

Mathématiques pour les Sciences expérimentales

H. A. Emamirad

Université de Poitiers

Table des matières

1	Espaces euclidien et hermitien	5
1.1	Espaces vectoriels	5
1.2	Formes linéaires et bilinéaires	6
1.3	Produit scalaire et la norme	8
1.4	Orthogonalité	11
2	Matrices	15
2.1	Transformations linéaires	15
2.2	Matrice associée à une transformation linéaire	16
2.3	Opérations sur les matrices	18
2.4	Inversibilité et le déterminant	20
2.5	Valeurs propres et vecteurs propres	26
2.6	Réduction d'une matrice	30
3	Systèmes différentiels linéaires	35
3.1	Systèmes différentiels linéaires du premier ordre	35
3.2	Matrice résolvante	37
3.3	Calcul de $\exp(tA)$	38
3.4	Formule de la variation des constantes	41
3.5	Analyse compartimentale	42
4	Séries de Fourier et applications	47
4.1	Fonctions périodiques	47
4.2	Fonctions paires et impaires	48
4.3	Séries de Fourier	49
4.4	Séries de Fourier à valeurs complexes	52
4.5	Théorèmes de convergences	53
4.6	L'identité de Parseval	54

4.7	Application à l'équation de la chaleur	55
4.8	Application à l'équation de la corde vibrante	58
5	Calcul différentiel et optimisation	61
5.1	Applications continues	61
5.2	Applications différentiable	62
5.3	Les dérivées partielles	64
5.4	Interprétation géométrique de la différentielle	67
5.5	Différentielle d'une fonction composée.	68
5.6	Dérivées partielles d'ordre supérieur.	71
5.7	Maxima Minima.	74
5.8	Multiplicateurs de Lagrange	75

Chapitre 1

Espaces euclidien et hermitien

1.1 Espaces vectoriels

Soit \mathbb{K} un corps commutatif d'éléments neutres 0 (pour l'addition) et 1 (pour la multiplication). On se restreint dans ce cours à $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

Définition 1.1.1

Soit E un ensemble muni de deux opérations : addition (+) et multiplication par un scalaire dans \mathbb{K} . L'ensemble E a une structure d'**espace vectoriel** sur \mathbb{K} , si

- (i) E est un groupe multiplicatif pour +.
- (ii) Pour tout $x, y \in E$, on a

$$\alpha(x + y) = \alpha x + \alpha y, \quad \forall \alpha \in \mathbb{K}.$$

- (iii) Pour tout $x \in E$, on a

$$\alpha(\beta x) = (\alpha\beta)x, \quad \forall \alpha, \beta \in \mathbb{K}.$$

Définition 1.1.2

Soit F un sous-ensemble d'un espace vectoriel E . On dira que F est un **sous-espace vectoriel** de E si et seulement si

- (i)

$$\forall x \in F \quad \text{et} \quad \forall y \in F, \quad \text{alors on a} \quad x - y \in F.$$

- (ii)

$$\forall x \in F \quad \text{et} \quad \forall \alpha \in \mathbb{K}, \text{ alors on a} \quad \alpha x \in F.$$

Remarquons qu'un sous-espace vectoriel de E n'est jamais vide, car il contient au moins 0 (en prenant $x = y$ dans Définition 1.1.2, (i)). Alors $\{0\}$ s'appelle **sous-espace trivial** de E et si un sous-espace vectoriel $F \neq \{0\}$, il s'appelle alors **sous-espace propre** de E .

Définition 1.1.3

Soit P une partie non-vide finie ou infinie d'un espace vectoriel E . on indexe les éléments de P par un ensemble d'indices I et on définit

$$[P] := \left\{ x \in E \mid \exists \{\lambda_i\}_{i \in I}, \lambda_i \in \mathbb{K} \text{ et } \{x_i\}_{i \in I} \subset P \text{ avec } x = \sum_{i \in I} \lambda_i x_i \right\}.$$

L'ensemble $[P]$ est un sous-espace propre de E et il s'appelle **sous-espace engendré** par P .

Définition 1.1.4

Soit E un espace vectoriel sur \mathbb{K} et $\mathcal{B} = \{e_i\}_{i \in I}$, $e_i \in E$ une famille d'éléments de E . On dira que \mathcal{B} forme une base dans E , si

- (i) La famille \mathcal{B} est **libre**, c'est-à-dire si

$$\sum_{i \in I} \alpha_i e_i = \mathbf{0} \in E, \quad \text{alors } \alpha_i = 0 \in \mathbb{K} \quad \forall i \in I. \quad (1.1)$$

- (ii) La famille \mathcal{B} est **engendrante** ou **génératrice**, c'est-à-dire si

$$\forall x \in E \quad \exists \{\alpha_i\}_{i \in I} \text{ telle que } x = \sum_{i \in I} \alpha_i e_i. \quad (1.2)$$

Dans une base \mathcal{B} si le nombre d'éléments de \mathcal{B} est fini (par exemple $= n$) on dira que E est un espace vectoriel de **dimension finie** et on écrit

$$\boxed{\dim E = n}$$

1.2 Formes linéaires et bilinéaires

Soient E un espace vectoriel sur \mathbb{K} et f une application de E dans \mathbb{K} .

Définition 1.2.1

On dira que f est une **forme linéaire** sur E si elle vérifie

- (i) Pour tout $x, y \in E$, on a $f(x + y) = f(x) + f(y)$;
(ii) Pour tout $x \in E$ et tout $\alpha \in \mathbb{K}$, on a $f(\alpha x) = \alpha f(x)$.

Si f une application de $E \times E$ dans \mathbb{K} . On dira que f est une **forme bilinéaire**, si elle vérifie

- (i) La forme $x \mapsto f(x, y)$ est linéaire pour tout $y \in E$;
(ii) La forme $y \mapsto f(x, y)$ est linéaire pour tout $x \in E$.

L'ensemble des formes linéaires définies sur un espace vectoriel E , s'appelle **dual** de E et il sera noté par E^* .

Proposition 1.2.1

Le dual E^* d'un espace vectoriel E est un espace vectoriel et si $\dim E = n$, alors $\dim E^* = n$.

Preuve. Si $f, g \in E^*$, alors

$$\begin{aligned}(f + g)(x + y) &= f(x + y) + g(x + y) \\ &= f(x) + f(y) + g(x) + g(y) \\ &= (f + g)(x) + (f + g)(y)\end{aligned}$$

et de même pour tout $\lambda \in \mathbb{K}$,

$$(\lambda f)(x) = \lambda f(x) = f(\lambda x).$$

Donc E^* est stable par l'addition et la multiplication par un scalaire. Supposons à présent que $\dim E = n$, alors il existe une base $\mathcal{E} = \{e_1, e_2, \dots, e_n\}$ qui engendre E . Pour chaque $i = 1, \dots, n$, on définit

$$f_i : E \mapsto \mathbb{K}, \quad \text{telle que} \quad f_i(e_j) = \delta_{ij},$$

où

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

est le **symbole de Kronecker**. Il est clair que si $x = \sum_{1 \leq i \leq n} x_i e_i$, alors $f_i(x) = x_i$. Montrons que $\{f_1, f_2, \dots, f_n\}$, forme une base de E^* . En effet, si

$$\beta_1 f_1 + \beta_2 f_2 + \dots + \beta_n f_n = 0,$$

alors en appliquant cette expression à e_i on trouve que $\beta_i = 0$ pour tout $i = 1, \dots, n$. Donc les f_i sont linéairement indépendantes. De plus elles sont engendrantes, car pour tout $f \in E^*$, on peut écrire $f = \sum_{1 \leq i \leq n} \beta_i f_i$ avec $\beta_i = f(e_i)$. En effet

$$\begin{aligned}f(x) &= f\left(\sum_{1 \leq i \leq n} x_i e_i\right) \\ &= \sum_{1 \leq i \leq n} x_i f(e_i) \\ &= \sum_{1 \leq i \leq n} f_i(x) f(e_i) \\ &= \left(\sum_{1 \leq i \leq n} \beta_i f_i\right)(x).\end{aligned}$$

pour tout $x \in E$. □

1.3 Produit scalaire et la norme

Supposons que \mathbb{K} soit le corps réel ($\mathbb{K} = \mathbb{R}$). On définit la forme

$$f(x, y) = \sum_{i=1}^n x_i y_i, \quad (1.3)$$

où x et y sont deux vecteurs de l'espace vectoriel E de dimension n sur \mathbb{R} et x_i et y_i sont les coordonnées de x et y dans une base $\mathcal{E} := \{e_1, e_2, \dots, e_n\}$.

Proposition 1.3.1

La forme f est une forme bilinéaire, symétrique, positive et non-dégénérée.

Preuve. Cette forme est en effet bilinéaire, car pour y fixé,

$$x \mapsto \sum_{i=1}^n x_i y_i$$

est linéaire et de même pour x fixé, l'application

$$y \mapsto \sum_{i=1}^n x_i y_i$$

est linéaire.

Elle est **symétrique**, c'est-à-dire $f(x, y) = f(y, x)$. En effet

$$f(x, y) = \sum_{i=1}^n x_i y_i = \sum_{i=1}^n y_i x_i = f(y, x).$$

Elle est **positive**, c'est-à-dire $f(x, x) \geq 0$. En effet

$$f(x, x) = \sum_{i=1}^n x_i^2 \geq 0.$$

Elle est **non-dégénérée**, c'est-à-dire

$$f(x, y) = 0 \quad \forall y \in \mathbb{R}^n \quad \implies \quad x = 0.$$

En effet, comme les coefficients de $e_i = (0, 0, \dots, \underbrace{1}_{i\text{ème colonne}}, \dots, 0)$ sont nuls apart $y_i = 1$, alors si $f(x, e_i) = 0$ on a $f(x, e_i) = x_i = 0$, pour tout $i = 1, \dots, n$. □

Définition 1.3.1

La forme bilinéaire (1.3) s'appelle **produit scalaire** et un espace vectoriel sur \mathbb{R} muni d'un produit scalaire est dit **espace euclidien**.

Pour noter le produit scalaire de deux vecteurs x et y dans E , suivant les auteurs les notations suivantes sont utilisées pour désigner $f(x, y)$,

$$\vec{x} \cdot \vec{y}, \quad (x | y), \quad \langle x, y \rangle, \quad \langle x | y \rangle.$$

Dans ce cours nous gardons la notation $f(x, y) = (x | y)$.

On peut généraliser cette définition à un espace vectoriel E de dimension finie sur \mathbb{C} . On définit alors le produit scalaire de x par y , par

$$f(x, y) = \sum_{i=1}^n x_i \overline{y_i}, \quad (1.4)$$

où \bar{z} est le conjugué de $z \in \mathbb{C}$. Avec cette définition la forme

$$\varphi_y : x \longrightarrow (x | y),$$

est bien linéaire, tandis que

$$\varphi_x : y \longrightarrow (x | y),$$

vérifie

$$\begin{aligned} \varphi_x(y_1 + y_2) &= (x | y_1 + y_2) = (x | y_1) + (x | y_2) \\ &= \varphi_x(y_1) + \varphi_x(y_2), \end{aligned}$$

mais

$$\begin{aligned} \varphi_x(\lambda y) &= (x | \lambda y) = \sum_{i=1}^n x_i (\overline{\lambda y_i}) \\ &= \bar{\lambda} \sum_{i=1}^n x_i (\overline{y_i}) = \bar{\lambda} (x | y). \end{aligned}$$

Une telle forme est qualifiée de **semi-linéaire**, par conséquent la forme

$$(x, y) \in E \times E \mapsto (x | y)$$

est qualifiée de **sesqui-linéaire** (*sesqui* en latin vaut dire 1,5).

Définition 1.3.2

Un espace vectoriel de dimension finie sur \mathbb{C} muni d'un produit scalaire de la forme (1.4) est dit **espace hermitien**.

Définition 1.3.3

Soit E un espace vectoriel de dimension finie sur \mathbb{R} ou \mathbb{C} muni d'un produit scalaire de la forme (1.3) ou (1.4). Alors on peut définir sur E une application de la forme

$$\mathcal{N} : x \mapsto \|x\| = \sqrt{(x | x)}, \quad (1.5)$$

appelée **norme** d'un vecteur x .

Lemme 1.3.1 (Inégalité de Cauchy-Schwarz)

Soit E un espace euclidien ou hermitien et x, y deux vecteurs quelconques de E , alors on a

$$|(x | y)| \leq \|x\| \|y\|. \quad (1.6)$$

Preuve. On donnera ici la démonstration dans un espace euclidien et on laissera au lecteur à titre d'exercice la démonstration dans un espace hermitien.

Soient $x, y \in \mathbb{R}^n$ deux vecteurs quelconques et pour chaque $r \in \mathbb{R}$ on a

$$\begin{aligned} 0 \leq \|x + ry\|^2 &= (x + ry | x + ry) \\ &= \|x\|^2 + 2r(x | y) + r^2\|y\|^2 = p_2(r). \end{aligned}$$

Le polynôme $p_2(r)$ est de degré 2 et positif, donc son discriminant $\Delta = (x | y) - \|x\|^2\|y\|^2 \leq 0$. D'où (1.6). \square

On obtient ainsi trois propriétés essentielles sur l'application \mathcal{N} .

$$(\mathcal{N}1) \quad \boxed{\mathcal{N}(x) = 0 \implies x = 0.}$$

En effet, si $\mathcal{N}(x) = \|x\| = 0$, alors l'inégalité de Cauchy-Schwarz implique que

$$(x | y) = 0, \quad \forall y \in E.$$

Comme la forme $f(x, y) = (x | y)$ est non-dégénérée, alors $x = 0$.

$$(\mathcal{N}2) \quad \boxed{\mathcal{N}(\lambda x) = |\lambda| \mathcal{N}(x).}$$

En effet,

$$\begin{aligned} \mathcal{N}(\lambda x) &= \sqrt{(\lambda x | \lambda x)} = \sqrt{\lambda \bar{\lambda} (x | x)} \\ &= |\lambda| \sqrt{(x | x)} = |\lambda| \mathcal{N}(x). \end{aligned}$$

$$(\mathcal{N}3) \quad \boxed{\mathcal{N}(x + y) \leq \mathcal{N}(x) + \mathcal{N}(y).}$$

En effet,

$$\begin{aligned}
 \mathcal{N}(x+y)^2 &= (x+y | x+y) \\
 &= (x | x) + (x | y) + (y | x) + (y | y) \\
 &= \mathcal{N}(x)^2 + 2\operatorname{Re}(x | y) + \mathcal{N}(y)^2 \\
 &= \mathcal{N}(x)^2 + 2|(x | y)| + \mathcal{N}(y)^2
 \end{aligned}$$

d'après (1.6),

$$\begin{aligned}
 &\leq \mathcal{N}(x)^2 + 2\mathcal{N}(x)\mathcal{N}(y) + \mathcal{N}(y)^2 \\
 &= (\mathcal{N}(x) + \mathcal{N}(y))^2
 \end{aligned}$$

En prenant la racine carrée de deux membres, on obtient ($\mathcal{N}3$).

Plus généralement sur un espace vectoriel de dimension finie ou infinie, une application \mathcal{N} , vérifiant ($\mathcal{N}1$), ($\mathcal{N}2$) et ($\mathcal{N}3$) s'appelle toujours une norme.

1.4 Orthogonalité

Définition 1.4.1

Soient x et y deux éléments d'un espace vectoriel E , on dira qu'ils sont **orthogonaux** si et seulement si $(x | y) = 0$ et on notera $x \perp y$. Si A est un sous-ensemble de E , on notera A^\perp l'ensemble

$$A^\perp := \{x \in E \mid (x | y) = 0, \quad \forall y \in A\}.$$

On dira que $S = \{x_1, x_2, \dots, x_n\}, x_i \neq 0$ forme un système **orthogonal** si et seulement si

$$(x_i | x_j) = 0 \quad \forall i \neq j$$

et ce système est dit **orthonormal** si et seulement si

$$(x_i | x_j) = \delta_{ij}$$

où δ_{ij} est le symbole de Kronecker.

Si $S = \{x_1, x_2, \dots, x_n\}$ est un système orthogonal, alors $\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n = 0$ implique que $(\sum_{1 \leq i \leq n} \alpha_i x_i | x_j) = \alpha_j \|x_j\|^2 = 0$, comme $x_j \neq 0$ alors $\alpha_j = 0$ pour tout $j = 1, \dots, n$. Ainsi

Proposition 1.4.1

Un système orthogonal S forme une base dans $[S]$ le sous-espace vectoriel engendré par S .

Définition 1.4.2

Soit F_1 et F_2 deux sous-espaces vectoriels de E , on dit que $F = F_1 \oplus F_2$ est la **somme directe** de F_1 et F_2 , si $F_1 \cap F_2 = \{0\}$ et que tout vecteur $x \in F$ peut se mettre sous la forme $x = x_1 + x_2$ avec $x_1 \in F_1$ et $x_2 \in F_2$. Dans ce cas on dira que F_1 est le **complémentaire** de F_2 dans F et on note

$$\boxed{F = F_1 \oplus F_2}$$

Si $\mathcal{E} := \{e_1, e_2, \dots, e_n\}$ est une base orthonormée d'un espace vectoriel E de dimension n , alors les scalaires $x_i = (x | e_i)$ sont appelés les **coordonnées** du vecteur x et dans ce cas

$$x = \sum_{i=1}^n x_i e_i.$$

Lorsque la base $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ est quelconque dans V_n , il existe une procédure dite **procédure d'orthonormalisation de Gram-Schmidt** pour la rendre orthonormale. Pour cela on procède de la manière suivante :

On définit

$$\begin{aligned} \bar{e}_1 &= u_1 \\ \bar{e}_2 &= u_2 + \lambda_2^1 \bar{e}_1 \\ &\vdots \\ \bar{e}_k &= u_k + \lambda_k^1 \bar{e}_1 + \dots + \lambda_k^{k-1} \bar{e}_{k-1} \\ &\vdots \\ \bar{e}_n &= u_n + \lambda_n^1 \bar{e}_1 + \dots + \lambda_n^{n-1} \bar{e}_{n-1} \end{aligned} \tag{1.7}$$

et on déterminera les coefficients λ_j^i de sorte que $\bar{\mathcal{E}} := \{\bar{e}_1, \bar{e}_2, \dots, \bar{e}_n\}$ forme une base orthogonale. En multipliant scalairement la seconde ligne par \bar{e}_1 et en imposant que $(\bar{e}_2 | \bar{e}_1) = 0$, on obtient

$$\begin{aligned} 0 &= (\bar{e}_2 | \bar{e}_1) = (u_2 + \lambda_2^1 \bar{e}_1 | \bar{e}_1) \\ &= (u_2 | \bar{e}_1) + \lambda_2^1 \|\bar{e}_1\|^2 \end{aligned}$$

Par récurrence supposons que pour tout $m = 3, 4, \dots, k-1$, on ait trouvé les coefficients $\lambda_m^1, \lambda_m^2, \dots, \lambda_m^{m-1}$ tel que le vecteur \bar{e}_m est orthogonal à $\{\bar{e}_1, \bar{e}_2, \dots, \bar{e}_{m-1}\}$. Pour que \bar{e}_k soit orthogonal à $\{\bar{e}_1, \bar{e}_2, \dots, \bar{e}_{k-1}\}$, on multiplie scalairement (1.7) par $\bar{e}_j, j = 1, \dots, k-1$ et on trouve

$$\lambda_k^j = -\frac{(u_k | \bar{e}_j)}{\|\bar{e}_j\|^2}.$$

D'où la formule de Gram-Schmidt

$$\bar{e}_k = u_k - \sum_{i=1}^{k-1} \frac{(u_k | \bar{e}_i)}{\|\bar{e}_i\|^2} \bar{e}_i. \quad (1.8)$$

Les $\{\bar{e}_k\}_{k=1, \dots, n}$ ainsi définis sont deux à deux orthogonaux pour les rendre orthonormaux il suffit de considérer le système $\{e_1, e_2, \dots, e_n\}$, avec $e_k = \bar{e}_k / \|\bar{e}_k\|$.

Proposition 1.4.2

Soit F un sous-espace vectoriel de E , alors

$$E = F \oplus F^\perp.$$

Preuve. Soit $x \in F \cap F^\perp$, alors $\|x\| = (x|x) = 0$, d'où $x = 0$. Par ailleurs si $z \in E$ et $\mathcal{E} := \{e_1, e_2, \dots, e_p\}$ est une base de F , que l'on peut la supposer orthonormale à l'aide le la procédé de Gram-Schmidt. Alors en prenant $x = \sum_{1 \leq i \leq p} (z|e_i)e_i$ et $y = z - x$, on a évidemment $z = x + y$ et pour tout e_j , on a $(y|e_j) = (z|e_j) - \sum_{1 \leq i \leq p} (z|e_i)(e_i|e_j) = (z|e_j) - (z|e_j) = 0$. Alors, pour tout $j = 1, \dots, p$, y est perpendiculaire à tout élément de la base de F , il est alors dans F^\perp . \square

Chapitre 2

Matrices

2.1 Transformations linéaires

Etant donné deux espace vectoriels E et F sur le même corps \mathbb{K} ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C}). On appelle **transformation linéaire** de E dans F toute application $T : E \mapsto F$ telle que

- $T(x + y) = T(x) + T(y), \quad \forall x, y \in E;$
- $T(\lambda x) = \lambda T(x), \quad \forall x \in E \text{ et } \forall \lambda \in \mathbb{K}.$

L'ensemble des transformation linéaire de E dans F est noté par $\mathcal{L}(E, F)$ et si $E = F$, $\mathcal{L}(E, E)$ sera noté par $\mathcal{L}(E)$. Les éléments de $\mathcal{L}(E)$ sont qualifiés d'**endomorphismes** de E .

On note par

$$\text{Ker } T = \{x \in E \mid T(x) = 0\} \quad \text{et} \quad \text{Im } T = \{y \in F \mid \exists x \in E \mid y = T(x)\}$$

$\text{Ker } T$ est appelé **noyau** de T et $\text{Im } T$ **image** de T .

Théorème 2.1.1

Ker T et Im T sont deux sous espaces vectoriels de E et F.

Preuve. Soient x_1 et x_2 deux éléments de $\text{Ker } T$, alors comme $T(x_1 - x_2) = \underbrace{T(x_1)}_{=0} - \underbrace{T(x_2)}_{=0} = 0$, alors $x_1 - x_2 \in \text{Ker } T$ et de même si $x \in \text{Ker } T$ et $\lambda \in \mathbb{K}$ alors $T(\lambda x) = \lambda T(x) = 0$. Donc $\text{Ker } T$ est un sous espace vectoriel de E .

