\ e_2 = -v_1 + 3v_2 \end{cases}$. D'où $p\left(\binom{x}{y}\right) = \binom{3x-3y}{2x-2y}$ et $s\left(\binom{x}{y}\right) = \binom{5x-6y}{4x-5y}$.

II. Utilisation des matrices

Dans cette partie, on fixe trois $\mathbb{K}$-espaces vectoriels E, F et G de dimension finie.
On se donne des bases $\mathscr{B} = (v_1, \ldots, v_p)$ de E, $\mathscr{C} = (w_1, \ldots, w_n)$ de F et $\mathscr{D}$ de G.

1. Matrice d'une application linéaire

Définition-Proposition

Soit $f \in \mathscr{L}(E, F)$.

(a) *La matrice de f dans les bases $\mathscr{B}$ et $\mathscr{C}$* est : $\mathfrak{Mat}_{\mathscr{B},\mathscr{C}} f := \begin{array}{c} \begin{array}{ccc} f(v_1)/\mathscr{C} & \cdots & f(v_p)/\mathscr{C} \end{array} \\ \left(\begin{array}{ccc} \Box & \cdots & \Box \end{array} \right) \end{array}$.

Il s'agit de l'unique matrice $A \in \mathfrak{M}_{n,p}(\mathbb{K})$ telle que :

$$f : v_{/\mathscr{B}} \Big| X \longmapsto w_{/\mathscr{C}} \Big| Y \quad \text{où} \quad \boxed{Y = AX}.$$

(b) Quand $E = F$ et $\mathscr{B} = \mathscr{C}$, on note : $\mathfrak{Mat}_{\mathscr{B}} f := \mathfrak{Mat}_{\mathscr{B},\mathscr{B}} f$.

Démonstration

Soit $A \in \mathfrak{M}_{n,p}(\mathbb{K})$. L'application $g : v_{/\mathscr{B}} \Big| X \mapsto w_{/\mathscr{C}} \Big| AX$ est linéaire (facile).

Les vecteurs $g(v_1)$, ..., $g(v_p)$ ont pour coordonnées $A\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$, ..., $A\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}$ dans $\mathscr{C}$.

Les colonnes de A sont donc $g(v_1)/\mathscr{C}$, ..., $g(v_p)/\mathscr{C}$.

Si $f = g$, alors vu ce qui précède $A = \mathfrak{Mat}_{\mathscr{B},\mathscr{C}} f$.

Si $A = \mathfrak{Mat}_{\mathscr{B},\mathscr{C}} f$, alors f coïncide avec g en chacun des vecteurs $v_1, ..., v_p$, puis $f = g$. □

Exemples

(1) On a, par la définition du (a) : $\mathfrak{Mat}_{\mathscr{B}}\, \mathrm{id}_E = \begin{pmatrix}1 & & 0\\ & \ddots & \\ 0 & & 1\end{pmatrix} = I_p$ où $p = \dim E$.

(2) Soit $A \in \mathfrak{M}_{n,p}(\mathbb{K})$. On sait que l'application $f : \begin{array}{rcl}\mathbb{K}^p & \to & \mathbb{K}^n\\ X & \mapsto & AX\end{array}$ est linéaire.

Par la caractérisation du (a), la matrice de f dans les bases canoniques est égale à A.

(3) Soit $\theta \in \mathbb{R}$. La rotation $r : \begin{array}{rcl}\mathbb{R}^2 & \to & \mathbb{R}^2\\ \binom{x}{y} & \mapsto & \binom{x'}{y'}\end{array}$ déterminée par $\begin{cases} x' = (\cos\theta)\, x - (\sin\theta)\, y \\ y' = (\sin\theta)\, x + (\cos\theta)\, y\end{cases}$

s'écrit $\binom{x}{y} \mapsto \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\binom{x}{y}$. On a donc : $\mathfrak{Mat}_{\text{base canonique de } \mathbb{R}^2}\, r = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$.

(4) On considère l'application $\widetilde{f} : \begin{array}{rcl}\mathbb{R}_2[X] & \to & \mathbb{R}[X]\\ P & \mapsto & (X+1)P' - P\end{array}$ qui est clairement linéaire.

On cherche, à l'aide d'un calcul matriciel, une base de $\mathrm{Ker}\, \widetilde{f}$ et une base de $\mathrm{Im}\, \widetilde{f}$.
On a : $\widetilde{f}(1) = -1$, $\widetilde{f}(X) = 1$ et $\widetilde{f}(X^2) = X^2 + 2X$.
Par conséquent : $\mathrm{Im}\, \widetilde{f} = \mathrm{Vect}(\widetilde{f}(1), \widetilde{f}(X), \widetilde{f}(X^2)) \subseteq \mathbb{R}_2[X]$.
Cela permet de définir l'application $f : \begin{array}{rcl}\mathbb{R}_2[X] & \to & \mathbb{R}_2[X]\\ P & \mapsto & (X+1)P' - P\end{array}$ qui est encore linéaire.

On a : $\mathrm{Ker}\, \widetilde{f} = \mathrm{Ker}\, f$ et $\mathrm{Im}\, \widetilde{f} = \mathrm{Im}\, f$. On recherche des bases de $\mathrm{Ker}\, f$ et $\mathrm{Im}\, f$ en coordonnées.

Par définition, la matrice $\mathfrak{Mat}_{(1,X,X^2)} f$ est : $A := \begin{pmatrix}-1 & 1 & 0\\ 0 & 0 & 2\\ 0 & 0 & 1\end{pmatrix}$.

On prend en compte l'autre caractérisation donnée pour $\mathfrak{Mat}_{(1,X,X^2)} f$:
$\mathrm{Ker}\, f = \{P \in \mathbb{R}^2[X] \mid f(P) = 0\} = \left\{ P_{/(1,X,X^2)} \Bigg| \begin{matrix}a_0\\a_1\\a_2\end{matrix} \in \mathbb{R}^2[X] \;\Big|\; A\begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}\right\}$.

On résout l'équation $A\begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$ d'inconnue $\begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix} \in \mathbb{R}^3$ par la méthode de Gauss :

$$\begin{array}{ccc} a_0 & a_1 & a_2 \end{array}\begin{pmatrix}-1 & 1 & 0\\ 0 & 0 & 2\\ 0 & 0 & 1\end{pmatrix} \xrightarrow[C_2 \leftrightarrow C_3]{} \begin{array}{ccc} a_0 & a_2 & a_1 \end{array}\begin{pmatrix}-1 & 0 & 1\\ 0 & 2 & 0\\ 0 & 1 & 0\end{pmatrix} \xrightarrow[L_3 \leftarrow 2L_3 - L_2]{} \begin{array}{ccc} a_0 & a_2 & a_1 \end{array}\begin{pmatrix}-1 & 0 & 1\\ 0 & 2 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

Donc $\mathrm{Ker}\, f$: $\begin{cases} a_0 = t\\ a_1 = t\\ a_2 = 0\end{cases}$, $t \in \mathbb{R}$, puis $\mathrm{Ker}\, f = \{t(X+1)\ ;\ t \in \mathbb{R}\}$ et $\mathrm{Ker}\, f$ a pour base $(X+1)$.

