Cours d'Algèbre pour les étudiants de MPCI

Par

Oumar SALL

Maître de conférences au

Département de Mathématiques

U.F.R. des Sciences et Technologies

Université de Ziguinchor - SENEGAL

E-mail: oumarsfr@yahoo.fr

05/10/2007

Contents

Preface

Chapter 1

Notions préliminaires

1.1 Ensembles et sous-ensembles

George Cantor, le fondateur de la théorie des ensembles définissait un ensemble comme " un groupement d'objets déterminés et bien distincts, de notre perception ou de notre entendement, et que l'on appelle les éléments de l'ensemble". Nous considèrerons la notion d'ensemble comme intuitive en gardant néanmoins en mémoire le fait qu'on ne peut pas considérer "n'importe quoi" comme ensemble si l'on veut éviter les contradictions.

1.1.1 Notion d'ensemble - Elément d'un ensemble

Un ensemble est une collection d'objets satisfaisant un certain nombre de propriétés et chacun de ces objets est appelé élément de cet ensemble.

Dans la pratique il y a deux façons de construire ou d'écrire des ensembles:

- par extension c'est-à-dire en donnant la liste des éléments, par exemple $E := \{0, 1, 5, 6, 9\}$ est un ensemble;

- par compréhension c'est-à-dire en décrivant une caractérisation des éléments, par exemple $A :=\{x \mid x \text{ est un jour de la semaine}\}$ est un ensemble.

Si x est un élément d'un ensemble E, on dit aussi que x appartient à E et on note $x \in E$. Si x n'appartient pas à E, on note $x \notin E$. Parmi les ensembles les plus importants on peut citer: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, et $\mathbb{C}$ qui désignent respectivement l'ensemble des entiers naturels, des entiers relatifs, des nombres rationnels, des nombres réels et des nombres complexes. On admet l'existence d'un ensemble n'ayant aucun élément; cet ensemble est appelé ensemble vide et peut être défini par $\{x \mid x \neq x\}$ et il est noté $\varnothing$ ou $\{\ \}$.

1.1.2 Inclusion

On dit qu'un ensemble A est inclus dans un ensemble B si chaque élément de A est un élément de B. On dit aussi " A est contenu dans B " ou " A est un sous-ensemble de B " et on note $A \subset B$.

Remarques

Soient A, B, C trois ensembles

- La relation d'inclusion est réflexive: $A \subset A$
- La relation d'inclusion est transitive: Si $A \subset B$ et $B \subset C$, alors $A \subset C$
- La relation d'inclusion est antisymétrique: ($A \subset B$ et $B \subset A$) si et seulement si $A = B$

Exemples

- $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$
- $\{x \in \mathbb{R} \,|0 \leq x \leq 5\} \subset \mathbb{R}_+$
- Les sous-ensembles d'un ensemble E forment un ensemble appelé ensembles des parties de E et noté $\mathcal{P}(E)$.

Si $E = \{1, 2\}$ alors $\mathcal{P}(E) = \{\varnothing, \{1\}, \{2\}, E\}$.

Les trois assertions $x \in E$, $\{x\} \subset E$ et $\{x\} \in \mathcal{P}(E)$ sont équivalentes.

1.1.3 Intersection et réunion (ou union)

Soient A et B deux sous-ensembles d'un ensemble E, l'ensemble

$$\{x \mid x \in A \text{ et } x \in B\}$$

est appelé l'intersection des ensembles A et B et est noté $A \cap B$. Si $A \cap B = \varnothing$, on dit que A et B sont disjoints.

L'ensemble

$$\{x \mid x \in A \text{ ou } x \in B\}$$

est appelé l'union des ensembles A et B et est noté $A \cup B$. On a:

- $A \cap \varnothing = \phi$ et $A \cup \varnothing = A$
- $A \cap B \subset A$ et $A \cap B \subset B$
- $A \subset A \cup B$ et $B \subset A \cup B$
- $A \cup B = A$ si et seulement si $B \subset A$
- $A \cap B = A$ si et seulement si $A \subset B$

Propriétés de $\cap$ et $\cup$

Soient A, B, C trois sous-ensembles d'un ensemble E. On a:

- L'union est commutative:

$$A \cup B = B \cup A$$

- L'intersection est commutative:

$$A \cap B = B \cap A$$

- L'union est associative:

$$A \cup (B \cup C) = (A \cup B) \cup C$$

- L'intersection est associative:

$$A \cap (B \cap C) = (A \cap B) \cap C$$

- L'union est distributive par rapport à l'intersection:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

- L'intersection est distributive par rapport à l'union:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Remarque

Ne pas oublier les parenthèses. Par exemple, $A \cap B \cup C$ n'a pas de sens.
Si $A = [0,\ 1]$, $B = [1,\ 2]$ et $C = [2,\ +\infty[$, on a
$(A \cap B) \cup C = \{1\} \cup [2,\ +\infty[$, et $A \cap (B \cup C) = \{1\}$.

Généralisation

Si A_1, A_2,..., A_n sont des sous-ensembles d'un ensemble E, on définit de même la réunion

$$A_1 \cup A_2 \cup \ldots \cup A_n$$

comme l'ensemble des x qui appartiennent à au moins l'un des ensembles A_1, A_2,...ou A_n,

et l'intersection

$$A_1 \cap A_2 \cap \ldots \cap A_n$$

comme l'ensemble des x qui appartiennent à tous les ensembles A_1, A_2,..., A_n:

$$A_1 \cup A_2 \cup \ldots \cup A_n = \{x \mid \exists i \in \{1, 2, \ldots, n\},\ x \in A_i\}$$
$$A_1 \cap A_2 \cap \ldots \cap A_n = \{x \mid \forall i \in \{1, 2, \ldots, n\},\ x \in A_i\}$$

1.1.4 Complémentaire d'un ensemble

Soit A un sous-ensemble d'un ensemble E.

L'ensemble

$$\{x \mid x \in E \text{ et } x \notin A\}$$

est appelé le complémentaire de A dans E et est noté $C_E A$ ou encore s'il n'y a pas d'ambiguïté sur E, CA, ${}^cA$, A^c ou $\bar{A}$.

Lois de De Morgan

Soient A et B deux sous-ensembles d'un ensemble E, on démontre facilement les propriétés suivantes:

- $C_E\,(C_E A) = A$
- $A \subset B$ si et seulement si $C_E B \subset C_E A$
- $C_E\,(A \cup B) = (C_E A) \cap (C_E B)$
- $C_E\,(A \cap B) = (C_E A) \cup (C_E B)$

1.1.5 Différence de deux ensembles

Soient A et B deux sous-ensembles d'un ensemble E.

On appelle différence de A et B l'ensemble noté $A \setminus B$ défini par:

$$A \setminus B = A \cap \overline{B}$$

c'est donc l'ensemble

$$\{x \in E \mid x \in A \text{ et } x \notin B \}.$$

1.1.6 Différence symétrique de deux ensembles

Soient A et B deux sous-ensembles d'un ensemble E.

On appelle différence symétrique de A et B l'ensemble noté $A \triangle B$ défini par:

$$A \triangle B = (A \setminus B) \cup (B \setminus A)$$

1.1.7 Partition

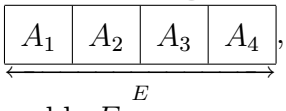
Soient A_1, A_2,..., A_n des sous-ensembles d'un ensemble E.

On dit que ces sous-ensembles forment une partition de E si les trois conditions suivantes sont vérifiées:

1) La réunion des A_i est égale à E : $E = A_1 \cup A_2 \cup \ldots \cup A_n$

2) Les A_i sont deux à deux disjoints : si $i, j \in \{1, 2, \ldots, n\}$ et $i \neq j$ alors $A_i \cap A_j = \varnothing$

3) Les A_i sont non vides : pour tout $i \in \{1, 2, \ldots, n\}$, $A_i \neq \varnothing$

Sur le dessin 

, les ensembles A_1, ..., A_4 forment une partition de l'ensemble E.

Exemples

- Soient $E = \mathbb{N}$, A_1 le sous- ensemble formé des entiers pairs, A_2 le sous-ensemble formé des entiers impairs.

Alors A_1 et A_2 forment une partition de E.

- Soient $E = \mathbb{R}$, $A_1 = \mathbb{R}_+^*$, $A_2 = \mathbb{R}_-^*$, $A_3 = \{0\}$. Alors A_1, A_2 et A_3 forment une partition de E.

1.1.8 Produit cartésien

Le produit cartésien de deux ensembles E et F, est l'ensemble noté $E \times F$ définit par

$E \times F = \{(x, y) \mid x \in E \text{ et } y \in F\}$

Formulaire

Soient A, B, C et D des ensembles, on a les formules suivantes dont les démonstrations sont laissées aux lecteurs:

- $A \times (B \cap C) = (A \times B) \cap (A \times C)$
- $A \times (B \cup C) = (A \times B) \cup (A \times C)$
- Si $(A \subset B)$ *et* $(C \subset D)$ alors $A \times C \subset B \times D$

1.2 Relation binaire sur un ensemble

1.2.1 Vocabulaire et notations usuelles

Définitions: Soient E et F deux ensembles.

On appelle relation binaire de E vers F un sous-ensemble de $E \times F$.

Soit $\mathcal{R}$ une relation binaire de E vers F, on dit qu'un couple $(x, y) \in E \times F$ vérifie la relation $\mathcal{R}$ si $(x, y) \in \mathcal{R}$ et on note alors $x\mathcal{R}y$.

On appelle relation binaire sur E une relation binaire de E vers E.

Définitions: Soit $\mathcal{R}$ une relation binaire sur un ensemble E. On dit que $\mathcal{R}$ est

· réflexive lorsque pour tout $x \in E$, on a $x\mathcal{R}x$

· symétrique lorsque pour tout couple $(x, y) \in E \times E$, $x\mathcal{R}y \Rightarrow y\mathcal{R}x$

· antisymétrique lorsque pour tout couple $(x, y) \in E \times E$,

$$(x\mathcal{R}y \text{ et } y\mathcal{R}x) \Rightarrow x = y$$

· transitive lorsque pour tout triplet $(x, y, z) \in E^3$,

$$(x\mathcal{R}y \text{ et } y\mathcal{R}z) \Rightarrow x\mathcal{R}z$$

1.2.2 Relations d'ordre

Définitions: Soit $\mathcal{R}$ une relation binaire sur un ensemble E. On dit que $\mathcal{R}$ est une relation d'ordre sur E lorsqu'elle est à la fois réflexive, antisymétrique et transitive. Une relation d'ordre $\mathcal{R}$ sur E est dite relation d'ordre total lorsque pour tout couple $(x, y) \in E^2$ on a $x\mathcal{R}y$ ou $y\mathcal{R}z$. Une relation d'ordre qui n'est pas d'ordre total est appelée relation d'ordre partiel.

Exemples:

· La relation d'inclusion " $\subset$ " est une relation d'ordre partiel sur $\mathcal{P}(E)$.

· La relation " divise " est une relation d'ordre partiel sur $\mathbb{N}$.

· La relation " $\leq$ " est une relation d'ordre total sur $\mathbb{R}$.

1.2.3 Relation d'équivalence

Définitions: Soit $\mathcal{R}$ une relation binaire sur un ensemble E. On dit que $\mathcal{R}$ est une relation d'équivalence sur E lorsqu'elle est à la fois réflexive, symétrique et transitive.

Si $\mathcal{R}$ est une relation d'équivalence sur E, on dit que les éléments x et y de E sont équivalents losque $x\mathcal{R}y$.

On appelle classe d'équivalence (modulo $\mathcal{R}$) d'un élément x de E, le sous-ensemble noté $\mathcal{R}(x)$ ou $\dot{x}$, des éléments de E qui sont équivalents à x:

$$\mathcal{R}(x) = \dot{x} = \{y \in E \mid y\mathcal{R}x\}.$$

Avec les notations ci-dessus, on appelle ensemble-quotient de E par $\mathcal{R}$, le sous-ensemble de $\mathcal{P}(E)$ noté $E/\mathcal{R}$ formé des classes d'équivalence des éléments de E.

Propriétés: Soit $\mathcal{R}$ une relation d'équivalence sur un ensemble E.

· Pour tout x de E, $\mathcal{R}(x)$ est non vide.

En effet, $\mathcal{R}$ étant réflexive, $x \in \mathcal{R}(x)$.

· L'ensemble des classes d'équivalence modulo $\mathcal{R}$ constitue une partition de E.

En effet, de $x \in \mathcal{R}(x)$, on déduit que $E = \underset{x\in E}{\cup}\{x\} \subset \underset{x\in E}{\cup}\mathcal{R}(x)$, et donc que $E = \underset{x\in E}{\cup}\mathcal{R}(x)$.

· On a $x\mathcal{R}y$ si et seulement si $\mathcal{R}(x) = \mathcal{R}(y)$.

En effet, lorsque $\mathcal{R}(x) \cap \mathcal{R}(y) \neq \varnothing$, il existe $z \in \mathcal{R}(x) \cap \mathcal{R}(y)$. On a alors $z \in \mathcal{R}(x)$ et $z \in \mathcal{R}(y)$; donc $z\mathcal{R}x$ et $z\mathcal{R}y$. La symétrie et la transitivité de $\mathcal{R}$ assurent que $x\mathcal{R}y$ et en fin $\mathcal{R}(x) = \mathcal{R}(y)$.

1.3 Applications

Une application (ou fonction) définie sur un ensemble X et à valeurs dans un ensemble Y est une loi qui, à tout élément de X fait correspondre un unique élément de Y. Si on note f cette application, l'élément associé à x par f est noté $f(x)$. On donne la notation suivante:

$$f : X \longrightarrow Y,\ x \longmapsto f(x)$$

X est l'ensemble de départ ou source, Y est l'ensemble d'arrivée ou but, $f(x)$ est l'image de x par f, x est l'antécédent de $f(x)$ par f.

1.3.1 Graphe

Une fonction $f : X \longrightarrow Y$ peut être définie par son graphe, un sous-ensemble Γ de $X \times Y$ qui possède la propriété suivante:

$\forall x \in X$, $\exists y \in Y$, $(x, y) \in \Gamma$ et de plus si $(x, y) \in \Gamma$ et $(x, y') \in \Gamma$ alors $y = y'$. On a

$$\Gamma = \{(x, y) \in X \times Y \mid y = f(x)\} = \{(x, f(x)) \mid x \in X\}.$$

1.3.2 Exemples

Soient X et Y deux ensembles.

1) L'application qui à tout $x \in X$ associe x s'appelle application identique et se note id_X.

2) Si $f : X \longrightarrow Y$ est une application et si $X' \subset X$, on peut définir f' la restrictionde f à X' par $\forall x \in X'$, $f'(x) = f(x)$.

3) Si $f : X \longrightarrow Y$ et g: $Y \longrightarrow Z$ sont deux applications, on peut définir la composée de f et g par $(g \circ f\ )(x) = g\,(f(x))$.

Une propriété importante de la composition des applications est l'associativité.

Injection

On dit qu'une application $f : X \longrightarrow Y$ est injective (ou que f est une injection) si tout élément de Y est l'image d'au plus un élément de X. En d'autres termes tout élément de Y a au plus un antécédent.

Caractérisation: Une application $f : X \longrightarrow Y$ est injective si et seulement si pour tout $(x, y) \in X^2$ l'égalité $f(x) = f(y)$ entraîne $x = y$.

Surjection

On dit qu'une application $f : X \longrightarrow Y$ est surjective (ou que f est une surjection) si tout élément de Y est l'image d'au moins un élément de X. En d'autres termes tout élément de Y a au moins un antécédent.

Caractérisation: Une application $f : X \longrightarrow Y$ est surjective si et seulement si pour tout $y \in Y$ il existe $x \in X$ tel que $y = f(x)$.

Bijection

On dit qu'une application $f : X \longrightarrow Y$ est bijective (ou que f est une bijection) si elle est à la fois injective et surjective. En d'autres termes tout élément de Y a exactement un antécédent.

On appelle bijection réciproque d'une bijection f et on note f^{-1} l'application caractérisée par $x = f^{-1}(y)$ si et seulement si $y = f(x)$.

1.3.3 Image directe, image réciproque

Soit $f : X \longrightarrow Y$ une application.

Définition. Si A est une partie de X, on appelle image directe de A par f et on note $f(A)$, l'ensemble

$$f(A) = \{y \in Y \mid \exists x \in A,\ f(x) = y\}$$

Définition. Si B est une partie de Y, on appelle image réciproque de B par f et on note $f^{-1}(B)$, l'ensemble

$$f^{-1}(B) = \{x \in X \mid \ f(x) \in B\}$$

Remarques

1) On prendra bien garde à ne pas confondre l'application

$f^{-1} : \mathcal{P}(Y) \longrightarrow \mathcal{P}(X)$ ainsi définie (qui existe pour toute fonction f) avec la bijection réciproque

$f^{-1} : Y \longrightarrow X$ (qui n'existe que si f est bijective).

2) On pourra vérifier en exercice que:
(i) f est surjective si et seulement si $Y = f(X)$.
(ii) f est injective si et seulement si $f : X \longrightarrow f(X)$ est une injection.

Formulaire

Soit $f : X \longrightarrow Y$ une application, on a les formules suivantes:
1) Pour toutes parties A, B de X on a
a) $f(A \cup B) = f(A) \cup f(B)$
b) $f(A \cap B) \subset f(A) \cap f(B)$
c) Si $A \subset B$ alors $f(A) \subset f(B)$
2) Pour toutes parties A, B de Y on a
a) $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$
b) $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$
c) Si $A \subset B$ alors $f^{-1}(A) \subset f^{-1}(B)$
d) $f^{-1}(C_Y A) = C_X(f^{-1}(A))$

Démonstration:

1) a) $y \in f(A \cup B) \Longleftrightarrow y = f(x)$ avec $x \in A \cup B$
$\Longleftrightarrow y = f(x)$ avec $x \in A$ ou $x \in B$
$\Longleftrightarrow y \in f(A)$ ou $y \in f(B)$
$\Longleftrightarrow y \in f(A) \cup f(B)$
Ainsi on a bien $f(A \cup B) = f(A) \cup f(B)$.
b) $y \in f(A \cap B) \Longleftrightarrow y = f(x)$ avec $x \in A \cap B$
$\Longleftrightarrow y = f(x)$ avec $x \in A$ et $x \in B$
$\Longrightarrow y \in f(A)$ et $y \in f(B)$ (la réciproque de cette implication n'est vraie que si f est injective)
$\Longleftrightarrow y \in f(A) \cap f(B)$
Ainsi on a bien $f(A \cap B) \subset f(A) \cap f(B)$, il y a égalité si f est injective.
L'exemple suivant montre qu'on a pas en général égalité:
Prenons E= $\{a, b\}$, F= $\{c\}$, $f(a) = f(b) = c$, $A = \{a\}$ et $B = \{b\}$. Alors $\varnothing = f(A \cap B) \neq f(A) \cap f(B) = \{c\}$.
c) $y \in f(A) \Longrightarrow y = f(x)$ avec $x \in A$
$\Longrightarrow y = f(x)$ avec $x \in B$
$\Longrightarrow y \in f(B)$.
Ainsi on a bien si $A \subset B$ alors $f(A) \subset f(B)$
2) a) $x \in f^{-1}(A \cup B) \Longleftrightarrow f(x) \in A \cup B$
$\Longleftrightarrow f(x) \in A$ ou $f(x) \in B$
$\Longleftrightarrow x \in f^{-1}(A)$ ou $x \in f^{-1}(B)$

$$\Longleftrightarrow x \in f^{-1}(A) \cup f^{-1}(B).$$

Ainsi on a bien $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$.

b) $x \in f^{-1}(A \cap B) \Longleftrightarrow f(x) \in A \cap B$
$\Longleftrightarrow f(x) \in A$ et $f(x) \in B$
$\Longleftrightarrow x \in f^{-1}(A)$ et $x \in f^{-1}(B)$
$\Longleftrightarrow x \in f^{-1}(A) \cap f^{-1}(B)$.

Ainsi on a bien $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$.

c) $x \in f^{-1}(A) \Longrightarrow f(x) \in A$
$\Longrightarrow f(x) \in B$
$\Longrightarrow x \in f^{-1}(B)$.

Ainsi on a bien si $A \subset B$ alors $f^{-1}(A) \subset f^{-1}(B)$.

d) $x \in f^{-1}(C_Y A) \Longleftrightarrow f(x) \in C_Y A$
$\Longleftrightarrow f(x) \notin A$
$\Longleftrightarrow \urcorner (f(x) \in A)$
$\Longleftrightarrow \urcorner (x \in f^{-1}(A))$
$\Longleftrightarrow x \notin f^{-1}(A)$
$\Longleftrightarrow x \in C_X(f^{-1}(A))$.

Ainsi on a bien $f^{-1}(C_Y A) = C_X(f^{-1}(A))$.

CQFD

1.4 Fonction caractéristique d'un sous-ensemble

1.4.1 Notions de base

Définition: Soit A un sous-ensemble d'un ensemble E. On appelle fonction caractéristique de A l'application notée φ_A définie par

$$\varphi_A : E \longrightarrow \mathbb{R},\ \varphi_A(x) = \begin{cases} 1 \text{ quand } x \in A \\ 0 \text{ quand } x \notin A \end{cases}$$

Remarques:

· φ_A appartient au sous-ensemble $\{0, 1\}^E$ des applications de E dans $\{0, 1\}$.

· $\varphi_A = \varphi_B \Longleftrightarrow A = B$, ce qui justifie l'expression " fonction caractéristique " d'un sous-ensemble.

Propriétés:

· L'application $A \mapsto \varphi_A$ de $\mathcal{P}(E)$ dans $\{0, 1\}^E$ est une bijection.

· Quels que soient les sous-ensembles A et B de E, on a:
(i) $\varphi_A\varphi_A = \varphi_A$
(ii) $\varphi_{A\cap B} = \varphi_A\varphi_B$
(iii) $\varphi_{A\cup B} = \varphi_A + \varphi_B - \varphi_A\varphi_B$; si $A \cap B = \varnothing$ alors $\varphi_{A\cup B} = \varphi_A + \varphi_B$.
(iv) $\varphi_{\overline{A}} = 1 - \varphi_A$
(v) $\varphi_{A\setminus B} = \varphi_A - \varphi_A\varphi_B = \sup(0,\ \varphi_A - \varphi_B)$
(vi) $\varphi_{A\triangle B} = \varphi_A + \varphi_B - 2\varphi_A\varphi_B$
(vii) $A \subset B \Longleftrightarrow \varphi_A(x) \leq \varphi_B(x),\ \forall x \in E$.

Démonstration:
Par exemple:

$$\begin{aligned}\varphi_{A\setminus B} &= \varphi_{A\cap\overline{B}} \\ &= \varphi_A\varphi_{\overline{B}} \quad \text{(d'après (ii))} \\ &= \varphi_A\left(1-\varphi_B\right) \quad \text{(d'après (iv))} \\ &= \varphi_A - \varphi_A\varphi_B.\end{aligned}$$

La démonstration des autres propriétés est laissée en exercices.

1.4.2 Exercices d'application

Exercice 1: Trouver une condition nécessaire et suffisante pour que les sous-ensembles A, B, C de E vérifient:

$$\begin{cases} A \cap B = A \cap C \\ A \cup B = A \cup C \end{cases}$$

Solution:

$$\begin{cases} A \cap B = A \cap C \\ A \cup B = A \cup C \end{cases} \Longleftrightarrow \begin{cases} \varphi_A\varphi_B = \varphi_A\varphi_C \\ \varphi_A + \varphi_B - \varphi_A\varphi_B = \varphi_A + \varphi_C - \varphi_A\varphi_C \end{cases}$$
$$\Longleftrightarrow \varphi_B = \varphi_C$$
$$\Longleftrightarrow B = C$$

Exercice 2: Soient A, B, C trois sous-ensembles d'un ensemble E. Montrer que

$$A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$$

Solution:

$$\begin{aligned}\varphi_{(A\cap B)\setminus(A\cap C)} &= \varphi_{A\cap B} - \varphi_{A\cap B}\varphi_{A\cap C} \\ &= \varphi_A\varphi_B - \varphi_A\varphi_B\varphi_A\varphi_C \\ &= \varphi_A\varphi_B - \varphi_A\varphi_B\varphi_C \\ &= \varphi_A\left[\varphi_B\left(1-\varphi_C\right)\right]\end{aligned}$$

$$\begin{aligned} &= \varphi_A \varphi_{B \setminus C} \\ &= \varphi_{A \cap (B \setminus C)} \end{aligned}$$

Ce qui montre que $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$.

Exercices d'application. Démontrer, en utilisant les fonctions caractéristiques, les propriétés suivantes:

· La fonction caractéristique de $A \triangle B$ est:

$$\begin{aligned} \varphi_{A \triangle B} &= \varphi_A + \varphi_B - 2\varphi_A \varphi_B \\ &= (\varphi_A - \varphi_B)^2 \\ &= |\varphi_A - \varphi_B| \end{aligned}$$

· La différence symétrique est commutative, associative, admet $\varnothing$ pour élément neutre et $A \triangle A = \varnothing$ pour tout $A \in \mathcal{P}(E)$.

· $A \triangle B = (A \cup B) \setminus (A \cap B)$ pour tout couple (A, B).

· $A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C)$ pour tout triplet (A, B, C).

Chapter 2

Eléments de logique

Ce chapitre est consacré essentiellement à la compréhension d'un énoncé, car un énoncé est parfois difficile à comprendre. Il peut contenir des mots utilisés dans un sens différent du sens courant, ou des mots spécifiques au langage mathématique. S'il n'est pas écrit avec assez de soin il peut devenir ambigu.

2.1 Vocabulaire de la logique

2.1.1 Langage

Le langage mathématique utilise à la fois le langage courant convenablement précisé et des signes mathématiques quelques fois tirés de l'usage courant (lettres, parenthèses, crochets, accolades, etc) ou spécifiquement mathématiques ($=$, $\leq$, $\cap$, $\times$, etc).

Le langage mathématique est essentiellement un langage symbolique. Un assemblage formé avec des mots du langage courant et des signes mathématiques n'a pas forcément un sens. Nous dirons que dans un contexte donné un assemblage est cohérent si dans ce contexte on peut lui donner une signification mathématique. Dans la suite nous ne considèrerons que des assemblages cohérents. On peut distinguer:

- les termes qui sont les mots du langage courant,

- les énoncés qui sont en quelque sorte les phrases du langage courant.

Exemples*:*

2, π, $6x$ sont des termes.

$\sqrt{3}$ n'est pas un nombre rationnel, $3 \leq 9$ sont des énoncés.

2.1.2 Classification des énoncés

Il n'est pas immédiat qu'un énoncé donné coïncide avec l'idée intuitive de cet énoncé. Une des tâches essentielles en mathématique est donc de chercher à s'assurer que tel ou tel énoncé est vrai ou faux.

Un énoncé mathématique (nous dirons simplement énoncé) est une phrase ayant un sens mathématique précis; par exemple

(A): $2 = 0$

(B): Pour tout nombre réel x on a $x^2 \geq 0$

(C): x est un entier et $x < 4$

sont des énoncés; (A) est faux, (B) est vrai, la véracité de (C) dépend de la valeur de la variable x. Par contre, " les fraises sont des fruits délicieux " n'est pas un énoncé mathématique, car la réponse est un jugement subjectif (certains trouvent les fraises délicieuses, d'autres pas). Nous ne chercherons pas à définir précisément la différence entre énoncé mathématique et énoncé non mathématique. Confronté à un énoncé, un mathématicien souhaite pouvoir décider dans quelles conditions il est vrai ou faux. Cet énoncé peut contenir une ou plusieurs variables.

Une proposition est un énoncé dont on peut décider s'il est vrai ou faux lorsque (si nécessaire) toutes les variables ont été remplacées par des valeurs connues; par exemple (A), (B) et (C) sont des propositions:

(A) est une proposition fausse, (B) une proposition vraie, (C) est une proposition vraie lorsque $x = 1$ et fausse lorsque $x = 8$.

Une conjecture est un énoncé dont la véracité n'est pas démontrée; par exemple, Pierre de Fermat a énoncé vers 1640 que:

" si n est un entier strictement supérieur à 2, il n'existe pas d'entiers non nuls a, b, c tels que $a^n + b^n = c^n$". Le sens de cet énoncé qui porte le nom de théorème de Fermat, est parfaitement clair, mais il a cependant fallu attendre plus de 3 siècles pour être sûr qu'il était vrai, grâce à la démonstration d' Andrew Wiles (1994).

Une assertion est un énoncé pour lequel on peut répondre sans ambiguïté et sans renseignement complémentaire à la question est-il vrai ou est-il faux? Par exemple (A) et (B) sont des assertions, par contre (C) n'est pas une assertion: pour répondre il faut préciser la valeur de x. L'énoncé (C) effectue une classification dans $\mathbb{N}$: les éléments pour lesquels $x < 2$ est un énoncé vrai, et tous les autres pour lesquels il est faux. Un tel énoncé est parfois appelé **prédicat**.

N.B: certains emploient proposition avec le sens du mot assertion.

Une hypothèse est un énoncé sur lequel on construit le point de départ de la démonstration.

Un axiome est une hypothèse considérée vraie à priori (i.e par convention).

Un théorème est un énoncé vrai en mathématique qui peut être paraphrasé de la manière suivante: " sous les hypothèses suivantes:..., la chose suivante:... est toujours vraie".

Un lemme est un théorème dont la démonstration précède et rend possible celle d'un théorème plus général.

Un corollaire est un théorème immédiatement déduit d'un autre théorème plus général.

2.2 Connecteurs - tables de vérité

Si une proposition P est vraie on dit que sa valeur de vérité est V (ou 1) et l'on note val(P)=V (ou val(P)=1); dans le cas contraire on dit que la valeur de vérité de P est F (ou 0) et l'on note val(P)=F (ou val(P)=0).

A partir de certaines propositions, on peut à l'aide des connecteurs logiques en fabriquer d'autres définies par leurs valeurs de vérité. Les démonstrations "intéressantes" en mathématiques sont souvent longues. Une manière simple (mais fastidieuse) de montrer qu'une proposition est vraie ou fausse, est de faire une table dite table de vérité avec les diverses possibilités: chaque proposition est vraie ou fausse. Par exemple, pour une proposition construite à partir de 2 autres propositions différentes, il y a 4 possibilités. Plus généralement, pour une proposition construite à partir de n autres propositions différentes, il y a 2^n possibilités. les connecteurs logiques les plus fréquemment utilisés sont: la négation, la conjonction, la disjonction, l'implication, l'équivalence.

2.2.1 Négation

Soit P une proposition. La négation de P est une proposition notée $\neg$P ou $\overline{P}$ et on lit " non P ". Elle est vraie lorsque P est fausse et fausse lorsque P est vraie. Ainsi on a val($\neg$P)=V lorsque val(P)=F, et val($\neg$P)=F lorsque val(P)=V. La négation est définie par la table de vérité suivante:

P	⌉P
V	F
F	V

La négation est un connecteur unaire: portant sur une proposition.

Exemples: 1) Soit x un réel. La proposition $(\rceil (x = 1))$ prend les mêmes valeurs de vérité que la proposition $(x \neq 1)$.

2) Soit n un entier naturel. La proposition $(\rceil (n \leq 3))$ prend les mêmes valeurs de vérité que la proposition $(n > 3)$

2.2.2 Conjonction et disjonction

Soient P et Q deux propositions. La conjonction de P et Q est une proposition notée P∧Q et on lit " P et Q ". Elle est vraie lorsque P et Q sont vraies, et fausse lorsque l'une au moins des propositions P et Q est fausse. Ainsi val(P∧Q)=V dans le seul cas où on a val(P)=V et val(Q)=V.

La disjonction de P et Q est une proposition notée P∨Q et on lit " P ou Q " . Elle est vraie lorsque l'une au moins des propositions P ou Q est vraie, et elle est fausse sinon. Ainsi val(P∨Q)=F dans le seul cas ou on a val(P)=F et val(Q)=F. La conjonction et la disjonction sont des connecteurs binaires; leur table de vérité est:

P	Q	P∧Q	P∨Q
V	V	V	V
V	F	F	V
F	V	F	V
F	F	F	F

N.B: Dans le langage courant, le mot " ou " peut prendre deux sens différents: le sens inclusif qui est celui donné plus haut et le sens exclusif, qu'on précise parfois par " ou bien ": (P ou bien Q) est vraie si l'une des deux propositions est vraie et l'autre est fausse. Lorsque P et Q sont simultanément vraies, la proposition (P∨Q) est vraie, mais (P ou bien Q) est fausse. Nous n'utiliserons que le " ou " inclusif.