Si $y_1, y_2 \in \text{Im } T$, alors il existe x_1 et x_2 deux éléments de E , tels que $y_1 = T(x_1)$ et $y_2 = T(x_2)$, ainsi $x_1 - x_2 \in E$ vérifie $y_1 - y_2 = T(x_1 - x_2)$, ce qui montre que $y_1 - y_2 \in \text{Im } T$. De même si $y \in \text{Im } T$ et $\lambda \in \mathbb{K}$ alors $\lambda y = T(\lambda x)$, pour un certain x qui vérifie $y = T(x)$. Donc $\text{Im } T$ est un sous espace vectoriel de F . \square

Soit $T \in \mathcal{L}(E, F)$, on appelle **rang** de T et on note $\text{rg}T$, la dimension de $\text{Im } T$.

Théorème 2.1.2

Soient E et F deux espaces vectoriels de dimensions finies et $T \in \mathcal{L}(E, F)$, alors

$$\dim(\text{Ker } T) + \text{rg}T = \dim(E).$$

Il faut remarquer que l'on n'a pas toujours $\text{Ker } T \cap \text{Im}T = \{0\}$. Prenez par exemple un endomorphisme non-nul nilpotent ($T^2 = 0$). Dans ce cas on trouvera $\{0\} \neq \text{Im}T \subset \text{Ker } T$.

Exercice 2.1

Montrer que si E est un espace vectoriel de dimension finie et si $T \in \mathcal{L}(E)$, alors $\text{Ker } T \oplus \text{Im}T = E$ si et seulement si $\text{Im } T = \text{Im } T^2$.

Définition 2.1.1

Soient E et F deux espace vectoriels et T une transformation linéaire de E dans F .

- On dit que T est une injection si $\text{Ker } T = \{0\}$.
- On dit que T est une surjection si $\text{Im } T = F$.
- On dit que T est une bijection si elle est injective et surjective. Une bijection linéaire est appelée **isomorphisme**. Si T est un isomorphisme alors il existe une application linéaire dite **inverse** de T et notée T^{-1} de F sur E telle que $T^{-1}(T(x)) = x$ pour tout $x \in E$.
- Si T est un isomorphisme et que $E = F$, alors T est qualifié d'**automorphisme**.

Théorème 2.1.3

Soient V_n un espace vectoriel de dimension n et T une transformation de V_n dans V_n , les propriétés suivantes sont équivalentes

- (a) T est une injection.
- (b) T est une surjection.
- (c) T est une bijection.

2.2 Matrice associée à une transformation linéaire

Soit V_n un espace vectoriel sur \mathbb{K} de dimension n rapporté à une base canonique

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

et V_m un autre espace vectoriel de dimension m sur \mathbb{K} rapporté à la base canonique

$$f_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad f_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad f_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad f_m = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

A toute transformation linéaire T de V_n dans V_m , nous pouvons associer une matrice A ; c'est un tableau de m ligne et n colonnes de scalaires a_{ij} , i.e.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

On dira qu'une telle matrice est **de type** (m, n) . Les coefficients a_{ij} ($i = 1, 2, \dots, m$ et $j = 1, 2, \dots, n$) sont définis par les formules :

$$T(e_i) = \sum_{j=1}^m a_{ji} f_j, \quad \text{pour } i = 1, 2, \dots, n. \quad (2.1)$$

Exemple 2.2.1

Soit $T : V_3 \mapsto V_2$ défini par $T(x, y, z) = (x + 2y + 3z, -3x - 2y - z)$. C'est une application linéaire, car

$$\begin{aligned} T(x_1 + x_2, y_1 + y_2, z_1 + z_2) &= ((x_1 + x_2) + 2(y_1 + y_2) + 3(z_1 + z_2), \\ &\quad -3(x_1 + x_2) - 2(y_1 + y_2) - (z_1 + z_2)) \\ &= (x_1 + 2y_1 + 3z_1, -3x_1 - 2y_1 - z_1) + \\ &\quad (x_2 + 2y_2 + 3z_2, -3x_2 - 2y_2 - z_2) \\ &= T(x_1, y_1, z_1) + T(x_2, y_2, z_2) \end{aligned}$$

et que

$$T(\lambda x, \lambda y, \lambda z) = \lambda(x + 2y + 3z, -3x - 2y - z) = \lambda T(x, y, z).$$

La matrice associée est une matrice de la forme

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

Pour trouver les coefficients a_{11} et a_{21} , on remplace (x, y, z) par $(1, 0, 0)$. Ce qui donne

$$T(1, 0, 0) = (1, -3) = (1)(1, 0) + (-3)(0, 1) \quad \text{d'où} \quad a_{11} = 1 \text{ et } a_{21} = -3.$$

Pour trouver les coefficients a_{12} et a_{22} , on remplace (x, y, z) par $(0, 1, 0)$. Ce qui donne

$$T(0, 1, 0) = (2, -2) = (2)(1, 0) + (-2)(0, 1) \quad \text{d'où} \quad a_{12} = 2 \text{ et } a_{22} = -2.$$

Pour trouver les coefficients a_{13} et a_{23} , on remplace (x, y, z) par $(0, 0, 1)$. Ce qui donne

$$T(0, 0, 1) = (3, -1) = (3)(1, 0) + (-1)(0, 1) \quad \text{d'où} \quad a_{13} = 3 \text{ et } a_{23} = -1.$$

D'où

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

Nous allons à présent définir quelques propriétés sur les transformations linéaires :

- (1) Si $T : V_n \mapsto V_m$ et $S : V_m \mapsto V_p$ sont deux transformations linéaires, alors $S \circ T$ définie par $S \circ T(x) = S(T(x))$ est aussi une transformation linéaire de V_n dans V_p ;
- (2) Si T est une transformation linéaire, on a obligatoirement $T(0) = 0$. En effet $T(-x) = -T(x)$ et pour $x = 0$ on a $T(0) = 0$.

2.3 Opérations sur les matrices

L'ensemble des matrices de type (m, n) sera noté par $\mathcal{M}_{m,n}(\mathbb{R})$ ou $\mathcal{M}_{m,n}(\mathbb{C})$, selon que les éléments $a_{i,j}$ sont réels ou complexes. Lorsque $m = n$, c'est-à-dire que les matrices sont carrées, alors on note $\mathcal{M}_{n,n}(\mathbb{K}) = \mathcal{M}_n(\mathbb{K})$.

Addition : Soient S et T deux transformations linéaires d'un espace vectoriel V_n de dimension n dans l'espace vectoriel V_m de dimension m . On vérifie facilement que la transformation $S + T$ définie par $(S + T)(x) = S(x) + T(x)$ est linéaire et si A , B et C sont les matrices associées respectivement à S , T et $S + T$, alors on a

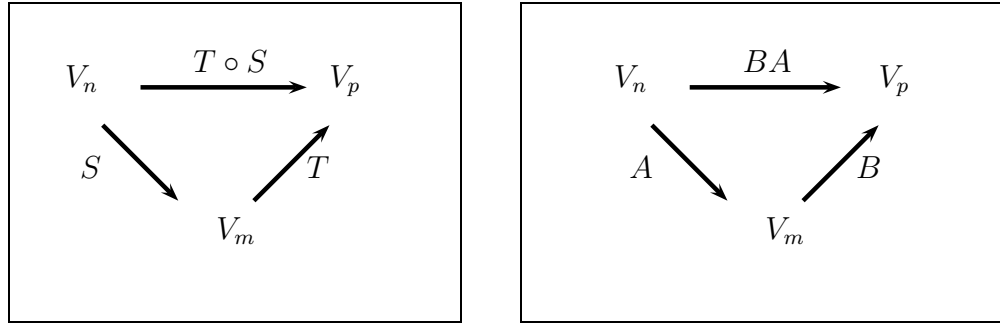
$$c_{ij} = a_{ij} + b_{ij} \quad i = 1, 2, \dots, m \quad \text{et} \quad j = 1, 2, \dots, m. \quad (2.2)$$

Multiplication par un scalaire : Soit A la matrice associée à S une transformation linéaire de V_n dans V_m et $\lambda \in \mathbb{K}$ un scalaire, on peut définir la matrice λA , dont les éléments sont λa_{ij} .

Multiplication : Soit A la matrice associée à S une transformation linéaire de V_n dans V_m et B la matrice associée à T une transformation linéaire de V_m dans V_p . La

matrice $C = BA$ associée à la transformation $T \circ S$ est une matrice de type (p, n) définie par

$$c_{ij} = \sum_{k=1}^m b_{ik}a_{kj} \quad i = 1, 2, \dots, p \quad \text{et} \quad j = 1, 2, \dots, n. \quad (2.3)$$



Dans le cas des matrices carrées $m = n$ on a

Théorème 2.3.1

L'ensemble $\mathcal{M}_n(\mathbb{K})$ forme une algèbre unitaire.

Dire qu'un ensemble $\mathcal{M}_n(\mathbb{K})$ muni de trois opérations addition

$$(A, B) \in \mathcal{M}_n(\mathbb{K}) \times \mathcal{M}_n(\mathbb{K}) \mapsto A + B \in \mathcal{M}_n(\mathbb{K}),$$

multiplication par un scalaire sur un corps \mathbb{K} ,

$$(A, \lambda) \in \mathcal{M}_n(\mathbb{K}) \times \mathbb{K} \mapsto \lambda A \in \mathcal{M}_n(\mathbb{K}),$$

et multiplication

$$(A, B) \in \mathcal{M}_n(\mathbb{K}) \times \mathcal{M}_n(\mathbb{K}) \mapsto AB \in \mathcal{M}_n(\mathbb{K}),$$

forme une algèbre vaut dire que

- $\mathcal{M}_n(\mathbb{K})$ est un espace vectoriel pour les lois d'addition et la multiplication par un scalaire.
- La multiplication est distributive par rapport à l'addition à gauche et à droite. i.e. Pour tout A, A', B et $B' \in \mathcal{M}_n(\mathbb{K})$ on a

$$A(B + B') = AB + AB'$$

$$(A + A')B = AB + A'B$$

- On dira que l'algèbre $\mathcal{M}_n(\mathbb{K})$ est **unitaire** s'il existe une matrice dite la **matrice d'identité** ou **unité** d'ordre n , I_n ($I_n = (\delta_{ij})$) où

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

est appelé **symbole de Kronecker**), telle que

$$AI_n = I_n A = A, \quad \forall A \in \mathcal{M}_n(\mathbb{K}).$$

- Attention, cette algèbre n'est pas commutative, i.e. si A et B sont deux matrices quelconques de $\mathcal{M}_n(\mathbb{K})$, alors on n'a pas nécessairement $BA = AB$.
- Attention, cette algèbre n'est pas un corps, i.e. si $A \in \mathcal{M}_n(\mathbb{K})$ il n'existe pas nécessairement une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $BA = AB = I_n$. Dans le cas où une telle matrice existe, on dira que la matrice A est **inversible** et on notera la matrice B par A^{-1} . Si la matrice A n'est pas inversible on dira que A est **singulier**.
- Si le corps $\mathbb{K} = \mathbb{C}$, alors l'espace vectoriel V_n sur lequel la transformation linéaire T est défini, est isomorphe à \mathbb{C}^n sur lequel on peut définir le produit scalaire

$$(u | v) = \sum_{i=1}^n u_i \overline{v_i},$$

où $u = (u_1, u_2, \dots, u_n)$ et $v = (v_1, v_2, \dots, v_n)$ sont les vecteurs de \mathbb{C}^n . Etant donné $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{C})$, on note $A^* \in \mathcal{M}_n(\mathbb{C})$ la matrice **adjointe** de A , définie de façon unique par la relation

$$(Au | v) = (u | A^*v), \quad \forall u, v \in \mathbb{C}^n.$$

Cette relation entraîne que

$$A^* = (a_{ij}^*) \quad \text{avec} \quad a_{ij}^* = \overline{a_{ji}}.$$

- Si le corps $\mathbb{K} = \mathbb{R}$, on dira que A^* est la matrice transposée de A et on la note ${}^t A$, dans ce cas on a

$${}^t A = ({}^t a_{ij}) \quad \text{avec} \quad {}^t a_{ij} = a_{ji}.$$

- Si $A = A^*$ on dira que A est **hermitienne** ou **auto-adjointe** et si $A = {}^t A$, on dira que A est **symétrique**.
- Une matrice A est **orthogonale** si elle est réelle et $A {}^t A = {}^t A A = I_n$.
- Une matrice A est **unitaire** si $AA^* = A^*A = I_n$.
- Une matrice A est **normale** si $AA^* = A^*A$.

2.4 Inversibilité et le déterminant

Soit \mathcal{P}_n le groupe de permutation de $\{1, 2, \dots, n\}$. Une permutation $\sigma_e \in \mathcal{P}_n$ est dite **élémentaire** si elle est de la forme

$$\sigma_e : \{1, 2, \dots, i, \dots, j, \dots, n\} \mapsto \{1, 2, \dots, j, \dots, i, \dots, n\}, \quad (2.4)$$

autrement dit, $\sigma_e(i) = j, \sigma_e(j) = i$ et $\sigma_e(k) = k$ pour tout $k \neq i$ et j . Il est clair que toute permutation $\sigma \in \mathcal{P}_n$ est la composition de k permutations élémentaires :

$$\sigma = \sigma_e^1 \circ \sigma_e^2 \circ \dots \circ \sigma_e^k. \quad (2.5)$$

On désigne par ε_σ la **signature** de σ , c'est le nombre $+1$ ou -1 selon que k le plus petit entier qui intervient dans (2.5) soit pair ou impair.

Exemple 2.4.1

Soit $n = 3$, alors \mathcal{P}_3 possède 6 éléments

$$\begin{aligned} \sigma_0 : \{1, 2, 3\} &\mapsto \{1, 2, 3\} & \sigma_3 : \{1, 2, 3\} &\mapsto \{3, 2, 1\} \\ \sigma_1 : \{1, 2, 3\} &\mapsto \{2, 1, 3\} & \sigma_4 : \{1, 2, 3\} &\mapsto \{3, 1, 2\} \\ \sigma_2 : \{1, 2, 3\} &\mapsto \{2, 3, 1\} & \sigma_5 : \{1, 2, 3\} &\mapsto \{1, 3, 2\}. \end{aligned}$$

On remarque que dans ce cas $\varepsilon_{\sigma_0} = \varepsilon_{\sigma_2} = \varepsilon_{\sigma_4} = 1$ et $\varepsilon_{\sigma_1} = \varepsilon_{\sigma_3} = \varepsilon_{\sigma_5} = -1$.

Définition 2.4.1

Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée d'ordre n de coefficients $a_{i,j}$. Le déterminant de A est un scalaire défini par

$$\det A = \sum_{\sigma \in \mathcal{P}_n} \varepsilon_\sigma a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}. \quad (2.6)$$

Exemple 2.4.2

Si $A \in \mathcal{M}_3(\mathbb{K})$, d'après l'exemple 2.4.1

$$\begin{aligned} \det(A) &= \varepsilon_{\sigma_0} a_{1,\sigma_0(1)} a_{2,\sigma_0(2)} a_{3,\sigma_0(3)} + \varepsilon_{\sigma_1} a_{1,\sigma_1(1)} a_{2,\sigma_1(2)} a_{3,\sigma_1(3)} + \varepsilon_{\sigma_2} a_{1,\sigma_2(1)} a_{2,\sigma_2(2)} a_{3,\sigma_2(3)} \\ &\quad + \varepsilon_{\sigma_3} a_{1,\sigma_3(1)} a_{2,\sigma_3(2)} a_{3,\sigma_3(3)} + \varepsilon_{\sigma_4} a_{1,\sigma_4(1)} a_{2,\sigma_4(2)} a_{3,\sigma_4(3)} + \varepsilon_{\sigma_5} a_{1,\sigma_5(1)} a_{2,\sigma_5(2)} a_{3,\sigma_5(3)} \\ &= a_{1,1} a_{2,2} a_{3,3} - a_{1,2} a_{2,1} a_{3,3} + a_{1,2} a_{2,3} a_{3,1} - a_{1,3} a_{2,2} a_{3,1} + a_{1,3} a_{2,1} a_{3,2} - a_{1,1} a_{2,3} a_{3,2}. \end{aligned}$$

En regroupant les permutations paires et impaires, on aura

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - (a_{13} a_{22} a_{31} + a_{11} a_{23} a_{32} + a_{12} a_{21} a_{33}).$$

On remarque que l'expression ci-dessus nous indique le calcul de déterminant par la **méthode de diagonale**. En effet, en mettant une copie de la matrice A à côté de A . Ainsi on obtient une matrice trois lignes et six colonnes. La somme de produit de trois premier diagonal que dans la figure ci-dessous sont indiqués par les flèches continues forment la partie paire de déterminant. Ceci doit être prise en compte avec un signe $+$. De même la somme de produit de trois premier diagonal que dans la figure ci-dessous sont indiqués par les flèches pointillées forment la partie impaire de déterminant. Ceci doit être prise en compte avec un signe $-$.

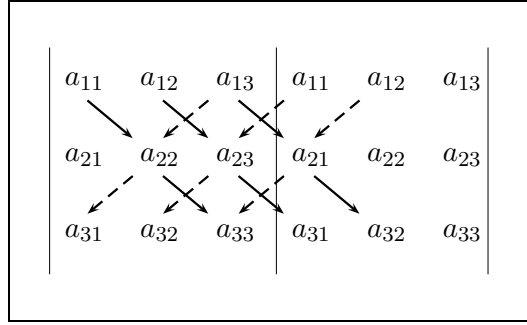


Figure 1.

Il existe une autre méthode pour calculer le déterminant d'une matrice et c'est le développement d'un déterminant suivant une ligne ou une colonne. Soient $A \in \mathcal{M}_n(\mathbb{K})$ et Δ_{ij} le **mineur relatif** au terme a_{ij} . C'est le déterminant de la matrice $A_{ij} \in \mathcal{M}_{n-1}(\mathbb{K})$ obtenue en supprimant dans A , la i -ième ligne et j -ième colonne. Alors

$$\det A = \sum_{k=1}^n (-1)^{k+j} a_{kj} \Delta_{kj} = \sum_{k=1}^n (-1)^{i+k} a_{ik} \Delta_{ik}. \quad (2.7)$$

Exemple 2.4.3

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \end{aligned}$$

Notons par \widetilde{a}_{ij} le **cofacteur** de a_{ij} , défini par $\widetilde{a}_{ij} = (-1)^{i+j} \Delta_{ij}$, alors la formule (2.7) s'écrit

$$\det A = \sum_{k=1}^n a_{kj} \widetilde{a}_{kj} = \sum_{k=1}^n a_{ik} \widetilde{a}_{ik}. \quad (2.8)$$

On peut voir que si \widetilde{A} est la matrice des cofacteurs de A , i.e. $(\widetilde{A})_{ij} = \widetilde{a}_{ij}$, alors (2.8) s'écrit

$$A {}^t \widetilde{A} = (\det A) I_n. \quad (2.9)$$

Ce qui donne dans le cas où A est inversible l'expression de l'inverse de A :

$$A^{-1} = \frac{{}^t(\widetilde{A})}{\det A}. \quad (2.10)$$

ou bien si on pose $A^{-1} = [\alpha_{ij}]$

$$\alpha_{ij} = \frac{(-1)^{i+j} \Delta_{ji}}{\det A} = \frac{\widetilde{a_{ji}}}{\det A}. \quad (2.11)$$

Ceci donne une règle très pratique pour calculer l'inverse d'une matrice (2, 2). Soit

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Calculons les cofacteurs de A

$$\widetilde{a_{11}} = (-1)^2 a_{22}, \quad \widetilde{a_{12}} = (-1)^3 a_{21}, \quad \widetilde{a_{21}} = (-1)^3 a_{12}, \quad \widetilde{a_{22}} = (-1)^4 a_{11}.$$

Donc

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Nous allons à présent indiquer quelques propriétés de déterminant.

(P1) Soit I_n la matrice identité de \mathbb{C}^n

$$\boxed{\det I_n = 1, \quad \forall n \in \mathbb{N}^*}$$

Cette propriété est évident à partir de la définition (2.6). En effet, $a_{i,\sigma(i)} = 0$ sauf si $(i, \sigma(i)) = (i, i)$ et ceci se produit seulement lorsque la permutation est identique, i.e. $\sigma_0(i) = i$. Donc il existe seulement un seul terme dans la somme et c'est le produit des éléments diagonaux qui a la signature +1.

(P2) Si $A \in \mathcal{M}_n(\mathbb{R})$ et ${}^t A$ sa transposée, alors

$$\boxed{\det(A) = (\det {}^t A)}.$$

En effet,

$$\sum_{\sigma \in \mathcal{P}_n} \varepsilon_\sigma a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} = \sum_{\sigma \in \mathcal{P}_n} \varepsilon_\sigma a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}.$$

(P3) Si $A, B \in \mathcal{M}_n(\mathbb{K})$, alors

$$\boxed{\det(AB) = (\det A)(\det B)}.$$

Cette propriété peut se démontrer à partir de (2.6) et (2.3), que l'on omettra les détails.

(P4) Soit $A \in \mathcal{M}_n(\mathbb{K})$

A est inversible, si et seulement si $\det A \neq 0$ et on a $\det A^{-1} = (\det A)^{-1}$.

Si A est inversible alors $AA^{-1} = I_n$, d'après (P3) on a $\det(AA^{-1}) = (\det A) \times (\det A^{-1}) = 1$, ce qui implique que $\det A \neq 0$. Réciproquement si $\det A \neq 0$, la formule (2.10) montre que A^{-1} existe.

A présent nous allons donner une méthode pratique pour inverser une matrice $A \in \mathcal{M}_n(\mathbb{K})$ inversible. Si $B \in \mathcal{M}_n(\mathbb{K})$ est telle que $BA = I_n$ et si $\tilde{A} = [A; I_n]$ est matrice de n lignes et $2n$ colonnes, en juxtaposant la matrice identité à droite de la matrice A . alors

$$B\tilde{A} = B[A; I_n] = [I_n; B].$$

Ainsi, comme $B = A^{-1}$ en appliquant des opérations de la méthode de Gauss sur A pour elle devient la matrice identité les mêmes opérations rendent la matrice identité en A^{-1} .

