MATHEMATIQUES POUR L'INGENIEUR

Support du cours donné en 1^{re} année

par Marietta Manolessou

EISTI - Département Mathématiques

Année 2007-2008

Table des matières

Table des figures

Chapitre 1

Espaces vectoriels - Matrices - Endomorphismes - Réduction (A)

Rappels, remarques références sur les :
- **1. Espaces Vectoriels**
- **2. Matrices**
- **3. Applications linéaires**
- **4. Matrices et Applications linéaires**

1 Espaces Vectoriels

1 Espaces Vectoriels

Définition 1.1 (Structure d'un espace vectoriel).
Supposons que
i) $E = \{u_1, u_2, \cdots, u_n, \cdots\}$ *est un ensemble d'éléments (fini ou infini).*
ii) On a défini une opération interne "+" ("addition") entre les éléments de E*, telle que* E *en forme un groupe abélien.*
iii) $\mathbb{K}$ *est un corps commutatif.*
iv) On a défini une opération externe "•" ("multiplication") sur les éléments de E*, ayant comme opérateurs ("scalaires") des éléments de* $\mathbb{K}$*. Cette opération est telle que pour tout* $u \in E$ *et tout* λ *et* $\mu \in \mathbb{K}$*,on ait :*
$\lambda \cdot (\mu \cdot u) = (\lambda \cdot \mu) \cdot u$ *associativité*
$e \cdot u = u$ *(*e *élément "neutre" de la multiplication dans* $\mathbb{K}$*)*
v) L'opération externe est distributive par rapport à l'addition dans $\mathbb{K}$ *et par rapport à "l'addition" dans* E*. autrement dit :*
$(\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u$
$\lambda \cdot (u_1 + u_2) = \lambda \cdot u_1 + \lambda \cdot u_2$

Si toutes ces conditions sont vérifiées E *est un espace vectoriel sur* $\mathbb{K}$*.*

Remarque 1.1. Dans nos applications on aura souvent :

$$\mathbb{K} = \mathbb{R} \quad \textit{ou} \quad \mathbb{K} = \mathbb{C}.$$

Exemples 1.1. *a)* L'ensemble des vecteurs du plan $(\mathbb{R}^2)$ notés $\vec{x}$ ou $\vec{y}$ muni des opérations $\vec{x}+\vec{y}$ et $\alpha\vec{x}$ avec $\alpha \in \mathbb{R}$ est un espace vectoriel sur $\mathbb{R}$.

b) L'ensemble des fonctions numériques continues sur un intervalle $I = [a,b] \subset \mathbb{R}$, muni des opérations $f+g,\ \alpha f \quad \alpha \in \mathbb{R}$, est un espace vectoriel *sur* $\mathbb{R}$

c) $\mathbb{R}$ (resp. $\mathbb{C}$) muni des opérations habituelles est un espace vectoriel sur lui-même.

2 Sous espaces vectoriels

Définition 1.2 (Sous espaces vectoriels).

Soit E espace vectoriel sur $\mathbb{K}$.

Une partie non vide F de E est un sous-espace vectoriel de E ssi elle a la même structure d'espace vectoriel que E c'est-à-dire :

$$\left.\begin{array}{ll}(\forall\, x \in F) & (\forall\, y \in F)\\ (\forall\, \lambda \in \mathbb{K}) & (\forall\, \mu \in \mathbb{K})\end{array}\right\} \lambda \cdot x + \mu \cdot y \in F$$

Remarque 1.2.

a) Les opérations interne "+" et externe "•" sont celles induites par la structure de E, donc F est un sous-groupe du groupe additif de E.

b) Pour simplifier les notations on écrit souvent par la suite $\left.\begin{array}{l}(\forall\, u \in E)\\ (\forall\, \lambda \in \mathbb{K})\end{array}\right\} \lambda\, u$ *au lieu de $\lambda \cdot u$.*

Exemple 1.1. L'ensemble de toutes les combinaisons linéaires : $\nu = \sum_{i=1}^{p} \lambda_i u_i$ de p éléments de E espace vectoriel sur $\mathbb{K}$, forme un sous-espace vectoriel de E.

3 Isomorphismes

Définition 1.3 (Isomorphisme d'espaces vectoriels).

Soient E, E' deux espaces vectoriels sur le même corps $\mathbb{K}$.

*On appelle **isomorphisme** de E sur E' toute bijection f de E sur E' telle que :*

$$\begin{array}{ll}(\forall\, u \in E)(\forall\, \nu \in E) f(u+\nu) & = f(u) + f(\nu)\\ (\forall\, u \in E)(\forall\, \nu \in E) f(\lambda u) & = \lambda\, f(u)\end{array}$$

* *E et E' sont alors des **espaces vectoriels isomorphes**.*

Remarque 1.3. Si $E = E'$, on dit que f est un *automorphisme*.

4 Espaces produits

Définition 1.4 (Espace vectoriel produit).

Soient n espaces vectoriels sur $\mathbb{K}$, $E_1,\ E_2, \cdots,\ E_n$. On définit l'espace vectoriel produit sur $\mathbb{K}$ et on note $E = E_1 \times E_2 \times \cdots \times E_n$ par les égalités suivantes :

Si, $x_i \in E_i,\ y_i \in E_i,\ \forall i$ alors $(x_1, \cdots, x_n) \in E,\ (y_1, \cdots, y_n) \in E$, et

$$(x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n) = (x_1+y_1, x_2+y_2, \cdots, x_n+y_n)$$
$$\lambda(x_1, x_2, \cdots, x_n) = (\lambda x_1, \lambda x_2, \cdots, \lambda x_n)$$

Remarque 1.4. *Si on prend $E_1 = E_2 = \cdots E_n = E$ alors on note par E^n l'espace vectoriel produit de E par lui-même (n fois)*

Exemples 1.2.

a) $\mathbb{R}^n$ *espace vectoriel produit sur* $\mathbb{R}$ *(de* $\mathbb{R}$ *par lui même* $2n$ *fois)*

b) $\mathbb{C}^n$ *est un espace vectoriel produit sur* $\mathbb{R}$ *isomorphe à* $\mathbb{R}^{2n}$*, (espace vectoriel produit de* $\mathbb{R}$ *par lui même* $2n$ *fois).*
$\mathbb{C}^n$ *est un espace vectoriel produit sur* $\mathbb{C}$ *(de* $\mathbb{C}$ *par lui même* n *fois).*

5 Bases

Définition 1.5 (Base d'un espace vectoriel).

Soit E *espace vectoriel sur* $\mathbb{K}$.

a) *Famille génératrice* G *: c'est une partie* $G = \{x_1, \cdots, x_p\} \subset E$ *telle que tout élément* $x \in E$ *peut étre engendré comme combinaison linéaire des éléments de* G.

b) *Famille libre* L *: c'est une partie finie d'éléments de* $E, L = \{y_1, y_2, \cdots, y_n\}$ *telle que toute équation du type :*
$\sum_{i=1}^{n} \lambda_i y_i = 0$
implique

$$\lambda_1 = \lambda_2 = \lambda_3 = \cdots = \lambda_n = 0$$

. On dit aussi que les éléments $y_1, \cdots, y_n$ *sont linéairement indépendants.*

c) *Base d'un espace vectoriel* E *sur* K *: c'est une famille de* E *libre et génératrice.*

Remarque 1.5.

a) *Lorsque* E *est engendré par un nombre fini de ses éléments (resp. nombre infini d'éléments on dit que c'est un espace vectoriel de dimension finie (resp.dimension infinie).*

b) *Une famille qui n'est pas libre est dite liée.*

c) *Définitions équivalentes pour une base* B *de* E.
B base $\Leftrightarrow$ B partie ***génératrice minimale*** de E.
B base $\Leftrightarrow$ B partie ***libre maximale*** de E.

d) *On appelle* ***base canonique*** *de* $\mathbb{R}^n$ *l'ensemble suivant :*

$$\varepsilon(n) = \{e_i\}_{i=1,\cdots,n} \text{ avec } e_i = (\delta_i^1, \cdots, \delta_i^i, \cdots, \delta_i^n) \text{ et } \delta_i^j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{ailleurs} \end{cases}$$

On vérifie facilement que $\varepsilon(n)$ *forme une base pour* $\mathbb{R}^n$.

Exemple 1.2.

Base canonique de $\mathbb{R}^3$ *:* $\vec{e}_1 = (1, 0, 0)$; $\vec{e}_2 = (0, 1, 0)$; $\vec{e}_3 = (0, 0, 1)$

Théorème 1.1.

a) *Tout espace vectoriel de dimension finie admet une base.*

b) *Toutes les bases d'un espace vectoriel, de dimension finie* E*, ont même nombre d'éléments ; ce nombre est appelé* ***la dimension de l'espace vectoriel*** E *(notation :* $\dim_{\mathbb{K}} E$*).*

Théorème 1.2. ***(Th. de la base incomplète)***

L *et* G *étant respectivement une partie libre et une partie génératrice d'un espace vectoriel* E *sur* $\mathbb{K}$*, il existe une partie* H *de* G *t.q.* $L \cup H$ *soit une base de* E.

Théorème 1.3.

a) *Tout espace vectoriel de dimension finie sur* $\mathbb{K}$ *:* $\dim_{\mathbb{K}} E = n$ *est isomorphe à* $\mathbb{K}^n$.

b) Si $E = E_1 \times E_2 \times \cdots \times E_m$ et $\forall i = 1, \cdots m \quad \dim_{\mathbb{K}} E_i = n_i$ (finie) $\forall i = 1, \cdots m$ alors
$\dim_{\mathbb{K}} E = \underline{\sum_{i=1}^{m} n_i}$.

Exemple 1.3.
$\dim_{\mathbb{R}} \mathbb{C} = 2$; $\dim_{\mathbb{R}} \mathbb{R} = 1$ $\dim_{\mathbb{R}} \mathbb{R}^n = n$

6 Somme directe

Définition 1.6 (Somme directe - Sous-espaces supplémentaires).
Soit $E = E_1 + E_2$ (ou $E = E_1 \cup E_2$) la somme de deux espaces vectoriels sur $\mathbb{K}$, les propriétés suivantes sont équivalentes

*a) $E = E_1 \oplus E_2$ (c'est-à-dire E espace vectoriel sur $\mathbb{K}$ est la **somme directe** de E_1 et E_2)*
b) $E_1 \cap E_2 = \{0\}$
*c) $\forall \ u \in E$ la décomposition $u = u_1 + u_2$ avec $u_1 \in E_1$, $u_2 \in E_2$ est **unique**.*

* *On appelle alors E_1 et E_2 sous-espaces vectoriels **supplémentaires** par rapport à E.*

Théorème 1.4. *L'espace produit $E_1 \times E_2$ est isomorphe à l'espace $E_1 \oplus E_2$*

Théorème 1.5.
a) Tout sous-espace vectoriel F de E (esp. vect. sur $\mathbb{K}$) admet au moins un supplémentaire G par rapport à E et

$$E = F \oplus G \Rightarrow \underline{\dim_K E = \dim F + \dim G}$$

b) Tous les supplémentaires de F par rapport à E ont même dimension appelée la <u>codimension de F</u> par rapport à E.

Exemples 1.3.
a) $\mathbb{R}^3 = \mathbb{R}_x \oplus \mathbb{R}_y \oplus \mathbb{R}_z$; $\dim_{\mathbb{R}} \mathbb{R}^3 = 3$; $\dim_{\mathbb{R}} \mathbb{R}_x = \dim_{\mathbb{R}} \mathbb{R}_y = \dim_{\mathbb{R}} \mathbb{R}_z = 1$
b) $\mathbb{C} = \mathbb{R}_x \oplus j\mathbb{R}_y$ $(j^2 = -1)$; $\dim_{\mathbb{R}} \mathbb{C} = 2$

7 Rang d'un système de vecteurs

Définition 1.7 (Rang d'un système de vecteurs d'un esp. vectoriel).
Soit un système $S = \{u_1, \cdots, u_s\}$ de vecteurs. La <u>dimension</u> (finie) du sous-espace F de E engendré par ce système de vecteurs est appelée le <u>rang de S</u> et on note $\underline{rg(S)}$.

Exemple 1.4. *Dans R^n on considère les vecteurs e_i $i = 1, 2, 3$ de la base canonique de R^n et soit $\lambda \in \mathbb{R}$ et $S = \{\lambda e_i\}_{i=1,2,3} \Rightarrow rg(S) = 3$*

2 Matrices

1 Généralités

Remarque 1.6.
Dans ce paragraphe on supposera connues certaines définitions fondamentales comme celles sur : les matrices sur un corps $\mathbb{K}$, les matrices carrées, rectangulaires, triangulaires, diagonales ; l'égalité entre matrices ; les opérations sur les matrices (multiplication par un scalaire $\alpha \in \mathbb{K}$, l'addition et multiplication des matrices) ; l'inversion d'une matrice carrée.

On supposera aussi connus :

a). La théorie des systèmes d'équations linéaires et les théorèmes associés à leurs solutions.

b). Les propriétés et les opérations sur les matrices en blocs.

Définition 1.8 (Transposition des matrices).

a) Soit $A = \{a_{ij}\}_{j=1,\cdots,m}^{i=1,\cdots,n}$ une matrice sur un corps $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) (de n lignes et m colonnes).
On appelle transposée de A et on note A^t la matrice :
$A^t = \{a_{ji}\}_{i=1,\cdots,n}^{j=1,\cdots,m}$
(les m lignes de A^t sont les m colonnes de A, les n colonnes de A^t sont les n lignes de de A)

b) Une matrice A est dite symétrique si : $A^t = A$
et antisymétrique si : $A^t = -A$

Exemples 1.4.

$$\begin{pmatrix} 1 & 3 & 5 \\ 0 & 2 & 4 \end{pmatrix}^t = \begin{pmatrix} 1 & 0 \\ 3 & 2 \\ 5 & 4 \end{pmatrix}$$

$$\textit{matrice symétrique : } A = \begin{pmatrix} 1 & 3 & 5 \\ 3 & 2 & 4 \\ 5 & 4 & 7 \end{pmatrix}$$

$$\textit{matrice antisymétrique : } B = \begin{pmatrix} 0 & -0,87 & 23 \\ 0,87 & 0 & -4 \\ -23 & 4 & 0 \end{pmatrix}$$

Théorème 1.6.

i) $(A + B)^t = A^t + B^t$
ii) $(A^t)^t = A$
iii) $(kA)^t = kA^t \quad \forall k \in \mathbb{K}$
iv) $(AB)^t = B^t A^t$

2 Rang d'une matrice

Définition 1.9 (Rang d'une matrice).

*a) On appelle **rang d'une matrice** $A = \{a_{ij}\}_{j=1,\cdots,m}^{i=1,\cdots,n}$ sur $\mathbb{K}$ le rang du système de ses vecteurs colonnes dans $\mathbb{R}^n$ (ou le rang du système de ses vecteurs lignes) dans $\mathbb{R}^n$).*
Notation : $rg(A)$.
On appelle aussi espace-colonnes (ou espace lignes) le sous-espace de $\mathbb{R}^n$ engendré par les colonnes (ou par les lignes) de A, linéairement indépendantes-famille libre- dans $\mathbb{R}^n$ (ou dans $\mathbb{R}^m$).

b) Deux matrices A_1, A_2 sont dites équivalentes lignes (ou équivalentes colonnes) si elles ont le même espace lignes (ou le même espace colonnes).

Théorème 1.7.

Deux matrices équivalentes lignes ont le même rang.

Définition 1.10 (Matrices échelonnées). *Une matrice $A = \{a_{ij}\}_{j=1,\cdots,m}^{i=1,\cdots,n}$ est dite* *échelonnée (ou mise **sous forme échelonnée**) si il existe des éléments :*

$$\begin{cases} a_{1j_1}, a_{2j_2}, \cdots, a_{rj_r} & \text{où } j_1 < j_2 < \cdots j_r \\ \text{avec } a_{ij} = 0 & \forall i \leq r,\ j < j_i, \text{ et pour } i > r \end{cases}$$

Autrement dit :

A est échelonnée, si le nombre de zéros qui précède le premier élément non nul d'une ligne augmente de ligne en ligne jusqu'à ce qu'il n'y ait plus de ligne. On appelle les premiers éléments non nuls des lignes d'une matrice échelonnée les éléments distingués ou remarquables.

Exemple 1.5.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 6 \\ 0 & 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & \mathbf{1} & 0 & 4 & -6 & 7 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 20 & 8 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{pmatrix}$$

Quand les éléments distingués des matrices échelonnées sont tous égaux à 1 et, qu'en plus, il sont les seuls éléments non nuls de leurs colonnes respectives (v. exemple B) on les appelle **matrices échelonnées réduites par les lignes (**$e.r.\ell.$**)**

Définition 1.11.

Algorithme d'échelonnage sur les lignes.

- 1^{re} étape
  Appelons j_1 la première colonne avec un élément non nul. Pour simplifier, supposons que cet élément se trouve à la première ligne L_1 (sinon on permute simplement les lignes) d'où $a_{1j_1} \neq 0$.
- 2^{me} étape
  Pour chaque ligne L_i avec $i > 1$ on applique la combinaison :

  $$L_i \to -a_{ij_1} L_1 + a_{1j_1} L_i \qquad (E)$$

  On répète la 1ère et la 2ème étape avec la sous-matrice formée par toutes les lignes sauf la première et on continue cette méthode jusqu'à ce que la matrice soit sous forme échelonnée.

Exemple 1.6.

$$A_0 = \begin{pmatrix} 1 & 2 & -3 & 0 \\ 2 & 4 & -2 & 2 \\ 3 & 6 & -4 & 3 \end{pmatrix} \Rightarrow A' \text{ avec } \begin{array}{rcl} L'_1 & = & L_1 \\ L'_2 & = & -2L_1 + L_2 \\ L'_3 & = & -3L_1 + L_3 \end{array} \Rightarrow$$

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

$$A' \Rightarrow A'' \text{ avec } \begin{array}{rcl} L''_1 & = & L'_1 = L_1 \\ L''_2 & = & L'_2 \\ L''_3 & = & -5L_2 + 4L_3 \end{array} \Rightarrow A'' = \begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

A" est déjà échelonnée. Pour qu'elle soit e.r. ℓ on divise chaque ligne par son élément distingué donc :

$$A'' \to A''' = \begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 1 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(et ensuite on recommence une procédure analogue pour annuler -3 et 1/2 sur les colonnes).

Remarque 1.7. La forme échelonnée des matrices est particulièrement adéquate pour trouver directement leur rang, car on a :

Théorème 1.8.

Deux matrices transformées l'une de l'autre par une transformation du type (E) de l'algorithme de réduction sous forme échelonnée, ont le même espace ligne donc le même rang.

Exemple 1.7. Considérons les matrices A, B, A_0, A', A'', A''' ci-dessus. On a :

$$rg(A) = 2, \quad rg(B) = 4, \quad rgA_0 = rgA' = rgA'' = rgA''' = 3$$

Remarque 1.8 (Quelques cas particuliers de matrices échelonnées).

i) On a les matrices (carrées) **triangulaires supérieures** (resp. triang. inférieures).

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & 0 & \ddots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \qquad (\textit{resp.} \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{12} & a_{22} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \vdots & \cdots \\ a_{1n} & \cdots & \cdots & \cdots & a_{nn} \end{pmatrix})$$

* *L'algorithme de **"Pivot de Gauss"** équivaut à l'algorithme ci-dessus de réduction à la forme échelonnée.*

ii) On a les matrices **diagonales**.

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & & 0 \\ \vdots & & \ddots & \cdots \\ 0 & & & a_n \end{pmatrix}$$

et plus particulièrement la **matrice identité**

$$I_n = \begin{pmatrix} 1 & & & \\ & 1 & & 0 \\ 0 & & \ddots & \\ & & & 1 \end{pmatrix} \quad \textit{(n lignes et n colonnes)}$$

et les matrices scalaires $S = \lambda I_n$ *avec* $\lambda \in K$

* *On sait que pour toute matrice carrée* $A = \{a_{ij}\}_{\substack{i=1,\cdots,n \\ j=1,\cdots,m}}$ *on a :*

$$AI_n = I_n A = A$$

3 Utilisation des matrices échelonnées pour la résolution des systèmes d'équations linéaires.

Remarque 1.9. Tout système d'équations linéaires de la forme :

$$\left.\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \cdots\cdots\cdots & \cdots & \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{nn}x_n & = & b_m \end{array}\right\} \qquad (S_1)$$

est équivalent à une équation matricielle : $\underline{AX = B}$ *(M_1)*

$$\textit{où } A = \begin{pmatrix} a_{11} & a_0 & \cdots & a_{1n} \\ a_{21} & & & \vdots \\ \vdots & & & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{pmatrix} \;;\; X = \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} \;;\; B = \begin{pmatrix} b_1 \\ \vdots \\ \vdots \\ b_n \end{pmatrix} \;;$$

On considère la matrice "augmentée" (ou complète)

$$(A,B) = A = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_n \\ a_{21} & & & \vdots \\ \vdots & & & \vdots \\ a_{m1} & & a_{mn} & b_m \end{pmatrix}$$

alors :

Théorème 1.9.

a) Deux systèmes (S_1) et (S_2) sont équivalents ssi leurs matrices augmentées ont le même rang.

b) On peut toujours remplacer (M_1) par une équ. (M_1') où la matrice complète est sous forme échelonnée (Equivalence lignes)

c) (M_1) a une solution ssi A et (A,B) ont même rang.

4 Polynômes de matrices

Définition 1.12 (Polynômes des matrices carrées).

On peut toujours définir les puissances A^n d'une matrice carrée.
On associe à tout polynôme : $P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ avec $a_i \in K, \ \forall i$ la matrice $P[A]$:

$$P[A] = a_0I_n + a_1A + a_2A_2 + \cdots + a_nA^n \qquad (P_1)$$

*Si $P[A] = 0$ (matrice nulle) alors on dit que A est une **racine** ou un **zéro** du polynôme $f(x)$.*

Exemple 1.8. Soit $A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix}$; $P_1(x) = 2x^2 - 3x + 5$; $P_2(x) = x^2 + 3x - 10$

On a :

$P_1[A] = \begin{pmatrix} 16 & -18 \\ -27 & 61 \end{pmatrix}$ *et* $P_2[A] = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

Donc A est une racine du polynôme $P_2(x)$.

5 Espace Vectoriel des matrices

On vérifie facilement le théorème suivant :

Théorème 1.10 (Espace vectoriel des matrices).

a) L'ensemble $\underline{\mathcal{M}_{\mathbb{K}}(n,m)}$ des matrices à n lignes et m colonnes, est un espace vectoriel sur $\underline{\text{le corps } \mathbb{K}}$.

b) L'ensemble $\underline{\mathcal{M}_{\mathbb{K}}(m,n)}$ est un espace vectoriel sur $\mathbb{K}$ isomorphe à $\mathcal{M}_{\mathbb{K}}(n,m)$ et $\dim_{\mathbb{K}} \mathcal{M}_{\mathbb{K}}\overline{(n,m) = \dim_{\mathbb{K}}} \mathcal{M}_{\mathbb{K}}(m,n)$

c) La base canonique dans $\mathbb{R}^{nm}$ de $\mathcal{M}_{\mathbb{K}}(n,m)$ est l'ensemble $\{E_i^j\}_{\substack{i=1,\cdots n \\ j=1,\cdots m}}$ avec :

$$E_i^j = \begin{pmatrix} & & 0 & & \\ & 0 & \vdots & 0 & \\ 0 & & 1 & & 0 \\ & 0 & 0 & 0 & \\ & & 0 & & \end{pmatrix} \quad \to \text{ligne } i$$

6 Matrices équivalentes - Matrices semblables

Définition 1.13. *Soient les matrices $A \in \mathcal{M}_{\mathbb{K}}(n,m)$ et $B \in \mathcal{M}_{\mathbb{K}}(n,m)$ telles qu'il existe deux matrices carrées inversibles $R,\ S$ avec : $\underline{B = R\,A\,S}$ alors A et B sont appelées $\underline{\text{équivalentes}}$.*

Si $A,\ \overline{B}$ sont deux matrices carrées et $R = S^{-1}$ (donc : $\underline{B = S^{-1}\,AS}$) alors A et B sont appelées $\underline{\text{semblables}}$.

3 Applications linéaires

1 Definition-Exemples

Définition 1.14 (Applications linéaires).
Soient E, F deux espaces vectoriels sur le corps K et soit $f : E \to F$ une application telle que $\forall u_1, u_2, \in E$ et $\forall \lambda \in K$ les deux conditions d'homomorphisme des deux espaces vectoriels sont verifiées :

$$\begin{aligned} f(u_1 + u_2) &= f(u_1) + f(u_2) \\ f(\lambda u_1) &= \lambda f(u_1) \end{aligned}$$

f est alors une application linéaire de E dans F.

Remarque 1.10. Lorsque $E = F$ alors l'application linéaire $f : E \to E$ est appelée un endomorphisme de E ou un opérateur linéaire agissant sur E.

Exemple 1.9.
a) Considérons $\mathbb{R}^3$ et soit :

$$\begin{aligned} f_1 : \quad \mathbb{R}^3 &\to \mathbb{R}^2 \\ (x, y, z) &\to (x, y) \end{aligned}$$

$$\begin{aligned} f_2 : \quad \mathbb{R}^3 &\to \mathbb{R}^2 \\ (x, y, z) &\to (y, z) \end{aligned}$$

f_1 et f_2 définissent deux applications linéaires de $\mathbb{R}^3$ dans $\mathbb{R}^2$ connues sous le nom de projections de $\mathbb{R}^3$ dans les plans $P_{(x,y)}$ et $P_{(y,z)}$ respectivement.

b) Soit E l'espace vectoriel des fonctions numériques continues sur $I = [a, b] \subset \mathbb{R}$
L'application $f : E \to \mathbb{R}$ telle que $\forall x \in E$

$$x \mapsto \int_0^1 x(t)dt$$

est une application linéaire de E sur $\mathbb{R}$.

Théorème 1.11.
a) L'application composée de deux applications linéaires est une application linéaire.
b) Si $f : E \to F$ application linéaire alors :
b.1) $f(\{0\}_E) = \{0\}_F$
b.2) A sous espace vectoriel de $E \Rightarrow f(A)$ sous-espace vectoriel de F.
b.3) B sous espace vectoriel de $F \Rightarrow f^{-1}(B)$ sous-espace vectoriel de E.

Remarque 1.11.
On supposera par la suite que E et F sont deux espaces vectoriels donnés sur $\mathbb{K}$ et on vérifie facilement que l'ensemble des applications linéaires $f : E \to F$ est un espace vectoriel sur $\mathbb{K}$. Notation : $\mathcal{L}_K(E, F)$

2 Image et Noyau

Définition 1.15 (Image et Noyau d'une application linéaire).

*Soit $f : E \to F$ une application linéaire $f \in \mathcal{L}_K(E, F)$ on appelle **image** de f et on note Imf, le sous-espace $f(E)$ de F : $Im\ f = f(E) \subset F$.*

*Le **noyau de** $f \in \mathcal{L}_K(E, F)$ noté par $\mathcal{K}er\ f$ est défini par l'ensemble :*

$\underline{\mathcal{K}er\ f = f^{-1}(0) \subset E}$

qui est un sous-espace vectoriel de E.

Exemple 1.10. Les noyaux des applications f_1 et f_2 respectivement données dans l'exemple 1.9 sont :

$$\begin{aligned} \mathcal{K}er\ f_1 &= \mathbb{R}(z)(\textit{ l'axe des "z"}). \\ \mathcal{K}er\ f_2 &= \mathbb{R}(x)(\textit{ l'axe des "x"}). \end{aligned}$$

alors que les images sont $Im\ f_1 = P_{(x,y)}, \quad Im\ f_2 = P_{(y,z)}$ respectivement.

Théorème 1.12.

a) $f \in \mathcal{L}_{\mathbb{K}}(E, F)$ est <u>injective</u> ssi $\mathcal{K}er\ f = \{0\}_E$

b) $f \in \mathcal{L}_{\mathbb{K}}(E, F)$ est <u>surjective</u> ssi $Im\ f = F$.

c) Si $f \in \mathcal{L}_{\mathbb{K}}(E, F)$ est un isomorphisme de E dans F

$\Rightarrow f^{-1}$ isomorphisme de F dans E.

Théorème 1.13. *Soit $f \in \mathcal{L}_{\mathbb{K}}(E, F)$ et E_2 supplémentaire de $\mathcal{K}er\ f$ par rapport à E : $\mathcal{K}er\ f \oplus E_2 = E$.*

$\Rightarrow f(E_2) = Im\ f$, et la <u>restriction g</u> de f à E_2 est un <u>isomorphisme</u> de E_2 sur $f(E)$.

Définition 1.16 (Rang d'une application).

On appelle <u>rang</u> d'une application $f \in \mathcal{L}_{\mathbb{K}}(E, F)$ et on note $\underline{rgf}$ la dimension (finie) de $f(E)$ (lorsque $\dim_{\mathbb{K}} E < +\infty$) et on a le résultat suivant :

Théorème 1.14 (Théorème du rang).

$$rg(f) = \dim_{\mathbb{K}} E - \dim_{\mathbb{K}}(\mathcal{K}er\ f)$$

En plus, on peut construire une <u>base</u> dans $\underline{f(E) = Im\ f}$ si on se donne une base de E d'après le :

Théorème 1.15.

Soit : $f \in \mathcal{L}_{\mathbb{K}}(E, F)$ et $rg(f) = r,\ \dim_{\mathbb{K}} E = n$;

Soit $\{a_1, \cdots, a_n\}$ = base de E telle que $\{a_{r+1}, \cdots, a_n\}$ est une base de $\mathcal{K}er\ f$

$\Rightarrow \{f(a_1), f(a_2), \cdots, f(a_r)\}$ est une base de $Im\ f$.

4 Applications linéaires et Matrices

1 Relation entre Applications linéaires et Matrices

Il existe un isomorphisme entre l'espace vectoriel sur $\mathbb{K}$, $\mathcal{L}_K(E, F)$

et l'espace vectoriel sur $\mathbb{K}$, $\mathcal{M}_{\mathbb{K}}(m, n)$.

Théorème 1.16.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E,F)$ et soit $\{a_i\}_{i=1,\cdots,n}$ une base de $E(\dim_{\mathbb{K}} E = n)$, $\{b_j\}_{j=1,\cdots,m}$ une base de $F(\dim_{\mathbb{K}} F = m)$ respect. alors l'application :

$$f \to M(f\{a_i\},\{b_j\}) = \{\alpha_{ij}\}_{\substack{i=1,\cdots,m \Rightarrow \text{ lignes} \\ j=1,\cdots,n \Rightarrow \text{ colonnes}}}$$

définie par : $\quad f(a_i) = \sum_{j=1}^{m} \alpha_{ji} b_j \quad (i = 1,\cdots,n)$
est une <u>bijection</u> de $\mathcal{L}_{\mathbb{K}}(E,F)$ sur $\mathcal{M}_{\mathbb{K}}(m,n)$. Chaque colonne de la matrice M a comme <u>éléments</u> les <u>coordonnées</u> dans la base $\{b_j\}_{j=1,\cdots,m}$ de l'image $f(a_i)$ de chacun des vecteurs a_i (de la base dans E).

** M est appelée <u>matrice associée</u> à l'application linéaire f, ou <u>représentation matricielle</u> de f.*

Remarque 1.12.