2.2.3 Implication

Soient P et Q deux propositions. Alors P$\Longrightarrow$Q est une proposition; elle est fausse lorsque P est vraie et Q est fausse, elle est vraie dans tous les autres cas. Ainsi val(P$\Longrightarrow$Q)=F dans le seul cas où on a val(P)=V et val(Q)=F. On lit:

" P implique Q "
" si P, alors Q "
" P entraîne Q "
" pour que P, il faut que Q "
" pour que Q, il suffit que P "
" une condition nécessaire pour que P est que Q "
" une condition suffisante pour que Q est que P ".

L'implication est définie par la table de vérité suivante:

P	Q	P$\Longrightarrow$Q
V	V	V
V	F	F
F	V	V
F	F	V

Dans P$\Longrightarrow$Q, P est l'antécédent et Q le conséquent. La proposition Q$\Longrightarrow$P est la réciproque (ou converse) de P$\Longrightarrow$Q. La proposition $\urcorner$Q$\Longrightarrow\urcorner$P est la contraposée de P$\Longrightarrow$Q.

N.B: Dans le langage courant, on emploie souvent " il faut " à la place de " il suffit ". Par exemple, on dit " pour traverser la rivière, il faut prendre le bateau ", alors que c'est en fait suffisant, mais pas nécessaire. On peut le faire à la nage.

2.2.4 Equivalence

Soient P et Q deux propositions. Alors P$\Longleftrightarrow$Q est une proposition. Elle est vraie lorsque Pet Q sont toutes les deux vraies ou toutes les deux fausses. Elle est fausse dans les autres cas. Ainsi val(P$\Longleftrightarrow$Q)=V lorsque val(P)=val(Q).

La proposition P$\Longleftrightarrow$Q s'énonce:

" P et Q sont équivalentes "

" P si et seulement si Q "
" pour que P, il faut et il suffit que Q "
" une condition nécessaire et suffisante pour que Q est que P ".
L'équivalence est définie par la table de vérité suivante:

P	Q	P$\Longleftrightarrow$Q
V	V	V
V	F	F
F	V	F
F	F	V

N.B: Dans P$\Longleftrightarrow$Q, les deux propositions sont interchangeables; " P$\Longleftrightarrow$Q " signifie que " P$\Longrightarrow$Q " et " Q$\Longrightarrow$P " sont vraies. Aussi affirmer " P$\Longleftrightarrow$Q" demande deux justifications, une pour chaque sens: la condition nécessaire qui correspond à l'implication directe P$\Longrightarrow$Q, et la condition suffisante qui correspond à l'implication indirecte P$\Longleftarrow$Q. Ainsi le connecteur " $\Longleftrightarrow$ " est appelé équivalence logique ou double implication ou condition nécessaire et suffisante.

la proposition P$\Longleftrightarrow$P est toujours vraie: on dit que c'est une tautologie. Une proposition toujours fausse est appelée une contradiction ou une antilogie.

Dire que deux propositions sont équivalentes, c'est dire que ces propositions ont mêmes valeurs de vérité i.e qu'elles signifient la même chose.

2.3 Quantificateurs

Pour désigner une proposition qui contient une variable x, on adopte souvent une notation de la forme P(x) pour en faciliter la lecture et la compréhension. Par exemple pour n un entier naturel, la proposition " n est pair " pourra être notée P(n).

Soit P(x) une proposition qui contient une variable x, où x désigne un élément d'un ensemble E. La valeur de vérité de P(x) peut dépendre de x et il est souvent important de savoir si elle est vraie pour tous les éléments x de E, pour au moins un, pour un et un seul etc. Les mathématiciens ont introduit des symboles pour exprimer ces idées; on les appelle des quantificateurs.

2.3.1 Différentes sortes de quantificateurs

Soit P(x) une proposition qui contient une variable x, où x désigne un élément d'un ensemble E.

La proposition " quelque soit x de E, on a P(x) " notée ($\forall x \in$ E, P(x)) est vraie si P(x) est vraie pour tous les éléments x de E; elle est fausse sinon. On lit aussi " pour tout x de E, on a P(x) ".

La proposition " il existe au moins un x de E tel que P(x) " notée ($\exists x \in$ E, P(x)) est vraie s'il existe au moins un élément x de E tel que P(x) soit vraie; elle est fausse sinon.

La proposisition " il existe un et un seul x de E tel que P(x) " notée ($\exists ! x \in$ E, P(x)) est vraie s'il existe un élément x de E et un seul tel que P(x) soit vraie; elle est fausse sinon.

Les quantificateurs permettent donc de transformer un énoncé contenant une variable en un énoncé "absolu ". Nous utiliserons exclusivement:

- le quantificateur universel (en symbole $\forall$)
- le quantificateur existentiel (en symbole $\exists$).

2.3.2 Quelques remarques

Dans le langage courant l'expression " il existe un x tel que... " est ambiguë, car elle signifie parfois " il existe exactement un x tel que... ", (ce qui correspond au symbole $\exists !$). C'est pourquoi, on préfère préciser " il existe au moins un x tel que... " ou " il existe un et un seul x tel que... ". On rencontre aussi l'expression " il existe au plus un x tel que... ".

Lorsqu'une proprosition contient des quantificateurs $\forall$ et $\exists$, leur ordre est essentiel. Par exemple, la proposition

$$(\forall x \in \mathbb{R},\ (\exists y \in \mathbb{R},\ y = x^2))$$

est vraie, mais la proposition

$$(\exists x \in \mathbb{R},\ (\forall y \in \mathbb{R},\ y = x^2))$$

est fausse.

Les propositions ($\forall x \in$ E, P(x)) et ($\exists x \in$ E, P(x)) ne traduisent pas des propriétés de x, mais de l'ensemble E; on dit que x est une variable muette. On obtient une proposition équivalente en remplaçant partout x par une autre lettre. Mais il est plus prudent de choisir une lettre qui n'a pas déjà été utilisée.

2.4 Négation d'une proposition

1) Soient P et Q des propositions. On peut montrer, à l'aide des tables de vérité par exemple, que les équivalences suivantes sont toujours vraies:

$\rceil(\rceil P) \Longleftrightarrow P$

$\rceil(P \wedge Q) \Longleftrightarrow \rceil(P) \vee \rceil(Q)$

$\rceil(P \vee Q) \Longleftrightarrow \rceil(P) \wedge \rceil(Q)$

$\rceil(P \Longrightarrow Q) \Longleftrightarrow P \wedge \rceil Q$

Preuve: Montrons par exemple que l'équivalence

$$\rceil(P \Longrightarrow Q) \Longleftrightarrow P \wedge \rceil Q$$

est toujours vraie.

Pour faciliter l'écriture dans la table de vérité, on pose

$$A = \rceil(P \Longrightarrow Q) \Longleftrightarrow P \wedge \rceil Q$$

P	Q	$\rceil Q$	$P \wedge \rceil Q$	$P \Longrightarrow Q$	$\rceil(P \Longrightarrow Q)$	A
V	V	F	F	V	F	V
V	F	V	V	F	V	V
F	V	F	F	V	F	V
F	F	V	F	V	F	V

La table de vérité nous montre que $\rceil(P \Longrightarrow Q) \Longleftrightarrow P \wedge \rceil Q$ est toujours vraie (cf colonne A). On peut aussi se limiter à la table de vérité sans la colonne A, et remarquer que les propositions $\rceil(P \Longrightarrow Q)$ et $P \wedge \rceil Q$ ont mêmes valeurs de vérité (données dans le même ordre) i.e qu'elles sont équivalentes.

CQFD

Exercice d'application:

Soient P et Q des propositions.

a) Montrer que les équivalences suivantes sont toujours vraies:

$(P \Longrightarrow Q) \Longleftrightarrow (\rceil P \vee Q) \Longleftrightarrow (\rceil Q \Longrightarrow \rceil P)$.

b) Utiliser a) pour construire deux phrases synonymes à la phrase " s'il est français alors il porte un béret ".

Contruire la négation de la phrase " s'il est français alors il porte un béret ".

Solution:

a)

P	Q	⌉P	⌉Q	P $\Longrightarrow$ Q	(⌉P $\vee$ Q)	⌉Q $\Longrightarrow$⌉P
V	V	F	F	V	V	V
V	F	F	V	F	F	F
F	V	V	F	V	V	V
F	F	V	V	V	V	V

La table de vérité montre que les propositions (P $\Longrightarrow$ Q), (⌉P $\vee$ Q) et (⌉Q $\Longrightarrow$⌉P) ont mêmes valeurs de vérité (données dans le même ordre) i.e qu'elles sont équivalentes.

b) Posons P: " il est français " et Q: " il porte un béret ". Ainsi, la phrase " s'il est français alors il porte un béret " se symbolise par (P $\Longrightarrow$ Q), et d'après a) cette phrase doit être synonyme aux phrases symbolisées par (⌉P $\vee$ Q) et (⌉Q $\Longrightarrow$⌉P). On a alors les phrases suivantes qui sont synonymes:

(P $\Longrightarrow$ Q): " s'il est français alors il porte un béret "

(⌉P $\vee$ Q): " il n'est pas français ou il porte un béret "

(⌉Q $\Longrightarrow$⌉P): "s'il ne porte pas de béret alors il n'est pas français "

On sait que ⌉(P $\Longrightarrow$ Q) $\Longleftrightarrow$P$\wedge$⌉Q, donc la négation de la phrase " s'il est français alors il porte un béret " est:

P$\wedge$⌉Q: " il est français et il ne porte pas de béret ".

2) Soit P(x) une proposition qui contient une variable x, où x désigne un élément d'un ensemble E. On a les équivalences suivantes:

⌉($\forall x \in$ E, P(x)) $\Longleftrightarrow$ ($\exists x \in$ E, ⌉P(x))

⌉($\exists x \in$ E, P(x)) $\Longleftrightarrow$ ($\forall x \in$ E, ⌉P(x))

N.B: Les règles précédentes peuvent se combiner. Soit par exemple la proposition

P(n): $\exists l \in \mathbb{R}$, $\forall \varepsilon \in \mathbb{R}_+^*$, $\exists N \in \mathbb{N}$, $\forall n \in \mathbb{N}$, ($n \geq N \Longrightarrow |u_n - l| < \varepsilon$).

Alors

⌉P(n): $\forall l \in \mathbb{R}$, $\exists \varepsilon \in \mathbb{R}_+^*$, $\forall N \in \mathbb{N}$, $\exists n \in \mathbb{N}$, ($n \geq N$ et $|u_n - l| \geq \varepsilon$).

2.5 Quelques types classiques de raisonnement

Quand on rédige une démonstration, la proposition que l'on cherche à démontrer s'appelle la conclusion. Au cours de la démonstration, on utilise des résultats déjà connus, par exemple, des définitions, des théorèmes ou des propositions déjà démontrés dans le cours; mais aussi les prémices offerts

par l'énoncé ou des résultats intermédiaires que l'on a obtenus préalablement. Toutes ces propositions s'appellent des données. Une des raisons de la compexité des démonstrations est que l'on peut être amené pendant la démonstration à ajouter provisoirement de nouvelles données qui ne font pas partie de celles dont nous venons de parler et aussi de nouveaux objets qui ne sont pas décrits dans l'énoncé. Leur introduction est signalée par des expressions comme " Supposons que...", " Soit...". Il faut noter qu'une démonstration ne contient pas que des propositions vraies. L'exemple le plus évident en est celui des raisonnements par l'absurde où l'on ajoute aux données une proposition ($\neg$P) afin de montrer que c'est P qui est vraie. Dans ce paragraphe, on essaiera d'expliquer tout ceci, en explicitant les principales règles d'écriture sous-jacentes.

2.5.1 Règle de l'hypothèse auxiliaire

Cette règle s'applique aux énoncés du type " montrer que si l'on a P, alors on a Q ", c'est-à-dire " démontrer (P $\Longrightarrow$ Q) ". Dans la pratique, on commence souvent en écrivant " on suppose P ", c'est-à-dire qu'on considère P comme une nouvelle donnée, puis on cherche à démontrer Q. La démonstration s'achève par une phrase du genre " on a montré Q " et il peut être bon de rappeler la conclusion globale " on a montré (P $\Longrightarrow$ Q) ". On peut expliciter cette démarche par la règle suivante:

Pour démontrer (P $\Longrightarrow$ Q):

1) on ajoute P aux données,

2) on démontre que Q est vraie.

Exemple: Soit n un entier. Montrer que si n est pair, alors n^2 est pair.

Commentaire: Soit P la proposition " n est pair " et Q la proposition " n^2 est pair ". La proposition à démontrer est (P $\Longrightarrow$ Q).

Démonstration:

On suppose que n est pair. Ainsi on a

n est pair $\Longrightarrow n = 2k$ avec $k \in \mathbb{Z}$

$\Longrightarrow n^2 = 2t$ avec $t = 2k^2$

donc n^2 est pair. On a montré que " si n est pair, alors n^2 est pair ".

CQFD

2.5.2 Règle du quelque soit

Cette règle s'applique aux énoncés du type $(\forall x \in \mathrm{E},\ \mathrm{P}(x))$. Dans la pratique, on commence souvent en écrivant une expression du style " Soit x un élément de E ", ou " Considérons un élément x de E ", puis on écrit une démonstration de $\mathrm{P}(x)$. On peut expliciter cette démarche par la règle suivante:

Pour démontrer $(\forall x \in \mathrm{E},\ \mathrm{P}(x))$:

1) on ajoute la donnée $x \in$E.

2) on démontre que $\mathrm{P}(x)$ est vraie.

Exemple: Montrer que pour tout x de $\mathbb{R}$, on a $(x \leq 2 \Longrightarrow 3x + 1 \leq 7)$.

Commentaire: Ici $\mathrm{P}(x)$ est la proposition $(x \leq 2 \Longrightarrow 3x + 1 \leq 7)$ qui est une implication de la forme $(\mathrm{P}_1 \Longrightarrow \mathrm{Q}_1)$. Pour la démontrer on utilise la règle de l'hypothèse auxiliaire, en ajoutant P_1 aux données et on démontre Q_1. Ainsi " x réel " et " $x \leq 2$ " sont ajoutées aux données, on peut les condenser en écrivant " Soit x un réel tel que $x \leq 2$ ".

Démonstration: Soit x un réel tel que $x \leq 2$. Alors on a:

$x \leq 2 \Longrightarrow 3x \leq 6$

$\Longrightarrow 3x + 1 \leq 7$

donc $3x + 1 \leq 7$.

On a montré que pour tout x de $\mathbb{R}$, on a $(x \leq 2 \Longrightarrow 3x + 1 \leq 7)$.

CQFD

2.5.3 Règle des cas

Pour démontrer une proposition P, on peut vouloir se servir d'une proposition déjà démontrée de la forme (A ou B). Pour cela:

1) on suppose A vraie et on démontre P;

2) on suppose B vraie (sans hypothèse sur A) et on démontre P.

La règle peut être annoncée au début par une phrase du genre " deux cas se présentent ", ou mise en évidence par deux paragraphes, un pour chaque cas; le point 1) se traduit par une expression du genre " Supposons d'abord A ", le point 2) par " supposons maintenant B " ou " dans le cas où B ". Souvent B n'est autre que la proposition $\rceil$A.

Exemple: Soient x, y, λ des nombres réels, avec $\lambda < 0$. Montrer que l'on a $\max(\lambda x,\ \lambda y) = \lambda \min(x,\ y)$.

Commentaire: On peut considérer deux cas où la proposition A est " $x \leq y$ " et la proposition B est " $x > y$ " qui est équivalentes à $\rceil$A.

Démonstration: Deux cas se présentent: le cas $x \leq y$ et le cas $x > y$;

Supposons d'abord $x \leq y$. Alors comme $\lambda < 0$, on a:
$x \leq y \Longrightarrow \lambda x \geq \lambda y$
$\Longrightarrow \max(\lambda x, \lambda y) = \lambda x$
$\Longrightarrow \max(\lambda x, \lambda y) = \lambda \min(x, y)$
Supposons maintenant $x > y$. Alors comme $\lambda < 0$, on a:
$x > y \Longrightarrow \lambda x < \lambda y$
$\Longrightarrow \max(\lambda x, \lambda y) = \lambda y$
$\Longrightarrow \max(\lambda x, \lambda y) = \lambda \min(x, y)$.
Dans les deux cas, on a l'égalité demandée.
CFDQ

2.5.4 Nommer un objet

Pour démontrer une proposition Q en utilisant une donnée de la forme $(\exists x \in \mathrm{E}, \mathrm{P}(x))$:

1) on choisit une lettre x_0 par exemple, qui n'a pas déjà été utilisée et surtout qui n'apparait pas dans la proposition Q;

2) on ajoute aux hypothèses la proposition ($x_0 \in$ E et $\mathrm{P}(x_0)$);

3) on démontre Q.

En pratique, les points 1) et 2) sont traduits par une expression de la forme " Soit x_0 un élément de E tel que $\mathrm{P}(x_0)$ " ou " On sait que $(\exists x \in \mathrm{E}, \mathrm{P}(x))$; soit x_0 un tel élément ". Il est souvent bon de justifier le " on sait que " par la référence à une définition ou à un théorème.

Exemple: Soit a un réel. Montrer que si $\big([0, 1] \cap \big[a - \frac{1}{2}, a + \frac{1}{2}\big] \neq \emptyset\big)$, alors $\frac{-1}{2} \leq a \leq \frac{3}{2}$.

Commentaire: Ici E= $\big([0, 1] \cap \big[a - \frac{1}{2}, a + \frac{1}{2}\big] \neq \emptyset\big)$. Dire que E est non vide, c'est dire $(\exists x \in \mathrm{E})$; on utilise la règle " nommer un objet " en considérant l'un des éléments de E (il peut y en avoir beaucoup) et en l'appelant x_0.

Démonstration: On suppose que E est non vide. Soit x_0 un élément de E. On a alors
$x_0 \in$E $\Longrightarrow 0 \leq x_0 \leq 1$ et $a - \frac{1}{2} \leq x_0 \leq a + \frac{1}{2}$
$\Longrightarrow 0 \leq x_0 \leq a + \frac{1}{2}$ et $a - \frac{1}{2} \leq x_0 \leq 1$
$\Longrightarrow 0 \leq a + \frac{1}{2}$ et $a - \frac{1}{2} \leq 1$
$\Longrightarrow -\frac{1}{2} \leq a \leq \frac{3}{2}$.
On a bien le résultat demandé.
CQFD

2.5.5 Raisonnement par l'absurde

Pour démontrer une proposition P, à partir de certaines hypothèses, on suppose que P est fausse et l'on essaie d'arriver avec les hypothèses à une proposition absurde (contradiction avec les hypothèses).

Dans la pratique, on peut d'abord annoncer " raisonnons par l'absurde " et on commence le raisonnement par " supposons non P ", et la fin du raisonnement est indiquée par une expression comme " il y a contradiction " ou " c'est impossible " ou " ce qui est absurde ". On conclut en écrivant " on a donc P ". La démonstration contient souvent des phrases au conditionnel.

Exemple: Montrer que 0 n'est pas racine de $A(x) = x^4 + 12x - 1$.

Commentaire: La proposition P est ici

$$\text{" 0 n'est pas racine de } A(x) = x^4 + 12x - 1 \text{ ".}$$

Démostration: Raisonnons par l'absurde. Supposons que 0 soit racine de $A(x)$. Par définition, on aurait donc $A(0) = 0$; or le calcul montre que $A(0) = -1$, d'où $-1 = 0$; ce qui est absurde. On a donc montré que 0 n'est pas racine de $A(x) = x^4 + 12x - 1$.

CQFD

2.5.6 Raisonnement par contraposition

Pour démontrer $(\text{P} \Longrightarrow \text{Q})$, on peut démontrer sa contraposée $(\rceil\text{Q} \Longrightarrow \rceil\text{P})$ en lui appliquant la règle de l'hypothèse auxiliaire. On peut annoncer au début " On va raisonner par contraposition ".

Exemple: Soit x un nombre réel. Montrer que $(x^3 = 2 \Longrightarrow x < 2)$.

Commentaire: Ici, P est la proposition $(x^3 = 2)$ et Q la proposition $(x < 2)$. La contraposée est donc la proposition $(x \geq 2 \Longrightarrow x^3 \neq 2)$, on lui applique la règle de l'hypothèse auxiliaire.

Démonstration: On va raisonner par contraposition. Soit x un réel tel que $x \geq 2$. On a alors

$x \geq 2 \Longrightarrow x^3 \geq 8$

$\Longrightarrow x^3 \neq 2$

On a donc $x^3 \neq 2$. On a montré que $(x \geq 2 \Longrightarrow x^3 \neq 2)$. Ainsi, on a bien $(x^3 = 2 \Longrightarrow x < 2)$.

CQFD

2.5.7 Démonstration d'une équivalence

Pour démontrer (P $\Longleftrightarrow$ Q):

- On peut démontrer d'abord (P $\Longrightarrow$ Q), puis (Q $\Longrightarrow$ P).

- On peut aussi démontrer d'abord (P $\Longrightarrow$ Q), puis ($\urcorner$P $\Longrightarrow\urcorner$Q).

Exemple: Montrer que $\forall n \in \mathbb{N}$,(n pair $\Longleftrightarrow n^2$ est divisible par 4).

Commentaire: Ici, P est la proposition " n pair " et Q la proposition " n^2 est divisible par 4 ".

Démonstration: Soit n un entier pair. Alors on a:

n pair $\Longrightarrow n = 2k$, $k \in \mathbb{N}$

$\Longrightarrow n^2 = 4t$, $t = k^2$

Donc n^2 est divisible par 4.

Réciproquement, Soit n un entier impair. Alors on a:

n impair $\Longrightarrow n = 2k + 1$, $k \in \mathbb{N}$

$\Longrightarrow n^2 = 4k^2 + 4k + 1$

$\Longrightarrow n^2 = 4\left(k^2 + k\right) + 1$

Donc, le reste de la division de n^2 par 4 est égal à 1, d'où n^2 n'est pas divisible par 4.

On a alors montré que $\forall n \in \mathbb{N}$,(n pair $\Longleftrightarrow n^2$ est divisible par 4).

CQFD

2.5.8 Démonstration de $(\mathbf{Q} \vee \mathbf{R})$

Dans le cas où l'une des propositions est vraie, alors (Q $\vee$ R) est automatiquement vraie. On suppose donc que l'une des propositions est fausse et on montre que l'autre est vraie.

N.B: On doit donc montrer une des propositions ($\urcorner$Q $\Longrightarrow$ R) ou ($\urcorner$R $\Longrightarrow$ Q) en considérant la négation la plus simple.

Exemple: Montrer que $\forall n \in \mathbb{N}$,(n est impair ou n^2 est pair).

Commentaire: Ici, Q est la proposition " n est impair " et R la proposition " n^2 est pair ". On peut appliquer ($\urcorner$Q $\Longrightarrow$ R).

Démonstration: Soit n un entier naturel pair. Alors on a:

n pair $\Longrightarrow n = 2k$, $k \in \mathbb{N}$

$\Longrightarrow n^2 = 2t$, $t = 2k^2$

Donc n^2 est pair. On a donc bien $\forall n \in \mathbb{N}$,(n est impair ou n^2 est pair).

2.5.9 Raisonnement par récurrence

Le raisonnement par récurrence s'applique aux propositions dont l'énoncé dépend d'un entier naturel n. Il est une conséquence de la construction de l'ensemble des entiers naturels $\mathbb{N}$ (basée sur les axiomes de Peano). Ce raisonnement peut prendre différentes formes:

- Récurrence simple: C'est le type le plus utilisé. Si les hypothèses suivantes son vérifiées:

Hypothèse de départ: la propriété est vraie en n_0,

Hypothèse de récurrence (ou d'hérédité): lorsque la propriété est vraie pour $k \geq n_0$, elle est vraie pour $k+1$,

alors la propriété est vraie pour tout entier $n \geq n_0$.

Exercice: Montrer que $\forall n \in \mathbb{N}, (2^n > n)$.

- Récurrence forte: C'est la forme la plus générale du raisonnement par récurrence. Si les hypothèses suivantes son vérifiées:

Hypothèse de départ: la propriété est vraie en n_0,

Hypothèse de récurrence (ou d'hérédité): lorsque la propriété est vraie pour tout entier p tel que $n_0 \leq p \leq k$, elle est vraie pour $k+1$,

alors la propriété est vraie pour tout entier $n \geq n_0$.

Exercice: Soit $(x_n)_{n\in\mathbb{N}}$ la suite réelle définie par:

$$x_0 = 1 \text{ et } x_{n+1} = x_0 + x_1 + \ldots + x_n \text{ pour tout entier } n > 0.$$

Montrer que $\forall n \in \mathbb{N}$, $(x_n \leq 2^n)$.

2.5.10 D'autres règles

Il y a beaucoup d'autres règles qui peuvent intervenir dans une démonstration. Par exemple, la règle dite du *modus ponens*: " si on a P et $(\mathrm{P} \Longrightarrow \mathrm{Q})$, alors on a Q ", ou des *syllogismes*: " si on a $(\forall x \in \mathrm{E}, \mathrm{P}(x))$ et si a est un élément de E, alors on a $\mathrm{P}(a)$ ".L'exemple célèbre de syllogisme est " Tous les hommes sont mortels et Socrate est un homme; donc Socrate est mortel ".

Exercice: Expliciter le syllogisme précédent à l'aide de quantificateurs.

Chapter 3

Structures algébriques

La formalisation des structures algébriques - groupes, anneaux, corps, espaces vectoriels - est relativement récente, mais l'idée est présente partout dans les sciences et en particulier en mathématiques. Il s'agit grosso modo d'extraire des règles opératoires, valables indépendamment de la nature des objets considérés. Par exemple les règles pour faire la somme de deux nombres, la somme de deux vecteurs du plan ou la composition de deux rotations sont les mêmes. L'idée sous-jacente à la notion de groupe est celle de la symétrie.

Nous consacrerons un chapitre entier aux espaces vectoriels à cause de son importance.

3.1 Groupes

3.1.1 Lois de composition internes

Définitions et exemples

Définition: Soit E un ensemble. Une loi de composition interne sur E est une application de $E \times E$ dans E.

Si on note

$$E \times E \longrightarrow E$$

$$(a, b) \longmapsto a * b$$

on parle de la loi $*$ et on dit que $a * b$ est le composé de a et b pour la loi $*$.

Définition: Soient E un ensemble muni d'une loi de composition interne notée " $*$ ", et H un sous-ensemble de E. On dit que H est stable pour la loi $*$ si $\forall\,(x,\,y) \in H^2$, $x * y \in H$; et on dit alors que la restriction de " $*$ " à H^2 est la loi induite sur H par " $*$ ".

Exemples: l'addition ou la multiplication sont des lois de composition internes sur $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$. La soustraction définit une loi de composition interne sur $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$ (mais pas sur $\mathbb{N}$).

La réunion ou l'intersection sont des lois de composition internes sur $\mathcal{P}(E)$.

Soit X un ensemble. On note $E = \mathcal{F}(X)$ l'ensemble des applications de X dans X.

La composition des applications

$$E \times E \longrightarrow E$$
$$(f,\,g) \longmapsto f \circ g$$

est une loi de composition interne sur E.

Propriétés usuelles des lois de composition internes

Définitions: Soit $*$ une loi de composition interne sur un ensemble E. On dit que

1) la loi $*$ est commutative si $\forall\,(x,\,y) \in E^2$, $x * y = y * x$.

2) la loi $*$ est associative si $\forall\,(x,\,y,\,z) \in E^3$, $(x * y) * z = x * (y * z)$.

Définitions: Soit $*$ une loi de composition interne sur un ensemble E. Un élément e de E est dit:

· neutre à droite pour la loi $*$ si $\forall x \in E$, $x * e = x$,

· neutre à gauche pour la loi $*$ si $\forall x \in E$, $e * x = x$,

· neutre s'il est neutre à droite et à gauche.

Si la loi $*$ est commutative, les énoncés " neutre à droite ", " neutre à gauche " et " neutre " sont équivalents.

Exemples:- L'addition et la multiplication dans $\mathbb{Z}$ sont commutatives et associatives. Ce n'est pas le cas de la soustraction dans $\mathbb{Z}$ (montrez le).

- La composition des applications dans $\mathcal{F}(X)$ est associative, mais n'est pas en général commutative (trouvez un contre-exemple).

Définition: Soit $*$ une loi de composition interne sur un ensemble E, possédant un élément neutre e, et soit a un élément de E. On dit que a admet:

· un symétrique à droite pour la loi $*$, si $\exists\ a' \in E$, $a * a' = e$.

· un symétrique à gauche pour la loi $*$, si $\exists\ a^{\text{J}} \in E$, $a^{\text{J}} * a = e$.

· un symétrique s'il existe un élément de E qui soit à la fois symétrique à droite et à gauche de a.

Si la loi $*$ est commutative, les énoncés " symétrique à droite ", "symétrique à gauche " et " symétrique " sont équivalents.

Ainsi, si la loi $*$ est commutative, alors a admet un symétrique $b \in E$ pour la loi $*$ si l'une des conditions équivalentes suivantes est vérifiée:

(i) $a * b = e$

(ii) $b * a = e$.

Exemple: Dans $\mathbb{R}$, chaque élément a possède un symétrique pour l'addition qui est son opposé $-a$. Mais a n'a un symétrique pour la multiplication que s'il est non nul; son symétrique est alors l'inverse $\dfrac{1}{a}$ de a.

Définition: Soit E un ensemble muni de deux lois de composition internes $*$ et $\perp$. On dit que $*$ est distributive par rapport à $\perp$ si

$$\forall\,(x,\,y,\,z) \in E^3,\ x * (y \perp z) = (x * y) \perp (x * z) \text{ et}$$
$$(x \perp y) * z = (x * z) \perp (y * z).$$

Exemples:- Dans $\mathbb{R}$, la multiplication est distributive par rapport à l'addition. Dans $\mathbb{R}$, l'addition est-elle distributive par rapport à la multiplication?

- Dans l'ensemble $\mathcal{P}(X)$ des parties d'un ensemble X, chacune des opérations $\cap$ et $\cup$ est distributive par rapport à l'autre.

Proposition: Soit $*$ une loi de composition interne sur un ensemble E, si $*$ possède un élément neutre, il est unique.

Démonstration: Supposons que $*$ ait deux éléments neutres e et e'. On a alors

$e = e * e'$ (car e' est élément neutre)

$\quad = e'$ (car e est élément neutre).

On en déduit que $e = e'$.

CQFD

Définition: Soit G un ensemble muni d'une loi de composition interne $*$. On dit que $(G, *)$ est un groupe si la loi $*$ satisfait aux trois conditions suivantes:

i) Elle est associative,

ii) Elle admet un élément neutre,

iii) Chaque élément de G admet un symétrique pour $*$.

Si de plus, la loi $*$ est commutative, on dit que le groupe $(G, *)$ est commutatif ou abélien (du nom du mathématicien Abel).

Exemples: - Les ensembles $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ munis de l'addition sont des groupes. Noter que $(\mathbb{N}, +)$ n'est pas un groupe car ne vérifie pas iii).

- Les ensembles $\mathbb{Q}^*$, $\mathbb{R}^*$ et $\mathbb{C}^*$ munis de la multiplication sont des groupes. Noter que $(\mathbb{Z}^*, \times)$ n'est pas un groupe car ne vérifie pas iii).

Remarques:

- On parlera souvent de groupe G au lieu de $(G, *)$ s'il n'y a pas de risque de confusion des lois utilisées.

- Pour calculer dans un groupe, on omettra souvent le signe $*$ et on écrira xy au lieu de $x * y$.

- Le symétrique d'un élément x est souvent noté x^{-1}.

- L'élément neutre est son propre symétrique.

Proposition: Dans un groupe, tout élément admet un symétrique unique.

Démonstration: Soient G un groupe et x un élément de G. Supposons que x ait deux symétriques x' et x^{J}. On a alors

$$\begin{aligned} x' &= x'e \text{ (ou } e \text{ est l'élément neutre pour la loi de groupe sur } G \text{)} \\ &= x'\left(xx^{\text{J}}\right) \\ &= \left(x'x\right)x^{\text{J}} \\ &= ex^{\text{J}} \\ &= x^{\text{J}} \end{aligned}$$

ce qui montre l'unicité du symétrique.