Ce calcul par Gauss montre aussi que $\mathrm{rg}\,(f(1), f(X), f(X^2)) = \mathrm{rg}\,(f(1), f(X^2)) = 2$.
Il en résulte que $\mathrm{Im}\, f$ a pour base $(f(1), f(X^2))$, c'est-à-dire $(-1, X^2 + 2X)$.

Définition-Proposition

(a) L'application $\begin{array}{rcl}\mathscr{L}(E,F) & \longrightarrow & \mathfrak{M}_{n,p}(\mathbb{K})\\ f & \longmapsto & \mathfrak{Mat}_{\mathscr{B},\mathscr{C}} f\end{array}$ est un isomorphisme de $\mathbb{K}$-espaces vectoriels.

Donc $\mathscr{L}(E,F)$ est de dimension finie et $\dim \mathscr{L}(E,F) = (\dim E)(\dim F)$.

Le $\mathbb{K}$-espace vectoriel E^* des formes linéaires sur E a d'ailleurs une base $\mathscr{B}^* = (v_1^*, ..., v_p^*)$, appelée *base duale de* $\mathscr{B}$, qui est déterminée par : $v_i^*(v_i) = 1$ et $v_i^*(v_j) = 0$ quand $j \neq i$.

(b) Soient $f \in \mathscr{L}(E,F)$ et $g \in \mathscr{L}(F,G)$.
On a : $\mathfrak{Mat}_{\mathscr{B},\mathscr{D}}(g \circ f) = \left(\mathfrak{Mat}_{\mathscr{C},\mathscr{D}}\, g\right)\left(\mathfrak{Mat}_{\mathscr{B},\mathscr{C}}\, f\right)$.

(c) Soit $f \in \mathscr{L}(E,F)$. On pose $A := \mathfrak{Mat}_{\mathscr{B},\mathscr{C}}\, f \in \mathfrak{M}_{n,p}(\mathbb{K})$.

L'application ${}^t\!f : \begin{array}{ccc} F^* & \longrightarrow & E^* \\ \varphi & \longmapsto & \varphi \circ f \end{array}$ est linéaire. De plus : $\mathfrak{Mat}_{\mathscr{C}^*,\mathscr{B}^*}({}^t\!f) = {}^t\!A$.

Démonstration

(a) Il est clair que l'application $\mathfrak{Mat}_{\mathscr{B},\mathscr{C}} : \begin{array}{ccc} \mathscr{L}(E,F) & \longrightarrow & \mathfrak{M}_{n,p}(\mathbb{K}) \\ f & \longmapsto & \mathfrak{Mat}_{\mathscr{B},\mathscr{C}} f \end{array}$ est linéaire.

Soit $A = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ & \cdots & \\ a_{n1} & \cdots & a_{np} \end{pmatrix} \in \mathfrak{M}_{n,p}(\mathbb{K})$. On note $\widetilde{w_1}, \ldots, \widetilde{w_p} \in F$ tels que $\underset{/\mathscr{C}}{\widetilde{w_1}}\left|\begin{matrix} a_{11} \\ \vdots \\ a_{n1} \end{matrix}\right.$ et ... et $\underset{/\mathscr{C}}{\widetilde{w_1}}\left|\begin{matrix} a_{1p} \\ \vdots \\ a_{np} \end{matrix}\right.$.
D'après une proposition du II 1, il existe $f \in \mathscr{L}(E,F)$ unique tel que $f(v_1) = \widetilde{w_1}$ et ... et $f(v_p) = \widetilde{w_p}$, c'est-à-dire $\mathfrak{Mat}_{\mathscr{B},\mathscr{C}} f = A$. Ainsi, l'application $\mathfrak{Mat}_{\mathscr{B},\mathscr{C}}$ est bijective.

Dans le cas où $F = \mathbb{K}$ et $\mathscr{C} = (1)$, les images des vecteurs $v_1^*, \ldots, v_p^*$ sont les vecteurs $E_{1,1}, \ldots, E_{1,p}$ de $\mathfrak{M}_{1,p}(\mathbb{K})$. Cela prouve que $\mathscr{B}^*$ est une base de E^*.

(b) On pose $A := \mathfrak{Mat}_{\mathscr{B},\mathscr{C}}\, f$ et $B := \mathfrak{Mat}_{\mathscr{C},\mathscr{D}}\, g$.
On obtient la décomposition $g \circ f : \underset{/\mathscr{B}}{u}\Big|X \xmapsto{f} \underset{/\mathscr{C}}{v}\Big|Y \xmapsto{g} \underset{/\mathscr{D}}{w}\Big|Z$ avec $Z = BY$ et $Y = AX$.
Donc $Z = (BA)X$, ce qui prouve que : $\mathfrak{Mat}_{\mathscr{B},\mathscr{D}}(g \circ f) = BA$.

(c) Il est clair que $\varphi \circ f \in E^*$ pour tout $\varphi \in F^*$, et ensuite que ${}^t\!f$ est linéaire.
Le coefficient de $\mathfrak{Mat}_{\mathscr{C}^*,\mathscr{B}^*}({}^t\!f)$ sur la i^{e} ligne et j^{e} colonne est la i^{e} coordonnée α_i de ${}^t\!f(w_j^*) = \sum\limits_{i=1}^{p} \alpha_i v_i^*$ suivant $(v_1^*, \ldots, v_p^*)$. Or : $\alpha_i = \left({}^t\!f(w_j^*)\right)(v_i) = w_j^*\left(\sum\limits_{k=1}^{n} a_{ki}\, w_k\right) = a_{ji}$.
D'où le résultat. $\square$

2. Utilisation du rang

Proposition

Soit $f \in \mathscr{L}(E,F)$. On pose $A := \mathfrak{Mat}_{\mathscr{B},\mathscr{C}}\, f \in \mathfrak{M}_{n,p}(\mathbb{K})$.

(a) On a : $\operatorname{rg}(A) = \operatorname{rg}(f)$.
D'où : $\left|\begin{array}{l} f \text{ est surjective si et seulement si } \operatorname{rg}(A) = n\,; \\ f \text{ est injective si et seulement si } \operatorname{rg}(A) = p. \end{array}\right.$

(b) L'application f est bijective si et seulement si : $n = p$ et A est inversible.
Dans ce cas : $\mathfrak{Mat}_{\mathscr{C},\mathscr{B}}(f^{-1}) = A^{-1}$.

(c) On a : $\operatorname{rg}({}^t\!f) = \operatorname{rg}(f)$, donc $\operatorname{rg}({}^t\!A) = \operatorname{rg}(A)$ (déjà vu).

Démonstration

((a) On sait calculer le rang d'une matrice et d'une famille de vecteurs par la méthode de Gauss : $\operatorname{rg} A = r = \operatorname{rg}\left(f(v_1), \ldots, f(v_p)\right) = \operatorname{rg} f$.
Pour le reste, on utilise : $\dim F = n$ et $\dim \operatorname{Ker} f = p - \operatorname{rg} f$ (théorème du rang).

b) D'après (a), f est bijective si et seulement si $n = p$ et $\operatorname{rg} A = n$, c'est à dire : $n = p$ et A est inversible. Dans ce cas, on sait que f^{-1} est linéaire. On a : $f^{-1} \circ f = \mathrm{id}_E$ donc $\mathfrak{Mat}_{\mathscr{C},\mathscr{B}}(f^{-1})\, A = I_n$, ce qui donne le résultat en multipliant à droite par A^{-1}.