Le théorème ci-dessus implique que : Les deux espaces vectoriels sur $\mathbb{K}$, $\mathcal{L}_{\mathbb{K}}(E,F)$ et $\mathcal{M}_{\mathbb{K}}(m,n)$ sont isomorphes et donc

$$\dim_{\mathbb{K}} \mathcal{L}_{\mathbb{K}}(E,F) = \dim_{\mathbb{K}} \mathcal{M}_{\mathbb{K}}(m,n) = nm = \dim_{\mathbb{K}} \mathcal{M}_{\mathbb{K}}(n,m)$$

<u>En plus</u> :

Théorème 1.17.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E,F)$ et M <u>matrice associée</u> à f, dans $\mathcal{M}_{\mathbb{K}}(m,n)$,

$$\Rightarrow$$

$$rg\ f = rgM$$

Exemple 1.11.

Soit $f \in \mathcal{L}_{\mathbb{R}}(\mathbb{R}^3,\mathbb{R}^2)$ définie par :

$$f(x,y,z) = (3x+2y-4z, x-5y+3z)$$

Cherchons la représentation matricielle de f dans les bases :

$$B^{(3)} = \{(1,1,1);(1,1,0);(1,0,0)\} \text{ de } R^3 \text{ et } B^{(2)} = \{(1,3),(2,5)\} \text{ de } \mathbb{R}^2.$$

On cherche d'abord les coordonnées d'un vecteur qui p.rapport à la base canonique de $\mathbb{R}^2$ est noté : $u = (a,b)$). On a :

$$\begin{array}{l} (a,b) = x(1,3) + y(2,5) = (x+2y, 3x+5y) \\ \Rightarrow \left| \begin{array}{l} x+2y = a \\ 3x+5y = b \end{array} \right. \Rightarrow \left| \begin{array}{l} x = 2b-5a \\ y = 3a-b \end{array} \right. \\ \Rightarrow \underline{(a,b) = (2b-5a)(1,3) + (3a-b)(2,5)} \end{array}$$

Donc en appliquant f sur les vecteurs de B^3

$$\begin{aligned}
f(1,1,1) &= (3+2-4, 1-5+3) = (1,-1) \text{ \textit{(p.rapport à la base canonique de)} } \mathbb{R}^2 \\
&= (-2-5)(1,3) + (3+1)(2,5) \\
&\Rightarrow \underline{f(1,1,1) = -7(1,3) + 4(2,5)} \qquad (C.1.) \\
f(1,1,0) &= (3+2, 1-5) = (5-4) \\
&= (-8-25)(1,3) + (15+4)(2,5) \\
&\Rightarrow \underline{f(1,1,0) = (-33)(1,3) + 19(2,5)} \qquad (C.2.) \\
f(1,0,0) &= (3,1) = (2-15)(1,3) + (9-1)(2,5) \\
&= -13(1,3) + 8(2,5) \\
&\Rightarrow \underline{f(1,0,0) = -13(1,3) + 8(2,5)} \qquad (C.3.)
\end{aligned}$$

De (C.1.) et (C.2.) et (C.3.) on obtient les colonnes de la matrice M associée à f :

$$M^{(2)}_{(3)} = \begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix}$$

$$\left(\begin{array}{lcl} \textit{attention : : } \dim_{\mathbb{R}} \mathbb{R}^3 & = & \textit{nombre de colonnes = 3} \\ \dim_{\mathbb{R}} \mathbb{R}^2 & = & \textit{nombre de lignes = 2} \end{array} \right)$$

2 Changement de base - Matrice de passage

Théorème 1.18. *[Changement de base]*

Soit E espace vectoriel sur $\mathbb{K}$ avec $\dim_{\mathbb{K}} E = n$; supposons qu'on a deux bases $\{a_i\}_{i=1,\cdots,n}$ et $\{a'_i\}_{i=1,\cdots,n}$ de E. La matrice qui a comme colonne de numéro i les coordonnées de a'_i sur la base $\{a_i\}_{i=1,\cdots,n}$, est une matrice inversible P et de plus, si on associe à un vecteur $x \in E$ les matrices unicolonnes X et X' par rapport à la base $\{a_i\}$ et $\{a'_i\}$ respectivement on a :

$$\begin{array}{ccccc} X & = & PX' & ; & X' = P^{-1}X \\ \downarrow & & & & \downarrow \\ \textit{anciennes coordonnées} & & & & \textit{nouvelles coordonnées} \end{array}$$

Remarque 1.13.

La matrice P s'appelle **matrice de passage** de la base $\{a_i\}$ à la base $\{a'_i\}$ et est associée à l'application identité de $E \to E$ autrement dit :

$$P = M(Id_E, (a'_i), (a_i)) \textit{ et } P^{-1} = M(Id_E, (a_i), (a'_i))$$

Et voici, pour compléter ce cours, le théorème qui exprime l'action d'un **changement de base** sur la représentation matricielle d'une application linéaire.

Théorème 1.19.

a) Soit $E,\ F$ espaces vectoriels de dimension finie sur $\mathbb{K}$ et soit $\{a_i\},\ \{a'_i\}$ deux bases de E et $\{b_j\},\ \{b'_j\}$ deux bases de F.
Si P est la matrice de passage de $\{a_i\}$ à $\{a'_i\}$ et Q la matrice de passage de $\{b_j\}$ à $\{b'_j\}$;
Si $f \in \mathcal{L}_{\mathbb{K}}(E, F)$ et A est sa représentation matricielle par rapport à $\{a_i\}$ $\{b_j\}$ alors :
$\Rightarrow$ La matrice associée à f par rapport à $\{a'_i\}$ et $\{b'_j\}$ est de la forme :

$$\underline{A' = Q^{-1}AP} \qquad (a)$$

b) Si f est un endomorphisme de E alors :

$$A' = P^{-1}AP \qquad (b)$$

(P étant la matrice de passage de $\{a_i\}$ à $\{a'_i\}$)

Remarque 1.14.

Les matrices $A,\ A'$ dans (a) (resp. (b)) sont équivalentes (resp. semblables).

Chapitre 2

Matrices - Endomorphismes - Réduction (B)

1 Invariants par similitude d'un endomorphisme ou de sa matrice associée

1 Généralités

Soient :
$\mathcal{L}_{\mathbb{K}}(E)$ l'espace vectoriel des endomorphismes de $E \to E$ (esp. vec. sur $\mathbb{K}$) et
$\mathcal{M}_n(\mathbb{K})$ l'espace vectoriel des matrices carrées $n \times n$ sur $\mathbb{K}$.

L'isomorphisme entre ces deux espaces permettra dans tout ce qui suit de parler indifféremment d'un endomorphisme ou de sa matrice associée.

Définition 2.1.

On appelle ***invariant par similitude*** *d'une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ (ou de son endomorphisme $f \in \mathcal{L}_{\mathbb{K}}(\mathrm{E})$ associé) toute application $H : \mathcal{M}_n(\mathbb{K}) \to \mathcal{M}_n(\mathbb{K})$ qui associe la même valeur à A et à toute matrice semblable $B(= P^{-1}AP)$ de A : $H(A) = H(B)$.*

Les invariants qui nous intéressent sont :

Le déterminant, la trace, le polynôme caractéristique et le polynôme minimal d'une matrice (ou de son endomorphisme).

2 Déterminant

Définition 2.2 (Permutations).

Soit l'ensemble $X = (1, 2, ..., n)$.

On appelle <u>*permutation*</u> *σ toute bijection de X sur lui-même et on note :*

$$\sigma = \begin{pmatrix} 1, 2, ..., n \\ j_1 j_2 ... j_n \end{pmatrix}$$

ou :

$$\sigma = j_1, j_2, ..., j_n$$

On note par S_n l'ensemble des permutations σ de X, et on a :

$$card S_n = n!$$

*On dira que $\sigma \in S_n$ est **paire** s'il existe un nombre pair de couples du type "p" :" $(i,k) \in \sigma$ tel que $i > k$ mais i précède k dans σ."*

Exemple 2.1.
Pour $\sigma \in S_5$ définie par : $\sigma = (3,5,1,4,2)$ on a 6 couples du type "p" qui sont : $(3,1)\,(3,2)\,(5,1)\,(5,4)\,(5,2)\,(4,2) \Rightarrow \sigma$ **permutation paire**

Par contre pour $\sigma' \in S_5$ définie par : $\sigma' = (5,3,1,4,2)$ on a 7 couples du type "p" qui sont : $(5,3)\,(5,1)\,(5,4)\,(5,2)\,(3,1)\,(3,2)\,(4,2) \Rightarrow \sigma'$ **permutation impaire**

Rappel : **La signature** $sgn(\sigma)$ (ou la **parité**) d'une permutation σ est définie par :

$$sgn(\sigma) = \left\{ \begin{array}{cc} 1 & \text{si } \sigma \text{ paire} \\ -1 & \text{si } \sigma \text{ impaire} \end{array} \right\}$$

Définition 2.3. *Déterminant $|A|$ d'une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$*

Soit $A = \{a_{ij}\}_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$

On appelle <u>déterminant</u> noté $\det(A)$ ou $|A|$ l'application :

$$\mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$$

$$A \mapsto |A| = \sum_{\sigma \in S_n} (sgn(\sigma)) a_{1j_1} a_{2j_2} \ldots a_{nj_n}$$

et on écrit aussi :$|A| = \begin{vmatrix} a_{11} & \ldots & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \ldots & \ldots & a_{nn} \end{vmatrix}$

Exemple 2.2.
Comme

$$S_2 = \left\{ \begin{array}{cc} \sigma_+ = (1,2) & \text{paire} \\ \sigma_- = (2,1) & \text{impaire} \end{array} \right\}$$

alors $|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$

De même si :

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

alors

$$S_3 = \left\{ \begin{array}{cc} \sigma_+^1 = (1,2,3)\,;\, \sigma_+^2 = (2,3,1)\,;\, \sigma_+^3 = (3,1,2) & \text{paires} \\ \sigma_-^1 = (3,2,1)\,;\, \sigma_-^2 = (2,1,3)\,;\, \sigma_-^3 = (1,3,2) & \text{impaires} \end{array} \right\}$$

donc

$$|A| = \sum_{\sigma+} - \sum_{\sigma-} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

L'application de la définition 2.3 devient très difficile si on veux calculer le déterminant pour $n > 3$. Rappelons la méthode la plus pratique donnée par le théorème suivant :

Théorème 2.1. *Soit :*

$$A = \{a_{ij}\} \begin{array}{l} i = 1, ..., n \\ j = 1, ..., n \end{array}$$

alors

$$|A| = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} |M_{ij}|$$

*où M_{ij} est le **mineur associé à l'élément** a_{ij} autrement dit : M_{ij} est la sous-matrice carrée de A obtenue en supprimant la i-ème ligne et la j-ème colonne de A.*

Rappel.
Par la suite on utilisera aussi **le cofacteur** qui est égal à :

$$(-1)^{i+j} |M_{ij}|$$

Théorème 2.2.
Propriétés des déterminants

i) *Soit $A \in \mathcal{M}_n(\mathbb{K})$*

a) Si A a une ligne (ou colonne) de zéros alors $|A| = 0$

b) Si A a deux lignes (ou colonnes) identiques alors $|A| = 0$

c) Si A est une matrice triangulaire $\Rightarrow |A| = \prod_{i=1}^{n} a_{ii}$ (produit des éléments diagonaux)

En particulier $|I_n| = 1$

ii) *Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $B \in \mathcal{M}_n(\mathbb{K})$ et supposons que B est obtenue de A :*

a) par multiplication d'une ligne (ou colonne) de A par $k \in \mathbb{K}$
$\Rightarrow |B| = k|A|$

b) en échangeant deux lignes (ou colonnes) de A,
$\Rightarrow |B| = -|A|$

c) en additionnant un multiple d'une ligne (ou colonne) de A à une autre
$\Rightarrow |B| = |A|$

d) par transposition $B = A^t \Rightarrow |B| = |A|$

iii) *Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $B \in \mathcal{M}_n(\mathbb{K})$*

$$\Rightarrow |AB| = |A||B|$$

iv) $|A^{-1}| = |A|^{-1}$

v) *Invariance par similitude (ou par changement de base) Si $B = P^{-1}AP$ (B est semblable à A)*

$$\Rightarrow \quad |B| = |A|$$

Remarque 2.1.

C'est justement la troisième propriété (invariance) qui nous permet d'associer un **unique déterminant** à tout endomorphisme $f \in \mathcal{L}_{\mathbb{K}}(E)$ qui est précisément le déterminant de toutes les représentations matricielles A_f de l'endomorphisme f.

Théorème 2.3. *Autres propriétés d'un déterminant*

Soit $A \in \mathcal{M}_n(\mathbb{K})$, alors on a équivalence entre :

i) *A est inversible (A^{-1} existe)*

ii) *A est **non singulière**, c'est-à-dire $AX = 0 \quad (X \in \mathbb{K}^n)$ admet seulement la <u>solution nulle</u>, ou $rgA = n$*

iii) $\det(A) \neq 0$

Applications d'un déterminant

Théorème 2.4. *Calcul de l'inverse d'une matrice*

*Soit M la **comatrice** (matrice des cofacteurs A_{ij}) d'une matrice carrée*

$$A = a_{ij} \begin{array}{l} i \in X \\ j \in X \end{array}$$

La matrice inverse A^{-1} de A est égale à : $A^{-1} = \dfrac{M^T}{|A|}$ où M^T est la transposée de M)

Pour l'application aux systèmes linéaires on a :

Théorème 2.5.

i) ***Règle de Cramer***

Soit (S_1) un système linéaire de n équations à n inconnues : $AY = b \quad (S_1)$

avec $\left.\begin{array}{l} Y = \{Y_1 ... Y_n\}^T \\ b = \{b_1 ... b_n\}^T \end{array}\right\}$ *vecteurs colonnes dans $\mathbb{K}^n$ et $A \in \mathcal{M}_n(\mathbb{K})$*

Alors (S_1) a une solution unique $\bar{Y}$ ssi $|A| \neq 0$.

ii) *Le système homogène $\underline{AY = 0}$ admet au moins une solution non triviale ssi $|A| = 0$*

Remarque 2.2.

D'après les théorèmes (2.3.ii) et (2.5ii) le système homogène $AY = 0$ de la partie i) du théorème 2.5 n'admet que la solution triviale : (0,0,0,...,0). Le système complet (S_1) est alors appelé **système de Cramer**. Pour le calcul de $\bar{Y}$ (solution unique de (S_1) on a :

Théorème 2.6.

Un système de Cramer (S_1) a pour solution unique :

$$\bar{Y}_i = (\det B_i)(\det A)^{-1}$$

où B_i est la matrice obtenue en remplaçant dans A le vecteur colonne de coefficients de y_i par le vecteur colonne b des seconds membres.

Remarque 2.3.

Autre méthode de calcul de l'inverse A^{-1} differente de celle du théorème (2.4)

D'après les théorèmes 2.5 et 2.6, pour trouver la solution $\bar{Y}$ de S_1 il suffit de connaître A^{-1} puisque $\bar{Y} = A^{-1}b$.

Réciproquement pour trouver l'inverse d'une matrice A^{-1} avec $|A| \neq 0$ on résoud d'abord un système du type (S_1) par triangularisation - méthode du pivot de Gauss ou méthode de Jordan-Gauss, "échelonnage" en e.r.l (voir chapitre précédent). Supposons donc qu'on a obtenu la solution $\bar{Y} = \bar{y}_1, \bar{y}_2, ..., \bar{y}_n$ en fonction des seconds membres, c'est-à-dire :

$$\left.\begin{array}{rl} \bar{y}_1 = & \alpha_{11}b_1 + \alpha_{12}b_2 + ... + \alpha_{1n}b_n \\ \vdots & \qquad\qquad \vdots \\ \bar{y}_n = & \alpha_{n1}b_1 + \alpha_{12}b_2 + ... + \alpha_{nn}b_n \end{array}\right\} \Leftrightarrow \bar{Y} = A^{-1}b$$

donc

$$A^{-1} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & ... & \alpha_{1n} \\ \alpha_{21} & ... & ... & ... \\ ... & ... & ... & ... \\ \alpha_{n1} & ... & ... & \alpha_{nn} \end{pmatrix}$$

Exemple 2.3.

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

On vérifie que A est inversible : $\det A \neq 0$. Ensuite on résoud le système suivant : $AY = b$ où $b = (b_1, b_2, b_3)^T \in \mathbb{K}^3$ est supposé connu. Autrement dit, on cherche $\bar{Y} = (x, y, z)^T$ tel qu'il soit solution de :

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 2 \\ 4 & 3 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \Leftrightarrow \left\{\begin{array}{rl} x + 2y + z = b_1 & (L_1) \\ x + 2y + 2z = b_2 & (L_2) \\ 4x + 3y + 5z = b_3 & (L_3) \end{array}\right\} (S_1)$$

Pour les transformations suivantes sur les lignes (Pivot de Gauss) (ou échelonnage) :

$$\left\{\begin{array}{l} L_2' = L_2 - L_1 \\ L_3' = L_3 - 4L_1 \end{array}\right. \text{ et } \begin{array}{l} L_2" = L_3' \\ L_3" = L_2' \end{array} \Rightarrow \left\{\begin{array}{l} x + 2y + z = b_1 \\ -5y + z = b_3 - 4b_1 \\ z = b_2 - b_1 \end{array}\right\} (S_2)$$

Le système S_2 est déjà sous forme triangulaire et la solution de S_2 (et de S_1) est obtenue facilement :

$$\bar{Y} \equiv \left\{\begin{array}{l} x = \frac{4}{5}b_1 - \frac{7}{5}b_2 + \frac{2}{5}b_3 \\ \\ y = \frac{3}{5}b_1 + \frac{1}{5}b_2 - \frac{1}{5}b_3 \\ \\ z = -b_1 + b_2 \end{array}\right\}$$

donc comme $\bar{Y} = A^{-1}b$, on a

$$A^{-1} = \begin{pmatrix} \frac{4}{5} & -\frac{7}{5} & \frac{2}{5} \\ \\ \frac{3}{5} & \frac{1}{5} & -\frac{1}{5} \\ \\ -1 & 1 & 0 \end{pmatrix}$$

3 La trace d'une matrice ou de l'endomorphisme associé

Définition 2.4.
*On appelle **trace** d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$, (avec $A = \{a_{ij}\}_{\substack{i \in X \\ j \in X}}$, l'application de : $\begin{array}{l}\mathcal{M}_n(\mathbb{K}) \to \mathbb{R} \\ A \mapsto Tr[A]\end{array}$, qui à toute matrice A associe le scalaire, qui est égal à la somme de tous les éléments diagonaux a_{ii} de A :*

$$Tr[A] = \sum_{i \in X} a_{ii}$$

Remarque 2.4.
A partir de cette définition on démontre facilement les propriétés de "commutativité des matrices à l'intérieur d'une trace et de l'additivité des traces :

Théorème 2.7. *Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $B \in \mathcal{M}_n(\mathbb{K})$ deux matrices carrées*

i) $Tr[AB] = Tr[BA]$

ii) $Tr[A + B] = Tr[A] + Tr[B]$

En utilisant la propriété i) de ce théorème, on obtient :

Théorème 2.8. *[L'invariance par similitude de la trace]*
Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $B \in \mathcal{M}_n(\mathbb{K})$ deux matrices semblables ($B = P^{-1}AP$)

$$\Rightarrow Tr[A] = Tr[B]$$

Exemple 2.4.
Soit

$$A = \begin{pmatrix} 2 & 0 & 4 \\ 3 & -4 & 12 \\ 1 & -2 & 5 \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

et

$$P = \begin{pmatrix} -4 & -4 & 2 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \Rightarrow P^{-1} = \begin{pmatrix} -1/2 & 1 & -2 \\ 1 & -2 & 5 \\ 3/2 & -2 & 6 \end{pmatrix}$$

On définit $B = P^{-1}AP$ et on obtient :

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \Rightarrow Tr[B] = 1 + 2 = 3$$

On a : $Tr[A] = 2 - 4 + 5 = 3$. Donc $Tr[A] = Tr[B]$ le théorème est vérifié.

Remarque 2.5.

L'invariance pour similitude de la trace d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ permet d'associer à tout endomorphisme $f \in \mathcal{L}_{\mathbb{K}}(E)$ un unique scalaire la trace de f qui est précisément la trace de toute représentation matricielle A de f :

$$Tr[f] = Tr[A]$$

4 Le polynôme caractéristique d'un endomorphisme (ou de sa matrice associée)

$$P_A(x) \quad \text{ou} \quad P_f(x)$$

Soit $A \in \mathcal{M}_n(\mathbb{K})$

Définition 2.5.

On considère la matrice $(xI_n - A)$, $x \in \mathbb{K}$.

On appelle ***polynôme caractéristique*** *de* A *le déterminant de cette matrice et on note :*

$$\det[xI_n - A] \equiv P_A(x) \tag{2.1.1}$$

D'après la propriété d'invariance par similitude d'un déterminant, on obtient immédiatement :

Théorème 2.9.

Soit A *et* B *deux matrices semblables dans* $\mathcal{M}_n(\mathbb{K})$

$\Rightarrow$

$$P_A(x) = P_B(x)$$

(invariance par similitude du polynôme caractéristique)

Remarque 2.6.

L'invariance par similitude permet à nouveau d'associer à chaque endomorphisme f un <u>unique</u> polynôme qui s'appelle **polynôme caractéristique de** $f \in \mathcal{L}_{\mathbb{K}}(E)$ et qui est égal à $P_A(x)$ de toute représentation matricielle de f c'est-à-dire : $P_A(x) = P_f(x)$

Remarque 2.7.

Toutes les propriétés présentées précédemment à propos des déterminants sont évidemment vérifiées par le polynôme caractéristique.

De plus, si $A = \{a_{ij}\}_{\substack{i \in X \\ j \in X}} \in \mathcal{M}_n(\mathbb{K})$, on obtient en développant (2.1.1)

$P_A(x) = (x - a_{11})(x - a_{22})...(x - a_{nn}) +$ (termes avec plus de $(n-2)$ facteurs de la forme $(f - a_{ii})$)

ou

$P_A(x) = x^n - (a_{11} + a_{22}) + ... + a_{nn})x^{n-1} +$ (termes de degré inférieur)

Donc le polynôme caractéristique est un polynôme normalisé (unitaire) et *le **coefficient du terme** x^{n-1} **est égal à*** $(-1)Tr[A]$.
Le terme constant (x^0) est obtenu par $P_A(x = 0)$.
Donc d'après la définition (2.1.1), $\underline{\alpha_0 = (-1)^n |A|}$.

Rappelons pour terminer le théorème fondamental et quelques-unes de ses applications (voir les valeurs propres et la diagonalisation) qui sont parmi les plus importantes de l'Algèbre linéaire.

Théorème 2.10. *de Cayley-Hamilton*
Toute matrice (ou tout endomorphisme) est une racine de son polynôme caractéristique :

$$P_A(A) = [0] \qquad \text{ou} \qquad P_f(f) = [0]$$

(Attention ! la matrice $[0] \in \mathcal{M}_n(\mathbb{K})$*)*

Exemple 2.5.
Soit

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

$$\Rightarrow$$

$$P_A(x) = |xI_2 - A| = \begin{vmatrix} x-1 & 2 \\ 3 & x-2 \end{vmatrix} = (x-1)(x-2) - 6 = x^2 - 3x - 4$$

et

$$A^2 - 3A - 4I_2 = \begin{pmatrix} 7 & 6 \\ 9 & 10 \end{pmatrix} - 3\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} - 4\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Donc le théorème 2.10 de Cayley-Hamilton est vérifié.

5 Polynôme minimal d'un endomorphisme (ou de sa matrice associée)

Définition 2.6.
*On appelle **polynôme minimal** d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ le polynôme noté $m_A(x)$ **unitaire (normalisé)**, de plus bas degré parmi tous les polynômes $P(x)$ tels que $P[A] = [0]$ (c'est-à-dire A est racine de son polynôme minimal).*

Théorème 2.11.

i) $\forall A \in \mathcal{M}_n(\mathbb{K})$ *le polynôme minimal $m_A(x)$ existe et il est <u>unique</u>.*

ii) *le polynôme minimal $m_A(x)$, $\forall A \in \mathcal{M}_n(\mathbb{K})$ divise tous les polynômes $P \neq 0$ tel que $P[A] = 0$. En particulier, il divise le polynôme caractéristique $P_A(x)$.*

iii) $\forall A \in \mathcal{M}_n(\mathbb{K})$ *, la polynôme minimal* $m_A(x)$ *et le polynôme caractéristique ont les mêmes facteurs irréductibles.*

Exemple 2.6. Soit

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

On a :

$$P_A[x] = \det[I_4 x - A] = \begin{vmatrix} x-2 & -1 & 0 & 0 \\ 0 & x-2 & 0 & 0 \\ 0 & 0 & x-2 & 0 \\ 0 & 0 & 0 & x-5 \end{vmatrix}$$

Et d'après le théorème , on a directement : $\underline{(x-2)^3(x-5)}$.
Les diviseurs de $P_A(x)$ sont :

$$\begin{cases} m_1(x) = (x-2)(x-5) \\ m_2(x) = (x-2)^2(x-5) \\ m_1(x) = (x-2)^3(x-5) \end{cases}$$

On vérifie que

$$\begin{cases} m_1(A) \neq 0 \\ m_2(A) = 0 \\ m_3(A) = 0 \end{cases}$$

Donc comme $m_3(A) = P_A[x] \Rightarrow$ Le polynôme minimal $m_A(x) = m_2(x)$

Exemple 2.7.

Soit A une matrice 3×3 sur le corps $\mathbb{R}$ et soit le polynôme $P(x) = x^2 + 1$.

On montre que : A **ne peut pas être racine de** $P(x)$.

D'après le théorème 2.10 de Cayley-Hamilton, A est un zéro de $P_A(x)$ et $P_A(x)$ est de **degré 3**.

Donc il a au moins une **racine réelle**.

Si on suppose que A est une racine de $P(x)$, comme ce dernier est irréductible sur $\mathbb{R}$, il doit être le polynôme minimal : $P(x) = m_A(x)$ de A.

Mais $P(x)$ n'a pas de racine réelle.

D'où <u>contradiction</u> avec la partie iii) du théorème précédent 2.11.

Conclusion : A n'est pas un zéro de $P(x)$

Au contraire, on peut vérifier que la matrice $\tilde{A}$ ci-dessous, définie sur le corps des complexes, est un zéro du polynôme $P(x) = x^2 + 1$:

$$\tilde{A} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & i \end{pmatrix} \in \mathcal{M}_3(\mathbf{C})$$

et

$$P(\tilde{A}) = \tilde{A}^2 + I_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Théorème 2.12 (Invariance par similitude du polynôme minimal $m_A(x)$).
Pour tout couple A, B *de matrices carrées semblables, on a :* $m_A(x) = m_B(x)$.

Remarque 2.8.
L'invariance par similitude du polynôme minimal de $A \in \mathcal{M}_n(\mathbb{K})$ permet d'associer à tout endomorphisme $f \in \mathcal{L}_{\mathbb{K}}(E)$ un unique polynôme minimal $m_f(x)$ qui est le polynôme minimal $m_A(x)$ de toute représentation matricielle A de f et qu'on appellera polynôme minimal de l'endomorphisme f.

2 Valeurs propres - Vecteurs propres

1 Valeurs propres- Vecteurs propres

Dans tout ce qui suit, on supposera que E est un espace vectoriel de **dimension finie** sur $\mathbb{K}$. On notera par f tout endomorphisme $f \in \mathcal{L}_{\mathbb{K}}(E)$ (ou opérateur linéaire de E) et par A_f la matrice carrée dans $\mathcal{M}_n(\mathbb{K})$ associée à f.
$\mathbb{K}$ sera toujours **un corps commutatif**.

Définition 2.7.
On dit que F sous-espace vectoriel de E est un ***sous-espace f-invariant*** *de E si :*
$\forall v \in F \quad f(v) \in F$.
Autrement dit : $f(F) \subset F$.
On dit aussi que F est un ***sous-espace stable par*** *f.*

Définition 2.8.
On appelle ***valeur propre*** *λ de f (et respectivement de A_f) tout scalaire $\lambda \in \mathbb{K}$ tel qu'il existe un vecteur* ***non nul*** *$x \in E$ pour lequel : $f(x) = \lambda x$.*
Le vecteur x qui se transforme par f de cette façon "homothétique" est appelé ***vecteur propre*** *de f associé à la valeur propre λ.*
Le ***spectre de*** *f : $S_p(f) \subset \mathbb{K}$ est l'ensemble des valeurs propres de f.*

Remarque 2.9.

a) Dans la littérature, on rencontre souvent les valeurs ou vecteurs propres sous le nom de **valeurs ou vecteurs caractéristiques**.

b) Le vecteur $\{0\} \in E$ est un vecteur propre $\forall f \in \mathcal{L}_{\mathbb{K}}(E)$ puisque :

$$f(0) = \lambda.0 = 0 \quad \forall \lambda \in \mathbb{K}$$

c) Les deux définitions précédentes sont valables même si $dim_{\mathbb{K}} R = \infty$

d) Si $\lambda = 0$ est valeur propre de f alors :
$\exists x \in E \quad x \neq 0$ tel que $f(x) = 0$
Donc : $Ker f \neq \{0\}$

2 Sous-espace propre

Résultat fondamental :

Théorème 2.13.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$, alors :

i) *Pour tout vecteur propre x de f, il existe une valeur propre unique (associée au vecteur propre x).*

ii) *A toute valeur propre λ de f correspond un sous-espace $V(\lambda)$ de E engendré par tous les vecteurs propres associés à λ.*

$V(\lambda)$ est appelé ***sous-espace propre associé à*** *λ.*

C'est un sous-espace f-invariant de E et il est distinct de $\{0\}$.

iii) *Si $\lambda_1,\ \lambda_2$ sont deux valeurs propres distinctes de f avec $\lambda_1 \neq \lambda_2$, les sous-espaces vectoriels $V(\lambda_1)$, $V(\lambda_2)$ associés n'ont en commun que le vecteur nul de E :*

$$V(\lambda_1) \cap V(\lambda_2) = \{0_E\}.$$

iv) *Si f admet m valeurs propres distinctes (deux à deux) $\{\lambda_i\}_{i=1,\ldots,m}$ alors la famille $\{x_i\}_{i=1,\ldots,m}$ de vecteurs propres associés est libre, et*

$$\bigcup_i V(\lambda_i) = \oplus_i V(\lambda_i)$$

$\Leftrightarrow$ La somme des sous-espaces propres est une somme directe.

v) *Si $dim_{\mathbb{K}} E = n$ alors tout $f \in \mathcal{L}_{\mathbb{K}}(E)$ a* ***au plus*** *n* ***valeurs propres distinctes****.*

Remarque 2.10.

a) Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$.

Si $V(\mu)$ est l'ensemble des vecteurs $x \in E$ tel que $f(x) = \mu x$ alors la condition $\underline{V(\mu) = \{0\}}$ implique que μ n'est pas une valeur propre de f.

b) Si λ est valeur propre de $f \in \mathcal{L}_{\mathbb{K}}(E)$,

on montre facilement que λ^k **est valeur propre de** $f^k \quad \forall k \in \mathbf{N}$.

Si de plus, f est inversible, λ^k est valeur propre de $f^k \quad \forall k \in \mathbf{Q}$.

c) Les résultats i), ii), iii) et iv) du théorème précédent (2.13) sont également vérifiés dans le cas plus général où $dim_{\mathbb{K}} E = \infty$.