CQFD

Propriétés: Soient x, y deux éléments d'un groupe G. Alors

i) $(xy)^{-1} = y^{-1}x^{-1}$

ii) $\left(x^{-1}\right)^{-1} = x$

Démonstration: Notons e l'élément neutre pour la loi de groupe sur G

i) On a $(xy)\, y^{-1}x^{-1} = x\left(yy^{-1}\right)x^{-1} = xex^{-1} = xx^{-1} = e$

$$y^{-1}x^{-1}(xy) = y^{-1}\left(x^{-1}x\right)y = y^{-1}ey = y^{-1}y = e.$$

Ce qui montre que $y^{-1}x^{-1}$ est le symétrique de xy c'est-à-dire qu'on a $(xy)^{-1} = y^{-1}x^{-1}$d'après l'unicité du symétrique.

ii) On a $\left(x^{-1}\right)^{-1}x^{-1} = \left(xx^{-1}\right)^{-1} = e^{-1} = e$

$$x^{-1}\left(x^{-1}\right)^{-1} = \left(x^{-1}x\right)^{-1} = e^{-1} = e.$$

Ce qui montre que $\left(x^{-1}\right)^{-1}$ est le symétrique de x^{-1} c'est-à-dire qu'on a $\left(x^{-1}\right)^{-1} = x$ d'après l'unicité du symétrique.

CQFD

3.1.2 Sous-groupe

Définition: On appelle sous-groupe d'un groupe $(G,*)$, toute partie H de G qui est stable pour la loi $*$ et qui est un groupe pour la loi induite sur H par la loi $*$.

Exemples: Si $(G,*)$ est un groupe et e l'élément neutre pour la loi $*$, alors $\{e\}$ et G sont des sous-groupes de G.

Définition: H est un sous-groupe propre d'un groupe G si H est un sous-groupe de G distinct de $\{e\}$ et G.

Proposition: Une partie H d'un groupe G est un sous-groupe de G si et seulement si elle satisfait:

i) $e \in H$, e étant l'élément neutre pour la loi de groupe sur G,

ii) $\forall\,(x,\,y) \in H^2$, $xy^{-1} \in H$.

Démonstration:

C.N $\Longrightarrow$) On a

$\cdot$ H sous-groupe de $G \Longrightarrow \exists x \in H$

$$\Longrightarrow xx^{-1} = e \in H$$

d'où i),

$\cdot$ $x,\, y \in H \Longrightarrow x,\, y^{-1} \in H$

$$\Longrightarrow xy^{-1} \in H.$$

C.S $\Longleftarrow$) On a

$\cdot$ $y \in H \Longrightarrow ey^{-1} \in H$

$$\Longrightarrow y^{-1} \in H$$

d'où tout élément de H est symétrisable.

$\cdot$ $x,\, y \in H \Longrightarrow x,\, y^{-1} \in H$

$$\Longrightarrow x\,(y^{-1})^{-1} \in H$$

$$\Longrightarrow xy \in H$$

d'où H est stable pour la loi de groupe sur G,

$\cdot$ l'associativité dans H résulte de celle dans G,

$\cdot$ par hypothèse, e est l'élément neutre dans H.

CQFD

Exercice: Montrer que

- Pour la loi d'addition, les inclusions suivantes sont les inclusions de sous-groupes: $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$;

- Pour la loi de multiplication, les inclusions suivantes sont les inclusions de sous-groupes: $\mathbb{Q}^* \subset \mathbb{R}^* \subset \mathbb{C}^*$; l'ensemble $\mathbb{R}_+^*$ (mais pas $\mathbb{R}_-^*$) est un sous-groupe de $\mathbb{R}$.

Définitions: Un homomorphisme de groupes $(G, *)$ et $(H, \perp)$ est une application $f : G \longrightarrow H$ telle que

$\forall\,(x, y) \in G^2,\ f(x * y) = f(x) \perp f(y)$.

Si de plus f est une bijection, on dit que f est un isomorphisme de groupes et que G et H sont isomorphes.

Exemple: Soient G un groupe et $a \in G$; définissons par récurrence $a^0 = e$ (e étant l'élément neutre pour la loi de groupe sur G) et $a^{n+1} = aa^n\ \forall n \in \mathbb{N}$, et enfin $a^{-n} = (a^n)^{-1}\ \forall n \in \mathbb{N}$. Ainsi $(\mathbb{Z}, +)$ et $(G, .)$ sont des groupes.

L'application

$$f : \mathbb{Z} \longrightarrow G$$

$$n \longmapsto a^n$$

est un homomorphisme de groupes. En effet:

$\forall\,(m, n) \in \mathbb{Z}^2,\ f(m + n) = a^{m+n} = a^m a^n = f(m)\, f(n)$.

Propriétés: Si $f : G \longrightarrow H$ est un homomorphisme de groupes, alors:

i) $f(e_G) = e_H$ où e_G (resp. e_H) désigne l'élément neutre pour la loi de groupe sur G (resp. sur H).

ii) $\forall x \in G,\ f(x^{-1}) = f(x)^{-1}$.

Démonstration:

i) $f(e_G) = f(e_G\ e_G) = f(e_G)\ f(e_G)$, et $f(e_G) = e_H\ f(e_G)$;

d'où

$f(e_G)\ f(e_G) = e_H\ f(e_G) \Longrightarrow (\ f(e_G)\, f(e_G))\ f(e_G)^{-1} = (e_H f(e_G))\ f(e_G)^{-1} \Longrightarrow f(e_G) = e_H$

ii) $f(x^{-1})\ f(x) = f(x^{-1}x) = f(e_G) = e_H$, d'où $f(x^{-1}) = f(x)^{-1}$.

CQFD

Définition: Le noyau d'un homomorphisme de groupes $f : G \longrightarrow H$ est l'ensemble

$$f^{-1}(\{e_H\}) = \{x \in G \mid f(x) = e_H\ \}.$$

On le note $Ker(f)$ à cause de l'allemand " Kern ".

L'importance du noyau vient du théorème suivant:

Théorème: Un homomorphisme de groupes $f : G \longrightarrow H$ est injectif si et seulement si $Ker(f) = \{e_G\}$. Le noyau de f est toujours un sous-groupe de G.

Démonstration: Soient $x, y \in G$, on a:

$f(x) = f(y) \iff f(x)f(y)^{-1} = e_H$
$\iff f(xy^{-1}) = e_H$
$\iff xy^{-1} \in Ker(f)$.

Nous voulons démontrer que: f injectif $\iff Ker(f) = \{e_G\}$.

C.N: $\Longrightarrow$) Utilisons le raisonnement par contraposée i.e

$Ker(f) \neq \{e_G\} \Longrightarrow f$ non injectif?

Si $Ker(f)$ contient un élément $a \neq e_G$, alors $f(a) = e_H = f(e_G)$ et f n'est pas injectif.

C.S: $\Longleftarrow$) Soient x, $y \in G$, on a:

$f(x) = f(y) \Longrightarrow xy^{-1} \in Ker(f)$
$\Longrightarrow xy^{-1} = e_G$ par hypothèse
$\Longrightarrow x = y$,

donc f est injectif.

Pour la deuxième affirmation, on a:

$\cdot\ e_G \in Ker(f)$, en effet
$f(e_G) = f(xx^{-1}) = f(x)f(x^{-1}) = f(x)f(x)^{-1} = e_H$.
$\cdot\ x, y \in Ker(f) \Longrightarrow f(xy^{-1}) = f(x)f(y^{-1}) = f(x)f(y)^{-1} = e_H\ e_H^{-1} = e_H$.
$\Longrightarrow xy^{-1} \in Ker(f)$.

CQFD

3.1.3 Groupes cycliques

Proposition: Soit A une partie d'un groupe G. Alors il existe un plus petit sous-groupe H de G contenant A; on l'appelle sous-groupe engendré par A et on le note $\langle A \rangle$.

Preuve: Il suffit de prendre $\langle A \rangle$ comme l'intersection de tous les sous-groupes de G contenant A.

NB: On peut décrire $\langle A \rangle$ comme l'ensemble des produits $x_1 \dots x_n$ ou chaque x_i vérifie: $x_i \in A$ ou $x_i^{-1} \in A$; si A est vide on prend $\langle A \rangle = \{e_G\}$.

Définition: Soient G un groupe et g élément de G.

On appelle ordre de g, le plus petit entier $n > 0$ (s'il existe) tel que $g^n = e_G$. Si $g^n \neq e_G$ pour tout $n > 0$, on dit que g est d'ordre infini.

Proposition: Soient G un groupe et g élément de G. Si $\langle g \rangle$ est infini, il est isomorphe à $\mathbb{Z}$; s'il est fini de cardinal n, il est isomorphe à $\mathbb{Z}\ /\ n\mathbb{Z}$. Dans les deux cas l'ordre de g est le cardinal de $\langle g \rangle$.

Preuve: Supposons que g est d'ordre infini. Alors l'application

$$\mathbb{Z} \longrightarrow \langle g \rangle$$
$$m \longmapsto g^m$$

est un homomorphisme surjectif; son noyau est trivial (car g est d'ordre infini, et $g^m = e_G$ est équivalent à $g^{-m} = e_G$ si m est un entier négatif) donc c'est un isomorphisme. Ainsi $\mathbb{Z}$ est isomorphe à $\langle g \rangle$.

supposons maintenant que g est d'ordre $n \in \mathbb{N}^*$. Comme $g^n = e_G$, l'application

$$\psi : \mathbb{Z} \,/\, n\mathbb{Z} \longrightarrow \langle g \rangle$$
$$\bar{m} \longmapsto g^m$$

est bien définie, et c'est un homomorphisme surjectif par définition de $\langle g \rangle$. Soit $\bar{m} \in \ker \psi$, la division euclidienne de m par n donne $m = nq + r$ avec $0 \leq r < n$. On obtient alors $g^r = 1$ et par suite $r = 0$ par définition de l'ordre. Ainsi $\bar{m} = 0$, ce qui montre que ψ est un isomorphisme de $\mathbb{Z} \,/\, n\mathbb{Z}$ sur $\langle g \rangle$.

Définition: Un groupe est dit monogène s'il est engendré par un seul élément, cyclique s'il est de plus fini.

Exemple: Le groupe $(\mathbb{Z}^n, +)$ est engendré par
$(1, 0, \ldots, 0)$, $(0, 1, \ldots, 0)$, ..., $(0, 0, \ldots, 1)$.

3.1.4 Sous-groupes distingués, groupes quotients

Définition: Soient G un groupe et g élément de G.

Alors l'application notée $int\ g$ définie par:

$$int\ g : G \longrightarrow G$$
$$h \longmapsto ghg^{-1}$$

est appelée automorphisme intérieur associé à g.

Définition: Un sous-groupe H de G est dit distingué ou normal et l'on note $H \lhd G$ s'il est laissé stable par tout automorphisme intérieur, i.e.: pour tout $g \in G$ et tout $h \in H$, on a $ghg^{-1} \in H$.

Remarques:

· On a $H \lhd G \Leftrightarrow \forall g \in G,\, gHg^{-1} \subset H \Leftrightarrow \forall g \in G,\, gHg^{-1} = H$

En effet la première équivalence est la définition; la deuxième équivalence s'obtient à partir de $gHg^{-1} \subset H$ en échangeant g et g^{-1} et en composant à gauche par g et à droite par g^{-1}.

· Si G est abélien, tout sous-groupe de G est distingué.

· $\{e_G\}$ et G sont toujours des sous-groupes distingués de G.

Proposition: Soit H un sous-groupe de G.

Alors la relation $\mathcal{R}_g(x,\, y)$ (resp. $\mathcal{R}_d(x,\, y)$) si et seulement si $x^{-1}y \in H$ (resp. $xy^{-1} \in H$) est une relation d'équivalence sur G. L'ensemble quotient s'appelle ensemble des classes à gauche (resp. à droite) modulo H, et est noté $G \,/\, H$ (resp. $H \setminus G$). Ses éléments sont de la forme aH (resp. Ha) avec $a \in H$. En particulier H est la classe de e_G.

Preuve: On démontre pour les classes à gauche.

$\mathcal{R}_g(x,\, x)$ est clair. Si $x^{-1}y \in H$, alors $(x^{-1}y)^{-1} = y^{-1}x \in H$ d'où la symétrie. si $x^{-1}y \in H$ et $y^{-1}z \in H$, alors $(x^{-1}y)\,(y^{-1}z) = x^{-1}z \in H$ d'où la transitivité.

Soit $a \in H$. Alors

$x \in aH \Rightarrow x = ay$ avec $y \in H$

$\Rightarrow a^{-1}x = y \in H$

et $\mathcal{R}_g(x,\, a)$. Réciproquement, $\mathcal{R}_g(a,\, x) \Rightarrow a^{-1}x \in H \Rightarrow x \in aH$. Ainsi la classe de a dans $G \,/\, H$ est bien aH.

Corollaire (Théorème de Lagrange): Si G est un groupe d'ordre fini, l'ordre de tout sous-groupe H de G divise l'ordre de G.

Preuve: Les classes à gauche constituent une partition de G et le cardinal de aH est le même que celui de H puisque les translations à gauche sont des bijections de G sur G.

Soient G un groupe et H un sous-groupe distingué de G. Par définition d'un sous-groupe distingué pour tout $a \in G$ on a $aHa^{-1} \subset H$ et $a^{-1}Ha \subset H$, d'où $aH \subset Ha$ et $Ha \subset aH$ i.e. $aH = Ha$ et par suite $G \,/\, H = H \setminus G$.

3.1.5 groupe symétrique

Structure

Définitions: soit E un ensemble fini. On appelle permutation de E, une bijection de E sur E. On note $S(E)$ l'ensemble des permutations de E. On note S_n l'ensemble des permutations de $\{1, \ldots, n\}$. Une permutation est souvent notée σ et représentée par un tableau.

Exemple:

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

représente la permutation de $\{1, 2, 3, 4\}$ telle que $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$, $\sigma(4) = 4$.

La première ligne du tableau représente les antécédents et la deuxième ligne les images respectives. On a $\sigma^{-1}(2) = 1$, $\sigma^{-1}(3) = 2$, $\sigma^{-1}(1) = 3$, $\sigma^{-1}(4) = 4$, σ^{-1} est la bijection réciproque de σ et on a $card(S_n) = n!$.

Les permutations étant des applications on peut les composer entre elles par la compée des applications $\circ$; la notation $\circ$ est parfois omise.

Exemples sur $\{1, \ldots, 6\}$:

$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 2 & 1 & 6 & 3 & 4 \end{pmatrix}$, $\theta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 1 & 2 & 3 \end{pmatrix}$ alors

$\sigma \circ \theta = \sigma\theta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 6 & 5 & 2 & 1 \end{pmatrix}$

Propriétés de $(S_n, \circ)$:

i) La loi $\circ$ est une loi de composition interne dans S_n.

En effet, la composée de deux bijections est une bijection.

ii) La loi $\circ$ est associative.

En effet la composition des aplications bijectives ou non, sur des ensembles finis ou non, est associative.

iii) La loi $\circ$ possède un élément neutre, à savoir

$Id = \begin{pmatrix} 1 & \ldots & n \\ 1 & \ldots & n \end{pmatrix}$ i.e $\sigma(i) = i$, $\forall i = 1, \ldots, n$.

iv) Tout $\sigma \in S_n$ est symétrisable et a pour symétrique σ^{-1}.

Ces quatre propriétés font de $(S_n, \circ)$ un groupe appelé groupe symétrique.

Décomposition d'une permutation

Définition: Soient σ une permutation de S_n et k un élément de $\{1, \ldots, n\}$. On appelle orbite de k suivant σ et l'on note $\mathcal{O}_\sigma(k)$, l'ensemble défini par $\mathcal{O}_\sigma(k) = \{\sigma^p(k) \mid p \in \mathbb{N}\}$.

On dit qu'une orbite est triviale si elle est réduite à un élément.

Exemple: $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 2 & 1 & 6 & 3 & 4 \end{pmatrix}$; $\mathcal{O}_\sigma(1) = \{1, 5, 3\}$, $\mathcal{O}_\sigma(2) = \{2\}$, $\mathcal{O}_\sigma(4) = \{4, 6\}$.

Remarques:

· L'orbite de i est de la forme $\{i, \sigma(i), \ldots, \sigma^p(i)\}$; c'est un ensemble formé d'éléments distincts avec $\sigma^{p+1}(i) = i$.

· Deux permutations différentes peuvent avoir même orbite, par exemple $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, $\theta = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$

Définition: On appelle tansposition de i et j et l'on note souvent τ_{ij}, la permutation qui échange simplement i et j et laisse les autres éléments fixes. En d'autres termes $\tau_{ij}(i) = j$, $\tau_{ij}(j) = i$ et $\tau_{ij}(k) = k$ si $k \neq i$ et $k \neq j$.

Définitions: On appelle cycle une permutation σ admettant une seule orbite non triviale $\mathcal{O}$. S'il en est ainsi, le cardinal de $\mathcal{O}$ est appelé longueur de σ et l'ensemble $\mathcal{O}$ est le support de σ.

On appelle $r-$cycle, un cycle de longueur r.

Exemples:

· Les transpositions sont des cycles à deux éléments.

· Dans S_6, la permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 3 & 6 & 5 & 1 \end{pmatrix}$$

est un cycle, dont le support est $\{1, 4, 6\}$ et de longueur 3.

Définition: On appelle permutation circulaire, une permutation constituée d'une seule orbite.

Exemple: Dans S_5 la permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 3 & 1 \end{pmatrix}$$

est une permutation circulaire.

Remarques:

· Un $r-$cycle s'écrit de r façons différentes.

· Un $r-$cycle c est d'ordre r (i.e $c^r = Id$).

· Un $2-$cycle est une transposition, et un $n-$cycle de S_n est une permutation circulaire.

Proposition: Toute permutation se décompose en produits de cycles disjoints. Ces cycles sont constitués des éléments des orbites non réduites à un élément.

On peut remarquer que ces cycles n'ayant aucun élément en commun commutent entre eux. La décomposition à l'ordre près des cycles, est unique.

Exemple: $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 3 & 1 & 6 & 5 \end{pmatrix} = (1,\, 2,\, 4)\,(5,\, 6)$

Proposition: Tout cycle se décompose en produit de transpositions.

Preuve: Cela découle de $(x_1,\, x_2,\, \ldots,\, x_p) = (x_1,\, x_2)\,(x_2,\, x_3)\ldots(x_{p-1},\, x_p)$. La démonstration est constructive.

Corollaire: Toute permutation se décompose en produit de transpositions.

Preuve: Cela résulte du fait que toute permutation se décompose en produit de cycles et que tout cycle se décompose en produit de transpositions.

Exemple: $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 3 & 1 & 6 & 5 \end{pmatrix} = (1,\, 2,\, 4)\,(5,\, 6)$

$$= (1,\, 2)\,(2,\, 4)\,(5,\, 6)$$

Remarques:

· Un cycle se décompose de plusieurs manières en produits de transpositions. On a par exemple

$$(x_1,\, x_2,\, x_3,\, \ldots,\, x_p) = (x_1,\, x_2)\,(x_2,\, x_3)\ldots(x_{p-1},\, x_p)$$
$$= (x_p,\, x_{p-1})\,(x_p,\, x_{p-2})\ldots(x_p,\, x_1)$$

· Le nombre de transpositions pour décomposer une permutation peut varier, mais la parité de ce nombre est conservée pour une permutation donnée.

Signature d'une permutation

Proposition: Soit σ une permutation de S_n. Il y a égalité entre les quatre quantités suivantes:

i) $(-1)^T$ où T est le nombre de transpositions dans une décomposition de σ comme produit de transpositions.

ii) $(-1)^D$ où D est la différence entre n et le nombre d'orbites suivant σ

iii) $(-1)^I$ où I est le nombre d'inversions de σ, i.e le nombre de couples $(i,\, j)$ avec $i < j$ et $\sigma(i) > \sigma(j)$

iv) $\underset{i<j}{\prod} \dfrac{\sigma(i) - \sigma(j)}{i-j}$

Cette quantité commune s'appelle signature de σ et se note $\mathrm{sgn}(\sigma)$ ou $\varepsilon(\sigma)$. Si $\mathrm{sgn}(\sigma) = 1$, on dit que σ est paire, sinon, elle est dite impaire.

Preuve:

$\cdot$ iv) = iii):

Numérateur et dénominateur comportent tous les produits $i-j$, au signe près, du fait de la bijectivité de σ. En valeur absolue, la quantité iv) vaut donc 1; son signe est donné par le nombre de couples $(i,\, j)$, tel que $i < j$ et $\sigma(i) > \sigma(j)$ i.e par le nombre d'inversions de σ.

$\cdot$ ii) = i):

Notons $\varepsilon(\sigma)$ la quantité ii). On remarque que pour toute permutation σ et toute transposition τ, on a:

$\varepsilon(\sigma\tau) = -\varepsilon(\sigma)$, car si $\tau = \tau_{ij}$, il y a deux cas à considérer:

a) i et j sont dans la même orbite suivant σ

Exemple: $\sigma = (1,\, 2,\, 4,\, 6,\, 7)\,(3,\, 5)$ et $\tau = (2,\, 6)$; on obtient

$\sigma\tau = (1,\, 2,\, 7)\,(3,\, 5)\,(4, 6)$.

b) i et j sont dans deux orbites différentes suivant σ

Exemple: $\sigma = (1,\, 2,\, 4,\, 6,\, 7)\,(3,\, 5)$ et $\tau = (2,\, 5)$; on obtient $\sigma\tau = (1, 2, 3, 5, 4, 6, 7)$.

On remarque que, pour une transposition τ, $\varepsilon(\tau) = -1$ car il y a $n-1$ orbites. On complète en montrant par récurrence que $\varepsilon(\sigma) = i$).

$\cdot$ iii) = i): Posons $\varepsilon'(\sigma) = \underset{i\neq j}{\prod} \dfrac{\sigma(i) - \sigma(j)}{i-j}$ on a:

$$\begin{aligned}
\varepsilon'(\sigma\sigma') &= \prod_{i\neq j} \frac{\sigma\sigma'(i) - \sigma\sigma'(j)}{i-j} \\
&= \prod_{i\neq j} \frac{\sigma\sigma'(i) - \sigma\sigma'(j)}{\sigma(i) - \sigma(j)} \prod_{i\neq j} \frac{\sigma(i) - \sigma(j)}{i-j} \\
&= \varepsilon'(\sigma)\,\varepsilon'(\sigma')
\end{aligned}$$

En fin $\varepsilon'(\tau) = -1$ pour une transposition; en effet, si $\tau = (i,\, j)$ alors il y a $2|j-i| - 1$ inversions. Donc par récurrence sur le nombre de transpositions, $\varepsilon'(\sigma)$ est égale à la formule i).

3.2 Anneaux et corps

3.2.1 Anneaux

Définition: Un anneau est la donnée d'un ensemble A et de deux lois de composition internes dans A notées $+$ (addition) et $\times$ (multiplication) telles que:

(i) $(A, +)$ est un groupe commutatif. L'élément neutre de $+$ est noté 0.

(ii) La loi $\times$ est associative,

(iii) La loi $\times$ est distributive par rapport à la loi $+$.

Un anneau tel anneau est noté $(A, +, \times)$. Un anneau $(A, +, \times)$ est dit commutatif (resp. unitaire) si la loi $\times$ est commutative (resp. admet un élément neutre).

Remarques:

· Dans la suite on ne considèrera que des anneaux unitaires, et on utilisera le mot anneau pour anneau unitaire. L'élément neutre de $\times$ sera noté 1 et appelé l'élément unité de l'anneau..

· Souvent on note xy au lieu de $x \times y$ pour x, y éléments de l'anneau.

Convention: Un anneau est donc un triplet $(A, +, \times)$, l'ensemble A s'appelle l'ensemble sous-jacent à l'anneau. On parle souvent de l'anneau A en sous-entendant les lois $+$ et $\times$ quand il est clair dans le contexte de quelles lois il s'agit.

Définition: Si l'élément x d'un anneau possède un symétrique pour la deuxième loi de cet anneau, on dira que x est inversible et on notera x^{-1} son inverse.

Propriétes: Soit $(A, +, \times)$ un anneau. Pour tous $x, y \in A$, on a:

1. $0x = 0$
2. $(-1)\, x = -x$
3. $(-1)\,(-1) = 1$
4. $(-x)\,(y) = -xy$

Preuve:

1.$0x = (0+0)\, x = 0x + 0x$, d'où $0x = 0$.

2.$0 = 0x = (1-1)\, x = x + (-1)\, x$, d'où $(-1)\, x = -x$

3.On multiplie par -1 l'égalité $(-1) + 1 = 0$; ce qui donne $(-1)\,(-1) + (-1)\, 1 = 0$, donc $(-1)\,(-1) + (-1) = 0$ et par suite $(-1)\,(-1) = 1$

4.$xy + (-x)\, y = (x + (-x))\, y = (x - x)\, y = 0y = 0$, d'où $(-x)\,(y) = -xy$.

Définitions: On dit qu'un élément a d'un anneau A est un diviseur de 0 s'il existe un élément non nul $b \in A$ tel que $ab = ba = 0$.

Un anneau est dit intègre s'il ne contient pas de diviseur de 0 autre que 0 lui-même.

Ainsi, un anneau A n'est pas intègre s'il existe des éléments x, y dans A tous deux non nuls tels que $xy = 0$.

Définition: Une partie B d'un anneau A, est un sous-anneau de A, si B munie des lois de A restreintes à B possède lui aussi une structure d'anneau.

Proposition: Une partie B d'un anneau A, est un sous-anneau de$(A, +, \times)$ si et seulement si:

· $(B, +)$ est un sous-groupe de $(A, +)$

· $1_A \in B$

· La loi $\times$ induit une loi de composition interne dans B.

3.2.2 Idéaux

Définitions: Soient $(A, +, \times)$ un anneau et I une partie de A. On dit que I est un idéal à gauche (resp. à droite) de A si et seulement si:

· $(I, +)$ est un sous-groupe abélien de $(A, +)$

· $\forall\, (a, x) \in A \times I$, $a \times x$ (resp. $x \times a$) est un élément de I.

On dit que I est un idéal bilatère de A lorsque I est à la fois un idéal à gauche et un idéal à droite de A. On utilise de manière générale le mot idéal pour idéal bilatère.

NB: On omettra souvent le signe $\times$ pour écrire xy au lieu de $x \times y$.

Définitions: Soient A un anneau et I un idéal (bilatère) de A.

· On dit que I est un idéal premier de A si et seulement si I n'est pas égal à A tout entier et si I vérifie:

$$\forall x,\, y \in A,\, xy \in I \Rightarrow x \in I \text{ ou } y \in I$$

· On dit que I est un idéal principal de A si et seulement si I est engendré par un unique élément a de A. Autrement dit:

$$I = \{xa;\, x \in A\}.$$

On notera dans ce cas (a), l'idéal engendré par l'élément a de A. L'idéal (0) est appelé l'idéal nul de A.

· On dit que I est strict ou propre dans A s'il n'est pas égal à A tout entier.

· On dit que I est un idéal maximal de A s'il est strict dans A, et si il n'est contenu dans aucun idéal de A autre que l'anneau A tout entier.

Proposition: Si un idéal d'un anneau A contient l'élément unité de A, alors cet idéal est égal à A tout entier.

Preuve: Supposons que l'idéal I de A contienne 1. Alors pour tout $a \in A$ on a $a = a1$ et est par définition d'un idéal un élément de I. Donc $A \subset I$ et par suite $I = A$.

Définition: Un anneau est dit principal s'il est intègre et que tous ses idéaux sont principaux.

3.2.3 Corps

Définition: Un corps est la donnée d'un ensemble K et de deux lois de composition internes notées $+$ (addition) et $\times$ (multiplication) telles que:

(i) $(K, +, \times)$ est un anneau,

(ii) $(K^*, \times)$ est un groupe, avec $K^* = K - \{0\}$.

Un corps est donc un anneau unitaire dont tous les éléments non nuls sont inversibles.

Un corps $(K, +, \times)$ est dit commutatif si la loi $\times$ est commutative.

Notations: Soit $(K, +, \times)$ un corps.

- L'élément neutre pour l'addition est noté 0_K ou simplement 0 s'il n'y a pas de risque de confusion des lois utilisées.

- L'élément neutre pour la multiplication est noté 1_K ou simplement 1 s'il n'y a pas de risque de confusion des lois utilisées.

- Le symétrique de $x \in K$ pour $+$ (opposé de x) est noté $-x$.

- Le symétrique de $x \in K$ pour $\times$ (inverse de x) est noté x^{-1}.

Exemples: Le triplet $(\mathbb{Z}, +, \times)$ est un anneau commutatif unitaire, mais ce n'est pas un corps car les seuls éléments de $\mathbb{Z}$ possédant un inverse pour la multiplication sont $+1$ et -1. Les corps les plus importants que nous étudierons sont $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$, $(\mathbb{C}, +, \times)$; tous ces corps sont commutatifs.

Convention: Un corps est donc un triplet $(K, +, \times)$, l'ensemble K s'appelle l'ensemble sous-jacent au corps. On parle souvent de corps K en sous-entendant les lois $+$ et $\times$ quand il est clair dans le contexte de quelles lois il s'agit.

Proposition: Un corps ne possède pas de diviseurs de 0.

Preuve: Supposons qu'il existe x et y dans un corps K tels que $xy = 0$; supposons de plus que x n'est pas nul. Alors x est inversible et on a $y = x^{-1}xy = x^{-1}0 = 0$, donc x et y ne sont pas des diviseurs de 0.

Proposition fondamentale: Les seuls idéaux d'un corps sont l'idéal nul et le corps lui-même. Réciproquement, si A est un anneau n'ayant comme seuls idéaux que l'idéal nul et lui même alors A est un corps.

Preuve:

· Supposons que K est un corps. Soient I un idéal non nul de A et x un élément non nul de I, donc x est inversible par définition d'un corps; on a donc $x^{-1}x \in I$ i.e $1 \in I$ et par suite $I = K$.

· Supposons maintenant que les seuls idéaux de l'anneau A sont l'idéal nul et A lui-même. Il suffit de montrer que tout élément non nul de A est inversible. Soit x un élément non nul de A, alors (x) l'idéal engendré par x n'est pas nul non plus et donc il est égal à A. L'unité de A est donc élément de (x), il existe alors $y \in A$ tel que $xy = 1$, donc x est inversible d'inverse y.

Définition: Soit K un corps. Soit A le sous-anneau de K engendré par l'élément unité 1 de K. Les éléments de A sont donc de la forme $\underbrace{1+1+\ldots+1}_{n \text{ fois}}$. Si A est de cardinal fini alors la caractéristique de K est le cardinal de A; sinon on dit que la caractéristique de K est nulle. Remarquons que si K est de caractéristique n alors $\underbrace{1+1+\ldots+1}_{n \text{ fois}} = 0$.

Chapter 4

Polynômes et fractions rationnelles

Dans ce chapitre, K désigne un corps commutatif
(le plus souvent $\mathbb{R}$ ou $\mathbb{C}$).

4.1 Polynômes à coefficients dans K

4.1.1 Présentation des polynômes

Définition: Soit $a = (a_k)_{k\geq 0}$ une suite à valeurs dans K. On appelle support de a l'ensemble (éventuellement vide) des indices k tels que $a_k \neq 0$. On note $K^{(\mathbb{N})}$ l'ensemble des suites à valeurs dans K qui sont à support fini.

Remarques:

· La suite $a = (a_k)_{k\geq 0}$ est à support fini si et seulement si $\exists\, n \in \mathbb{N} : \forall k > n$, $a_k = 0$. On peut alors noter symboliquement

$a = (a_0, a_1, \ldots, a_n, 0, 0, \ldots)$

· $\left(K^{(\mathbb{N})},\ +\ ,\ \times\right)$ est un anneau commutatif, où $+$ et $\times$ sont définies de la manière suivante:

soient $a = (a_k)_{k\geq 0}$ et $b = (b_k)_{k\geq 0}$ deux éléments de $K^{(\mathbb{N})}$. On a

$a + b = (a_k + b_k)_{k\geq 0}$ et $a \times b = c$ en posant: $\forall n \in \mathbb{N}$, $c_n = \sum\limits_{i+j=n} a_i \times b_j = \sum\limits_{j=0,\ldots,n} a_{n-j} \times b_j = \sum\limits_{i=0,\ldots,n} a_i \times b_{n-i}$.

En particulier $c_0 = a_0 \times b_0$, $c_1 = a_0 \times b_1 + a_1 \times b_0$, etc.

NB: Dans la suite on écrira xy au lieu de $x \times y$

Définition: Les éléments de l'anneau $K^{(\mathbb{N})}$ sont appelés polynômes à coefficients dans K.

Notation définitive des polynômes

Posons $X^0 = 1$, $X^1 = X = (0, 1, 0, 0, \dots)$; avec cette notation on $X^n = (0, \dots, 0, 1, 0, 0, \dots)$ où tous les coefficients sont nuls sauf celui de la $(n+1)^{i\text{éme}}$ position qui vaut 1. Le polynôme $P = (a_k)_{k\geq 0}$ s'écrit donc

$P = \sum_{k\geq 0} a_k X^k = a_0 + a_1 X + \ldots + a_n X^n + \ldots$. Une telle somme, toujours finie, représente P de façon unique (à l'ordre près). Si n est un entier tel que $\forall k > n$, $a_k = 0$, on notera aussi $P = \sum_{k=0}^{n} a_k X^k = a_0 + a_1 X + \ldots + a_n X^n$

On dit que les a_k sont les coefficients de P et X est l'indéterminée. On notera $K[X]$ l'anneau des polynômes à coefficients dans K.