Variante (sans utiliser le rang)

Pour tout $g \in \mathscr{L}(F,E)$, on a :

$$\begin{aligned} f \text{ bijective et } f^{-1} = g &\iff g \circ f = \mathrm{id}_E \text{ et } f \circ g = \mathrm{id}_F \\ &\iff \underbrace{n = p}_{\text{nécessaire pour que } E \text{ et } F \text{ soient isomorphes}} \text{ et } \mathfrak{Mat}_{\mathscr{B},\mathscr{B}}(g \circ f) = I_n \text{ et } \mathfrak{Mat}_{\mathscr{C},\mathscr{C}}(f \circ g) = I_n \\ &\iff n = p \text{ et } \left(\mathfrak{Mat}_{\mathscr{C},\mathscr{B}}\, g\right) A = I_n \text{ et } A\left(\mathfrak{Mat}_{\mathscr{C},\mathscr{B}}\, g\right) = I_n \\ &\iff A \text{ inversible et } A^{-1} = \mathfrak{Mat}_{\mathscr{C},\mathscr{B}}\, g \end{aligned}$$

Si f est bijective, on sait que f^{-1} est linéaire, et en appliquant ce qui précède à $g := f^{-1}$, on en déduit que A est inversible d'inverse $\mathfrak{Mat}_{\mathscr{C},\mathscr{B}}(f^{-1})$.

Si A est inversible, on sait que son inverse est la matrice dans les bases $\mathscr{C}$ et $\mathscr{B}$ d'une application linéaire g, et on déduit de ce qui précède que f est bijective.

(c) D'après le théorème du rang, on a : $\operatorname{rg}({}^t\!f) = \dim F - \dim \operatorname{Ker}({}^t\!f)$.

Par ailleurs : $\operatorname{Ker}({}^t\!f) = \{\varphi \in F^* \mid \varphi_{|\operatorname{Im} f} = 0\}$. Soit $(u_1, ..., u_r)$ une base de $\operatorname{Im} f$. On la complète en une base $(u_1, ..., u_n)$ de F. On note $(u_1^*, ..., u_n^*)$ sa base duale.

Pour tout $\varphi = \sum\limits_{i=1}^{n} \alpha_i u_i^* \in F^*$, on a : $\varphi_{|\operatorname{Im} f} = 0 \iff \alpha_1 = ... = \alpha_r = 0$.

Ainsi $\operatorname{Ker}({}^t\!f)$ a pour base $(u_{r+1}^*, ..., u_n^*)$, ce qui donne l'égalité $\operatorname{rg}({}^t\!f) = \operatorname{rg}(f)$.

Finalement : $\operatorname{rg}({}^t\!A) = \operatorname{rg}({}^t\!f) = \operatorname{rg}(f) = \operatorname{rg}(A)$. □

Remarque

Soit $A \in \mathfrak{M}(n, \mathbb{K})$. On lui associe l'application linéaire $f : \begin{array}{ccc} \mathbb{K}^n & \to & \mathbb{K}^n \\ X & \mapsto & AX \end{array}$.

Pour montrer que A est inversible et calculer A^{-1}, il suffit de montrer que f est bijective et calculer f^{-1}. Pour cela on peut tenter de résoudre à $Y \in \mathbb{K}^n$ fixé le système $AX = Y$ d'inconnue $X \in \mathbb{K}^n$ et de prouver qu'il a une unique solution qu'on écrira sous la forme $X = BY$ avec $B \in \mathfrak{M}(n, \mathbb{K})$. Dans ce cas, on aura : $\underbrace{A \text{ est inversible}}_{\text{car } f \text{ est bijective}}$ et $A^{-1} = \underbrace{B}_{\text{matrice de } f^{-1} : Y \mapsto BY}$.

3. Changement de base

On se donne deux nouvelles bases, $\mathscr{B}' = (v'_1, ..., v'_p)$ de E et $\mathscr{C}' = (w'_1, ..., w'_n)$ de F.

Définition-Proposition

(a) *La matrice de passage de* $\mathscr{B}$ *à* $\mathscr{B}'$ est : $\begin{array}{c} \begin{matrix} v'_1/\mathscr{B} & \cdots & v'_p/\mathscr{B} \end{matrix} \\ \begin{pmatrix} \Box & \cdots & \Box \end{pmatrix} \end{array}$.

(b) Il s'agit de l'unique matrice $P \in \mathfrak{M}(p, \mathbb{K})$ vérifiant :

$$\text{si} \quad \underset{/\mathscr{B}}{v} \Big| X \in E \quad \text{alors on a} \quad \underset{/\mathscr{B}'}{v} \Big| X' \quad \text{avec} \quad \underbrace{\boxed{\mathrm{X = PX'}}}_{\text{attention !}}.$$

(c) On a : P est inversible et P^{-1} est la matrice de passage de $\mathscr{B}'$ à $\mathscr{B}$.

Démonstration

(b) Comme la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est égale à $\mathfrak{Mat}_{\mathscr{B}',\mathscr{B}} \operatorname{id}_E$, il suffit d'appliquer le résultat analogue du paragraphe 1.

(c) Puisque id_E est bijective de réciproque id_E, la matrice $\mathfrak{Mat}_{\mathscr{B}',\mathscr{B}} \operatorname{id}_E$ est inversible et son inverse est égal à $\mathfrak{Mat}_{\mathscr{B},\mathscr{B}'} \operatorname{id}_E$. Compte tenu de l'égalité du début de cette démonstration, cela signifie que P est inversible et d'inverse la matrice de passage de $\mathscr{B}'$ à $\mathscr{B}$. □

Remarque

Soient $\widetilde{v}_1, ..., \widetilde{v}_p \in E$. On pose $\widetilde{\mathscr{B}} := (\widetilde{v}_1, ..., \widetilde{v}_p)$ et $\widetilde{P} := \begin{array}{c} \begin{matrix} \widetilde{v}_1/\mathscr{B} & \cdots & \widetilde{v}_p/\mathscr{B} \end{matrix} \\ \begin{pmatrix} \Box & \cdots & \Box \end{pmatrix} \end{array}$.

Si on dispose d'une expression explicite des vecteurs v_1, ..., v_p comme combinaisons linéaires de $\widetilde{v}_1$, ..., $\widetilde{v}_p$, alors : $E = \operatorname{Vect}(v_1, ..., v_p) \subseteq \operatorname{Vect}(\widetilde{v}_1, ..., \widetilde{v}_p)$ ce qui donne $E = \operatorname{Vect}(\widetilde{v}_1, ..., \widetilde{v}_p)$, puis $(\widetilde{v}_1, ..., \widetilde{v}_p)$ engendre E et $\operatorname{rg} \widetilde{P} = p$, enfin $\widetilde{\mathscr{B}}$ est une base de E et $\widetilde{P}$ est inversible.

On inverse $\widetilde{P}$ grâce au (c) de la proposition-définition précédente : $\widetilde{P}^{-1} = \begin{array}{c} \begin{matrix} v_1/\widetilde{\mathscr{B}} & \cdots & v_p/\widetilde{\mathscr{B}} \end{matrix} \\ \begin{pmatrix} \Box & \cdots & \Box \end{pmatrix} \end{array}$.