Exemples 2.1.

a) Soit $E = \mathbb{R}^2$ et $f \in \mathcal{L}_{\mathbb{R}}(\mathbb{R}^2)$ l'opérateur linéaire qui fait subir à tout vecteur $u \in \mathbb{R}^2$ une rotation d'un angle $\theta = 90°$. On vérifie immédiatement qu'aucun vecteur non nul n'est transformé en un multiple de lui-même. Donc f n'a **aucune** valeur propre, donc aussi **aucun** vecteur propre.

b) Soit E l'espace vectoriel des fonctions définies et différentiables sur $\mathbb{R}$, et soit la famille de fonctions :

$$\mathcal{F} = \left\{e^{a_1 t}, e^{a_2 t}, ..., e^{a_n t}\right\} \subset E$$

où $a_i \in \mathbb{R}^* \quad \forall i = 1, ..., n$. Si *D* est l'opérateur différentiel sur E, ($D \in \mathcal{L}_{\mathbb{R}}(E)$), on voit que :

$$D(e^{a_k t}) = a_k e^{a_k t} \quad \forall k = 1, ..., n.$$

Donc $\mathcal{F}$ constitue un ensemble de :
vecteurs propres de *D* associés aux **valeurs propres distinctes** $a_1, ..., a_n$ **de *D***, qui de plus sont linéairement indépendants ; autrement dit, les propriétés i) et iv) du théorème précédent sont vérifiées.

3 Le rôle du polynôme caractéristique

Définition 2.9.
*On appelle **multiplicité géométrique** d'une valeur propre $\lambda \in S_P(f)$, la dimension du sous-espace prore $V(\lambda)$ associé.*

Le rôle essentiel du polynôme caractéristique P_f pour la détermination des valeurs propres de f est donné par le théorème suivant.

Théorème 2.14.
Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ et $dim_{\mathbb{K}} E = n$

i) *Les valeurs propres de f sont les **racines** de son polynôme caractéristique $P_f(x)$*
*Si $\mathbb{K}$ est **algébriquement clos** alors f possède n valeurs propres distinctes ou confondues.*

ii) *Soit λ une racine multiple de $P_f(x)$ d'ordre k (appelé aussi **multiplicité algébrique**) alors :*

$$1 \leq dimV(\lambda) \leq k$$

Autrement dit : la multiplicité géométrique d'une valeur propre λ n'est **pas supérieure** à sa multiplicité algébrique.

Tenant compte de l'isomorphisme établi entre $\mathcal{M}_n(\mathbb{K})$ et $\mathcal{L}_{\mathbb{K}}(E)$, on parlera d'une manière équivalente des valeurs propres $\lambda(A)$ (ou du spectre $S_P(A)$ de la matrice A_f et de ses vecteurs propres associés. Les résultats précédents se résument sous forme d'un théorème d'équivalence comme ci-dessous :

Théorème 2.15.
Soit $A \in \mathcal{M}_n(\mathbb{K})$.
Pour tout $\lambda \in \mathbb{K}$, il y a équivalence entre les propriétés 1, 2 et 3 :

1) *λ est valeur propre de A.*

2) *$\lambda I_n - A$ **n'est pas une matrice inversible** .*

3) $\det(\lambda I_n - A) = 0 = P_A(\lambda)$

Voici deux propriétés (faciles à montrer) très intéressantes établissant une relation entre la trace ou le déterminant d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ et ses valeurs propres.

Théorème 2.16.

$A \in \mathcal{M}_n(\mathbb{K})$ et soit $\{\lambda_1(A)...\lambda_1(A)\} = S_{P_A}$ (le spectre de A) alors :

i) $TrA = \sum_{i=1}^{n} \lambda_i(A)$

ii) $\det A = \prod_{i=1}^{n} \lambda_i(A)$

3 Diagonalisation d'un endomorphisme ou de sa matrice associée.

1 Diagonalisation

Remarque 2.11.

La réduction d'un endomorphisme f ou de sa matrice associée A_f constitue l'objectif principal de ce cours. Il s'agit de résultats d'études réalisées par des mathématiciens éminents du 19^{me} et 20^{me} siècle. Ces études ont comme but de trouver la représentation matricielle A_f **la plus simple** grâce aux possibilités fournies par les **transformations de similitude** ou changement de base, dans un espace vectoriel E.

La forme simple de A_f facilitera par la suite ses applications (voir Analyse Numérique - Recherche opérationnelle etc.) et l'étude approfondie de ses propriétés .

La forme la plus simple d'une matrice étant la **diagonale**, on s'intéressera d'abord à l'étude des **conditions de diagonalisabilité** d'un endomorphisme ou de sa matrice associée. Les possibilités de réduction des matrices (ou des endomorphismes) non diagonalisables sous d'autres formes fera l'objet du prochain cours.

Définition 2.10.

*On dit que $f \in \mathcal{L}_{\mathbb{K}}(E)$ (resp. $A_f \in \mathcal{M}_n(\mathbb{K})$ **est diagonalisable**, s'il existe une base de E telle que la matrice associée $\tilde{A}_f$ soit diagonale. (resp. il existe une matrice carrée $P \in \mathcal{M}_n(\mathbb{K})$ inversible telle que $\tilde{A}_f = P^{-1}A_f P$ -la matrice semblable - soit diagonale.)*

On donne deux formes différentes de condition nécessaire et suffisante de diagonalisation, qu'on pourra choisir chaque fois suivant l'aspect du problème présenté.

Théorème 2.17.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ (resp. $A_f \in \mathcal{M}_{\mathbb{K}}(n)$ $dim_{\mathbb{K}}E = n$ alors f (resp. A_f) est diagonalisable **ssi** il existe une base de E formée de vecteurs propres de f (resp. de A_f).

Théorème 2.18.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ (resp. $A_f \in \mathcal{M}_{\mathbb{K}}(n)$) et $dim_{\mathbb{K}}E = n$ alors f (resp. A_f) est diagonalisable <u>ssi</u> les conditions suivantes sont satisfaites :

c.1) *Le polynôme caractéristique P_f (resp. P_A) a ses racines (distinctes ou confondues) dans $\mathbb{K}$*

c.2) *Pour toute racine λ_i de P_f (resp. de P_A) d'ordre k_i :*
$dimv(\lambda_i) = k_i \Rightarrow$ *la multiplicité géométrique de λ_i égale sa multiplicité algébrique.*

Une condition **suffisante** de diagonalisabilité, et utile pour les applications, est donnée par le théorème suivant :

Théorème 2.19.
Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ (resp. $A_f \in \mathcal{M}_{\mathbb{K}}(n)$) et $dim_{\mathbb{K}} E = n$ alors :

i) *f (resp. A_f) est diagonalisable **si** ses n valeurs propres sont **toutes distinctes** deux à deux.*

ii) *Si $\mathbb{K}$ **est algébriquement clos** alors* f *(resp. A_f) est diagonalisable, **si** toutes **les racines** du polynôme caractéristique P_f (resp. P_A) sont **simples.** est scindé en facteurs irréductibles linéaires.*

Remarque 2.12.
Dans le cas où les conditions des théorèmes précédents sont vérifiés, la matrice diagonale $\tilde{A}_f = P^{-1} A_f P$ a comme **éléments diagonaux les valeurs propres** de f (resp. de A_f). La matrice de passage P (ou matrice de transformation de similitude ou matrice de changement de base) est la matrice dont les colonnes sont les n vecteurs propres linéairement indépendants de f (resp. de A_f).

2 Exemples-Applications

Exemple 2.8 (Vérification du théorème 2.19).
Soit $f \in \mathcal{L}_{\mathbb{R}}(\mathbb{R}^3)$ tel que A_f a la forme suivante (dans la base canonique de $\mathbb{R}^3$

$$A_f = \begin{pmatrix} 2 & 0 & 4 \\ 3 & -4 & 12 \\ 1 & -2 & 5 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

On étudie la diagonalisabilité de f (ou de A_f, dont on déterminera d'abord les valeurs propres et les vecteurs propres.

On obtient :

a) *Pour les valeurs propres*

$$P_f = P_A = \begin{vmatrix} x-2 & 0 & 4 \\ 3 & x+4 & 12 \\ 1 & -2 & x-5 \end{vmatrix}$$

$$P_f = (2-x)(4+x)(5-x) - 24 - 4(4+x) - 24(2-x) \Rightarrow P_f = x(x-1)(x-2)$$

On a trois racines, donc trois valeurs propres qui appartiennent à $\mathbb{R}$:

$$\begin{cases} \lambda_1 = 0 \\ \lambda_2 = 1 \\ \lambda_3 = 2 \end{cases}$$

et par le théorème précédent on peut affirmer que f est diagonalisable.

b) Pour les vecteurs propres : On cherche les coordonnées du vecteur $\vec{r_i} = (u_i, v_i, w_i)$ vérifiant : $A_f \vec{r_i} = \lambda_i$

b.1) Pour $\lambda_1 = 0$, on obtient le système :

$$A_f \vec{r}_1 = \vec{0} \Leftrightarrow \left\{ \begin{array}{l} 2u_1 + 4w_1 = 0 \\ 3u_1 - 4v_1 + 12w_1 = 0 \\ u_1 - 2v_1 + 5w_1 = 0 \end{array} \right\}$$

$$\Rightarrow \frac{u_1}{-4} = \frac{w_1}{2} = \frac{v_1}{3} \Rightarrow dim V(\lambda_1) = 1$$

b.2) Pour $\lambda_2 = 1$, on obtient le système :

$$A_f \vec{r}_2 = \vec{r}_2 \Leftrightarrow \left\{ \begin{array}{l} 2u_2 + 4w_2 = u_2 \\ 3u_2 - 4v_2 + 12w_2 = v_2 \\ u_2 - 2v_2 + 5w_2 = w_2 \end{array} \right\}$$

$$\Rightarrow \begin{array}{l} u_2 = -4w_2 \\ v_2 = 0 \end{array} \Rightarrow dim V(\lambda_2) = 1$$

et finalement

b.3) Pour $\lambda_2 = 3$, on obtient le système :

$$A_f \vec{r}_3 = 2\vec{r}_3 \Leftrightarrow \left\{ \begin{array}{l} 2u_3 + 4w_3 = 2u_3 \\ 3u_3 - 4v_3 + 12w_3 = 2v_3 \\ u_3 - 2v_3 + 5w_3 = 2w_3 \end{array} \right\}$$

$$\Rightarrow \begin{array}{l} w_3 = 0 \\ u_3 = 2v_3 \end{array} \Rightarrow dim V(\lambda_3) = 1$$

D'après (b.1), (b.2) et (b.3), on peut choisir une base de $\mathbb{R}^3$ fermée des vecteurs propres de A_f (et de f :

$$\vec{r}_1 = \begin{pmatrix} -4 \\ 3 \\ 2 \end{pmatrix} \quad \vec{r}_2 = \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix} \quad \vec{r}_3 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

c) On définit la matrice de passage :$P = \begin{pmatrix} -4 & -4 & 2 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$

Avec une des méthodes connues, on calcule l'inverse P^{-1}.

On obtient $P^{-1} = \dfrac{1}{2} \begin{pmatrix} -1 & 2 & -4 \\ 2 & -4 & 10 \\ 3 & -4 & 12 \end{pmatrix}$.

Ensuite, on vérifie que la matrice semblable :

$$\tilde{A}_f = P^{-1}A_fP = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Donc effectivement $\tilde{A}_f$ est une matrice diagonale ayant comme éléments diagonaux les trois valeurs propres.
Le théorème est vérifié.

Exemple 2.9.

Soit la matrice $A = \begin{pmatrix} -4 & 0 & -2 \\ 0 & 1 & 0 \\ 5 & 1 & 3 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$
On obtient avec la même méthode que précédemment : $P_A(x) = (x-1)^2(x+2)$.
On trouve que $\dim V(1) = 1$ et $\dim V(-2) = 1$.
Donc $\dim V(1) < k_1$
$\Rightarrow$ D'après le théorème A n'est pas diagonalisable dans $\mathcal{M}_3(\mathbb{R})$

Exemple 2.10.

On étudie la diagonalisabilité :

a) dans $\mathcal{M}_3(\mathbb{R})$

b) dans $\mathcal{M}_3(\mathbb{C})$

de la matrice

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

Par la même méthode on obtient : $P_A(x) = (x-3)(x^2+1)$
Donc pour :

a) $A \in \mathcal{M}_3(\mathbb{R})$
$\lambda_1 = 3$: c'est la seule valeur propre de A dans $\mathcal{M}_3(\mathbb{R})$
Donc A n'est pas diagonalisable dans $\mathcal{M}_3(\mathbb{R})$.

b) $A \in \mathcal{M}_3(\mathbf{C})$ (A matrice dans le corps des complexes)
alors : $\lambda_1 = 3 \quad \lambda_2 = j \quad \lambda_3 = -j$
Donc A est diagonalisable dans $\mathcal{M}_3(\mathbf{C})$.

Exemple 2.11.

Etudions la <u>diagonalisabilité d'une matrice triangulaire $A \in \mathcal{M}_n(\mathbb{K})$</u>

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & ... & a_{1n} \\ 0 & a_{22} & ... & ... & a_{2n} \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & a_{nn} \end{pmatrix}$$

Puisque A est triangulaire et xI_n diagonale, la matrice $xI_n - A$ est aussi triangulaire ayant comme éléments diagonaux a_{ii} :

$$xI_n - A = \begin{pmatrix} x-a_{11} & a_{12} & ... & ... & a_{1n} \\ 0 & x-a_{22} & ... & ... & a_{2n} \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & a_{nn} \end{pmatrix}$$

Donc *:*

$$P_A(x) = \det|xI_n - A| = (x-a_{11})(x-a_{22})...(x-a_{nn})$$

$\Rightarrow$ Condition suffisante de diagonalisabilité de A *:*

$$a_{ii} \in \mathbb{K} \qquad \textit{et} \qquad a_{ii} \neq a_{jj} \quad i \neq j$$

Par exemple, la matrice triangulaire :

$$A = \begin{pmatrix} 2 & 6 & 0 & 3 \\ 0 & 4 & 1 & 5 \\ 0 & 0 & j & 530j \\ 0 & 0 & 0 & 7 \end{pmatrix}$$

est *diagonalisable* dans $\mathcal{M}_4(\mathbf{C})$ *mais elle n'est pas diagonalisable dans* $\mathcal{M}_4(\mathbb{R})$.

4 Références

Pour une lecture plus riche

1. G. GAGNAC - E. RAMIS : 1. Algèbre (MASSON)
2. F. R. GANTMACHER : Matrix Theory (Vol.I,II) (Chelsea Publ. Co. N.York)
3. R. GODEMENT : Cours d'Algèbre (Hermann)
4. S. : LANG : Algebra (Addison Wesley Publ. Co.)
5. LELONG - FERRAND - ARNAUDIER : Tome 1
6. S. LIPSCHUTZ : (Series Schaum - Mc Graw Hill N.York)
   – a) Theory and Problems of linear Algebra
   – b) Matrix theory
7. PISOT - ZAMANSKY : Mathématiques Générales
8. M. QUEYSANNE : Algèbre (Ed. Armand COLIN)
9. Toutes vos références utilisées pour l'Algèbre en 1er cycle.

Chapitre 3

Formes canoniques d'une matrice ou d'un endomorphisme

Dans ce chapitre on continue la réduction des endomorphismes et de leurs matrices associées. On s'intéressera aux endomorphismes (ou matrices) **qui ne sont pas** diagonalisables. On cherchera donc toutes les possibilités de *simplification* d'une représentation matricielle en la réduisant à l'une des formes appelées canoniques (différentes de la forme diagonale) suivantes :

- **1. Forme Triangulaire**
- **2. Décomposition en sommes directes**
- **3. Décompostion primaire**
- **4. Réduction des endomorphismes nilpotents**
- **5. Forme canonique de Jordan**
- **6. Forme rationnelle canonique**

On supposera toujours que :

a) $\mathbb{K}$=Corps Commutatif

b) E=Espace Vectoriel sur $\mathbb{K}$ de dimension finie $(dim_{\mathbb{K}}E = n)$

c) $f \in \mathcal{L}_{\mathbb{K}}(E), et\ A_f \in \mathcal{M}_n(\mathbb{K})$ matrice associée à f.

1 Forme Triangulaire

1 Forme Triangulaire simple

On a vu que l'échelonnage d'une matrice carrée, ou la méthode du pivot de Gauss (v. chapitre précédent) constituent avec leurs transformations successives sur les lignes (ou les colonnes) les algorithmes appropriés pour trouver la matrice de passage P de la base initiale à la base où la forme matricielle est triangulaire.

Par ailleurs, au chapitre précédent, on avait constaté que, si on a une matrice triangulaire $A \in \mathcal{M}_n(\mathbb{K})$ dont les éléments diagonaux sont les scalaires $(a_{11}, a_{22}, ..., a_{nn})$ dans le corps $\mathbb{K}$, alors le polynôme caractéristique $P_A(x)$ a la forme réduite (en facteurs linéaires) suivante :

$$P_A(x) = (x - a_{11})(x - a_{22}) \dots (x - a_{nn}) \tag{3.1.1}$$

Le résultat qu'on présente est la réciproque de cette propriété, et donne une condition suffisante de triangulation.

Théorème 3.1. (Triangulation)
Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ (resp $\mathcal{A}_f \in \mathcal{M}_{\mathbb{K}}(n)$).
Si le polynôme caractéristique $\mathcal{P}_f(x) = \mathcal{P}_{\mathcal{A}}(x)$ a toutes ses racines dans $\mathbb{K}$
$\Rightarrow$ $\exists$ une base dans E (resp. une matrice de passage non sigulière P) t.q. la représentation matricielle de f, $\tilde{\mathcal{A}}_f \in \mathcal{M}_{\mathbb{K}}(n)$ et matrice semblable de $\mathcal{A}_f$ avec $\tilde{\mathcal{A}}_f = \mathcal{P}^{-1}\mathcal{A}_f\mathcal{P}$, soit triangulaire.
Les éléments diagonaux de $\tilde{\mathcal{A}}_f$ sont les valeurs propres de f(resp. de $\mathcal{A}_f$ ou de $\tilde{\mathcal{A}}_f$)

Remarque 3.1. *Si* $\mathcal{P}_{\mathcal{A}}(x)$ a toutes ses racines dans $\mathbb{K}$, alors on écrira :

$$\mathcal{P}_{\mathcal{A}}(x) = (x - a_{11})(x - a_{22}) \dots (x - a_{nn}) \tag{3.1.2}$$

alors f(resp. de $\mathcal{A}_f$) sera *diagonalisable* si $a_{ii} \neq a_{jj}$, et on retrouve le résultat du théorème 2.19 du chapitre précédent.

2 Forme triangulaire par blocs

Dans le chapitre précédent, on a vu la définition des sous-espaces f-invariants de E pour tout $f \in \mathcal{L}_{\mathbb{K}}(E)$. Ces sous-espaces jouent un rôle essentiel pour la réduction de f (ou $\mathcal{A}_f$, sous forme matricielle **triangulaire** ou **diagonale par blocs**.

On présente d'abord la triangulation par blocs fournie par le théorème ci-dessous et ensuite dans le paragraphe suivant, on considérera le cas plus important de la décomposition ou *diagonalisation en sommes directes invariantes*.

Théorème 3.2. *Soient : $f \in \mathcal{L}_{\mathbb{K}}(E)$ ($\mathcal{A}_f \in \mathcal{M}_n(\mathbb{K})$), W sous-espace f-invariant de E et f_W la restriction de f à W (resp. $\mathcal{A}_{f_W}$).*
Alors f admet une représentation matricielle $\tilde{\mathcal{A}}_f$ triangulaire par blocs de la forme :

$$\tilde{\mathcal{A}}_f = \begin{pmatrix} \mathcal{A}_{f_W} & B \\ 0 & C \end{pmatrix} \tag{3.1.3}$$

Remarque 3.2. Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ et $W = \mathcal{K}er\ f(E) \neq \{0\}$ alors ;

$$\forall x \in W : \ f(x) = \{0\} \in W$$

donc $f(W) \subset W$
$\Leftrightarrow \mathcal{K}er\ f$ est un sous-espace f-invariant

Exercice (facile) :

Montrer que : $\{0\}$, E, Imf, et $\mathcal{K}er[P(f)]$ (pour un polynôme quelconque $P(x)$ sur $\mathbb{K}$), sont des sous-espaces f-invariants de E.

En utilisant le théorème précédent pour $W = Kerf$ on pourra écrire :

$$\tilde{\mathcal{A}}_f = \begin{pmatrix} 0 & B \\ 0 & C \end{pmatrix} \tag{3.1.4}$$

avec

$$rgC = rgf \tag{3.1.5}$$

2 Décomposition en somme directe

Etudions maintenant le cas où l'espace vectoriel E sur $\mathbb{K}$, peut être décomposé en somme directe de sous-espaces invariants pour un $f \in \mathcal{L}_{\mathbb{K}}(E)$. La condition nécessaire et suffisante d'une telle décomposition de E est donnée par le théorème suivant :

1 Somme directe de sous-espaces invariants

Théorème 3.3. *Soient : E espace vectoriel sur $\mathbb{K}$, $f \in \mathcal{L}_{\mathbb{K}}(E)$, $\{W_i\}_{i=1,\ldots,r}$ une famille de sous-espaces f-invariants de E, et $B = \left\{\{W_{i_1}, \ldots, W_{i_n}\}_{i=1,\ldots,r}\right\}$ l'ensemble des bases correspondantes de$\{W_i\}_{i=1,\ldots,r}$, alors*

*$E = \oplus_i W_i$ **ssi** B est une base pour E.*

Supposons que les conditions ci-dessus sont vérifiées pour E :

$E = W_1 \oplus W_2 \oplus \ldots W_r$ avec $f(W_i) \subset W_i$ et soit f_i la restriction de f à W_i alors, on dit que :

- E est *décomposable en somme directe de sous-espaces f-invariants*
- f est *décomposable en somme directe de f_i : $f = \underset{i}{\oplus} f_i$*

et on a le résultat important de *diagonalisation par blocs* de $\mathcal{A}_f$

2 Décomposition en "somme directe de matrices - blocs"

Théorème 3.4.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ ($\mathcal{A}_f \in \mathcal{M}_n(\mathbb{K})$).

Supposons que E est décomposable en somme directe de sous-espaces f-invariants : $E = \overset{r}{\underset{i=1}{\oplus}} W_i$ et $\forall\, i = 1, \ldots, r$ soit f_i la restriction de f sur W_i .

$\Rightarrow \exists$ une base de E t.q. la représentation matricielle $\tilde{\mathcal{A}}_f$ ait la forme diagonale par blocs suivante :

$$\tilde{\mathcal{A}}_f = \begin{pmatrix} \mathcal{A}_{f_1} & 0 & \ldots & 0 \\ 0 & \mathcal{A}_{f_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathcal{A}_{f_r} \end{pmatrix} \tag{3.2.6}$$

où $\mathcal{A}_{f_i}$ matrice associée à f_i

Remarque 3.3. Souvent on écrit : $\tilde{\mathcal{A}}_f = \oplus_{i=1}^{r} \mathcal{A}_{f_i}$ et on appelle $\tilde{\mathcal{A}}_f$ la somme directe des matrices $\mathcal{A}_{f_i}$.

Exemple 3.1. (Application)

Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ et supposons que $E = E_1 \oplus E_2$ avec $f(E_1) \subset E_1$, $f(E_2) \subset E_2$ et $f = f_1 \oplus f_2$ (f_i la restriction de f sur E_i, $i = 1, 2$). Alors on peut montrer que le polynôme caractéristique de f se factorise en produit :

$P_f = P_{f_1} P_{f_2}$ *(où* P_{f_i} est le polynôme caractéristique de f_i).

En effet, par application du théorème *3.4,* on a la possibilité d'avoir comme représentation matricielle de f une forme diagonale par blocs :

$$\mathcal{A}_f = \begin{pmatrix} \mathcal{A}_{f_1} & 0 \\ 0 & \mathcal{A}_{f_2} \end{pmatrix}$$

donc,

$$P_f = P_{\mathcal{A}_f} = \begin{vmatrix} xI_{E_1} - \mathcal{A}_{f_1} & 0 \\ 0 & xI_{E_2} - \mathcal{A}_{f_2} \end{vmatrix}$$

et par application de la propriété des déterminants des matrices en blocs :

$$det \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = det(A) det(B)$$

On obtient

$$P_f = det[xI_{E_1} - \mathcal{A}_{f_1}] det[xI_{E_2} - \mathcal{A}_{f_2}] = P_{f_1}.P_{f_2}$$

c.q.f.d.

3 Décomposition primaire

Il s'agit d'une conséquence du théorème 3.4 (Théorème de la décomposition en somme directe) et de la propriété de f-invariance du $KerP(f)$ pour un polynôme quelconque P sur $\mathbb{K}$. Présentons d'abord un résultat utile pour l'énoncé fondamental de la décomposition primaire.

Proposition 3.1. *Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ et soit $P(x)$ un polynôme t.q.*

$$P(x) = P_1(x) P_2(x), \quad P(f) = 0$$

avec P_1 et P_2 polynômes premiers entre eux, alors :

– *i) $E = E_1 \oplus E_2$ avec E_1, E_2 sous-espaces f-invariants et*

$$E_1 = KerP_1(f), \quad E_2 = KerP_2(f)$$

– *ii) Si $P = m_f$ (polynôme minimal de f) et si $P_1,\ P_2$ sont normalisés*

$$\Rightarrow \underline{P_1 = m_{f_1}},\quad \underline{P_2 = m_{f_2}.}$$

Autrement dit : $P_1,\ P_2$ sont les polynômes minimaux des restrictions f_1 et f_2 de f sur E_1 et E_2 respectivement

$$\Leftrightarrow \underline{m_f = m_{f_1} m_{f_2}}$$

Comme généralisation de ce théorème, on obtient le théorème de la décomposition primaire :

Théorème 3.5. *(de la décomposition primaire)*
Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ avec :

$$m_f(t) = P_1(t)^{n_1} P_2(t)^{n_2} \dots P_r(t)^{n_r}$$

où les $P_i(t)$ sont les polynômes normalisés irréductibles distincts, alors :
– *i) $E = \oplus W_i$ avec W_i sous-espace invariant de E défini par :*

$$W_i = KerP_i(f)^{n_i} \quad \forall\, i = 1, \dots, r$$

– *ii) $P_i(f)^{n_i}$ est le polynôme minimal de la restriction f_i de f à W_i*

$$\Leftrightarrow \qquad \underline{m_{f_i}(t) = P_i(t)^{n_i}}$$

<u>Conclusion</u>

D'après le théorème 3.4 et le théorème 4.1 du chapitre précédent, la représentation matricielle A_f de f se réduit à une forme diagonale en blocs :

$$\mathcal{A}_f = \begin{pmatrix} \mathcal{A}_{f_1} & 0 & 0 & 0 \\ 0 & \mathcal{A}_{f_2} & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mathcal{A}_{f_r} \end{pmatrix}$$

où $\mathcal{A}_{f_i}$ est la matrice associée à la restriction f_i de f sur chaque sous-espace f-invariant W_i défini $\forall\, i = 1, 2 \dots r$ par :

$$W_i = KerP_i(f)^{n_i}$$

Avec le corollaire suivant on trouve une autre forme de la diagonalisabilité d'un endomorphisme f (ou A_f).On obtient ce résultat par le théorème 3.3 en posant $degP_i = 1 \quad \forall i$ et $n_i = 1$.

Corollaire 3.1. *Soient*

$$f \in \mathcal{L}_{\mathbb{K}}(E) \quad (\mathcal{A}_f \in \mathcal{M}_n(K))$$

$\Rightarrow$

f (ou A_f) <u>est diagonalisable ssi</u> son polynôme minimal m_f est un produit de polynômes <u>linéaires distincts</u>.

Exemple 3.2. *Soit $A_f \neq I$ et telle que $A_f^3 = I$.*

On étudie la diagonalisabilité de A_f *On a :*

$$A_f^3 - I = 0$$

donc A_f est une racine du polynôme suivant :

$$P(t) = t^3 - 1 = (t-1)(t^2+t+1)$$

Donc le polynôme minimal m_f doit être de la forme :
$m_f^{(a)}(t) = t^2 + t + 1$ *ou* $m_f^{(b)}(t) = t^3 - 1$
(Il est impossible d'avoir $m_f^{(c)}(t) = t - 1$ car par hypothèse $A \neq I$)

- *a)* Dans $\mathbb{R}$ aucun de ces polynômes n'est linéaire $\Rightarrow \mathcal{A}_f$ n'est pas diagonalisable.
- *b)* Dans $\mathbf{C}$ chacun de ces polynômes est produit de polynômes linéaires $\Rightarrow A_f$ *est diagonalisable sur* $\mathbf{C}$.

4 Réduction des endomorphismes nilpotents

1 Endomorphisme nilpotent

Définition 3.1.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ (resp. $\mathcal{A}_f \in \mathcal{M}_n(\mathbb{K})$).

- *a) On dit que f est un endomorphisme nilpotent (resp. $\mathcal{A}_f$ matrice nilpotente) si $f^n = 0$ pour un n donné entier positif.*
- *b) L'entier $k \geq 1$ est appelé indice de nilpotence de f (ou de $\mathcal{A}_f$) endomorphisme nilpotent, si il est le plus petit des entiers tel que $f^k = 0$ et $f^{k-1} \neq 0$.*

D'après la définition précédente $\lambda = 0$ est la seule valeur propre d'un endomorphisme nilpotent ; on en déduit que le polynôme minimal $m_f(t)$ (resp. $m_{\mathcal{A}_f}(t)$) a la forme :

$$m_f(t) = t^k = m_{\mathcal{A}_f}(t)$$

Comme conséquence immédiate on a le :

Théorème 3.6. *Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ endomorphisme nilpotent d'indice k*

$\Rightarrow$ il existe une représentation matricielle $\tilde{\mathcal{A}}_f$ canonique sous forme matricielle diagonale en blocs :

$$\tilde{\mathcal{A}}_f = \begin{pmatrix} N_1 & 0 & \dots & 0 \\ 0 & N_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & N_r \end{pmatrix} \textit{ avec } N_i = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix} \tag{3.4.7}$$

- *Les éléments matrices carrées N_i d'ordre r_i ont tous leurs éléments zéro, sauf ceux qui sont au-dessus de la diagonale principale et qui sont égaux à* 1.
- *Il existe au moins une matrice N_i d'ordre k, les autres sont d'ordre $\leq k$.*
- *Le nombre des matrices N_i dépend uniquement de l'endomorphisme f.*
- *Ce nombre total égale précisément la nullité (dimension du noyau) de f :*

$$Card\{N_i\} = dimKerf$$

Remarquons que :

- a) Chacune des matrices N_i est une matrice nilpotente ayant commme indice de nilpotence son ordre r_i.
- b) La matrice nilpotente d'ordre 1 est la matrice 1×1 qui est $[\mathcal{O}_{(1)}]$

.

Exemple 3.3.