Définitions: Soit $P = \sum_{k=0}^{n} a_k X^k = a_0 + a_1 X + \ldots + a_n X^n$ un polynôme non nul de $K[X]$.

· On dit que P est un monôme s'il est de la forme λX^n avec $\lambda \in K$ et $n \in \mathbb{N}$.

· On dit que P est un polynôme constant si $P = \lambda$, avec $\lambda \in K$.

· On appelle degré de P et on note $\deg(P)$ le plus grand entier k tel que $a_k \neq 0$; ce coefficient a_k est appelé coefficient dominant de P.

· On appelle valuation de P et on note $\mathrm{val}(P)$ le plus petit entier k tel que $a_k \neq 0$.

Par convention on pose $\deg(0) = -\infty$ et $\mathrm{val}(0) = +\infty$.

· On dit que P est normalisé (ou encore unitaire) si son coefficient dominant est égal à 1.

Définitions: Soit A un polynôme non nul et λ le coefficient dominant de A. Le polynôme $A^* = \dfrac{A}{\lambda}$ est appelé le normalisé de A. Deux polynômes non nuls sont dits associés s'ils ont même normalisé.

Exemple: Soit $P = X^4 + 3X + 1$, on a $\mathrm{val}(P) = 0$, $\deg(P) = 4$ et P est normalisé.

Proposition: Soient P et Q deux éléments de $K[X]$. On a:

· $\deg(P+Q) \leq \max(\deg(P), \deg(Q))$, avec égalité si $\deg(P) \neq \deg(Q)$

· $\mathrm{val}(P+Q) \geq \min(\deg(P), \deg(Q))$, avec égalité si $\deg(P) \neq \deg(Q)$

· $\deg(PQ) = \deg(P) + \deg(Q)$

· $\mathrm{val}(PQ) = \mathrm{val}(P) + \mathrm{val}(Q)$

Preuve: la démonstration est laissée au lecteur.

Corollaire: $K[X]$ est un anneau intègre.

Preuve: L'égalité $\deg(PQ) = \deg(P)+\deg(Q)$ montre que si $P \neq 0$ et $Q \neq 0$ alors $PQ \neq 0$, i.e

$(PQ = 0) \Rightarrow (P = 0$ ou $Q = 0)$.

Remarque: Le polynôme $P = \sum_{k=0}^{n} a_k X^k = a_0+a_1X+\ldots+a_nX^n$ est aussi noté $P(X)$. On distingue le polynôme $P(X)$ (qui, par construction, est nul si et seulement si tous ses coefficients sont nuls) de la fonction polynômiale associée:

$$P : K \longrightarrow K,\ x \longmapsto a_0 + a_1x + \ldots + a_nx^n = P(x).$$

Celle-ci est nulle si et seulement si: $\forall x \in K,\ P(x) = 0$.

On a évidemment l'implication: $P(X) = 0 \Rightarrow \forall x \in K,\ P(x) = 0$.

4.1.2 Division euclidienne

Théorème: Soient $A,\ B \in K[X]$ avec B non nul. Il existe un unique couple $(Q,\ R)$ de polynômes dans $K[X]$ tels que:

$$A = QB + R \quad \text{et} \quad \deg(R) < \deg(B).$$

Le passage du couple $(A,\ B)$ au couple $(Q,\ R)$ s'appelle division euclidienne de A par B. Dans cette division, A est le dividende, B le diviseur, Q le quotient et R le reste.

N.B: - Lorsque $R = 0$, on dit que B divise A, que A est divisible par B, que B est un diviseur de A ou que A est un multiple de B.

- Avant d'effectuer une division euclidienne, il faut toujours vérifier que les deux polynômes sont ordonnés suivant les puissances décroissantes.

Preuve:

· Existence

- Si $\deg(A) < \deg(B)$, on peut prendre $Q = 0$ et $R = A$.

- Si $\deg(A) \geq \deg(B)$, on note $Q_1 = \frac{a_n}{b_m}X^{n-m}$ le quotient des termes de plus haut degré. En posant $A_1 = A - BQ_1$, on a $\deg(A_1) < \deg(A)$.

- si $\deg(A_1) < \deg(B)$, on peut prendre $Q = Q_1$ et $R = A_1$.

- si $\deg(A_1) \geq \deg(B)$, on recommence avec A_1 et B: on obtient Q_2 et A_2, et $\deg(A_2) < \deg(A_1)$.

- si $\deg(A_2) < \deg(B)$, on peut prendre $Q = Q_1 + Q_2$ et $R = A_2$.

- si $\deg(A_2) \geq \deg(B)$, on recommence...

Puisque la suite de degrés $\deg(A) > \deg(A_1) > \deg(A_2) > \ldots$ est une suite strictement décroissante de nombres entiers, il n'y a qu'un nombre fini k d'étapes,on peut prendre $Q = Q_1 + \ldots + Q_k$ et $R = A_k$

· Unicité

Supposons que $A = QB + R = Q_1B + R_1$ avec $\deg(R) < \deg(B)$ et $\deg(R_1) < \deg(B)$. On en tire:

$$R - R_1 = (Q_1 - Q)\, B \text{ et } \deg(R - R_1) < \deg(B).$$

Par l'absurde: si $Q_1 \neq Q$, alors

$\deg((Q_1 - Q)\, B) = \deg(Q_1 - Q) + \deg(B) \geq \deg(B)$ ce qui est contradictoire.

Donc $Q_1 = Q$, d'où $QB + R = Q_1B + R_1 = QB + R_1$ et par suite $R = R_1$. CQFD

4.1.3 Algorithme d'Euclide - Pgcd - Ppcm

Soient A, $B \in K[X]$; on note $B \mid A$ pour exprimer que B est un diviseur de A. L'ensemble des diviseurs de A sera noté $\mathcal{D}(A)$, et l'ensemble des multiples de A sera noté sera noté $AK[X]$.

Proposition et définitions: Soient A, $B \in K[X]$.

Il existe un unique polynôme normalisé ou nul D tel que

$$\mathcal{D}(A) \cap \mathcal{D}(B) = \mathcal{D}(D).$$

Autrement dit, pour tout $P \in K[X]$, on a

$$(P \mid A \text{ et } P \mid B) \Leftrightarrow P \mid D.$$

On dit que D est le pgcd de A et B, et on note $D = \text{pgcd}(A, B)$, ou $D = A \wedge B$.

Il existe un couple (U, V) d'éléments de $K[X]$ tels que

$$AU + BV = A \wedge B.$$

On dit que (U, V) est un couple de coefficients de Bezout du couple (A, B).

Il existe un unique polynôme normalisé ou nul M tel que

$$AK[X] \cap BK[X] = MK[X].$$

On dit que M est le ppcm de A et B, et on note $M = \text{ppcm}(A, B)$ ou $M = A \vee B$.

Preuve et remarques:

· Si $A = B = 0$, alors $\mathcal{D}(A) = \mathcal{D}(B) = K[X]$. Seul le polynôme $D = 0$ vérifie $K[X] = \mathcal{D}(D)$, d'où $0 \wedge 0 = 0$. Dans ce cas l'égalité $AU + BV = A \wedge B$ est vérifiée pour tout couple (U, V) d'éléments de $K[X]$.

· L'unicité (si existence) de $D = A \wedge B$ vient du fait que si $\mathcal{D}(D_1) = \mathcal{D}(D_2)$ alors D_1 et D_2 sont associés; or deux polynômes associés et unitaires sont égaux.

Pour démontrer la propostion quand $(A, B) \neq (0, 0)$, on s'inspire de l'algorithme d'Euclide dans $\mathbb{Z}$ et on forme une succession de divisions euclidiennes partant du couple (A, B) jusqu'à obtenir un reste non nul. Le pgcd de A et B est alors le normalisé du dernier reste non nul.

Principe de l'algorithme:

· Quitte à échanger A et B on peut supposer $B \neq 0$.

On pose $R_0 = A$ et $R_1 = B$. Le polynôme R_1 est le premier " reste ": il est non nul.

On effectue la division euclidienne de R_0 par R_1.

Notons $R_0 = Q_1 R_1 + R_2$ cette division. On a bien sûr $\deg(R_2) < \deg(R_1)$.

· Si $R_2 = 0$, alors le procédé s'arrête, et R_1 est le dernier reste non nul obtenu.

Sinon on effectue la division de de R_1 par R_2: $R_1 = Q_2 R_2 + R_3$ avec $\deg(R_3) < \deg(R_2)$.

· Si $R_3 = 0$, alors le procédé s'arrête, et R_2 est le dernier reste non nul obtenu.

Sinon on effectue la division de de R_2 par R_3: $R_2 = Q_3 R_3 + R_4$ avec $\deg(R_4) < \deg(R_3)$.

$\vdots$

· La $k-$ième étape de cet algorithme est une division $R_{k-1} = Q_k R_k + R_{k+1}$ avec $\deg(R_{k+1}) < \deg(R_k)$.

A ce stade on a: $\deg(R_0) > \deg(R_1) > \ldots > \deg(R_k) > \deg(R_{k+1})$; ce qui montre que l'algorithme doit s'arrêter à un certain moment car la suite des degrés est strictement décroissante dans $\mathbb{N}$.

Il existe donc une $n-$ième étape lors de laquelle $R_{n-1} = Q_n R_n$ c'est-à-dire $R_{n+1} = 0$; le polynôme R_n est donc le dernier reste non nul obtenu par cette méthode. Le normalisé de R_n est le pgcd de A, B.

Remarque: Soient A, $B \in K[X]$, l'algorithme d'Euclide permet aussi de déterminer (U, V) éléments de $K[X]$ tels que

$$AU + BV = A \wedge B$$

Pour cela, il suffit de remonter les étapes de l'algorithme en remplaçant les restes par les expressions correspondantes obtenues.

Définition: Soient A, $B \in K[X]$.

On dit que A, B sont premier entre eux si $A \wedge B = 1$.

Ce qui équivaut à dire que les seuls diviseurs communs à A et B sont les constantes non nulles.

4.2 Fractions rationnelles

4.2.1 Le corps des fractions rationnelles

Définition: Une fraction rationnelle est le quotient $\dfrac{A}{B}$ de deux polynômes avec $B \neq 0$.

On note $K(X)$ l'ensemble des fractions rationnelles de polynômes à coefficients dans K. Il n'est pas difficile de vérifier qu'il s'agit d'un corps.

Ainsi $K(X)$ est le corps des fractions de l'anneau intègre $K[X]$.

Remarques:

· Le corps $K(X)$ contient l'anneau intègre $K[X]$, car tout polynôme P peut s'écrire $P = \dfrac{P}{1}$.

· Soient $F = \dfrac{A}{B}$ et $G = \dfrac{C}{D}$ dans $K(X)$. On a $F = G \Leftrightarrow AD = BC$. En particulier, pour tous polynômes A, B, Q avec $B \neq 0$ et $Q \neq 0$, on a $\dfrac{AQ}{BQ} = \dfrac{A}{B}$. Soient $F = \dfrac{A}{B} \in K(X)$ et $\Delta = A \wedge B \neq 0$; il existe C, D tels que $A = \Delta C$ et $B = \Delta D$, donc $F = \dfrac{C}{D}$ avec $C \wedge D = 1$.

On dit alors que F est écrit sous forme irréductible, ou simplifiée.

Définition: Soit $F = \dfrac{A}{B}$ une fraction rationnelle irréductible. Soient $\widetilde{A}$, $\widetilde{B}$ les fonctions polynômiales associées àA et B.

On appelle fonction rationnelle $\widetilde{F}$ associée à F, l'application

$$x \mapsto \widetilde{F}(x) = \frac{\widetilde{A}(x)}{\widetilde{B}(x)}.$$

Le domaine de définition de $\widetilde{F}$ est K privé de l'ensemble des racines de B.

4.2.2 Pôles et parties polaires

Définition: Soit $F = \dfrac{A}{B}$ une fraction rationnelle écrite sous forme irréductible.

Soient $\alpha \in K$ et $m \in \mathbb{N}^*$. On dit que α est un pôle de F, avec la multiplicité m, si α est une racine du polynôme B avec la multiplicité m.

Remarques:

α est un pôle de F avec la multiplicité m, signifie qu'on a $F = \dfrac{A}{(X-\alpha)^m Q}$, avec $A(\alpha) \neq 0$ et $Q(\alpha) \neq 0$.

On parle de pôle simple si $m = 1$, et de pôle multiple si $m > 1$. On parle de pôle double si $m = 2$, triple si $m = 3$, etc.

Définition: Soit $F = \dfrac{A}{B} \in K(X)$ admettant α comme pôle de multiplicité $m \geq 1$. Alors F s'écrit de manière unique sous la forme

$F = \sum_{k=1}^{m} \dfrac{\lambda_k}{(X-\alpha)^k} + G$, où $(\lambda_1, \ldots, \lambda_m) \in K^m$ et où G est une fraction rationnelle n'admettant pas α pour pôle.

On dit alors que $\sum_{k=1}^{m} \dfrac{\lambda_k}{(X-\alpha)^k}$ est la partie polaire de F relativement au pôle α.

N.B: Les pôles de G sont ceux de F (sauf α), avec les mêmes multiplicités respectives.

4.2.3 Méthodes pratiques

On considère $F = \sum_{k=1}^{m} \dfrac{\lambda_k}{(X-\alpha)^k} + G$, où $(\lambda_1, \ldots, \lambda_m) \in K^m$ et où G est une fraction rationnelle n'admettant pas α pour pôle.

On donne ici quelques méthodes permettant de calculer les coefficients.

Cas d'un pôle simple

Il existe $\lambda \in K$ tel que $F = \dfrac{\lambda}{X-\alpha} + G$. Voici trois méthodes permettant de calculer le coefficient λ.

· Posons $F = \dfrac{A}{(X-\alpha)\,Q}$, avec $A(\alpha) \neq 0$ et $Q(\alpha) \neq 0$. Alors $\lambda = \dfrac{A(\alpha)}{Q(\alpha)}$.

Dans la pratique, on multiplie F par $(X-\alpha)$, et après simplification on substitue α à X.

Cette méthode est très adaptée au cas où B est factorisé.

· Avec la fonction rationnelle associée à F, on écrit : $\lambda = \lim\limits_{x\to\alpha}(x-\alpha)\,\widetilde{F}(x)$.

Cette méthode est plutôt utilisée quand $K = \mathbb{R}$.

· Posons $F = \dfrac{A}{B}$, avec $A(\alpha) \neq 0$. alors $\lambda = \dfrac{A(\alpha)}{B'(\alpha)}$.

Cette méthode est très adaptée au cas où B n'est pas factorisé.

Un cas classique est $B = X^n - 1$: les pôles sont les racines $n-$ièmes de l'unité.

Cas d'un pôle double

Il existe λ, $\mu \in K$ tels que $F = \dfrac{\lambda}{(X-\alpha)^2} + \dfrac{\mu}{X-\alpha} + G$.

· Posons $F = \dfrac{A}{(X-\alpha)^2\,Q}$, avec $A(\alpha) \neq 0$ et $Q(\alpha) \neq 0$. Alors $\lambda = \dfrac{A(\alpha)}{Q(\alpha)}$.

· Avec la fonction rationnelle associée à F, on écrit :

$$\lambda = \lim_{x\to\alpha}(x-\alpha)^2\,\widetilde{F}(x).$$

· Posons $F = \dfrac{A}{B}$, avec $A(\alpha) \neq 0$. alors $\lambda = \dfrac{2A(\alpha)}{B^{\text{J}}(\alpha)}$.

· Une fois λ déterminé, on peut écrire:

$H = F - \dfrac{\lambda}{(X-\alpha)^2} = \dfrac{\mu}{X-\alpha} + G$.

Ou bien α n'est pas un pôle de H (donc $\mu = 0$ et c'est fini), ou bien α est un pôle simple de H et on est ramené à des méthodes connes.

Cas d'un pôle multiple

$F = \sum_{k=1}^{m} \frac{\lambda_k}{(X-\alpha)^k} + G.$

On peut facilement calculer le coefficient λ_m de $\frac{1}{(X-\alpha)^m}$:

· Posons $F = \frac{A}{(X-\alpha)^m Q}$, avec $A(\alpha) \neq 0$ et $Q(\alpha) \neq 0$. Alors $\lambda_m = \frac{A(\alpha)}{Q(\alpha)}$.

· On peut aussi écrire $\lambda_m = \lim_{x\to\alpha} (x-\alpha)^m \widetilde{F}(x)$.

· Posons $F = \frac{A}{B}$, avec $A(\alpha) \neq 0$. alors $\lambda_m = \frac{m!A(\alpha)}{B^{(m)}(\alpha)}$ (méthode rarement utilisée).

Une fois λ_m déterminé, on peut écrire:

$H = F - \frac{\lambda}{(X-\alpha)^m} = \sum_{k=1}^{m-1} \frac{\lambda_k}{(X-\alpha)^k} + G.$

On est alors en mesure de calculer λ_{m-1}, etc. Cette méthode est trop lourde pour m " grand ", et on préfère dans ce cas revenir à

$F = \frac{A}{(X-\alpha)^m Q}$, avec $A(\alpha) \neq 0$ et $Q(\alpha) \neq 0$.

L'égalité $F = \sum_{k=1}^{m} \frac{\lambda_k}{(X-\alpha)^k} + G$ devient $\frac{A}{Q} = \sum_{k=1}^{m} \lambda_k (X-\alpha)^{m-k} + (X-\alpha)^m G$.

La substitution qui consiste à remplacer X par $X+\alpha$ donne alors:

$A(\alpha+X) = (\lambda_m + \lambda_{m-1}X + \ldots + \lambda_1 X^{m-1}) Q(\alpha+X) + X^m G(X+\alpha)$.

On peut alors calculer successivement λ_m, ..., λ_2, λ_1 par une méthode de division suivant les puissances croissantes de $A(\alpha+X)$ par $Q(\alpha+X)$, mais cette méthode n'est pas au programme de L_1.

4.2.4 Décomposition en éléments simples

Proposition et définition: Tout élément $F \in K(X)$ s'écrit de manière unique $F = P + Q$, où P est un polynôme, et G une fraction rationnelle de degré < 0.

On dit que P est la partie entière de F.

Remarque: Posons $F = \frac{A}{B}$, et soit $A = BQ + C$ la division euclidienne de A par B. On a l'égalité

$F = Q + \dfrac{C}{B}$ avec $\deg(C) < \deg(B)$.

On admettra la proposition suivante:

Proposition (décomposition dans $\mathbb{C}(X)$): Soit $F = \dfrac{A}{B}$ une fraction rationnelle à coefficients complexes. Soient $\alpha_1, \ldots, \alpha_p$ les pôles distincts de F, avec les multiplicités respectives $r_1, \ldots, r_p$. Alors F s'écrit de manière unique

$$F = E + \sum_{k=1}^{p} \left(\sum_{j=1}^{r_k} \frac{\lambda_{k,\,j}}{(X-\alpha_k)^j} \right)$$

où E est la partie entière de F et où les $\lambda_{k,\,j}$ sont des éléments de $\mathbb{C}$.

Cette écriture est appelée décomposition en éléments simples de F dans $\mathbb{C}(X)$.

N.B: Dans cette écriture, chaque somme $\displaystyle\sum_{j=1}^{r_k} \frac{\lambda_{k,\,j}}{(X-\alpha_k)^j}$ est la partie polaire de F pour α_k.

Donnons maintenant quelques techniques facilitant la décomposition en éléments simples pour compléter les méthodes vues précédemment.

Décomposition dans $\mathbb{C}(X)$ d'une fraction à coefficients réels

Soit $F = \dfrac{A}{B} \in \mathbb{R}(X)$.

L'idée est de considérer F comme élément de $\mathbb{C}(X)$ et la décomposer en tant que telle.

Si α et $\overline{\alpha}$ sont deux pôles conjugués non réels de F, de multiplicité m, les parties polaires respectives s'écrivent:

$$\sum_{k=1}^{m} \frac{\lambda_k}{(X-\alpha)^k} \quad \text{et} \quad \sum_{k=1}^{m} \frac{\overline{\lambda_k}}{(X-\overline{\alpha})^k}$$

ce qui permet de diminuer de moitié environ le nombre d'inconnues.

On peut aussi faire la décomposition de F dans $\mathbb{C}(X)$ pour obtenir sa décomposition dans $\mathbb{R}(X)$ après regroupement des termes conjugués. Cette méthode n'est envisageable que si les pôles non réels sont de multiplicité 1.

Méthode de parité et de l'imparité

Si F est paire ou impaire, les transformations

$$X \mapsto F(-X) \text{ ou } X \mapsto -F(-X)$$

permettent de déduire des relations sur les coefficients à calculer (le nombre d'inconnues diminue environ de moitié).

Méthode d'injection des valeurs particulières

Quand il reste peu de coefficients à calculer, il peut être intéressant d'injecter, dans l'égalité entre F et sa décomposition, une ou plusieurs valeurs qui ne soient pas des pôles de F.

F étant élément de $\mathbb{R}(X)$, on peut injecter la valeur complexe i, et l'identification donnera deux relations entre les coefficients réels inconns.

Méthode de limite

On suppose ici que $\deg(F) < 0$ (la partie entière est donc nulle).

La décomposition de F fait apparaître des termes de type $\dfrac{\lambda_k}{X-\alpha_k}$ ou $\dfrac{\alpha_k X + b_k}{X^2+\beta_k X+\gamma_k}$.

Le calcul de $\lim\limits_{x\to\infty} x\widetilde{F}(x)$ donne une relation liant les coefficients λ_k et α_k.

Cette méthode est intéréssante quand il ne reste plus que un ou deux coefficients à calculer.

Chapter 5

Espaces vectoriels

Dans ce chapitre, $(K,\ +\ ,\ \times)$ désignera un corps commutatif; en particulier $K = \mathbb{R}$ ou $\mathbb{C}$. Comme application à des situations concrètes, on peut noter que la théorie des codes de télécommunication utilise largement les espaces vectoriels sur les corps finis $\mathbb{Z}/p\mathbb{Z}$.

5.1 Espaces vectoriels (e.v)

Définitions: Soit E un ensemble muni d'une loi de composition interne notée $*$. On dit que E a une structure d'espace vectoriel sur K si:

1) $(E,\ *)$ est un groupe abélien,

2) il existe une loi " . " externe à coefficients dans K (multiplication par un scalaire):

$$K \times E \longrightarrow E$$

$$(\alpha,\ x) \longmapsto \alpha.x$$

satisfaisant les quatre axiomes suivants:

(i) $\forall \alpha \in K,\ \forall\,(x,\ y) \in E^2,\ \alpha.\,(x * y) = (\alpha.x) * (\alpha.y)$,

(ii) $\forall\,(\alpha,\ \beta) \in K^2,\ \forall x \in E,\ (\alpha + \beta)\,.x = (\alpha.x) * (\beta.x)$,

(iii) $\forall\,(\alpha,\ \beta) \in K^2,\ \forall x \in E,\ (\alpha \times \beta)\,.x = \alpha.\,(\beta.x)$,

(iv) $\forall x \in E,\ 1_K.x = x$.

Si E a une structure d'espace vectoriel sur K, alors les éléments de E sont appelés vecteurs et ceux de K des scalaires.

Remarques:

- On dira souvent un K-espace vectoriel au lieu d'un espace vectoriel sur K,

- Généralement, la loi " $*$ " dans E est aussi notée $+$, et l'on omet le symbole $\times$ pour écrire le produit de deux éléments de K.

- Tout corps commutatif a une structure d'espace vectoriel sur lui-même lorsque l'on identifie la loi externe et la multiplication interne.

- Un espace vectoriel est donc un triplet $(E,\ *\ ,\ .)$, mais on parlera souvent d'un espace vectoriel E en sous-entendant les lois " $*$ " et " . " quand il est clair dans le contexte de quelles lois il s'agit.

Exemples: $K^n = \{(x_1,\ldots,\ x_n) \mid x_i \in K\},\ K^0 = \{0\}$ sont des K-espaces vectoriels; l'ensemble $\mathcal{F}(X,\ K)$ des applicatons de X dans K (où X est un ensemble) est un K-espace vectoriel; l'ensemble $K[X]$ des polynômes en une indéterminée et à coefficients dans K, forme un K-espace vectoriel.

Conséquences de la définition: Soit E un K-espace vectoriel. On a les propriétés suivantes:

1) $\forall \alpha \in K,\ \forall x \in E, (\alpha = 0_K \text{ ou } x = 0_E) \Longrightarrow \alpha.x = 0_E$

2) $\forall \alpha \in K,\ \forall (x,\ y) \in E^2,\ \alpha.(x - y) = \alpha.x - \alpha.y$;
en particulier $\alpha.(-y) = -\alpha.y$.

3) $\forall (\alpha,\ \beta) \in K^2,\ \forall x \in E,\ (\alpha - \beta).x = \alpha.x - \beta.x$,
en particulier $(-\beta).x = -\beta.x$

4) $\forall \alpha \in K,\ \forall x \in E,\ (\alpha.x = 0_E) \Longleftrightarrow (\alpha = 0_K \text{ ou } x = 0_E)$.

Démonstration: Soient $(E,\ *\ ,\ .)$ un K-espace vectoriel, $(\alpha,\ \beta) \in K^2$ et $(x,\ y) \in E^2$; on a:

1) $0_K.x = (0_K + 0_K).x = 0_K.x + 0_K.x$, d'où $0_K.x = 0_E$
$\alpha.0_E = \alpha.(0_E + 0_E) = \alpha.0_E + \alpha.0_E$, d'où $\alpha.0_E = 0_E$

2) $\alpha.(x - y) + \alpha y = \alpha.(x - y + y) = \alpha.x$, d'où $\alpha.(x - y) = \alpha.x - \alpha.y$.
En faisant $x = 0_E$ on obtient $\alpha.(-y) = -\alpha.y$.

3) $(\alpha - \beta).x + \beta x = (\alpha - \beta + \beta).x = \alpha.x$, d'où $(\alpha - \beta).x = \alpha.x - \beta.x$.
En faisant $\alpha = 0_K$ on obtient $(-\beta).x = -\beta.x$.

4) C.N$\Longrightarrow$) Supposons $\alpha.x = 0_E$.
Si $\alpha \neq 0_K$ alors α^{-1} existe et on a

$$x = 1.x = (\alpha^{-1}\alpha).x = \alpha^{-1}(\alpha.x) = \alpha^{-1}.0_E = 0_E.$$

Si $x \neq 0_E$ alors $\alpha = 0_K$ sinon

$$(\alpha \neq 0_K,\ \alpha.x = 0_E) \Longrightarrow x = 1.x = (\alpha^{-1}\alpha).x = \alpha^{-1}(\alpha.x) = \alpha^{-1}.0_E = 0_E,$$

ce qui contredit l'hypothèse.
C.S$\Longleftarrow$) Déjà démontrée: c'est la conséquence 1).
CQFD

5.2 Sous-espaces vectoriels (s.e.v)

Définition: Un sous-ensemble F d'un espace vectoriel E est un sous-espace vectoriel de E si muni des deux lois de E (restreintes à F) il devient un espace vectoriel.

Proposition: Un sous-ensemble F d'un espace vectoriel E est un sous-espace vectoriel de E si et seulement si on a:
(i) $0_E \in F$
(ii) F est stable pour l'addition dans E i.e

$$\forall\,(x,\,y) \in F^2,\, x+y \in F$$

(iii) F est stable pour la multiplication externe sur E à coefficients dans K i.e

$$\forall\,(\alpha,\,x) \in K \times F,\, \alpha x \in F$$

Preuve: La condition nécessaire est évidente; la condition suffisante résulte du fait que l'addition et la multiplication par un scalaire (dans E) définissent bien deux lois $F \times F \longrightarrow F$ et $K \times F \longrightarrow F$ et les axiomes d'espace vectoriel sont alors vérifiés dans F puisqu'ils le sont dans E. CQFD

Remarque: Les conditions (ii) et (iii) de la proposition précédente peuvent se résumer en une condition appelée stabilité par combinaison linéaire de la manière suivante:

$$\forall\,(\alpha,\,\beta) \in K^2,\, \forall\,(x,\,y) \in F^2,\, (\alpha x+\beta y) \in F.$$

On peut alors reformuler la proposition de la manière suivante:

Proposition: Un sous-ensemble F d'un espace vectoriel E est un sous-espace vectoriel de E si et seulement si on a:
(i) $0_E \in F$
(ii) F est stable par combinaison linéaire i.e

$$\forall\,(\alpha,\,\beta) \in K^2,\, \forall\,(x,\,y) \in F^2,\, (\alpha x+\beta y) \in F.$$

Exemples: $\{0_E\}$, E sont des sous-espaces vectoriels d'un espace vectoriel E.

Le sous-ensemble des fonctions continues, ou dérivables, ou ... est un sous-espace vectoriel de $\mathcal{F}(\mathbb{R}, \mathbb{C})$.

Théorème: L'intersection $F \cap G$ de deux sous-espaces vectoriels F et G d'un K-espace vectoriel E est un sous-espace vectoriel de E.

Démonstration: Soient $\alpha \in K$ et $(x, y) \in (F \cap G)^2$; on a:

$\cdot$ F et G s.e.v de E, donc $0_E \in F \cap G$

$\cdot$ $\left.\begin{array}{l} x, y \in F \Longrightarrow x + y \in F \\ x, y \in G \Longrightarrow x + y \in G \end{array}\right\} \Longrightarrow x + y \in F \cap G$

$\cdot$ $(\alpha x \in F$ et $\alpha x \in G) \Longrightarrow \alpha x \in F \cap G$.

CQFD

On montrera en exercice que l'union de deux sous-espaces vectoriels est un sous-espace vectoriel si et seulement si l'un contient l'autre.

5.3 Bases et dimension

Dans tout ce paragraphe on travaille dans un espace vectoriel E sur un corps K que l'on pourra prendre égal à $\mathbb{R}$. Ce paragraphe donne une définition précise de la notion de dimension d'un espace vectoriel et est fondamental pour la suite. Le point clef est qu'un K-espace vectoriel E (de dimension finie) est toujours isomorphe à K^n. Ceci généralise la notion de repère du plan (deux vecteurs) et de l'espace (trois vecteurs).

Définition: Une famille (ou système) de vecteurs est une suite (finie dans la pratique) de vecteurs indexés par un ensemble I

(dans la pratique on prend souvent $I = \{1, \ldots, n\}$); on note $(e_i)_{i \in I}$ ou $\{e_i\}_{i \in I}$ une telle famille.

Définitions: Soit A une partie de E.

Un vecteur u de E est combinaison linéaire d'éléments de A s'il existe a_1, ..., $a_n \in A$, λ_1, ..., $\lambda_n \in K$ tels que

$$u = \lambda_1 a_1 + \ldots + \lambda_n a_n.$$

L'ensemble de ces combinaisons linéaires est un sous-espace vectoriel de E contenant A, appelé le sous-espace vectoriel engendré par A et noté $\mathcal{L}in(A)$.

On dit que A est une partie génératrice (ou un sous-ensemble générateur) de E si $\mathcal{L}in(A) = E$.

Si F et G sont deux sous-espaces vectoriels de E, l'ensemble

$$F + G = \{u \in E \mid u = u_1 + u_2 \text{ avec } u_1 \in F \text{ et } u_2 \in G \}$$

est le sous-espace vectoriel engendré par $F \cup G$.

Une famille $\{e_i\}_{i \in I}$ de vecteurs d'un espace vectoriel E est génératrice de E si tout vecteur de E est une combinaison linéaire des e_i.

Soit $\{u_1, \ldots, u_m\}$ une famille de vecteurs de E; $\mathcal{L}in\{u_1, \ldots, u_m\}$ désigne l'ensemble des combinaisons linéaires des vecteurs $u_1, \ldots, u_m$ i.e

$$\mathcal{L}in\{u_1, \ldots, u_m\} =$$
$$\{u \in E \mid \exists (\lambda_1, \ldots, \lambda_m) \in K^m : u = \lambda_1 u_1 + \ldots + \lambda_m u_m\}.$$

Proposition: Soit A une partie non vide de E. Le sous-espace de E engendré par A est l'intersection des sous-espaces vectoriels de E contenant A.

Le sous-espace de E engendré par A est donc le plus petit (au sens de l'inclusion) sous-espace vectoriel de E contenant A.