Proposition (« formule de changement de base »)

Soit $f \in \mathscr{L}(E,F)$. On note $A := \mathfrak{Mat}_{\mathscr{B},\mathscr{C}}\, f$ et $A' := \mathfrak{Mat}_{\mathscr{B}',\mathscr{C}'}\, f$.

On a : $\boxed{A' = Q^{-1}AP}$ où $\left| \begin{array}{lll} P & \text{est la matrice de passage de} & \mathscr{B} \text{ à } \mathscr{B}' \,; \\ Q & \quad\quad\quad \text{,,} & \mathscr{C} \text{ à } \mathscr{C}'. \end{array} \right.$

Démonstration

On part de la décomposition suivante, au niveau des applications linéaires :

$$f : \underset{\text{muni de } \mathscr{B}'}{E} \xrightarrow{\mathrm{id}_E} \underset{\text{muni de } \mathscr{B}}{E} \xrightarrow{f} \underset{\text{muni de } \mathscr{C}}{F} \xrightarrow{\mathrm{id}_F} \underset{\text{muni de } \mathscr{C}'}{F}.$$

La formule qui donne la matrice d'une composée permet d'en déduire, au niveau des matrices :

$A' = \mathfrak{Mat}_{\mathscr{B}',\mathscr{C}'}\,(\mathrm{id}_F \circ f \circ \mathrm{id}_E) = \left(\mathfrak{Mat}_{\mathscr{C},\mathscr{C}'}\ \mathrm{id}_F\right)\left(\mathfrak{Mat}_{\mathscr{B},\mathscr{C}}\, f\right)\left(\mathfrak{Mat}_{\mathscr{B}',\mathscr{B}}\ \mathrm{id}_E\right) = Q^{-1}AP.$ □

Remarque

On suppose ici que $E = F$ avec $\mathscr{C} = \mathscr{B}$ et $\mathscr{C}' = \mathscr{B}'$.

Lorsque les vecteurs de la base $\mathscr{B}'$ s'expriment à l'aide d'un système triangulaire en fonction des vecteurs de la base $\mathscr{B}$, il est aussi commode de calculer A' par un calcul vectoriel, sans utiliser la formule de changement de base (méthode de l'exemple de la fin du paragraphe II 1).

Définition-Proposition

Soient $A, A' \in \mathfrak{M}(n,\mathbb{K})$.

(a) On dit que A et A' sont *semblables* s'il existe $P \in \mathfrak{M}(n,\mathbb{K})$ inversible tel que :

$$A' = P^{-1}AP. \quad \longleftarrow [\text{donc } A = PA'P^{-1} \text{ puis } A' \text{ et } A \text{ sont semblables}]$$

(b) Dans ce cas, on a : $\mathrm{rg}\,(A') = \mathrm{rg}\,(A)$.

Démonstration

On fait intervenir l'application linéaire $f : \begin{array}{rcl} \mathbb{K}^n & \to & \mathbb{K}^n \\ X & \mapsto & AX \end{array}$ et la base canonique $\mathscr{B}$ de $\mathbb{K}^n$.

Comme P est de rang n, ses colonnes forment une base $\mathscr{B}'$ de $\mathbb{K}^n$. D'après la formule de changement de base, la matrice de f dans la base $\mathscr{B}'$ est A'. D'où : $\mathrm{rg}\ A' = \mathrm{rg}\ f = \mathrm{rg}\ A$. □

Remarque

Soit $A \in \mathfrak{M}_{n,p}(\mathbb{K})$. L'idée qui a permis de démontrer l'égalité du (b), dans le cas des matrices carrées, va nous permettre de redémontrer une 3e fois que ${}^tA$ a même rang que A.

On introduit $f : \begin{array}{rcl} \mathbb{K}^p & \to & \mathbb{K}^n \\ X & \mapsto & AX \end{array}$ ainsi que les bases canonique $\mathscr{B}$ de $\mathbb{K}^p$ et $\mathscr{C}$ de $\mathbb{K}^n$.

On complète à gauche une base $(v'_{q+1}, \ldots, v'_p)$ de $\mathrm{Ker}\, f$ en une base $\mathscr{B}' = (v'_1, \ldots, v'_p)$ de $\mathbb{K}^p$. D'après la proposition qui précède le théorème du rang, la famille $(f(v'_1), \ldots, f(v'_q))$ est une base de $\mathrm{Im}\, f$. On la complète à droite en une base $\mathscr{C}' = (w'_1, \ldots, w'_n)$ de $\mathbb{K}^n$.

La matrice A' de f dans les bases $\mathscr{B}'$ et $\mathscr{C}'$ s'écrit donc : $A' = \underbrace{\left(\begin{array}{c|c} I_q & 0 \\ \hline 0 & 0 \end{array}\right)}_{p} \Big\updownarrow n.$

D'après la formule de changement de base, on a aussi :

$A' = Q^{-1}AP$ où $\left| \begin{array}{lll} P & \text{est la matrice de passage de} & \mathscr{B} \text{ à } \mathscr{B}' \,; \\ Q & \quad\quad\quad \text{,,} & \mathscr{C} \text{ à } \mathscr{C}'. \end{array} \right.$

Donc ${}^t(A') = {}^tP\, {}^tA\, ({}^tQ)^{-1}$ avec ${}^t(A') = \underbrace{\left(\begin{array}{c|c} I_q & 0 \\ \hline 0 & 0 \end{array}\right)}_{n} \Big\updownarrow p.$

On note $g : \begin{array}{rcl} \mathbb{K}^n & \to & \mathbb{K}^p \\ X & \mapsto & {}^tAX \end{array}$, $\mathscr{B}''$ la base de $\mathbb{K}^p$ dont les vecteurs sont les colonnes de $({}^tP)^{-1}$ et $\mathscr{C}''$ la base de $\mathbb{K}^n$ dont les vecteurs sont les colonnes de $({}^tQ)^{-1}$.

D'après la formule de changement de base, la matrice de g dans les bases $\mathscr{C}''$ et $\mathscr{B}''$ est ${}^t(A')$.

La méthode de Gauss (par exemple) montre que A' et ${}^t(A')$ ont toutes deux pour rang q.

D'où le résultat : $\mathrm{rg}\ {}^tA = \mathrm{rg}\ g = \mathrm{rg}\ {}^t(A') = q = \mathrm{rg}\ A' = \mathrm{rg}\ f = \mathrm{rg}\ A$.

Exemple

Soit $f: \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ \binom{x}{y} & \mapsto & \binom{x'}{y'} \end{array}$ avec $\begin{cases} x' = 4x - 3y \\ y' = 2x - y \end{cases}$.

On pose : $e_1 := \binom{1}{0}$ et $e_2 := \binom{0}{1}$, $\mathscr{B} := (e_1, e_2)$; $e'_1 := \binom{3}{2}$ et $e'_2 := \binom{1}{1}$, $\mathscr{B}' := (e'_1, e'_2)$.
On va voir que $\mathscr{B}'$ est une base de $\mathbb{R}^2$ pour laquelle $D := \mathfrak{Mat}_{\mathscr{B}'} f$ est « simple ».

(a) D'après l'exemple 2 du paragraphe 1, la matrice $A := \mathfrak{Mat}_{\mathscr{B}} f$ s'écrit : $A = \begin{pmatrix} 4 & -3 \\ 2 & -1 \end{pmatrix}$.