- *a) Soit $f \in \mathcal{L}_{\mathbf{R}}(\mathbb{R}^5)$ t.q.*

$$\mathcal{A}_f = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \text{ on a } \mathcal{A}_f^2 = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \text{ et } \mathcal{A}_f^3 = [\mathcal{O}_{(5)}]$$

  donc f (ou $\mathcal{A}_f$) est un endomorphisme nilpotent d'indice $\underline{k=3}$ donc la représentation matricielle $\tilde{\mathcal{A}}_f$ (semblable à $\mathcal{A}_f$) en bloc-diagonale contient au moins un bloc N_1 d'ordre 3 et aucun d'ordre supérieur à 3.
  De plus, $rg\mathcal{A}_f = 2 \Rightarrow dimKer\mathcal{A}_f = 5 - 2 = \underline{3}$

$$\tilde{\mathcal{A}}_f = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

  Donc au total, on a 3 blocs diagonaux :
  1 d'ordre 3 et deux d'ordre 1 qui sont les matrices nilpotentes $[\mathcal{O}_{(1)}]$.

- *b)* Soit $f \in \mathcal{L}_{\mathbb{R}}(\mathbb{R}^5)$ t.q.

$$\mathcal{B}_f = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

  on a $\mathcal{B}_f^2 = [\mathcal{O}_{(5)}]$
  $\Rightarrow f$ (ou $\mathcal{B}_f$) est nilpotent ayant comme indice de nilpotence $\underline{k=2}$
  De plus,

$$rg\mathcal{B}_f = 1 \Rightarrow dimKer\mathcal{B}_f = 5 - 1 = \underline{4}$$

$$\tilde{\mathcal{B}}_f = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Il en résulte qu'on peut avoir une matrice $\tilde{\mathcal{B}}_f$ (semblable à $\mathcal{B}_f$) diagonale en blocs d'un nombre égal à 4 dont un est d'ordre 2 et les 3 autres sont d'ordre 1, c'est à dire $[\mathcal{O}_{(1)}]$.

5 Forme canonique de Jordan

1 Factorisation du polyn.caractéristique et minimal

La forme canonique en blocs diagonale de Jordan est réalisable chaque fois qu'on peut écrire les polynômes, minimal m_f et caractéristique P_f d'un endomorphisme $f \in \mathcal{L}_{\mathbb{K}}(E)$ sous forme de produits en facteurs linéaires. Plus précisément :

Théorème 3.7.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ (resp. $\mathcal{A}_f \in \mathcal{M}_n(K)$) t.q. les polynômes caractéristique et minimal associés ont la forme :

$$P_f(t) = \prod_{i=1}^{r} (t - \lambda_i)^{n_i} \tag{3.5.8}$$

et

$$m_f(t) = \prod_{i=1}^{r} (t - \lambda_i)^{m_i} \tag{3.5.9}$$

$\Rightarrow$ f admet (resp. $\mathcal{A}_f$ est semblable à $\tilde{\mathcal{A}}_f = P^{-1}\mathcal{A}_f P$) une repr. matricielle $\tilde{\mathcal{A}}_f$ ***diagonale en blocs :***

$$\tilde{\mathcal{A}}_f = \begin{pmatrix} J_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & J_2 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & J_i & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 0 & 0 & J_n \end{pmatrix}$$

Les blocs (matrices carrées) J_i ont la forme suivante :

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 1 \\ 0 & 0 & 0 & 0 & \lambda_i \end{pmatrix} \Leftrightarrow \underline{J_i = \lambda_i I_{\tilde{n}_i} + N_{\tilde{n}_i}}$$

où $N_{\tilde{n}_i}$ matrice nilpotente d'ordre $\tilde{n}_i$.

Les blocs J_i vérifient les propriétés suivantes :

- *i) Il y a au moins une matrice J_i d'ordre m_i, toutes les autres J_i sont d'ordre $\leq m_i$.*
- *ii) La somme des ordres des J_i égale n_i.*
- *iii) Le nombre des J_i est égale à l'ordre de multiplicité géométrique ($dim V(\lambda_i)$) des λ_i.*
- *iv) Le nombre des J_i de chaque ordre possible est uniquement déterminé par f.*

Remarque 3.4. Si $f \in \mathcal{L}_{\mathbb{K}}(E) \quad (\mathcal{A}_f \in \mathcal{M}_n(\mathbb{K}))$ est tel que $\forall\, i, n_i = 1$ (ou $\forall\, i \ m_i = 1$), alors on retrouve le corollaire de "diagonalisabilité" 3.1 car $J_i = \lambda_i$.

Exemple 3.4.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ t.q.

$$P_f(t) = (t-2)^4(t-3)^3$$

et

$$m(t) = (t-2)^2(t-3)^2$$

Alors, la représentation matricielle de f peut avoir deux formes canoniques de Jordan possibles :

$$\tilde{\mathcal{A}}_f^{(1)} = \begin{pmatrix} 2&1&0&0&0&0&0\\ 0&2&0&0&0&0&0\\ 0&0&2&1&0&0&0\\ 0&0&0&2&0&0&0\\ 0&0&0&0&3&1&0\\ 0&0&0&0&0&3&0\\ 0&0&0&0&0&0&3 \end{pmatrix} \text{ ou } \tilde{\mathcal{A}}_f^{(2)} = \begin{pmatrix} 2&1&0&0&0&0&0\\ 0&2&0&0&0&0&0\\ 0&0&2&0&0&0&0\\ 0&0&0&2&0&0&0\\ 0&0&0&0&3&1&0\\ 0&0&0&0&0&3&0\\ 0&0&0&0&0&0&3 \end{pmatrix}$$

La première forme $\tilde{\mathcal{A}}_f^{(1)}$ est obtenue dans le cas où $\underline{dim V(2) = 2}$. La deuxième forme est obtenue dans le cas où $\underline{dim V(2) = 3}$ (et $dim V(3) = 2$).

6 Forme rationnelle canonique

1

Définition 3.2. *a) Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ et soit $\{P_{\mathbb{K}}(t)\}$ l'ensemble de tous les polynômes de K ; on appelle sous-espace <u>f-cyclique de E engendré par le vecteur $v \in E$</u>, l'ensemble de vecteurs $\underline{\{P_{\mathbb{K}}(f)\}(v)}$. C'est un <u>sous-espace f-invariant de E</u>, et on le note : $\underline{Z(v, f)}$.*

On notera par f_v la restriction de f à $\underline{Z(v,f)}$.

b) Soit : $\underline{m_v(t) = t^k + a_{k-1}t^{k-1} + \cdots + a_1 t + a_0}$

Le polynôme unique normalisé, de plus bas degré t.q. $\underline{m_v(f)(v) = 0}$. On appelera <u>f-annihilateur de</u> $\underline{v}$ <u>et</u> $\underline{Z(v,f)}$.

Théorème 3.8.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E), Z(v,f)$ le sous-espace f-cyclique invariant engendré par $v \in E$, f_v la restriction de f à $Z(v,f)$, alors si $m_v(t)$ est le f-annihilateur de v, on a :

i) L'ensemble $\mathcal{B}_v = \{v, f(v), \ldots, f^{k-1}(v)\}$ (où k = degré de $m_v(t)$) est une <u>base</u> de $Z(v,f)$ $\underline{dim_{\mathbb{K}}(Z(v,f)) = k}$.

ii) Le polynôme <u>minimal</u> de f_v est $m_v(t)(t \in \mathbb{K})$

iii) La représentation matricielle de f_v dans la base $\mathcal{B}_v$ a la forme suivante :

$$\tilde{\mathcal{C}}_{f_v} = \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ddots & 0 & -a_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \ldots & 0 & 1 & -a_{k-1} \end{pmatrix} \tilde{C} \equiv \text{ matrice compagnon de } m_v(t) \text{(d'ordre } k)$$

La grande utilité de la réduction en forme canonique rationnelle, présentée par le théorème suivant, vient du fait qu'on peut l'appliquer même dans les cas où le polynôme minimal ne se factorise pas en polynômes linéaires.

Théorème 3.9. ***Forme rationnelle canonique***

i) Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$, et soit $\underline{m_f(t) = \big(P(t)\big)^n}t \in K$ (polynôme minimal de f) où $P(t)$ est un polynôme irréductible normalisé.

$$\Rightarrow E = \bigoplus_{i=1}^{r} Z(v_i, f)$$

et les sous-espaces f-invariants cycliques $Z(v_i, f)$ ont comme annihilateurs les polynômes :

$$\{P(t)^{n_1}, P(t)^{n_2}, \ldots, P(t)^{n_r} \text{ où } n = n_1 \geq n_2 \geq \cdots \geq n_r\}$$

qui sont uniquement déterminés par f. En plus, f admet une représentation matricielle en blocs diagonale unique :

$$\tilde{\mathcal{A}}_f = \begin{pmatrix} \tilde{\mathcal{C}}_{f_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \tilde{\mathcal{C}}_{f_r} \end{pmatrix}$$

où les $\mathcal{C}_{f_i}$ sont les matrices compagnons des polynômes $[P(t)]^{n_i}$.

ii) Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$; si le polynôme minimal de f est :

$$\underline{m_f(t) = [P_1(t)]^{m_1}[P_2(t)]^{m_2} \ldots [P_s(t)]^{m_s}} t \in \mathbb{K}$$

où les $P_i(t)$ sont des polynômes irréductibles distincts (normalisés)

$\Rightarrow$ *f admet une représentation matricielle <u>unique</u> en blocs diagonale de la forme :*

$$\tilde{\mathcal{A}}_f = \begin{pmatrix} \mathcal{C}_{11} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathcal{C}_{1r_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathcal{C}_{s_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathcal{C}_{sr_s} \end{pmatrix}$$

Les $\mathcal{C}_{ij}$ sont les matrices compagnons des polynômes $\underline{P_1(t)^{n_{ij}}}$ où

$$\begin{pmatrix} m_1 = n_{11} \geq n_{12} \geq \cdots \geq n_{1r_1} \\ \vdots \\ m_s = n_{s1} \geq n_{s2} \geq \cdots \geq n_{sr_s} \end{pmatrix}$$

$\forall\, i, j$, on appelle $P_i(t)^{n_{ij}}$ ***les diviseurs élémentaires de*** *f.*

Exemple 3.5. Soit $f \in \mathcal{L}_{\mathbb{R}}(\mathbb{R}^6)$ et soit $m_f(t) = (t^2 - t + 3)(t-2)^2$ son polynôme minimal.

D'après le théorème précédent, si on écrit :

$$m_f(t) = [P_1(t)]^{m_1}[P_2(t)]^{m_2} \textit{avec} \begin{Bmatrix} P_1(t) = t^2 - t + 3 \\ m_1 = 1 \end{Bmatrix} \textit{ et } \begin{Bmatrix} P_2(t) = (t-2) \\ m_2 = 2 \end{Bmatrix}$$

$$\Rightarrow$$

Les diviseurs élémentaires de f ne peuvent être qu'une des suites de polynômes suivants :

(a) : $\left\{ \big(P_1(t)\big)^1, \big(P_1(t)\big)^1, \big(P_2(t)\big)^2 \right\}$

(b) : $\left\{ \big(P_1(t)\big)^1, \big(P_2(t)\big)^2, \big(P_2(t)\big)^2 \right\}$

(c) : $\left\{ \big(P_1(t)\big)^1, \big(P_2(t)\big)^2, \big(P_2(t)\big), \big(P_2(t)\big) \right\}$

avec les $\mathcal{C}_{ij}$ matrices compagnons correspondantes :

$$t^2 - t + 3 = \big(P_1(t)\big)^1 \Rightarrow \mathcal{C}_1 \equiv \mathcal{C}[t^2 - t + 3] = \begin{pmatrix} 0 & -3 \\ 1 & 1 \end{pmatrix}$$

$$t^2 - 4t + 4 - \big(P_2(t)\big)^2 \Rightarrow \mathcal{C}_2 \equiv \mathcal{C}[t^2 - 4t + 4] = \begin{pmatrix} 0 & -4 \\ 1 & 4 \end{pmatrix}$$

$$t - 2 = \big(P_2(t)\big)^1 \Rightarrow \mathcal{C}_{22} \equiv [t_2] = (2)$$

Donc f admettra une représentation matricielle ayant une des trois formes suivantes (diagonales en blocs) correspondantes :

$$\textit{(a) :} \begin{pmatrix} 0 & -3 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix} \textit{(b) :} \begin{pmatrix} 0 & -3 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 1 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix} \textit{(c) :} \begin{pmatrix} 0 & -3 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 1 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

Chapitre 4

formes linéaires - formes bilinéaires

1 Formes linéaires

1 Rappel de notations

- $\mathbb{K}$= corps commutatif.
- $E,\ F$ = espaces vectoriels sur $\mathbb{K}$ ($\dim_{\mathbb{K}} E = n\ ;\ \dim_{\mathbb{K}} F = m$) .
- $\mathcal{L}_{\mathbb{K}}(E, F)$ = esp. vectoriel des applications linéaires $f\ : E \to F$ ($\dim_{\mathbb{K}} \mathcal{L}_{\mathbb{K}}(E, F) = nm$)
- $\mathcal{M}_{\mathbb{K}}(m, n)$ = Esp. vectoriel des matrices (de $\dim nm$ sur $\mathbb{K}$) isomorphe à $\mathcal{L}_{\mathbb{K}}(E, F)$
- $\mathcal{L}_{\mathbb{K}}(E)$ = esp. vectoriel des endomorphismes sur E ($\dim \mathcal{L}_{\mathbb{K}}(E) = n^2$)
- $\mathcal{M}_n(\mathbb{K})$ = esp. vectoriel des matrices carrées de dimension n^2 isomorphe à $\mathcal{L}_{\mathbb{K}}(E)$.

2 Formes (ou fonctionnelles) linéaires

Définition 4.1.

*On appelle **forme linéaire** ou **fonctionnelle linéaire** toute application linéaire $\phi \in \mathcal{L}_{\mathbb{K}}(E, \mathbb{K})$ définie sur E,*

$$\phi : E \to \mathbb{K}$$

.

Remarque 4.1. Les propriétés de linéarité qui caractérisent les applications linéaires de $\mathcal{L}_{\mathbb{K}}(E, F)$ sont par conséquent vraies pour tout $\phi \in \mathcal{L}_{\mathbb{K}}(E, \mathbb{K})$ (où simplement $F = \mathbb{K}$) :

$$\forall (x, y) \in E^2, \quad \forall\, \lambda \in \mathbb{K}, \quad \begin{aligned} \phi(x + y) &= \phi(x) + \phi(y) \\ \phi(\lambda x) &= \lambda\phi(x) \end{aligned}$$

3 Dualité

L'espace vectoriel de toutes les formes linéaires sur E est un espace particulièrement intéressant et on l'appelle le **dual** de E noté E^* :

Définition 4.2.

a) Le dual de E est $E^* = \mathcal{L}_{\mathbb{K}}(E, K)$; c'est un espace vectoriel sur $\mathbb{K}$ et

$$\dim_{\mathbb{K}} E^* = \dim_{\mathbb{K}} \mathcal{L}_{\mathbb{K}}(E, \mathbb{K}) = n \times 1 = \dim_{\mathbb{K}} E.$$

b) Le bidual de E est : $E^{**} = \mathcal{L}_{\mathbb{K}}(E^*, \mathbb{K})$, c'est-à-dire :

$$y \in E^{**} \Leftrightarrow y : \mathcal{L}_{\mathbb{K}}(E, \mathbb{K}) \to \mathbb{K}$$

Remarque 4.2.

a) On vérifie que $\underline{E^{**} \text{ et } E}$ sont isomorphes. Par la suite, on identifiera le bidual de E avec E lui même : $E^{**} \equiv E$

b) Dans diverses références on utilise souvent la notation x^* (ou y^*) pour un élément de E sans qu'il y ait une relation quelconque avec le vecteur $x \in E$ (ou $y \in E$) et la forme linéaire x^* (ou y^*).

Exemple 4.1.

a) Soit $\{a_1, \cdots, a_n\}$ une base de $\mathbb{K}^n$. Toute combinaison linéaire : $\phi(x) = \sum_{i=1}^{n} a_i\, x_i$ qui exprime le vecteur $x \in \mathbb{K}^n$ en terme des vecteurs de la base $\mathbb{K}^n$ est une fonctionnelle (ou forme) linéaire ; historiquement, c'est la <u>première forme</u> à laquelle on a attribué le nom de la forme linéaire.

b) Soit $\{a_i\}$ une base de E et $x \in E$: $\underline{x = \sum_{i=1}^{n} \alpha^{(i)} a_i}$: alors l'application

$$\phi^{(i)} : \begin{array}{c} E \to \mathbb{K} \\ x \mapsto \phi^{(i)}(x) = \alpha^{(i)} \end{array}$$

(où $\alpha^{(i)}$ *est la* $(i)^{\textit{ème}}$ coordonnée de $x \in E$ par rapport à la base $\{a_i\}$) est une fonctionnelle linéaire sur E, autrement dit $\phi^{(i)} \in E^*$ ($\phi^{(i)}$ appartient donc à l'espace dual $\mathcal{L}_{\mathbb{K}}(E, K)$ de E).

Cette forme linéaire est appelée souvent : la $(i)^{\textbf{\textit{ème}}}$ ***projection de*** E ***sur l'axe*** a_i ***et on note***

$$\phi^{(i)} \equiv \pi^{(i)}.$$

c) Soit $E = \mathcal{C}[0, 1]$ l'espace vectoriel des **fonctions numériques continues** $y(t)$ **sur [0,1]**. Alors l'application :

$$\begin{array}{c} \phi : \mathcal{C}[0, 1] \to \mathbb{R} \\ \textit{avec } y \mapsto \int_0^1 y(t)dt \equiv \phi(y) \end{array}$$

est une forme linéaire sur $\mathcal{C}[0, 1]$.

d) Soit $\{a_{ij}\}_{\substack{i=1,\cdots,n \\ j=1,\cdots,n}} = A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée, alors la trace

$$Tr[A] = \sum_{i=1}^{n} a_{ii}$$

est l'image de A par une fonctionnelle linéaire,
$\mathcal{T}r : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$
Autrement dit, **la trace d'une matrice carrée est un élément du dual** $\mathcal{M}_n^*(\mathbb{K})$ **de l'espace vectoriel des matrices carrées.**

Remarque 4.3. ***Représentation matricielle d'une*** $\phi \in E^*$

D'après le théorème général de l'isomorphisme (v. cours n^0 1) :

$$\mathcal{L}_{\mathbb{K}}(E,F) \sim \mathcal{M}_{\mathbb{K}}(n,m) \text{ on a aussi } \underline{\mathcal{L}_{\mathbb{K}}(E,\mathbb{K}) \sim \mathcal{M}_{\mathbb{K}}(n,1)}$$

c'est-à-dire chaque forme linéaire $\phi \in E^*$, admet une représentation matricielle sous forme de **matrice-ligne**, qui est la transposée de la matrice colonne correspondante, et qui est uniquement déterminée par les images par ϕ de tous les vecteurs de la base de E : $\phi = (a_1, a_2, \cdots, a_n)$ (vecteur ligne).

Donc si $x = (x_1, \cdots, x_n)^T \in E$ alors : $\phi(x) = (a_1, a_2, \cdots, a_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} a_i \, x_i$.

Dans l'espace vectoriel E^* on peut définir une (ou plusieurs) base qui s'appellera base duale. Le théorème suivant établit une **construction canonique d'une base duale** à partir d'une base donnée de E.

Théorème 4.1. *Base duale*

Soit $\{a_i\}_{i=1,\cdots,n}$ une base de E sur $\mathbb{K}$. Alors on définit une base de E^* appelée ***base duale*** *de $\{a_i\}$:*

$$\textit{par la famille suivante } \phi_i(a_j) = \delta_{ij} = \begin{cases} 1 & \textit{si } i = j \\ 0 & \textit{si } i \neq j \end{cases}$$

Exemple 4.2.

Soit $\mathcal{B} = \{\nu_1, \nu_2\}$ une base de $\mathbb{R}^2$ définie par : $\mathcal{B} = \{\nu_1 = (2,1), \nu_2 = (3,1)\}$. Construction de la base $\mathcal{B}^*$ de $\mathbb{R}^{2*}$[1]

D'après le théor. 1.1.1. il faut

$$\underline{\phi_1(\nu_1) = 1,\ \phi_1(\nu_2) = 0, \phi_2(\nu_1) = 0,\ \phi_2(\nu_2) = 1} \quad (\mathcal{B}^*.0)$$

On pose : $\phi_1(x,y) = ax + by,\ \phi_2(x,y) = cx + dy$ avec $(a,b,c,d) \in \mathbb{R}^4$.

Pour déterminer les $a,\ b,\ c,\ d$, on établit les deux systèmes satisfaisant les conditions $(\mathcal{B}^*.0)$ de la base duale associée à $\mathcal{B}$:

$$\left.\begin{array}{lcl} \phi_1(2,1) = 1 & \Leftrightarrow & 2a + b = 1 \\ \phi_1(3,1) = 0 & \Leftrightarrow & 3a + b = 0 \end{array}\right\} \Rightarrow \quad a = -1, \quad b = 3\}$$

$$\left.\begin{array}{lcl} \phi_2(2,1) = 0 & \Leftrightarrow & 2c + d = 0 \\ \phi_2(3,1) = 1 & \Leftrightarrow & 3c + d = 1 \end{array}\right\} \Rightarrow \quad c = 1, \qquad d = -2\}$$

$\Rightarrow$ *Base duale de* $\mathcal{B} = \{\nu_1, \nu_2\}$*Ê*

$\mathcal{B}^* = \{\phi_1(x,y) = -x + 3y \ ; \ \phi_2(x,y) = x - 2y\}$

4 Annihilateur

Définition 4.3.

a) Soit $E^* = \mathcal{L}_{\mathbb{K}}(E,K)$ et soit $W \subset E$ (sous-ensemble de E et pas forcément sous-espace de E).

On dit que $\phi \in E^*$ est ***un annihilateur*** *de W si $\forall w \in W$ on a $\phi(w) = \{0\}$.*

[1] Attention : à ne jamais confondre la notation * du dual (example $\mathbb{R}^{2*}$ avec le symbole * qu'on utilise souvent pour indiquer le complémentaire de $\{0\}$ (example $\mathbb{R}^*_+$. On précisera par la suite dans lequel des deux cas on situe l'étude.

b) L' ensemble de tous les annihilateurs de W est un sous-espace de E^* appelé l' ***(espace) annihilateur de*** W *et noté* W^0.

c) D' une manière analogue on définit l'annihilateur de l'annihilateur, noté W^{00}, *par :*
$W^{00} = \{x \in E \ : \phi(x) = 0 \quad \forall\, \phi \in W^0\}$.

On a le résultat suivant :

Théorème 4.2.
Si $\dim_{\mathbb{K}} E = n < \infty$ *et si* W *est un sous-espace de* E *alors :*

$$i) \quad W^{00} = W \qquad\qquad ii) \quad \dim_{\mathbb{K}} W + \dim_{\mathbb{K}} W^0 = \dim_{\mathbb{K}} E$$

Remarque 4.4.
a) A ne pas confondre : pour $f \in \mathcal{L}_{\mathbb{K}}(E)$. $\mathcal{K}er\ f$ et W^0 (avec $W = Im\ f$)

$$\textit{car } \mathcal{K}er\ f \in E \textit{ et } W^0 \subset E^*$$

b) Il y a des références qui appellent W^0 l'orthogonal de W.

Exemple 4.3.
Soit W le sous-espace de $\mathbb{R}^4$ engendré par $\nu_1(1, 2, -3, 4)$ et $\nu_2(0, 1, 4, -1)$. Cherchons une base de l'annihilateur de W. On vérifie facilement que si $\phi \in \mathbb{R}^{4*}$ annule ν_1 et ν_2 (vecteurs de base de W). Alors il annule toute combinaison linéaire de ν_1 et ν_2. Par conséquent, on cherchera une base de l'ensemble des formes linéaires :

$$\phi(x, y, z, w) = ax + bycz + dw \quad \textit{t.q.}$$
$$\phi(\nu_1) = 0 \textit{ et } \phi(\nu_2) = 0$$

$$\Leftrightarrow \begin{cases} \phi(1, 2, -3, 4) &= a + 2b - 3c + 4d = 0 \\ \phi(0, 1, 4, -1) &= b + 4c - d = 0 \end{cases}$$

On obtient donc un système échelonné par rapport à $a,\ b,\ c,\ d$.

On a 2 paramètres libres alors :

1) si on pose $c = 1,\ d = 0$
on obtient la solution $\left| \begin{array}{ll} a = 11, & b = -4 \\ c = 1, & d = 0 \end{array} \right.$
donc $\phi_1(x, y, z, w) = 11x - 4y + z$

2) si on pose $c = 1,\ d = -1$
on obtient la solution $a = 6,\ b = -1,\ c = 0,\ d = -1$
donc $\phi_2(x, y, z, w) = 6x - y - z$

Conclusion :
L'ensemble $\{\phi_1 = 11x - 4y + z,\ \phi_2 = 6x - y - w\}$ est une base de l'annihilateur W^0 de W.

5 Transposée d'une application linéaire

Définition 4.4.

Soit $f\ : E \to F(f \in \mathcal{L}_{\mathbb{K}}(E,F))$. Pour tout $\phi \in F^*$ (forme linéaire sur F), on définit l'application linéaire composée :
$f^t = \phi \circ f$ appelée la **transposée** $f^t(\phi)$ de f. Autrement dit,

$$f^t\ :\ E \xrightarrow{f} F \xrightarrow{\phi} \mathbb{K} \text{ et } f^t \in E^*$$

$$\phi \circ f = f^t(\phi)$$

Remarque 4.5.

a) D'après la définition 4.4 on a $\phi \in F^* \mapsto f^t(\phi) \in E^*$.
Autrement dit, l'application f^t (transposée de f) est une application de F^* (dual de F) dans E^* (dual de E) et on vérifie facilement que f^t est une application linéaire :

$$f^t \in \mathcal{L}_{\mathbb{K}}(F^*, E^*)$$

b) On vérifie facilement que : si f_1, f_2 sont 2 applications linéaires

$$\Rightarrow (f_1 \circ f_2)^t = f_2^t \circ f_1^t.$$

Tenant compte des isomorphismes :

$$\begin{cases} \mathcal{L}_{\mathbb{K}}(E,F) \sim \mathcal{M}_{\mathbb{K}}(n,m) \\ \mathcal{L}_{\mathbb{K}}(F^*,E^*) \sim \mathcal{M}_{\mathbb{K}}(m,n) \end{cases}$$

On obtient le résultat suivant concernant la représentation matricielle de f^t.

Théorème 4.3.

Soit $\{a_i\}_{i=1,\cdots,n}$ une base de E ($\dim_{\mathbb{K}} E = n$) et $\{b_j\}_{j=1,\cdots,m}$ une base de F ($\dim_{\mathbb{K}} F = m$). Si $f \in \mathcal{L}_{\mathbb{K}}(E,F)$, avec $A_f \in \mathcal{M}_{\mathbb{K}}(n,m)$, une matrice associée alors :

La représentation matricielle de f^t est égale à la transposée de A_f :

$$A_{f^t} = A_f^T$$

Exemple 4.4. ***(IMPORTANT ! ! ! Application - corollaire de plusieurs définitions et théorèmes)***

Soient $E,\ F$ 2 espaces vectoriels avec $\dim_{\mathbb{K}} E = n$ et $\dim_{\mathbb{K}} F = m$ et soit $f \in \mathcal{L}_{\mathbb{K}}(E,\ F)$. On montre que

$$rg\ f = rg\ f^t.$$

a) On montre d'abord que $\mathcal{K}er\ f^t = (\text{Im}\ f)^0$. En effet,

<u>a.1.</u> Soit $\phi \in \mathcal{K}er\ f^t \subset F^* \Leftrightarrow f^t(\phi) = \phi \circ f = 0$
et soit $\underline{u \in Im\ f} \Leftrightarrow \exists\, \nu \in E$ t.q. $f(\nu) = u \in F$
$\Rightarrow \phi(u) = \phi(f(\nu)) = (\phi \circ f)(\nu) = 0$
et $\phi(u) = 0\ \forall u \in Im\ f \Rightarrow \phi \in (Im\ f)^0 \Rightarrow \mathcal{K}er\ f^t \subset (Im\ f)^0$ (a.1.)

<u>a.2.</u> Soit $\sigma \in (Im\ f)^0 \subset F^* \Leftrightarrow \sigma(Im\ f) = \{0\}$
$\Rightarrow \forall\, \nu \in E\ f^t(\sigma)(\nu) = (\sigma \circ f)(\nu) = \sigma(f(\nu)) = 0$
$\Rightarrow f^t(\sigma) = 0 \Rightarrow\ \ \sigma \in \mathcal{K}er\ f^t \Rightarrow (Im\ f)^0 \subset \mathcal{K}er\ f^t.$ (a.2.)

Des deux conclusions (a.1.) et (a.2.) $\Rightarrow \mathcal{K}er\ f^t = (Im\ f)^0$ *c.q.f.d.*

b) Par application du th. 4.2 pour l'annihilateur $(Im\ f)^0$ on a :

$$\dim_{\mathbb{K}}(Im\ f)^0 = \dim_{\mathbb{K}} F - \dim_{\mathbb{K}}(Im\ f) = m - rg\ f \qquad \text{(b.1.)}$$

En utilisant le théorème du rang (v. cours n^0 1) à l'application

$$f^t\ : F^* \to E^* \text{ on a : } rg\ f^t = \dim_K F^* - \dim_{\mathbb{K}} \mathcal{K}er\ f^t = m - \dim_{\mathbb{K}}\ \mathcal{K}er\ f^t$$

En utilisant le résultat

$$a) \Rightarrow rg\ f^t = m - \dim_{\mathbb{K}}(Im\ f)^0 \qquad \text{(b.2.)}$$

Application de (b.1.) au second membre de (b.2.)

$$\Rightarrow\ \ rg\ f^t = m - m + rg\ f = rg\ f \qquad \textit{c.q.f.d.}$$

Exemple 4.5.

Soit ϕ la forme linéaire de $\mathbb{R}^2$ définie par $\phi(x,y) = 3x - 2y$.
On cherche les images de ϕ par les transposées $f_1^t,\ f_2^t$ des applications linéaires $f_1,\ f_2$: $\mathbb{R}^3 \to \mathbb{R}^2$

$$\text{déf. par : } \begin{cases} f_1(x,y,z) &= (x+y, y+z) \\ f_2(x,y,z) &= (x+y+z, 2x-y) \end{cases}$$

On aura :

$$(\phi \circ f_1)(x,y,z) = \phi(f_1(x,y,z)) = \phi(x+y, y+z) = 3(x+y) - 2(y+z)$$

$$\Rightarrow \quad f_1^t(\phi) = 3x + y - 2z$$

et

$$(\phi \circ f_2)(x,y,z) = \phi(f_2(x,y,z)) = \phi(x+y+z, 2x-y)$$
$$= 3(x+y+z) - 2(2x-y) = -x + 5y + 3z$$
$$\Rightarrow \quad f_2^t(\phi) = -x + 5y + 3z$$

Terminons cette partie du chapitre par le théorème de changement de base dans E^* en utilisant l'isomorphisme $\mathcal{L}_{\mathbb{K}}(E^*) \sim \mathcal{M}_n(\mathbb{K})$.

Théorème 4.4.

*Soit $\{a_i\}_{i=1,\cdots,n}$ $\{b_j\}_{j=1,\cdots,n}$ deux bases de E et $\{\phi_i\}_{i=1,\cdots,-n}$ $\{\sigma_j\}_{j=1,\cdots,-n}$, leurs bases duales associées de E^*. Alors si l'application, qui transforme $\{a_i\}_{i=1,\cdots,n}$ en $\{b_j\}_{j=1,\cdots,n}$, a comme représentation matricielle **P** (matrice de passage), alors **l'application transposée** (qui transforme $\{\phi_i\}_{i=1,\cdots,n}$ en $\{\sigma_j\}_{j=1,\cdots,-n}$ a comme **représentation matricielle** $(P^{-1})^t$ (**matrice de passage** dans E^*).*

2 Formes Bilinéaires

1 forme bilinéaire

Définition 4.5.