Démonstration: On sait que tout sous-espace vectoriel est stable par combinaison linéaire, donc toute combinaison linéaire d'éléments de A appartient à l'intersection des sous-espaces vectoriels de E contenant A. Inversement, tout élément de l'intersection des sous-espaces vectoriels de E contenant A est aussi un élément du sous-espace de E engendré par A (car $\mathcal{L}in(A)$ est un sous-espace vectoriel de E contenant A). CQFD

Remarques: Soit A une partie d'un K-espace vectoriel E.

· $\mathcal{L}in\{\} = \{0_E\}$ i.e si A est vide alors $\mathcal{L}in(A) = \{0_E\}$.

· Si A est un sous-espace vectoriel de E alors $\mathcal{L}in(A) = A$.

Définitions: Une famille $\{u_1, \ldots, u_n\}$ de vecteurs de E est liée s'il existe $\lambda_1, \ldots, \lambda_n \in K$ non tous nuls tels que

$$\lambda_1 u_1 + \ldots + \lambda_n u_n = 0.$$

On dit aussi que les vecteurs u_i sont linéairement dépendants.

Une famille$\{u_1, \ldots, u_n\}$ qui n'est pas liée est dite famille libre ou que les vecteurs u_i sont linéairement indépendants.

Conséquences: On laisse au lecteur le soin de montrer que:

· toute famille contenant le vecteur nul est liée.

· toute famille contenant deux vecteurs égaux est liée. Plus généralement, une famille est liée si et seulement si l'un de ses vecteurs est combinaison linéaire des autres vecteurs.

· toute famille de laquelle on peut extraire une famille liée est liée.

· toute sous-famille d'une famille libre est libre.

Lemme: Soient $v_1, \ldots, v_n$ des vecteurs de E, et $y_1, \ldots, y_p$ des vecteurs qui sont combinaisons linéaires des v_i.

Si $p > n$ alors la famille $\{y_1, \ldots, y_p\}$ est liée.

Démonstration: Démontrons par récurrence sur n.

Hypothèse de départ: Supposons $n = 1$.

Soit $y_i \in \{y_1, \ldots, y_p\}$, si $y_i = 0_E$ alors la famille $\{y_1, \ldots, y_p\}$ est liée; si $y_i \neq 0_E$ alors pour tout $y_j \in \{y_1, \ldots, y_p\}$ avec $i \neq j$ on a $y_i = \alpha_i v_1$ et $y_j = \alpha_j v_1$ avec $(\alpha_i, \alpha_j) \in K^* \times K$ d'où $y_j = \alpha_j v_1 = \dfrac{\alpha_j}{\alpha_i}(\alpha_i v_1) = \dfrac{\alpha_j}{\alpha_i}\, y_i$ et par suite y_j et y_i sont linéairement dépendants; donc la famille $\{y_1, \ldots, y_p\}$ est liée.

Hypothèse de récurrence: Supposons que la propriété est vraie jusqu'au rang $(n-1)$ avec $n > 1$ et montrons qu'elle est vraie au rang n.

On écrit $y_1 = \lambda_1 v_1 + \ldots + \lambda_n v_n$ avec $\lambda_1, \ldots, \lambda_n \in K$; on peut supposer que $\lambda_n \neq 0$ quitte à changer la notation des indices ou l'ordre des termes. Alors v_n est combinaison linéaire de $v_1, \ldots, v_{n-1}$ et y_1. Donc il existe $\alpha_2, \ldots, \alpha_p \in K$ tels que $y_2 - \alpha_2 y_1, \ldots, y_p - \alpha_p y_1$ soient combinaisons linéaires de $v_1, \ldots, v_{n-1}$. Par hypothèse de récurrence ils sont linéairement dépendants, ce qui entraîne que $y_1, \ldots, y_p$ sont linéairement dépendants. CQFD

Définition: Une base de E est une famille $\{u_1, \ldots, u_n\}$ d'éléments de E qui est libre et génératrice. Autrement dit, tout vecteur u de E s'écrit de manière unique

$$u = \lambda_1 u_1 + \ldots + \lambda_n u_n.$$

On dit alors que $(\lambda_1, \ldots, \lambda_n)$ est le n-uplet de coordonnées de u dans la base $\{u_1, \ldots, u_n\}$.

Exemples: Dans K^n, posons $e_i = (0, \ldots, 1, \ldots, 0)$ où 1 se trouve à la $i^{\text{è}}$ position. Alors $(e_1, \ldots, e_n)$ est une base de K^n, dite base canonique. Par exemple la base canonique de $\mathbb{R}$ est $\{e\}$ avec $e = 1$; la base canonique de $\mathbb{R}^2$ est $\{e_1, e_2\}$ avec $e_1 = (1, 0)$ et $e_2 = (0, 1)$.

La base canonique de l'ensemble $\mathbb{R}_n[X]$ des polynômes à une variable à coefficients dans $\mathbb{R}$ et de degré au plus n est $\{1, x, \ldots, x^n\}$.

Théorème (de la base incomplète): Soient $\{u_1, \ldots, u_p\}$ une famille libre d'éléments de E et S une partie génératrice finie de E. Il existe alors une base $\{u_1, \ldots, u_n\}$ de E avec $u_{p+1}, \ldots, u_n \in S$.

Démonstration: Si $u_1, \ldots, u_p$ et les éléments de S sont linéairement indépendants, on a gagné; sinon un élément u de S est combinaison linéaire des autres et des u_i, on peut donc recommencer avec $S - \{u\}$. CQFD

Remarque: Le nom du théorème de la base incomplète provient du fait qu'il indique que l'on peut compléter une famille libre en une base (à l'aide d'éléments d'une famille génératrice). On pourrait aussi l'appeler " théorème de la base extraite "puisqu'il dit aussi que l'on peut extraire une base de toute famille génératrice.

Définition: On dit que E est de dimension finie s'il admet une famille génératrice finie.

Exemples: L'espace vectoriel $K[X]$ n'est pas de dimension finie sur K. En effet, si $\{P_1, \ldots, P_n\}$ est une famille finie de polynômes, soit $d = \max(\deg(P_i))$, alors toute combinaison linéaire des P_i est un polynôme de degré $\leq d$. Les espaces K^n et $\mathbb{R}_n[X]$ sont de dimension finie (cf exemples de bases canoniques).

Corollaire: Tout espace vectoriel de dimension finie admet une base (finie).

Démonstration: C'est une conséquence immédiate du théorème de la base incomplète; on prend une partie génératrice finie et on peut en extraire une base. CQFD

Proposition: Toutes les bases d'un même espace vectoriel ont le même cardinal.

Démonstration: Soient $\{u_1, \ldots, u_n\}$ et $\{v_1, \ldots, v_m\}$ deux bases d'un même espace vectoriel. Si on avait $m > n$, comme les v_i sont combinaisons des u_i, on déduirerait du lemme prédédent que $\{v_1, \ldots, v_m\}$ est liée, ce qui n'est pas. On a donc établit que $m \leq n$, et par symétrie $n \leq m$ et donc $m = n$. CQFD

On peut donc définir

Définitions: Le cardinal d'une base de E s'appelle la dimension de E et l'on note $\dim_K E$ ou simplement $\dim E$ s'il n'y a pas de risque de confusion sur K.

La dimension de l'espace engendré par un système de vecteurs $\{u_1, \ldots, u_n\}$ s'appelle le rang du système i.e

$$\mathrm{rg}(\{u_1, \ldots, u_n\}) = \dim \mathcal{L}in\{u_1, \ldots, u_m\}.$$

Remarques:

· $\dim\{0_E\} = 0$.

· La dimension est l'invariant le plus important d'un espace vectoriel. La dimension dépend du corps considéré. Par exemple on a $\dim_{\mathbb{R}}\mathbb{R}^n = n$, $\dim_{\mathbb{C}}\mathbb{C}^n = n$, $\dim_{\mathbb{R}}\mathbb{C}^n = 2n$, $\dim_{\mathbb{R}}\mathbb{R}_n[X] = n+1$.

Théorème: Supposons que $\dim E = n$, alors:

(i) Toute famille libre d'éléments de E est de cardinal au plus n

(avec égalité si et seulement si c'est une base de E).

(ii) Toute famille génératrice de E est de cardinal au moins n

(avec égalité si et seulement si c'est une base de E).

Démonstration: (i) Une famille libre peut être complétée en une base de cardinal n et a donc un cardinal $\leq n$ avec égalité si il n'y a pas besoin de compléter, i.e si c'est une base.

(ii) Si une famille est génératrice, on peut en extraire une base de cardinal n et a donc un cardinal $\geq n$ avec égalité si il n'y a pas besoin de compléter, i.e si c'est une base. CQFD

En particulier, dans un espace de dimension n, le rang d'un système $\{u_1, \ldots, u_m\}$ est toujours inférieur ou égal à $\min(m, n)$ et on a:

· $\mathrm{rg}(\{u_1, \ldots, u_m\}) = m$ si et seulement si $\{u_1, \ldots, u_m\}$ est libre

· $\mathrm{rg}(\{u_1, \ldots, u_m\}) = n$ si et seulement si $\{u_1, \ldots, u_n\}$ est générateur.

Corollaire: Soit $\{u_1, \ldots, u_n\}$ une famille d'éléments d'un espace vectoriel E de dimension n. Alors les conditions suivantes sont équivalentes:

(i) La famille $\{u_1, \ldots, u_n\}$ est libre,

(ii) La famille $\{u_1, \ldots, u_n\}$ est génératrice de E,

(iii) La famille $\{u_1, \ldots, u_n\}$ est une base de E.

On peut énoncer le corollaire de la manière suivante:

Une famille génératrice de n vecteurs d'un espace vectoriel E de dimension n constitue une base de E; Une famille libre de n vecteurs d'un espace vectoriel E de dimension n constitue une base de E.

Démonstration: C'est une conséquence immédiate du théorème précédent.

Théorème: Soient F et G deux $K-$espaces vectoriels. Alors $F \subset G$ si et seulement si tous les éléments d'une famille génératrice de F appartiennent à G.

Démonstration:

C.N$\Longrightarrow$) Soit $\{u_1, \ldots, u_n\}$ une famille génératrice de F. Il est évident que tous les u_i appartiennent à G car ils appartiennent à F qui est par hypothèse une partie de G.

C.S$\Longleftarrow$)Soit $\{u_1, \ldots, u_n\}$ une famille génératrice de F, et supposons que tous les u_i appartiennent à G. On a alors

$u \in F \Longrightarrow \exists\,(\lambda_1,\ldots,\lambda_n) \in K^n$: $u = \lambda_1 u_1 + \ldots + \lambda_n u_n$
$\Longrightarrow u$ est combinaison linéaire des éléments de G,
$\Longrightarrow u \in G$ qui est un espace vectoriel donc stable par combinaison linéaire.

CQFD.

5.4 Somme directe; Supplémentaires.

Définition: La somme $F_1 + \ldots + F_p$ ($p \geq 2$) de p sous-espaces vectoriels d'un $K-$espace vectoriel E est dite directe si la décomposition de tout vecteur u de cette somme en la somme d'un vecteur de chaque sous-espace est unique.

On note alors $F_1 \oplus \ldots \oplus F_p$ ou $\overset{p}{\underset{i=1}{\oplus}} F_i$ au lieu de $F_1 + \ldots + F_p$, pour traduire cette particularité.

Caractérisation:

1) Si $p = 2$, la somme $F_1 + F_2$ est directe si et seulement si $F_1 \cap F_2 = \{0_E\}$ i.e

$$\boxed{F_1 \oplus F_2 \Longleftrightarrow F_1 \cap F_2 = \{0_E\}}$$

2) Si $p > 2$, la somme $F_1 + \ldots + F_p$ est directe si et seulement si

$$\begin{cases} F_1 \cap F_2 = \{0_E\} \\ (F_1 + F_2) \cap F_3 = \{0_E\} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (F_1 + F_2 + \ldots + F_{p-1}) \cap F_p = \{0_E\} \end{cases}$$

N.B: On ne commettra pas l'erreur de croire que pour $p > 2$, la somme $F_1 + \ldots + F_p$ est directe si et seulement si, pour tout $i \neq j$, $F_i \cap F_j = \{0_E\}$. Cette condition est nécessaire mais pas suffisante.

Définition: Soient F, G deux sous-espaces vectoriels d'un $K-$espace vectoriel E.

On dit que E est somme directe de F et G et on écrit $E = F \oplus G$, si tout vecteur u de E s'écrit de façon unique $u = u_1 + u_2$ avec $u_1 \in F$ et $u_2 \in G$; cela équivaut à dire qu'on a $F \cap G = \{0_E\}$ et $E = F + G$, ou encore

$F \cap G = \{0_E\}$ et $\dim E = \dim F + \dim G$ si $\dim E$ est finie. On peut résumer en écrivant:

$$E = F \oplus G \Longleftrightarrow \begin{cases} F \cap G = \{0_E\} \\ E = F + G \end{cases}$$

$$\Longleftrightarrow \begin{cases} F \cap G = \{0_E\} \\ \dim E = \dim F + \dim G \end{cases} \quad (\text{ si } \dim E \text{ est finie }).$$

Lorsque $E = F \oplus G$, on dit que G est supplémentaire de F dans E.

Proposition: Soit F un sous-espace vectoriel d'un espace vectoriel E de dimension finie, alors il existe un autre sous-espace vectoriel G de E tel que $E = F \oplus G$.

Démonstration: On choisit une base de F, disons $\{e_1, \ldots, e_m\}$ que l'on complète en une base de E, disons $\{e_1, \ldots, e_n\}$; alors le sous-espace vectoriel G engendré par $\{e_{m+1}, \ldots, e_n\}$ répond à la question. CQFD

Remarque: Un supplémentaire n'est pas unique; un décompte des cardinaux des diverses bases montre que la dimension du supplémentaire est indépendante du choix du supplémentaire, et si $E = F \oplus G$ alors

$\dim E = \dim F + \dim G$.

Proposition: Soient E et F deux $K-$espaces vectoriels de dimension finie.

1) Si $F \subset E$ alors $\dim F \leq \dim E$, avec égalité si et seulement si $E = F$.

2) $\dim (E \oplus F) = \dim E + \dim F$

3) $\dim (E \times F) = \dim E + \dim F$.

Démonstration: Ces trois énoncés peuvent se démontrer en dénombrant des bases de chacun des espaces vectoriels. Ainsi

1) On complète une base $\{e_1, \ldots, e_m\}$ de F en une base $\{e_1, \ldots, e_n\}$ de E et bien sûr $m \leq n$, l'égalité ayant lieu si $E = F$.

2) On constate que l'union d'une base de E et d'une base de F est disjointe et donne une base de $E \oplus F$.

3) On observe que si $\{e_1, \ldots, e_n\}$ est une base de E et $\{f_1, \ldots, f_m\}$ est une base de F alors

$\{(e_1, 0_F), \ldots, (e_n, 0_F), (0_E, f_1), \ldots, (0_E, f_m)\}$ est une base de $E \times F$. CQFD

Corollaire: Soient E et F deux $K-$espaces vectoriels de dimension finie et de base repective $\mathcal{B}_E$ et $\mathcal{B}_F$.

1) $\mathcal{B}_E \cup \mathcal{B}_F$ est une famille génératrice de $E + F$.

2) $(rg(\mathcal{B}_E \cup \mathcal{B}_F) = \dim E + \dim F) \Longrightarrow (E + F \text{ est directe})$.

Démonstration:

1) Soit u un élément de $E+F$, il existe alors $u_1 \in E$ et $u_2 \in F$ tels que $u = u_1 + u_2$; or u_1 est combinaison linéaire des éléments de $\mathcal{B}_E$ et u_2 est combinaison linéaire des éléments de $\mathcal{B}_F$, d'où u est combinaison linéaire des éléments de $\mathcal{B}_E \cup \mathcal{B}_F$ qui est donc une famille génératrice de $E+F$.

2) D'après 1) on a $E+F = \mathcal{L}in\,(\mathcal{B}_E \cup \mathcal{B}_F)$ d'où par définition

$$rg\,(\mathcal{B}_E \cup \mathcal{B}_F) = \dim\,(E+F)$$

et par suite

$$rg\,(\mathcal{B}_E \cup \mathcal{B}_F) = \dim\,(E+F) = \dim E + \dim F - \dim\,(E \cap F).$$

Ainsi

$(rg\,(\mathcal{B}_E \cup \mathcal{B}_F) = \dim E + \dim F) \Longrightarrow \dim\,(E \cap F) = 0$

$\Longrightarrow (F+G$ est directe).

5.5 L'algorithme du pivot

Définition: Soit S un système de vecteurs. Une opération élémentaire sur S consiste à:

· changer l'ordre des éléments de S,

· multiplier un élément de S par un scalaire non nul,

· ajouter à un élément de S une combinaison linéaire des autres éléments de S.

Théorème: Soit S un système de vecteurs.

Le sous-espace engendré par S (i.e $\mathcal{L}in(S)$) ne change pas si on effectue sur S une ou plusieurs des opérations élémentaires.

Démonstration: Effectuer ces opérations élémentaires donne toujours des combinaisons linéaires des éléments de S, d'où l'énoncé. CQFD

On admettra la proposition suivante:

Proposition: Soient K^n un espace vectoriel, et $S' = \{u_1, u_2, \ldots, u_m\}$ un système de vecteurs non nuls de K^n. Supposons les vecteurs de S' ordonnés comme suit: pour chaque vecteur u_i, si sa première composante non nulle est x_{ij}, les vecteurs u_k $(k > i)$ ont tous la composante x_{kj} nulle. Autrement dit, si $u_i = (x_{i1}, x_{i2}, x_{i3}, \ldots, x_{in})$ avec $1 \leq i \leq m$, on a un tableau de type (quitte à changer l'ordre des indices):

$$\left|\begin{array}{ccccccc|c} x_{11} & x_{12} & x_{13} & \dots & x_{1m} & \dots & x_{1n} & u_1 \\ 0 & x_{22} & x_{23} & \dots & x_{2m} & \dots & x_{2n} & u_2 \\ 0 & 0 & x_{33} & \dots & x_{3m} & \dots & x_{3n} & u_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & x_{mm} & \dots & x_{mn} & u_m \end{array}\right.$$

S'il en est ainsi, alors S' est libre.

Pour transformer un système donné S en un système de type S' par la méthode d'algorithme du pivot, on peut adopter le principe suivant:

Principe: La transformation d'un système donné S en un système de type S' par la méthode d'algorithme du pivot peut exiger plusieurs étapes dont chacune comprend:

a) une bonne disposition: on consigne les composantes des vecteurs dans un tableau en commençant par les vecteurs dont la première composante non nulle est la plus située à gauche.

b) une annulation de certains coefficients du tableau: on choisit la première ligne du tableau (qu'on note l_i pour faciliter la compréhension) dont le premier coefficient non nul x_{ij} est le plus situé à gauche et tel qu'il existe d'autres coefficients x_{kj} non nuls avec $k > i$. On annule alors les x_{kj} coefficients des lignes l_k en appliquant la formule $l_k \longrightarrow l_k + \alpha l_i$ (i.e on remplace l_k par $l_k + \alpha l_i$ où α est un scalaire convenablement choisi).

On réitère ainsi les étapes jusqu'à ce que le premier coefficient non nul de la $(i+1)$-éme ligne soit à droite de celui de la i-éme ligne.

Conséquences et applications: Soient S un système de vecteurs de K^n, et S' le système obtenu après transformation de S par la méthode d'algorithme du pivot. Alors:

1) Tous les systèmes obtenus dans les différents tableaux lors de la transformation de S en S' sont équivalents. Donc si l'un des systèmes est libre (resp. lié) alors tous les autres sont libres (resp. liés).

2) Si S' ne contient pas de ligne nulle alors S est libre (il en est de même pour S' et les autres systèmes équivalents).

Exemples: Les systèmes

$S_1 = \{(0,\ -2,\ 1)\,,\ (1,\ 3,\ 2)\,,\ (3,\ 3,\ 1)\}$ et

$S_2 = \{(1,\ -2,\ 1,\ 2)\,,\ (0,\ 3,\ 2,\ 0)\,,\ (2,\ -1,\ 4,\ 4)\,,\ (-1,\ 5,\ 1,\ -2)\}$

sont-ils libres?

Solution:

· Pour S_1 on a:

$$\begin{array}{ccc|l}
1 & 3 & 2 & l_1 \\
3 & 3 & 1 & l_2 \\
0 & -2 & 1 & l_3 \\
\hline
1 & 3 & 2 & l_1 \\
0 & -6 & -5 & l_2^{(1)} = l_2 - 3l_1 \\
0 & -2 & 1 & l_3 \\
\hline
1 & 3 & 2 & l_1 \\
0 & -2 & 1 & l_3 \\
0 & -6 & -5 & l_2^{(1)} = l_2 - 3l_1 \\
\hline
1 & 3 & 2 & l_1 \\
0 & -2 & 1 & l_3 \\
0 & 0 & -8 & l_2^{(2)} = l_2^{(1)} - 3l_3
\end{array}$$

Aucune ligne du dernier tableau n'est nulle, donc S_1 est libre.

· Pour S_2 on a:

1	−2	1	2	l_1
2	−1	4	4	l_2
−1	5	1	−2	l_3
0	3	2	0	l_4

1	−2	1	2	l_1
0	3	2	0	$l_2^{(1)} = l_2 - 2l_1$
0	3	2	0	$l_3^{(1)} = l_3 + l_1$
0	3	2	0	l_4

1	−2	1	2	l_1
0	3	2	0	$l_2^{(1)} = l_2 - 2l_1$
0	0	0	0	$l_3^{(2)} = l_3^{(1)} - l_2^{(1)}$
0	0	0	0	$l_4^{(1)} = l_4 - l_2^{(1)}$

Le dernier tableau contient des lignes nulles, donc S_2 est lié.

3) Si S est lié, on peut déterminer les relations de dépendance linéaire entre ses vecteurs en explicitant par " remontée " les lignes nulles du dernier tableau.

Exemple: L'exemple précédent montre que S_2 est lié; déterminons les relations de dépendance linéaire entre ses vecteurs.

· $l_4^{(1)} = 0 \iff l_4 - l_2^1 = 0 \iff l_4 - (l_2 - 2l_1) = 0$, d'où la relation $2l_1 - l_2 + l_4 = 0$.

· $l_3^{(2)} = 0 \iff l_3^{(1)} - l_2^{(1)} = 0 \iff (l_3 + l_1) - (l_2 - 2l_1) = 0$, d'où la relation $3l_1 - l_2 + l_3 = 0$.

Ainsi, les relations de dépendance linéaire entre les vecteurs de S_2 sont:

$$\boxed{2l_1 - l_2 + l_4 = 0 \text{ et } 3l_1 - l_2 + l_3 = 0}.$$

4) L'algorithme de pivot permet de déterminer le rang de S (i.e $dim\mathcal{L}in(S)$), et aussi une base de $\mathcal{L}in(S)$: le rang de S est le nombre de lignes non nulles de S', et une base de $\mathcal{L}in(S)$ est la famille constituée des vecteurs non nuls de S'.

Exemples: En reprenant les exemples traités au 2), on obtient:

$$rg\,(S_1) = dim\mathcal{L}in(S_1) = 3 \;\text{ et }\; rg\,(S_2) = dim\mathcal{L}in(S_2) = 2;$$

une base de $\mathcal{L}in(S_1)$ est $\{(1,\,3,\,2)\,,\ (0,\ -2,\,1)\,,\ (0,\,0,\ -8)\}$, et comme S_1 est libre c'est aussi une base de $\mathcal{L}in(S_1)$;

une base de $\mathcal{L}in(S_2)$ est $\{(1,\ -2,\,1,\,2)\,,\ (0,\,3,\,2,\,0)\}$.

5) L'algorithme de pivot permet de déterminer les équations de $\mathcal{L}in(S)$ i.e exprimer que des vecteurs appartiennent à $\mathcal{L}in(S)$ sous la forme d'équations linéaires reliant les composantes de ces vecteurs; pour cela les vecteurs de S sont disposés verticalement. Deux cas se présentent:

· Si la dernière étape de l'algorithme ne contient pas de ligne nulle alors $\mathcal{L}in(S) = K^n$, et il n'y a pas d'équation particulière à déterminer.

· Si la dernière étape contient des lignes nulles, alors on explicite par remontée ces lignes pour déterminer les équations linéaires reliant les composantes des vecteurs de S.

Exemples: Soient les systèmes:

$$S_1 = \{\ (1,\,3)\,,\ (3,\,0)\} \text{ et } S_2 = \{(1,\ -2,\,1)\,,\ (0,\,2,\,2)\,,\ (2,\ -2,\,4)\}.$$

Déterminer si possible les équations de $\mathcal{L}in(S_1)$ et $\mathcal{L}in(S_2)$.

Solution:

Pour $\mathcal{L}in(S_1)$ on a

$$\begin{array}{cc|l} 1 & 3 & x \\ 3 & 0 & y \\ \hline 1 & 3 & x \\ 0 & -9 & y^{(1)} = y - 3x \end{array}$$

Aucune ligne n'est nulle, donc $\mathcal{L}in(S_1) = \mathbb{R}^2$.

Pour $\mathcal{L}in(S_2)$ on a:

$$\begin{array}{ccc|l} 1 & 0 & 2 & x \\ -2 & 2 & -2 & y \\ 1 & 2 & 4 & z \end{array}$$

$$\begin{array}{ccc|l} 1 & 0 & 2 & x \\ 0 & 2 & 2 & y^{(1)} = y + 2x \\ 0 & 2 & 2 & z^{(1)} = z - x \end{array}$$

$$\begin{array}{ccc|l} 1 & 0 & 2 & x \\ 0 & 2 & 2 & y^{(1)} \\ 0 & 0 & 0 & z^{(2)} = z^{(1)} - y^{(1)} \end{array}$$

La dernière étape contient une ligne nulle, on peut alors déterminer une équation de $\mathcal{L}in(S_2)$ en explicitant " par remontée " cette ligne. On a: $z^{(2)} = 0 \Longleftrightarrow z^{(1)} - y^{(1)} = 0$

$\Longleftrightarrow (z - x) - (y + 2x) = 0,$

d'où l'équation $-3x - y + z = 0$.

Ainsi

$\mathcal{L}in(S_2) = \{(x,\, y,\, z) \in \mathbb{R}^3 \mid -3x - y + z = 0\}$.

6) Soient F et G deux sous-espaces vectoriels de K^n. L'algorithme du pivot permet de déterminer $F + G$, $F \cap G$ et aussi de voir si $K^n = F \oplus G$. Dans ce cas, on travaille avec la famille constituée de l'union des bases de F et G.

Exemple1: Soient les systèmes:

$S_1 = \{(0,\ -2,\ 1)\,,\ (1,\ 3,\ 2)\,,\ (3,\ 3,\ 1)\}$ et

$S_2 = \{(1,\ -2,\ 1)\,,\ (0,\ 3,\ 2)\,,\ (2,\ -1,\ 4)\}$.

On pose $F = \mathcal{L}in(S_1)$ et $G = \mathcal{L}in(S_2)$.

a) Déterminer si possible une base de $F + G$ et celle de $F \cap G$.

b) A-t-on $R^3 = F \oplus G$?

Solution:

a) D'après 4) une base de F est $\mathcal{B}_F = \{(1,\ 3,\ 2)\,,\ (0,\ -2,\ 1)\,,\ (0,\ 0,\ -8)\}$.

Détermination d'une base de G

$$
\begin{array}{ccc|l}
1 & -2 & 1 & l_1 \\
2 & -1 & 4 & l_2 \\
0 & 3 & 2 & l_3 \\
\hline
1 & -2 & 1 & l_1 \\
0 & 3 & 2 & l_2^{(1)} = l_2 - 2l_1 \\
0 & 3 & 2 & l_3 \\
\hline
1 & -2 & 1 & l_1 \\
0 & 3 & 2 & l_2^{(1)} = l_2 - 2l_1 \\
0 & 0 & 0 & l_3^{(1)} = l_3 - l_2^{(1)}
\end{array}
$$

Une base de G est donc $\mathcal{B}_G = \{(1,\ -2,\ 1)\,,\ (0,\ 3,\ 2)\}$.

Considérons le système:

$S = \mathcal{B}_F \cup \mathcal{B}_G = \{(1,\ 3,\ 2)\,,\ (0,\ -2,\ 1)\,,\ (0,\ 0,\ -8)\,,\ (1,\ -2,\ 1)\,,\ (0,\ 3,\ 2)\}$

$$
\begin{array}{ccc|l}
1 & 3 & 2 & l_1 \\
1 & -2 & 1 & l_2 \\
0 & -2 & 1 & l_3 \\
0 & 3 & 2 & l_4 \\
0 & 0 & -8 & l_5
\end{array}
$$

$$
\begin{array}{ccc|l}
1 & 3 & 2 & l_1 \\
0 & -5 & -1 & l_2^{(1)} = l_2 - l_1 \\
0 & -2 & 1 & l_3 \\
0 & 3 & 2 & l_4 \\
0 & 0 & -8 & l_5
\end{array}
$$

$$
\begin{array}{ccc|l}
1 & 3 & 2 & l_1 \\
0 & -5 & -1 & l_2^{(1)} \\
0 & 0 & \frac{7}{5} & l_3^{(1)} = l_3 - \frac{2}{5} l_2^{(1)} \\
0 & 0 & \frac{7}{5} & l_4^{(1)} = l_4 + \frac{3}{5} l_2^{(1)} \\
0 & 0 & -8 & l_5
\end{array}
$$

$$
\begin{array}{ccc|l}
1 & 3 & 2 & l_1 \\
0 & -5 & -1 & l_2^{(1)} \\
0 & 0 & -8 & l_5 \\
0 & 0 & \frac{7}{5} & l_3^{(1)} = l_3 - \frac{2}{5} l_2^{(1)} \\
0 & 0 & \frac{7}{5} & l_4^{(1)} = l_4 + \frac{3}{5} l_2^{(1)}
\end{array}
$$

$$
\begin{array}{ccc|l}
1 & 3 & 2 & l_1 \\
0 & -5 & -1 & l_2^{(1)} \\
0 & 0 & -8 & l_5 \\
0 & 0 & 0 & l_3^{(2)} = l_3^{(1)} + \frac{7}{40} l_5 \\
0 & 0 & 0 & l_4^{(2)} = l_4^{(1)} + \frac{7}{40} l_5
\end{array}
$$

Une base de $F+G$ est la famille constituée des vecteurs lignes non nuls du dernier tableau. Ainsi une base de $F+G$ est $\{(1,\,3,\,2)\,,\ (0,\ -5,\ -1)\,,\ (0,\,0,\ -8)\}$. La dimension de $F \cap G$ est le nombre de lignes nulles du dernier tableau, et la dimension de $F+G$ est le nombre de lignes non nulles du dernier tableau.

Ainsi

$$\dim(F \cap G) = 2 \ \text{ et } \ \dim(F+G) = 3.$$

Pour déterminer une base de $F \cap G$ on explicite chaque ligne nulle du dernier tableau et on garde dans chaque membre de l'égalité les vecteurs appartenant à la même base. On a:

$$\begin{aligned}
l_4^{(2)} = 0 &\Longleftrightarrow l_4^{(1)} + \frac{7}{40} l_5 = 0 \\
&\Longleftrightarrow \left(l_4 + \frac{3}{5} l_2^{(1)}\right) + \frac{7}{40} l_5 = 0 \\
&\Longleftrightarrow 40 l_4 + 24\,(l_2 - l_1) + 7 l_5 = 0 \\
&\Longleftrightarrow 24 l_1 - 7 l_5 = 24 l_2 + 40 l_4 = (24,\,72,\,104) \\
l_3^{(2)} = 0 &\Longleftrightarrow l_3^{(1)} + \frac{7}{40} l_5 = 0 \\
&\Longleftrightarrow l_3 - \frac{2}{5} l_2^{(1)} + \frac{7}{40} l_5 = 0 \\
&\Longleftrightarrow l_3 - \frac{2}{5}\,(l_2 - l_1) + \frac{7}{40} l_5 = 0 \\
&\Longleftrightarrow 16 l_1 + 40 l_3 + 7 l_5 = 16 l_2 = (16,\ -32,\,16).
\end{aligned}$$

Une base de $F \cap G$ est alors $\{(24,\,72,\,104)\,,\ (16,\ -32,\,16)\}$ ou d'après le théorème précédent

$$\left\{\frac{1}{8}\,(24,\,72,\,104)\,,\ \frac{1}{16}\,(16,\ -32,\,16)\right\} = \{(3,\,9,\,13)\,,\ (1,\ -2,\,1)\}.$$

b) On a $F \cap G = \mathcal{L}in\,\{(3,\,9,\,13)\,,\ (1,\ -2,\,1)\} \neq \{0_{\mathbb{R}^3}\}$, donc on a pas $R^3 = F \oplus G$.