Exemple 2.4.4

Soit

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

Pour trouver l'inverse de P , on écrit

$$\begin{array}{l} \ell_2 - \ell_1 \rightarrow \ell_2 \\ \ell_3 - \ell_1 \rightarrow \ell_3 \end{array} \left| \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 & 1 & 0 \\ 1 & 0 & -2 & 0 & 0 & 1 \end{array} \right| \quad (\text{i.e. multiplier par } Q_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix})$$

$$\begin{array}{l} \ell_1 + \ell_3 \rightarrow \ell_1 \\ \ell_3 - \frac{1}{2}\ell_2 \rightarrow \ell_3 \end{array} \left| \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -2 & 0 & -1 & 1 & 0 \\ 0 & -1 & -3 & -1 & 0 & 1 \end{array} \right| \quad (\text{i.e. multiplier par } Q_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix})$$

$$\ell_1 - \frac{2}{3}\ell_3 \rightarrow \ell_1 \quad \left| \begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & 0 & 1 \\ 0 & -2 & 0 & -1 & 1 & 0 \\ 0 & 0 & -3 & -\frac{1}{2} & -\frac{1}{2} & 1 \end{array} \right| \quad (\text{i.e. multiplier par } Q_3 = \begin{pmatrix} 1 & 0 & -\frac{2}{3} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix})$$

$$\begin{array}{l} -\frac{1}{2}\ell_2 \rightarrow \ell_2 \\ -\frac{1}{3}\ell_3 \rightarrow \ell_3 \end{array} \quad \left| \begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & -2 & 0 & -1 & 1 & 0 \\ 0 & 0 & -3 & -\frac{1}{2} & -\frac{1}{2} & 1 \end{array} \right| \quad (\text{i.e. multiplier par } Q_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix})$$

$$\left| \begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & \frac{1}{6} & \frac{1}{6} & -\frac{1}{3} \end{array} \right|$$

Ainsi

$$P^{-1} = Q_4 Q_3 Q_2 Q_1 = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{6} & -\frac{1}{3} \end{pmatrix}.$$

Une autre application de déterminant c'est de reconnaître les matrices symétriques définie positive.

Définition 2.4.2

Une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est dite **définie positive** si pour tout $\mathbf{x} \in \mathbb{R}^n$, on a $(A\mathbf{x}|\mathbf{x}) > 0$, on dira qu'elle est **semi-définie positive** si $(A\mathbf{x}|\mathbf{x}) \geq 0$ et finalement **définie négative** [resp. **semi-définie négative**] si $-A$ est **définie positive** [resp. **semi-définie positive**].

Définition 2.4.3

Soit $A \in \mathcal{M}_n(\mathbb{C})$. Nous appellerons **mineur** de la matrice A , tout déterminant d'une matrice carrée extraite de A et nous appellerons **mineur principal d'ordre k** de la

matrice A , et nous le notons Δ_k , le déterminant d'une matrice carrée obtenue en éliminant les $n - k$ dernières lignes et colonnes de la matrice A .

Théorème 2.4.1

Soient $A \in \mathcal{M}_n(\mathbb{K})$, une matrice symétrique,

- (i) elle est définie positive si et seulement si tous les mineurs principaux sont strictement positifs. Autrement dit, il existe n nombres $\{a_1, \dots, a_n\}, a_k > 0, k = 1, \dots, n$ tels que $\Delta_k = (a_k)^k, k = 1, \dots, n$.
- (ii) elle est définie négative si et seulement si il existe n nombres $\{a_1, \dots, a_n\}, a_k > 0, k = 1, \dots, n$ tels que $\Delta_k = (-a_k)^k, k = 1, \dots, n$. Autrement dit, $\Delta_1 < 0, \Delta_2 > 0, \dots$

2.5 Valeurs propres et vecteurs propres

Soit A une matrice carrée d'un endomorphisme d'un espace vectoriel V_n sur \mathbb{C} . On appelle **valeur propre** de A tout scalaire $\lambda \in \mathbb{K}$ tel qu'il existe un vecteur $u \neq 0$ de V_n vérifiant

$$Au = \lambda u. \quad (2.12)$$

Le vecteur u ainsi défini s'appelle le **vecteur propre** associé à la valeur propre λ .

Il est clair que (2.12) est équivalent à $(\lambda I_n - A)u = 0$ ce qui implique que $\text{Ker}(\lambda I_n - A) \neq \{0\}$, autrement dit la matrice $\lambda I_n - A$ n'est pas injective donc, elle n'est pas inversible. ceci montre que $\det(\lambda I_n - A) = 0$.

Théorème 2.5.1

Soit A une matrice carrée d'ordre n . $P_n(\lambda) = \det(\lambda I_n - A)$ est un polynôme d'ordre n , appelé **polynôme caractéristique de A** . Les racines de ce polynôme sont les valeurs propres de la matrice A . Ainsi toute matrice possède exactement n valeurs propres dans le corps \mathbb{C} .

Définition 2.5.1

Soit $A \in \mathcal{M}_n(\mathbb{K})$, on appelle **spectre** de A et on note $\sigma(A)$ l'ensemble des valeurs propres de A . Autrement dit

$$\lambda \in \sigma(A) \iff \exists u \neq 0 \text{ tel que } Au = \lambda u.$$

Si $A = [a_{ij}]$ est une matrice triangulaire supérieure ou inférieure (i.e. $a_{ij} = 0$ si $i > j$ ou $a_{ij} = 0$ si $i < j$), alors $\lambda I_n - A$ est aussi une matrice triangulaire supérieure ou inférieure et par conséquent $\det(\lambda I_n - A) = \prod_{i=1}^n (\lambda - a_{ii})$. Ceci montre que dans ce cas les éléments

diagonaux sont les valeurs propres de la matrice A . En particulier si A est une matrice diagonale, alors

$$P_n(\lambda) = \det(\lambda I_n - A) = \prod_{i=1}^n (\lambda - a_{ii}) = \lambda^n - \alpha_1 \lambda^{n-1} + \cdots + \alpha_n. \quad (2.13)$$

où dans cette formule on a

$$\alpha_1 = \sum_{i=1}^n a_{ii} \quad \text{et} \quad \alpha_n = \det(-A) = (-1)^n \det A. \quad (2.14)$$

Comme les valeurs propres de A sont les racines de $P_n(\lambda)$, alors

$$\begin{aligned} P_n(\lambda) &= (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n) \\ &= \lambda^n - \alpha_1 \lambda^{n-1} + \cdots + \alpha_n. \end{aligned}$$

alors on a aussi

$$\alpha_1 = \sum_{i=1}^n \lambda_i \quad \text{et} \quad \alpha_n = (-1)^n \prod_{i=1}^n \lambda_i. \quad (2.15)$$

Définition 2.5.2

On appelle **trace** de la matrice A et on la note $\text{Tr}(A)$, la somme des éléments diagonaux,

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

Proposition 2.5.1

Même si deux matrices $A, B \in \mathcal{M}_n(\mathbb{K})$ ne commutent pas on a

$$\text{Tr}(AB) = \text{Tr}(BA).$$

Définition 2.5.3

Deux matrices A et $B \in \mathcal{M}_n(\mathbb{K})$, sont **semblables** si il existe une matrice inversible $Q \in \mathcal{M}_n(\mathbb{K})$, telle que $B = Q^{-1}AQ$.

Proposition 2.5.2

Supposons que deux matrices $A, B \in \mathcal{M}_n(\mathbb{K})$ sont semblable alors

- (i) leurs spectres sont égaux, i.e. $\sigma(A) = \sigma(B)$;
- (ii) Leurs polynômes caractéristiques sont identiques.

Preuve.

- (i) Si $\lambda \in \sigma(A)$, alors il existe $u \in \mathbb{K}^n$ tel que $Au = \lambda u$ et comme $Q^{-1}A = BQ^{-1}$, en appelant $v = Q^{-1}u$ on aura $Bv = BQ^{-1}u = Q^{-1}Au = Q^{-1}(\lambda u) = \lambda v$. Donc $\lambda \in \sigma(B)$ c'est-à-dire $\sigma(A) \subset \sigma(B)$. En intervertissant A et B le même argument implique que $\sigma(B) \subset \sigma(A)$;
- (ii) Comme $\lambda I - B = \lambda Q^{-1}Q - Q^{-1}AQ = Q^{-1}(\lambda I - A)Q$ on a donc,

$$\det(\lambda I - B) = \det(Q^{-1}) \det(\lambda I - A) \det(Q) = \det(\lambda I - A).$$

□

Définition 2.5.4

Soit λ une valeur propre de A . On appelle **sous-espace propre** associé à de la valeur propre λ ,

$$V(\lambda) = \{u \in E \mid Au = \lambda u\}$$

Remarquons que $V(\lambda)$ est bien un sous-espace vectoriel de E . En effet,

$$V(\lambda) = \text{Ker}(\lambda I - A).$$

Définition 2.5.5

On appelle **multiplicité algébrique** de la valeur propre λ_i , la multiplicité de la racine correspondante du polynôme caractéristique et on appelle **multiplicité géométrique** la dimension de $V(\lambda)$.

Théorème 2.5.2

Soient $A \in \mathcal{M}_n(\mathbb{K})$ et λ une valeur propre de A . Alors la multiplicité algébrique de λ est plus grand que la multiplicité géométrique de λ .

Exemple 2.5.1

(a) Soient $\alpha \in \mathbb{C}$ et

$$A_1 = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}.$$

Dans ce cas le polynôme caractéristique de A_1 est $p_3(\lambda) = \det(\lambda I_3 - A_1) = (\lambda - \alpha)^3$, donc α est la valeur propre A_1 de multiplicité algébrique 3. Maintenant si $u = (x, y, z) \in V(\alpha) = \text{Ker}(\alpha I_3 - A_1)$, alors u vérifie le système

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad (2.16)$$

ce qui donne $y = 0$ et $z = 0$, c'est-à-dire $V(\alpha) = [(1, 0, 0)]$ est un sous-espace engendré par le vecteur $(1, 0, 0)$. Ainsi la multiplicité géométrique de α est égale à 1 ;

(b) Soit

$$A_2 = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}.$$

Dans ce cas on a toujours le polynôme caractéristique est $p_3(\lambda) = (\lambda - \alpha)^3$ et α est la valeur propre A_2 de multiplicité algébrique 3. Mais cette fois le système (2.16) devient

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad (2.17)$$

et $u \in V(\alpha)$ si et seulement si $u = (x, 0, z)$, donc $V(\alpha) = [(1, 0, 0), (0, 0, 1)]$ est un sous-espace engendré par les vecteurs $(1, 0, 0)$ et $(0, 0, 1)$. Ainsi la multiplicité géométrique de α est égale à 2 ;

(c) Finalement pour

$$A_3 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix},$$

le polynôme caractéristique reste toujours $p_3(\lambda) = (\lambda - \alpha)^3$, mais $V(\alpha) = \text{Ker}(\alpha I_3 - A_1) = \text{Ker } \mathbf{0} = \mathbb{R}^3$. donc la multiplicité géométrique de α est égale à la multiplicité algébrique =3.

Théorème 2.5.3

Soient $A \in \mathcal{M}_n(\mathbb{K})$, $\{\lambda_1, \dots, \lambda_p\}$, p valeurs propres de A deux à deux distincts et $\{u_1, \dots, u_p\}$, p vecteurs propres associés. Alors on a

- (i) La famille $\{u_1, \dots, u_p\}$ est libre ;
- (ii) $F = V(\lambda_1) \oplus \dots \oplus V(\lambda_p)$ est un sous espace vectoriel de E .

Preuve.

- (i) Supposons que $\alpha_1 u_1 + \dots + \alpha_p u_p = 0$. Pour montrer que $\alpha_1 = \alpha_2 = \dots = \alpha_p = 0$, on procède par récurrence. Soit m un entier entre 1 et p . Le résultat est vrai pour $m = 1$. Supposons le vrai pour $m - 1$. c'est-à-dire :

$$\alpha_1 u_1 + \dots + \alpha_{m-1} u_{m-1} = 0 \implies \alpha_1 = \alpha_2 = \dots = \alpha_{m-1} = 0. \quad (2.18)$$

Si $\beta_1 u_1 + \dots + \beta_m u_m = 0$, en appliquant la matrice A à cette équation et en la multipliant par λ_m on obtient deux équations dont la soustraction donne $(\lambda_1 - \lambda_m)\beta_1 u_1 + (\lambda_2 - \lambda_m)\beta_2 u_2 + \dots + (\lambda_{m-1} - \lambda_m)\beta_{m-1} u_{m-1} = 0$. En prenant dans (2.18) $\alpha_k = \beta_k(\lambda_k - \lambda_m)$, on trouve que $\beta_k(\lambda_k - \lambda_m) = 0$ et comme $\lambda_k \neq \lambda_m$ pour

tout $k = 1, \dots, m - 1$, alors on a $\beta_k = 0$ pour tout $k = 1, \dots, m - 1$. En reportant dans la première relation de (2.18) on obtient aussi $\beta_m = 0$.

- (ii) Pour que la somme directe de $V(\lambda_1), \dots, V(\lambda_p)$ soit défini il faut que pour tout $u_k \in V(\lambda_k), k = 1, \dots, p$, si

$$u_1 + u_2 + \dots + u_p = 0 \quad (2.19)$$

on ait $u_k = 0$ pour tout $k = 1, \dots, p$. En effet, si ce n'est pas ainsi, (2.19) montre que les vecteurs u_k sont liés, ce qui est impossible d'après (i).

□

2.6 Réduction d'une matrice

Dans l'étude des transformations linéaires on fait souvent l'usage de la représentation matricielle pour prouver les théorèmes ou faire les calculs numériques. On sait que la représentation d'une matrice dépend de la base choisie. Si la base initiale est mal choisie on peut avoir une représentation peu malléable ou malcommode. Pour pallier ce problème on essaye de trouver une base dans laquelle une matrice a une représentation agréable. Par "représentation agréable" on entend une matrice triangulaire ou diagonale. En effet, on peut démontrer qu'il existe toujours une base dans laquelle une matrice a une forme triangulaire supérieure. Mais pour avoir la forme diagonale, on verra qu'il faut imposer certaines conditions que l'on va les invoquer dans cette section.

Théorème 2.6.1 (Changement de base)

Soient $\mathcal{E} = \{e_1, e_2, \dots, e_n\}$ et $\mathcal{F} = \{f_1, f_2, \dots, f_n\}$ deux bases différentes qu'elles engendrent un espace vectoriel E de dimension n . Soient A et B deux matrices dans $\mathcal{M}_n(\mathbb{K})$ qu'elles représentent T une transformation linéaire de E dans E dans les bases \mathcal{E} et \mathcal{F} . Alors il existe une matrice P , dite la **matrice de passage** telle que

$$B = P^{-1}AP. \quad (2.20)$$

Preuve. Soit P la matrice de passage, cette une matrice qui définit la base \mathcal{F} par rapport à la base \mathcal{E} : $P(e_i) = f_i, i = 1, \dots, n$. Alors $AP(e_i) = Af_i$ et d'après (2.1) c'est la représentation de la transformation T , mais dans la base \mathcal{E} . Pour la voir dans la base \mathcal{F} il faut multiplier les deux membre par la matrice de passage inverse P^{-1} . D'où (2.20). □

Exemple 2.6.1

Soit $T : \mathbb{R}^3 \mapsto \mathbb{R}^3$ une transformation linéaire définie par

$$T(x, y, z) = (y + z, x + z, y + x).$$

Calculons

- (i) la matrice A représentant de T dans la base canonique \mathcal{E} de \mathbb{R}^3 .
(ii) la matrice B représentant de T dans la base

$$\mathcal{F} = \{(1, 1, 1), (1, -1, 0), (1, 1, -2)\}$$

de \mathbb{R}^3 .

Dans la base canonique \mathcal{E} de \mathbb{R}^3 ,

$$T(1, 0, 0) = (0, 1, 1)$$

$$T(0, 1, 0) = (1, 0, 1)$$

$$T(0, 0, 1) = (1, 1, 0).$$

On a donc

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

Dans la base \mathcal{F} ,

$$T(f_1) = T(1, 1, 1) = (2, 2, 2) = 2f_1 + 0f_2 + 0f_3,$$

$$T(f_2) = T(1, -1, 0) = (-1, 1, 0) = 0f_1 - f_2 + 0f_3$$

$$T(f_3) = T(1, 1, -2) = (-1, -1, 2) = 0f_1 + 0f_2 - f_3.$$

On a donc

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

Ainsi on remarque la grande différence de ces deux représentations. Si on se donne la matrice A et on veut trouver la matrice B sans connaître explicitement la transformation T , alors on calcule la matrice de passage P de la base \mathcal{E} à la base \mathcal{F} , i.e.

$$P(\mathcal{E}) = \mathcal{F} \quad \text{c'est-à-dire} \quad P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix}.$$

Ce qui donne (voir Exemple 2.4.4),

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix} \quad \text{et} \quad P^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{6} & -\frac{1}{3} \end{pmatrix}.$$

Par conséquent $B = P^{-1}AP$ et on peut vérifier que

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{6} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix}.$$

Définition 2.6.1

On appelle **bloc de Jordan** de taille k , une matrice $J(\lambda, k) \in \mathcal{M}_k(\mathbb{K})$ de la forme

$$J(\lambda, k) = \begin{pmatrix} \lambda & \epsilon & 0 & \cdots & 0 \\ 0 & \lambda & \epsilon & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \epsilon \\ 0 & 0 & \cdots & \cdots & \lambda \end{pmatrix},$$

où λ un scalaire se trouve sur la diagonale et les termes sur-diagonales sont $\epsilon = 0$ ou 1 . On appelle **matrice de Jordan** une matrice diagonale par bloc, dont chaque bloc est un bloc de Jordan. Ainsi elle est de la forme

$$J = \begin{pmatrix} J(\lambda_1, k_1) & \mathbf{O} & \cdots & \mathbf{O} \\ \mathbf{O} & J(\lambda_2, k_2) & \cdots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \cdots & J(\lambda_p, k_p) \end{pmatrix}$$

Il est clair qu'une matrice de Jordan est triangulaire supérieur et si les ϵ sont tous nuls alors la matrice de Jordan est diagonale.

Théorème 2.6.2 (Théorème de Jordan)

Quelque soit $A \in \mathcal{M}_n(\mathbb{K})$, il existe une matrice de passage P telle que $J = P^{-1}AP$ est une matrice de Jordan, dont les $\lambda_1, \lambda_2, \dots, \lambda_p$ sont les valeurs propres distincts de A avec la multiplicité k_1, k_2, \dots, k_p . (Démonstration omise)

Ce théorème montre qu'en choisissant une base convenable toute matrice peut se mettre sous la forme d'une matrice triangulaire mais pour qu'elle soit **diagonalisable** il faut imposer quelques conditions supplémentaires.

Théorème 2.6.3

Pour qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ soit diagonalisable il faut et il suffit que pour chaque valeur propre λ de A , la multiplicité algébrique et géométrique de λ coïncident.

Dans l'Exemple 2.5.1 (c), on a rencontré un tel cas.

Corollaire 2.6.1

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

- (i) La matrice A est diagonalisable.
- (ii) Les n vecteurs propres de A forment une base dans \mathbb{K}^n .
- (iii) Si $\{\lambda_1, \lambda_2, \dots, \lambda_p\}$ sont p valeurs propres distinctes de A et d_i la multiplicité géométrique de chaque λ_i , ($i = 1, \dots, p$), alors $\sum_{1 \leq i \leq p} d_i = n$.

Remarquons que si $\{u_1, u_2, \dots, u_n\}$ sont n vecteurs propres de n valeurs propres $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ distinctes ou confondues et si ils forment une base de \mathbb{K}^n alors la matrice dont les colonnes sont $\{u_1, u_2, \dots, u_n\}$ est inversible et c'est exactement la matrice de passage P . Car, pour passer de la base canonique $\{e_1, e_2, \dots, e_n\}$ à la base $\{u_1, u_2, \dots, u_n\}$, on forme $PI_n = P$.

Exemple 2.6.2

Reprenons l'Exemple 2.6.1 avec

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

Le polynôme caractéristique de cette matrice est

$$\begin{aligned} P_3(\lambda) &= \begin{vmatrix} \lambda & -1 & -1 \\ -1 & \lambda & -1 \\ -1 & -1 & \lambda \end{vmatrix} \\ &= (\lambda^3 - 1 - 1) - (\lambda + \lambda + \lambda) \\ &= \lambda^3 - 3\lambda - 2 = (\lambda - 2)(\lambda + 1)^2. \end{aligned}$$

Pour trouver u_1 le vecteur propre associé à la valeur propre $\lambda_1 = 2$, il faut résoudre le système

$$\begin{aligned} y + z &= 2x \\ x + z &= 2y \\ x + y &= 2z. \end{aligned}$$

Ce qui donne $x = y = z$, d'où $u_1 = (1, 1, 1)$. Pour trouver u_2 et u_3 deux vecteurs propres associés à la valeur propre double $\lambda_2 = \lambda_3 = -1$, il faut résoudre le système

$$\begin{aligned} y + z &= -x \\ x + z &= -y \\ x + y &= -z. \end{aligned}$$

Ce qui donne $x + y + z = 0$, d'où $u_2 = (1, -1, 0)$ et $u_3 = (1, 1, -2)$. On a donc la matrice de passage

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

et on a vu dans l'Exemple 2.6.1 que

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{6} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix}.$$

Théorème 2.6.4 (Hamilton–Cayley)

Soit A une matrice carrée d'ordre n et $P_n(\lambda)$ son polynôme caractéristique. Alors en remplaçant λ par la matrice A dans $P_n(\lambda)$ on obtient la matrice identiquement nulle, i.e.

$$P_n(A) = \mathbf{0}.$$

Ce Théorème nous conduit à une formule pratique de l'inverse d'une matrice. En effet, d'après (2.13)

$$P_n(A) = A^n - \alpha_1 A^{n-1} + \alpha_2 A^{n-2} \cdots + \alpha_n I_n,$$

alors si on considère A inversible on aura $\alpha_n = (-1)^n \det(A) \neq 0$ et par conséquent

$$\begin{aligned} I_n &= \frac{-1}{\alpha_n} (A^n - \alpha_1 A^{n-1} + \alpha_2 A^{n-2} \cdots + \alpha_{n-1} A) \\ &= \frac{-A}{\alpha_n} (A^{n-1} - \alpha_1 A^{n-2} + \alpha_2 A^{n-3} \cdots + \alpha_{n-1} I_n). \end{aligned}$$

D'où

$$A^{-1} = \frac{-1}{\alpha_n} (A^{n-1} - \alpha_1 A^{n-2} + \alpha_2 A^{n-3} \cdots + \alpha_{n-1} I_n).$$

Chapitre 3

Systèmes différentiels linéaires

3.1 Systèmes différentiels linéaires du premier ordre

Un système de n équations différentielles du premier ordre est dit **linéaire**, s'il est de la forme

$$\begin{cases} x_1'(t) = a_{11}(t)x_1 + a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n + b_1(t) \\ x_2'(t) = a_{21}(t)x_1 + a_{22}(t)x_2 + \cdots + a_{2n}(t)x_n + b_2(t) \\ \vdots \\ x_n'(t) = a_{n1}(t)x_1 + a_{n2}(t)x_2 + \cdots + a_{nn}(t)x_n + b_n(t) \end{cases} \quad (3.1)$$

où l'on suppose que les fonctions $a_{ij}(t)$ et $b_i(t)$ ($i, j = 1, 2, \dots, n$) sont continues dans l'intervalle $I =]\alpha, \beta[$.