Soient $E,\ F$ espaces vectoriels sur $\mathbb{K}$. On appelle ***forme bilinéaire*** *sur $\underline{E \times F}$.*
Toute application

$\dfrac{\phi\ :\ E \times F \to \mathbb{K}}{(u,\nu) \mapsto \phi(u,\nu)}$ t.q. $\forall\, u, u_i \in E,\ \nu, \nu_i \in F \quad i = 1,2$ *et* $\forall\,(a,b) \in \mathbb{K}^2$ on a :

$$\left|\begin{array}{lcl} i) \quad \phi(au_1 + bu_2, \nu) & = & a\,\phi(u_1,\nu) + b\,\phi(u_2,\nu) \\ ii) \quad \phi(u, a\nu_1 + b\nu_2) & = & a\,\phi(u,\nu_1) + b\,\phi(u,\nu_2) \end{array}\right.$$

Autrement dit :

ϕ ***est linéaire*** *par rapport à la première variable (vecteur)*

ϕ ***est linéaire*** *par rapport à la deuxième variable (vecteur)*

Remarque 4.6.

* Cas particulier : quand $E = F$ la forme bilinéaire $\phi\ :\ E \times E \to \mathbb{K}$ est définie par i) et ii) sur E^2

Exemple 4.6.

Soit $\{a_{ij}\}_{\substack{i=1,\cdots,n \\ j=1,\cdots,m}} = A \in \mathcal{M}_n(\mathbb{K})$ et E espace vectoriel sur $\mathbb{K}$, alors $\forall\,(x,y) \in E^2$ on associe à A le polynôme :

$$f(x,y) = \sum_{i,j=1}^{n} a_{ij}\ x_i y_j = a_{11}x_1y_1 + a_{12}x_1y_2 + \cdots + a_{nn}x_ny_n,$$

qui est une forme bilinéaire sur E^2. On écrit aussi :

$$f(x,y) = X^T A Y = (x_1, \cdots, x_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \cdots & \cdots & a_m \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & \cdots & a_m \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{pmatrix}$$

où X^T et Y représentent les vecteurs ligne et colonne associés à x et y respectivement.

Exemple 4.7.

Soit $\phi_1 \in E^*$ et $\phi_2 \in E^*$ (deux formes linéaires sur E) alors l'application :

$$f\ :\ E \times E \to \mathbb{K}$$
$$(u,\nu) \mapsto \quad f(u,\nu) = \phi_1(u)\phi_2(\nu)$$

définie une forme bilinéaire sur E^2 et est appelée **produit tensoriel**.

Exemple 4.8.

Soit $E = \mathcal{C}[0,1]$ (espace vectoriel des fonctions numériques continues sur l'intervalle [0,1]). Alors l'application :

$$f\ :\ \mathcal{C}[0,1] \times \mathcal{C}[0,1] \to \mathbb{R}$$
$$(x,y) \mapsto \int_0^1 x(t)\ y(t)\ dt$$

est une forme bilinéaire sur E^2.

Remarque 4.7.

On vérifie facilement que l'ensemble de toutes les formes bilinéaires définies sur $E \times F$ (resp. sur $E \times E$) est un espace vectoriel sur $\mathbb{K}$ et on le note :

$$\mathcal{L}_2(E, F, \mathbb{K}) \text{ (resp. } \mathcal{L}_2(E, \mathbb{K}) \text{ et } (\dim_{\mathbb{K}} \mathcal{L}_2(E, F, \mathbb{K}) = nm)$$

car

$$\{\forall\, f,\ g \in \mathcal{L}_2(E, \mathbb{K}), (u \in E, \nu \in F), (f+g)(u,\nu) = f(u,\nu) + g(u,\nu)$$

$$\{\forall\, \lambda \in \mathbb{K},\ \forall\, f \in \mathcal{L}_2(E, F, \mathbb{K})) \quad (\lambda f)(u,\nu) = \lambda f(u,\nu)$$

Théorème 4.5. *[Base de l'espace vectoriel des formes bilinéaires]*

Soit $\{\phi_1, \cdots, \phi_n\}$ une base du dual E^* et $\{\sigma_1, \cdots, \sigma_m\}$ une base de F^* ; alors on définit une base de $\mathcal{L}_2(E, F, \mathbb{K})$ par la famille :

$$f_{ij}(u,\nu) = \phi_i(u)\sigma_j(\nu) \forall \begin{array}{rcl} i & = & 1, \cdots, n \\ j & = & 1, \cdots, m \end{array}$$

Le résultat suivant établit un isomorphisme entre l'espace vectoriel $\mathcal{L}_2(E, F, \mathbb{K})$ et l'espace vectoriel des matrices $\mathcal{M}_{\mathbb{K}}(n,m)$, et donne la représentation matricielle associée à tout $f \in \mathcal{L}_2(E, F, \mathbb{K})$ ainsi que la transformation de cette représentation matricielle par un changement de base.

Théorème 4.6.

Soit $\{a_i\}_{i=1,\cdots,n}$, une base de E (espace vectoriel sur $\mathbb{K}$) et $\{b_i\}_{i=1,\cdots,m}$ une base de F (espace vectoriel sur $\mathbb{K}$) $\Rightarrow$

i) L'espace vectoriel $\mathcal{L}_2(E, F, \mathbb{K})$ des formes bilinéaires sur ($E \times F$) est isomorphe à $\mathcal{M}_{\mathbb{K}}(n,m)$ l'espace vectoriel des matrices $n \times m$ sur $\mathbb{K}$:

$$\mathcal{L}_2(E, F, \mathbb{K}) \sim \mathcal{M}_{\mathbb{K}}(n,m)$$

(resp. pour $E = F$: $\mathcal{L}_2(E, \mathbb{K}) \sim \mathcal{M}_{\mathbb{K}}(n)$ espace vectoriel des matrices carrées n^2)

ii) a) A tout $f \in \mathcal{L}_2(E, F, \mathbb{K})$ on peut associer une représentation matricielle :

$$A_f = \{\alpha_{ij}\}_{\substack{i=1,\cdots,n \\ j=1,\cdots,m}} \text{ avec } \alpha_{ij} = f(a_i, b_j)$$

*(resp. à tout $f \in \mathcal{L}_2(E, \mathbb{K})$, on associe une **matrice carrée***

$$A_f = \{\alpha_{ij}\}_{\substack{i=1,\cdots,n \\ j=1,\cdots,\ m}}$$

b) Si X et Y représentent les matrices unicolonnes de $x \in E$ et $y \in F$ respectivement (relativement aux bases $\{a_i\}_i$ et $\{b_j\}_j$ alors :

$$f(x,y) = Y^t A_f X = X^t A_f^t Y$$

*iii) La matrice A'_f associée à f après un **changement de base** :*

$$\{a_i\}_{i=1,\cdots,n} \xrightarrow{P} \{a'_i\}_{i=1,\cdots,n} \text{ (dans } E) \text{ (ou } X = PX' \text{ et } P \text{ non singulière)}$$

$$QP \longrightarrow \{b'_j\}_{j=1,\cdots,m} \text{ (dans } F) \text{ (ou } Y = QY' \text{ et } Q \text{ non singulière)}$$

est donnée par la formule :

$$A'_f = Q^t AP$$

Remarque 4.8.

a) A constater que $f(x,y)$ est une “matrice” 1×1 donc elle est égale à sa transposée :

$$f(x,y) = X^t A_f^t Y = f^t(x,y)$$

b) A comparer le résultat iii) du théorème $A'_f = Q^t AP$ avec la formule de "passage" $A'_f = Q^{-1}AP$ dans le cas où f est une application linéaire $f \in \mathcal{L}_{\mathbb{K}}(E,F)$.

c) Si $F = E$ alors le changement de base donne comme représentation matricielle de $f \in \mathcal{L}_2(E,F)$:

$$A'_f = P^t AP$$

* A'_f et A_f sont appelées matrices congruentes.

Terminons ce chapitre avec une définition sur le rang de $f \in \mathcal{L}_2(E,F)$.

Définition 4.6.

a) On appelle **rang d’une forme bilinéaire** sur $E \times E$, $(f \in \mathcal{L}_2(E,F))$ le rang de la matrice A_f associée.

b) Une forme bilinéaire, $f \in \mathcal{L}_2(E,F)$ est **non dégénérée** (resp. **dégénérée**) si $rg\ f = \dim_{\mathbb{K}} E$. (resp. si $\underline{rg\ f < \dim_{\mathbb{K}} E}$).

Chapitre 5

Formes bilinéaires symétriques Formes quadratiques - Formes hermitiennes

Rappels - Définitions - Propriétés - Remarques sur les :
- **3. Formes bilinéaires symétriques - Formes quadratiques**
- **4. Formes bilinéaires alternées**
- **5. Formes hermitiennes**

Rappel de notations :
- $\underline{\mathbb{K}}$= corps commutatif.
- $E,\ F$ = espaces vectoriels sur $\mathbb{K}$ ($\dim_{\mathbb{K}} E = n$; $\dim_{\mathbb{K}} F = m$) .
- $\mathcal{L}_{\mathbb{K}}(E, F)$ = esp. vectoriel des applications linéaires $f\ : E \to F$ ($\dim_{\mathbb{K}} \mathcal{L}_{\mathbb{K}}(E, F) = nm$)
- $\mathcal{L}_{\mathbb{K}}(E)$ = esp. vectoriel des endomorphismes sur E ; $f\ : E \to E$ ($\dim \mathcal{L}_{\mathbb{K}}(E) = n^2$)
- $\mathcal{L}_2(E, F, K)$ = esp. vectoriel des formes bilinéaires sur $E \times F$ ($\dim_{\mathbb{K}} \mathcal{L}_2(E, F, \mathbb{K}) = nm$)
- $\mathcal{L}_2(E, \mathbb{K})$ = esp. vectoriel des formes bilinéaires sur $E \times E$; ($\dim_{\mathbb{K}} \mathcal{L}_2(E, \mathbb{K}) = n^2$)
- $\mathcal{M}_{\mathbb{K}}(n, m)$ = Esp. vectoriel des dimensions nm sur $\mathbb{K}$ et **isomorphe** à $\mathcal{L}_{\mathbb{K}}(E, F)$ et **isomorphe** à $\mathcal{L}_2(E, F, \mathbb{K})$
- $\mathcal{M}_n(\mathbb{K})$ = esp. vectoriel des matrices carrées de dimension $\underline{n^2}$ sur $\mathbb{K}$ et **isomorphe** à $\mathcal{L}_{\mathbb{K}}(E)$ et **isomorphe** à $\mathcal{L}_2(E, \mathbb{K})$.

3

1 Formes bilinéaires Symétriques

Définition 5.1.

*Soit $f \in \mathcal{L}_2(E, \mathbb{K})$. On dit que f est une **forme bilinéaire symétrique**,*

$$f\ :\quad \left.\begin{array}{c} E\times E\to\mathbb{K} \\ (x,y)\mapsto f(x,y)\end{array}\right\}\ \textbf{\textit{ssi}}$$

$$f(x,y)=f(y,x)\quad \forall\,(x,y)\in E^2$$

Remarque 5.1.
D'après l'isomorphisme : (v. th. 4.6)

$$\mathcal{L}_2(E,\mathbb{K})\sim\mathcal{M}_n(K)$$

Toute représentation matricielle A_f d'une forme bilinéaire symétrique $f\in\mathcal{L}_2(E,\mathbb{K})$ est une matrice carrée symétrique car on a :

$$\begin{array}{c} f(x,y)=X^tA_f^tY=Y^tA_fX=f(y,x)=Y^tA_f^tX=X^tA_fY \\ \Leftrightarrow X^tA_f^tY=X^tA_fY\Leftrightarrow A_f^t=A_f\end{array}$$

Le résultat suivant établit la **diagonalisation** de toute matrice A_f associée à une forme bilinéaire symétrique f.

Théorème 5.1.
Soit $f\in\mathcal{L}_2(E,\mathbb{K})$ symétrique, et soit $A_f\in\mathcal{M}_n(K)$ sa matrice carrée associée. Alors, il existe une matrice $P\in\mathcal{M}_n(\mathbb{K})$ ***non singulière*** *($\exists\ P^{-1}$) t.q. la matrice :*

$$\tilde{A}_f=P^tA_fP\qquad \textbf{\textit{soit diagonale}}$$

Important !

Remarque 5.2.
a) D'après la remarque du chapitre n^0 4 la matrice A_f est congruente à la matrice diagonale associée.
b) D'après le procédé d'échelonnage (v. cours n^0 1) d'une matrice carrée, on peut déduire qu'une matrice P inversible du théorème n'est rien d'autre que le produit de transformations (matrices) élémentaires. Donc, une façon d'obtenir P^t (ou P^tA_fP) est d'effectuer une suite d'opérations élémentaires d'échelonnage sur les lignes et d'effectuer la même suite d'opérations sur les colonnes.

Exemple 5.1.
Soit $f\in\mathcal{L}_2(E,\mathbb{K})$ symétrique avec $A_f\in\mathcal{M}_n(K)$ définie par :

$$A_f=\begin{pmatrix}1&2&-3\\2&5&-4\\-3&-4&8\end{pmatrix}\qquad \textit{(matrice symétrique)}$$

On utilise la matrice "bloc"

$$\begin{array}{cc}(A_f,I)=\begin{pmatrix}1&2&-3&\vdots&1&0&0\\2&5&-4&\vdots&0&1&0\\-3&-4&8&\vdots&0&0&1\end{pmatrix} & \begin{array}{c}(L_1)\\(L_2)\\(L_3)\end{array}\end{array}$$

En appliquant les transformations :

$$(A_f,I)\xrightarrow{\mathcal{I}_1^L}(A'_f,P^{(1)})\ \textit{avec}\ (\mathcal{I}_1^{(L)})\ :\left\{\begin{array}{rcl}L'_1&=&L_1\\L'_2&=&-2L_1+L_2\\L'_3&=&3L_1+L_3\end{array}\right.$$

et

$$(A'_f, P^{(1)}) \xrightarrow{\mathcal{I}_1^C} (A''_f, P^{(1)}) \text{ avec } (\mathcal{I}_1^{(C)}) \ : \left\{ \begin{array}{lcl} C''_1 & = & C'_1 \\ C''_2 & = & -2C'_1 + C'_2 \\ C''_3 & = & 3C'_1 + C'_3 \end{array} \right.$$

On obtient :

$$(A'_f, P^{(1)}) = \begin{array}{c} (C'_1)(C'_2)(C'_3) \\ \begin{pmatrix} 1 & 2 & -3 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 2 & \vdots & -2 & 1 & 0 \\ 0 & 2 & -1 & \vdots & 3 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} (L'_1) \\ (L'_2) \\ (L'_3) \end{array}$$

et

$$(A''_f, P^{(1)}) = \begin{array}{c} (C''_1)(C''_2)(C''_3) \\ \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 2 & \vdots & -2 & 1 & 0 \\ 0 & 2 & -1 & \vdots & 3 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} (L''_1) \\ (L''_2) \\ (L''_3) \end{array}$$

Ensuite on applique Ies transformations :

$$(A''_f, P^{(1)}) \xrightarrow{\mathcal{I}_2^L} (A'''_f, P^{(2)}) \text{ avec } (\mathcal{I}_2^{(L)}) \ : \left\{ \begin{array}{lcl} L'''_1 & = & L''_1 \\ L'''_2 & = & L''_2 \\ L'''_3 & = & -2L''_2 + L''_3 \end{array} \right.$$

et

$$(A'''_f, P^{(2)}) \xrightarrow{\mathcal{I}_2^C} (A^{(i\nu)}_f, P^{(2)}) \text{ avec } (\mathcal{I}_2^{(C)}) \ : \left\{ \begin{array}{lcl} C_1^{(i\nu)} & = & C'''_1 \\ C_2^{(i\nu)} & = & C'''_2 \\ C_3^{(i\nu)} & = & -2C'''_2 + C'''_3 \end{array} \right.$$

On obtient finalement :

$$(A'''_f, P^{(2)}) = \begin{array}{c} (C'''_1)(C'''_2)(C'''_3) \\ \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 2 & \vdots & -2 & 1 & 0 \\ 0 & 0 & -5 & \vdots & 7 & -2 & 1 \end{pmatrix} \end{array}$$

et

$$(A^{(i\nu)}_f, P^{(2)}) = \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 0 & \vdots & -2 & 1 & 0 \\ 0 & 0 & -5 & \vdots & 7 & -2 & 1 \end{pmatrix}$$

Donc la matrice de passage (non singulière) P et sa transposée P^t sont définies par :

$$P^t = \mathcal{I}_2^{(L)} \mathcal{I}_1^{(L)} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 7 & -2 & 1 \end{pmatrix}$$

et

$$P = (\mathcal{I}_1^{(L)})^t (\mathcal{I}_2^{(L)})^t \equiv (P^t)^t = \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

et la matrice diagonale $\tilde{A}_f$ associée à A_f est :

$$\tilde{A}_f = A_f^{(i\nu)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix} = P^t A_f P$$

Exemple 5.2.
Soit $f \in \mathcal{L}_2(E, \mathbb{K})$ avec :

$$A_f = \begin{pmatrix} 1 & -2 & 3 \\ -2 & 6 & -9 \\ 3 & -9 & 4 \end{pmatrix}$$

Par la même méthode détaillée ci-dessus, on obtient successivement :

$$(A_f, I) = \begin{pmatrix} 1 & -2 & 3 & \vdots & 1 & 0 & 0 \\ -2 & 6 & -9 & \vdots & 0 & 1 & 0 \\ 3 & -9 & 4 & \vdots & 0 & 0 & 1 \end{pmatrix}$$

$$(A_f I) \xrightarrow{\mathcal{I}_1^{(L)}} (A'_f, P^{(1)}) = \begin{pmatrix} 1 & -2 & 3 & \vdots & 1 & 0 & 0 \\ 0 & 2 & -3 & \vdots & 2 & 1 & 0 \\ 0 & -3 & -5 & \vdots & -3 & 0 & 1 \end{pmatrix} ;$$

$$(A'_f, P^{(1)}) \xrightarrow{\mathcal{I}_1^{(C)}} (A''_f, P^{(1)}) = \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 2 & -3 & \vdots & 2 & 1 & 0 \\ 0 & -3 & -5 & \vdots & -3 & 0 & 1 \end{pmatrix} ;$$

$$(A''_f, P^{(1)}) \xrightarrow{\mathcal{I}_2^{(L)}} (A'''_f, P^{(2)}) = \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 2 & -3 & \vdots & 2 & 1 & 0 \\ 0 & 0 & -19 & \vdots & 0 & 3 & 2 \end{pmatrix} ;$$

$$(A'''_f, P^{(2)}) \xrightarrow{\mathcal{I}_2^{(C)}} (A_f^{(i\nu)}, P^{(2)}) = \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 2 & 0 & \vdots & 2 & 1 & 0 \\ 0 & 0 & -38 & \vdots & 0 & 3 & 2 \end{pmatrix} ;$$

Donc finalement :

$$P^t = \mathcal{I}_2^{(L)} \mathcal{I}_1^{(L)} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 3 & 2 \end{pmatrix}$$

et

$$P^t = (\mathcal{I}_1^{(L)})^t (\mathcal{I}_2^{(L)})^t = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

et

$$\tilde{A}_f = A_f^{(i\nu)} = P^t A P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -38 \end{pmatrix}$$

2 Formes Quadratiques

Définition 5.2.

Pour toute forme bilinéaire $f \in \mathcal{L}_2(E, \mathbb{K})$ symétrique on définit la **forme quadratique associée** q_f par l'application :

$$\begin{aligned} q_f \ : \quad & E \to \mathbb{K} \\ & \nu \to q_f(\nu) = f(\nu, \nu) \end{aligned}$$

Remarque 5.3 (Forme polaire d'une forme bilinéaire symétrique).

D'après la définition précédente on aura : $\forall (u, \nu) \in E^2$

$$\begin{aligned} q(u+\nu) &= f(u+\nu, u+\nu) \\ q(u) &= f(u,u) \\ q(\nu) &= f(\nu,\nu) \end{aligned}$$

En "additionnant" ces égalités et par application de la bilinéarité et la symétrie de f on aura :

$$\begin{aligned} q(u+\nu) - q(u) - q(\nu) \quad &= f(u,u) + f(u,\nu) + f(\nu,u) + f(\nu,\nu) \\ &-f(u,u) - f(\nu,\nu) = 2f(u,\nu) \end{aligned}$$

$$\Rightarrow \quad f(u,\nu) = \frac{1}{2}[q(u+\nu) - q(\nu) - q(u)] \qquad (III.2.1)$$

On appelle cette expression de f la **forme polaire** de l'application bilinéaire symétrique.

Tenant compte des résultats précédents on a le :

Théorème 5.2 (Représ. matricielle et f.canonique d'une f.quadratique).

i) Soit $f \in \mathcal{L}_2(E, \mathbb{K})$ symétrique et soit $A_f = \{a_{ij}\}_{\substack{i=1,\cdots,n \\ j=1,\cdots,n}} = \{a_{ji}\}_{\substack{i=1,\cdots,n \\ j=1,\cdots,n}} \in \mathcal{M}_n(K)$ sa matrice carrée symétrique associée alors :
$\Rightarrow$ la forme quadratique q_f associée à f, ($\forall\ x \in E$) est donnée par :

$$q = X^t A_f X = \sum_{i,j=1}^{n} a_{ij} x_i x_j = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{i<j} a_{ij} x_i x_j \qquad (III.2.2)$$

$\{a_{ij}\} = A_f$ est la représentation matricielle de q_f et $\overline{rg\ q_f = rg\ A_f = r}$

*ii) Deux formes quadratiques sur $\mathbb{K}$ sont **équivalentes ssi** $\overline{\text{leurs matrices sont}}$ **congruentes** sur $\mathbb{K}$:*

$$rg\ q' = rg B^t A_f\ B = rg\ A_f = rg\ q$$

*où B est une matrice **non singulière**.*

iii) Il existe une matrice inversible B telle que $B^tA_fB = \tilde{A}_f$ matrice associée à la nouvelle forme quadratique q'_f soit diagonale.
$\Leftrightarrow$ la nouvelle forme quadratique associée q'_f s'écrit sous forme canonique :

$$\underline{q'_f = \sum_{i=1}^{n} a'_{ii} x_i^2} \qquad (III.2.3)$$

Remarque 5.4.

a) L'expression $q = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$ est un polynôme homogène par rapport aux variables x_i et est appelé le polynôme quadratique correspondant à la matrice A_f.
Donc sachant la forme du polynôme homogène q, on peut si f n'est pas déterminée, trouver la représentation matricielle A_f (ou A_q) en n'oubliant pas dans le procédé inverse diviser par 2 les coefficients a_{ij} avec $i \neq j$

b) Si $r = (rg\ f = rg\ q_f = rg\ A_f) = n$ alors q_f est dite **non singulière**
(si $r < n \Rightarrow q_f$ est **singulière**).

Exemple 5.3.

a) Soit : $q(x) = x_1^2 + 2x_2^2 - 7x_3^2 - 4x_1x_2 + 8x_1x_3$ alors on a :

$$q = X^tA_fX = X^t \begin{bmatrix} 1 & -2 & 4 \\ -2 & 2 & 0 \\ 4 & 0 & -7 \end{bmatrix} X \text{ donc : } A_f = A_q = \begin{bmatrix} 1 & -2 & 4 \\ -2 & 2 & 0 \\ 4 & 0 & -7 \end{bmatrix}$$

est la représentation matricielle de q.

b) On peut diagonaliser cette matrice par la même méthode d'éehelonnage (v. exemples précédents) expliquée auparavant, et trouver la matrice non singulière $B \in \mathcal{U}_3(\mathbb{R})$ t.q. $\tilde{A}_f = B^tA_fB$ **soit diagonale** donc t.q. $q'(y)$ prenne sa forme canonique par rapport à Y avec $X = BY$. Autrement dit :

$$\underline{X^t = Y^tB^t} \text{ et } \underline{X^tA_fX = Y^t(B^tA_fB)Y = Y^t\tilde{A}_fY}$$

On écrit

$$(A_f, I) = \begin{pmatrix} 1 & -2 & 4 & \vdots & 1 & 0 & 0 \\ -2 & 2 & 0 & \vdots & 0 & 1 & 0 \\ 4 & 0 & -7 & \vdots & 0 & 0 & 1 \end{pmatrix}$$

Par des transformations successives sur les lignes et les colonnes, on obtient :

$$\begin{aligned} (A_f, I) \quad &\mapsto (A'_f, P^{(1)}) = \begin{pmatrix} 1 & -2 & 4 & \vdots & 1 & 0 & 0 \\ 0 & -2 & 8 & \vdots & 2 & 1 & 0 \\ 0 & 8 & -23 & \vdots & -4 & 0 & 1 \end{pmatrix} \\ &\mapsto (A''_f, P^{(1)}) = \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & -2 & 8 & \vdots & 2 & 1 & 0 \\ 0 & 8 & -23 & \vdots & -4 & 0 & 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned}(A''_f, P^{(1)}) &\mapsto (A'''_f, P^{(2)}) = \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & -2 & 8 & \vdots & 2 & 1 & 0 \\ 0 & 0 & 9 & \vdots & 4 & 4 & 1 \end{pmatrix} \\ &\mapsto (A^{(i\nu)}_f, P^{(2)}) = \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & -2 & 0 & \vdots & 2 & 1 & 0 \\ 0 & 0 & 9 & \vdots & 4 & 4 & 1 \end{pmatrix} \\ &\Rightarrow \tilde{A}_f = A^{(i\nu)}_f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 9 \end{pmatrix}\end{aligned}$$

et

$$B^t = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 4 & 1 \end{pmatrix},\ B = (B^t)^t = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\Leftrightarrow \tilde{A}_f = B^t A_f\ B$$

Donc finalement : $q(x) = q'(y) = y_1^2 - 2y_2^2 + 9y_3^2$ **forme canonique** de q et q est **non singulière**.

Remarque 5.5.

On peut réaliser la réduction d'une forme quadratique à une forme canonique par le procédé de **Gauss-Lagrange** qui consiste essentiellement à compléter successivement les "débuts" des carrées.

Exemple 5.4 (la même forme quadratique q (Ex. 5.3)).

$$\begin{array}{rcl} q &=& x_1^2 + 2x_2^2 - 7x_3^2 - 4x_1x_2 + 8x_1x_3 \\ &=& x_1^2 - 4x_1[x_2 - 2x_3] + 2x_2^2 - 7x_3^2 \\ && [x_1^2 - 4x_1(x_2 - 2x_3) + 4(x_2 - 2x_3)^2] + 2x_2^2 - 7x_3^2 - 4(x_2 - 2x_3)^2 \\ &=& (x_1 - 2x_2 + 4x_3)^2 - 2(x_2^2 - 8x_2x_3) - 23x_3^2 \\ &=& (x_1 - 2x_2 + 4x_3)^2 - 2(x_2^2 - 8x_2x_3) + 16x_3^2) + 9x_3^2 \\ &=& (x_1 - 2x_2 + 4x_3) - 2(x_2 - 4x_3)^2 + 9x_3^2 \end{array}$$

On pose :

$$\left.\begin{array}{rcl} y_1 &=& x_1 - 2x_2 + 4x_3 \\ y_2 &=& x_2 - 4x_3 \\ y_3 &=& x_3 \end{array}\right\} \Leftrightarrow \begin{array}{rcl} x_1 &=& y_1 + 2y_2 + 4y_3 \\ x_2 &=& y_2 + 4y_3 \\ x_3 &=& y_3 \end{array} \quad \Leftrightarrow X = BY$$

(où on reconnait la transf. associée à B)

Donc :

$$q(X) = q'(y) = y_1^2 - 2y_2^2 + 9y_3^2$$

Par la même méthode on réduit la forme quadratique sur K^3,

$$Q(x, y, z) = x^2 + y^2 - 2z^2 + xy + 6xz$$

à sa forme canonique par des transformations successives suivantes :

$$\begin{aligned}
Q(x,y,z) &= x^2 + 2x \cdot \left(\frac{y}{2} + 3z\right) + y^2 - 2z^2 \\
&= \left(x + \frac{y}{2} + 3z\right)^2 - \left(\frac{y}{2} + 3z\right)^2 + y^2 - 2z^2 \\
&= \left(x + \frac{y}{2} + 3z\right)^2 + \frac{3}{4}y^2 - 11z^2 - 3zy \\
&= \left(x + \frac{y}{2} + 3z\right)^2 + \frac{3}{4}(y^2 - 4yz) - 11z^2 \\
&= \left(x + \frac{y}{2} + 3z\right)^2 + \frac{3}{4}(y - 2z)^2 - 3z^2 - 11z^2 \\
&= \left(x + \frac{y}{2} + 3z\right)^2 + \frac{3}{4}(y - 2z)^2 - 14z^2
\end{aligned}$$

On pose :

$$\left.\begin{array}{lcl}
w_1 &=& x + \frac{y}{2} + 3z \\
w_2 &=& y - 2z \\
w_3 &=& z
\end{array}\right\} \Leftrightarrow Q'(w) = w_1^2 + \frac{3}{4}w_2^2 - 14w_3^2$$

On obtient donc une forme canonique de $Q(x,y,z)$

3 Loi d'inertie de Sylvester

Si $K = \mathbb{R}$ on a Ie résultat suivant plus particulier pour une réduction canonique "normalisée" qui a comme conséquence l'unicité de cette forme canonique (**Loi d'inertie de Sylvester**).

Théorème 5.3.

Soit $q = X^t AX$ une forme quadratique réelle. Alors il existe une application réelle non singulière B t.q. $X = BW$ et t.q. la forme canonique $q'(w)$ soit donnée par :

$$q'(w) = q(x) = X^t AX = W^t(B^t AB)W = \sum_{i=1}^{P} w_1^2 - \sum_{P+1}^{r} w_j^2 \qquad (III.3.1.)$$

où P est le nombre de termes positifs et $r = rg\ q$

Remarque 5.6.

Supposons que d'après les méthodes du paragraphe précédent on ait obtenu pour q la forme canonique :

$$q(y) = \sum_{i=1}^{P} S_i y_i^2 - \sum_{j=P+1}^{r} S_j y_j^2 \text{ avec } S_i > 0 \quad \forall\, i = 1, \cdots, r$$

L'application Q non singulière t.q. :

$$\begin{cases} Z_i = \sqrt{S_i} y_i & i = 1, \cdots, r \\ Z_j = y_j & j = r + 1 \cdots n \end{cases}$$

Donc

$$\left.\begin{array}{rcl} Y & = & QZ \\ Y^t & = & Z^t Q^t \end{array}\right\} \text{ avec } Q = \begin{pmatrix} \dfrac{1}{\sqrt{S_1}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \dfrac{1}{\sqrt{S_r}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

transforme q en la forme canonique $q'(Z) = Z^t(Q^t A_q Q)Z$

$$= Z^t \tilde{A} Z$$

ou $q'(Z) = \sum\limits_{i=1}^{P} Z_i^2 - \sum\limits_{j=1}^{N} Z_j^2$ avec $\underline{rg\ \tilde{A} = rg\ A = r}$

Définition 5.3.

On appelle **indice**, le nombre P de termes positifs d'une **forme quadratique réelle réduite en sa forme canonique** (III.3.1.).