Exemple 2: Soient les systèmes $S_1 = \{(1,\ -2)\,,\ (1,\,3)\}$ et $S_2 = \{(1,\,2)\}$.

On pose $F = \mathcal{L}in(S_1)$ et $G = \mathcal{L}in(S_2)$.

a) Déterminer si possible une base de $F+G$ et celle de $F \cap G$.

b) A-t-on $R^2 = F \oplus G$?

Solution:

a) On montre facilement que $S_1 = \{(1,\ -2)\,,\ (1,\,3)\}$ est une base de F, mais il est préférable de considérer la base obtenue par l'algorithme:

$$\begin{array}{cc|l} 1 & -2 & l_1 \\ 1 & 3 & l_2 \\ \hline 1 & -2 & l_1 \\ 0 & 5 & l_2^{(1)} = l_2 - l_1 \end{array}$$

Ainsi $\mathcal{B}_F = \{(1,\ -2)\,,\ (0,\ 1)\}$ i.e $\left\{(1,\ -2)\,,\ \frac{1}{5}\,(0,\ 5)\right\}$.
Une base de G est $\mathcal{B}_G = \{(1,\ 2)\}$. Considérons la famille
$S = \mathcal{B}_F \cup \mathcal{B}_G = \{(1,\ -2)\,,\ (0,\ 1)\,,\ (1,\ 2)\}$, on a:

$$\begin{array}{cc|l} 1 & -2 & l_1 \\ 1 & 2 & l_2 \\ 0 & 1 & l_3 \\ \hline 1 & -2 & l_1 \\ 0 & 4 & l_2^{(1)} = l_2 - l_1 \\ 0 & 1 & l_3 \\ \hline 1 & -2 & l_1 \\ 0 & 4 & l_2^{(1)} \\ 0 & 0 & l_3^{(1)} = l_3 - \frac{1}{4} l_2^{(1)} \end{array}$$

Ainsi une base de $F+G$ est $\{(1,\ -2)\,,\ (0,\ 4)\}$ ou encore $\{(1,\ -2)\,,\ (0,\ 1)\}$.
Une base de $F \cap G$ est obtenue en explicitant $l_3^{(1)} = 0$. On a:

$$\begin{aligned} l_3^{(1)} = 0 &\Longleftrightarrow l_3 - \frac{1}{4} l_2^{(1)} = 0 \\ &\Longleftrightarrow l_3 - \frac{1}{4}\,(l_2 - l_1) = 0 \\ &\Longleftrightarrow l_1 + 4l_3 = l_2 = (1,\ 2), \end{aligned}$$

donc une base de $F \cap G$ est $\{(1,\ 2)\}$.

b) $\dim(F \cap G) = 1 \neq 0 = \dim(\{0_{\mathbb{R}^2}\})$, d'où $F \cap G \neq \{0_{\mathbb{R}^2}\}$ et par suite on a pas $R^2 = F \oplus G$.

Exemple 3: Soient les systèmes
$S_1 = \{(1,\ -2,\ 1,\ 2)\,,\ (1,\ 0,\ -1,\ 1)\}$ et $S_2 = \{(1,\ 2,\ 0,\ 0)\,,\ (0,\ 1,\ 2,\ -1)\}$.

On pose $F = \mathcal{L}in(S_1)$ et $G = \mathcal{L}in(S_2)$.

a) Déterminer si possible une base de $F + G$ et celle de $F \cap G$.

b) A-t-on $R^4 = F \oplus G$?

Solution:

a)

$$\begin{array}{cccc|l} 1 & -2 & 1 & 2 & l_1 \\ 1 & 0 & -1 & 1 & l_2 \\ \hline 1 & -2 & 1 & 2 & l_1 \\ 0 & 2 & -2 & -1 & l_2^{(1)} = l_2 - l_1 \end{array}$$

d'où $\mathcal{B}_F = \{(1, -2, 1, 2), (0, 2, -2, -1)\}$.

Une base de G est $\mathcal{B}_G = S_2 = \{(1, 2, 0, 0), (0, 1, 2, -1)\}$.

Considérons la famille

$S = \mathcal{B}_F \cup \mathcal{B}_G = \{(1, -2, 1, 2), (0, 2, -2, -1), (1, 2, 0, 0), (0, 1, 2, -1)\}$,

on a:

$$
\begin{array}{cccc|l}
1 & -2 & 1 & 2 & l_1 \\
1 & 2 & 0 & 0 & l_2 \\
0 & 2 & -2 & -1 & l_3 \\
0 & 1 & 2 & -1 & l_4 \\
\hline
1 & -2 & 1 & 2 & l_1 \\
0 & 4 & -1 & -2 & l_2^{(1)} = l_2 - l_1 \\
0 & 2 & -2 & -1 & l_3 \\
0 & 1 & 2 & -1 & l_4 \\
\hline
1 & -2 & 1 & 2 & l_1 \\
0 & 1 & 2 & -1 & l_4 \\
0 & 2 & -2 & -1 & l_3 \\
0 & 4 & -1 & -2 & l_2^{(1)} \\
\hline
1 & -2 & 1 & 2 & l_1 \\
0 & 1 & 2 & -1 & l_4 \\
0 & 0 & -6 & 1 & l_3^{(1)} = l_3 - 2l_4 \\
0 & 0 & -9 & 2 & l_2^{(2)} = l_2^{(1)} - 4l_4 \\
\hline
1 & -2 & 1 & 2 & l_1 \\
0 & 1 & 2 & -1 & l_4 \\
0 & 0 & -6 & 1 & l_3^{(1)} = l_3 - 2l_4 \\
0 & 0 & 0 & \frac{1}{2} & l_2^{(3)} = l_2^{(2)} - \frac{9}{6} l_3^{(1)}
\end{array}
$$

Une base de $F + G$ est
$\{(1, -2, 1, 2), (0, 1, 2, -1), (0, 0, -6, 1), (0, 0, 0, 1)\}$.

Aucune ligne du dernier tableau n'est nulle, donc $\dim(F \cap G) = 0$, d'où $F \cap G = \{0_{\mathbb{R}^4}\}$ qui n'admet pas de base.

b) Aucune ligne du dernier tableau n'est nulle, donc $R^4 = F \oplus G$.

On appliquera la méthode d'algorithme du pivot pour résoudre plusieurs

autres problèmes d'algèbre linéaire.

Chapter 6

Applications linéaires

Dans tout le chapitre E et F désigneront des espaces vectoriels sur le même corps K.

6.1 Le K-espace vectoriel $\mathcal{L}_K(E, F)$

Définitions: Soient E et F deux espaces vectoriels sur le même corps K. Une application f de E dans F est dite K-linéaire (ou morphisme de K-espaces vectoriels) si

(i) $\forall (x, y) \in E^2$, $f(x+y) = f(x) + f(y)$

(ii) $\forall \lambda \in K$, $\forall x \in E$, $f(\lambda.x) = \lambda.f(x)$

ou encore si

$\forall (\lambda, \mu) \in K^2$, $\forall (x, y) \in E^2$, $f(\lambda.x + \mu.y) = \lambda.f(x) + \mu.f(y)$.

Notations et cas particuliers:

1) Si $f : E \longrightarrow F$ est une application K-linéaire, on dira souvent que f est une application linéaire (s'il n'y a pas ambiguïté sur le corps K) ou que f est un homomorphisme de E dans F. L'ensemble des applications K-linéaires de E dans F est noté $\mathcal{L}_K(E, F)$ ou simplement $\mathcal{L}(E, F)$ s'il n'y a pas ambiguïté sur le corps K.

Un homomorphisme bijectif $f : E \longrightarrow F$ est appelé isomorphisme de E sur F, et on dit que E et F sont isomorphes.

2) Soit $f \in \mathcal{L}_K(E, F)$; si $E = F$ on dit que f est un endomorphisme de E et on pose $\mathcal{L}_K(E, E) = \mathcal{L}_K(E)$. Un endomorphisme bijectif de E est appelé automorphisme de E. L'ensemble des automorphismes de E est noté $GL_K(E)$.

3) Soit $f \in \mathcal{L}_K(E, F)$; si $F = K$ on dit que f est une forme linéaire sur E.

Conséquences des définitions: Soient f, $g \in \mathcal{L}_K(E, F)$ et $\lambda \in K$. On montrera en exercice que:

1) Muni des lois + et ., $f + g$ et $\lambda.f$ étant définies par

$\forall x \in E$, $(f + g)(x) = f(x) + g(x)$

$\forall x \in E$, $(\lambda.f)(x) = \lambda.f(x)$

l'ensemble $\mathcal{L}_K(E, F)$ est un K-espace vectoriel.

2) $f(0_E) = 0_F$ et $\forall x \in E$, $f(-x) = -f(x)$

3) Si $f \in \mathcal{L}_K(E, F)$ et $g \in \mathcal{L}_K(F, G)$ alors $g \circ f \in \mathcal{L}_K(E, G)$. Ainsi la composée de deux applications linéaires est une application linéaire.

Exemples:

1) $f : K \longrightarrow K$, $x \longmapsto ax$ avec $a \in K$ est une application linéaire; elle est bijective si $a \neq 0$.

2) Soient $a_1, \ldots, a_n$ des réels, l'application

$$p : \mathbb{R}^n \longrightarrow \mathbb{R},\ x = (x_1, \ldots, x_n) \longmapsto a_1x_1 + \ldots + a_nx_n$$

est linéaire.

Remarque: Soit $\{u_1, \ldots, u_n\}$ une base de E, se donner $f \in \mathcal{L}_K(E, F)$ revient à se donner les n vecteurs $v_i = f(u_i)$ de F. En effet, pour tout vecteur $u = \lambda_1u_1 + \ldots + \lambda_nu_n$ de E on a alors $f(u) = \lambda_1v_1 + \ldots + \lambda_nv_n$. On a ainsi le théorème suivant:

Théorème: Soient $\{u_1, \ldots, u_n\}$ une base de E et $f \in \mathcal{L}_K(E, F)$. Alors

(i) f est injective $\Longleftrightarrow \{f(u_1), \ldots, f(u_n)\}$ est libre.

(ii) f est surjective $\Longleftrightarrow \{f(u_1), \ldots, f(u_n)\}$ est génératrice.

(iii) f est bijective $\Longleftrightarrow \{f(u_1), \ldots, f(u_n)\}$ est une base de F.

Démonstration:

(i) C.N $\Longrightarrow$) Supposons f injective.

Soient $\lambda_1, \ldots, \lambda_n$ des scalaires tels que $\lambda_1 f(u_1) + \ldots + \lambda_n f(u_n) = 0_F$, on a alors

$$\begin{aligned}\lambda_1 f(u_1) + \ldots + \lambda_n f(u_n) = 0_F &\Longrightarrow f(\lambda_1u_1 + \ldots + \lambda_nu_n) = 0_F \\ &\Longrightarrow (\lambda_1u_1 + \ldots + \lambda_nu_n) \in \ker f \\ &\Longrightarrow (\lambda_1u_1 + \ldots + \lambda_nu_n) \in \{0_E\} \\ &\Longrightarrow (\lambda_1u_1 + \ldots + \lambda_nu_n) = 0_E \\ &\Longrightarrow \lambda_1 = \ldots = \lambda_n = 0.\end{aligned}$$

C.S $\Longleftarrow$) Supposons $\{f(u_1), \ldots, f(u_n)\}$ libre.

Soit $u \in \ker f \subset E$, il existe alors des scalaires $\lambda_1, \ldots, \lambda_n$ tels que

$$u = \lambda_1 u_1 + \ldots + \lambda_n u_n.$$

On a

$f(u) = 0_F \Longrightarrow f(\lambda_1 u_1 + \ldots + \lambda_n u_n) = 0_F$

$\Longrightarrow \lambda_1 f(u_1) + \ldots + \lambda_n f(u_n) = 0_F$

$\Longrightarrow \lambda_1 = \ldots = \lambda_n = 0$

d'où $u = 0_E$; ce qui montre que $\ker f = \{0_E\}$ i.e que f est injective.

(ii) C.N $\Longrightarrow$) Supposons f surjective.

Soit $v \in F$, il existe alors $u = (\lambda_1 u_1 + \ldots + \lambda_n u_n) \in E$ tel que $v = f(u)$;

d'où $v = \lambda_1 f(u_1) + \ldots + \lambda_n f(u_n)$ ce qui montre que $\{f(u_1), \ldots, f(u_n)\}$ est génératrice.

C.S $\Longleftarrow$) Supposons que $\{f(u_1), \ldots, f(u_n)\}$ est génératrice.

Soit $v \in F$, il existe alors des scalaires $\lambda_1, \ldots, \lambda_n$ tels que

$$v = \lambda_1 f(u_1) + \ldots + \lambda_n f(u_n);$$

d'où $v = f(\lambda_1 u_1 + \ldots + \lambda_n u_n) = f(u)$ avec $u = (\lambda_1 u_1 + \ldots + \lambda_n u_n) \in E$ et par suite f est surjective.

(iii) f est bijective $\Longleftrightarrow$ f est injective et surjective

$\Longleftrightarrow \{f(u_1), \ldots, f(u_n)\}$ est libre et génératrice

$\Longleftrightarrow \{f(u_1), \ldots, f(u_n)\}$ est une base de F.

CQFD

6.2 Image, noyau d'une application linéaire

Soit $f \in \mathcal{L}_K(E, F)$; on montre que:

· si E_1 est un sous-espace vectoriel de E, $f(E_1)$ est un sous-espace vectoriel de F

· si F_1 est un sous-espace vectoriel de F, $f^{-1}(F_1)$ est un sous-espace vectoriel de E.

En particulier, nous obtenons les définitions suivantes:

Définitions: Soit $f \in \mathcal{L}_K(E, F)$.

L'ensemble $f(E) = \{y \in F \mid \exists x \in E, y = f(x)\}$ est un sous-espace vectoriel de F, appelé image de f et noté $\operatorname{Im} f$. Ainsi

$$\operatorname{Im} f = \{y \in F \mid \exists x \in E, y = f(x)\}$$

L'ensemble $f^{-1}(\{0_F\}) = \{x \in E \mid f(x) = 0_F\}$ est un sous-espace vectoriel de E, appelé noyau de f et noté $\ker f$. Ainsi

$$\mathrm{Ker}f = \{x \in E \mid f(x) = 0_F\}$$

Théorème des trois dimensions (ou théorème noyau-image): Soit $f \in \mathcal{L}_K(E, F)$. Si E est de dimension finie, alors

$$\dim E = \dim \ker f + \dim \mathrm{Im}\, f$$

Démonstration: Supposons $\dim E = n$. Comme $\mathrm{Ker}f$ est une partie de E, $\mathrm{Ker}f$ est aussi de dimension finie.

Posons $\dim\mathrm{ker}f = r \leq n$, et montrons que $\dim\mathrm{Im}f = n - r$.

Soit $\{w_1, \ldots, w_r\}$ une base de $\mathrm{Ker}f$; étendons-la à une base

$$\{w_1, \ldots, w_r, v_1, \ldots, v_{n-r}\}$$

de E.

Le théorème est démontré si nous montrons que $\mathcal{B} = \{f(v_1), \ldots, f(v_{n-r})\}$ est une base de $\mathrm{Im}f$.

- Montrons que $\mathcal{B}$ engendre $\mathrm{Im}f$.

$u \in \mathrm{Im}\, f \Longrightarrow \exists v \in E : u = f(v)$. Comme $v \in E$, il existe des scalaires a_i et b_i tels que

$$v = a_1 w_1 + \ldots + a_r w_r + b_1 v_1 + \ldots + b_{n-r} v_{n-r};$$

d'où $u = b_1 f(v_1) + \ldots + b_{n-r} f(v_{n-r})$ car les $w_i \in \mathrm{Ker}f$.

- Montrons que $\mathcal{B}$ est libre.

Soient des scalaires $\lambda_1, \ldots, \lambda_{n-r}$ tels que

$$\lambda_1 f(v_1) + \ldots + \lambda_{n-r} f(v_{n-r}) = 0;$$

on a alors

$$f(\lambda_1 v_1 + \ldots + \lambda_{n-r} v_{n-r}) = 0,$$

d'où $\lambda_1 v_1 + \ldots + \lambda_{n-r} v_{n-r} \in \mathrm{Ker}f$ et par suite il existe des scalaires μ_1, $\ldots$, μ_r tels que

$$\lambda_1 v_1 + \ldots + \lambda_{n-r} v_{n-r} = \mu_1 w_1 + \ldots + \mu_r w_r$$

i.e $\lambda_1 v_1 + \ldots + \lambda_{n-r} v_{n-r} - \mu_1 w_1 - \ldots - \mu_r w_r = 0$.

Comme $\{w_1, \ldots, w_r, v_1, \ldots, v_{n-r}\}$ est une base on déduit que

$$\lambda_1 = \ldots = \lambda_{n-r} = \mu_1 = \ldots = \mu_r = 0.$$

Ainsi $\mathcal{B} = \{f(v_1), \ldots, f(v_{n-r})\}$ est une base de Imf, donc

$$\dim \operatorname{Im} f = n - r = \dim E - \dim \ker f,$$

i.e

$$\dim E = \dim \ker f + \dim \operatorname{Im} f.$$

CQFD

Corollaire: Soit f un endomorphisme de E, et on suppose que dim$E$ est finie. Alors les énoncés suivants sont équivalents:

(i) f est injectif

(ii) f est surjectif

(iii) f est bijectif

Démonstration: on a $\dim E = \dim \ker f + \dim \operatorname{Im} f$.

f est injectif$\Longleftrightarrow \ker f = \{0_E\}$

$\Longleftrightarrow \dim \ker f = 0$

$\Longleftrightarrow \dim \operatorname{Im} f = \dim E$

$\Longleftrightarrow f$ est surjectif

$\Longleftrightarrow f$ est bijectif.

CQFD

Définition: Soit $f \in \mathcal{L}_K(E, F)$ avec dim$E$ finie, et F de dimension quelconque.

On appelle rang de f, l'entier noté rg(f) défini par:

$$\operatorname{rg}(f) = \dim \operatorname{Im} f = \dim f(E)$$

Autrement dit, si $\{u_1, \ldots, u_n\}$ est une base de E, alors

$$\operatorname{rg}(f) = \operatorname{rg}(\{f(u_1), \ldots, f(u_n)\}).$$

Conséquences du théorème noyau-image: On montre en exercice que:

1) Soit $f \in \mathcal{L}_K(E, F)$ où E et F sont de dimensions finies. On a

a) rg$(f) = \dim E$ si et seulement si f est injective.

b) rg$(f) = \dim F$ si et seulement si f est surjective.

Et comme situation particulière on a

2) Soit $f \in \mathcal{L}_K(E, F)$ où E et F sont de dimensions finies tels que $\dim E = \dim F$.

Alors les propriétés suivantes sont équivalentes:

a) f est injective

b) f est surjective

c) f est bijective

d) $\ker f = \{0_E\}$

e) $\mathrm{Im} f = F$

g) $\mathrm{rg}(f) = \dim E = \dim F$.

Chapter 7

Matrices

Ce chapitre introduit un outil de calcul très commode: les matrices. Celles-ci définies comme un tableau de nombres mais on en verra une interprétation plus abstraite. La technique centrale de ce chapitre est celle dite du " pivot de Gauss " qui est à la fois simple et efficace.

Dans tout le chapitre K désignera un corps commutatif.

7.1 Généralités

Définitions: Soient m, n deux entiers ≥ 1.

On appelle matrice $m \times n$ (on dit aussi matrice de type (m, n)) à coefficients dans K un tableau à m lignes et n colonnes d'éléments de K:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \text{ (avec des parenthèses),}$$

$$\text{ou encore } A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \text{ (avec des crochets).}$$

On note parfois brièvement $A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ et on dit que $m \times n$ est la taille de A. Les a_{ij} s'appellent les coefficients de la matrice A; on dit que

a_{ij} est le coefficient d'indices $(i,\, j)$ de A, i est le numéro de la ligne et j le numéro de la colonne. Le vecteur $(a_{1j},\, \ldots,\, a_{mj})$ de K^m est appelé $j-i$ème vecteur colonne de A; définition analogue pour les vecteurs lignes.

Si $m = n$, on dit que A est une matrice carrée d'ordre n. Si $n = 1$, A est une matrice colonne; si $m = 1$, une matrice ligne.

L'ensemble $\mathcal{M}_{m,n}(K)$ des matrices de type $(m,\, n)$ à coefficients dans K s'identifie à K^{mn}; il est donc muni d'une structure d'espace vectoriel sur K, de dimension mn.

Cas particuliers:

· La matrice nulle $m \times n$ est la matrice dont tous les coefficients sont nuls; on la note O_{mn} ou simplement O si le contexte rend clair sa taille.

· La matrice identité $m \times m$ est la matrice dont tous les coefficients sont nuls sauf ceux situés sur la diagonale principale qui valent 1; on la note I_m ou simplement I si le contexte rend clair sa taille. Ainsi

$$I = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \text{ (où les coefficients laissés en blanc sont nuls)}$$

· Une matrice diagonale $m \times m$ est une matrice dont tous les coefficients sont nuls sauf ceux situés sur la diagonale principale. Ainsi

$$D = \begin{pmatrix} a_{11} & & 0 \\ & a_{22} & \\ & \ddots & \\ 0 & & a_{mm} \end{pmatrix} \text{ (où les coefficients laissés en blanc sont nuls)}$$

· Une matrice triangulaire supérieure (resp. inférieure) $m \times m$ est une matrice dont tous les coefficients sont nuls sauf ceux situés sur la diagonale principale et au-dessus (resp. au-dessous) de la diagonale principale. Ainsi

$$T_1 = \underbrace{\begin{pmatrix} a_{11} & * & * & * \\ & a_{22} & * & * \\ & & \ddots & * \\ 0 & & & a_{mm} \end{pmatrix}}_{\text{Triangulaire supérieure}}, \quad T_2 = \underbrace{\begin{pmatrix} a_{11} & & & 0 \\ * & a_{22} & & \\ * & * & \ddots & \\ * & * & * & a_{mm} \end{pmatrix}}_{\text{Triangulaire inférieure}}$$

(où les coefficients laissés en blanc sont nuls et les étoiles désignent des coefficients quelconques).

Exemples:

$A = \begin{pmatrix} 1 & -5 & 0 & 3 \\ 6 & 8 & 4 & 2 \end{pmatrix}$ est une matrice 2×4 (on dit encore une matrice de type $(2, 4)$);
le coefficient a_{22} est égal à 8, alors que $a_{14} = 3$.

7.2 Opérations sur les matrices

7.2.1 Addition

Si $A = (a_{ij})$ et $B = (b_{ij})$ sont des matrices $m \times n$ alors $A + B$ est une matrice $m \times n$ dont les coefficients sont donnés par $c_{ij} = a_{ij} + b_{ij}$.

Exemple:

$$\begin{pmatrix} 1 & -5 & 2 \\ 6 & 8 & 4 \end{pmatrix} + \begin{pmatrix} -5 & 0 & 3 \\ 8 & 4 & 2 \end{pmatrix} = \begin{pmatrix} -4 & -5 & 5 \\ 14 & 12 & 6 \end{pmatrix}$$

7.2.2 Multiplication par un scalaire

si $A = (a_{ij})$ est une matrice $m \times n$ et λ un scalaire, alors λA est une matrice $m \times n$ dont les coefficients sont donnés par $c_{ij} = \lambda a_{ij}$.

Exemple:

$$3 \begin{pmatrix} 1 & -5 & 2 \\ 6 & 8 & 4 \end{pmatrix} = \begin{pmatrix} 3 & -15 & 6 \\ 18 & 24 & 12 \end{pmatrix}$$

7.2.3 Multiplication de deux matrices

Cette opération est la plus intéressante (et la plus compliquée).

Si $A = (a_{ij})$ et $B = (b_{ij})$ sont des matrices $m \times n$ et $n \times p$ respectivement, alors AB est une matrice $m \times p$ dont les coefficients sont donnés par $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

La multiplication de deux matrices n'est donc définie que si le nombre de colonnes de la première est égal au nombre de lignes de la seconde; et pour simplifier on écrit

$$\boxed{(m,\, n)\,(n,\, p) = (m,\, p)}$$

Le produit de deux matrices s'obtient en faisant le produit de chaque ligne de la première matrice par les colonnes de la seconde.

$$\begin{pmatrix} a_{11} & \dots & \dots & a_{1n} \\ \vdots & & & \\ \boxed{a_{i1} & \dots & \dots & a_{in}} \\ \vdots & & & \\ a_{m1} & \dots & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} \dots & \boxed{b_{1j}} & \dots b_{1p} \\ \vdots & \boxed{\vdots} & \vdots \\ \vdots & \boxed{\vdots} & \vdots \\ b_{n1} \dots & \boxed{b_{nj}} & \dots b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1p} \\ & & \\ & \boxed{c_{ij}} & \\ & & \\ c_{m1} & \dots & c_{mp} \end{pmatrix}$$

Exemple:

$$\begin{pmatrix} 1 & 2 \\ 0 & -1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 5 & -3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 9 & -1 \\ -2 & -1 \\ 10 & -6 \end{pmatrix}$$

Remarques:

· On peut faire le produit de matrices par blocs (de tailles compatibles):

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix} = \begin{pmatrix} AA_1 + BC_1 & AB_1 + BD_1 \\ CA_1 + DC_1 & CB_1 + DD_1 \end{pmatrix}$$

où A, B, C, D, A_1, B_1, C_1, D_1 sont des matrices, en les multipliant comme des scalaires, mais en faisant attention à ne pas les faire commuter car la multiplication des matrices n'est pas commutative: en général $AB \neq BA$; par exemple:

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

· Pour toute matrice A de taille $n \times p$ on a $O_{mn}A = O_{mp}$, et pour toute matrice B de taille $s \times m$ on a $BO_{mn} = O_{sn}$

· Pour toute matrice A de taille $m \times n$ on a $I_m A = A = AI_n$.

Proposition: Soient A, B, C des matrices et λ un scalaire. Les opérations (quand elles sont définies) sur les matrices vérifient les règles suivantes

a) Distributivité

$A(B+C) = AB + AC$ et $(A+B)C = AC + BC$

b) Associativité

$A(BC) = A(BC)$

c) Compatibilité

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

Démonstration: La vérification ne présente pas de difficulté.

7.3 Anneau des matrices carrées

On note simplement $\mathcal{M}_n(K)$ l'espace $\mathcal{M}_{n,n}(K)$ des matrices carrées d'ordre n à coefficients dans K. Il est muni d'une loi de produit qui est associative, distributive par rapport à l'addition et possède un élément unité, à savoir la matrice I_n. L'ensemble $\mathcal{M}_n(K)$ muni de l'addition et de la multiplication est donc un anneau; si $n = 1$ c'est un corps "égal" à K, si $n \geq 2$ c'est un anneau non commutatif qui n'est pas un corps. En effet, on a déjà vu que l'anneau n'est pas commutatif si $n = 2$ et de même ce n'est pas un corps car

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = O_{22}$$

est nul sans qu'aucun facteur ne soit nul.

Il est donc intéressant de chercher si une matrice a un inverse. Une matrice A est dite inversible s'il existe une matrice B telle que $AB = BA = I$; on pose alors $B = A^{-1}$. En fait l'une des égalités $AB = I$ ou $BA = I$ suffit à entraîner l'autre. Un intérêt clair du calcul de l'inverse d'une matrice (quand elle existe) est la résolution des systèmes d'équations linéaires associés: en effet, si A est inversible alors $AX = B$ équivaut à $X = A^{-1}B$. Il est néanmoins

rare qu'on ait recours à cette méthode: tout d'abord elle ne s'adapte qu'à des cas particuliers et ensuite il existe une méthode permettant de traiter tous les cas.

7.4 La méthode du pivot

On décrit une procédure de résolution des systèmes linéaires: l'idée est de se ramener à un système triangulaire que l'on peut ensuite résoudre facilement.

7.4.1 Matrices échelonnées

Définition: Une opération élémentaire sur les lignes d'une matrice consiste à:

· Remplacer une ligne L_i par la ligne $L_i + L_j$ (avec $i \neq j$).

· Echanger deux lignes.

· Multiplier une ligne par un scalaire non nul.

Remarque: en combinant ces opérations, on voit qu'on peut remplacer L_i par $L_i + \alpha L_j$.

Définition: Une matrice est échelonnée si

(i) Le premier coefficient non nul d'une ligne est 1 (on dit que c'est un pivot).

(ii) Le premier coefficient non nul de la $(i+1)-$éme ligne est à droite de celui de la $i-$éme ligne.

(iii) Les coefficients au-dessus d'un pivot sont nuls.

Exemples:

$$\begin{pmatrix} 1 & 8 & 0 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 & 0 & 0 & -6 & 8 & -4 \\ 0 & 1 & 0 & 3 & -2 & 0 \\ 0 & 0 & 1 & 0 & 9 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \text{ sont échelonnées.}$$

Remarques: Dans une matrice échelonnée, une ligne est soit nulle soit de la forme $(0, \ldots, 0, 1, \star, \ldots, \star)$; si une ligne est nulle, toutes les lignes situées en dessous sont nulles.

Indiquons comment en pratique, on peut appliquer des transformations élémentaires à n'importe quelle matrice pour la rendre échelonnée: on choisit un coefficient non nul situé le plus à gauche possible; par multiplication par

un scalaire et échange des lignes on se ramène au cas où ce coefficient est 1 et est situé sur la première ligne. En ajoutant à chacune des lignes un multiple adéquat de la première ligne on fait apparaître des zéros en dessous du pivot 1 de la première ligne. On répète l'opération sans toucher la première ligne et au bout d'un certain temps on arrive à une matrice du type:

$$\begin{pmatrix} 0\ldots\ 0\ 1\ \ \star\ldots\ \ \ldots\ldots\star \\ 0\ldots\ 0\ 0\ 1\star\ldots\ \ldots\ldots\star \\ 0\ldots\ 0\ 0\ 0\ldots 0\ 1\star\ldots\star \end{pmatrix}$$

Et il ne reste qu'à faire apparaître des zéros au dessus de chaque pivot, ce qui peut se faire en retranchant un multiple adéquat de la ligne du pivot.

7.4.2 Application de la méthode du pivot

Résolution des systèmes d'équations linéaires

On peut juxtaposer une matrice $m \times n$ et une matrice $m \times r$ pour obtenir une matrice $m \times (n+r)$. Si A et B sont les deux matrices initiales, on note la nouvelle matrice $(A \mid B)$.

Exemples:

Si $A = \begin{pmatrix} 1 & 8 & -2 & 2 \\ 0 & 7 & -5 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 6 & 9 & 0 \\ 9 & 0 & -7 \end{pmatrix}$ alors $(A \mid B) = \left(\begin{array}{cccc|ccc} 1 & 8 & -2 & 2 & 6 & 9 & 0 \\ 0 & 7 & -5 & 0 & 9 & 0 & -7 \end{array}\right)$.

Les matrices permettent de compacter des formules et donc de les manipuler de façon plus efficace; une application de cela est la résolution des systèmes d'équations linéaires, i.e des systèmes du type:

$$\mathcal{S} : \begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

On peut réécrire un tel système en posant $A = (a_{ij})$ c'est la matrice associée au système $\mathcal{S}$,

$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ alors le système $\mathcal{S}$ équivaut à :

$$AX = B$$

Proposition: Considérons le système linéaire

$$\mathcal{S} : AX = B \quad (i)$$

Supposons que l'on passe de la matrice $A_1 = (A \mid B)$ à la matrice $A_2 = (A' \mid B')$ par une succession d'opérations élémentaires, alors les solutions du système linéaire

$$\mathcal{S}' : A'X = B' \quad (ii)$$

sont les mêmes que celles du système $\mathcal{S}$; on dit que les deux systèmes sont équivalents.