On introduit $P := \begin{array}{c} \begin{matrix} e'_1/\mathscr{B} & e'_2/\mathscr{B} \end{matrix} \\ \begin{pmatrix} \square & \square \end{pmatrix} \end{array} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$.

On montre que P est inversible et calcule son inverse par la méthode de Gauss-Jordan :

$$(P|I) = \left(\begin{array}{cc|cc} 3 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array}\right) \xrightarrow[L_2 \leftarrow 3L_2 - 2L_1]{} \left(\begin{array}{cc|cc} 3 & 1 & 1 & 0 \\ 0 & 1 & -2 & 3 \end{array}\right) \xrightarrow[L_1 \leftarrow L_1 - L_2]{} \left(\begin{array}{cc|cc} 3 & 0 & 3 & -3 \\ 0 & 1 & -2 & 3 \end{array}\right) \xrightarrow[L_1 \leftarrow \frac{1}{3} L_1]{} \left(\begin{array}{cc|cc} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 3 \end{array}\right).$$

Donc P est inversible et $P^{-1} = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix}$. ⟵[reprise de l'exemple de la fin du II 1]

Par suite, $\mathscr{B}'$ est une base de $\mathbb{R}^2$ car $\operatorname{rg}(e'_1, e'_2) = \operatorname{rg} P = 2$. Finalement :

$$D = P^{-1}AP = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 4 & -3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

(Cela se vérifie aussi après coup : $f(e'_1) = A\binom{3}{2} = \binom{6}{4} = 2e'_1$ et $f(e'_2) = A\binom{1}{1} = \binom{1}{1} = e'_2$.)

(b) L'égalité $A = PDP^{-1}$ montre que A est semblable à une matrice diagonale.
Elle va aussi permettre de calculer facilement par exemple A^{100}.
On a tout d'abord, par une récurrence immédiate, l'égalité $D^n = \begin{pmatrix} 2^n & 0 \\ 0 & 1 \end{pmatrix}$ pour tout $n \in \mathbb{N}$.

$$\left(\text{En effet : } \left| \begin{array}{l} \text{(i) } D^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2^0 & 0 \\ 0 & 1 \end{pmatrix}; \\ \text{(ii) si } D^n = \begin{pmatrix} 2^n & 0 \\ 0 & 1 \end{pmatrix}, \text{ alors } D^{n+1} = D^n D = \begin{pmatrix} 2^n & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2^{n+1} & 0 \\ 0 & 1 \end{pmatrix}. \end{array} \right. \right)$$

On en déduit que : $A^{100} = \underbrace{(PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1})(PDP^{-1})}_{\text{produit de 100 termes } (PDP^{-1}) \text{ dans lequel les } P^{-1})(P \text{ peuvent être enlevés}} = PD^{100}P^{-1}$ puis

$$A^{100} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2^{100} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 3{\times}2^{100} & 1 \\ 2{\times}2^{100} & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 3{\times}2^{100} - 2 & -3{\times}2^{100} + 3 \\ 2{\times}2^{100} - 2 & -2{\times}2^{100} + 3 \end{pmatrix}.$$

<u>Précisions</u>

L'égalité $A = PDP^{-1}$ permet de même de calculer d'autres choses comme $\exp(A)$, à définir.

Mais pour calculer A^{100}, il est plus simple de constater que $A^2 = 3A - 2I$. En effet, la division euclidienne de X^{100} par $X^2 - 3X + 2$ s'écrit $X^{100} = (X^2 - 3X + 1)Q + aX + b$ avec $Q \in \mathbb{R}[X]$ et $a, b \in \mathbb{R}$; le calcul des valeurs en 1 et 2 montre que $a = 2^{100} - 1$ et $b = -2^{100} + 2$. Cela redonne le résultat du (b) : $A^{100} = aA + bI = (2^{100} - 1)A + (-2^{100} + 2)I$.

Ch. 5. Suites récurrentes d'ordre deux

Dans tout ce chapitre, on note $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et fixe $p \in \mathbb{N} \setminus \{0\}$.

I. Suites récurrentes d'ordre p

1. Généralités (hors programme)

Définition

Soit $(u_n)_{n\geq 0} \in \mathbb{R}^{\mathbb{N}}$. Une *relation de récurrence d'ordre p* est une équation de la forme

$$(E)\colon\ u_{n+p} = \underbrace{f\big(n, u_n, \ldots, u_{n+p-1}\big)}_{\text{avec la condition que } (n, u_n, \ldots, u_{n+p-1}) \in U} \ \text{ pour tout } n \in \mathbb{N}, \text{ avec } \ f : \underbrace{U}_{\text{partie de } \mathbb{R}^{p+1}} \longrightarrow \mathbb{R}.$$

Proposition

Soient A un ensemble, $(f_n)_{n\geq 0}$ une suite d'applications de A dans A, et $\alpha \in A$.

Il existe une unique suite $(x_n)_{n\geq 0}$ d'éléments de A telle que :

$$x_0 = \alpha \qquad \text{et} \qquad x_{n+1} = f_n(x_n) \quad \text{pour tout } n \in \mathbb{N}$$

« suite définie par une relation de récurrence ».

Démonstration (idée)

Une récurrence permet de montrer que pour tout $n \in \mathbb{N}$, la proposition suivante est vraie : $\mathscr{P}(n)$: il existe $(X_0, ..., X_n) \in A^{n+1}$ unique tel que $X_0 = \alpha$ et $X_{k+1} = f_k(X_k)$ pour $0 \leq k \leq n-1$. On constate que pour chaque $k \in \mathbb{N}$, le terme X_k dans $(X_0, X_1, ..., X_n)$ ne dépend pas de $n \geq k$. On le note x_k. La suite $(x_n)_{n\geq 0}$ convient et elle la seule à convenir. $\square$

Corollaire

On suppose que $U = \mathbb{N} \times \mathbb{R}^p$ dans l'équation (E) et fixe $\alpha_0, ..., \alpha_{p-1} \in \mathbb{R}$.

On change d'inconnue en remplaçant $(u_n)_{n\geq 0}$ par la suite $(X_n)_{n\geq 0}$ de vecteurs définie par :

$$X_n = (u_n, ..., u_{n+p-1}) \text{ pour } n \in \mathbb{N}.$$

L'équation (E) équivaut à : $X_{n+1} = \underbrace{\big(u_{n+1}, ..., u_{n+p-1}, f(n, X_n)\big)}_{\text{noté } f_n(X_n)}$ pour tout $n \in \mathbb{N}$.

En choisissant $A = \mathbb{R}^p$, on obtient : il existe une unique solution $(u_n)_{n\geq 0}$ de (E) qui vérifie la condition initiale suivante : $u_0 = \alpha_0$ et ... et $u_{p-1} = \alpha_{p-1}$.

2. Cas linéaire (au programme avec $p = 2$)

Définition

Soit $(u_n)_{n\geq 0} \in \mathbb{K}^{\mathbb{N}}$. Une *relation de récurrence linéaire d'ordre p* est une équation

$$(L)\colon\ u_{n+p} + a_n^{(p-1)} u_{n+p-1} + \cdots + a_n^{(0)} u_n = b_n \ \text{ pour tout } n \in \mathbb{N},$$

où $(a_n^{(0)})_{n\geq 0}$, ..., $(a_n^{(p-1)})_{n\geq 0}$, et $(b_n)_{n\geq 0}$ sont des suites fixées d'éléments de $\mathbb{K}$.