Théorème 5.4 (Loi d'inertie de Sylvester).

Toute forme quadratique réelle sur E avec $\dim_R E = n$ admet une représentation canonique du type (III.3.1.) ***unique****.*
Autremcnt dit, si q admet deux expressions canoniques du type III.3.1. alors ces deux formes ***ont le même rang*** *r* ***et le même indice*** *P.*

Remarque 5.7.

On appelle signature Sg :

a) d'après certaines références, le couple : $\underline{Sg = (P; N)}$ (N Ie nombre de termes négatifs) et

b) d'après d'autres références la différence : $\underline{\tilde{Sg} = P - N}$.

Exemple 5.5.

La forme quadratique des exemples III.2.1. et III.2.2 (définie sur $\mathbb{R}^3$)

$q(x) = x_1^2 + 2x_2^2 - 7x_3^2 - 4x_1x_2 + 8x_1x_3$ *a été réduite à :*
$q' = y_1^2 - 2y_2^2 + 9y_3^2$

L'application non singulière, définie par (f_1) :

$$\mathbb{R}^3 \to \mathbb{R}^3 \begin{cases} y_1 = z_1 \\ y_2 = z_3 \\ y_3 = z_2 \end{cases} \qquad \text{donne : } \begin{cases} q' \mapsto q'' \\ q'' = z_1^2 + 9z_8^2 - 2z_3^2 \end{cases}$$

et l'application non singulière, définie par (f_2) :

$$\mathbb{R}^3 \to \mathbb{R}^3 \begin{cases} z_1 = w_1 \\ z_2 = w_2/3 \\ z_3 = w_3/\sqrt{2} \end{cases} \qquad \text{donne : } q'' \mapsto q''' \textit{ avec } q''' = w_1^2 + W_2^2 - w_3^2$$

Le produit de ces deux applications est une application $g \in \mathcal{L}_{\mathbb{R}}(\mathbb{R}^3)$ avec matrice associèe B définie par : $X = B_gW$ ou

$$\begin{cases} x_1 &=& w_1 + \frac{4}{3}w_2 + \sqrt{2}w_3 \\ x_2 &=& \frac{4}{3}w_2 + \frac{\sqrt{2}}{2}w_3 \\ x_3 &=& \frac{1}{3}w_2 \end{cases} \qquad \text{donc } B_g = \begin{bmatrix} 1 & 4/3 & \sqrt{2} \\ 0 & 4/3 & \sqrt{2}/2 \\ 0 & 1/3 & 0 \end{bmatrix}$$

Cette application non singulière réduit q en sa forme canonique ***unique** :*

$$q \xrightarrow{B_g} q'''(w) = w_1^2 + w_2^2 - w_3^2 \Leftrightarrow \tilde{A}_q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Donc la forme quadratique q est d'indice 2 de rang 3 et de signature $Sg = (2; -1)$ (ou $\tilde{Sg} = (1)$)

Définition 5.4 (Formes Quadratiques définies positives).

*Une forme quadratique réelle est **définie positive** (resp. **définie négative**) si son indice P est **égal à son rang** et det. $A_q \neq 0$ donc q = non singulière - autrement dit $P = r = n$ (resp. si $P = 0$ et $n = r$)*

Remarque 5.8.

D'après la forme canonique de la loi d'inertie, une forme quadratique réelle **définie positive** (resp. **définie négative**) se réduit en une **forme canonique positive** (resp. **négative**)pour tout ensemble de valeurs réelles de $y \in \mathbb{R}^n$.

$$\tilde{q}_p = \sum_{i=1}^{n} y_i^2 > 0 \left(\textit{resp. } \tilde{q}_n = -\sum_{j=1}^{n} y_j^2 < 0 \right)$$

Définition 5.5.

Une forme quadratique réelle est semi-définie positive (resp. semi-définie négative) si $r < n$ et $P = r$ (resp. si $p = 0$ et $r < n$).

Remarque 5.9.

Par analogie avec la remarque III.3.3. et déf. III.3.2., une forme quadratique réelle **semi-définie positive** (resp. **semi-définie négative**) est singulière et se réduit en une forme canonique :

$$\tilde{q}_{sp} = \sum_{i=1}^{r} z_i^2 \geq 0,\ r < n\ ;\ \left(\textit{resp. } \tilde{q}_{sn} = -\sum_{j=1}^{r} z_j^2 \geq 0 \right)$$

On a le :

Théorème 5.5.

Si $q = X^t A_q X$ est une forme quadratique ***définie positive*** *alors* $\det A_q > 0$.
Et voici quelques résultats sur les matrices associées.

Définition 5.6.

Une matrice A_q associée à une forme quadratique réelle q est dite définie ou semi-définie pos. (resp. déf. négative ou semi déf négative) selon que la forme quadratique est définie ou semi-définie positive (resp. définie ou semi-définie négative).
Le critère de positivité pour les matrices est fourni par le :

Théorème 5.6.

i) Une matrice ***réelle symétrique*** A ***est définie positive*** *si $\exists$ une matrice inversible C t.q. :*

$$A = C^t C$$

ii) Une matrice réelle symétrique de rang r est ***semi-définie positive ssi*** *$\exists$ une matrice C* ***de rang*** *r t.q. :*

$$A = C^t C$$

4 Formes Bilinéaires Alternées (Antisymétriques)

1 Formes Antisymétriques

Définition 5.7.

Une forme bilinéaire $f \in \mathcal{L}_2(E, \mathbb{K})$ est alternée ou antisymétrique si :

$$\underline{f(\nu, \nu) = 0} \quad \forall\, \nu \in E \textit{ (Ant. i)}$$

ou

$$\Leftrightarrow f(u, \nu) = -f(\nu, u) \quad \forall\, (u, \nu) \in E^2 \textit{ (Ant. ii)}$$

Remarque 5.10.

On vérifie facilement que la condition (Ant. i) implique (Ant. ii) et inversement.
On a le résultat très utile pour la réduction des formes bilinéaires alternées (antisymétriques)(ou leurs matrices associées) donné par le :

Théorème 5.7. ***(Décomposition en blocs d'une matrice antisymétri-que)***
Soit $f \in \mathcal{L}_2(E, \mathbb{K})$ alternée alors :
i) la matrice associée à f, $\underline{A_f}$ est <u>antisymétrique</u> : $A_f = -A_f^t$
ii) pour toute matrice A_f (antisymétrique) il existe une matrice P non singulière t.q.

$$\underline{\tilde{A}_f = P^t A_f P}$$

soit de la forme canonique <u>diagonale par blocs</u> suivante :

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

Le nombre d_r des matrices ***blocs antisymétriques*** *égale :*

$$d_r = \frac{1}{2} r \quad \textit{où} \quad r = rg\ f = rg\ A_f = rg \tilde{A}_f$$

Remarque 5.11.
a) On vérifie facilement que si A est antisymétrique alors toute matrice :

$$B = P^t A P$$

congruente à A est une matrice antisymétrique car :

$$B^t = (P^t A P)^t = P^t A^t P = -P^t A P = -B$$

b) D'après le th. III.4.1. toute forme bilinéaire alternée (antisymétrique a un rang r **= entier pair** (puisque $r = 2d_r \quad d_r \in \mathbf{N}^*$*).*

Exemple 5.6.
On considère la matrice antisymétrique :

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

Par une méthode analogue (à celle des ex. III.1.1, et III.2.1.), on cherche une matrice P non singulière t.q. $\tilde{A} = P^t A P$ soit dans la forme canonique du théorème III.4.1.

1$^{\text{ère}}$ transformation *On échange la 3$^{\text{ème}}$ et la 2$^{\text{ème}}$ ligne de*
$(A.I) \to (A', P^{(1)})$.

2ème transformation *On échange la 3ème et la 2ème colonne de* $(A', P^{(1)})$.

$$(A', P^{(1)}) \to (A'', P^{(1)}) = \begin{bmatrix} 0 & 2 & 0 & 4 & \vdots & 1 & 0 & 0 & 0 \\ -2 & 0 & -1 & -2 & \vdots & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -3 & \vdots & 0 & 1 & 0 & 0 \\ -4 & 2 & 3 & 0 & \vdots & 0 & 0 & 0 & 1 \end{bmatrix}$$

3ème transformation *On multiplie la 1ère ligne par* $\frac{1}{2} \to (A''', P^{(2)})$.

4ème transformation *On multiplie la 1ère colonne*

$$\frac{1}{2} \to (A^{(i\nu)}, P^{(2)}) = \begin{bmatrix} 0 & 1 & 0 & 2 & \vdots & 1/2 & 0 & 0 & 0 \\ -1 & 0 & -1 & -2 & \vdots & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -3 & \vdots & 0 & 1 & 0 & 0 \\ -2 & 2 & 3 & 0 & \vdots & 0 & 0 & 0 & 1 \end{bmatrix}$$

5ème transformation *Echelonnage (entre 1ère et 3ème ligne)*

6ème transformation *Echelonnage (entre 2ème et 4ème ligne)*

7ème transformation *Echelonnage (entre 1ère et 4ème ligne)*

8ème transformation *Echelonnage (entre 2ème et 4ème colonne)*

9ème transformation *Echelonnage (entre 1ère et 3ème colonne)*

$$\Rightarrow (A^{(ix)}, P^{(4)}) = \begin{bmatrix} 0 & 1 & 0 & 0 & \vdots & 1/2 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & \vdots & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -5 & \vdots & -1/2 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 & \vdots & -1 & 0 & -2 & 1 \end{bmatrix}$$

10ème transformation *On multiplie la 3ème ligne par* $-\frac{1}{5}$

11ème transformation *On multiplie la 3ème colonne par* $-\frac{1}{5}$

$$\Rightarrow (A^{(xi)}, P^{(5)}) = \begin{bmatrix} 0 & 1 & 0 & 0 & \vdots & 1/2 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & \vdots & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & \vdots & 1/10 & -1/5 & 0 & 0 \\ 0 & 0 & -1 & 0 & \vdots & -1 & 0 & -2 & 1 \end{bmatrix}$$

Donc

$$P^t = \begin{bmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1/10 & -1/5 & 0 & 0 \\ -1 & 0 & -2 & 1 \end{bmatrix} \text{ et } \tilde{A} = A^{(xi)} = P^t A P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

5 Formes Hermitiennes

On supposera que $\mathbb{K} = \mathbb{C}$ et E = espace vectoriel sur $\mathbb{C}$ avec $\dim_C E = n$.

Définition 5.8.
*On appelle **forme hermitienne** h, sur E toute application :*

$$\begin{array}{rcl} h \ : \ E \times E & \to & \mathbb{C} \\ (u, \nu) & \to & h(u, \nu) \quad \textit{t.q.} \end{array}$$

*h soit linéaire par rapport à la "première" variable (u) et **anti-linéaire** par rapport à la "seconde" variable (ν).*
Autrement dit : $\forall\, a, b, \in \mathbb{C}, \quad \forall\, u, \nu, u_1, u_2, \nu_1, \nu_2 \in E$ on a :
(i) $h(au_1 + bu_2, \nu) = ah(u_1, \nu) + bh(u_2, \nu)$
(ii) $h(u, a\nu_1 + b\nu_2) = \bar{a}h(u, \nu_1) + \bar{b}h(u, \nu_2)$
* la notation $\bar{a}, \bar{b}$ signifie "complexe conjugué" : $\left\{ \begin{array}{rcl} \mathit{Re}(a) & = & \mathit{Re}(\bar{a}) \\ \mathit{Im}(a) & = & -\mathit{Im}(\bar{a}) \end{array} \right.$

Remarque 5.12.
a) On peut définir une forme hermitienne sur E en prenant un couple de conditions équivalentes aux (i) et (ii) :

$$(i) \ \textit{et} \ (ii) \ \ \Leftrightarrow (i)\textit{et} \ \ (\tilde{ii}) \ : \ h(u, \nu) = \overline{h(\nu, u)}$$

b) La propriété (ii) donne :$h(u, u) = \overline{h(u, u)}$ donc

$$\begin{array}{rcl} \mathit{Im}(h(u, u)) & = & \mathit{Im}(\overline{h(u, u)}) \quad = \quad -\mathit{Im}(h(u, u)) \\ & & \mathit{Im}(h(u, u)) \quad = \quad 0 \\ & & h(u, u) \in \mathbb{R} \end{array}$$

Définition 5.9.
a) *Une matrice $A \in \mathcal{M}_n(\mathbb{C})$ avec $A = \{a_{ij}\}_{\substack{i=1,\cdots,n \\ j=1,\cdots,n}}$ est dite **matrice hermitienne** si*

$$a_{ij} = \bar{a}_{ji}$$

(Conséquence : $a_{ii} = \bar{a}_{ii} \Rightarrow a_{jj} \in \mathbb{R}$).
b) *On définit la **matrice adjointe** par :*

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

Remarque 5.13.
Une matrice hermitienne satisfait : $A^* = A$ (A est égale à sa matrice adjointe A^*).

Remarque 5.14.
Deux matrices hermitiennes $A, \ \tilde{A}$ sont **congruentes au sens de Hermite** si il existe une matrice non singulière P t.q.

$$\tilde{A} = P^* A \cdot P$$

Remarque 5.15.
Soit A une matrice hermitienne ($A = A^* = \bar{A}^t$). Montrons que d'après les définitions qui précèdent, la forme :

$$h(X, Y) = X^t A \bar{Y} \quad \forall\, (x, y) \in E \times E \quad (E = \mathbb{C}^n)$$

définit une forme hermitienne sur $\mathbb{C}^x$ car : soit $(a;b) \in \mathbb{C}^2$

$$\begin{aligned} h((aX_1 + bX_2), Y) &= (aX_1 + bX_2)^t A\bar{Y} = (aX_1^t + bX_2^t)A\bar{Y} \\ &= aX_1^t A\bar{Y} + bX_2^t A\bar{Y} = ah(X_1, Y) + bh(X_2, Y) \end{aligned} \tag{5.5.1}$$

(linéaire par rapport à la 1ère variable).

De même, si on utilise le fait que $X^t AY$ est un scalaire donc égal à son transposé :

$$(X^t A\bar{Y})^t = (X^t A\bar{Y})$$

On aura aussi :

$$\overline{h(XY)} = \overline{(X^t A\bar{Y})} = \overline{(X^t A\bar{Y})}^t = \overline{\bar{Y}^t A^t X} = Y^t A^* \bar{X} = Y^t A\bar{X} = h(Y, X) \tag{5.5.2}$$

(5.5.1) et (5.5.2) montrent que (i) et $(\tilde{ii})$ de la définition d'une forme hermitienne (v. remarque 5.12) sont vérifiées.

$$\Rightarrow h(X, Y) \equiv X^t A\bar{Y} \qquad \textit{est une forme hermitienne}$$

Remarque 5.16. *D'une manière analogue que pour les formes bilinéaires , on définit une* ***forme quadratique hermitienne*** $q(\nu)$ *associée à une* f *hermitienne* h *par :*

$$q(\nu) = h(\nu, \nu) \quad \forall\, \nu \in E$$

D'après la remarque 5.12 b)

$$q(\nu) \in \mathbb{R}$$

Pour la représentation matricielle d'une forme hermitienne, sa forme polaire et sa réduction en forme canonique on a le résultat suivant :

Théorème 5.8.

Soit h une forme hermitienne sur E alors :

i) Pour toute base $B = \{e_1, \cdots, e_n\}$ de E, la représentation matricielle $H = \{n_{ij}\}_{\substack{i=1,\cdots,n \\ j=1,\cdots,n}}$ de h est donnée par :

$$n_{ij} = h(e_i, e_j)\ ;$$

H est une matrice hermitienne (ayant donc des éléments diagonaux réels : $n_{ij} \in \mathbb{R}$ ou

$$n_{ii} = h(e_i, e_i) = \overline{h(e_i, e_i)})$$

ii) La forme polaire d'une forme hermitienne en termes de sa forme quadratique hermitienne associée est :

$$h(u, \nu) = \frac{1}{4}(q(u + \nu) - q(u - \nu)) + \frac{i}{4}(q(u + i\nu) - q(u - i\nu)).$$

iii) Il existe une base (donc une matrice P non singulière) dans E t.q. la représentation matricielle H associée soit diagonale :

$$(\tilde{H} = P^* HP)$$

La forme canonique de H est :

$$\tilde{C} = \begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{array}{l} P = \text{ indice de } H \\ r = \text{ rang de } H \end{array}$$

Exemple 5.7.

Soit $H = \begin{bmatrix} 1 & 1+2i & 2-3i \\ 1-2i & 5 & -4-2i \\ 2+3i & -4+2i & 13 \end{bmatrix}$ par la méthode d'échelonnage utilisée aux paragraphes précédents, on détermine une matrice P t.q. $\tilde{C} = P^*HP$ a la forme canonique du théorème. Les étapes (en résumé) sont :

$$(A, I) = \left[\begin{array}{ccc|ccc} 1 & 1+2i & 2-3i & 1 & 0 & 0 \\ 1-2i & 5 & -4-2i & 0 & 1 & 0 \\ 2+3i & -4+2i & 13 & 0 & 0 & 1 \end{array}\right] \rightarrow$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 5i & -1+2i & 1 & 0 \\ 0 & -5i & 0 & -2-3i & 0 & 1 \end{array}\right] = [A', P^{(1)}]$$

$$(A', P^{(2)}) \rightarrow (A'', P^{(2)})) = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 10 & 5i & 2 & 1 & i \\ 0 & -5i & 0 & -2-3i & 0 & 1 \end{array}\right] \rightarrow$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 10 & 0 & 2 & 1 & i \\ 0 & 0 & -25 & -20-20i & 5i & 5 \end{array}\right] = [A''', P^{(3)}]$$

$$(A''', P^{(3)}) \rightarrow (A^{(i\nu)}, P^{(4)}) = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 10 & 0 & 2 & 1 & i \\ 0 & 0 & -250 & -20-20i & 5i & 5 \end{array}\right]$$

$$\Rightarrow \tilde{C} = A^{(i\nu)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & -250 \end{bmatrix}$$

Donc :

$$\underline{\tilde{C} = P^*HP} \Leftarrow \text{ et } P^* = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & i \\ -20(1+i) & 5i & 5 \end{bmatrix}$$

Remarque 5.17.

a) D'après le théorème précédent toute forme quadratique hermitienne $h = X^*HX$ admet une forme canonique $\forall\, w \in E\ (E = \mathbb{C}^n)$.

$$h(w,w) = \sum_{i=1}^{P} \bar{w}_i w_i - \sum_{j=p+1}^{r} \bar{w}_j w_j \qquad IV.1.0$$

où P est l'indice et r le rang de la forme quadratique hermitienne, ou de la matrice H associée.

b) Deux formes hermitiennes sont équivalentes ssi elles ont le même indice et le même rang r.

Exemple 5.8.

Soit h la forme quadratique hermitienne sur $\mathbb{C}^3$ qui admet comme représentation matricielle $H \in \mathcal{M}_3(\mathbb{C})$ une matrice définie par l'ex. précédent. On peut trouver la forme canonique de h (associée à $H \in \mathcal{M}_3(\mathbb{C})$) suivante $\forall\, x \in \mathbb{C}^3$:

$$h(x,x) = \bar{x}_1 x_1 + \bar{x}_2 x_2 - \bar{x}_3 x_3 = |x_1|^2 + |x_2|^2 - |x_3|^2$$

Définition 5.10. *Formes hermitiennes définies positives*

*On dit qu'une forme hermitienne h non singulière $(r = n)$ sur $\mathbb{C}^n$ est **définie positive** si la f. quadratique associée vérifie :*

$$h(x,x) = X^*HX > 0$$

Autrement dit, si $r = p = n$ (L'indice est égal au rang).
La matrice H associée est dite aussi définie positive.
* *Une déf. analogue est valable pour une f. hermitienne **définie négative**.*

On a le résultat suivant :

Théorème 5.9. *Une forme hermitienne est définie positive **ssi il existe une matrice non singulière** $C \in \mathcal{M}_n(\mathbb{C})$ t.q. $H = C^*C$*

Définition 5.11. *Formes hermitiennes semi-définies positives (ou positives)*

Une forme hermitienne est semi-définie positive ou positive si la forme quadratique associée satisfait $h(x,x) = X^*HX \geq 0$ (ou si $r = p < n$). On dit que la matrice associée est positive (ou semi-définie positive).

* Une définition analogue est valable pour les f. hermitiennes semi-définies négatives.

Définition 5.12.

a) Une application $h\ :\ E \times E \to \mathbb{C}$

$$(u,\nu) \mapsto h(u,\nu),\ (u,\nu) \in E^2$$

est dite **antihermitienne** si : $i)$ *h est linéaire par rapport à u*
$ii)$ $h(u,\nu) = -h(\nu,u)$

b) Une **matrice antihermitienne** H satisfait :

$$H^* = -H$$

Remarque 5.18.

Si H est hermitienne alors on vérifie que :

$$H = -iA \text{ ou } A = -iH \text{ avec } A \text{ antihermitienne.}$$

D'après cette propriété et la réduction d'une matrice hermitienne ou sa forme hermitienne associée, on a le :

Théorème 5.10.

i) Si h est une forme antihermitienne alors toute représentation matricielle par rapport à une base donnée, est une matrice antihermitienne A_h.

ii) Il existe une base (ou une matrice P non singulière) dans E t.q. la représentation matricielle $\tilde{A}_h$ soit diagonale : $\tilde{A}_h = P^* A_h P$.

La forme canonique $\tilde{B}_h$ est :

$$\tilde{B}_h = \begin{bmatrix} iI_p & & 0 \\ 0 & -iI_{r-p} & 0 \\ 0 & 0 & 0 \end{bmatrix} \left\{ \begin{array}{l} \text{où} \\ P = \text{ indice de } -iA_h \\ r = \text{ rang de } A_h \end{array} \right.$$

Chapitre 6

Espaces Préhilbertiens Espaces Vectoriels Normés

Rappels - Définitions - Propriétés - Remarques sur les :
I. **Espaces préhilbertiens (Unitaires - Euclidiens)**
II. **Espaces Vectoriels Normés**

1 Espaces Préhilbertiens - Orthogonalité

1 Produit scalaire - Espaces Préhilbertiens (Euclidiens-Unitaires)

Définition 6.1.
Une forme hermitienne h est non dégénérée : $h(u, u) = 0$ ssi $u = 0$

Définition 6.2.
Soit E un espace vectoriel sur le corps (commutatif) $\mathbb{K}$.
*On appelle **produit scalaire** une forme hermitienne définie positive non dégénérée, notée $\langle\ ,\ \rangle$ définie sur E, autrement dit, un produit scalaire sur E vérifie : $\forall\ a, b \in \mathbb{K} \quad u_1, u_2, u, \nu \in E$*

$$\left|\begin{array}{ll} (P.S.1) & \langle au_1 + bu_2, \nu\rangle = a\langle u_1, \nu\rangle + b\langle u_2, \nu\rangle \\ (P.S.2) & \langle u, \nu\rangle = \overline{\langle \nu, u\rangle} \\ (P.S.3) & \langle u, u\rangle \geq 0 \text{ et } \langle u, u\rangle = 0 \text{ ssi } u = 0 \end{array}\right.$$

Un espace vectoriel E muni d'un produit scalaire $\langle\ ,\ \rangle$ est un espace **préhilbertien.**
Si $\mathbb{K} = \mathbb{R}$, le produit scalaire est alors une forme bilinéaire symétrique sur E ; l'espace E est très souvent appelé espace Euclidien.
Un espace Préhilbertien sur $\mathbb{K} = \mathbb{C}$ est appelé aussi espace Unitaire.

Remarque 6.1.
D'après les propriétés (P.S.1) et (P.S.2) on vérifie l'antilinéarité du produit scalaire.

$$\langle u, a\nu_1 + b\nu_2\rangle = \bar{a}\langle u, \nu_1\rangle + \bar{b}\langle u, \nu_2\rangle \left\{\begin{array}{ll} \forall & u, \nu_1, \nu_2 \in E \\ \forall & a, b \in \mathbb{K} \end{array}\right.$$

Exemple 6.1.

a) $E = \mathbb{R}^n$ Produit scalaire habituel :

$$\begin{aligned} &\forall \quad x = \{x_1, \ldots, x_n\} \in \mathbb{R}^n,\ y = \{y_1, \ldots, y_n\} \in \mathbb{R}^n \\ &\Rightarrow \langle x, y\rangle \equiv x \cdot y = \sum_{i=1}^{n} x_i y_i \\ &\Rightarrow \underline{(\mathbb{R}^n,\ \langle\ ,\ \rangle)\ \text{espace Euclidien.}} \end{aligned}$$

b) $E = \mathbb{C}^n$ Produit scalaire habituel :

$$\begin{aligned} &\forall \quad u = \{a_1, \ldots, a_n\} \in \mathbb{C}^n, \nu = \{b_1, \ldots, b_n\} \in \mathbb{C}^n \\ &\Rightarrow \langle u, \nu\rangle = u \cdot \nu = \sum_{i=1}^{n} a_i \bar{b}_i \\ &\Rightarrow (\mathbb{C}^n,\ \langle\ ,\ \rangle)\ \underline{\text{espace Unitaire.}} \end{aligned}$$

Exemple 6.2.

a) Soit $E = \mathcal{M}_{\mathbb{R}}(m, n)$. Alors la trace (somme des éléments diagonaux) du produit $(B^t A) \quad \forall\, A, B \in \mathcal{M}_{\mathbb{R}}(m, n))$ définit un produit scalaire :

$$\underline{\langle A, B\rangle = Tr(B^t A)} \Rightarrow (\mathcal{M}_{\mathbb{R}}(m, n)\langle\ ,\ \rangle)\ \underline{\text{espace Euclidien}}$$

Notons que $(B^t A) \in \mathcal{M}_n(\mathbb{R})$ donc la trace $Tr(B^t A)$ a un sens.

b) Soit $E = \mathcal{M}_{\mathbb{C}}(m, n)$. Alors la trace de $(B^* A)$ définit un produit scalaire :

$$\underline{\langle A, B\rangle = Tr(B^* A)} \text{ sur } \mathcal{M}_{\mathbb{C}}(m, n)\ (\mathcal{M}_{\mathbb{C}}(m, n),\ \langle\ ,\ \rangle)\ \underline{\text{espace Unitaire}}.$$

Exemple 6.3.

a) On considère l'espace vectoriel $\mathcal{C}[a, b]$ des fonctions numériques continues sur $[a, b] \subset \mathbb{R}$ (avec $0 < a < b$) alors on définit la forme bilinéaire symétrique :

$$\langle f, g\rangle = \int_a^b f(t)g(t)\ dt$$

qui est un produit scalaire sur $\mathcal{C}[a, b]$ donc ($\mathcal{C}[a, b],\ \langle\ ,\ \rangle$) est un espace Euclidien.

b) De même, si f et g sont des fonctions continues complexes, sur l'intervalle réel $[a, b] \subset \mathbb{R}$, on définit la forme hermitienne définie positive :

$$\langle f, g\rangle = \int_a^b f(t)\bar{g}(t)\ dt \textit{ avec } b > a \geq 0$$

C'est un produit scalaire sur $\mathcal{C}_{\mathbb{C}}[a, b]$ l'espace vectoriel des fonctions continues complexes sur l'intervalle réel $[a, b]$

$$\Rightarrow\ (\mathcal{C}_{\mathbb{C}}[a, b],\ \langle\ ,\ \rangle)\ \underline{\textit{espace unitaire}}$$

Exercice : A vérifier que dans tous ces exemples on a bien un produit scalaire.

Exemple 6.4.

Espace ℓ^2 : {Espace des suites $X = \{X_i\}_{i \in N}$ vérifiant $x_i \in \mathbb{R} \quad \sum_{i=1}^{\infty} |x_i|^2 \langle +\infty$ (Séries de carré-sommables)}

La forme bilinéaire symètrique : $\underline{\langle X, Y\rangle = \sum_{i=1}^{\infty} x_i y_i}$ définit un produit scalaire sur ℓ^2 donc ($\ell^2\ \langle\ \rangle$) est un espace Euclidien.

On a les résultats fondamentaux suivants connus sous le nom d'inégalités de "Schwartz" et resp."triangulaire" qui sont très utiles pour les applications et les démonstrations d'autres propriétés des espaces préhilbertiens.

Théorème 6.1. *Soit* $(E, \langle\ ,\ \rangle)$ *un espace préhilbertien , alors :* $\forall\ (x, y) \in E^2$

i) $|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$ *(In. de Schwartz)*

ii) $\sqrt{\langle x+y, x+y\rangle} \leq \sqrt{\langle x, x\rangle} + \sqrt{\langle y, y\rangle}$

2 Orthogonalité

Définition 6.3.

Comme en géométrie ordinaire de $\mathbb{R}^3$ ou $\mathbb{R}^2$, on dit qu'un élément $x \in E$ $(E, \langle\ \rangle$ espace préhilbertien) est orthogonal à $y \in E$ si $\langle x, y\rangle = 0$. Notation $x \perp y$.

rappel *:* en langage élémentaire $\cos\theta = 0 \quad \theta = \frac{\pi}{2}$ dans $\mathbb{R}^3$

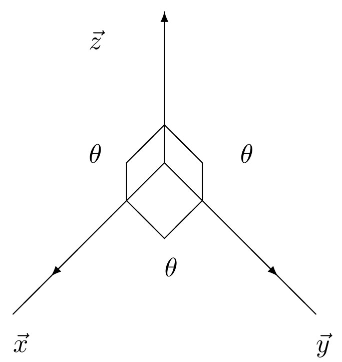

FIG. 6.1 – $\langle\vec{x}, \vec{z}\rangle = 0;\ \langle\vec{x}, \vec{y}\rangle = 0;\ \langle\vec{y}, \vec{z}\rangle = 0.$

Toujours avec l'hypothèse : $(E, \langle\ ,\ \rangle)$ espace préhilbertien.

Définition 6.4.

a) On appelle orthogonal de $x \in E$ et on note $x^{\perp}$ l'ensemble des éléments de E orthogonaux à x

$$x^{\perp} = \{y \in E\ :\ \langle x, y\rangle = 0\}$$

b) Si $M =$ sous espace vectoriel de $(E, \langle\ ,\ \rangle)$, alors l'ensemble :

$$M^{\perp} = \{y \in E :\ \langle x, y\rangle = 0 \quad \forall\, x \in M\} = \bigcap_{x \in M} x^{\perp}$$

est appelé l'orthogonal de M.

Remarque 6.2.

Si $x \in M$ et $x \in M^{\perp}$ on a $\langle x, x\rangle = 0$ et d'après la propriété du produit scalaire ceci équivaut à $x = 0$ d'où le :

Théorème 6.2. *Soit $(E, \langle\ ,\ \rangle)$ espace préhilbertien et M sous-espace vectoriel de E, alors :*

$$\underline{M \bigcap M^{\perp} = \{0\}}$$

2 Espaces Vectoriels Normés

1 Norme d'un vecteur

Définition 6.5.