Démonstration: Echanger l'ordre de deux équations ou en multiplier une par un scalaire non nul, ne change pas les solutions; passer de deux équations $L_1 = L_2 = 0$ à deux autres équations $L_1 + \alpha L_2 = L_2 = 0$ n'en change pas les solutions. CQFD

Exemple: Soit le système d'équations linéaires

$$\mathcal{S} : \begin{cases} x_1 + 2x_3 + 3x_4 = 4 \\ x_1 + x_2 - x_3 + x_4 = 0 \\ 2x_1 + x_2 - 3x_3 + 7x_4 = 2 \end{cases}$$

Transformons la matrice

$$(A \mid B) = \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 4 \\ 1 & 1 & -1 & 1 & 0 \\ 2 & 1 & -3 & 7 & 2 \end{array}\right)$$

Indiquons par une flèche le fait de passer à une autre matrice (par une opération élémentaire).

$$(A \mid B) = \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 4 \\ 1 & 1 & -1 & 1 & 0 \\ 2 & 1 & -3 & 7 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 4 \\ 0 & 1 & -3 & -2 & -4 \\ 0 & 1 & -7 & 1 & -6 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 4 \\ 0 & 1 & -3 & -2 & -4 \\ 0 & 0 & -4 & 3 & -2 \end{array}\right) = (A' \mid B')$$

Le système $AX = B$ possède les mêmes solutions que $A'X = B'$ i.e

$$\begin{cases} x_1 + 2x_3 + 3x_4 = 4 \\ x_2 - 3x_3 - 2x_4 = -4 \\ -4x_3 + 3x_4 = -2 \end{cases} \text{ d'où } \begin{cases} x_1 = \dfrac{9}{2}x_4 + 3 \\ x_2 = \dfrac{17}{4}x_4 - \dfrac{5}{2} \\ x_3 = \dfrac{3}{4}x_4 + \dfrac{1}{2} \end{cases}$$

Voyons maintenant quelle est la forme la plus simple que l'on puisse obtenir pour un système linéaire.

Théorème: Soit $A'X = B'$ un système d'équations linéaires tel que la matrice $(A' \mid B')$ soit échelonnée.

(i) Le système possède une solution si et seulement si il n'y a pas de pivot sur la dernière colonne.

(ii) S'il n'y a pas de pivot sur la dernière colonne, la solution générale du système s'obtient en fixant arbitrairement chacune des inconnues x_i telles que la i-ème colonne ne contienne pas de pivot et en calculant chacune des autres x_i en fonction de celles-là grâce à l'équation correspondant à la i-ème ligne.

Exemples: Soit à résoudre les systèmes précédents par des matrices échelonnées:

$$\mathcal{S} : \begin{cases} x_1 + 2x_3 + 3x_4 = 7 \\ x_1 + x_2 - x_3 + x_4 = 2 \\ 2x_1 + x_2 - 3x_3 + x_4 = 3 \end{cases}, \quad \mathcal{S}_1 : \begin{cases} x_1 + 2x_2 + 3x_3 = 3 \\ x_1 + x_2 - x_3 = -1 \\ 2x_1 + x_2 - x_3 = 0 \end{cases} \text{ et}$$

$$\mathcal{S}_2 : \begin{cases} x_1 + 2x_2 + 3x_3 = 3 \\ x_1 + x_2 - x_3 = -1 \\ 2x_1 + 3x_2 + 2x_3 = 0 \end{cases}$$

· Pour $\mathcal{S}$ on a:

$$(A \mid B) = \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 7 \\ 1 & 1 & -1 & 1 & 2 \\ 2 & 1 & -3 & 1 & 3 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 7 \\ 0 & 1 & -3 & -2 & -5 \\ 0 & 1 & -7 & -5 & -11 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 7 \\ 0 & 1 & -3 & -2 & -5 \\ 0 & 0 & -4 & -3 & -6 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 7 \\ 0 & 1 & -3 & -2 & -5 \\ 0 & 0 & 1 & \frac{3}{4} & \frac{3}{2} \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{3}{2} & 4 \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{1}{2} \\ 0 & 0 & 1 & \frac{3}{4} & \frac{3}{2} \end{array}\right) = (A' \mid B')$$

La matrice $(A' \mid B') = \left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{3}{2} & 4 \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{1}{2} \\ 0 & 0 & 1 & \frac{3}{4} & \frac{3}{2} \end{array}\right)$ est échelonnée et il n'y a pas de pivot sur la dernière colonne; le système $\mathcal{S} : AX = B$ est donc compatible et possède les mêmes solutions que $\begin{cases} x_1 + \frac{3}{2}x_4 = 4 \\ x_2 + \frac{1}{4}x_4 = -\frac{1}{2} \\ x_3 + \frac{3}{4}x_4 = \frac{3}{2} \end{cases}$ d'où

$$\begin{cases} x_1 = -\frac{3}{2}x_4 + 4 \\ x_2 = -\frac{1}{4}x_4 - \frac{1}{2} \\ x_3 = -\frac{3}{4}x_4 + \frac{3}{2} \end{cases}.$$

L'ensemble des solutions du système $\mathcal{S}$ est donc

$$\left\{\left(-\frac{3}{2}\alpha + 4,\ -\frac{1}{4}\alpha - \frac{1}{2},\ -\frac{3}{4}\alpha + \frac{3}{2},\ \alpha\right) \mid \alpha \in \mathbb{R}\right\}.$$

· Pour $\mathcal{S}_1$ on a:

$$(A_1 \mid B_1) = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 1 & 1 & -1 & -1 \\ 2 & 1 & -1 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & -1 & -4 & -4 \\ 0 & -3 & -7 & -6 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & -1 & -4 & -4 \\ 0 & 0 & 5 & 6 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 1 & 4 & 4 \\ 0 & 0 & 1 & \frac{6}{5} \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & -5 & -5 \\ 0 & 1 & 4 & 4 \\ 0 & 0 & 1 & \frac{6}{5} \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -\frac{4}{5} \\ 0 & 0 & 1 & \frac{6}{5} \end{array}\right) = (A_1' \mid B_1')$$

La matrice $(A_1' \mid B_1') = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -\frac{4}{5} \\ 0 & 0 & 1 & \frac{6}{5} \end{array}\right)$ est échelonnée et il n'y a pas de pivot sur la dernière colonne; le système $\mathcal{S}_1 : A_1X = B_1$ est donc compatible et possède les mêmes solutions que $\begin{cases} x_1 = 1 \\ x_2 = -\dfrac{4}{5} \\ x_3 = \dfrac{6}{5} \end{cases}$.

L'ensemble des solutions du système $\mathcal{S}$ est donc le singleton $\left\{\left(1, -\frac{4}{5}, \frac{6}{5}\right)\right\}$.

· Pour $\mathcal{S}_2$ on a:

$$(A_2 \mid B_2) = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 1 & 1 & -1 & -1 \\ 2 & 3 & 2 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & -1 & -4 & -4 \\ 0 & -1 & -4 & -6 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & -1 & -4 & -4 \\ 0 & 0 & 0 & -2 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 1 & 4 & 4 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & -5 & -5 \\ 0 & 1 & 4 & 4 \\ 0 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & -5 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) = (A_2' \mid B_2')$$

La matrice $(A_2' \mid B_2') = \left(\begin{array}{ccc|c} 1 & 0 & -5 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$ est échelonnée et il y a un pivot sur la dernière colonne,donc le système $\mathcal{S}_2 : A_2X = B_2$ est incompatible.

Calcul de l'inverse d'une matrice inversible.

Définition: Deux matrices carrées A et A' sont dites semblables s'il existe une matrice inversible P telle que

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

Pour déterminer l'inverse d'une matrice carrée A de type $n \times n$, on réduit par opérations élémentaires la matrice $(A \mid I_n)$ à une matrice échelonnée; si la matrice A est inversible la matrice échelonnée obtenue sera $(I_n \mid A^{-1})$.

Exemples: Soient les matrices

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

· Pour A on a:

$$(A \mid I_3) = \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 0 & 1 & 5 & -2 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 0 & 1 & 5 & -2 & 1 & 0 \\ 0 & 0 & -9 & 4 & -2 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 0 & 0 \\ 0 & 1 & 5 & -2 & 1 & 0 \\ 0 & 0 & 1 & -\frac{4}{9} & \frac{2}{9} & -\frac{1}{9} \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{15}{7} & -\frac{4}{7} & \frac{2}{7} \\ 0 & 1 & 5 & -2 & 1 & 0 \\ 0 & 0 & 1 & -\frac{4}{9} & \frac{2}{9} & -\frac{1}{9} \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{9} & \frac{4}{9} & -\frac{2}{9} \\ 0 & 1 & 0 & \frac{2}{9} & \frac{19}{9} & -\frac{5}{9} \\ 0 & 0 & 1 & -\frac{4}{9} & \frac{2}{9} & -\frac{1}{9} \end{array}\right) = (I_3 \mid A^{-1})$$

Ainsi l'on a $A^{-1} = \begin{pmatrix} \frac{1}{9} & \frac{4}{9} & -\frac{2}{9} \\ \frac{2}{9} & \frac{19}{9} & -\frac{5}{9} \\ -\frac{4}{9} & \frac{2}{9} & -\frac{1}{9} \end{pmatrix}$.

· Pour B on a:

$$(B \mid I_3) = \left(\begin{array}{ccc|ccc} 2 & 1 & 4 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ -1 & 2 & -3 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 0 & 1 & 0 \\ 2 & 1 & 4 & 1 & 0 & 0 \\ -1 & 2 & -3 & 0 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & -5 & 2 & 1 & -2 & 0 \\ 0 & 5 & -2 & 0 & 1 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & -5 & 2 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & 1 & -\frac{2}{5} & -\frac{1}{5} & \frac{2}{5} & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & \frac{11}{5} & \frac{3}{5} & -\frac{1}{5} & 0 \\ 0 & 1 & -\frac{2}{5} & -\frac{1}{5} & \frac{2}{5} & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & \frac{11}{5} & 0 & \frac{2}{5} & -\frac{3}{5} \\ 0 & 1 & -\frac{2}{5} & 0 & \frac{1}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 1 & -1 & 1 \end{array}\right)$$

Ainsi B n'est pas inversible.

Détermination du rang d'une matrice.

Le rang d'une matrice A que l'on note $\mathrm{rg}(A)$ est le nombre de lignes non nulles obtenu après les transformations par la méthode du pivot.

Exemples:

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \\ 0 & 2 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 7 \end{pmatrix} \text{ d'où } \mathrm{rg}(A) = 3.$$

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

$$\longrightarrow \begin{pmatrix} 1 & 3 & 1 \\ 0 & -5 & 2 \\ 0 & 0 & 0 \end{pmatrix} \text{ ce qui montre que } \mathrm{rg}(B) = 2.$$

Conséquences: Soit

$$\mathcal{S}:\begin{cases} a_{11}x_1+\ldots+a_{1n}x_n=b_1 \\ \vdots \\ a_{m1}x_1+\ldots+a_{mn}x_n=b_m \end{cases}$$

un système d'équations linéaire.

· Si le système est homogène ie $b_1=\ldots=b_m=0$, alors le système est toujours compatible, en particulier $X=0\in K^n$ est solution; et il y en a d'autres si et seulement si $\mathrm{rg}(A)<n$, n étant le nombre de variables et A la matrice associée au système $\mathcal{S}$.

· Si le système n'est pas homogène, alors il est compatible si et seulement si

$$\mathrm{rg}(A)=\mathrm{rg}(A\mid B)$$

avec $B=\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$.

Exercices:

· Dire si les systèmes suivants admettent des solutions autres que la solution nulle:

$$\mathcal{S}:\begin{cases} x_1+2x_2=0 \\ x_2-x_3=0 \\ x_1+3x_2-x_3=0 \end{cases},\quad \mathcal{S}_1:\begin{cases} x_1+x_2-x_3=0 \\ x_1-x_2-x_3=0 \\ 2x_1+x_2+x_3=0 \end{cases}$$

· Dire si les systèmes suivants sont compatibles:

$$\mathcal{S}:\begin{cases} x_1-x_2+3x_3+2x_4=1 \\ x_1+2x_2-x_3+x_4=0 \\ 2x_1+x_2+2x_3+3x_4=-1 \end{cases},\quad \mathcal{S}_1:\begin{cases} x_1+2x_2-x_3=0 \\ x_1-x_2+x_3=1 \\ 2x_1+x_2+x_3=2 \end{cases}$$

7.5 Matrice d'une application linéaire

Un moyen simple de construire une application linéaire est de fixer l'image des éléments d'une base de l'espace de départ $f(e_1),\ldots,f(e_n)$. Si l'on dispose d'une base de l'espace d'arrivée on peut exprimer les vecteurs $f(e_i)$ comme

combinaisons linéaires des vecteurs de la base. On arrive à une description de f par un ensemble de nombres qui se range naturellement dans un tableau, ie une matrice. Cette construction est l'inverse de celle qui à une matrice A associe l'application linéaire $X \longrightarrow AX$ et est extrêmement utile pour l'étude des applications linéaires et des matrices.

Définition: Soient E et F deux espaces vectoriels de dimensions respectives n et p sur le même corps commutatif K; $n, p \in \mathbb{N}^*$ et soit $f \in \mathcal{L}_K(E, F)$. Si $\mathcal{A} = \{a_1, \ldots, a_n\}$ et $\mathcal{B} = \{b_1, \ldots, b_p\}$ sont des bases respectives de E et F, on appelle matrice de f relativement aux bases $\mathcal{A}$ et $\mathcal{B}$ la matrice définie par:

$\mathcal{M}(f, \mathcal{A}, \mathcal{B}) = (x_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ où x_{ij} est la coordonnée de $f(a_j)$ sur b_i. On écrira

$$\mathcal{M}(f, \mathcal{A}, \mathcal{B}) = \begin{array}{c} \begin{array}{ccc} f(a_1) & \ldots & f(a_n) \end{array} \\ \begin{pmatrix} x_{11} & \ldots & x_{1n} \\ \vdots & \vdots & \vdots \\ x_{p1} & \ldots & x_{pn} \end{pmatrix} \end{array} \begin{array}{c} b_1 \\ \vdots \\ b_p \end{array}$$

Les colonnes de $\mathcal{M}(f, \mathcal{A}, \mathcal{B})$ sont les coordonnées des vecteurs $f(a_j)$ dans la base $\mathcal{B} = \{b_1, \ldots, b_p\}$.

Si $\mathcal{A} = \mathcal{B}$ on parlera de la matrice de f relativement à la base $\mathcal{A}$ et l'on note $\mathcal{M}(f, \mathcal{A})$.

La correspondance entre matrices et applications linéaires est si naturelle que la composition d'applications linéaires correspond à la multiplication des matrices. On a:

· $\mathcal{M}(g \circ f) = \mathcal{M}(g)\,\mathcal{M}(f)$

· $\mathcal{M}(f + g) = \mathcal{M}(f) + \mathcal{M}(g)$

· $\mathcal{M}(\lambda f) = \lambda \mathcal{M}(f)$, $\lambda \in K$.

Exemples:

· Soient l'application linéaire

$$\begin{array}{c} f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2 \\ (x, y, z) \longmapsto (2x + y, x - y + z) \end{array},$$

$\mathcal{A} = \{(1, 0, -1), (2, 1, 1), (0, 1, 1)\}$ une base de $\mathbb{R}^3$ et

$\mathcal{B} = \{(-1,\, 3)\,,\ (2,\, 1)\}$ une base de $\mathbb{R}^2$.

On note $\mathcal{E} = \{e_1,\, e_2\}$ la base canonique de $\mathbb{R}^2$.

Déterminer $\mathcal{M}\,(f,\, \mathcal{A},\, \mathcal{E})$, $\mathcal{M}\,(f,\, \mathcal{A},\, \mathcal{B})$.

Posons $a_1 = (1,\, 0,\, -1)$, $a_2 = (2,\, 1,\, 1)$, $a_3 = (0,\, 1,\, 1)$ et $b_1 = (-1,\, 3)$, $b_2 = (2,\, 1)$. On a

$$\mathcal{M}\,(f,\, \mathcal{A},\, \mathcal{E}) = \begin{array}{c} \begin{array}{ccc} f\,(a_1) & f\,(a_2) & f\,(a_3) \end{array} \\ \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{pmatrix} \\ \end{array} \begin{array}{c} \\ e_1 \\ e_2 \end{array}$$

$f\,(a_1) = f\,(1,\, 0,\, -1) = (2,\, 0)$, $f\,(a_2) = f\,(2,\, 1,\, 1) = (5,\, 2)$, $f\,(a_3) = f\,(0,\, 1,\, 1) = (1,\, 0)$.

Ainsi on a $\mathcal{M}\,(f,\, \mathcal{A},\, \mathcal{E}) = \begin{pmatrix} 2 & 5 & 1 \\ 0 & 2 & 0 \end{pmatrix}$.

$$\mathcal{M}\,(f,\, \mathcal{A},\, \mathcal{B}) = \begin{array}{c} \begin{array}{ccc} f\,(a_1) & f\,(a_2) & f\,(a_3) \end{array} \\ \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{pmatrix} \\ \end{array} \begin{array}{c} \\ b_1 \\ b_2 \end{array}$$

$f\,(a_1) = x_{11}b_1 + x_{21}b_2$ i.e $(2,\, 0) = x_{11}\,(-1,\, 3) + x_{21}\,(2,\, 1)$, d'où

$$x_{11} = -\frac{2}{7} \text{ et } x_{21} = \frac{6}{7}.$$

$f\,(a_2) = x_{12}b_1 + x_{22}b_2$ i.e $(5,\, 2) = x_{12}\,(-1,\, 3) + x_{22}\,(2,\, 1)$, d'où

$$x_{12} = -\frac{1}{7} \text{ et } x_{22} = \frac{17}{7}.$$

$f\,(a_3) = x_{13}b_1 + x_{23}b_2$ i.e $(1,\, 0) = x_{13}\,(-1,\, 3) + x_{23}\,(2,\, 1)$, d'où

$$x_{13} = -\frac{1}{7} \text{ et } x_{23} = \frac{3}{7}.$$

Ainsi on a

$$\mathcal{M}(f, \mathcal{A}, \mathcal{B}) = \frac{1}{7}\begin{pmatrix} -2 & -1 & -1 \\ 6 & 17 & 3 \end{pmatrix}.$$

· Soient l'application linéaire

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x, y) \longmapsto (2x + y, x - y)$$

et $\mathcal{A} = \{(1, -1), (1, 1)\}$ une base de $\mathbb{R}^2$.
Déterminer $\mathcal{M}(f, \mathcal{A})$ la matrice de f relativement à la base $\mathcal{A}$.
Posons $a_1 = (1, -1)$, $a_2 = (1, 1)$, on a

$$\mathcal{M}(f, \mathcal{A}) = \begin{matrix} \\ \\ \\ \end{matrix}\begin{matrix} f(a_1) & f(a_2) \end{matrix}\begin{matrix} \\ a_1 \\ a_2 \end{matrix}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$$

$f(a_1) = x_{11}a_1 + x_{21}a_2$ i.e $(1, 2) = x_{11}(1, -1) + x_{21}(1, 1)$, d'où

$$x_{11} = -\frac{1}{2} \text{ et } x_{21} = \frac{3}{2}.$$

$f(a_2) = x_{12}a_1 + x_{22}a_2$ i.e $(3, 0) = x_{11}(1, -1) + x_{21}(1, 1)$, d'où

$$x_{11} = \frac{1}{2} \text{ et } x_{21} = \frac{5}{2}.$$

Ainsi on a

$$\mathcal{M}(f, \mathcal{A}) = \frac{1}{2}\begin{pmatrix} -1 & 1 \\ 3 & 5 \end{pmatrix}.$$

7.6 Matrice de changement de base

Définition: Soit E un K-espace vectoriel de dimension n, $n \in \mathbb{N}^*$, et soient $\mathcal{A} = \{a_1, \ldots, a_n\}$ et $\mathcal{A}' = \{a'_1, \ldots, a'_n\}$ deux bases de E. On appelle matrice de passage (ou matrice de changement de base) de $\mathcal{A}$ à $\mathcal{A}'$ la matrice $P = (x_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ dont la j-ième colonne est formée des coordonnées de a'_j dans la base $\mathcal{A}$. Ainsi

$$P = \mathcal{M}(id, \mathcal{A}', \mathcal{A}) = \begin{array}{c} \begin{array}{ccc} a'_1 & \ldots & a'_n \end{array} \\ \begin{pmatrix} x_{11} & \ldots & x_{1n} \\ \vdots & \vdots & \vdots \\ x_{n1} & \ldots & x_{nn} \end{pmatrix} \end{array} \begin{array}{c} a_1 \\ \vdots \\ a_n \end{array}$$

Notation:

Nous noterons souvent $\mathcal{P}(\mathcal{A}, \mathcal{A}')$ la matrice de passage $\mathcal{M}(id, \mathcal{A}', \mathcal{A})$ de $\mathcal{A}$ à $\mathcal{A}'$;

id étant l'application identité.

Exemples: Soient $\mathcal{A} = \{(1, 2), (1, 0)\}$ et $\mathcal{A}' = \{(-2, 1), (3, 2)\}$ deux bases de $\mathbb{R}^2$; on note $\mathcal{E} = \{e_1, e_2\}$ la base canonique de $\mathbb{R}^2$.

Déterminer $\mathcal{P}(\mathcal{E}, \mathcal{A}')$, $\mathcal{P}(\mathcal{A}, \mathcal{E})$ et $\mathcal{P}(\mathcal{A}, \mathcal{A}')$.

On a

$$\mathcal{P}(\mathcal{E}, \mathcal{A}') = \begin{array}{c} \begin{array}{cc} a'_1 & a'_2 \end{array} \\ \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \end{array} \begin{array}{c} e_1 \\ e_2 \end{array}$$

d'où

$$\mathcal{P}(\mathcal{E}, \mathcal{A}') = \begin{pmatrix} -2 & 3 \\ 1 & 2 \end{pmatrix};$$

$$\mathcal{P}(\mathcal{A}, \mathcal{E}) = \begin{array}{c} \begin{array}{cc} e_1 & e_2 \end{array} \\ \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \end{array} \begin{array}{c} a_1 \\ a_2 \end{array},$$

$e_1 = x_{11}a_1 + x_{21}a_2$ i.e $(1, 0) = x_{11}(1, 2) + x_{21}(1, 0)$, d'où

$$x_{11} = 0 \text{ et } x_{21} = 1;$$

$e_2 = x_{12}a_1 + x_{22}a_2$ i.e $(0,\, 1) = x_{12}\,(1,\, 2) + x_{22}\,(1,\, 0)$, d'où

$$x_{12} = \frac{1}{2} \text{ et } x_{22} = -\frac{1}{2}.$$

Ainsi on a

$$\mathcal{P}\,(\mathcal{A},\, \mathcal{E}) = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix}.$$

$$\mathcal{P}\,(\mathcal{A},\, \mathcal{A}') = \begin{array}{c} \begin{array}{cc} a'_1 & a'_2 \end{array} \\ \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \end{array} \begin{array}{c} \\ \begin{array}{c} a_1 \\ a_2 \end{array} \end{array},$$

$a'_1 = x_{11}a_1 + x_{21}a_2$ i.e $(-2,\, 1) = x_{11}\,(1,\, 2) + x_{21}\,(1,\, 0)$, d'où

$$x_{11} = \frac{1}{2} \text{ et } x_{21} = -\frac{5}{2};$$

$a'_2 = x_{12}a_1 + x_{22}a_2$ i.e $(3,\, 2) = x_{12}\,(1,\, 2) + x_{22}\,(1,\, 0)$, d'où

$$x_{12} = 1 \text{ et } x_{22} = 2.$$

Ainsi on a

$$\mathcal{P}\,(\mathcal{A},\, \mathcal{A}') = \begin{pmatrix} \frac{1}{2} & 1 \\ -\frac{5}{2} & 2 \end{pmatrix}.$$

Remarque: Soit $u \in E$; écrivons

$$u = \sum_{1 \le i \le n} x_i a_i = \sum_{1 \le j \le n} x'_j a'_j$$

i.e $u_{\mathcal{A}} = (x_1,\, \ldots,\, x_n)$ (coordonnées de u dans la base $\mathcal{A}$) et
$u_{\mathcal{A}'} = (x'_1,\, \ldots,\, x'_n)$ (coordonnées de u dans la base $\mathcal{A}'$). On a alors

$$u_{\mathcal{A}} = \mathcal{P}\,(\mathcal{A},\, \mathcal{A}')\, u_{\mathcal{A}'}$$

En particulier $\mathcal{P}(\mathcal{A}, \mathcal{A}')$ est inversible et son inverse est $\mathcal{P}(\mathcal{A}', \mathcal{A})$; ainsi

$$u_{\mathcal{A}'} = \mathcal{P}(\mathcal{A}', \mathcal{A})\, u_{\mathcal{A}}.$$

Exemples: Soient $\mathcal{A} = \{(1, 2),\ (1, 0)\}$ et $\mathcal{A}' = \{(-2, 1),\ (3, 2)\}$ deux bases de $\mathbb{R}^2$, et soient $u = (2, 3)$, $v_{\mathcal{A}} = (-1, 1)$.

Déterminer $u_{\mathcal{A}}$ et v. Notons $\mathcal{E}$ la base canonique de $\mathbb{R}^2$.

On a $u_{\mathcal{A}} = \mathcal{P}(\mathcal{A}, \mathcal{E})\, u$, or dans l'exemple précédent on montré que

$$\mathcal{P}(\mathcal{A}, \mathcal{E}) = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix},$$

donc la forme matricielle donne

$u_{\mathcal{A}} = \mathcal{P}(\mathcal{A}, \mathcal{E})\, u = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, et par suite on a

$u_{\mathcal{A}} = \left(\frac{3}{2}, \frac{1}{2}\right)$.

On a $v = \mathcal{P}(\mathcal{E}, \mathcal{A})\ v_{\mathcal{A}}$; le calcul donne $\mathcal{P}(\mathcal{E}, \mathcal{A}) = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$, d'où

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

Ainsi $v = (0,\ -2)$.

Théorème: Soit $f : E \longrightarrow F$ une application linéaire et soient
$\mathcal{A} = \{a_1, \ldots, a_n\}$ et $\mathcal{A}' = \{a'_1, \ldots, a'_n\}$ deux bases de E, et
$\mathcal{B} = \{b_1, \ldots, b_m\}$, $\mathcal{B}' = \{b'_1, \ldots, b'_m\}$ deux bases de F;
notons
$A = \mathcal{M}(f, \mathcal{A}, \mathcal{B})$ et $A' = \mathcal{M}(f, \mathcal{A}', \mathcal{B}')$, $P = \mathcal{P}(\mathcal{A}, \mathcal{A}')$ et $Q = \mathcal{P}(\mathcal{B}, \mathcal{B}')$.
Alors

$$A' = Q^{-1} A P$$

Démonstration: Considérons la suite d'applications:

$$(E, \mathcal{A}') \xrightarrow{id_E} (E, \mathcal{A}) \xrightarrow{f} (F, \mathcal{B}) \xrightarrow{id_F} (F, \mathcal{B}')$$

on peut en déduire l'égalité des matrices:

$$\mathcal{M}(f, \mathcal{A}', \mathcal{B}') = \mathcal{M}(id_F, \mathcal{B}, \mathcal{B}')\,\mathcal{M}(f, \mathcal{A}, \mathcal{B})\,\mathcal{M}(id_E, \mathcal{A}', \mathcal{A})$$

ce qui correspond à l'égalité $A' = Q^{-1}AP$. CQFD

Définition: Le rang d'une matrice A est le rang de la suite de ses vecteurs colonnes.

Proposition: Soient E, F deux K-espaces vectoriels munis respectivement de la base $\mathcal{A} = \{a_1, \ldots, a_n\}$, $\mathcal{B} = \{b_1, \ldots, b_p\}$.

Si $f \in \mathcal{L}_K(E, F)$, alors $\mathrm{rg}(\mathcal{M}(f, \mathcal{A}, \mathcal{B})) = \mathrm{rg}(f)$.

Démonstration: $\mathrm{rg}(f) = \dim \mathrm{Im}\, f$. Or $\mathrm{Im}\, f$ est engendré par $\{f(a_1), \ldots, f(a_n)\}$ qui s'identifie dans la base $\mathcal{B}$ à la suite des vecteurs colonnes de $\mathcal{M}(f, \mathcal{A}, \mathcal{B})$. CQFD

Chapter 8

Déterminants

8.1 définition par récurrence

Définition: Soit A une matrice carrée d'ordre n

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \in \mathcal{M}_n(K)$$

On définit, par récurrence, une application:

$$\begin{aligned} \text{dét: } \mathcal{M}_n(K) &\longrightarrow K \\ A &\longmapsto détA \end{aligned}$$

de la manière suivante:

· si $n = 1$, i.e $A = (a)$, on pose $détA = a$

· si $n > 1$, notons A_{ij} la matrice obtenue de A en supprimant la $i - éme$ ligne et la $j - éme$ colonne (i.e la ligne et la colonne qui passent par a_{ij}). On pose alors (puisque $A_{ij} \in \mathcal{M}_{n-1}(K)$):

$$détA = a_{11}détA_{11} + \ldots + (-1)^{k+1} a_{1k}détA_{1k} + \ldots + (-1)^{n+1} a_{1n}détA_{1n}$$

Le scalaire $détA$ est appelé déterminant de A et on note

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

le déterminant de la matrice $\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$.

Exemples:

1) $\begin{vmatrix} 4 & -1 \\ 3 & 2 \end{vmatrix} = (4 \times 2) + (-1)^3 (-1)(3) = 11.$

Plus généralement on a:

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

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

$$= (6) + (6) + (27) = 39.$$

Plus généralement on peut appliquer la règle de SARRUS pour le calcul d'un déterminant d'ordre 3:

Le déterminant d'une matrice d'ordre 3 est la somme de six termes, trois affectés du signe + et trois du signe −:

· Les produits affectés du signe + contiennent soit les trois termes de la diagonale principale, soit deux termes parallèles à cette diagonale.

· Pour les produits affectés du signe −, on procède de même en changeant de diagonale.

On peut utiliser les dispositions suivantes:

a) On ajoute à droite les deux premières colonnes:

Calculons $\begin{vmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \\ 1 & 5 & 1 \end{vmatrix}$,

on a $\begin{vmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \\ 1 & 5 & 1 \end{vmatrix} \begin{matrix} 1 & -2 \\ 2 & 1 \\ 1 & 5 \end{matrix}$

d'où

$$\begin{vmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \\ 1 & 5 & 1 \end{vmatrix} = (1)(1)(1) + (-2)(-1)(1) + (3)(2)(5) - (1)(1)(3) - (5)(-1)(1) - (1)(2)(-2) = 39.$$

b) On ajoute en bas les deux premières lignes:

Calculons $\begin{vmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \\ 1 & 5 & 1 \end{vmatrix}$,

on a $\begin{vmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \\ 1 & 5 & 1 \end{vmatrix} \begin{matrix} 1 & -2 \\ 2 & 1 \end{matrix}$ d'où

$$\begin{vmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \\ 1 & 5 & 1 \end{vmatrix} = (1)(1)(1)+(2)(5)(3)+(1)(-2)(-1)-(1)(1)(3)-(1)(5)(-1)-(2)(-2)(1) = 39.$$

Définition: Soit $A = (a_{ij}) \in \mathcal{M}_n(K)$.
On appelle cofacteur de l'élément a_{ij} le scalaire

$$cof(a_{ij}) = (-1)^{i+j} \, dét A_{ij}$$

Exemples: $A = \begin{pmatrix} 0 & 2 \\ 1 & 5 \end{pmatrix}$,

$cof(a_{11}) = (-1)^2(5) = 5$, $cof(a_{12}) = (-1)^3(1) = -1$, $cof(a_{21}) = (-1)^3(2) = -2$, $cof(a_{22}) = (-1)^4(0) = 0$.

$$B = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 1 & -1 \\ 1 & 5 & 1 \end{pmatrix},$$

$$cof(b_{23}) = (-1)^5 \begin{vmatrix} 1 & -2 \\ 1 & 5 \end{vmatrix} = -7,\ cof(b_{31}) = (-1)^4 \begin{vmatrix} -2 & 3 \\ 1 & -1 \end{vmatrix} = -1.$$

On peut alors reformuler la définition du déterminant de la façon suivante:

Soit $A = (a_{ij}) \in \mathcal{M}_n(K)$, alors:

· Le développement du déterminant suivant la $i - éme$ ligne est donné par:

$$détA = a_{i1}cof(a_{i1}) + a_{i2}cof(a_{i2}) + \ldots + a_{in}cof(a_{in})$$

· Le développement du déterminant suivant la $j - éme$ colonne est donné par:

$$détA = a_{1j}cof(a_{1j}) + a_{2j}cof(a_{2j}) + \ldots + a_{nj}cof(a_{nj})$$

Exemples: Soit $A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 2 \\ -1 & 3 & 0 \end{pmatrix}$

- Développement suivant la $2^{éme}$ ligne:

$$détA = (2)(-1)^3 \begin{vmatrix} 1 & -1 \\ 3 & 0 \end{vmatrix} + (2)(-1)^5 \begin{vmatrix} 1 & 1 \\ -1 & 3 \end{vmatrix}$$
$$= (2)(-1)(3) + (2)(-1)(4) = -14.$$

- Développement suivant la $3^{éme}$ colonne:

$$détA = (-1)(-1)^4 \begin{vmatrix} 2 & 0 \\ -1 & 3 \end{vmatrix} + (2)(-1)^5 \begin{vmatrix} 1 & 1 \\ -1 & 3 \end{vmatrix}$$
$$= (-1)(1)(6) + (2)(-1)(4) = -14.$$

Remarque: Par transposition, les formules précédentes donnent des formules de développement par rapport à une colonne ou à une ligne quelconque:

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

Pour se souvenir des signes de ces deux formules, on peut remarquer que la distribution des signes + et − avec la formule $(-1)^{i+j}$ est analogue à la distribution des cases noires et blanches sur un damier:

$$\begin{vmatrix} + & - & + & - & \cdots \\ - & + & - & + & \ldots \\ + & - & + & - & \ldots \\ - & + & - & + & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$

8.2 Formes n-linéaires alternées

Définition: Soient E un k-espace vectoriel et $n \geq 1$ un entier, une application $f : E \times \ldots \times E \longrightarrow K$ est n-linéaire si elle est linéaire par rapport à chaque variable, i.e si

$$f(x_1, \ldots, x_{i-1}, \alpha y_i + \beta x_i, x_{i+1}, \ldots, x_n) =$$
$$\alpha f(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n) + \beta f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$$

Elle est symétrique si elle est invariante par permutation des facteurs, et antisymétrique ou alternée si l'interversion de deux facteurs change le signe, i.e si

$$\forall (x_1, \ldots, x_n) \in E^n, \forall i < j,$$
$$f(x_1, \ldots, x_i, \ldots x_j, \ldots, x_n) = -\, f(x_1, \ldots, x_j, \ldots x_i, \ldots, x_n)$$

Remarques: On déduit de la définition que:

1) si f est alternée, alors $f(\ldots, x, \ldots, x, \ldots) = 0$; et plus généralement que si x_i est combinaison linéaire des autres vecteurs $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ alors $f(x_1, \ldots, x_n) = 0$.