Soit $\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_n(t))$ une fonction inconnue à valeurs vectorielles, alors le système (3.1) peut se mettre sous la forme

$$\mathbf{x}' = A(t)\mathbf{x}(t) + \mathbf{b}(t) \quad (3.2)$$

où $A(t)$ est la matrice $\{a_{ij}(t)\}_{1 \leq i, j \leq n}$ et $\mathbf{b}(t) = (b_1(t), b_2(t), \dots, b_n(t))$.

Le système (3.2) s'appelle **système linéaire non-homogène** et toute équation linéaire non-homogène d'ordre n ,

$$y^{(n)}(t) + a_1(t)y^{(n-1)} + \cdots + a_n(t)y(t) = b(t) \quad (3.3)$$

peut se mettre sous cette forme. En effet en choisissant

$$x_1(t) = y(t), \quad x_2(t) = y'(t), \quad \cdots, \quad x_n(t) = y^{(n-1)}(t)$$

l'équation (3.3) peut s'écrire sous la forme d'un système du type (3.2), où

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n(t) & -a_{n-1}(t) & -a_{n-2}(t) & \cdots & -a_1(t) \end{pmatrix}, \quad \mathbf{b}(t) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}.$$

Exemple 3.1.1

(a) Considérons le système de dimension 2,

$$\begin{cases} x' = t^2x + ty + \sin t \\ y' = t^3x + t^4y + \cos t. \end{cases}$$

Il peut s'écrire sous la forme

$$\frac{d}{dt} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} t^2 & t \\ t^3 & t^4 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} + \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}.$$

(b) L'équation différentielle d'ordre 3

$$y''' + ty'' + t^2y' + t^3y = \sin t.$$

peut se mettre sous la forme

$$\frac{d}{dt} \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -t^3 & -t^2 & -t \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \sin t \end{pmatrix}.$$

Théorème 3.1.1 (Théorème d'existence et l'unicité)

Soient $A(t)$ une matrice d'ordre n et $\mathbf{b}(t)$ une fonction à valeurs vectorielles dont les coefficients $a_{ij}(t)$ et $b_i(t)$, $1 \leq i, j \leq n$ sont des fonctions continues dans $I =]\alpha, \beta[$. Alors le système (3.2) avec la donnée initiale

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad t_0 \in I \tag{3.4}$$

admet une solution unique.

Désormais nous considérons $A(t)$ une fonction matricielle continue pour tout $t \in \mathbb{R}$, de sorte que $I = \mathbb{R}$. Pour calculer la solution de (3.2) et (3.4), on considère d'abord le système homogène

$$\begin{cases} \mathbf{x}' = A(t)\mathbf{x}(t) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases} \tag{3.5}$$

3.2 Matrice résolvante

Définition 3.2.1

Chaque système de la forme (3.5) admet une **matrice résolvante** $\Phi(t, t_0)$, c'est une matrice dépendante de t et t_0 telle que la solution du système (3.5) est donnée par $\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}_0$.

Les Propriétés de la matrice résolvante :

- (1) **Propriété de la matrice identité.** Comme $\mathbf{x}_0 = \mathbf{x}(t_0) = \Phi(t_0, t_0)\mathbf{x}_0$ et ceci pour tout \mathbf{x}_0 , alors

$$\Phi(t_0, t_0) = I;$$

- (2) **Propriété de la dérivée.** Comme $\mathbf{x}(t) = \Phi(t, t_0)$ résoud le système (3.5), alors

$$\frac{\partial}{\partial t}\Phi(t, t_0) = A(t)\Phi(t, t_0); \quad (3.6)$$

- (3) **Propriété de la transitivité.**

$$\Phi(t_2, t_0) = \Phi(t_2, t_1)\Phi(t_1, t_0) \quad \forall t_0, t_1, t_2 \in \mathbb{R};$$

- (4) **Propriété de l'inversion.**

$$\Phi(t_0, t_1) = \Phi(t_1, t_0)^{-1};$$

- (5) **Propriété de la séparation.** Il existe une matrice $V(t)$ dite **matrice fondamentale** telle que

$$\Phi(t_1, t_0) = V(t_1)V(t_0)^{-1}, \quad \forall t_0, t_1 \in \mathbb{R};$$

- (6) **Propriété du déterminant.**

$$\det \Phi(t_0, t_1) = e^{\int_{t_0}^{t_1} \text{tr}(A(s))ds}.$$

Le calcul de la matrice résolvante pour le cas général (3.5), où la matrice A dépend de la variable t n'est pas dans le programme de ce cours. Néanmoins, on a le théorème suivant

Théorème 3.2.1

Lorsque $A(t)$ commute avec $A(s)$ pour tout $t, s \in \mathbb{R}$, alors on a

$$\Phi(t_0, t) = e^{\int_{t_0}^t A(s)ds}. \quad (3.7)$$

Dans ce cas la matrice fondamentale est

$$V(t) = e^{\int_0^t A(s)ds}.$$

Un cas particulière où la situation ci-dessus se présente est lorsque la matrice A est constante et dans ce cas $V(t) = \exp(\int_0^t A ds) = \exp(tA)$. Dans le paragraphe suivant on verra comment définir l'exponentiel d'une matrice.

3.3 Calcul de $\exp(tA)$

Dans le cas où la fonction matricielle $A(t)$ est constante par rapport à t , on dira que le système

$$\begin{cases} \mathbf{x}' = A\mathbf{x}(t) \\ \mathbf{x}(0) = \mathbf{x}_0 \end{cases} \quad (3.8)$$

est *autonome*. On sait que la solution de

$$\begin{cases} x' = ax(t) \\ x(0) = x_0 \end{cases}$$

est donnée par $x(t) = x_0 e^{at}$. De la même manière que

$$e^{at} = \sum_{n=0}^{\infty} \frac{(at)^n}{n!}, \quad \forall t \in \mathbb{R}$$

Théorème 3.3.1

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ et pour tout $t \in \mathbb{R}$, la série

$$\exp(tA) = \sum_{n=0}^{\infty} \frac{t^n A^n}{n!} = I + tA + \frac{t^2}{2} A^2 + \dots \quad (3.9)$$

converge.

La démonstration est omise, car pour démontrer ce Théorème il faut définir la norme matricielle qui n'est pas dans le programme de ce cours.

Si A est une matrice constante $V(t) = \exp\left(\int_0^t A ds\right) = \exp(tA)$. On peut vérifier que $\Phi(t, t_0) = V(t)V(t_0)^{-1} = e^{(t-t_0)A}$ est la matrice résolvante et elle vérifie toutes les propriétés de cette matrice. Ainsi par unicité $\exp(tA)$ est la matrice fondamentale du système (3.8). Pour calculer $\exp(tA)$, on traite d'abord le cas d'une matrice diagonale

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Dans ce cas on a

$$\begin{aligned} \exp(\Lambda) &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} + \begin{pmatrix} t\lambda_1 & 0 & \cdots & 0 \\ 0 & t\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t\lambda_n \end{pmatrix} + \begin{pmatrix} \frac{t^2}{2}\lambda_1^2 & 0 & \cdots & 0 \\ 0 & \frac{t^2}{2}\lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{t^2}{2}\lambda_n^2 \end{pmatrix} + \cdots \\ &= \begin{pmatrix} e^{t\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{t\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{t\lambda_n} \end{pmatrix} \end{aligned}$$

Traïtons à présent le cas où la matrice A est diagonalisable à l'aide d'une matrice de passage P . Dans ce cas il existe une matrice diagonale Λ telle que $A = P\Lambda P^{-1}$. Remarquons que

$$\begin{aligned} A^2 &= P\Lambda P^{-1}P\Lambda P^{-1} = P\Lambda^2 P^{-1}, \\ A^3 &= P\Lambda^2 P^{-1}P\Lambda P^{-1} = P\Lambda^3 P^{-1}, \\ &\vdots \\ A^n &= P\Lambda^{n-1} P^{-1}P\Lambda P^{-1} = P\Lambda^n P^{-1} \end{aligned}$$

et par conséquent

$$\begin{aligned} \exp(tA) &= I + tA + \frac{t^2}{2}A^2 + \cdots \\ &= PP^{-1} + tP\Lambda P^{-1} + \frac{t^2}{2}P\Lambda^2 P^{-1} + \cdots \\ &= Pe^{t\Lambda} P^{-1}. \end{aligned}$$

La même formule est valable lorsque $A(t)$ et $A(s)$ commutent pour tout $t, s \in \mathbb{R}$ et que la matrice $\Lambda(t) = P^{-1}A(t)P$ est diagonale, alors on a

$$V(t) = Pe^{\int_0^t \Lambda(s) ds} P^{-1} \quad \text{et} \quad \Phi(t, t_0) = Pe^{\int_{t_0}^t \Lambda(s) ds} P^{-1}.$$

Exemple 3.3.1

Calculons $\exp(tA)$ où

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

Pour cela les valeurs propres de A sont les racines de polynôme caractéristique $P(\lambda) = \det(\lambda I - A) = \lambda^2 + 2\lambda + \frac{1}{2}$ qui sont $\lambda_1 = -1 + \frac{\sqrt{2}}{2}$ et $\lambda_2 = -1 - \frac{\sqrt{2}}{2}$. Comme ces racines

sont distinctes la matrice A est diagonalisable par intermédiaire de la matrice de passage P où

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

Pour trouver l'inverse de la matrice P , on se réfère à la formule

$$\begin{pmatrix} a, b \\ c, d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d, -b \\ -c, a \end{pmatrix}$$

d'où

$$P^{-1} = \begin{pmatrix} 1 & 1 \\ 1 + \sqrt{2} & 1 - \sqrt{2} \end{pmatrix}^{-1} = \frac{\sqrt{2}}{4} \begin{pmatrix} \sqrt{2} - 1 & 1 \\ \sqrt{2} + 1 & -1 \end{pmatrix}.$$

Ainsi on vérifie que

$$D = P^{-1}AP = \begin{pmatrix} -1 + \frac{\sqrt{2}}{2}, 0 \\ 0, -1 - \frac{\sqrt{2}}{2} \end{pmatrix}.$$

Cette diagonalisation nous permet de calculer

$$\begin{aligned} \exp(tA) &= P \exp(t\Lambda) P^{-1} \\ &= \frac{\sqrt{2}}{4} \begin{pmatrix} 1 & 1 \\ 1 + \sqrt{2} & 1 - \sqrt{2} \end{pmatrix} \begin{pmatrix} e^{(-1 + \frac{\sqrt{2}}{2})t} & 0 \\ 0 & e^{(-1 - \frac{\sqrt{2}}{2})t} \end{pmatrix} \begin{pmatrix} \sqrt{2} - 1 & 1 \\ \sqrt{2} + 1 & -1 \end{pmatrix} \\ &= \frac{\sqrt{2}}{4} \begin{pmatrix} e^{(-1 + \frac{\sqrt{2}}{2})t} & e^{(-1 - \frac{\sqrt{2}}{2})t} \\ (1 + \sqrt{2})e^{(-1 + \frac{\sqrt{2}}{2})t} & (1 - \sqrt{2})e^{(-1 - \frac{\sqrt{2}}{2})t} \end{pmatrix} \begin{pmatrix} \sqrt{2} - 1 & 1 \\ \sqrt{2} + 1 & -1 \end{pmatrix} \\ &= e^{-t} \begin{pmatrix} \cosh \frac{\sqrt{2}}{2}t - \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t & \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t \\ \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t & \cosh \frac{\sqrt{2}}{2}t + \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t \end{pmatrix}. \end{aligned}$$

On peut en déduire immédiatement que la solution du système

$$\begin{cases} x' &= -\frac{3}{2}x + \frac{1}{2}y \\ y' &= \frac{1}{2}x - \frac{1}{2}y \end{cases}$$

avec les données initiales

$$x(0) = 0, \quad y(0) = 1$$

est donnée par

$$\begin{aligned} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} &= e^{-t} \begin{pmatrix} \cosh \frac{\sqrt{2}}{2}t - \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t & \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t \\ \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t & \cosh \frac{\sqrt{2}}{2}t + \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= e^{-t} \begin{pmatrix} \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t \\ \cosh \frac{\sqrt{2}}{2}t + \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t \end{pmatrix} \end{aligned}$$

3.4 Formule de la variation des constantes

Nous allons considérer à présent le cas d'un système non-homogène de type (3.2). Ce système avec une condition initiale s'écrit sous la forme

$$\begin{cases} \mathbf{x}' = A(t)\mathbf{x}(t) + \mathbf{b}(t) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases} \quad (3.10)$$

La méthode de la variation des constantes consiste à considérer la solution de ce système sous la forme de $\mathbf{x}(t) = \Phi(t, t_0)\mathbf{c}(t)$ où \mathbf{c} est une fonction vectorielle de la variable t et $\mathbf{c}(t_0) = I\mathbf{c}(t_0) = \Phi(t_0, t_0)\mathbf{c}(t_0) = \mathbf{x}(t_0) = \mathbf{x}_0$. En dérivant cette solution on trouve que $\mathbf{x}'(t) = \Phi'(t, t_0)\mathbf{c}(t) + \Phi(t, t_0)\mathbf{c}'(t)$. D'après la propriété de la dérivée, on sait que $\Phi'(t, t_0) = A(t)\Phi(t, t_0)$. D'où $\mathbf{x}'(t) = A(t)\Phi(t, t_0)\mathbf{c}(t) + \Phi(t, t_0)\mathbf{c}'(t) = A(t)\mathbf{x}(t) + \Phi(t, t_0)\mathbf{c}'(t)$ et en identifiant avec (3.10), on obtient $\mathbf{b}(t) = \Phi(t, t_0)\mathbf{c}'(t)$. En vertu de la propriété d'inversion, la matrice résolvante $\Phi(t, t_0)$ est inversible et on a $\mathbf{c}'(t) = \Phi(t_0, t)\mathbf{b}(t)$. En intégrant on obtient $\mathbf{c}(t) = \int_{t_0}^t \Phi(t_0, s)\mathbf{b}(s)ds + \mathbf{c}(t_0)$, ainsi en tenant compte de la condition initiale $\mathbf{c}(t_0) = \mathbf{x}_0$ on obtient $\mathbf{x}(t) = \Phi(t, t_0)[\mathbf{x}_0 + \int_0^t \Phi(t_0, s)\mathbf{b}(s)ds]$, ce qui donne

Formule de la variation des constantes *La solution du système (3.10) est donnée par la formule*

$$\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}_0 + \int_0^t \Phi(t, s)\mathbf{b}(s)ds. \quad (3.11)$$

Dans le cas où A est une matrice indépendante de t , on peut écrire

$$\mathbf{x}(t) = e^{tA}\mathbf{x}_0 + \int_0^t e^{(t-s)A}\mathbf{b}(s)ds. \quad (3.12)$$

Exemple 3.4.1

Contrairement au cas homogène, dans le cas non-homogène même avec une donnée initiale zéro, si $\mathbf{b}(t) \neq 0$ on obtient une solution non-triviale. Considérons le cas non-homogène du exemple 3.3.1 :

$$\begin{cases} x' &= -\frac{3}{2}x + \frac{1}{2}y + \alpha e^{-t} \\ y' &= \frac{1}{2}x - \frac{1}{2}y \end{cases}$$

avec les données initiales

$$x(0) = 0, \quad y(0) = 0.$$

D'après la formule (3.12)

$$\begin{aligned} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} &= e^{-t} \begin{pmatrix} \cosh \frac{\sqrt{2}}{2}t - \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t & \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t \\ \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t & \cosh \frac{\sqrt{2}}{2}t + \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}t \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &+ \int_0^t e^{-(t-s)} \begin{pmatrix} \cosh \frac{\sqrt{2}}{2}(t-s) - \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}(t-s) & \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}(t-s) \\ \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}(t-s) & \cosh \frac{\sqrt{2}}{2}(t-s) + \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}(t-s) \end{pmatrix} \begin{pmatrix} \alpha e^{-s} \\ 0 \end{pmatrix} ds \\ &= \alpha e^{-t} \int_0^t \begin{pmatrix} \cosh \frac{\sqrt{2}}{2}(t-s) - \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}(t-s) \\ \frac{\sqrt{2}}{2} \sinh \frac{\sqrt{2}}{2}(t-s) \end{pmatrix} ds \end{aligned}$$

Comme

$$\begin{aligned} \int_0^t \cosh \frac{\sqrt{2}}{2}(t-s) ds &= \left[-\sqrt{2} \sinh \frac{\sqrt{2}}{2}(t-s) \right]_0^t = \sqrt{2} \sinh \frac{\sqrt{2}}{2}t \\ \int_0^t \sinh \frac{\sqrt{2}}{2}(t-s) ds &= \left[-\sqrt{2} \cosh \frac{\sqrt{2}}{2}(t-s) \right]_0^t = \sqrt{2} \left(\cosh \frac{\sqrt{2}}{2}t - 1 \right), \end{aligned}$$

alors

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \alpha e^{-t} \begin{pmatrix} \sqrt{2} \sinh \frac{\sqrt{2}}{2}t - \cosh \frac{\sqrt{2}}{2}t + 1 \\ \cosh \frac{\sqrt{2}}{2}t - 1 \end{pmatrix}$$

3.5 Analyse compartimentale

Cette technique de modélisation est très utilisée en biologie et en médecine. On s'en sert pour suivre l'évolution, au cours du temps, de substances biochimiques (médicaments, hormones,...) et de leurs métabolites. On peut envisager l'utilisation d'analyse compartimentale dans bien d'autres circonstances. C'est ainsi, par exemple, que la gestion d'ateliers, où des produits (non nécessairement des substances chimiques) vont passer d'un atelier à l'autre, relève de l'analyse compartimentale. A chaque produit on associe un compartiment numéroté de 1 à n et on suppose qu'il existe des liens à l'intérieur des compartiments et entre le milieu extérieur et les compartiments. Ainsi les **compartiments** d'un système est une suite finie de classe définies par leurs propriétés physiques ou biologiques.

Pour décrire les principes de l'analyse compartimentale il faut d'abord imposer les hypothèses sur la nature des échanges. Si les échanges dépendent du temps on obtient un système non-autonome et dans le cas où les échanges du système sont indépendants du temps on obtient un système autonome. Selon que le système soit **ouvert** (c'est-à-dire en liaison avec le milieu extérieur) ou **fermé** (sans liaison avec le milieu extérieur) on a un système non-homogène ou homogène.

Par exemple, dans l'étude du cycle de l'eau sur la terre, on sait que trois formes sont possibles : l'eau liquide, l'eau solide (glacier) et l'eau sous la forme de gaz (vapeur

d'eau). Cela conduit à définir trois compartiments que l'on représente généralement par des disques.

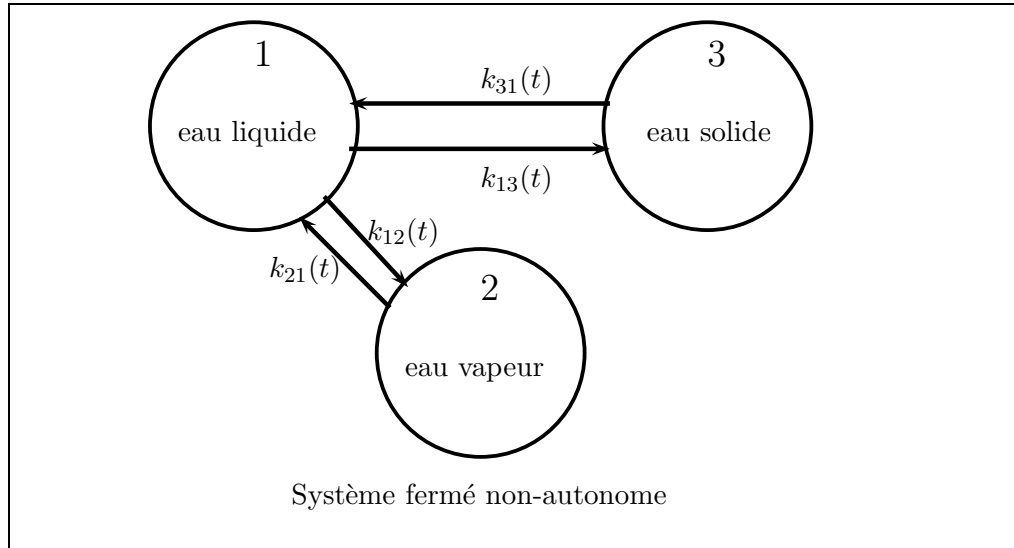


Figure 1.

Sous l'influence de paramètres climatiques, des "échanges" vont avoir lieu entre ces différents compartiments, c'est ainsi que ces échanges sont temporels, ils varient selon des saisons. Ainsi, dans le système de la figure 1, nous sommes dans le cadre d'un système fermé non-autonome.

On notera dans la suite $x_i(t)$ la quantité, à l'instant t , contenue dans le compartiment i . Pour faciliter l'exposé nous allons désormais considérer seulement de l'analyse compartimentale linéaire autonome. Cela implique que la quantité allant du compartiment i vers le compartiment j par unité de temps est égale à : $k_{ij}x_i(t)$. La constante k_{ij} s'appelle la constante de proportionnalité ; c'est une constante positive ayant deux indices i et j . Le premier indice i désigne toujours le compartiment du départ et le second indice le compartiment d'arrivée. A chaque système compartimental on attribue un schéma constitué par des cercles numérotés 1 à n . Lorsque $k_{ij} > 0$, il existe une flèche entre le compartiment i et le compartiment j dans le sens de i vers j et lorsque $k_{ij} = 0$ ces deux compartiments ne sont pas liés.