Soit $x \in E$ (E espace vectoriel sur $\mathbb{K}$)
*On appelle **norme** de x sur E, toute fonction (notée $\| \cdot \|$)*

$$\| \cdot \| : \quad E \to \mathbb{R}^+$$
$$x \mapsto \|x\| \text{ qui vérifie les conditions suivantes :}$$

$\mathcal{N}.1.$ $\quad \|x\| = 0 \Leftrightarrow x = 0$

$\mathcal{N}.2.$ $\quad \|\lambda x\| = |\lambda| \, \|x\| \,, \, \forall \, (\lambda \in \mathbb{K}, \; x \in E)$

$\mathcal{N}.3.$ $\quad \|x + y\| \leq \|x\| + \|y\| \; \forall \, (x, y) \in E^2$

* *E muni d'une norme, noté $(E, \| \cdot \|)$ s'appelle espace vectoniel normé.*

Exemple 6.5.

a) $E = \mathbb{R}$. La fonction $x \mapsto |x|$ est une norme sur $\mathbb{R}$. Donc $\mathbb{R}$ muni de la norme "valeur absolue" est un espace vectoriel normé.

b) $E = \mathbb{R}^k$. On définit $N_i \; : \mathbb{R}^k \to \mathbb{R}^+, \; i = 1, 2, 3,$ par

$$\begin{array}{lcl} x & \equiv & (x_1, x_2, \ldots x_k) \xrightarrow{N_1} \sqrt{x_1^2 + x_2^2 + \ldots + x_k^2} = N_1(x) \\ x & \equiv & (x_1, x_2, \ldots x_k) \xrightarrow{N_2} |x_1| + |x_2| + \ldots + |x_k| = N_2(x) \\ x & \equiv & (x_1, x_2, \ldots x_k) \xrightarrow{N_3} \sup\{|x_1|, |x_2|, \ldots |x_k|\} = N_3(x) \end{array}$$

On montre facilement que N_1, N_2, N_3 définissent 3 normes sur $\mathbb{R}^k$, et que :

$$N_3(x) \leq N_1(x) \leq N_2(x) \leq k \, N_3(x)$$

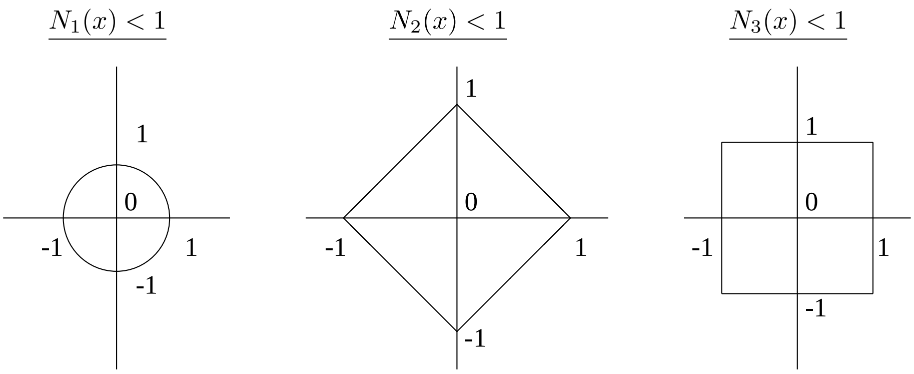

FIG. 6.2 – Représentation graphique des normes N_1, N_2, N_3 sur $\mathbb{R}^2$

Définition 6.6. *Normes équivalentes*

Deux normes N_1 et N_2 définies sur un espace vectoriel sont équivalentes ssi $\exists\, k_1, k_2 \in \mathbb{R}^{*+}$ t.q.

$$\underline{k_1 N_2 \leq \mathbf{N}_1 \leq k_2 N_2}$$

Exemple 6.6. Les normes N_1, N_2, N_3 définies sur $\mathbb{R}^k$ dans l'exemple 6.5 sont trois normes équivalentes (deux à deux).

Dans un espace préhilbertien (espace vectoriel muni d'un produit scalaire), on peut toujours définir une norme :

Définition 6.7. *Norme dans un espace préhilbertien*

Soit $(E, \langle\ ,\ \rangle)$ espace préhilbertien, alors l'application :

$$E \to \mathbb{R}^+$$
$$x \mapsto \|x\| = \sqrt{\langle x, x\rangle}$$

définit une norme sur $(E, \langle\ ,\ \rangle)$ donc $(E, \|\ \cdot\ \|)$ est un espace vectoniel normé.

Exemple 6.7. *(A vérifier)*

a)

$$E = \mathbb{C}^n \quad \forall\, u \in \mathbb{C}^n \qquad \textit{(v. ex. 1.1.2.b)}$$
$$\|u\| = \sqrt{\textstyle\int_i |u_i|^2}$$

b)

$$E = \mathcal{U}_R(m,n) \text{ avec } \langle A, B\rangle = Tr(B^T A) \qquad \textit{(v. ex. 1.2.a)}$$
$$\Rightarrow\ \forall\, A \in \mathcal{U}_R(m,n) \quad \underline{\|A\| = \sqrt{Tr(A^T A)}}$$

c)

$$E = \mathcal{M}_{\mathbb{C}}(m,n) \text{ avec } \langle A, B\rangle = Tr(B^* A) \qquad \textit{(v. ex. 1.2.b)}$$
$$\Rightarrow\ \forall\, A \in \mathcal{M}_{\mathbb{C}}(m,n) \quad \underline{\|A\| = \sqrt{Tr(A^* A)}}$$

d)

$$E = \mathcal{C}[a,b] \text{ avec } \langle f,g\rangle = \int_a^b f(t)g(t)dt \quad \text{(v. ex. 1.3.a)}$$
$$\Rightarrow \ \forall\, f \in \mathcal{C}[a,b], \underline{\|f\| = \sqrt{\int_a^b f^2(t)dt}}$$

e)

$$E = \mathcal{C}_c[a,b] \text{ avec } \langle f,g\rangle = \int_a^b f(t)\bar{g}(t)dt \quad \text{(v. ex. 1.3.b)}$$
$$\Rightarrow \ \forall\, f \in \mathcal{C}_{\mathbb{C}}[a,b], \underline{\|f\| = \sqrt{\int_a^b |f(t)|^2 dt}}$$

f)

$E = \ell^2$ *(v. ex. 1.4)* *Espace de suites, avec* $X = \{x_i\}_{i \in N}$ $\sum_i |x_i|^2 < +\infty$

avec

$$\langle X,Y\rangle = \sum_{i=1}^{\infty} x_i y_i \Rightarrow \|X\| = \sqrt{\sum_{i=1}^{\infty} |x_i|^2}$$

Chapitre 7

Normes matricielles Opérateurs Normaux - Autoadjoints - Unitaires Appl. de l'Algèbre linéaire à l'Analyse numérique

1 Systèmes Orthonormaux

1

Définition 7.1.

Soit $E, \langle\ \rangle$ un espace préhilbertien de dimension quelconque. La famille d'éléments de E :

$$\mathcal{F} = \{u_i \ \ i \in I \subset \mathbf{N} \ \ |u_i \in E, \quad \forall\, i \in \ I\}$$

est un système orthonormal pour E ssi

$$\langle u_i, u_j \rangle = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \tag{7.1.1}$$

Autrement dit les éléments de $\mathcal{F}$ sont de norme 1 et ils sont orthogonaux deux à deux. Si en plus les u_i sont linéairement indépendants et engendrent l'espace E tout entier, on dit que $\mathcal{F}$ est une base orthonormale pour E.

Etant donné une base quelconque $\mathcal{B} \subset E$, on peut toujours en construire une autre qui serait orthonormale, par le procédé d'orthogonalisation de Gram-Schmidt qu'on présente ci-dessous, ainsi que l'algorithme associé :

Théorème 7.1. *Orthonormalisation de "Gram-Schmidt"*

Soit $(E, \langle\ \rangle)$ un espace préhilbertien, et soit la suite (finie ou infinie) de vecteurs linéairement indépendants de E :

$\mathcal{U} = \{u_0, u_1, \ldots, u_n, \ldots\}$. Pour tout n soit L_n le sous-espace de E engendré par $\mathcal{U}_n = \{u_0, u_1, \ldots, u_n\}$. Si :

$\psi_0 = u_0$ et $\forall\, n \geq 1$ $\underline{\psi_{n+1} = u_{n+1} - P_{L_n}(u_{n+1})}$ où P_{L_n} est la projection orthogonale de u_{n+1} sur L_n alors :
$\Rightarrow$ La suite $\{\psi_n\}$ est un système orthogonal et pour tout n les vecteurs $\psi_0, \psi_1, \ldots, \psi_n$ engendrent L_n.

On en déduit l'algorithme pratique suivant :

Algorithme 7.1.
Algorithme d'orthonormalisation (Gram-Schmidt)
Soit $(E, \langle\ \rangle)$ espace préhilbertien.
A partir de $\mathcal{U} = \{u_0, u_1, \ldots, u_n, \ldots\}$ système de vecteurs linéairement indépendants, on construit :

a) $\psi = \{\psi_0, \psi_1, \ldots, \psi_n, \ldots\}$ *système de vecteurs orthogonaux*
et b) $\phi = \{\phi_0, \phi_1, \ldots, \phi_n, \ldots\}$ *système de vecteurs orthonormaux*

On pose

$$\psi_0 = u_0 \qquad \phi_O = \frac{\psi_0}{\|\psi_0\|}$$

et pour tout $n \geq 1$

$$\phi_n = \frac{\psi_n}{\|\psi_n\|} \text{ avec } \psi_n = u_n + \sum_{i=0}^{n-1} \lambda_{in}\phi_i$$

détermination des $\underset{j=0,1,\ldots,n-1}{\lambda_{jn}}$ *, (comme on doit vérifier l'orthogonalité) :*

$$\langle \phi_n, \phi_j\rangle = 0\ \forall\, j = 1, \ldots, n-1 \ \Leftrightarrow 0 = \langle u_n, \phi_j\rangle + \lambda_{jn} \Rightarrow \underline{\lambda_{jn} = -\langle u_n, \phi_j\rangle}$$

$\Rightarrow$

$$\boxed{\phi_0 = \frac{u_0}{\|u_0\|}\ ;\ \phi_n = \frac{u_n - \sum\limits_{j=0}^{n-1}\langle u_n, \phi_j\rangle\phi_j}{\left\| u_n - \sum\limits_{j=0}^{n-1}\langle u_n, \phi_j\rangle\phi_j \right\|} \equiv \frac{\psi_n}{\|\psi_n\|}} \tag{7.1.2}$$

Exemple 7.1. *Polynômes orthogonaux -Cas des Polynômes de Legendre*
On considère l'espace L_2 muni du produit scalaire : $\langle f, g\rangle = \int_{-1}^{+1} f(x)g(x)dx$.
Soit le système $\mathcal{U} = \{u_0(x) = 1, u_1(x) = x, u_2(x) = x^2, \ldots, u_n(x) = x^n, \ldots\}$.
Application de l'algorithme d'orthonormalisation donne :

$$u_O = 1 \text{ donc } \phi_0 = \frac{1}{\left[\int\limits_{-1}^{+1} dx\right]^{1/2}} = \frac{1}{\sqrt{2}} \tag{7.1.3}$$

$$\psi_1 = u_1 - \langle u_1, \phi_0\rangle\phi_0 = x - \int\limits_{-1}^{+1} \frac{x}{\sqrt{2}}dx\frac{1}{\sqrt{2}} = x - 0$$
$$\Rightarrow \psi_1 = x \text{ et } \phi_1 = \frac{\psi_1}{\|\psi_1\|} = \frac{x}{\left[\int\limits_{-1}^{+1} x^2dx\right]^{1/2}} = \frac{x}{\sqrt{\frac{2}{3}}} \quad \Rightarrow \phi_1 = \sqrt{\frac{3}{2}}x \tag{7.1.4}$$

En continuant de la même façon on trouve :

$$\psi_2 = x^2 - \langle x^2, \phi_1 \rangle \phi_1 - \langle x^2, \phi_0 \rangle \phi_0$$

$$\text{on a :} \quad \langle x^2, \phi_1 \rangle = \sqrt{\frac{3}{2}} \int_{-1}^{+1} x^3 dx = 0 \text{ et } \langle x^2, \phi_0 \rangle = \frac{1}{\sqrt{2}} \int_{-1}^{+1} x^2 dx = \frac{\sqrt{2}}{3}$$

$$\text{a)} \quad \Rightarrow \underline{\psi_2(x) = x^2 - \frac{1}{3}} \text{ et } \|\psi_2(x)\|^2 = \int_{-1}^{+1} \left(x^2 - \frac{1}{3} \right)^2 dx = \frac{8}{45}$$

$$\text{b)} \quad \Rightarrow \phi_2(x) = \frac{\psi_2(x)}{\|\psi_2(x)\|} = \frac{x^2 - \frac{1}{3}}{\frac{2}{3}\sqrt{\frac{2}{3}}} \Rightarrow \phi_2(x) = \sqrt{\frac{5}{2}} \, \frac{3x^2 - 1}{2} \tag{7.1.5}$$

$$\text{et} \quad \phi_3(x) = \sqrt{\frac{7}{2}} \, \frac{1}{2}(5x^3 - 3x) \quad \text{(A vérifier !)} \tag{7.1.6}$$

C'est la méthode (pas la plus rapide !) canonique pour engendrer les polynômes de Legendre normalisés :

$$\phi_n(x) = \sqrt{\frac{2n+1}{2}} P_n(x) \tag{7.1.7}$$

2 Normes $\| \; \|_p$ - Inégalité de Hölder

Présentons sans démonstration le résultat suivant

Théorème 7.2.
Soit E un espace vectoriel avec $\dim_{\mathbb{K}} E = n < \infty$.
i) Pour tout nombre réel $p \geq 1$, l'application $\| \; \|_p$ définie par :

$$\begin{aligned} \| \cdot \|_p \; : \quad & E \to \mathbb{R}^+ \\ & u \mapsto \|\nu\|_p = \left[\sum_i |\nu_i|^p \right]^{1/q} \end{aligned}$$

est une norme sur E.
ii) Soit $p > 1$ et $q > 1$ et t.q. $\frac{1}{p} + \frac{1}{q} = 1$ alors :
$\forall \, (u, \nu) \in E^2$ on a la généralisation suivante de l'inégalité Cauchy-Schwartz :

$$\sum_{i=1}^{n} |u_i \nu_i| \leq \left[\sum_{i=1}^{n} |u_i|^p \right]^{1/p} \left[\sum_{i=1}^{n} |\nu_i|^q \right]^{1/q} \tag{7.2.8}$$

(Inégalité de Hölder)

Remarque 7.1.
a) Pour $p = q = 2$ on retrouve la norme "Euclidienne" et l'inégalité Cauchy Schwartz.

b) L'inégalité de la 3ème propriété de la norme $\|\cdot\|_p$ (sous-additivité) est souvent utilisée sous le nom d'inégalité Minkowski :

$$\|u+\nu\|_p \leq \|u\|_p + \|\nu\|_p$$

$$\left[\sum_i |u_i+\nu_i|^p\right]^{1/p} \leq \left[\sum_i |u_i|^p\right]^{1/p} + \left[\sum_i |\nu_i|^p\right]^{1/p} \qquad (7.2.9)$$

c) Les normes $\|\cdot\|_p$ sont toutes équivalentes car :

Théorème 7.3.

Dans un espace de dimension finie toutes les normes sont équivalentes.

3 Normes matricielles

En ajoutant une propriété supplémentaire aux 3 propriétés fondamentales définissant une norme sur un espace vectoriel, on peut munir $\mathcal{M}_n(\mathbb{K})$ d'une norme matricielle.

Définition 7.2.

Soit $A \in \mathcal{M}_n(\mathbb{K})$. L'application $\|\cdot\| : \mathcal{M}_n(\mathbb{K}) \to \mathbb{R}^+$ est une norme matricielle si les propriétés suivantes sont vérifiées : $\forall\, A, B \in \mathcal{M}_n(\mathbb{K})$, $\forall\, \alpha \in \mathbb{K}$,

$$\begin{aligned}
&\|A\| \geq 0; \quad \textit{et} \quad \|A\| = 0 \Leftrightarrow A = 0\\
&\|\alpha A\| = |\alpha|\, \|A\|\\
&\|A+B\| \leq \|A\| + \|B\|\\
&\|AB\| \leq \|A\|\|B\|
\end{aligned}$$

Exemple 7.2.

On considère $\mathbb{C}^n$ espace vectoriel normé, muni des normes (équivalentes) vectorielles :

$$\forall\, \nu \in \mathbb{C}^n \mapsto \begin{cases} \|\nu\|_1 = \sum\limits_i |\nu_i| \\ \|\nu\|_2 = \sqrt{\sum\limits_i |\nu_i|^2} = \langle \nu, \nu\rangle^{1/2} \\ \|\nu\|_\infty = \max\limits_i |\nu_i| \end{cases}$$

alors on montre que les applications suivantes définissent des normes matricielles dans$\mathcal{M}_n(\mathbb{C})$, et $\forall\, A \in \mathcal{M}_n(\mathbb{C})$:

– *i)*

$$\|A\|_1 \equiv \sup_{\nu \neq 0} \frac{\|A\nu\|_1}{\|\nu\|_1}$$

– *ii)*

$$\|A\|_2 \equiv \sup_{\nu \neq 0} \frac{\|A\nu\|_2}{\|\nu\|_2}$$

– *iii)*

$$\|A\|_\infty \equiv \sup_{\nu \neq 0} \frac{\|A\nu\|_\infty}{\|\nu\|_\infty}$$

On présente le résultat qui donne un moyen simple pour construire une norme matricielle à partir d'une norme vectorielle.

Théorème 7.4. *(Norme matricielle subordonnée)*

Soit $\mathbb{C}^n$ l'espace vectoriel normé, (isomorphe à tout espace vectoriel sur $\mathbb{C}$, de dimension n) muni de la norme vectorielle $\|\cdot\|$,

$\Rightarrow$ L'application

$$\begin{array}{lccl} \|\cdot\|_{\mathcal{M}} : & \mathcal{M}_n(\mathbb{C}) & \to & \mathbb{R}^+ \\ & A & \mapsto & \|A\|\|_{\mathcal{M}} \end{array}$$

où

$$\|A\|_{\mathcal{M}} = \sup_{\substack{\nu \in \mathbb{C}^n \\ \|\nu\| \leq 1}} \|A\nu\| = \sup_{\substack{\nu \in \mathbb{C}^n \\ \|\nu\| = 1}} \|A\nu\|$$

(i) est une norme matricielle appelée norme matricielle subordonnée à la norme vectorielle $\|\cdot\|$

(ii) $\|A\nu\| \leq \|A\|_{\mathcal{M}}\|\nu\| \quad \forall\, \nu \in \mathbb{C}^n$.

Exemple 7.3.

1. Les normes définies par Ex.7.2 sont précisément des normes matricielles subordonnées aux normes vectorielles associées.

2. L'application $\|\cdot\|_E : \mathcal{M}_n(\mathbb{K}) \to \mathbb{R}$ définie par

$$\|A\|_E = \left\{ \sum_{ij} |a_{ij}|^2 \right\}^{1/2}$$

est une norme matricielle mais non-subordonnée à une quelconque des normes vectorielles de $\mathbb{C}^n$ (ou $E \sim \mathbb{C}^n$).

4 Adjoint d'un opérateur

1

Dans tout ce paragraphe (4) on considère un espace préhilbertien $(E, \langle\ \rangle)$ de dimension finie : $\dim_K E = n < \infty$.

Définition 7.3.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$ (opérateur linéaire ou endomorphisme sur E). On appelle opérateur adjoint de l'opérateur f et on note f^*, l' opérateur unique qui vérifie :

$$\underline{\langle f(u), \nu\rangle = \langle u, f^*(\nu)\rangle} \quad \forall\, (u, \nu) \in E^2 \tag{7.4.10}$$

Remarque 7.2.

Si l'espace préhilbertien E est de dimension infinie, on ne peut pas toujours définir l'adjoint d'un opérateur Iinéaire sur E. On montre facilement le :

Théorème 7.5.

Soient $f,\ g \in \mathcal{L}_{\mathbb{K}}(E)$ alors :

i)

$$(f+g)^* = f^* + g^* \ ;\ (\lambda f)^* = \bar{\lambda} f^* \quad \forall\, \lambda \in \mathbb{K}$$
$$(\mathbb{K} = \mathbb{R} \text{ ou } \mathbb{C})$$
$$(fg)^* = g^* f^* \ ;\ (f^*)^* = f \ ;\ rg\ f^* = rg\ f$$

ii) Si $A_f \in \mathcal{M}_n(\mathbb{C})$ est la matrice associée à f par rapport à une base orthonormale,
$\Rightarrow$ la matrice A_{f^*} associée à f^* par rapport à la même base orthononnale est la "matrice adjointe"
($\equiv$ transposée de la complexe conjuguée) de $A_f \Leftrightarrow \underline{A_{f^*} = A_f^* \overset{\text{déf}}{\equiv} \overline{A}_f^T}$

Exemple 7.4.

Soit $f \in \mathcal{L}_{\mathbb{C}}(\mathbb{C}^3)$ défini par :
$f(x,y,z) = (x + jz,\ y + 6jz,\ 4x + (2-j)y + 5z)$ alors dans la "base canonique" de $\mathbb{C}^3$ on a :

$$A_f = \begin{pmatrix} 1 & 0 & j \\ 0 & 1 & 6j \\ 4 & 2-j & 5 \end{pmatrix} \Rightarrow A_{f^*} = \overline{A}_f^T = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & 2+j \\ -j & -6j & 5 \end{pmatrix}$$

$$\Rightarrow\ f^*(x,y,z) = (x + 4z,\ y + (2+j)z,\ -jx - 6jy + 5z)$$

2 Opérateur Normal

Définition 7.4.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$. On dit que f est un opérateur normal si il commute avec son adjoint :

$$\underline{f\ f^* = f^*\ f}$$

Exemple 7.5.

Soit $f \in \mathcal{L}_{\mathbb{K}}(E)$, alors l'opérateur $f\ f^*$ (resp. $f^*\ f$ est un opérateur normal car :

$$(f\ f^*)^* = f\ f^* \Rightarrow (f\ f^*)^*(f\ f^*) = (f\ f^*)(f\ f^*)^*$$

3 Opérateur Autoadjoint (ou Hermitien)

Définition 7.5.

$f \in \mathcal{L}_{\mathbb{K}}(E)$ est un opérateur autoadjoint (ou opérateur hermitien) si il est égal à son adjoint : $f = f^*$.

Exemple 7.6.

L'opérateur $f\ f^*$ (ou f^*f) de l'ex. 7.5 est un opérateur autoadjoint puisqu'on a montré que $(f\ f^*)^* = f\ f^*$.

4 Opérateur Unitaire

Définition 7.6.

$U \in \mathcal{L}_{\mathbb{K}}(E)$ est un opérateur unitaire si il est non singulier et il vérifie :

$$U^{-1} = U^*$$

Proposition 7.1.

Dans un espace vectoriel normé, tout opérateur unitaire conserve les normes car :

$$(E, \|\ \cdot\ \|) \quad \forall\ u \in E \text{ et } U^{-1} = U^*$$
$$\Rightarrow\ \|Uu\|^2 = \langle Uu, Uu\rangle = \langle u, U^*Uu\rangle = \langle u, u\rangle = \|u\|^2$$

Remarque 7.3.

Dans les espaces Euclidiens (préhilbertiens réels car $K = \mathbb{R}$) les opérateurs autoadjoints (resp. unitaires) sont les opérateurs symétriques (resp. orthogonaux) et les matrices associées hermitiennes $A_f = A_{f^*} = A_f^*$ (resp. les matrices associées unitaires $U^{-1} = U^*$) sont les matrices symétriques $A = A^T$ (resp. orthogonales $O^{-1} = O^T$).

On peut montrer facilement le théorème suivant :

Théorème 7.6.

Dans un espace préhilbertien $(E, \langle\ \rangle)$ (resp. dans $\mathcal{M}_n(\mathbb{K})$).

*i) Tout opérateur autoadjoint (resp. toute matrice hermitienne) est un opérateur normal (resp. est une **matrice normale** $A_fA_{f^*} = A_{f^*}A_f$),*

ii) Tout opérateur unitaire (resp. toute matrice unitaire) est un opérateur normal (resp. est une matrice normale).

Résumons les résultats précédents sur le tableau ci-dessous

<table>
<tr><td colspan="2">Espaces Préhilbertiens $(E, \langle\ \rangle)$</td></tr>
<tr><td>Espaces Unitaires $\mathbb{K} = \mathbb{C}$</td><td>Espaces Euclidiens $\mathbb{K} = \mathbb{R}$</td></tr>
<tr><td colspan="2">Opérateurs, $f \in \mathcal{L}_{\mathbb{K}}(E)$ normaux $f\ f^* = f^*f$</td></tr>
<tr><td colspan="2">(Matrices $A_f \in \mathcal{M}_n(K)$ normales)</td></tr>
<tr><td>Opérateurs Autoadjoints</td><td>$\Leftrightarrow$ Opérateurs Symétriques</td></tr>
<tr><td>(Matrices Hermitiennes)</td><td>$\Leftrightarrow$ (Matrices Symétriques)</td></tr>
<tr><td>Opérateurs Unitaires</td><td>$\Leftrightarrow$ Opérateurs Orthogonaux</td></tr>
<tr><td>(Matrices Unitaires)</td><td>$\Leftrightarrow$ (Matrices Orthogonales)</td></tr>
</table>

En Physique (Mécanique Quantique) et en Analyse Numérique, on étudie et on utilise très souvent des opérateurs (ou des matrices) particuliers comme ci-dessus d'où l'importance du théorème fondamental de la réduction des matrices (ou leurs endomorphismes associés) normales, hermitiennes, symétriques, unitaires ou orthogonales, qu'on présente sans démonstration (Exercice !).

Théorème 7.7.

Etant donné une <u>matrice normale</u> $A \in \mathcal{M}_n(\mathbb{K})$,

i) Il existe une matrice unitaire $U \in \mathcal{M}_n(\mathbb{K})$, telle que la matrice :
$\tilde{A} = U^{-1}AU$ soit diagonale.

ii) a) Les éléments diagonaux de la matrice $\tilde{A}$ du i) sont les <u>valeurs propres</u> de la matrice normale A.

*b) Si A est une **matrice hermitienne** (resp. unitaire) les **valeurs propres** sont réelles (resp. **complexes** de module 1).*

iii) Les vecteurs colonnes de la matrice de passage unitaire U sont les vecteurs propres de A et forment une base orthonormale de $\mathbb{C}^n$.

iv) Si $A \in \mathcal{M}_n(\mathbb{K})$ est hermitienne les vecteurs propres associés à la même valeur propre $\lambda \in \mathbb{R}$ forment une base orthonormale du sous-espace propre associé V_λ.

5 Opérateurs Positifs

Les opérateurs hermitiens (ou matrices hermitiennes) fournissent aussi des informations sur la "positivité" des opérateurs dans un espace préhilbertien, par le théorème-définition suivant :

Théorème 7.8. *(Opérateurs Positifs)*

Soit $P \in \mathcal{L}_{\mathbb{K}}(E)$ les conditions suivantes sont équivalentes.

(a) P est positif (resp. défini positif).

(b) $P = T^2$ où T est un opérateur autoadjoint (resp. $P = T^2$ où T est autoadjoint non singulier)

(c) $P = S^*S$ pour un certain S. (resp. $P = S^*S$ où S est non singulier)

(d) P est un opérateur autoadjoint et $\langle P(u), u\rangle \geq 0 \quad \forall\, u \in E$ (resp. P autoadjoint et $\langle P(u), u\rangle > 0, \quad \forall\, u \neq \{0\} \in E$)

6 Rayon Spectral - Théorème Spectral

Définition 7.7.

On appelle rayon spectral d'un opérateurf $f \in \mathcal{L}_{\mathbb{K}}(E)$ (ou d'une matrice $A_f \in \mathcal{M}_n(K)$) le plus grand des modules des valeurs propres de f :

$$\underline{\rho(A) = \max_i\{|\lambda_i(A_f)| \perp\leq i \leq n\} \quad (= \textit{rayon spectral de } A_{f^*})}$$

On peut montrer (v. TD. 5) la propriété utile suivante pour les normes matricielles $\|A\|_2$ définies plus haut.

Théorème 7.9.

Si $A \in \mathcal{M}_n(K)$, alors :

i) $\|A\|_2 \overset{déf}{=} \dfrac{\sup \|Au\|_2}{\|u\|_2} = \sqrt{\rho(A^*A)} = \sqrt{\rho(AA^*)}$
ii) Si $\| \cdot \|$ *est une norme matricielle (subordonnée ou pas), alors :*

$$\rho(A) \leq \|A\|$$

iii) La norme $\|A\|_2$ *est invariante par transf. unitaire.*
iv) Si la matrice A *est normale alors* $\|A\|_2 = \rho(A)$.

On présente finalement un résultat important pour ses applications en Mécanique Quantique et Analyse Numérique, le théorème qui établit la décomposition d'un opérateur (ou de sa matrice associée) en projections orthogonales.

Théorème 7.10. *(Th. spectral)*
Soit f *un opérateur normal sur* $(E, \langle\,\rangle)$ *(resp. matrice normale* A_{f^*}*)*
$\Rightarrow$ *Il existe des opérateurs "projections orthogonales"* $f_1, f_2, \ldots, f_r$ *sur* E *et des scalaires* $\lambda_1, \lambda_2, \ldots, \lambda_r$ $(\in \mathbb{K})$ *tels que :*
(i) $f = \lambda_1 f_1 + \lambda_2 f_2 + \ldots \lambda_r f_r$
(ii) $f_1 + f_2 + \ldots f_r = I$
(iii) $f_i f_j = 0 \quad \forall\, i \neq j$

Exemple 7.7.
Supposons avoir effectué la réduction d'un endomorphisme f normal. Considérons la matrice associée diagonale $\tilde{A}_f$:

$$\tilde{A}_f = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2j & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

On définit :

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \; ; A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \; ; A_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

On vérifie facilement que :

$$A_i^2 = A_i \text{ (projection) et } \left\{ \begin{array}{l} A_f = 3A_1 + 2jA_2 - 2A_3 \\ \Sigma A_i = I \; ; \; A_i A_j = 0 \quad \forall\, i \neq j \end{array} \right.$$

5 Algèbre linéaire et Analyse Numérique

On présente deux des plus importants (et utiles) résultats pour les études matricielles en Analyse Numérique.

1 Valeurs Singulières ou "Diagonalisation" des matrices "Rectangulaires"

Définition 7.8.
Pour une matrice carrée $A \in \mathcal{M}_n(K)$*, on appelle valeurs singulières de* A*, les racines carrées positives des valeurs propres de l'opérateur autoadjoint* A^*A *(resp.matrice carrée hermitienne).*

On fait une extension de cette notion pour une matrice rectangulaire $A \in \mathcal{M}_{\mathbb{C}}(m,n)$ par le théorème suivant, qui fournit simultanément la diagonalisation de la matrice rectangulaire A.

Théorème 7.11.
Soit $A \in \mathcal{M}_{\mathbb{C}}(m,n)$
$\Rightarrow \exists$ une matrice unitaire $U \in \mathcal{M}_m(\mathbb{C})$ et une autre matrice unitaire $V \in \mathcal{M}_n(\mathbb{C})$ telles que :

$$U^*AV = \begin{pmatrix} \mu_1 & & 0 & 0 & \vdots & 0 \\ 0 & \mu_2 & 0 & 0 & \vdots & 0 \\ 0 & 0 & \ddots & 0 & \vdots & 0 \\ \cdots & \cdots & \cdots & \mu_r & \vdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{avec } \mu_i > 0 \ \forall\, 1 \leq i \leq r \quad \textit{et } r = rg\ A$$

Les nombres réels positifs μ_i s'appellent les valeurs singulières de la matrice A, et sont les racines carrées positives des valeurs propres de la matrice carrée autoadjointe A^*A.