2) si $\{e_1, \ldots, e_n\}$ est une base de E et si f est une forme n-linéaire alternée non nulle, alors $f(e_1, \ldots, e_n) \neq 0$.

Définition: Soit $\{e_1, \ldots, e_n\}$ une base de E, le déterminant par rapport à la base $\{e_1, \ldots, e_n\}$ est l'unique application n-linéaire alternée

$$dét : E^n \longrightarrow K$$

telle que $dét\,(e_1, \ldots, e_n) = 1$.

La propriété fondamentale du déterminant est:

Théorème: Le déterminant de n vecteurs $x_1, \ldots, x_n$ dans K^n est nul si et seulement si ces vecteurs sont linéairement dépendants.

Démonstration: On sait que si les vecteurs sont linéairement dépendants, le déterminant est nul; inversement si les vecteurs $x_1, \ldots, x_n$ sont linéairement indépendants ils forment une base de K^n (puisque $\dim K^n = n$) et donc $dét\,(x_1, \ldots, x_n) \neq 0$. CQFD

Définition: Le déterminant d'une matrice carrée est le déterminant de ses vecteurs colonnes par rapport à la base canonique de K^n.

Le théorème précédent se traduit alors ainsi:

Théorème: Soit A une matrice carrée, alors A est inversible si et seulement si $détA \neq 0$.

Théorème: Le déterminant des matrices est multiplicatif, i.e si A et B sont deux matrices carrées de même ordre n on a:

$$dét\,(A.B) = détA.détB$$

Démonstration**:** Soient $(x_1, \ldots, x_n) \in K^n$, les applications n-linéaires alternées de K^n vers K définies par

$$(x_1, \ldots, x_n) \longmapsto dét\,(Ax_1, \ldots, Ax_n) \text{ et } (x_1, \ldots, x_n) \longmapsto dét\,(x_1, \ldots, x_n)$$

sont proportionnelles; donc $dét\,(Ax_1, \ldots, Ax_n) = \alpha dét\,(x_1, \ldots, x_n)$. En applicant ceci aux vecteurs de la base canonique, on obtient que $\alpha = détA$. Ainsi $dét\,(Ax_1, \ldots, Ax_n) = détA.dét\,(x_1, \ldots, x_n)$. Maitenant on peut calculer:

$$\begin{aligned} dét\,(AB)\,.dét\,(x_1, \ldots, x_n) &= dét\,(ABx_1, \ldots, ABx_n) \\ &= détA.dét\,(Bx_1, \ldots, Bx_n) \\ &= détA.détB.dét\,(x_1, \ldots, x_n) \end{aligned}$$

d'où $dét(AB) = détA.détB$; puisque $dét\,(x_1, \ldots, x_n) \neq 0$. CQFD

8.3 Règles de calcul

Soit A une matrice carrée.

1) Si l'on échange deux colonnes de A alors $détA$ est transformé en son opposé.

Exemple: $A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 2 \\ -1 & 3 & 0 \end{pmatrix}$,

on montre que $\begin{vmatrix} 1 & 1 & -1 \\ 2 & 0 & 2 \\ -1 & 3 & 0 \end{vmatrix} = -14$ et si on échange par exemple la première et la deuxième colonne on obtient $\begin{vmatrix} 1 & 1 & -1 \\ 0 & 2 & 2 \\ 3 & -1 & 0 \end{vmatrix} = 14 = -détA$.

2) $détA$ reste inchangé si l'on ajoute à un vecteur colonne une combinaison linéaire des autres.

N.B: Cette règle permet de faire apparaître des zéros dans le calcul du déterminant et de réduire ainsi le nombre de termes, ce qui facilite le calcul.

Exemple: $A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 2 \\ -1 & 3 & 0 \end{pmatrix}$, on a

$$\begin{vmatrix} 1 & 1 & -1 \\ 2 & 0 & 2 \\ -1 & 3 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 2 & -2 & 4 \\ -1 & 4 & -1 \end{vmatrix}$$

$$\left(\begin{array}{c} C_2 \longrightarrow C_2 - C_1 : \text{ la deuxiéme colonne est remplacée par} \\ \text{la même colonne moins la première;} \\ \text{et } \ C_3 \longrightarrow C_3 + C_1 \text{: la troisième colonne est remplacée par} \\ \text{la même colonne plus la première.} \end{array} \right)$$

$$= (1) \begin{vmatrix} -2 & 4 \\ 4 & -1 \end{vmatrix} = -14$$

3) $détA$ est nul si l'un des vecteurs colonnes est combinaison linéaire des autres. En particulier, $détA = 0$ si un vecteur colonne est nul; $détA = 0$ si deux colonnes sont égaux.

Exemple: $B = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 2 & 4 & 2 & 1 \\ -1 & 1 & 2 & 1 \\ 0 & -2 & -2 & 2 \end{pmatrix}$, on constate que $C_3 = C_2 - C_1$ i.e la troisième colonne est égale à la deuxième moins la première; le calcul donnera certainement $détB = 0$.

4) En multipliant par un scalaire λ toutes les composantes d'un vecteur colonne quelconque de A, on obtient un déterminant égal à $\lambda détA$. En d'autres termes

$$\begin{vmatrix} a_{11} & \dots & a_{1(j-1)} & \lambda a_{1j} & a_{1(j+1)} & \dots & a_{1n} \\ \vdots & & & & & & \\ \vdots & & & & & & \\ a_{n1} & \dots & a_{n(j-1)} & \lambda a_{nj} & a_{n(j+1)} & \dots & a_{nn} \end{vmatrix} =$$

$$\lambda \begin{vmatrix} a_{11} & \dots & a_{1(j-1)} & a_{1j} & a_{1(j+1)} & \dots & a_{1n} \\ \vdots & & & & & & \\ \vdots & & & & & & \\ a_{n1} & \dots & a_{n(j-1)} & a_{nj} & a_{n(j+1)} & \dots & a_{nn} \end{vmatrix}$$

Exemple: $A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 2 \\ -1 & 3 & 0 \end{pmatrix}$, on a déjà montré dans 1) que $détA = -14$; si on multiplie la deuxième colonne (par exemple) par -2 le calcul donne $\begin{vmatrix} 1 & -2 & -1 \\ 2 & 0 & 2 \\ -1 & -6 & 0 \end{vmatrix} = 28 = (-2)\, détA$

5) Si toutes les composantes du $j-iéme$ vecteur colonne de A sont de la forme $cij = a_{ij} + a'_{ij}$ avec $i = 1, \ldots, n$, alors $détA$ est la somme de deux déterminants dont les $j-iémes$ colonnes sont composées repectivement des éléments a_{ij} et a'_{ij}, et toutes les autres colonnes sont identiques aux colonnes correspondantes de $détA$. En d'autres termes, on a

$$\begin{vmatrix} a_{11} & \ldots & a_{1(j-1)} & a_{1j}+a'_{1j} & a_{1(j+1)} & \ldots & a_{1n} \\ \vdots & & & & & & \\ \vdots & & & & & & \\ a_{n1} & \ldots & a_{n(j-1)} & a_{nj}+a'_{nj} & a_{n(j+1)} & \ldots & a_{nn} \end{vmatrix} =$$

$$\begin{vmatrix} a_{11} & \ldots & a_{1(j-1)} & a_{1j} & a_{1(j+1)} & \ldots & a_{1n} \\ \vdots & & & & & & \\ \vdots & & & & & & \\ a_{n1} & \ldots & a_{n(j-1)} & a_{nj} & a_{n(j+1)} & \ldots & a_{nn} \end{vmatrix} +$$

$$\begin{vmatrix} a_{11} & \ldots & a_{1(j-1)} & a'_{1j} & a_{1(j+1)} & \ldots & a_{1n} \\ \vdots & & & & & & \\ \vdots & & & & & & \\ a_{n1} & \ldots & a_{n(j-1)} & a'_{nj} & a_{n(j+1)} & \ldots & a_{nn} \end{vmatrix}$$

Exemple:

$$\begin{vmatrix} 1 & -2 & -1 \\ 2 & 0 & 2 \\ -1 & -6 & 0 \end{vmatrix} = \begin{vmatrix} 1 & (-1)+(-1) & -1 \\ 2 & 0+0 & 2 \\ -1 & (-4)+(-2) & 0 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & -1 & -1 \\ 2 & 0 & 2 \\ -1 & -4 & 0 \end{vmatrix} + \begin{vmatrix} 1 & -1 & -1 \\ 2 & 0 & 2 \\ -1 & -2 & 0 \end{vmatrix} = 18+10 = 28$$

Remarque: Les règles de calcul précédentes sont vraies si on remplace colonne par ligne.

8.4 Application des déterminants

8.4.1 Rang d'une matrice

Les déterminants peuvent servir à calculer le rang d'une matrice de taille quelconque.

Définition: Soit $r \leq \min(m,\ n)$, on appelle mineur d'ordre r d'une matrice $m \times n$ le déterminant d'une matrice extraite de taille $r \times r$.

Exemples: Les mineurs d'ordre 1 sont simplement les coefficients de la matrice. Les mineurs d'ordre 3 de la matrice

$$A = \begin{pmatrix} -1 & 4 & 5 \\ 2 & 2 & 1 \\ 1 & -6 & 4 \\ 0 & 1 & 2 \end{pmatrix}$$

sont

$$\begin{vmatrix} -1 & 4 & 5 \\ 2 & 2 & 1 \\ 1 & -6 & 4 \end{vmatrix}, \begin{vmatrix} -1 & 4 & 5 \\ 1 & -6 & 4 \\ 0 & 1 & 2 \end{vmatrix}, \begin{vmatrix} 2 & 2 & 1 \\ 1 & -6 & 4 \\ 0 & 1 & 2 \end{vmatrix}.$$

Théorème: Soit A une matrice $m \times n$ alors $rgA = r$ si et seulement si:
(i) on peut extraire de A un mineur non nul d'ordre r,
(ii) tous les mineurs de A d'ordre $r + 1$ sont nuls.

Démonstration: Si un mineur d'ordre r est non nul, alors il y a r vecteurs colonnes indépendants donc $rgA \geq r$. On a $rgA = r$ dès que (ii) est vérifiée. Inversement si $rgA = r$ on a évidemment (i) et (ii). CQFD

On voit donc que le rang de A est l'ordre maximal des mineurs non nuls extraits de A.

Définition: Soient A une matrice et δ un mineur d'ordre r extrait de A. On appelle bordant de δ tout mineur d'ordre $r + 1$ extrait de A, dont δ est un déterminant extrait.

Exemple: Soit $A = \begin{pmatrix} \boxed{\begin{matrix} 0 & -1 \\ 1 & 2 \end{matrix}} & 4 & 1 \\ 3 & -5 & 2 & 6 \\ 6 & -2 & 1 & 2 \end{pmatrix}$, les bordants de $\delta = \begin{vmatrix} 0 & -1 \\ 1 & 2 \end{vmatrix}$ sont:

· en bordant avec la $3-iéme$ ligne

$$\begin{vmatrix} \boxed{\begin{matrix} 0 & -1 \\ 1 & 2 \end{matrix}} & 4 \\ 3 & -5 & 2 \end{vmatrix}, \begin{vmatrix} \boxed{\begin{matrix} 0 & -1 \\ 1 & 2 \end{matrix}} & 1 \\ 3 & -5 & 6 \end{vmatrix}$$

· en bordant avec la $4-iéme$ ligne

$$\begin{vmatrix} \boxed{\begin{matrix} 0 & -1 \\ 1 & 2 \end{matrix}} & 4 \\ 6 & -2 & 1 \end{vmatrix}, \begin{vmatrix} \boxed{\begin{matrix} 0 & -1 \\ 1 & 2 \end{matrix}} & 1 \\ 6 & -2 & 2 \end{vmatrix}$$

Montrer qu'on retrouve les mêmes bordants de δ en bordant avec les colonnes.

On peut reformuler le théorème précédent de la façon suivante:

Théorème: Soit A une matrice $m \times n$. Le rang de A est r si et seulement si on peut extraire de A un mineur δ d'ordre r non nul et tous les bordants de δ dans A sont nuls.

Exemples: Déterminons le rang de $A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & 3 \\ -1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

· $A \neq O \Longrightarrow rgA \geq 1$

· $\begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} = 3 \neq 0 \Longrightarrow rgA \geq 2$

$$\cdot \begin{vmatrix} 1 & -1 & 0 \\ 2 & 1 & 3 \\ -1 & 2 & 1 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & -1 & 0 \\ 2 & 1 & 3 \\ 0 & 1 & 1 \end{vmatrix} = 0.$$

Donc $rgA = 2$.

8.4.2 Caractérisation d'une famille libre

Théorème: Soient $\{v_1, \ldots, v_r\}$ r vecteurs d'un espace vectoriel E de dimension n $(r \leq n)$,

et $A = ||v_1, \ldots, v_r||$ la matrice dont les colonnes sont les composantes des vecteurs $v_1, \ldots, v_r$ dans une base quelconque de E.

La famille $\{v_1, \ldots, v_r\}$ est libre si et seulement si on peut extraire de A un mineur d'ordre r non nul.

Ce théorème résulte du théorème suivant:

Théorème: Soit E un espace vectoriel de dimension n.

La famille $\{v_1, \ldots, v_n\}$ de vecteurs de E forme une base de E si et seulement si $dét\,||v_1, \ldots, v_r||_{\{e_i\}} \neq 0$, où $||v_1, \ldots, v_r||$ désigne la matrice dont les colonnes sont les composantes des vecteurs $v_1, \ldots, v_r$ dans une base $\{e_i\}$ de E.

Exemple: Les vecteurs $v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 3 \\ 5 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 4 \\ 0 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ 5 \\ 9 \\ -2 \\ 0 \end{pmatrix}$ forment une famille libre de $\mathbb{R}^5$, car dans la matrice

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 5 \\ \boxed{\begin{matrix} 3 & 2 & 9 \\ 3 & 4 & -2 \\ 5 & 0 & 0 \end{matrix}} \end{pmatrix}$$

le mineur encadré est non nul.

8.4.3 Appartenance d'un vecteur à une famille

Théorème: Soient $v_1, \ldots, v_r$ r vecteurs linéairement indépendants,

$A = ||v_1, \ldots, v_r||$ la matrice dont les colonnes sont les composantes de ces vecteurs dans une base quelconque, et δ un mineur non nul d'ordre r extrait de A.

Pour qu'un vecteur w appartienne à $\mathcal{L}in\{v_1, \ldots, v_r\}$, il faut et il suffit que tous les bordants de δ dans la matrice $B = ||v_1, \ldots, v_r, w||$ soient nuls.

Démonstration**:** Si $w \in \mathcal{L}in\{v_1, \ldots, v_r\}$ alors tous les bordants de δ dans B sont nuls. En effet, si l'un des bordants était non nul, la famille $\{v_1, \ldots, v_r, w\}$ serait libre, ce qui est exclu.

Réciproquement, si tous les bordants de δ sont nuls, alors les r premiers vecteurs lignes de B sont indépendants, car $\delta \neq 0$ et chacun des autres est lié aux r premiers. Ainsi $rgB = r$, et comme $\{v_1, \ldots, v_r\}$ est libre $w \in \mathcal{L}in\{v_1, \ldots, v_r\}$. CQFD

Exemple: Pour quelles valeurs de α, $\beta \in \mathbb{R}$ le vecteur $w = \begin{pmatrix} \alpha \\ 2 \\ 1 \\ \beta \end{pmatrix}$

appartient-il à $\mathcal{L}in\{v_1, v_2\}$ avec $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ et $v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$?

On a: $A = \begin{pmatrix} \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} \\ \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \end{pmatrix}$, $\delta = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \neq 0$ et $B = \begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & \beta \end{pmatrix}$

Les bordants de δ sont $\triangle_1 = \begin{vmatrix} 1 & 0 & \alpha \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{vmatrix} = 1-\alpha$ et $\triangle_2 = \begin{vmatrix} 1 & 0 & \alpha \\ 0 & 1 & 2 \\ 0 & 1 & \beta \end{vmatrix} = \beta-2$.

Donc $w \in \{v_1, v_2\}$ si et seulement si $\alpha = 1$ et $\beta = 2$.

8.4.4 Détermination de l'inverse d'une matrice inversible

Définition: On appelle transposée d'une matrice $A = (a_{ij}) \in \mathcal{M}_n(K)$ la matrice de $\mathcal{M}_n(K)$ notée t_A définie par:

$$t_A = (a_{ji}).$$

On a les propriétés suivantes:

Propriétés:

$t_{(A+B)} = t_A + t_B$; $t_{(\lambda A)} = \lambda t_A$ $(\lambda \in K)$; $t_{(AB)} = t_B t_A$.

Définition et théorème: Soit $A \in \mathcal{M}_n(K)$. On appelle matrice des cofacteurs de A et on note $cof(A)$, la matrice obtenue de A en remplaçant chaque élément par son cofacteur. Si A est inversible alors:

$$A^{-1} = \frac{1}{détA} t_{cof(A)}$$

Exemples:

· Soit $A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$; on a $détA = 5 \neq 0$, donc A est inversible.

$$cof(A) = \begin{pmatrix} cof(a_{11}) & cof(a_{12}) \\ cof(a_{21}) & cof(a_{22}) \end{pmatrix}$$

avec $cof(a_{11}) = 3$, $cof(a_{12}) = 1$, $cof(a_{21}) = -2$, $cof(a_{22}) = 1$.

Donc $cof(A) = \begin{pmatrix} 3 & 1 \\ -2 & 1 \end{pmatrix}$ et par suite

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

Plus généralement:
si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec $ad - bc \neq 0$, on a:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

· Soit $B = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 3 & 0 \\ 0 & 1 & -1 \end{pmatrix}$, $détB = -5 \neq 0$, donc B est inversible.
$cof(b_{11}) = -3$, $cof(b_{12}) = -1$, $cof(b_{13}) = -1$,
$cof(b_{21}) = 2$, $cof(b_{22}) = -1$, $cof(b_{23}) = -1$,
$cof(b_{31}) = 0$, $cof(b_{32}) = 0$, $cof(b_{33}) = 5$.

$$Cof(B) = \begin{pmatrix} -3 & -1 & -1 \\ 2 & -1 & -1 \\ 0 & 0 & 5 \end{pmatrix} \quad \text{et} \quad B^{-1} = -\frac{1}{5} \begin{pmatrix} -3 & 2 & 0 \\ -1 & -1 & 0 \\ -1 & -1 & 5 \end{pmatrix}.$$

Chapter 9

Systèmes d'équations linéaires

Considérons le système d'équations linéaire suivant:

$$\mathcal{S}: \begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

On sait que le système $\mathcal{S}$ peut s'écrire sous la forme matricielle

$$AX = B$$

avec $A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$.

On appelle rang du système $\mathcal{S}$ le rang de la matrice A associée à $\mathcal{S}$.

9.1 Systèmes de Cramer

Définition: On appelle système de Cramer un système linéaire dont la matrice associée est carrée et inversible.

Il s'agit donc d'un système de n équations en n inconnues de rang n:

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \qquad\qquad \vdots \qquad\qquad\qquad \text{avec } dét A \neq 0. \\ a_{n1}x_1 + \ldots + a_{nn}x_n = b_n \end{cases}$$

Sous forme matricielle on a:

$$AX = B \Longleftrightarrow A^{-1}(AX) = A^{-1}B \Longleftrightarrow (A^{-1}A)X = A^{-1}B$$

d'où

$$X = A^{-1}B$$

Ainsi, un système de Cramer admet une solution unique; ce qui donne une méthode de résolution du système.

9.1.1 Résolution par une matrice inverse

La formule

$$X = A^{-1}B$$

permet de résoudre le système $AX = B$.

Exemple: Résoudre le système $\begin{cases} x + 2y = 1 \\ -x + 3y = 0 \end{cases}$

Sous forme matricielle le système s'écrit $AX = B$, avec

$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \end{pmatrix}$, $B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$;

d'où

$X = A^{-1}B$ i.e $\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

La solution du système est donc $\left(\frac{3}{5}, \frac{1}{5}\right)$.

9.1.2 Résolution par les formules de Cramer

En pratique, au lieu de calculer A^{-1} on se sert des formules de Cramer.

Si D désigne le déterminant de la matrice associée au système, on notera Dx_i le déterminant obtenu de D en remplaçant la colonne des coefficients de x_i par les b_i respectifs. Ainsi la solution $(x_1, \ldots, x_n)$ du système est donnée par les formules suivantes dites formules de Cramer:

$x_i = \dfrac{Dx_i}{D}$ avec $i = 1, \ldots, n$.

Exemple: On considère le système:

$$\begin{cases} 2x - 5y + 2z = 7 \\ x + 2y - 4z = 3 \\ 3x - 4y - 6z = 5 \end{cases}$$

On a $D = \begin{vmatrix} 2 & -5 & 2 \\ 1 & 2 & -4 \\ 3 & -4 & -6 \end{vmatrix} = -46$

$$x = \frac{Dx}{D} = \frac{\begin{vmatrix} 7 & -5 & 2 \\ 3 & 2 & -4 \\ 5 & -4 & -6 \end{vmatrix}}{-46} = 5,\ y = \frac{Dy}{D} = \frac{\begin{vmatrix} 2 & 7 & 2 \\ 1 & 3 & -4 \\ 3 & 5 & -6 \end{vmatrix}}{-46} = 1,$$

$$z = \frac{Dz}{D} = \frac{\begin{vmatrix} 2 & -5 & 7 \\ 1 & 2 & 3 \\ 3 & -4 & 5 \end{vmatrix}}{-46} = 1.$$

La solution du système est donc (5, 1, 1).

9.2 Cas général

Considérons le système de m équations en n inconnes de rang r:

$$\mathcal{S}:\begin{cases} \boxed{a_{11}x_1+\ldots+a_{1r}x_r} & +\ldots+a_{1n}x_n=b_1 \\ \vdots & \\ \boxed{a_{r1}x_1+\ldots+a_{rr}x_r} & +\ldots+a_{rn}x_n=b_r \\ \vdots & \\ a_{m1}x_1+\ldots+a_{mr}x_r & +\ldots+a_{mn}x_n=b_m \end{cases}$$

On peut supposer, quitte à changer l'ordre des équations et la numérotation des inconnues, que le mineur δ encadré est non nul.

Définition: Posons $\delta=\begin{vmatrix} a_{11}\ldots a_{1r} \\ \vdots \\ a_{r1}\ldots a_{rr} \end{vmatrix}$, alors les déterminants

$$\triangle_s=\begin{vmatrix} \begin{vmatrix} a_{11}\ldots a_{1r} \\ \vdots \\ a_{r1}\ldots a_{rr} \end{vmatrix} & \begin{matrix} b_1 \\ \vdots \\ b_r \end{matrix} \\ a_{s1}\ldots a_{sr} & b_s \end{vmatrix} \quad s=r+1,\ldots,m,$$

sont appelés déterminants caractéristiques associés à δ.

Il résulte des théorèmes précédents le théorème suivant:

Théorème de Rouché-Fontené: Soit

$$\mathcal{S}:\begin{cases} \boxed{a_{11}x_1+\ldots+a_{1r}x_r} & +\ldots+a_{1n}x_n=b_1 \\ \vdots & \\ \boxed{a_{r1}x_1+\ldots+a_{rr}x_r} & +\ldots+a_{rn}x_n=b_r \\ \vdots & \\ a_{m1}x_1+\ldots+a_{mr}x_r & +\ldots+a_{mn}x_n=b_m \end{cases}$$

un système de m équations en n inconnes de rang r. On extrait de $\mathcal{S}$ un mineur δ d'ordre non nul (quitte à changer la numérotation, on peut supposer que δ est le mineur encadré):

$$\delta = \begin{vmatrix} a_{11} \dots a_{1r} \\ \vdots \\ a_{r1} \dots a_{rr} \end{vmatrix} \neq 0$$

1) Le système $\mathcal{S}$ est compatible si et seulement si tous les déterminants caractéristiques associés à δ sont nuls:

$$\triangle_s = \begin{vmatrix} \boxed{\begin{matrix} a_{11} \dots a_{1r} \\ \vdots \\ a_{r1} \dots a_{rr} \end{matrix}} & \begin{matrix} b_1 \\ \vdots \\ b_r \end{matrix} \\ \begin{matrix} a_{s1} \dots a_{sr} \end{matrix} & \begin{matrix} b_s \end{matrix} \end{vmatrix} = 0, \quad (s = r+1, \dots, m).$$

2) Si cette condition est réalisée, $\mathcal{S}$ est équivalent au système des " équations principales ":

$$\mathcal{S}' : \begin{cases} \boxed{\begin{matrix} a_{11}x_1 + \dots + a_{1r}x_r \\ \vdots \\ a_{r1}x_1 + \dots + a_{rr}x_r \end{matrix}} & + \dots + a_{1n}x_n = b_1 \\ & \\ & + \dots + a_{rn}x_n = b_r \end{cases}$$

Il admet alors une infinité de solutions dépendantes de $n - r$ paramètres. Les solutions se calculent en résolvant le système de Cramer obtenu en donnant aux " variables libres " $x_{r+1}, \dots, x_n$ des valeurs arbitraires:

$$x_{r+1} = \lambda_{r+1}, \dots, x_n = \lambda_n.$$

Le système $\mathcal{S}'$ s'écrit alors:

$$\begin{cases} \boxed{\begin{matrix} a_{11}x_1 + \dots + a_{1r}x_r \\ \vdots \\ a_{r1}x_1 + \dots + a_{rr}x_r \end{matrix}} & = b_1 - a_{1,r+1}\lambda_{r+1} - \dots - a_{1n}\lambda_n \\ & \\ & = b_r - a_{r,r+1}\lambda_{r+1} - \dots - a_{rn}\lambda_n \end{cases}$$

Exemple: Discuter, d'après les valeurs de k, α, β, $\gamma \in \mathbb{R}$ le système:

$$\begin{cases} 2x & + & y & - & z & = & \alpha \\ & & y & + & 3z & = & \beta \\ 2x & + & ky & + & 2z & = & \gamma \end{cases}$$

On étudie d'abord le rang du système. La matrice associée au système est:

$$A = \begin{pmatrix} \boxed{\begin{matrix} 2 & 1 \\ 0 & 1 \end{matrix}} & -1 \\ 2 \quad k & 2 \end{pmatrix}$$

et $dét A = 6\,(2-k)$.

Si $k \neq 2$, $rg(A) = 3$; si $k = 2$, $rg(A) = 2$ car $\begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} = 2 \neq 0$.

Ainsi deux cas se présentent:

a) $k \neq 2$, le système est alors un système de Cramer.

La solution s'obtient par les formules de Cramer:

$$x = \frac{Dx}{D},\ y = \frac{Dy}{D} \text{ et } z = \frac{Dz}{D} \text{ avec } D = dét A.$$

b) $k = 2$, le système est de rang 2. On étudie alors la compatibilité à l'aide des déterminants caractéristiques. On extrait un mineur d'ordre 2 non nul, par exemple:

$$\delta = \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix}$$

Le seul déterminant caractéristique associé à δ est

$$\triangle_3 = \begin{vmatrix} 2 & 1 & \alpha \\ 0 & 1 & \beta \\ 2 & 2 & \gamma \end{vmatrix} = 2\,(\gamma - \alpha - \beta).$$

Donc le système est compatible si $\gamma = \alpha + \beta$. Les solutions s'obtiennent alors en résolvant le système des équations principales:

$$\begin{cases} \boxed{\begin{array}{r} 2x + y \\ y \end{array}} & -z = \alpha \\ & +3z = \beta \end{cases}$$

La variable libre est z. En posant $z = \lambda$, on a:

$\begin{cases} 2x + y = \alpha + \lambda \\ y = \beta - 3\lambda \end{cases}$, d'où $x = \dfrac{\alpha - \beta + 4\lambda}{2}$, $y = \beta - 3\lambda$, $z = \lambda$.

En résumé:

(i) si $k \neq 2$, le système admet une et une seule solution;

(ii) si $k = 2$, deux cas se présentent:

· $\gamma \neq \alpha + \beta$: le système n'admet pas de solution;

· $\gamma = \alpha + \beta$: le système admet une infinité de solutions dépendants d'un paramètre.

9.3 Cas des systèmes homogènes

Définition: On appelle système homogène, un système de type

$$AX = 0$$

c'est-à-dire un système linéaire dont le deuxième membre est nul.

On voit immédiatement qu'un système homogène est toujours compatible: il admet au moins la solution nulle.

On déduit du paragraphe précédent la proposition suivante:

Proposition: Soit

$$\begin{cases} \boxed{\begin{array}{l} a_{11}x_1 + \ldots + a_{1r}x_r \\ \vdots \\ a_{r1}x_1 + \ldots + a_{rr}x_r \end{array}} & + \ldots + a_{1n}x_n = 0 \\ \vdots & \\ a_{m1}x_1 + \ldots + a_{mr}x_r + \ldots + a_{mn}x_n = 0 \end{cases}$$

un système linéaire homogène de rang r. On suppose que le mineur encadré est non nul.

Puisque les conditions de compatibilité sont trivialement satisfaites (cf théorème de Rouché-Fontené), le système est équivalent au système des équations principales:

$$\begin{cases} \boxed{\begin{array}{l} a_{11}x_1 + \ldots + a_{1r}x_r \\ \vdots \\ a_{r1}x_1 + \ldots + a_{rr}x_r \end{array}} & + \ldots + a_{1n}x_n = 0 \\ & \\ & + \ldots + a_{rn}x_n = 0 \end{cases}$$

Les solutions forment un espace vectoriel de dimension $n - r$.

Remarques:

· Les solutions d'un système linéaire non homogène $AX = B$ avec $B \neq 0$ ne forment pas un espace vectoriel. En effet, si X_1 et X_2 sont deux solutions, on a $AX_1 = B$ et $AX_2 = B$, d'où $A(X_1 + X_2) = 2B \neq B$: donc la somme de deux solutions n'est pas une solution.

· Un système homogène

$$\begin{cases} \boxed{\begin{array}{l} a_{11}x_1 + \ldots + a_{1r}x_r \\ \vdots \\ a_{r1}x_1 + \ldots + a_{rr}x_r \end{array}} & + \ldots + a_{1n}x_n = 0 \\ & \\ & + \ldots + a_{rn}x_n = 0 \\ \vdots & \\ a_{m1}x_1 + \ldots + a_{mr}x_r + \ldots + a_{mn}x_n = 0 \end{cases}$$

de rang r admet une solution non nulle si et seulement si $n > r$. C'est le cas, par exemple, s'il y a plus d'inconnues que d'équations.

· Un cas particulièrement important est celui où la matrice A associée au système linéaire homogène est carrée. Dans ce cas, la condition pour qu'il y ait des solutions non nulles $(n > r)$ est équivalente à $détA = 0$.

Ainsi, on a la proposition suivante:

Proposition: Soit $AX = 0$ un système linéaire homogène où A est une matrice carrée. Alors le système admet des solutions non nulles si et seulement si $détA = 0$.