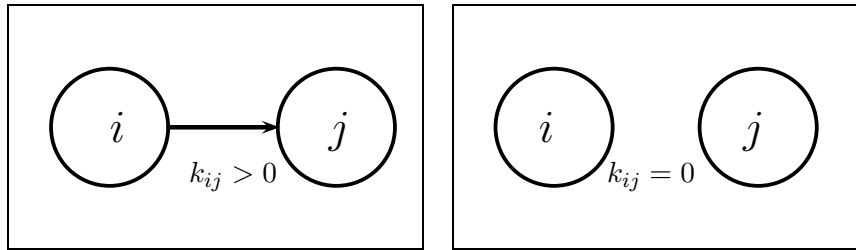


Figure 2.

Lorsque le système est ouvert, il existe d'autres types de liaisons

- une sortie vers l'extérieur du système à partir de compartiment i , matérialisée par une flèche k_{ie} .
- une entrée au niveau du compartiment i sous la forme d'une fonction $b_i(t)$.

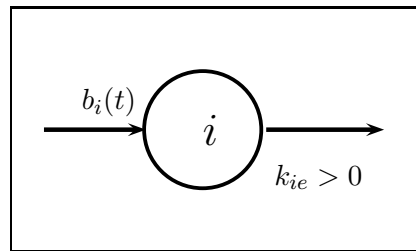


Figure 3.

Pour obtenir les équations différentielles d'un modèle compartimental, il suffira de faire un bilan de masse au niveau de chacun des compartiments. Autrement dit nous écrivons :

- La variation instantanée de quantité au niveau du compartiment i , qui est donnée par la dérivée de la fonction $x_i(t)$ est égale à la somme des quantités entrant dans i , moins la somme des quantités sortant de i , par unité de temps. D'où l'expression mathématique

$$\frac{dx_i}{dt} = x'_i = \left(\sum_{j=1 \text{ et } j \neq i}^n k_{ji} x_j(t) \right) - \left(x_i(t) \sum_{j=1 \text{ et } j \neq i}^n k_{ij} \right)$$

Exemple 3.5.1

Considérons le cas $n = 3$ et dressons le schéma d'un système compartimental autonome non-homogène où chaque compartiment admet une entrée et une sortie vers l'extérieur.

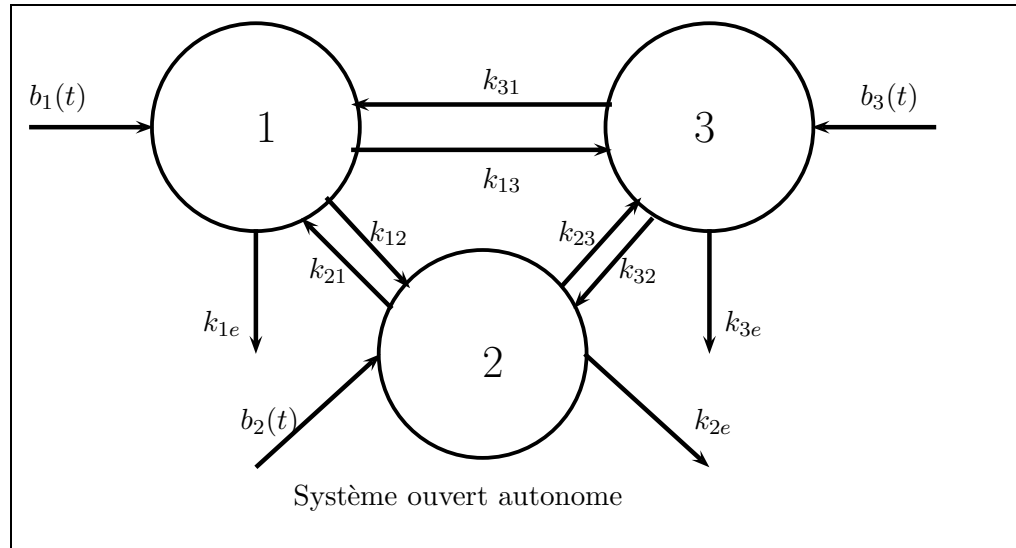


Figure 4.

Le système différentiel associé à ce schéma est le suivant :

$$\begin{cases} x'_1(t) = k_{21}x_2 + k_{31}x_3 - (k_{1e} + k_{12} + k_{13})x_1 + b_1(t) \\ x'_2(t) = k_{12}x_1 + k_{32}x_3 - (k_{2e} + k_{21} + k_{31})x_2 + b_2(t) \\ x'_3(t) = k_{13}x_1 + k_{23}x_2 - (k_{3e} + k_{31} + k_{32})x_3 + b_3(t) \end{cases}$$

Exemple 3.5.2

Considérons un modèle à deux compartiments qui est souvent utilisé en pharmacocinétique pour suivre l'évolution d'un médicament dans un organisme humain ou animal.

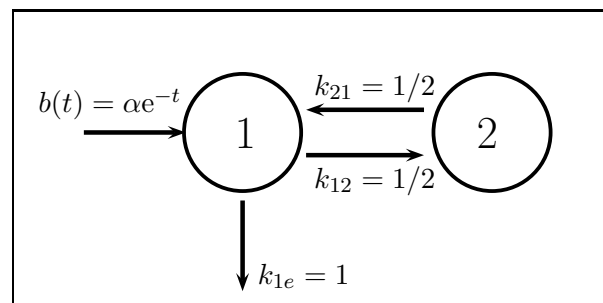


Figure 5.

Le compartiment (1) est le compartiment sanguin et le compartiment (2) est le compartiment où agit le produit (coeur, reins, poumons, tumeur, ...). $x(t)$ et $y(t)$ sont les quantités

ou les concentrations de la substance dans le compartiment (1) et (2). On suppose qu'à l'instant $t = 0$ il n'y a pas de substance dans les compartiments (1) et (2). On injecte la substance dans le sang d'une manière décroissante en unité de temps comme la fonction $b(t) = \alpha e^{-t}$ ($\alpha > 1$). Supposons que $\frac{1}{2}$ -unité de cette quantité passe dans le coeur qui sera de nouveau rejetée dans le sang et 1-unité de cette quantité va sortir du compartiment (2).

Le système différentiel associé à ce problème est exactement donné dans l'exemple 3.4.1, dont la solution a été calculée dans le paragraphe précédent.

Chapitre 4

Séries de Fourier et applications

4.1 Fonctions périodiques

Une fonction définie sur \mathbb{R} continue par morceaux est dite périodique s'il existe un nombre réel ω , tel que

$$f(x + \omega) = f(x), \quad \text{pour tout } x \in \mathbb{R}. \quad (4.1)$$

Le plus petit nombre réel ω tel que l'on ait (4.1) s'appelle *la période* de la fonction f . Si ω vérifie (4.1) tout multiple entier de ω , $n\omega$ vérifie l'aussi. En effet,

$$f(x) = f(x + \omega) = f(x + 2\omega) = \cdots = f(x + (n - 1)\omega) = f(x + n\omega).$$

Par ailleurs si \mathcal{P}_ω désigne l'ensemble des fonctions périodiques de période ω , alors \mathcal{P}_ω forme un espace vectoriel. i.e. si f et g sont deux fonctions ω -périodiques et α et β deux scalaires $\alpha f + \beta g$ est aussi une fonction ω -périodique.

Remarque *Si f est une fonction périodique de période ω_1 , alors*

$$g(x) = f\left(\frac{\omega_1 x}{\omega_2}\right)$$

est une fonction périodique de période ω_2 .

En effet

$$g(x + \omega_2) = f\left(\frac{\omega_1(x + \omega_2)}{\omega_2}\right) = f\left(\frac{\omega_1 x}{\omega_2} + \omega_1\right) = f\left(\frac{\omega_1 x}{\omega_2}\right) = g(x).$$

Une autre propriété remarquable des fonctions périodiques est que si f est ω -périodique

$$\int_a^b f(x) \, dx = \int_{a+\omega}^{b+\omega} f(x) \, dx. \quad (4.2)$$

En effet, par le changement de variable $y = x + \omega$,

$$\int_a^b f(x) dx = \int_a^b f(x + \omega) dx = \int_{a+\omega}^{b+\omega} f(y) dy.$$

De même pour toute fonction ω -périodique et tout $t \in \mathbb{R}$ on a

$$\int_a^{a+\omega} f(x) dx = \int_{a+t}^{a+t+\omega} f(x) dx. \quad (4.3)$$

En effet il suffit d'écrire

$$\int_a^{a+\omega} f(x) dx = \int_a^{a+t} f(x) dx + \int_{a+t}^{a+\omega} f(x) dx.$$

D'après (4.2)

$$\int_a^{a+t} f(x) dx = \int_{a+\omega}^{a+t+\omega} f(x) dx,$$

d'où

$$\int_a^{a+\omega} f(x) dx = \int_{a+t}^{a+\omega} f(x) dx + \int_{a+\omega}^{a+t+\omega} f(x) dx = \int_{a+t}^{a+t+\omega} f(x) dx.$$

4.2 Fonctions paires et impaires

Une fonction à valeurs réelles définie sur un intervalle $[-a, a]$ ($0 < a \leq \infty$) est dite *paire* si

$$f(x) = f(-x) \quad \forall x \in [-a, a]$$

et *impaire* si

$$f(-x) = -f(x) \quad \forall x \in [-a, a].$$

Autrement dit le graphe d'une fonction paire est symétrique par rapport à l'axe des y et celui d'une fonction impaire est symétrique par rapport à l'origine.

Comme les nombres entiers pairs et impairs qu'on a

• pair	+	pair	=	pair
• pair	+	impair	=	impair
• impair	+	impair	=	pair

Si on remplace les nombres entiers par les fonctions paires ou impaires et la loi $+$ par la loi \times on aura le tableau suivant

• paire	\times	paire	=	paire
• paire	\times	impaire	=	impaire
• impaire	\times	impaire	=	paire

pour les fonctions.

Une fonction peut ne pas être paire ou impaire, comme les fonctions e^x ou $1+x+x^2$, mais toute fonction définie sur un intervalle $[-a, a]$ peut s'écrire sous la forme d'une fonction paire et une fonction impaire. En effet

$$\begin{aligned} f(x) &= \frac{1}{2}[f(x) + f(-x)] + \frac{1}{2}[f(x) - f(-x)] \\ &= f_p(x) + f_i(x) \end{aligned}$$

où $f_p(x) = (f(x) + f(-x))/2$ est une fonction paire et $f_i(x) = (f(x) - f(-x))/2$ est une fonction impaire. Par exemple

$$\begin{aligned} e^x &= \frac{1}{2}[e^x + e^{-x}] + \frac{1}{2}[e^x - e^{-x}] \\ &= \cosh x + \sinh x \end{aligned}$$

et la fonction $f(x) = 1 + x + x^2$ est déjà sous cette forme

$$1 + x + x^2 = (1 + x^2) + (x).$$

cette propriété géométrique des fonction paires et impaire simplifie considérablement le calcul d'intégrale. En effet,

$$\int_{-a}^a f(x) \, dx = 2 \int_0^a f(x) \, dx \quad \text{si } f \text{ est paire}$$

et

$$\int_{-a}^a f(x) \, dx = 0 \quad \text{si } f \text{ est impaire.}$$

Ceci peut se vérifier en écrivant $\int_{-a}^a = \int_{-a}^0 + \int_0^a$ et en faisant le changement de variable x en $-x$. Mais on peut aussi se convaincre graphiquement.

4.3 Séries de Fourier

Soit $f(x)$ une fonction 2ω -périodique Riemann-Intégrable sur l'intervalle $[-\omega, \omega]$. Il s'agit de représenter cette fonction sous la forme d'une série

$$f(x) = \sum_{n=0}^{\infty} a_n \cos n\alpha x + b_n \sin n\alpha x. \quad (4.4)$$

Comme $\sin nx$ et $\cos nx$ sont des fonctions 2π -périodique, alors d'après la Remarque 4.1 $\sin \frac{n\pi x}{\omega}$ et $\cos \frac{n\pi x}{\omega}$ sont de période 2ω . Alors pour que le second membre de l'identité (4.4)

soit aussi de période 2ω , il suffit de prendre $\alpha = \pi/\omega$. Comme $\sin 0 = 0$ le coefficient b_0 n'existe pas et on écrit la série (4.4) sous la forme

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{\omega} + b_n \sin \frac{n\pi x}{\omega}. \quad (4.5)$$

Cette série s'appelle la *série de Fourier* de $f(x)$ et les coefficients $\{a_0, a_n, b_n\}$, s'obtiennent grâce aux relations suivantes :

$$\begin{aligned} \int_{-\omega}^{\omega} \sin \frac{n\pi x}{\omega} \sin \frac{m\pi x}{\omega} dx &= \frac{\omega}{\pi} \int_{-\pi}^{\pi} \sin ny \sin my dy \\ &= \frac{\omega}{2\pi} \int_{-\pi}^{\pi} \cos(n-m)y - \cos(n+m)y dy \\ &= \begin{cases} 0 & \text{si } n \neq m \\ \omega & \text{si } n = m, \end{cases} \end{aligned} \quad (4.6)$$

$$\begin{aligned} \int_{-\omega}^{\omega} \sin \frac{n\pi x}{\omega} \cos \frac{m\pi x}{\omega} dx &= \frac{\omega}{\pi} \int_{-\pi}^{\pi} \sin ny \cos my dy \\ &= \frac{\omega}{2\pi} \int_{-\pi}^{\pi} \sin(n-m)y + \sin(n+m)y dy \\ &= 0 \quad \text{pour tout } n \text{ et } m \end{aligned} \quad (4.7)$$

et

$$\begin{aligned} \int_{-\omega}^{\omega} \cos \frac{n\pi x}{\omega} \cos \frac{m\pi x}{\omega} dx &= \frac{\omega}{\pi} \int_{-\pi}^{\pi} \cos ny \cos my dy \\ &= \frac{\omega}{2\pi} \int_{-\pi}^{\pi} \cos(n-m)y + \cos(n+m)y dy \\ &= \begin{cases} 0 & \text{si } n \neq m \\ \omega & \text{si } n = m, \end{cases} \end{aligned} \quad (4.8)$$

Pour calculer le coefficient a_0 on intègre l'identité (4.5) entre $-\omega$ et ω . Les intégrales $\int_{-\omega}^{\omega} \cos \frac{n\pi x}{\omega} dx$ et $\int_{-\omega}^{\omega} \sin \frac{n\pi x}{\omega} dx$ sont nulles, donc on obtient $\int_{-\omega}^{\omega} f(x) dx = \frac{a_0}{2} \int_{-\omega}^{\omega} dx = \omega a_0$, d'où

$$a_0 = \frac{1}{\omega} \int_{-\omega}^{\omega} f(x) dx. \quad (4.9)$$

Pour calculer a_n on multiplie (4.5) par $\cos \frac{m\pi x}{\omega}$ et on l'intègre entre $-\omega$ et ω . D'après (4.7) et (4.8) tous les termes de la série (4.5) seront nuls sauf le coefficient de a_m . ce qui donne

$$\int_{-\omega}^{\omega} f(x) \cos \frac{m\pi x}{\omega} dx = a_m \int_{-\omega}^{\omega} \left(\cos \frac{m\pi x}{\omega} \right)^2 dx = \omega a_m.$$

d'où

$$a_m = \frac{1}{\omega} \int_{-\omega}^{\omega} f(x) \cos \frac{m\pi x}{\omega} dx. \quad (4.10)$$

De même pour calculer b_n on multiplie (4.5) par $\sin \frac{m\pi x}{\omega}$ et on l'intègre entre $-\omega$ et ω . D'après (4.6) et (4.7) tous les termes de la série (4.5) seront nuls sauf le coefficient de b_m . ce qui donne

$$\int_{-\omega}^{\omega} f(x) \sin \frac{m\pi x}{\omega} dx = b_m \int_{-\omega}^{\omega} \left(\sin \frac{m\pi x}{\omega} \right)^2 dx = \omega b_m.$$

d'où

$$b_m = \frac{1}{\omega} \int_{-\omega}^{\omega} f(x) \sin \frac{m\pi x}{\omega} dx. \quad (4.11)$$

Remarques (a) Comme $\sin \frac{m\pi x}{\omega}$ est une fonction impaire sur $] -\omega, \omega[$, alors si $f(x)$ est une fonction paire sur le même intervalle le produit $f(x) \sin \frac{m\pi x}{\omega}$ sera impaire et les coefficients

$$b_m = \frac{1}{\omega} \int_{-\omega}^{\omega} f(x) \sin \frac{m\pi x}{\omega} dx = 0 \quad \text{pour tout } m = 1, 2, \dots$$

(b) De même comme $\cos \frac{m\pi x}{\omega}$ est une fonction paire sur $] -\omega, \omega[$, alors si $f(x)$ est une fonction impaire sur le même intervalle le produit $f(x) \cos \frac{m\pi x}{\omega}$ sera impaire et les coefficients

$$a_m = \frac{1}{\omega} \int_{-\omega}^{\omega} f(x) \cos \frac{m\pi x}{\omega} dx = 0 \quad \text{pour tout } m = 0, 1, 2, \dots$$

Exemple (a) Soit $f(x)$ une fonction périodique paire de période 2, définie sur $]0, 1[$ par

$$f(x) = \begin{cases} 1 & \text{si } 0 < x < \frac{1}{2} \\ 0 & \text{si } \frac{1}{2} < x < 1. \end{cases}$$

D'après la remarque (a), comme $f(x)$ est une fonction paire, $b_n = 0$ pour tout $n = 1, 2, \dots$

et d'après (4.9) et (4.10)

$$\begin{aligned} a_0 &= \frac{2}{\omega} \int_0^\omega f(x) \, dx = \frac{2}{1} \int_0^{1/2} dx \\ &= 1 \\ a_n &= \frac{2}{\omega} \int_0^\omega f(x) \cos n\pi x \, dx \\ &= \frac{2}{1} \int_0^{1/2} \cos n\pi x \, dx \\ &= \frac{2}{n\pi} \sin \frac{n\pi}{2} \end{aligned}$$

Ainsi la série de Fourier de f est donnée par

$$\begin{aligned} f(x) &= \frac{1}{2} + \sum_{n=1}^{\infty} \frac{2}{(2n-1)\pi} (-1)^{n-1} \cos(2n-1)\pi x \\ &= \frac{1}{2} + \frac{2}{\pi} \left(\cos \pi x - \frac{1}{3} \cos 3\pi x + \frac{1}{5} \cos 5\pi x - \dots \right). \end{aligned}$$

Exemple (b) Soit $f(x)$ une fonction périodique de période 4, définie sur $] -2, 2[$ par $f(x) = x$. D'après la remarque (b), comme $f(x)$ est une fonction impaire, $a_n = 0$ pour tout $n = 0, 1, 2, \dots$ et d'après (4.11)

$$\begin{aligned} b_n &= \frac{2}{\omega} \int_0^\omega f(x) \sin n\pi x/2 \, dx \\ &= \frac{2}{2} \int_0^2 x \sin n\pi x/2 \, dx \\ &= \left[-x \frac{2 \cos n\pi x/2}{n\pi} \right]_0^1 + \frac{2}{n\pi} \int_0^2 \cos n\pi x/2 \, dx \\ &= \frac{4}{n\pi} (-1)^{n+1} \end{aligned}$$

Ainsi la série de Fourier de f est

$$f(x) = \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin \frac{n\pi x}{2}.$$

4.4 Séries de Fourier à valeurs complexes

Parfois il est plus commode de représenter la fonction f par une série de Fourier à valeurs complexes. Etant

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i} \quad \text{et} \quad \cos x = \frac{e^{ix} + e^{-ix}}{2}$$

la série de Fourier (4.5) s'écrit

$$\begin{aligned}
 f(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{\omega} + b_n \sin \frac{n\pi x}{\omega} \\
 &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[\left(\frac{a_n - ib_n}{2} \right) e^{in\pi x/\omega} + \left(\frac{a_n + ib_n}{2} \right) e^{-in\pi x/\omega} \right] \\
 &= c_0 + \sum_{n=1}^{\infty} c_n e^{\frac{ni\pi x}{\omega}} + c_{-n} e^{-\frac{ni\pi x}{\omega}} \\
 &= \sum_{n=-\infty}^{\infty} c_n e^{\frac{ni\pi x}{\omega}}, \tag{4.12}
 \end{aligned}$$

où

$$c_0 = a_0/2 = \frac{1}{2\omega} \int_{-\omega}^{\omega} f(x) dx, \quad c_n = \frac{1}{2\omega} \int_{-\omega}^{\omega} f(x) e^{-\frac{ni\pi x}{\omega}} dx, \quad n \neq 0.$$

La série (4.12) s'appelle la *série de Fourier à valeurs complexe ou exponentielle* de f et les coefficients c_n sont les coefficients complexes de Fourier de f .

Exemple (a) Considérons l'extension 2π -périodique de la fonction $f(x) = e^x$ définie sur $] -\pi, \pi[$. Les coefficients complexes de Fourier de f sont :

$$\begin{aligned}
 c_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^x e^{-inx} dx \\
 &= \frac{1}{2\pi} \left[\frac{e^{(1-in)x}}{1-in} \right]_{-\pi}^{\pi} \\
 &= \frac{1}{2\pi} \left\{ \frac{e^{(1-in)\pi} - e^{-(1-in)\pi}}{1-in} \right\} \\
 &= \frac{(1+in)(-1)^n \sinh \pi}{\pi(1+n^2)}
 \end{aligned}$$

Ainsi la série de Fourier exponentielle de f est

$$f(x) = \sum_{n=-\infty}^{\infty} \frac{(-1)^n (1+in) \sinh \pi}{\pi(1+n^2)} e^{inx}.$$

4.5 Théorèmes de convergences

Pour simplifier l'écriture on suppose dans cette section que la fonction $f(x)$ est continue par morceaux et 2π -périodique. Dans ce cas la série de Fourier de f sera représentée par

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx.$$

Soit

$$s_n(x) = \frac{a_0}{2} + \sum_{k=1}^n a_k \cos kx + b_k \sin kx$$

les n premiers termes de cette série. Elle s'appelle la somme partielle de f .

Théorème de la convergence quadratique *Sous les hypothèses ci-dessus*

$$\int_{-\pi}^{\pi} |f(x) - s_n(x)|^2 dx \rightarrow 0$$

lorsque $n \rightarrow \infty$.