Dans ce cas, *la relation de récurrence linéaire homogène associée à (L)* est

$$(H)\colon\ u_{n+p} + a_n^{(p-1)} u_{n+p-1} + \cdots + a_n^{(0)} u_n = 0 \ \text{ pour tout } n \in \mathbb{N}.$$

Remarque

On pose à nouveau $X_n = (u_n, ..., u_{n+p-1})$ pour $n \in \mathbb{N}$.

Soient $\alpha_0, ..., \alpha_{p-1} \in \mathbb{K}$. Les arguments du corollaire du 1 s'adaptent quand $\mathbb{K} = \mathbb{C}$.
Il existe une unique solution $(u_n)_{n\geq 0}$ de (L) telle que : $u_0 = \alpha_0$ et ... et $u_{p-1} = \alpha_{p-1}$.

Cas particulier

Soient $a, b \in \mathbb{K}$. On considère l'équation $(H_{a,b})\colon u_{n+2} + au_{n+p-1} + bu_n = 0$ dans $\mathbb{K}^{\mathbb{N}}$.

Pour tout $(\alpha, \beta) \in \mathbb{K}^2$, il existe une solution $(u_n)_{n\geq 0}$ de $(H_{a,b})$ telle que $u_0 = \alpha$ et $u_1 = \beta$.
Toute solution $(u_n)_{n\geq 0}$ de $(H_{a,b})$ est déterminée par ses deux premiers termes u_0 et u_1.

Proposition (notations de la définition)

On se donne (il en existe) une solution $u_P \in \mathbb{K}^{\mathbb{N}}$ de (L), qu'on dira « particulière ».

L'ensemble $\mathscr{S}_L$ des solutions $(u_n)_{n\geq 0}$ de (L) s'écrit $\boxed{\mathscr{S}_L = \mathscr{S}_H + u_P}$.
Cela signifie que les solutions de (L) sont les suites $u := u_H + u_P$ avec u_H qui décrit $\mathscr{S}_H$.

Démonstration

On se donne une solution $u_P = ((u_P)_n)_{n\geq 0}$ de (L), par exemple l'unique solution $\widetilde{u}$ de (L) vérifiant $\widetilde{u}_0 = \widetilde{u}_1 = ... = \widetilde{u}_{p-1} = 0$ que fournit la remarque. Soit $u = (u_n)_{n\geq 0} \in \mathbb{K}^{\mathbb{N}}$.

On a : $(u_P)_{n+p} + a_n^{(p-1)}(u_P)_{n+p-1} + \cdots + a_n^{(0)}(u_P)_n = b_n$ pour tout $n \geq 0$. Donc :

$$\begin{aligned}
(L) \iff \forall n \in \mathbb{N} \quad & u_{n+p} + a_n^{(p-1)}u_{n+p-1} + \cdots + a_n^{(0)}u_n \\
& = (u_P)_{n+p} + a_n^{(p-1)}(u_P)_{n+p-1} + \cdots + a_n^{(0)}(u_P)_n \\
\iff \forall n \in \mathbb{N} \quad & (u - u_P)_{n+p} + a_n^{(p-1)}(u - u_P)_{n+p-1} + \cdots + a_n^{(0)}(u - u_P)_n = 0 \\
\iff u - u_P \in \mathscr{S}_H \quad & \iff \exists u_H \in \mathscr{S}_H \quad u = u_H + u_P.
\end{aligned}$$

Cela permet de conclure □

II. Équation homogène

1. Coefficients quelconques (au programme avec $p = 2$)

Proposition

Soient $(a_n^{(0)})_{n\geq 0}, ..., (a_n^{(p-1)})_{n\geq 0}, (b_n)_{n\geq 0} \in \mathbb{K}^{\mathbb{N}}$.

L'ensemble $\mathscr{S}_H$ des solutions $(u_n)_{n\geq 0}$ de $(H)\colon u_{n+p} + a_n^{(p-1)}u_{n+p-1} + \cdots + a_n^{(0)}u_n = 0$ est un sous-espace vectoriel de $\mathbb{K}^{\mathbb{N}}$ de dimension p.
De plus, des solutions $v^{(1)}$, ..., $v^{(p)}$ de (H) forment une base de $\mathscr{S}_H$ si et seulement si les vecteurs $\begin{pmatrix} v_0^{(1)} \\ v_1^{(1)} \\ \vdots \\ v_{p-1}^{(1)} \end{pmatrix}$, ..., $\begin{pmatrix} v_0^{(p)} \\ v_1^{(p)} \\ \vdots \\ v_{p-1}^{(p)} \end{pmatrix}$ forment une base de $\mathbb{K}^p$.

Démonstration

Il est immédiat que $\mathscr{S}_H$ est un sous-espace vectoriel de $\mathbb{K}^{\mathbb{N}}$.
On introduit l'application $l : \begin{array}{ccc} \mathscr{S}_H & \longrightarrow & \mathbb{K}^p \\ u & \longmapsto & (u_0, u_1, ..., u_{p-1}) \end{array}$.

Elle est clairement linéaire. D'après la remarque, elle est bijective.
Ainsi, l est un isomorphisme de $\mathscr{S}_H$ sur $\mathbb{K}^p$. Or $\mathbb{K}^p$ est de dimension finie et $\dim \mathbb{K}^p = p$.
Donc $\mathscr{S}_H$ est de dimension finie et $\dim \mathscr{S}_H = p$.

En outre, des vecteurs $v^{(1)}, ..., v^{(p)} \in \mathscr{S}_H$ forment une base de $\mathscr{S}_H$ si et seulement si les vecteurs $l(v^{(1)}), ..., l(v^{(p)})$ forment une base de $\mathbb{K}^p$. D'où le résultat. □

Cas particulier

Soient $a, b \in \mathbb{K}$. On considère l'équation $(H_{a,b})\colon u_{n+2} + au_{n+p-1} + bu_n = 0$ dans $\mathbb{K}^{\mathbb{N}}$.

L'ensemble $\mathscr{S}_{H_{a,b}}$ des solutions $(H_{a,b})$ est un $\mathbb{K}$-espace vectoriel de dimension 2.

Exemple

On considère l'équation $(H)\colon u_{n+1} - (n+1)u_n = 0$ d'inconnue $(u_n)_{n\geq 0} \in \mathbb{R}^{\mathbb{N}}$.
Elle équivaut à : $(u_n)_{n\geq 0} = (n!\, u_0)_{n\geq 0}$.
L'ensemble de ses solutions est donc la droite vectorielle de base $(n!)_{n\geq 0}$.

2. Coefficients constants (au programme avec $p = 2$)

On suppose dans ce paragraphe que dans l'équation (H) les suites $(a_n^{(0)})_{n\geq 0}$, ..., $(a_n^{(p-1)})_{n\geq 0}$ sont constantes. On note $a_0, \dots, a_{p-1} \in \mathbb{K}$ leurs valeurs prises.

Par conséquent : $(H)\colon u_{n+p} + a_{p-1}u_{n+p-1} + \cdots + a_0 u_n = 0$ pour tout $n \geq 0$.

On a vu que l'ensemble $\mathscr{S}_H$ des solutions (H) est un $\mathbb{K}$-espace vectoriel de dimension p.