2 Théor. de Gerschgorin - Hadamard

Quand on étudie la stabilité de certains systèmes et souvent en Analyse Numérique (en particulier quand la diagonalisation de certaines matrices est impossible ou très difficile) on se contente d'explorer certains voisinages D_ℓ (dans $\mathbb{C}$) des valeurs propres λ_ℓ ; ainsi au lieu de les déterminer complétement on essaie de les localiser d'une certaine manière. Le résultat suivant fournit précisément la “localisation” des valeurs propres de A. On utilise la notation $Sp(A)$ pour l'ensemble des valeurs propres (spectre) de A.

Théorème 7.12. *(Gerschgorin - Hadamard)*
Soit $A = \{a_{ij} \in \mathcal{M}_n(\mathbb{C})$. On définit les “disques de Hadamard” sur le plan $\mathbb{C}$ par :

$$D_\ell = \left\{ z \in \mathbb{C}\ ;\ |z - a_{\ell\ell}| \leq \sum_{j \neq \ell} |a_{\ell j}| \right\}$$

$$\Rightarrow Sp(A) \subset \bigcup_{\ell=1}^{n} D_\ell$$

Exemple 7.8.
Soit la matrice :

$$A = \begin{pmatrix} 7 & -2 & -1 & 0 \\ -2 & 7 & -1 & -1 \\ -1 & -1 & 7 & 2 \\ 0 & -1 & 2 & 7 \end{pmatrix}$$

d'après le théor. 7.12 toutes les valeurs propres (réelles car A symétrique réelle) se trouvent dans un “cercle” (intervalle) de centre 7 et de “rayon” 4.

Chapitre 8

Application de l'algèbre en optimisation - Programmation linéaire (Simplexe)

1 Méthode du "Simplexe" - Algorithme de Dantzig

Ce cours constitue une introduction simple aux méthodes d'optimisation connues sous le nom de **"programmation linéaire"** et plus particulièrement à la méthode de Dantzig (v.ref.G.DANTZIG) appelée autrement **méthode du "simplexe"**.

Les notions fondamentales pour ces méthodes et techniques sont d'une part la linéarité (V. Cours Algèbre 1, 4) (des fonctions objectifs à optimiser et des inégalités représentant les contraintes du problème) et d'autre part la convexité (qu'on va aborder lors d'un des cours de **topologie**). C'est la raison pour laquelle les théorèmes, les définitions et les détails de la programmation linéaire font appel aux résultats de **l'Algèbre Linéaire** (V. Cours Algèbre 1 et 4) d'une part, et aux propriétés **topologiques de la convexité** (V.Cours corres. Analyse) d'autre part.

La méthode du simplexe n'est pas récente, mais elle reste toujours moderne et pratique : Grâce à son aspect **algorithmique**, elle permet, par l'utilisation de l'ordinateur, de résoudre des problèmes très complexes en raison du nombre très grand de variables indépendantes et de contraintes. Cette méthode fait l'objet d'une étude très approfondie en deuxième année de vos études (matière d'Optimisation-Recherche Opérationnelle).

A présent, on abordera seulement les définitions et les idées fondamentales et on apprendra à appliquer la technique (sous sa forme la plus simple) des **tableaux du simplexe** (v.section 2 et3) . ainsi que la Méthode Géométrique (ou Graphique) (v.section 4) correspondante (pour le cas de deux variables indépendantes) sans aborder les techniques plus élaborées.

2 L'algorithme du simplexe

Problème $(P.O)$ (Problème initial d'optimisation)

Minimiser (ou maximiser) $f(x)$ sous les contraintes :

$$\begin{cases} g_i(x) & = & 0 & i \in I^0 & \subset \mathbf{N} \\ g_j(x) & \leq & 0 & j \in I^- & \subset \mathbf{N} \\ g_k(x) & \geq & 0 & k \in I^+ & \subset \mathbf{N},\ x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n \text{ avec } x_i \geq 0 \end{cases}$$

et, où $f,\ g_{\ell(\ell \in I^0 \cup I^- \cup I^+)}$ sont des fonctions linéaires des "variables" x_1,x_2,$\cdots$,x_n ou des formes linéaires définies sur l'espace vectoriel $\mathbb{R}^n$

1 Forme standard

Problème $(P.1)$

$$\text{Minimiser} \left(z = c \cdot x = \sum_{j=1}^{n} c_j x_j \right)$$

avec :

$$Ax = b; \quad x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n \quad x_i \in \mathbb{R}^+$$

où
n = nombre de variables indépendantes
m = nombre de contraintes.
$A \in \mathcal{M}(n, m)$, matrice des coefficients des contraintes ;
$c = (c_1, c_2\ \ldots,\ c_n)$ vecteur ligne ($\in \mathbb{R}^n$) des coûts.
$b = (b_1, \cdots, b_m)^T \in \mathbb{R}^m$ vecteur colonne des 2^{mes} membres.
$z = \sum_{j=1}^{n} c_j x_j \Leftrightarrow$ fonction à minimiser ou **"fonction objectif"**.

Proposition 8.1. *Forme standard*
*Un problème $(P.0)$ peut toujours se mettre sous **forme standard** $(P.1)$ par l'outil des "**variables d'écart**" (variables supplémentaires).*
(attention aux signes ! !)

Exemple 8.1.

$(P.O)$
$\max(f = -5x_1 + 3x_2)$

$$\textit{avec} \begin{cases} x_1 - x_2 \geq 2 \\ 2x_1 + 3x_2 \leq 4 \\ -x_1 + 6x_2 = 10 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

$\Leftrightarrow$

Forme standard $(P.1)$
$\min(z = 5x_1 - 3x_2)$

$$\textit{avec} \begin{cases} x_1 - x_2 - x_3 = 2 \\ 2x_1 + 3x_2 + x_4 = 4 \\ -x_1 + 6x_2 = 10 \\ x_1 \geq 0,\ \ \forall\, i = 1, 2, 3, 4 \\ \textit{avec : } (x_3, x_4) \textit{ variables d'écart} \end{cases}$$

Remarque 8.1.
On suppose par la suite que $rgA = m$

2 Base réalisable -Base optimale

Définition 8.1.
Polytope convexe *:* $X = \{x|Ax = b,\ x \geq 0\}$
$\Leftrightarrow$ ***Simplexe***
Polytope borné $\Leftrightarrow$ ***Polyèdre convexe***

Définition 8.2.
x ***Point extrême*** *de* X
$\Leftrightarrow$
x *ne peut pas être exprimé comme combinaison linéaire d'autres points de* X.

Définition 8.3.
Base *: toute sous matrice* $B \in \mathcal{M}_m(\mathbb{R})$ *de* A *qui a le même rang que la matrice* A *:*

$$rgA = rgB = m$$

donc :

$$A = [B, N] \ \text{ et } \ BX_B + NX_N = b$$

où X_B $\Leftrightarrow$ *l'ensemble des* ***"variables (vecteurs) de base"*** *et*
X_N $\Leftrightarrow$ *l'ensemble des* ***variables (vecteurs)"hors base"***

Définition 8.4.
Solution de base *: Elle est obtenue en posant* $X_N = 0$

$$\Rightarrow BX_B = b \ \Rightarrow \ X_B = B^{-1}b$$

Définition 8.5.
Soit B, base de $(P.1)$
Base réalisable *: si* $X_B \geq 0$ *ou si* $B^{-1}b \geq 0$

Théorème 8.1.
(i) L'ensemble des points extrêmes d'un polytope convexe X

$$\Leftrightarrow$$

L'ensemble de solutions de bases réalisables.
*(ii) L'****optimum de*** z *est atteint en au moins 1 point extrême de* X.

Théorème 8.2.
(i) Une Base de (P.1) est une ***base réalisable optimale***
ssi

$$\bar{C}_N = C_N - C_B B^{-1} \geq 0$$

($\bar{C}_N \equiv$ ***vecteur ligne des coûts réduits des variables hors base****).*
(ii) Si B *est une base réalisable quelconque, soit* x_0 *la solution correspondante. Si* $\exists\, x_h \in X_N$ *hors base t.q.* $\bar{c}_h < 0$ *alors ou bien optimum =* $-\infty$ *ou bien on met en évidence une autre base* $\hat{B}$ *(****changement de base*** *v. aussi cours 1 theorème 1.18) réalisable ayant comme solution correspondante* $\hat{x}$ *t.q*

$$z(\hat{x}) < z(x_0)$$

Remarque 8.2. Voir plus loin fig. 8.1 la représentation sous forme de "tableau" de tous ces vecteurs et matrices du simplexe.

3 Le tableau du simplexe

La méthode des tableaux du simplexe permet d'appliquer toutes les étapes de l'algorithme du simplexe sous forme d'une sequence de tableaux représentés par la figure 8.1.

Cette sequence converge vers la solution optimale lue sur le **dernier tableau** d'après le critère (i) du Théorème 8.2.

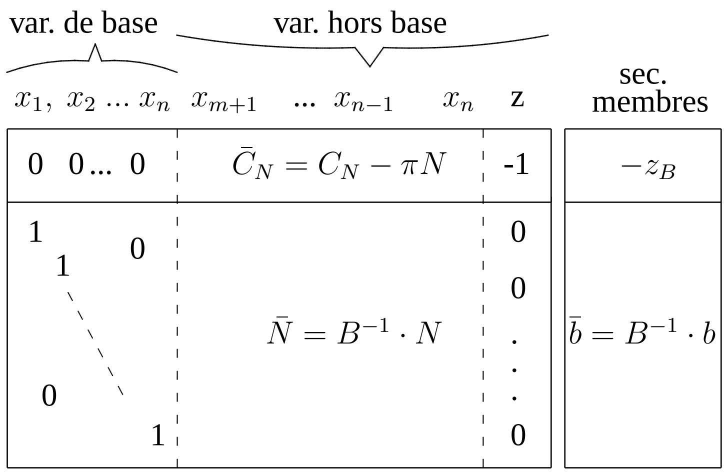

FIG. 8.1 – Schéma d'un tableau du simplexe à une étape k de l'algorithme

4 L'algorithme du simplexe $\Leftrightarrow$ la sequence des tableaux

On présente l'algorithme du simplexe (tel qu'on le retrouve dans toutes les références récentes (v. listes des références de la section 5)). Nous mettons parallèlement en évidence (entre parenthèse) son application directe sur la construction des **tableaux successifs**, du type de la figure 8.1.

Avant d' appliquer l'algorithme on doit mettre le problème d'optimisation sous forme standard avec tous les seconds membres $b_i \geq 0 \ \ \forall \ \ i = 1, 2 \dots, m$ et parallèlement avoir établi le premier tableau du simplexe sous la forme décrite par la figure précédente (v. fig. 8.1).

Soit B_0 une base réalisable de départ.

D'après cette hypothèse le tableau ci-dessus donne au départ $\bar{N} = N$, et $\bar{b} = b$.

Etapes de l' algorithme "primal" du simplexe

(a) B_0 *base réalisable de départ. Itération* $k = 0$.

($\Leftrightarrow$ 1^{er} **tableau du simplexe** et lecture de la solution de base réalisable B_0)

(b) $k \leftarrow k + 1$

($\Leftrightarrow$ **Nouveau tableau** après changement de base)

(c) *à l'itération k soit B la base suivante et $x = [x_B, x_N]$ la solution correspondante.*

Calculer : (v. fig. 8.1)

$$\left\|\begin{array}{rcl} \bar{b} & = & B^{-1}b \\ \pi & = & C_B B^{-1} \\ \bar{C}_N & = & C_N - \pi N \end{array}\right.$$

($\Leftrightarrow$ **Lecture sur le nouveau tableau** de la solution de base et du vecteur de coûts réduits $\bar{C}_N$)

(d) (1) *Si* $\bar{C}_N \geq 0, STOP$: *l'optimum est atteint.*

($\Leftrightarrow$ **Application du critère (i)** Th.8.2 du vec. de coûts réduits $\bar{C}_N$)

(2) *Si* $\exists\, x_e \in X_N$ t.q. $\bar{c}_e < 0$ *alors*

($\Leftrightarrow$ **Application du critère (ii)** Th.8.2)

(e) *Soit* A_e *la colonne* $|e|$ *de* A*. Calculer* $\bar{A}_e = B^{-1}A_e$;

si $\bar{a}_{ie} \leq 0 \;\; \forall\, i = 1, \cdots, m$ *STOP : optimum non borné* $(-\infty)$

sinon calculer : $\hat{\vartheta} = \hat{b}_s/\hat{a}_{se} = \min\limits_{\bar{a}_{ie}>0} \left\{\bar{b}_i/\bar{a}_{ie}\right\}$

$\Leftrightarrow$

Changement de base (sous-étape A) :

Variable entrante x_e t.q. $\bar{c}_e \leq \bar{c}_j \;\forall\, j = 1, 2 \ldots, n$

Variable sortante :

-Sur la colonne de x_e écrire le système d'équations de toutes les (contraintes)-lignes du tableau t.q. $\bar{a}_{ie} > 0$

ATTENTION ! ! ! *seulement* les variables de base, *doivent contribuer à ce système d'équations.*

-Evaluer le minimum des rapports des seconds membres avec les coefficients correspondants :

$$\hat{\vartheta} = \frac{\hat{b}_s}{\hat{a}_{se}} = \min_{\bar{a}_{ie}>0} \left\{ \frac{\bar{b}_i}{\bar{a}_{ie}} \right\}$$

La variable de base qui correspond à ce **minimum** $\dfrac{\hat{b}_s}{\hat{a}_{se}}$ sera la variable sortante.

-Si ce minimum n'existe pas (car $\bar{a}_{ie} \leq 0 \;\forall i = 1, 2 \ldots, m$ alors le tableau sera le dernier et la solution n'existe pas (minimum $-\infty$).)

(f) Soit x_s la variable de base correspondant à la $s^{\text{ième}}$ ligne de la base (et qui a fournit le minimum $\hat{\vartheta}$ de l'étape précédente), alors :

$$B^{-1}A_e = \hat{e}_s = s^{\text{ième}} \text{ ligne } \rightarrow \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \left\{ \begin{array}{l} x_s \\ x_e \end{array} \right.$$

avec : x_s variable sortante de la base

et : x_e variable entrante dans la base

Calcul de l'inverse de la nouvelle base et retour en (b).

(Changement de base **sous étape B**) :

-On détermine le **pivot** pour le changement de base :

C'est l'élément du tableau qui correspond à la **colonne de la variable entrante et la ligne de la variable sortante :** $\hat{a}_{se}$

- On applique l' échelonnage : (V. cours 1 Algèbre) sur les lignes du dernier tableau,

$$\left\| \begin{array}{c} L'_s = \dfrac{L_s}{\bar{a}_{se}} \\ \text{et} \\ \forall\, i = 1, 2 \ldots m+1 \quad \text{avec } m \neq s \\ L'_i = -\bar{a}_{ie} L'_s + L_i \end{array} \right.$$

et on obtient le nouveau tableau qui correspond à la nouvelle base,
et qui est le retour à l'étape b) de l'algorithme.)

5 Intérêt des "tableaux du simplexe"

L'algorithme devient plus commode avec l'usage des
"Tableaux du simplexe"8.1 car :

(1) La solution de base s'obtient par lecture directe : sur chaque ligne i du tableau (correspondant à la variable de base x_i^B) on lit $\underline{x_i^B = \bar{b}_i}$ (v.fig.8.1).

(2) La valeur $\underline{z_B}$ de la fonction objectif est contenue dans la case en haut et à droite du tableau (avec le signe -) (v.fig. 8.1)

(3) Les composantes du vecteur des coûts réduits des variables hors-base $\bar{C}_N$ sont obtenues par lecture directe de la première ligne du tableau de simplexe. Elles permettent en particulier de voir immédiatement si la solution de base courante est optimale. (rappel : il faut $\bar{C}_j \geq 0 \; \forall x_j$ hors base d'après le théorème 8.2)

3 Exemple d'application de la méthode de Dantzig par les tableaux successifs du simplexe

Soit le problème d'optimisation

$$\max(z = 8x_1 + 5x_2)$$

$$\text{avec :} \left\{ \begin{array}{rcl} 3x_1 + 6x_2 & \leq & 30 \\ 3x_1 + x_2 & \leq & 15 \\ 5x_1 + 6x_2 & \leq & 42 \\ x_1 \geq 0 & ; & x_2 \geq 0 \end{array} \right\} (P.0)$$

Etape ((O)) : Forme standard

$$\min(w = -z = -8x_1 - 5x_2)$$

$$\text{avec :} \left\{ \begin{array}{rcl} 3x_1 + 6x_2 + x_3 & = & 30 \\ 3x_1 + x_2 + x_4 & = & 15 \\ 5x_1 + 6x_2 + x_5 & = & 42 \\ \text{où } x_i \geq 0 \ ; \quad i & = & 1, 2, \cdots 5 \end{array} \right\} (P.1)$$

$\boxed{x_3, x_4, x_5}$ variables d'écart.

Etape ((1)) : 1^{er} tableau du simplexe

V. hors Base (x_1, x_2) — Var. de Base (x_3, x_4, x_5)

x_1	x_2	x_3	x_4	x_5	w	Second Membre	
-8	-5	0	0	0	-1	0	(L_1)
3	6	1	0	0	0	30	(L_2)
3	1	0	1	0	0	15	(L_3)
5	6	0	0	1	0	42	(L_4)

Base B_0

Base initiale réalisable :

$$B_0 = \{x_3; x_4; x_5\}$$

Solution de Base B_0

$$\left. \begin{array}{rcl} \tilde{x}_3 & = & 30 \\ \tilde{x}_4 & = & 15 \\ \tilde{x}_5 & = & 42 \\ \tilde{x}_1 & = & 0 \text{ car hors base} \\ \tilde{x}_2 & = & 0 \text{ car hors base} \end{array} \right\} \text{et } w = 0$$

Mais ! cette solution n'est pas optimale car :

$$\underset{-8}{\bar{C}_1 < 0} \quad \text{et} \quad \underset{-5}{\bar{C}_2 < 0}$$

Il faut changer la base :

a) variable “entrante” x_1
b) variable “sortante” ?
(trouver $\hat{\vartheta}$ qui minimise les contraintes de x_1)

$$\boxed{\text{Attention } x_2 = 0 \text{ toujours car il est hors base}}$$

$$\left\{\begin{array}{rcl} 3\vartheta + 6\times 0 + x_3 & = & 30 \\ 3\vartheta + 0 + x_4 & = & 15 \\ 5\vartheta + 0 + x_5 & = & 42 \end{array}\right. \left|\begin{array}{lcl} \Leftrightarrow \ \hat{\vartheta} & = & \min\limits_{r\geq 0}\{\frac{30}{3};\frac{15}{3};\frac{42}{5}\} \\ \Rightarrow \ \hat{\vartheta} & = & 5 \end{array}\right.$$

$$3\times 5 + x_4 = 15 \iff \underline{x_4 = 0}$$

$\Rightarrow x_4$ variable “sortante”

* Nouvelle base : $\underline{B_1 = \{x_1; x_3; x_5\}}$
* Ecrire le tableau du simplexe explicité par rapport à B_1.

Opérations sur les lignes du 1^{er} tableau ;

(il faut retrouver la matrice $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ qui correspond à B_1)

$\Rightarrow$ Echelonnage

* Pivot : l’élément de la colonne (x_1) qui correspond à la ligne (L_3) (car x_4 sort !)

donc :

$$\begin{array}{lcll} L'_3 & = & L_3/3 & \text{(pour avoir 1 à } (x_1)) \\ L'_1 & = & 8L_3/3 + L_1 & \text{(pour avoir 0 à } (x_1)) \\ L'_2 & = & -L_3 + L_2 & \text{(pour avoir 0 à } (x_1)) \\ L'_4 & = & -\dfrac{5}{3}L_3 + L_4 & \text{(pour avoir 0 à } (x_1)) \end{array}$$

2^{me} tableau du simplexe

x_1	x_2	x_3	x_4	x_5	w	Second Membre	
0	-7/3	0	8/3	0	-1	40	(L'_1)
0	5	1	-1	0	0	15	(L'_2)
1	1/3	0	1/3	0	0	5	(L'_3)
0	13/3	0	-5/3	1	0	17	(L'_4)

Base$B_1 = \{x_3; x_1; x_5\}$

Solution de Base B_1

$$\left.\begin{array}{lcl} \tilde{x}_3 & = & 15 \\ \tilde{x}_1 & = & 5 \\ \tilde{x}_5 & = & 17 \\ \tilde{x}_4 & = & 0 \text{ car hors base} \\ \tilde{x}_2 & = & 0 \text{ car hors base} \end{array}\right\} \text{ et } \underline{w = -40}$$

Mais ! cette solution n’est pas optimale car le coût réduit :

$$\bar{C}_2 = -\frac{7}{3} < 0$$

Il faut changer la base :
a) variable “entrante” x_2

b) variable “sortante” ?
(trouver $\hat{\vartheta}$ qui minimise les contraintes de x_2)

$$\begin{cases} 5\vartheta + x_3 + 0x_4 & = & 15 \\ x_1 + \vartheta/3 + 0x_3 + 0x_4 & = & 5 \\ 0x_1 + 13/3\,\vartheta + 0x_4 + x_5 & = & 17 \end{cases}$$

$$\Leftrightarrow \hat{\vartheta} = \min_{r \geq 0}\{\frac{15}{5}; 5 \times 3; \frac{17 \times 3}{13}\} = \frac{15}{5} = 3$$

$$3 \times 5 + x_3 = 15 \iff \underline{x_3 = 0}$$

$\Rightarrow x_3$ variable “sortante”

* Nouvelle base : $B_2 = (x_1; x_2; x_5)$
* Ecrire le tableau du simplexe explicité par rapport à B_2.
$\Rightarrow$ Echelonnage
* Pivot : l’élément de la colonne (x_2) qui correspond à la ligne (L_2) (car x_3 sort !)
donc :

$$\begin{array}{lcll} L'_2 & = & L_2/5 & \text{(pour avoir 1 à } (x_2)) \\ L'_1 & = & 7L_2/15 + L_1 & \text{(pour avoir 0 à } (x_2)) \\ L'_3 & = & -L_2/15 + L_3 & \text{(pour avoir 0 à } (x_2)) \\ L'_4 & = & -\frac{13}{15}L_2 + L_4 & \text{(pour avoir 0 à } (x_2)) \end{array}$$

3ème tableau du simplexe

x_1	x_2	x_3	x_4	x_5	w	Second Membre
0	0	7/15	11/5	0	-1	47
0	1	1/5	-1/5	0	0	3
1	0	-1/15	2/5	0	0	4
0	0	-13/5	-4/5	1	0	4

Solution de Base

$$\left.\begin{array}{lcl} x_1^* & = & 4 \\ x_2^* & = & 3 \\ x_5 & = & 4 \\ \tilde{x}_4 & = & 0 \text{ car hors base} \\ \tilde{x}_3 & = & 0 \text{ car hors base} \end{array}\right\} \Rightarrow -w^* = z^* = 47$$

$$\Rightarrow$$

Solution optimale car : $\boxed{\bar{C}_i \geq 0 \ \forall i}$
Fin de l’algorithme.

4 Méthode Géometrique-Cas à 2 dimensions.

1 Méthode Graphique

On présente maintenanant la **méthode Géométrique** (ou **Graphique**) (à 2 dimensions) de la **Programmation linéaire** et comparaison avec la méthode algébrique dite des tableaux du simplexe étudiée précédemment.

Problème (P.O)

$$\max(z = 8x_1 + 5x_2)$$

$$\text{avec : } \left\{ \begin{array}{lcl} 3x_1 + 6x_2 & \leq & 30 \\ 3x_1 + x_2 & \leq & 15 \\ 5x_1 + 6x_2 & \leq & 42 \\ x_1 \geq 0\,;x_2 & \geq & 0 \end{array} \right\}$$

- a) Sur le plan $(0x_1x_2)$, on trace les droites qui correspondent aux contraintes du problème :

$$\begin{array}{ll} D_1 : 3x_1 + 6x_2 = 30 & (A_1(0;5); B_1(10;0)) \\ D_2 : 3x_1 + x_2 = 15 & (A_2(0;15); B_2(5;0)) \\ D_3 : 5x_1 + 6x_2 = 42 & (A_3(0;7); B_3(6;2)) \end{array}$$

- b) La région de l'ensemble des solutions (S) v.fig. 8.2 est obtenue en vérifiant si l'origine $O = (0, 0)$ satisfait ou pas, les contraintes du problème.
  CONSEIL : Hachurer les régions interdites ! ! ! !
  Pour l'exemple présent le "simplexe" de la solution est défini par les sommets $\{O,\ A,\ K,\ B_2\}$ autrement dit : les "points extrêmes " du simplexe qui sont d'après le théorème 8.1 les solutions de bases réalisables.
- c) *Famille des droites parallèles à la fonction objectif :*

$$\underline{8x_1 + 5x_2 - z = 0}$$

On choisit le représentant pour $\underline{z = 0}$

$$8x_1 + 5x_2 = 0 : \underline{D_0} \left\{ \begin{array}{c} 0(0;0) \\ C(5;-8) \end{array} \right\}$$

La droite $D_{zmax} \parallel D_0$ passant par $\underline{K}$ fig.(8.2) (obtenue par translation parallèlement en elle même), maximise la fonction objectif z car elle a la plus grande ordonnée à l'origine parmi les droites parallèles à D_0 et

$$K = D_1 \cap D_2 \text{ donc :}$$
$$K(\vec{x}_1, \vec{x}_2) \text{ solution du système :}$$

$$\left. \begin{array}{l} 3x_1 + 6x_2 = 30 \\ 3x_1 + x_2 = 15 \end{array} \right\} \Rightarrow \begin{array}{l} \bar{x}_1 = 4 \\ \bar{x}_2 = 3 \end{array}$$

$\Rightarrow \underline{K(4;3)}$ sommet du "simplexe" OA_1KB_2 (v.fig. 8.2) solution des contraintes de (P.O).

2 Conclusions

a) $z_{max} = 8 \times 4 + 5 \times 3 = 32 + 15 = 47$
$\underline{z_{max} = 47}$ et $D_{zmax} \ : \ 8x_1 + 5x_2 = 47$
et ordonnée à l'origine de $D_{zmax} \ : \ x_2^0 = \dfrac{47}{5} = 9,4$ point $D(0;9,4)$ sur la figure 8.2

b) La solution est évidemment la même obtenue par la méthode des tableaux du simplexe (v.section 3)

c) Le déplacement de la droite représentative D_0 (parallèlement en elle-même) d'un sommet du "simplexe" à l'autre représente géométriquement le changement de bases réalisables par la méthode des tableaux du simplexe (v.section 2. et 3) .

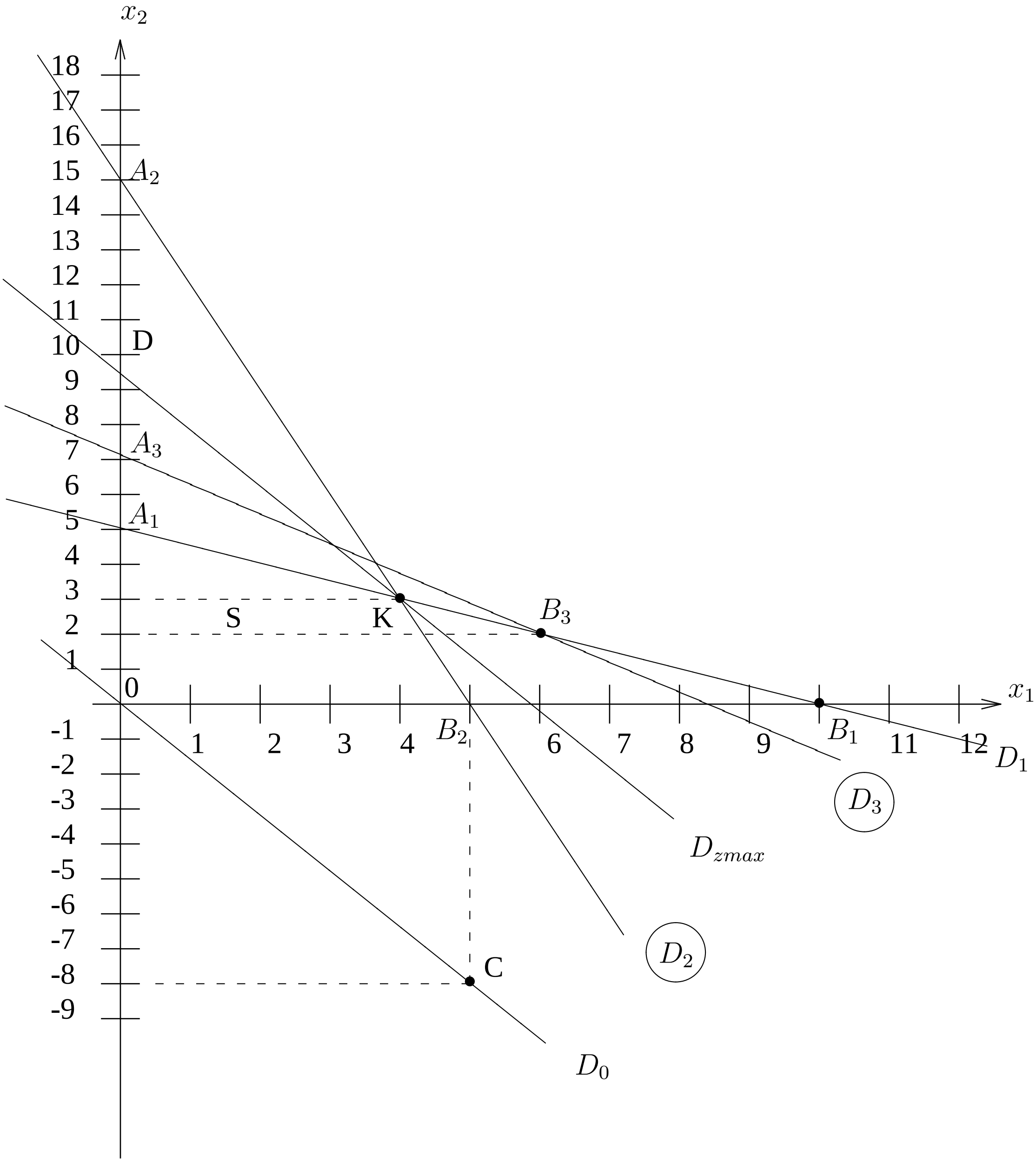

FIG. 8.2 – La solution géometrique du simplexe

5 Références

a) G.DANTZIG
"Linear programming and Extensions"
Princeton, N.J.Princeton, University Press, 1963

b) R.FAURE
"Précis de Recherche Opérationnelle ",
Dunod (Paris 1979)

c) S.GASS
"Linear Programming : Méthods and Applications", 5^{th} edition New York : Mc Graw-Hill 1985

d) M. MINOUX (1975)
Programmation Mathématique (Dunod)

e) M. MINOUX (1975)
Programmation Linéaire
(Cours de l'Ecole Nationale Supérieure des Télécommunications, Paris)

f) C. PAPADIMITRIOU and K.STEIGLITZ
"Combinatorial Optimization : Algorithms and Complexity" Englewood Cliffs , N.J. Prentice-Hall 1982

g) A. W. TUCKER
Recent advances in Mathematical Programming
(Mc GRAW-HILL, New York)

g) W.L.WINSTON
" Operations Research : Applications and Algorithmes" PWS-KENT 1991