Théorème de la convergence ponctuelle *Si f est 2π -périodique continue par morceaux ainsi que sa dérivée alors en chaque point $x \in [-\pi, \pi]$ on a*

$$s_n(x) \rightarrow \frac{f(x_+) + f(x_-)}{2}$$

lorsque $n \rightarrow \infty$.

Ce théorème implique que si f est continue au point x , alors $s_n(x) \rightarrow f(x)$ et si f admet une discontinuité au point x , alors $s_n(x)$ tend vers le milieu du saut de f au point x .

Exemple Dans l'exemple 4.3 (a) la fonction f est continue au point $x = 0$ et $f(0) = 1$. Alors, l'expression

$$f(x) = \frac{1}{2} + \frac{2}{\pi} \left(\cos \pi x - \frac{1}{3} \cos 3\pi x + \frac{1}{5} \cos 5\pi x - \dots \right), \quad (4.13)$$

pour $x = 0$ devient

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \dots$$

Par ailleurs au point $x = \frac{1}{2}$ il existe une discontinuité. comme $f((\frac{1}{2})_+) = 0$ et $f((\frac{1}{2})_-) = 1$, on a $(f((\frac{1}{2})_+) + f((\frac{1}{2})_-))/2 = \frac{1}{2}$ et en remplaçant dans (4.13) x par $\frac{1}{2}$ on retrouve la même valeur $\frac{1}{2}$.

4.6 L'identité de Parseval

En explicitant la valeur de

$$\int_{-\pi}^{\pi} [f(x) - s_n(x)]^2 dx$$

dans le théorème de la convergence quadratique on trouve

$$\int_{-\pi}^{\pi} [f(x) - s_n(x)]^2 dx = \int_{-\pi}^{\pi} f(x)^2 dx - \left\{ \frac{\pi a_0^2}{2} + \pi \sum_{k=1}^n a_k^2 + b_k^2 \right\} \quad (4.14)$$

Comme le premier membre de (4.14) tend vers zéro lorsque $n \rightarrow \infty$, alors à la limite on obtient

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2 dx = \frac{a_0^2}{2} + \sum_{k=1}^{\infty} a_k^2 + b_k^2. \quad (4.15)$$

Cette identité s'appelle l'*identité de Parseval* et elle peut être utilisée pour trouver la somme de certaine série numérique.

Lorsque la fonction $f(x)$ est de période 2ω , en utilisant la Remarque 4.1, on obtient

$$\frac{a_0^2}{2} + \sum_{n \geq 1} a_n^2 + b_n^2 = \frac{1}{\omega} \int_{-\omega}^{\omega} f(x)^2 dx \quad (4.16)$$

et lorsque la fonction est donnée par son expression en série de Fourier exponentielle l'identité de Parseval se lit de la façon suivante

$$\sum_{n=-\infty}^{\infty} |c_n|^2 = \frac{1}{2\omega} \int_{-\omega}^{\omega} |f(x)|^2 dx.$$

Exemple Dans l'exemple 4.3 (a) les coefficients de Fourier de la fonction f sont $a_0 = 1$, $a_n = \frac{2}{n\pi} \sin \frac{n\pi}{2}$ et $b_n = 0$ pour $n = 1, 2, \dots$. Le premier membre de l'identité de Parseval (4.16) est

$$\frac{1}{1} \int_{-1}^1 f(x)^2 dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} dx = 1$$

et le second membre est

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} a_k^2 + b_k^2 = \frac{1}{2} + \frac{4}{\pi^2} \left(1 + \frac{1}{9} + \frac{1}{25} + \dots \right)$$

d'où

$$1 + \frac{1}{9} + \frac{1}{25} + \dots = \frac{\pi^2}{8}.$$

4.7 Application à l'équation de la chaleur

Considérons une tige de longueur l suffisamment mince pour que la chaleur soit distribuée uniformément sur un section transversale de la tige à chaque instant t . La surface latérale de la tige est bien isolée de sorte qu'il n'ya pas de la perte de chaleur à travers de

la surface. La distribution de la température $u(x, t)$ au point $x \in [0, l]$ et à l'instant t est la solution du problème aux limites

$$(P) \quad \begin{cases} \frac{\partial}{\partial t} u(x, t) = k \frac{\partial^2}{\partial x^2} u(x, t) & 0 < x < l, \quad t > 0, \\ u(0, t) = u(l, t) = 0, & t > 0 \\ u(x, 0) = f(x), & x \in [0, l]. \end{cases}$$

Nous allons appliquer la méthode de la *séparation des variables* pour résoudre ce problème qui consiste à chercher une solution de la forme

$$u(x, t) = \phi(x)\psi(t)$$

elle vérifie donc

$$\phi(x)\psi'(t) = k\phi''(x)\psi(t),$$

ou bien

$$\frac{\phi''(x)}{\phi(x)} = \frac{\psi'(t)}{k\psi(t)} = \lambda. \quad (4.17)$$

Le signe de la constante λ est déterminé par une condition supplémentaire

$$|u(x, t)| \leq M, \quad \forall t > 0 \quad \text{et} \quad \forall x \in [0, l]. \quad (4.18)$$

Cette condition est physiquement acceptable, car la distribution de la température doit rester bornée et tendre vers un état d'équilibre.

A partir de la dernière identité de (4.17) on obtient $\psi'(t) = k\lambda\psi(t)$ est la solution de cette équation du premier ordre est de la forme

$$\psi(t) = ce^{k\lambda t}, \quad (4.19)$$

k étant positive λ doit être négative car sinon $\lim_{t \rightarrow \infty} \psi(t) = \infty$ ce qui contredit (4.18). Nous noterons $\lambda = -\alpha^2$ et la première identité de (4.17) implique que $\phi''(x) + \alpha^2\phi(x) = 0$, d'où la solution est de la forme

$$\phi(x) = a \cos \alpha x + b \sin \alpha x$$

Avec les condition aux limites

$$u(0, t) = \phi(0)\psi(t) = 0 \quad \text{et} \quad u(l, t) = \phi(l)\psi(t) = 0,$$

nous aurons $\phi(0) = \phi(l) = a = 0$ et $\phi(l) = b \sin \alpha l = 0$ implique que $\sin \alpha l = 0$ car en prenant $b = 0$, $\phi(x)$ devient identiquement nulle, ce qui n'est pas à prendre en considération.

Alors

$$\alpha = \frac{n\pi}{l}, \quad \text{et} \quad \lambda = -\alpha^2 = -\left(\frac{n\pi}{l}\right)^2, \quad \text{pour tout } n = 1, 2, \dots$$

et

$$\phi(x) = b \sin \alpha x \quad (4.20)$$

Ainsi en remplaçant α et λ par leurs valeurs dans (4.19) et (4.20), pour chaque $n \in \mathbb{N}$ on a une solution générale de la forme

$$u_n(x, t) = \phi_n(x)\psi_n(t) = b_n e^{-k\left(\frac{n\pi}{l}\right)^2 t} \sin \frac{n\pi x}{l},$$

où les b_n sont des constantes arbitraires.

Jusqu'à présent on n'a pas utilisé la condition initiale $u(x, 0) = f(x)$ du problème (P). Pour prendre en compte de cette condition on écrit le *principe de la superposition* qui affirme que toute combinaison linéaire des solutions générales est encore une solution générale. Or

$$u(x, t) = \sum_{n=1}^{\infty} b_n e^{-k\left(\frac{n\pi}{l}\right)^2 t} \sin \frac{n\pi x}{l}. \quad (4.21)$$

A l'instant $t = 0$ ceci donne

$$u(x, 0) = f(x) = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{l}.$$

Ainsi on peut déterminer les coefficients b_n en développant la fonction f en un série de Fourier sinus. Pour cela on prolonge cette fonction d'une manière impaire, ce qui donne

$$b_n = \frac{2}{l} \int_0^l f(x) \sin \frac{n\pi x}{l} dx. \quad (4.22)$$

Cette solution peut être modifier pour résoudre le problème suivant :

$$(PM) \quad \begin{cases} \frac{\partial v}{\partial t} = k \frac{\partial^2 v}{\partial x^2}, & x \in [0, l], t > 0, \\ v(0, t) = c_0, \\ v(l, t) = c_1, \\ v(x, 0) = v_0(x). \end{cases}$$

Ce problème correspond au cas où les extrémités de la tige sont maintenus à la température constante c_0 au point $x = 0$ et à la température c_1 au point $x = l$. En prenant $u(x, t) = v(x, t) - c_0 - \frac{x}{l}(c_1 - c_0)$ on remarque que $u(x, t)$ est la solution du problème (P). Donc

$$v(x, t) = c_0 + \frac{x}{l}(c_1 - c_0) + \sum_{n=1}^{\infty} b_n e^{-k\left(\frac{n\pi}{l}\right)^2 t} \sin \frac{n\pi x}{l}. \quad (4.23)$$

Ici la condition initiale

$$v_0(x) = c_0 + \frac{x}{l}(c_1 - c_0) + \sum_{n \geq 1} b_n \sin \frac{n\pi x}{l}.$$

détermine les coefficients b_n .

4.8 Application à l'équation de la corde vibrante

Considérons une corde vibrante de longueur l étendue sur l'axe des x fixée aux points 0 et l . Si $u(x, t)$ est le déplacement de la corde suivant l'axe des y au point $x \in [0, l]$ et à l'instant t , u est gouverné par le problème :

$$(CV) \quad \begin{cases} \frac{\partial^2}{\partial t^2} u(x, t) = c^2 \frac{\partial^2}{\partial x^2} u(x, t) & 0 < x < l, \quad t > 0, \\ u(0, t) = u(l, t) = 0, & t > 0 \\ u(x, 0) = f(x), & x \in [0, l] \\ \frac{\partial u}{\partial t}(x, 0) = g(x), & x \in [0, l]. \end{cases}$$

où c est la *célérité* ou la *vitesse de propagation de l'onde* et f et g désignent respectivement le déplacement initial et la vitesse initiale de la corde. Pour appliquer la méthode de la *séparation des variables* pour résoudre ce problème nous allons écrire la solution du problème (CV) de la forme $u(x, t) = \phi(x)\psi(t)$. La première équation de (CV) implique que

$$\phi(x)\psi''(t) = c^2\phi''(x)\psi(t),$$

ou bien

$$\frac{\phi''(x)}{\phi(x)} = \frac{\psi''(t)}{c^2\psi(t)} = \lambda.$$

et par conséquent

$$\phi''(x) - \lambda\phi(x) = 0 \quad \text{et} \quad \psi''(t) - \lambda c^2\psi(t) = 0. \quad (4.24)$$

De plus la condition aux limites $u(0, t) = 0$ implique que $\phi(0) = 0$ et la condition $u(l, t) = 0$ implique que $\phi(l) = 0$. Ainsi le système

$$\begin{cases} \phi'' - \lambda\phi = 0, \\ \phi(0) = \phi(l) = 0 \end{cases}$$

admet une solution unique selon le signe de la constante λ .

Si $\lambda > 0$, alors la solution générale est de la forme

$$\phi(x) = Ae^{-\sqrt{\lambda}x} + Be^{\sqrt{\lambda}x}. \quad (4.25)$$

Les constantes A et B seront déterminées à partir des conditions aux limites

$$\phi(0) = A + B = 0 \quad (4.26)$$

$$\phi(l) = Ae^{-\sqrt{\lambda}l} + Be^{\sqrt{\lambda}l} = 0 \quad (4.27)$$

En multipliant (4.26) par $-e^{-\sqrt{\lambda}l}$ et en additionnant avec (4.27), on trouve $2B \sinh \sqrt{\lambda}l = 0$. Comme les constantes λ et l ne sont pas nulles alors $B = 0$ et par conséquent $A = B = 0$ et finalement la fonction ϕ définie par (4.25) sera identiquement nulle et c'est ce qui est à éviter.

Si $\lambda = 0$, alors la solution générale est de la forme $\phi(x) = Ax + B$ et $\phi(0) = B = 0$, $\phi(l) = Al = 0$. Ce qui donne encore $A = B = 0$ et par conséquent on obtient encore une solution triviale.

Si $\lambda < 0$, alors la solution générale est de la forme

$$\phi(x) = A \cos \sqrt{-\lambda}x + B \sin \sqrt{-\lambda}x. \quad (4.28)$$

La condition $\phi(0) = 0$ implique que $A = 0$ et la condition $\phi(l) = 0$ donne $B \sin \sqrt{-\lambda}l = 0$. Si $B = 0$ on obtient une fois de plus la solution triviale. ainsi le seul cas possible est de prendre $\lambda = -\omega_n^2$ avec $\omega_n = n\pi/l$ est la solution non-triviale de (4.28) sera

$$\phi_n(x) = B_n \sin \frac{n\pi x}{l}, \quad n = 1, 2, \dots$$

Par ailleurs pour chaque $n \in \mathbb{N}$,

$$\psi_n(t) = \alpha_n \cos \frac{n\pi ct}{l} + \beta_n \sin \frac{n\pi ct}{l}.$$

Finalement le principe de la superposition implique que la solution du problème **(CV)** est donnée par

$$u(x, t) = \sum_{n \geq 1} \sin \frac{n\pi x}{l} \left(a_n \cos \frac{n\pi ct}{l} + b_n \sin \frac{n\pi ct}{l} \right).$$

où $a_n = B_n \alpha_n$ et $b_n = B_n \beta_n$. Les conditions initiales

$$f(x) = \sum_{n \geq 1} a_n \sin \frac{n\pi x}{l}, \quad g(x) = \sum_{n \geq 1} b_n \left(\frac{n\pi c}{l} \right) \sin \frac{n\pi x}{l}.$$

permettent de déterminer les coefficients a_n et b_n .

Lorsque la position initiale $f(x)$ est nulle et la vitesse initiale $g(x)$ n'est pas identiquement nulle alors le problème est un problème de *la corde frappée*. Par contre si la vitesse initiale est nulle mais la position initiale est différent de zéro le problème est qualifiée du problème de *la corde pincée*.

Chapitre 5

Calcul différentiel et optimisation

5.1 Applications continues

Considérons l'espace vectoriel \mathbb{R}^n muni de la norme euclidienne standard :

$$\|\mathbf{x}\| = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}$$

On appelle la boule ouverte de centre \mathbf{a} et de rayon r , notée $B(\mathbf{a}, r)$ l'ensemble

$$\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}\| < r\}.$$

Une réunion finie ou infinie des boules ouvertes, s'appelle **un ouvert** de \mathbb{R}^n .

Définition 5.1.1

Soit Ω un ouvert de \mathbb{R}^m . Une application $f : \Omega \mapsto \mathbb{R}^n$ est **continue** au point $\mathbf{a} \in \Omega$ si pour tout $\epsilon > 0$, il existe $\eta > 0$, tel que $f(B(\mathbf{a}, \eta)) \subset B(f(\mathbf{a}), \epsilon)$. Si E est un sous ensemble de Ω , on dira que f est continue sur E si f est continue en tout point de E .

Définition 5.1.2

Soit Ω un ouvert de \mathbb{R}^m et $\{\mathbf{x}_n\}$ une suite dans Ω on dira que la suite $\{\mathbf{x}_n\}$ converge vers $\mathbf{x} \in \Omega$ lorsque n tend vers infinie et on écrit

$$\lim_{n \rightarrow \infty} \mathbf{x}_n = \mathbf{x} \quad \text{ou} \quad \mathbf{x}_n \rightarrow \mathbf{x},$$

si pour tout $\epsilon > 0$ il existe $N \in \mathbb{N}$ assez grand, tel que pour tout $n \geq N$ on ait $\mathbf{x}_n \in B(\mathbf{x}, \epsilon)$, où $B(\mathbf{x}, \epsilon)$ est la boule ouvert de centre \mathbf{x} et de rayon ϵ .

Proposition 5.1.1

Soit Ω est un ouvert et $f : \Omega \mapsto \mathbb{R}^n$, alors elle est une application continue au point $\mathbf{x} \in \Omega$ si et seulement si pour toute suite $\{\mathbf{x}_n\} \subset \Omega$ qui converge vers \mathbf{x} , $f(\mathbf{x}_n)$ converge vers $f(\mathbf{x})$ dans \mathbb{R}^n .

Remarque 5.1.1

Si f est continue au point \mathbf{a} , on écrit

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a}).$$

Attention, dans cette écriture on permet de \mathbf{x} de tend vers \mathbf{a} dans n'importe quelle direction.

Exemple 5.1.1

Soit $f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$. On remarque que

$$\lim_{x \rightarrow 0} f(x, y) = -1 \quad \text{et} \quad \lim_{y \rightarrow 0} f(x, y) = 1,$$

donc

$$\lim_{x \rightarrow 0} \left[\lim_{y \rightarrow 0} f(x, y) \right] = \lim_{x \rightarrow 0} (-1) = -1$$

et

$$\lim_{y \rightarrow 0} \left[\lim_{x \rightarrow 0} f(x, y) \right] = \lim_{y \rightarrow 0} (1) = 1.$$

Par ailleurs si $x = y$ et $f(x, y) = 0$, donc $f(x, y) = 1$ sur l'axe des x , égale à -1 sur l'axe des y et finalement elle est nulle sur la première bissectrice du plan. Donc au point $(0, 0)$ elle est indéterminée.

5.2 Applications différentiable**Définition 5.2.1**

Soit Ω un ouvert de \mathbb{R}^m et $f : \Omega \mapsto \mathbb{R}^n$ une application définie sur Ω . On dit que f est différentiable au point $\mathbf{a} \in \Omega$, si pour tout $\mathbf{h} \in \mathbb{R}^m$ tel que $\mathbf{a} + \mathbf{h} \in \Omega$, il existe une application linéaire $L_{\mathbf{a}} : \mathbb{R}^m \mapsto \mathbb{R}^n$ telle que

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = L_{\mathbf{a}}(\mathbf{h}) + \varepsilon(\mathbf{h})\|\mathbf{h}\|,$$

où ε est une application de $-\mathbf{a} + \Omega$, telle que $\varepsilon(\mathbf{h}) \rightarrow 0$ lorsque $\mathbf{h} \rightarrow 0$.

Dans ce cas, l'application $L_{\mathbf{a}}$ est alors unique et elle s'appelle la **différentielle** de f au point \mathbf{a} , elle est noté par

$$Df(\mathbf{a}) \quad \text{ou} \quad df(\mathbf{a}) \quad \text{ou parfois} \quad f'(\mathbf{a}).$$

Lemme 5.2.1

Soit Ω un ouvert de \mathbb{R}^m et $f : \Omega \mapsto \mathbb{R}^n$ une application différentiable au point $\mathbf{a} \in \Omega$. Alors $Df(\mathbf{a})$ est uniquement déterminé par f .

Preuve. Soient L_1, L_2, ε_1 et ε_2 , tels que

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L_1\mathbf{h} = \varepsilon_1(\mathbf{h})\|\mathbf{h}\|$$

et

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L_2\mathbf{h} = \varepsilon_2(\mathbf{h})\|\mathbf{h}\|.$$

Prenons \mathbf{e} un vecteur unitaire de \mathbb{R}^m et $\mathbf{h} = \lambda\mathbf{e}$, alors $\|\mathbf{h}\| = |\lambda|$ et

$$\begin{aligned} \|L_1(\mathbf{e}) - L_2(\mathbf{e})\| &= \frac{\|L_1(\mathbf{h}) - L_2(\mathbf{h})\|}{\|\mathbf{h}\|} \\ &\leq \frac{\|f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L_1(\mathbf{h})\|}{\|\mathbf{h}\|} + \frac{\|f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L_2(\mathbf{h})\|}{\|\mathbf{h}\|} \\ &= \varepsilon_1(\mathbf{h}) + \varepsilon_2(\mathbf{h}) \rightarrow 0, \quad \text{lorsque } \|\mathbf{h}\| = |\lambda| \rightarrow 0. \end{aligned}$$

Ainsi $L_1(\mathbf{e}) = L_2(\mathbf{e})$, pour tout \mathbf{e} vecteur unitaire de \mathbb{R}^m et par conséquent pour tout $\mathbf{x} \in \Omega$, $L_1(\frac{\mathbf{x}}{\|\mathbf{x}\|}) = L_2(\frac{\mathbf{x}}{\|\mathbf{x}\|})$, implique que $L_1(\mathbf{x}) = L_2(\mathbf{x})$. \square

Corollaire 5.2.1

(a) Si $m = n = 1$, la notion de différentiabilité de f au point $a \in \mathbb{R}$, s'identifie à celle de la dérivabilité. Pour cela il suffit d'écrire

$$f(a + h) - f(a) = f'(a)h + \varepsilon(h)|h|$$

et la diviser par $h \neq 0$ et faire $h \rightarrow 0$. Ainsi on obtient

$$\lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h} = f'(a)$$

car $\lim_{h \rightarrow 0} \varepsilon(h) \frac{|h|}{h} = 0$.

(b) L'application $\mathbf{x} \mapsto \|\mathbf{x}\|^2$ de \mathbb{R}^n dans \mathbb{R}_+ est différentiable. En effet en tout point $\mathbf{a} \in \mathbb{R}^n$ on a

$$(\mathbf{a} + \mathbf{h}|\mathbf{a} + \mathbf{h}) - (\mathbf{a}|\mathbf{a}) = 2(\mathbf{a}|\mathbf{h}) + (\mathbf{h}|\mathbf{h}),$$

donc on peut prendre $L_{\mathbf{a}}$ l'application $\mathbf{h} \mapsto [Df(\mathbf{a})](\mathbf{h}) = 2(\mathbf{a}|\mathbf{h})$ et $\varepsilon(\mathbf{h}) = \|\mathbf{h}\|$.

(c) Une application linéaire A est toujours différentiable et sa différentielle est elle même. En effet, $A(\mathbf{a} + \mathbf{h}) - A(\mathbf{a}) = A(\mathbf{h})$, ainsi les conditions de la définition 5.2.1 sont satisfaites avec $L_{\mathbf{a}} = A$ et $\varepsilon = 0$.

Lemme 5.2.2

Soit Ω un ouvert de \mathbb{R}^m et $f : \Omega \mapsto \mathbb{R}^n$ une application différentiable au point $\mathbf{a} \in \Omega$. Alors f est continue. Plus précisément il existe une constante $k > 0$ et $\delta > 0$ tel que si $\|\mathbf{x} - \mathbf{a}\| \leq \delta$ on a

$$\|f(\mathbf{x}) - f(\mathbf{a})\| \leq k\|\mathbf{x} - \mathbf{a}\|.$$

Cette propriété s'appelle propriété de Lipschitz et une fonction qui a la propriété ci-dessus s'appelle localement lipschitzienne.