Proposition (cas particulier de la proposition du I. 1, obtenu directement)

Soient $p \in \mathbb{N}$, $M \in \mathfrak{M}(p, \mathbb{K})$ et $C \in \mathbb{K}^p$.

Il existe une unique suite $(X_n)_{n\geq 0}$ d'éléments de $\mathbb{K}^p$ telle que :

$$X_0 = C \quad \text{et} \quad X_{n+1} = MX_n \quad \text{pour } n \in \mathbb{N}.$$

En effet, cela équivaut à : $X_n = M^n C$ pour tout $n \in \mathbb{N}$.

Remarque (construction explicite d'une solution de (H))

Soit $(u_n)_{n\geq 0} \in \mathbb{K}^{\mathbb{N}}$. On pose encore $X_n = (u_n, \dots, u_{n+p-1})$ pour $n \in \mathbb{N}$.

Dans ce cas, (H) équivaut — en voyant X_n en colonne — à :

$$X_{n+1} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{p-1} \end{pmatrix}}_{\text{notée } M} X_n \quad \text{pour } n \in \mathbb{N}.$$

Soient $\alpha_0, \dots, \alpha_{p-1} \in \mathbb{K}$. D'après la proposition, la solution $(u_n)_{n\geq 0}$ de (H) qui vérifie $u_0 = \alpha_0$ et ... et $u_{p-1} = \alpha_{p-1}$ est déterminée par : $X_n = M^n \begin{pmatrix} \alpha_0 \\ \cdots \\ \alpha_{p-1} \end{pmatrix}$ pour tout $n \in \mathbb{N}$.

(On a : $(M^p + a_{p-1}\, M^{p-1} + \cdots + a_0 I)X_0 = X_p + a_{p-1}\, X_{p-1} + \cdots + X_0 = 0$.
Par choix successifs pour $X_0 = (\alpha_0, \dots, \alpha_{p-1})$ de chacune des colonnes de la matrice I (vues comme blocs), on obtient : $M^p + a_{p-1}\, M^{p-1} + \cdots + a_0 I = (M^p + a_{p-1}\, M^{p-1} + \cdots + a_0 I)I = 0$.
Mieux : le polynôme caractéristique C de M est $C := X^p + a_{p-1}\, X^{p-1} + \cdots + a_0$ (facile).)

Définition

Soient $a, b \in \mathbb{K}$. On considère l'équation $(H_{a,b})\colon u_{n+2} + au_{n+p-1} + bu_n = 0$ dans $\mathbb{K}^{\mathbb{N}}$.

L'*équation caractéristique* de $(H_{a,b})$ est l'équation $\boxed{r^2 + ar + b = 0}$ d'inconnue $r \in \mathbb{C}$.
[C'est une condition nécessaire et suffisante pour que $(r^n)_{n\geq 0}$ vérifie $(H_{a,b})$ quand $\mathbb{K} = \mathbb{C}$.]

Proposition

Soient $a, b \in \mathbb{K}$. On s'intéresse à $(H_{a,b})\colon u_{n+2} + au_{n+1} + bu_n = 0$ d'inconnue $(u_n)_{n\geq 0} \in \mathbb{K}^{\mathbb{N}}$. On note r_1 et r_2 les solutions dans $\mathbb{C}$ de son équation caractéristique.

(a) On obtient une base (v, w) de l'espace des solutions de $(H_{a,b})$ en posant pour $n \in \mathbb{N}$

– si $r_1, r_2 \in \mathbb{K}$: $v_n = r_1^n$ et $w_n = \begin{cases} r_2^n & \text{si } r_2 \neq r_1 \\ \text{ou} & \\ nr_1^{n-1} & \text{si } r_2 = r_1 \end{cases}$ $(*)$.

– sinon : $v_n = \rho^n \cos(n\theta)$ et $w_n = \rho^n \sin(n\theta)$ où $r_1 = \rho\, \mathrm{e}^{\mathrm{i}\theta}$ avec $\rho > 0$ et $\theta \in \mathbb{R}$, et $\overbrace{r_2 = \overline{r_1}}^{\text{car } a,b \in \mathbb{K} \overset{\text{ici}}{=} \mathbb{R}}$.

$(*)$ Dans ces formules, les notations r_1^n, r_2^n et nr_1^{n-1} sont obtenues par prolongement par continuité à n fixé quand $r_1 = 0$ ou $r_2 = 0$. Lorsque $r_1 = r_2$ et $r_1 \neq 0$, on peut choisir $(w_n)_{n\geq 0}$ avec nr_1^n à la place de nr_1^{n-1}.

(b) Soient $P \in \mathbb{C}[X]\setminus\{0\}$ de degré d et $r \in \mathbb{C}\setminus\{0\}$ (**).

L'équation $(L_{a,b})\colon u_{n+2} + a u_{n+1} + b u_n = P(n)\, r^n$ a une $\overbrace{\text{solution}}^{\text{unique}}$ $u_P \in \mathbb{K}^{\mathbb{N}}$ de la forme $u_P = \big(Q(n)\, n^m r^n\big)_{n\geq 0}$ avec $Q \in \mathbb{C}[X]\setminus\{0\}$ de degré d et $m = \begin{cases} 0 & \text{si } r \notin \{r_1, r_2\} \\ \text{ou } 1 & \text{si } r \in \{r_1, r_2\} \text{ et } r_1 \neq r_2 \\ \text{ou } 2 & \text{si } r = r_1 = r_2 \end{cases}$.

Démonstration

(a) On suppose d'abord que $r_1, r_2 \in \mathbb{K}$. Les suites v et w sont des solutions de $(H_{a,b})$, car : cela a été signalé dans la définition quand $r_1 \neq r_2$ et on utilise $2r_1 + a = 0$ quand $r_1 = r_2$.

On a : $\begin{pmatrix} v_0 & w_0 \\ v_1 & w_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ r_1 & r_2 \end{pmatrix}$ ou $\begin{pmatrix} 1 & 0 \\ r_1 & 1 \end{pmatrix}$ donc (v, w) est libre dans $\mathscr{S}_{H_{a,b}}$ qui est de dimension 2.

Pour terminer, on suppose que $\mathbb{K} = \mathbb{R}$, et que les racines r_1 et r_2 du polynôme réel $X^2 + a\,X + b$ sont non-réelles, donc de la forme $r_1 = \rho\, \mathrm{e}^{\mathrm{i}\theta}$ et $r_2 = \rho\, \mathrm{e}^{-\mathrm{i}\theta}$ pour certains $\rho > 0$ et $\theta \in \mathbb{R}$.
D'après ce qui précède, $v^{(1)} = (r_1^n)_{n\geq 0}$ et son conjugué $v^{(2)} = (r_2^n)_{n\geq 0}$ sont des solutions de $(H_{a,b})$ sur $\mathbb{C}$ qui sont linéairement indépendantes. Donc les parties réelle $v := \frac{v^{(1)} + v^{(2)}}{2}$ et imaginaire $w := \frac{v^{(1)} - v^{(2)}}{2\mathrm{i}}$ de $v^{(1)}$ sont des solutions de $(H_{a,b})$ sur $\mathbb{R}$ qui sont linéairement indépendantes (car $v^{(1)}, v^{(2)} \in \mathrm{Vect}_{\mathbb{C}}(v, w)$), et qui engendrent l'espace des solutions de $(H_{a,b})$ sur $\mathbb{R}$ (car une solution u de $(H_{a,b})$ sur $\mathbb{R}$ est solution de $(H_{a,b})$ sur $\mathbb{C}$, donc $u \in \mathrm{Vect}_{\mathbb{C}}(v^{(1)}, v^{(2)})$ puis $u = \mathrm{Re}\, u \in \mathrm{Vect}_{\mathbb{R}}(\mathrm{Re}\, v^{(1)}, \mathrm{Re}\, v^{(2)}, \mathrm{Im}\, v^{(1)}, \mathrm{Im}\, v^{(2)},))$.