Preuve. Rappelons que si L est une application linéaire de \mathbb{R}^m dans \mathbb{R}^n , il existe une constante M telle que $\|L(\mathbf{x})\| \leq M\|\mathbf{x}\|$ pour tout $\mathbf{x} \in \mathbb{R}^m$. Pour montrer la propriété de Lipschitz de f on choisit $\mathbf{h} \in \mathbb{R}^m$ tel que $\|\mathbf{h}\| \leq \delta$ alors pour $\mathbf{x} = \mathbf{a} + \mathbf{h}$ on a $\|\mathbf{x} - \mathbf{a}\| = \|\mathbf{h}\| \leq \delta$ et

$$\begin{aligned} \|f(\mathbf{x}) - f(\mathbf{a})\| &= \|Df(\mathbf{a})(\mathbf{h})\| + \varepsilon(\mathbf{h})\|\mathbf{h}\| \\ &\leq (M + \epsilon)\|\mathbf{h}\| = k\|\mathbf{x} - \mathbf{a}\|. \end{aligned}$$

Par ailleurs, une fonction lipschitzienne est toujours continue, car d'après la Proposition 5.1.1 si $\mathbf{x}_n \rightarrow \mathbf{a}$,

$$\|f(\mathbf{x}_n) - f(\mathbf{a})\| \leq k\|\mathbf{x}_n - \mathbf{a}\| \rightarrow 0,$$

lorsque $n \rightarrow \infty$, ainsi $f(\mathbf{x}_n) \rightarrow f(\mathbf{a})$ ce qui montre la continuité de f au point \mathbf{a} . \square

5.3 Les dérivées partielles

Définition 5.3.1

Soit \mathbf{u} un vecteur de \mathbb{R}^m et $f : \Omega \subset \mathbb{R}^m \mapsto \mathbb{R}^n$ une application différentiable au point $\mathbf{a} \in \Omega$. On appelle la **dérivée directionnelle** de f dans la direction de \mathbf{u} au point \mathbf{a} et on la note $\frac{\partial f}{\partial \mathbf{u}}(\mathbf{a})$, le vecteur défini par

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a})}{t},$$

si elle existe. De plus si $\{\mathbf{e}_i\}_{i=1, \dots, m}$ est une base canonique de \mathbb{R}^m et $\mathbf{u} = \mathbf{e}_i$, au lieu de $\frac{\partial f}{\partial \mathbf{e}_i}(\mathbf{a})$ on écrit $\frac{\partial f}{\partial x_i}(\mathbf{a})$ et on l'appelle la i -ème dérivée partielle de f au point \mathbf{a} .

Remarquons que la différentiabilité de f au point \mathbf{a} , implique que si $\mathbf{h} \in \mathbb{R}^m$, alors

$$\begin{aligned} f(\mathbf{a} + t\mathbf{h}) - f(\mathbf{a}) &= L_{\mathbf{a}}(t\mathbf{h}) + \varepsilon(t\mathbf{h})\|t\mathbf{h}\| \\ &= tL_{\mathbf{a}}(\mathbf{h}) + |t|\varepsilon(t\mathbf{h})\|\mathbf{h}\|. \end{aligned}$$

En divisant cette relation par $t \neq 0$ et en faisant $t \rightarrow 0$, on obtient

$$\frac{\partial f}{\partial \mathbf{h}}(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{h}) - f(\mathbf{a})}{t} = L_{\mathbf{a}}(\mathbf{h}) = [Df(\mathbf{a})](\mathbf{h})$$

et on peut écrire

$$\frac{\partial f}{\partial \mathbf{h}}(\mathbf{a}) = [Df(\mathbf{a})](\mathbf{h}) = [Df(\mathbf{a})] \left(\sum_{i=1}^m h_i \mathbf{e}_i \right) = \sum_{i=1}^m h_i \frac{\partial f}{\partial x_i}(\mathbf{a}).$$

Donc on vient de démontrer la proposition suivante :

Proposition 5.3.1

Une application différentiable en un point admet des dérivées partielles dans toutes les directions de ce point et on a

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{a}) = [Df(\mathbf{a})](\mathbf{u}) = \sum_{i=1}^m u_i \frac{\partial f}{\partial x_i}(\mathbf{a}),$$

pour tout $\mathbf{u} = \sum_{i=1}^m u_i \mathbf{e}_i$.

Par contre, l'existence des dérivées partielles n'implique pas nécessairement celle de la différentiabilité. L'exemple suivant peut constituer un contre-exemple de ce qu'on vient d'affirmer.

Exemple 5.3.1

Soit

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y} & \text{si } x^2 \neq -y, \\ 0 & \text{si } x^2 = -y \end{cases}$$

admet des dérivées partielles dans toutes les directions au point $(0, 0)$, car pour $\mathbf{h}(h_1, h_2)$

$$\begin{aligned} \frac{f((th_1, th_2)) - f(0, 0)}{t} &= \frac{f((th_1, th_2))}{t} = \frac{t^2 h_1 h_2}{t(t^2 h_1^2 + th_2)} \\ &= \frac{h_1 h_2}{th_1^2 + h_2} \rightarrow h_1, \end{aligned}$$

lorsque $t \rightarrow 0$. Ainsi les dérivées directionnelles existent dans toutes les directions, mais cette fonction n'est même pas continue au point $(0, 0)$. En effet, on peut choisir (x, y) au voisinage de $(0, 0)$ de sorte que $|x^2 + y| = \epsilon$ (i.e. $x^2 = -y + \epsilon$) et $\left|\frac{xy}{\epsilon}\right| \rightarrow \infty$ lorsque $(x, y) \rightarrow (0, 0)$. Prenons par exemple $\epsilon = \frac{1}{n}$ et $y = -n^{-1/4}$, d'où $x = \sqrt{n^{-1/4} + n^{-1}}$, alors $\frac{xy}{\epsilon} = -n^{3/4} \sqrt{n^{-1} + n^{-1/4}} = -\sqrt{n^{1/2} + n^{5/4}} \rightarrow -\infty$. Alors ceci montre que f n'est pas continue au point $(0, 0)$.

Par contre on a

Proposition 5.3.2

Soit Ω un ouvert de \mathbb{R}^m et $f : \Omega \mapsto \mathbb{R}^n$ une application telle que $\partial f_i / \partial x_j$ existent pour tout i, j et qu'elles sont continues dans Ω , alors f est une application différentiable dans Ω .

Soient $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m)$ et $(\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n)$ les bases canoniques de \mathbb{R}^m et \mathbb{R}^n . Soit $f : \Omega \mapsto \mathbb{R}^n$ une application différentiable dans Ω avec (f_1, f_2, \dots, f_n) ses composantes dans

$(\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n)$. Alors les dérivées partielles $\partial f_i / \partial x_j$ existent pour tout i, j et la matrice de $Df(\mathbf{a})$ par rapport aux bases $\{\mathbf{e}_i\}$ et $\{\mathbf{e}'_j\}$ est représentée par

$$Df(\mathbf{a}) = \begin{pmatrix} \frac{\partial f_1(\mathbf{a})}{\partial x_1} & \frac{\partial f_1(\mathbf{a})}{\partial x_2} & \dots & \frac{\partial f_1(\mathbf{a})}{\partial x_m} \\ \frac{\partial f_2(\mathbf{a})}{\partial x_1} & \frac{\partial f_2(\mathbf{a})}{\partial x_2} & \dots & \frac{\partial f_2(\mathbf{a})}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{a})}{\partial x_1} & \frac{\partial f_n(\mathbf{a})}{\partial x_2} & \dots & \frac{\partial f_n(\mathbf{a})}{\partial x_m} \end{pmatrix}$$

Cette matrice s'appelle **matrice jacobienne**

En effet, d'après la définition d'une matrice A associée à une application linéaire L , les éléments de A sont $a_{i,j} = (\mathbf{e}'_i | L\mathbf{e}_j)$, alors si on note $\mathbf{h} = t\mathbf{e}_j$

$$\begin{aligned} & \frac{\|f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - [Df(\mathbf{a})]\mathbf{h}\|}{\|\mathbf{h}\|} \\ &= \left\| \frac{f(a_1, a_2, \dots, a_j + t, \dots, a_m) - f(a_1, a_2, \dots, a_m) - t[Df(\mathbf{a})]\mathbf{e}_j}{|t|} \right\| \rightarrow 0 \end{aligned}$$

lorsque $t \rightarrow 0$ et en particulier la i -ème composante de ce vecteur qui est

$$\frac{1}{|t|} |f_i(a_1, a_2, \dots, a_j + t, \dots, a_m) - f_i(a_1, a_2, \dots, a_m) - ta_{i,j}| \rightarrow 0.$$

D'où

$$\frac{\partial f_i(\mathbf{a})}{\partial x_j} = \lim_{t \rightarrow 0} \frac{f_i(a_1, a_2, \dots, a_j + t, \dots, a_m) - f_i(a_1, a_2, \dots, a_m)}{t} = a_{i,j}.$$

- Lorsque $n = 1$, f s'appelle un **champ de vecteurs** et sa différentielle

$$\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right) = \text{grad}f$$

s'appelle le **gradient** du champ de vecteurs f et dans ce cas la dérivée directionnelle de f dans la direction de \mathbf{u} est donnée par

$$[Df(\mathbf{a})]\mathbf{u} = (\text{grad}f(\mathbf{a})|\mathbf{u}) = \sum_{i=1}^m \frac{\partial f}{\partial x_i} u_i.$$

- Si $n = m$, $[Df(\mathbf{a})]$ est une matrice carrée et son déterminant s'appelle le **jacobien**. Il est noté par

$$Jf(\mathbf{a}) \quad \text{où} \quad \frac{D(f_1, \dots, f_n)}{D(x_1, \dots, x_n)}.$$

5.4 Interprétation géométrique de la différentielle

Si f est une application différentiable sur un ouvert $\Omega \subset \mathbb{R}^m$, l'écriture $\mathbf{y} = f(\mathbf{x})$ signifie qu'à chaque vecteur $\mathbf{x} \in \mathbb{R}^m$, on fait correspondre un vecteur $\mathbf{y} \in \mathbb{R}^n$, donc $f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})$ est un accroissement de \mathbf{y} , correspondant à l'accroissement \mathbf{h} du vecteur \mathbf{x} ; Ceci est égal à $Df(\mathbf{x})\mathbf{h}$ plus un reste d'ordre $\varepsilon(\mathbf{h})\|\mathbf{h}\|$. Si on considère un accroissement infinitésimal du vecteur \mathbf{x} noté $d\mathbf{x}$, il lui correspond un accroissement infinitésimal $d\mathbf{y}$ de \mathbf{y} . Le reste étant négligeable devant le terme linéaire quand \mathbf{h} tend vers zéro. Ainsi on obtient la formule

$$d\mathbf{y} = Df(\mathbf{x})d\mathbf{x}. \quad (5.1)$$

Le second membre de (5.1) désigne l'action de la matrice $Df(\mathbf{x})$ sur le vecteur $d\mathbf{x}$. Cette formule généralise la formule bien connue de $dy = f'(x)dx$.

- Lorsque $m = 1$, $f :]\alpha, \beta[\mapsto \mathbb{R}^n$ représente une courbe dans \mathbb{R}^n . Dans ce cas

$$d\mathbf{y} = \begin{pmatrix} dy_1 \\ dy_2 \\ \vdots \\ dy_n \end{pmatrix} = \begin{pmatrix} f'_1(t) \\ f'_2(t) \\ \vdots \\ f'_n(t) \end{pmatrix} dt = f'(t)dt.$$

Alors $f'(t)$ est un vecteur et comme par définition il est la limite d'un déplacement divisé par un temps

$$f'(t) = \lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h},$$

alors il est appelé **vitesse instantanée du déplacement** ou tout simplement **le vecteur vitesse**. C'est un vecteur dont chaque composante est la pente de la tangente à la courbe

$$\tan \theta_i = f'_i(t), \quad \text{où } \theta_i = \arccos \left(\frac{(f'(t) | \mathbf{e}_i)}{\|f'(t)\|} \right).$$

- Le cas $f : \mathbb{R}^2 \mapsto \mathbb{R}$ représente une surface. Pour un vecteur quelconque \mathbf{u} dans le plan la dérivée directionnelle dans la direction de \mathbf{u} est

$$\frac{\partial f}{\partial u}(\mathbf{x}_0) = \tan \theta = \lim_{h \rightarrow 0} \frac{f(\mathbf{x}_0 + h\mathbf{u}) - f(\mathbf{x}_0)}{h}, \quad \text{où } \theta = \arccos \left(\frac{\left(\frac{\partial f}{\partial u}(\mathbf{x}_0) | \mathbf{u} \right)}{\left\| \frac{\partial f}{\partial u}(\mathbf{x}_0) \right\|} \right).$$

Lorsque \mathbf{u} varie dans le plan \mathbb{R}^2 , la tangente directionnelle décrit un plan de l'équation

$$z = f(\mathbf{x}_0) + Df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0). \quad (5.2)$$

appelé le **plan tangent** à f au point \mathbf{x}_0 .

Exemple 5.4.1

Le plan tangent à $f(x, y) = x^2 + y$ au point $(1, 2)$ est donné par

$$\begin{aligned} z &= f(1, 2) + (Df)(1, 2) \begin{pmatrix} x - 1 \\ y - 2 \end{pmatrix} \\ &= 3 + (2x, 1) \Big|_{(1,2)} \begin{pmatrix} x - 1 \\ y - 2 \end{pmatrix} \\ &= 3 + (2, 1) \begin{pmatrix} x - 1 \\ y - 2 \end{pmatrix} = 2x + y - 1. \end{aligned}$$

5.5 Différentielle d'une fonction composée.**Théorème 5.5.1 (Théorème de Leibnitz)**

Soit $f : U \mapsto \mathbb{R}^n$, une fonction différentiable sur un ouvert U de \mathbb{R}^m et $g : V \mapsto \mathbb{R}^p$, une autre fonction différentiable sur un ouvert V de \mathbb{R}^n , telle que $f(U) \subset V$. Alors la fonction composée $h = g \circ f$ est une fonction différentiable sur U à valeurs dans \mathbb{R}^p et pour tout point $\mathbf{a} \in U$ on a

$$D(h)(\mathbf{a}) = [Dg](f(\mathbf{a})) \circ [Df](\mathbf{a}). \quad (5.3)$$

Cette formule signifie que $Df(\mathbf{a})$ est une matrice $n \times m$ et $[Dg](f(\mathbf{a}))$ est une matrice $p \times n$, donc $D(g \circ f)(\mathbf{a}) = D(h)(\mathbf{a})$ est une matrice $p \times m$. Autrement dit :

$$Dh(\mathbf{a}) = \begin{pmatrix} \frac{\partial g_1(f(\mathbf{a}))}{\partial y_1} & \frac{\partial g_1(f(\mathbf{a}))}{\partial y_2} & \dots & \frac{\partial g_1(f(\mathbf{a}))}{\partial y_n} \\ \frac{\partial g_2(f(\mathbf{a}))}{\partial y_1} & \frac{\partial g_2(f(\mathbf{a}))}{\partial y_2} & \dots & \frac{\partial g_2(f(\mathbf{a}))}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_p(f(\mathbf{a}))}{\partial y_1} & \frac{\partial g_p(f(\mathbf{a}))}{\partial y_2} & \dots & \frac{\partial g_p(f(\mathbf{a}))}{\partial y_n} \end{pmatrix} \begin{pmatrix} \frac{\partial f_1(\mathbf{a})}{\partial x_1} & \frac{\partial f_1(\mathbf{a})}{\partial x_2} & \dots & \frac{\partial f_1(\mathbf{a})}{\partial x_m} \\ \frac{\partial f_2(\mathbf{a})}{\partial x_1} & \frac{\partial f_2(\mathbf{a})}{\partial x_2} & \dots & \frac{\partial f_2(\mathbf{a})}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{a})}{\partial x_1} & \frac{\partial f_n(\mathbf{a})}{\partial x_2} & \dots & \frac{\partial f_n(\mathbf{a})}{\partial x_m} \end{pmatrix}$$

La formule (5.3) peut se démontrer d'une manière simpliste. Comme

$$\begin{aligned} g(f(\mathbf{a} + \mathbf{h})) - g(f(\mathbf{a})) &= [Dg](f(\mathbf{a})) (f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a})) + \dots \\ &= [Dg](f(\mathbf{a})) [Df(\mathbf{a})\mathbf{h} + \varepsilon(\mathbf{h})\|\mathbf{h}\|] + \dots \\ &= [Dg](f(\mathbf{a})) [Df(\mathbf{a})\mathbf{h} + \varepsilon'(\mathbf{h})\|\mathbf{h}\|]. \end{aligned}$$

où les \dots représente les petits termes correctifs non-écrits. Donc $[Dg](f(\mathbf{a})) [Df(\mathbf{a})]$ est bien la partie linéaire dans le développement de $g(f(\mathbf{a} + \mathbf{h})) - g(f(\mathbf{a}))$.

La formule (5.3) nous aide aussi à représenter le plan tangent dans \mathbb{R}^3 . Soit \mathcal{S} une surface représentée implicitement par

$$f(x, y, z) = 0.$$

où f est une fonction différentiable dans un ouvert $\Omega \subset \mathbb{R}^3$ contenant \mathcal{S} . Soit $\mathbf{a} \in \mathcal{S}$ tel que $[Df](\mathbf{a}) = \text{grad } f(\mathbf{a}) \neq 0$. Considérons un arc tracé sur \mathcal{S} et passant par \mathbf{a} défini par une fonction

$$\gamma : I \mapsto \mathcal{S}, \quad \gamma(t_0) = \mathbf{a},$$

où I est un intervalle de \mathbb{R} , contenant t_0 et γ est différentiable au point t_0 . Comme f est nulle en tous points de \mathcal{S} , alors $f(\gamma(t)) = 0$, pour tout $t \in I$. Donc sa dérivée aussi en t_0 , c'est-à-dire

$$\left. \frac{d}{dt} f(\gamma(t)) \right|_{t=t_0} = 0 \quad \text{ou bien} \quad \text{grad } f(\mathbf{a}) \cdot \gamma'(t_0) = 0.$$

Explicitement, si $\gamma'(t_0) = (x_0, y_0, z_0)$, alors

$$\frac{\partial f}{\partial x}(\mathbf{a})x_0 + \frac{\partial f}{\partial y}(\mathbf{a})y_0 + \frac{\partial f}{\partial z}(\mathbf{a})z_0 = 0.$$

Le vecteur $\gamma'(t_0) = (x_0, y_0, z_0)$ étant tangent à la courbe γ au point t_0 appartient donc nécessairement au plan tangent en \mathbf{a} à \mathcal{S} . Alors pour définir l'équation de ce plan, il suffit donc de faire varier ce vecteur. i.e.

$$\frac{\partial f}{\partial x}(\mathbf{a})x + \frac{\partial f}{\partial y}(\mathbf{a})y + \frac{\partial f}{\partial z}(\mathbf{a})z = 0 \tag{5.4}$$

est l'équation du plan tangent au point \mathbf{a} .

Ainsi on vient de montrer que

Proposition 5.5.1

Si \mathbf{a} est un point sur la surface S alors le vecteur $\nabla f(\mathbf{a}) = \text{grad } f(\mathbf{a})$ est toujours perpendiculaire au plan tangent à la surface au point \mathbf{a} .

Cette formule montre que

Exemple 5.5.1

Calcul des dérivées partielles d'une fonction dans les coordonnées sphériques.

Soit f une fonction différentiable dans les coordonnées cartésiennes $(x, y, z) \in \mathbb{R}^3$. Le changement de coordonnées en coordonnées sphériques est donnée par la fonction $g : \mathbb{R}^3 \mapsto \mathbb{R}^3$, définie comme

$$g_1(r, \theta, \phi) = x = r \cos \theta \sin \phi$$

$$g_2(r, \theta, \phi) = y = r \sin \theta \sin \phi$$

$$g_3(r, \theta, \phi) = z = r \cos \phi.$$

Alors si $h(r, \theta, \phi) = f \circ g(r, \theta, \phi)$, on aura

$$\begin{aligned} \left(\frac{\partial h}{\partial r}, \frac{\partial h}{\partial \theta}, \frac{\partial h}{\partial \phi} \right) &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} & \frac{\partial z}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} & \frac{\partial z}{\partial \theta} \\ \frac{\partial x}{\partial \phi} & \frac{\partial y}{\partial \phi} & \frac{\partial z}{\partial \phi} \end{pmatrix} \\ &= \left(\cos \theta \sin \phi \frac{\partial f}{\partial x} - r \sin \theta \sin \phi \frac{\partial f}{\partial y} + r \cos \theta \cos \phi \frac{\partial f}{\partial z}, \right. \\ &\quad \left. \sin \theta \sin \phi \frac{\partial f}{\partial x} + r \cos \theta \sin \phi \frac{\partial f}{\partial y} + r \sin \theta \cos \phi \frac{\partial f}{\partial z}, \cos \phi \frac{\partial f}{\partial x} - r \sin \phi \frac{\partial f}{\partial z} \right). \end{aligned}$$

Théorème 5.5.2 (Théorème des accroissements finis.)

Soit $f : U \mapsto \mathbb{R}^n$, une fonction différentiable sur un ouvert U de \mathbb{R}^m et soient \mathbf{x} et \mathbf{y} deux points de U tels que le segment $[\mathbf{x}, \mathbf{y}] = \{t\mathbf{y} + (1-t)\mathbf{x} \mid t \in [0, 1]\}$ soit dans U . Alors il existe n points $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n$ sur ce segment tels que

$$f_i(\mathbf{y}) - f_i(\mathbf{x}) = Df_i(\mathbf{c}_i)(\mathbf{y} - \mathbf{x}), \quad \text{pour tout } i = 1, \dots, n.$$

Ce théorème découle du théorème des accroissements finis dans \mathbb{R} , qui doit être assez familier, à savoir :

Lemme 5.5.1

Soit $f : \mathbb{R} \mapsto \mathbb{R}$ une application différentiable sur \mathbb{R} . Alors pour tout $a, b \in \mathbb{R}$, il existe $c \in [a, b]$, tels que

$$f(b) - f(a) = f'(c)(b - a).$$

En effet, si $g : U \mapsto \mathbb{R}$, une fonction différentiable sur un ouvert U , Pour tout $\mathbf{x}, \mathbf{y} \in U$, on définit

$$[0, 1] \ni t \mapsto \gamma_{\mathbf{x}, \mathbf{y}}(t) = (1-t)\mathbf{x} + t\mathbf{y}$$

et $h(t) = g \circ \gamma_{\mathbf{x}, \mathbf{y}}(t)$ qui est différentiable sur $]0, 1[$. On a ainsi

$$h'(t) = \frac{d}{dt} g((1-t)\mathbf{x} + t\mathbf{y}) = [Dg]((1-t)\mathbf{x} + t\mathbf{y})(\mathbf{y} - \mathbf{x}).$$

D'après le lemme 5.5.1, il existe $t_0 \in [0, 1]$ tel que

$$\begin{aligned} g(\mathbf{y}) - g(\mathbf{x}) &= h(1) - h(0) \\ &= h'(t_0)(1 - 0) \\ &= [Dg](c)(\mathbf{y} - \mathbf{x}), \end{aligned}$$

où $c = (1 - t_0)\mathbf{x} + t_0\mathbf{y}$. Ainsi en remplaçant pour chaque $i = 1, \dots, n$, g par f_i et \mathbf{c} par \mathbf{c}_i on obtiendra le Lemme 5.5.2.