On suppose que $\mathbb{K} = \mathbb{C}$. On a : $\mathscr{S}_H = \mathrm{Ker}\, C(f)$ où $f\colon (u_n)_{n\geq 0} \in \mathbb{C}^{\mathbb{N}} \mapsto (u_{n+1})_{n\geq 0} \in \mathbb{C}^{\mathbb{N}}$.
On note $f_H\colon \mathscr{S}_H \xrightarrow{f} \mathscr{S}_H$, donc $f_H((u_n)_{n\geq 0}) = (v_n)_{n\geq 0}$ avec $(v_0, \dots, v_{p-1}) = M(u_0, \dots, u_{p-1})$.
On a : $\mathscr{S}_H$ est somme directe de $\mathrm{Ker}(f - r\,\mathrm{id})^\alpha$, de dimension α car solutions d'un $(H_{r,\alpha})$.
On trouve : $\underbrace{\big((r^n)_{n\geq 0}, (\tfrac{n}{1!}\, r^{n-1})_{n\geq 0}, \dots, \tfrac{(n(n-1)\dots(n-\alpha+1)}{\alpha!}\, r^{n-\alpha})_{n\geq 0}\big)}_{\substack{\text{la matrice des } \alpha \text{ premiers coefficients est triangulaire inférieure} \\ \text{et } f - r\,\mathrm{id} \text{ envoie l'un de ces vecteurs sur le précédent (le premier sur 0)}}}$ base de $\mathrm{Ker}(f - r\,\mathrm{id})^\alpha$.

b) Soient $P \in \mathbb{C}[X]\setminus\{0\}$ de degré d et $r \in \mathbb{C}\setminus\{0\}$. Cas général de l'ordre p (ici $p = 2$) :

$$(L)\colon u_{n+p} + a_{p-1} u_{n+p-1} + \cdots + a_0 u_n = P(n)\, r^n \text{ pour tout } n \geq 0.$$

On note : $C = X^p + a_{p-1} X^{p-1} + \cdots + a_0$ et m la multiplicité de r dans C ;
$f\colon (u_n)_{n\geq 0} \in \mathbb{C}^{\mathbb{N}} \mapsto (u_{n+1})_{n\geq 0} \in \mathbb{C}^{\mathbb{N}}$ et $i\colon U \in \mathbb{C}[X] \mapsto \big(U(n) r^n\big)_{n\geq 0} \in \mathbb{C}^{\mathbb{N}}$;
$\widetilde{f}\colon U \in \mathbb{C}[X] \mapsto r\, U(X+1) \in \mathbb{C}[X]$.

Soit $U \in \mathbb{C}[X]$. On pose $u := i(U)$. On a : $f(i(U)) = i(\widetilde{f}(U))$. D'où, comme i est injective : $(L) \iff C(f)(i(U)) = i(P) \iff i(C(\widetilde{f})(U)) = i(P) \iff C(\widetilde{f})(U) = P$.

Il reste à démontrer que $C(\widetilde{f})(X^m\, \mathbb{C}_d[X]) = \mathbb{C}_d[X]$. On pose pour cela :
$C = \prod_{k=1}^{q} (X - r_k)^{\alpha_k}$ avec $\underbrace{\{r_1, \dots, r_q\}}_{\text{distincts}} \subseteq \mathbb{C}$ et $\alpha_1, \dots, \alpha_q \in \mathbb{N}\setminus\{0\}$, donc $C(\widetilde{f}) = \prod_{k=1}^{q} (\widetilde{f} - r_k\,\mathrm{id})^{\alpha_k}$.

On a : $\deg((\widetilde{f} - r\,\mathrm{id})(X^{n+1})) = n$ pour $n \in \mathbb{N}$, donc $X^m \mathbb{C}_d[X] \xrightarrow{(\widetilde{f} - r\,\mathrm{id})^m} \mathbb{C}_d[X]$ est bijective.
Si $r_k \neq r$: $\deg((\widetilde{f} - r_k\,\mathrm{id})(X^n)) = n$ pour $n \in \mathbb{N}$, donc $\mathbb{C}_d[X] \xrightarrow{(f - r_k\,\mathrm{id})^{\alpha_k}} \mathbb{C}_d[X]$ est bijective.
En appliquant d'abord $(\widetilde{f} - r\,\mathrm{id})^m$, on en déduit que $X^m\, \mathbb{C}_d[X] \xrightarrow{C(\widetilde{f})} \mathbb{C}_d[X]$ est bijective. □

<u>Variante pour obtenir $C(\widetilde{f})(X^m\, \mathbb{C}_d[X]) = \mathbb{C}_d[X]$</u>

Soit $j \in \{0, 1, \dots, d\}$. Formule du binôme : $C(\widetilde{f})(X^{m+j}) = \sum_{k=0}^{m+j} \binom{m+j}{k} C_k(r)\, X^{m+j-k}$,
où $C_k := p^k X^p + a_{p-1}(p-1)^k X^{p-1} + \cdots + a_0 0^k$ est déterminé par $C_0 = C$ et $C_{k+1} = X C_k'$.
En particulier, par récurrence on a : $C_k \in X^k C^{(k)} + \mathrm{Vect}(X^{k-1} C^{(k-1)}, \dots, X^1 C', C)$.
Comme $C_0(r) = \dots = C_{m-1}(r) = 0$ et $C_m(r) \neq 0$, on en déduit que $\deg(C(\widetilde{f})(X^{m+j})) = j$.
Ainsi $\mathbb{C}_d[X]$ a pour base $(C(\widetilde{f})(X^m), \dots, C(\widetilde{f})(X^{m+d}))$ et $X^m\, \mathbb{C}_d[X] \xrightarrow{l} \mathbb{C}_d[X]$ est bijective.

Exemple

Soit $(F_n)_{n\geq 0}$ vérifiant : $F_{n+2} = F_{n+1} + F_n$ avec $F_0 = 0$ et $F_1 = 1$ « suite de Fibonacci ».
Voici ses premiers termes : 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987...

Les solutions de $r^2 - r - 1 = 0$ sont $r_1 = \frac{1+\sqrt{5}}{2}$ et $r_2 = \frac{1-\sqrt{5}}{2}$, donc $-1 < r_2 < 0$.

(**) Dans le cas $r = 0$, on choisirait $u_P = (u_n)_{n\geq 0}$ avec $u_n = 0$ pour $n \geq 1$ et $b u_0 = P(0)$.

Vu les conditions initiales, on obtient : $F_n = \frac{1}{\sqrt{5}}(r_1^n - r_2^n)$ pour tout $n \geq 0$.

Pour tout $n \geq 0$, on a : $\underbrace{\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix}}_{\text{noté } X_n} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, ce qui donne $F_n^2 + F_{n+1}^2 = {}^t X_n X_n = F_{2n+1}$.