Corollaire 5.5.1

Si U est un ouvert convexe et pour tout $\mathbf{a} \in U$, $Df(\mathbf{a}) = 0$, alors f est une fonction constante.

En effet, comme U est convexe alors pour chaque \mathbf{x} et \mathbf{y} dans U , le segment $[\mathbf{x}, \mathbf{y}] = \{t\mathbf{y} + (1-t)\mathbf{x} \mid t \in [0, 1]\}$ est entièrement contenu dans U et si on désigne $\mathbf{c}_i = (1-t_i)\mathbf{x} + t_i\mathbf{y}$, pour $i = 1, \dots, n$, l'identité

$$f_i(\mathbf{y}) - f_i(\mathbf{x}) = Df_i(\mathbf{c}_i)(y - x) = 0,$$

implique que chaque f_i est constante.

5.6 Dérivées partielles d'ordre supérieur.

Soit $f : U \mapsto \mathbb{R}^n$, une fonction différentiable sur un ouvert U de \mathbb{R}^m . Comme on a vu, pour tout $\mathbf{a} \in U$, $Df(\mathbf{a})$ est une application linéaire de \mathbb{R}^m dans \mathbb{R}^n , ou bien une matrice dans $\mathcal{M}_{n \times m}(\mathbb{R})$. L'ensemble des application linéaires de \mathbb{R}^m dans \mathbb{R}^n sera noté par $\mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$, qu'il peut s'identifier avec $\mathbb{R}^{n \times m}$. On peut penser à calculer la différentielle de cette application

$$Df : U \mapsto \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n) \simeq \mathbb{R}^{n \times m},$$

si elle existe, elle sera notée $D(Df(\mathbf{a})) = D^2f(\mathbf{a})$ et elle sera aussi linéaire de $\mathbb{R}^m \rightarrow \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$. Dans ce cas on dira que f est 2-fois différentiable et pour chaque $\mathbf{x} \in \mathbb{R}^m$, $[D^2f(\mathbf{a})](\mathbf{x})$ est encore une application linéaire de $\mathbb{R}^m \mapsto \mathbb{R}^n$. Ainsi

$$\Phi_{\mathbf{a}}(\mathbf{x}, \mathbf{y}) = [D^2f(\mathbf{a})(\mathbf{x})](\mathbf{y}) \quad (5.5)$$

définit une application bilinéaire dans $\mathbb{R}^m \times \mathbb{R}^m$.

D'une manière générale une application k -linéaire Φ sur k espaces vectoriels E_1, E_2, \dots, E_k est une application de $E_1 \times E_2 \times \dots \times E_k$ dans un espace vectoriel F telle qu'elle est linéaire par rapport à chaque variable, i.e. pour chaque $i = 1, \dots, k$ on a

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \alpha\mathbf{x}_i + \beta\mathbf{x}'_i, \dots, \mathbf{x}_k) = \alpha\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \dots, \mathbf{x}_k) + \beta\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}'_i, \dots, \mathbf{x}_k).$$

Lorsque $F = \mathbb{R}$, alors Φ est qualifiée d'une **forme k -linéaire**. Supposons que chaque E_i est de dimension n et sa base canonique est $\mathcal{E}_i = (\mathbf{e}_1^{(i)}, \dots, \mathbf{e}_n^{(i)})$, alors à chaque forme k -linéaire on peut associer un **tenseur d'ordre k** de la forme a_{i_1, i_2, \dots, i_k} , donné par

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k) = \sum_{1 \leq i_1, i_2, \dots, i_k \leq n} a_{i_1, i_2, \dots, i_k} x_{i_1}^{(1)} x_{i_2}^{(2)} \dots x_{i_k}^{(k)},$$

où $x_{i_j}^{(j)}$ sont les composant de \mathbf{x}_j dans la base \mathcal{E}_j .

Cas particulier. Si $k = 2$, alors un tenseur d'ordre 2 est tout simplement une matrice carrée d'ordre n et

$$\Phi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i y_j.$$

C'est une forme bilinéaire sur $\mathbb{R}^n \times \mathbb{R}^n$. Si $A = [a_{i,j}]$ alors on peut écrire $\Phi(\mathbf{x}, \mathbf{y}) = (\mathbf{Ax}|\mathbf{y})$.

Théorème 5.6.1

Soit $f : U \mapsto \mathbb{R}$, une fonction 2- fois différentiable sur un ouvert U de \mathbb{R}^m et à valeurs réelles. Alors la matrice associée à la forme bilinéaire (5.5) est donnée par

$$D^2 f(\mathbf{a}) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{a})}{\partial x_1 \partial x_1} & \frac{\partial^2 f(\mathbf{a})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{a})}{\partial x_1 \partial x_m} \\ \frac{\partial^2 f(\mathbf{a})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{a})}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{a})}{\partial x_2 \partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{a})}{\partial x_m \partial x_1} & \frac{\partial^2 f(\mathbf{a})}{\partial x_m \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{a})}{\partial x_m \partial x_m} \end{pmatrix} \quad (5.6)$$

Preuve. Il suffit de remarquer que

$$Df(\mathbf{a}) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_m} \right) (\mathbf{a}) = (g_1, \dots, g_m)(\mathbf{a}) = g(\mathbf{a}).$$

Donc

$$D^2 f(\mathbf{a}) = Dg(\mathbf{a}) = \begin{pmatrix} \frac{\partial g_1(\mathbf{a})}{\partial x_1} & \frac{\partial g_1(\mathbf{a})}{\partial x_2} & \cdots & \frac{\partial g_1(\mathbf{a})}{\partial x_m} \\ \frac{\partial g_2(\mathbf{a})}{\partial x_1} & \frac{\partial g_2(\mathbf{a})}{\partial x_2} & \cdots & \frac{\partial g_2(\mathbf{a})}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m(\mathbf{a})}{\partial x_1} & \frac{\partial g_m(\mathbf{a})}{\partial x_2} & \cdots & \frac{\partial g_m(\mathbf{a})}{\partial x_m} \end{pmatrix}$$

qui est identique à la matrice (5.6). □

On peut aussi démontrer (la démonstration est facile mais assez longue) que :

Théorème 5.6.2

Soit $f : U \mapsto \mathbb{R}$, une fonction 2- fois différentiable sur un ouvert U de \mathbb{R}^m et à valeurs réelles. Si les dérivées partielles $\frac{\partial f(\mathbf{a})}{\partial x_i}$, $\frac{\partial f(\mathbf{a})}{\partial x_j}$ et $\frac{\partial^2 f(\mathbf{a})}{\partial x_i \partial x_j}$ sont continues au point \mathbf{a} , alors $\frac{\partial^2 f(\mathbf{a})}{\partial x_j \partial x_i}$ existe et on a

$$\frac{\partial^2 f(\mathbf{a})}{\partial x_j \partial x_i} = \frac{\partial^2 f(\mathbf{a})}{\partial x_i \partial x_j} \quad (5.7)$$

pour tout $i, j = 1, \dots, m$.

Remarque 5.6.1

Si on a (5.7), alors la matrice (5.6) existe et elle est symétrique.

Définition 5.6.1

La matrice (5.6) s'appelle la **matrice hessienne** et si elle est continue dans un voisinage du point \mathbf{a} on dira que la fonction f est de classe C^2 . D'une manière générale une fonction f définie sur un ouvert Ω de \mathbb{R}^n à valeurs réelles, pour la quelle toutes les dérivées partielles jusqu'à l'ordre k existent et elles sont continues en tout point $\mathbf{a} \in \Omega$ est dite **de classe** C^k et on écrit $f \in C^k(\Omega)$ et si f est à valeurs dans un espace vectoriel F , on notera $f \in C^k(\Omega, F)$. Finalement une fonction est dite **lisse** ou **régulière** et on écrit $f \in C^\infty(\Omega)$ si elle est de classe C^k pour tout $k \in \mathbb{N}$.

Théorème 5.6.3 (Formule de Taylor-Young.)

Soit $f : \Omega \mapsto \mathbb{R}$, une fonction k - fois différentiable sur un ouvert convexe Ω de \mathbb{R}^n et soient \mathbf{x} et \mathbf{y} deux points de Ω . Alors il existe un points \mathbf{c} sur le segment $[\mathbf{x}, \mathbf{y}] = \{t\mathbf{y} + (1-t)\mathbf{x} \mid t \in [0, 1]\}$ tel que

$$f(\mathbf{y}) - f(\mathbf{x}) = \sum_{j=1}^{k-1} \frac{1}{j!} D^j f(\mathbf{x}) \underbrace{(\mathbf{y} - \mathbf{x}, \dots, \mathbf{y} - \mathbf{x})}_{j\text{-fois}} + \frac{1}{k!} D^k f(\mathbf{c}) \underbrace{(\mathbf{y} - \mathbf{x}, \dots, \mathbf{y} - \mathbf{x})}_{k\text{-fois}}. \quad (5.8)$$

où $D^k f(\mathbf{c}) \underbrace{(\mathbf{y} - \mathbf{x}, \dots, \mathbf{y} - \mathbf{x})}_{k\text{-fois}}$ est l'action de la forme k -linéaire $D^k f(\mathbf{c})$ sur le k -uples $(\mathbf{y} - \mathbf{x}, \dots, \mathbf{y} - \mathbf{x})$. Autrement dit

$$D^k f(\mathbf{c}) \underbrace{(\mathbf{y} - \mathbf{x}, \dots, \mathbf{y} - \mathbf{x})}_{k\text{-fois}} = \sum_{i_1, \dots, i_k=1}^n \frac{\partial^k f(\mathbf{c})}{\partial x_{i_1} \dots \partial x_{i_k}} (y_{i_1} - x_{i_1}) \dots (y_{i_k} - x_{i_k}). \quad (5.9)$$

En remplaçant \mathbf{y} par $\mathbf{x} + \mathbf{h}$, cette formule peut s'écrire sous la forme

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + Df(\mathbf{x})\mathbf{h} + \dots + \frac{1}{(k-1)!} D^{k-1} f(\mathbf{x}) \underbrace{(\mathbf{h}, \dots, \mathbf{h})}_{(k-1)\text{-fois}} + R_{k-1}(\mathbf{x}, \mathbf{h}), \quad (5.10)$$

où le reste $R_{k-1}(\mathbf{x}, \mathbf{h})$ vérifie

$$\frac{R_{k-1}(\mathbf{x}, \mathbf{h})}{\|\mathbf{h}\|^{k-1}} \rightarrow 0, \quad \text{lorsque } \mathbf{h} \rightarrow 0$$

et la formule (5.9) peut s'écrire sous la forme

$$D^k f(\mathbf{x}) \underbrace{(\mathbf{h}, \dots, \mathbf{h})}_{k\text{-fois}} = \left(\sum_{i=1}^n h_i \frac{\partial}{\partial x_i} \right)^k f(\mathbf{x}).$$

Cas particulier. Si $k = 3$ et $n = 2$.

$$\begin{aligned} D^3 f(x, y) \left[\begin{pmatrix} h \\ k \end{pmatrix}, \begin{pmatrix} h \\ k \end{pmatrix}, \begin{pmatrix} h \\ k \end{pmatrix} \right] &= \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^3 f(x, y) \\ &= \left(\frac{\partial^3 f}{\partial x^3} h^3 + 3 \frac{\partial^3 f}{\partial x^2 \partial y} h^2 k + 3 \frac{\partial^3 f}{\partial x \partial y^2} h k^2 + \frac{\partial^3 f}{\partial y^3} k^3 \right) (x, y). \end{aligned}$$

5.7 Maxima Minima.

L'une des applications fondamentales de la dérivée seconde d'une fonction est la localisation des maxima et des minima. Dans le cas d'une variable on sait que si f est de classe C^2 et $f'(x_0) = 0$, alors $f''(x_0) < 0$ implique que x_0 est un maximum local et si $f''(x_0) > 0$, x_0 est un minimum local. Notre objectif est de généraliser ce résultat à une fonction f définie sur un ouvert Ω de \mathbb{R}^n à valeurs réelles.

Définition 5.7.1

Soit Ω un ouvert de \mathbb{R}^n . et f une fonction de Ω dans \mathbb{R} . S'il existe un voisinage \mathcal{V} du point $\mathbf{x}_0 \in \Omega$, tel que $f(\mathbf{x}_0) \geq f(\mathbf{x})$ pour tout $\mathbf{x} \in \mathcal{V}$, alors nous dirons que $f(\mathbf{x}_0)$ est un **maximum local** de f . De même, si le point $\mathbf{x}_0 \in \Omega$, est tel que $f(\mathbf{x}_0) \leq f(\mathbf{x})$ pour tout $\mathbf{x} \in \mathcal{V}$, alors nous dirons que $f(\mathbf{x}_0)$ est un **minimum local** de f . Un point \mathbf{x}_0 est dit un **extrémum**, s'il est ou bien un maximum local ou un minimum local. Si f est différentiable dans Ω et $Df(\mathbf{x}_0) = 0$, alors on dit que \mathbf{x}_0 est un **point critique** de f .

Théorème 5.7.1

Soit Ω un ouvert de \mathbb{R}^n et f une fonction de Ω à valeurs réelles différentiable sur Ω . Si $\mathbf{x}_0 \in \Omega$ est un extrémum de f , alors \mathbf{x}_0 est un point critique pour f .

Définition 5.7.2

Un point critique qui n'est pas un extrémum est appelé **point selle** ou **point col**.

Définition 5.7.3

Une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est dite **définie positive** si pour tout $\mathbf{x} \in \mathbb{R}^n$, on a $(A\mathbf{x}|\mathbf{x}) > 0$, on dira qu'elle est **semi-définie positive** si $(A\mathbf{x}|\mathbf{x}) \geq 0$ et finalement **définie négative** [resp. **semi-définie négative**] si $-A$ est **définie positive** [resp. **semi-définie positive**].

Théorème 5.7.2

Soit Ω un ouvert de \mathbb{R}^n et f une fonction de classe \mathcal{C}^2 sur Ω .

- (i) Si \mathbf{x}_0 est un point critique telle que $H_{\mathbf{x}_0, f}$ la matrice hessienne de f est définie négative alors f admet un maximum local.
- (ii) Si f admet un maximum local en \mathbf{x}_0 alors $H_{\mathbf{x}_0, f}$ est semi-définie négative.

Le cas pour un minimum local s'obtient en changeant $H_{\mathbf{x}_0, f}$ par $-H_{\mathbf{x}_0, f}$. D'après le théorème 2.4.1 ce résultat admet le corollaire suivant :

Corollaire 5.7.1

Si $H_{\mathbf{x}_0, f}$ la matrice hessienne de f au point \mathbf{x}_0 et les mineurs principaux de cette matrice Δ_k vérifient

- (i) $\Delta_k = (-a_k)^k$ avec a_1, a_2, \dots, a_n tous des nombres positifs, alors f admet un maximum local au point \mathbf{x}_0 .
- (ii) $\Delta_k = (a_k)^k$ avec a_1, a_2, \dots, a_n tous des nombres positifs, alors f admet un minimum local au point \mathbf{x}_0 .

Exemple 5.7.1

Etudions la nature du point critique $(0, 0)$ de la fonction $f(x, y) = x^2 - xy + y^2$. La matrice hessienne de cette application au point $(0, 0)$ est

$$H_{(0,0),f} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

Ici $\Delta_1 = 2 > 0$ et $\Delta_2 = 4 - 1 = 3 > 0$, donc la matrice $H_{(0,0),f}$ est définie positive, par conséquent $(0, 0)$ est un minimum local.

5.8 Multiplicateurs de Lagrange

Soient Ω un ouvert de \mathbb{R}^n , f et g deux champs de vecteurs de classe \mathcal{C}^1 sur Ω et finalement $S = g^{-1}(c_0)$ où $c_0 \in \mathbb{R}$, un ensemble de niveau de g . Il s'agit de trouver un extrémum de la fonction f sur l'ensemble S .

Théorème 5.8.1

Supposons $\nabla g(x_0) \neq 0$, où \mathbf{x}_0 est un point de Ω avec $g(\mathbf{x}_0) = c_0$. Si $f|_S$ est la restriction de la fonction f à l'ensemble $S = \{\mathbf{x} \in \Omega \mid g(\mathbf{x}) = c_0\}$ admet un extrémum local en \mathbf{x}_0 , alors il existe un nombre réel λ tel que

$$\nabla f(x_0) = \lambda \nabla g(x_0). \quad (5.11)$$

Le nombre λ s'appelle le **multiplicateur de Lagrange**.

Preuve. D'après la Proposition 5.5.1, l'espace tangent à la surface S en \mathbf{x}_0 est l'espace orthogonal à $\nabla g(x_0)$. Donc si $\mathbf{c}(t)$ est une courbe dans S avec $\mathbf{c}(0) = \mathbf{x}_0$, alors $\mathbf{c}'(0)$ est le vecteur tangent en \mathbf{x}_0 à S . Comme

$$\frac{d}{dt}g(\mathbf{c}(t)) = \frac{d}{dt}c_0 = 0$$

et par ailleurs

$$\frac{d}{dt}g(\mathbf{c}(t)) = \nabla g(x_0) \cdot \mathbf{c}'(0) = 0,$$

donc $\mathbf{c}'(0)$ est orthogonal à $\nabla g(x_0)$.

D'autre part, si $f|_S$ a un extrémum en \mathbf{x}_0 , la fonction différentiable $h(t) = f(\mathbf{c}(t))$ a un extrémum local en $t = 0$. Donc

$$0 = h'(0) = \frac{d}{dt}f(\mathbf{c}(t))|_{t=0} = \nabla f(x_0) \cdot \mathbf{c}'(0).$$

Ainsi $\nabla f(x_0)$ est aussi orthogonal à S en \mathbf{x}_0 , c'est-à-dire $\nabla f(x_0)$ et $\nabla g(x_0)$ sont colinéaires. Comme $\nabla g(x_0) \neq 0$, alors il existe un $\lambda \in \mathbb{R}$ tel qu' on ait (5.11). \square

Exemple 5.8.1

Soit $S \subset \mathbb{R}^2$ la ligne passant par $(-1, 0)$ et $(0, 1)$. Trouver le minimum de $f : \mathbb{R}^2 \mapsto \mathbb{R}$, donnée par $f(x, y) = x^2 + y^2$ sur cette ligne.

Solution. L'ensemble S s'écrit $S = \{(x, y) \mid y - x - 1 = 0\}$. Notons $g(x, y) = y - x - 1$. Les extréma locaux de $f|_S$ doivent être parmi les points auxquels $\nabla f(x, y) = (2x, 2y)$ est orthogonal à S . Autrement dit, il existe λ tel que

$$\nabla f(x, y) = (2x, 2y) = \lambda \nabla g = \lambda(-1, 1).$$

D'où $2x = -\lambda$ et $2y = \lambda$ et $y - x - 1 = 0$. Ainsi on a trois équations et trois inconnues, qui donnent facilement $x = -\frac{1}{2}$, $y = \frac{1}{2}$ et $\lambda = 1$. Le point $(-\frac{1}{2}, \frac{1}{2})$ est l'intersection de la ligne S avec la ligne $L = \{(-\lambda, \lambda) \mid \lambda \in \mathbb{R}\}$ et ce point réalise un minimum de la fonction f , car

$$H_{(-\frac{1}{2}, \frac{1}{2}), f} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

est définie positive avec $\Delta_1 = 2$ et $\Delta_2 = 4$.

Si $g : \mathbb{R}^n \mapsto \mathbb{R}^k$ est une application de classe \mathcal{C}^1 et S est un ensemble de niveau dans \mathbb{R}^n défini par

$$\begin{cases} g_1(x_1, x_2, \dots, x_n) = c_1 \\ g_2(x_1, x_2, \dots, x_n) = c_2 \\ \vdots \\ g_k(x_1, x_2, \dots, x_n) = c_k. \end{cases}$$

Alors le Théorème 5.8.1 peut être généraliser de la façon suivante :

Théorème 5.8.2

Si la fonction f admet un extrémum local en $\mathbf{x}_0 \in S$, alors il existe k multiplicateurs de Lagrange $\lambda_1, \lambda_2, \dots, \lambda_k$ tels que

$$\nabla f(x_0) = \lambda_1 \nabla g_1(x_0) + \dots + \lambda_k \nabla g_k(x_0). \quad (5.12)$$

Exemple 5.8.2

Il s'agit de trouver les extréma locaux de la fonction $f(x, y, z) = x + y + z$ sous les contraintes $x^2 + y^2 = 2$ et $x + z = 1$.

Solution. Ici,

$$\begin{cases} g_1(x, y, z) = x^2 + y^2 - 2 = 0 \\ g_2(x, y, z) = x + z - 1 = 0. \end{cases}$$

Pour avoir $\nabla f(x, y, z) = \lambda_1 \nabla g_1(x, y, z) + \lambda_2 \nabla g_2(x, y, z)$, il faut que l'on ait

$$\begin{cases} 1 &= 2\lambda_1 x + \lambda_2 \\ 1 &= 2\lambda_1 y \\ 1 &= \lambda_2, \end{cases}$$

avec les contraintes

$$\begin{cases} x^2 + y^2 &= 2 \\ x + z &= 1. \end{cases}$$

On trouve ainsi 5 équations et 5 inconnues. Si $\lambda_2 = 1$ première équation donne $2\lambda_1 x = 0$ avec $\lambda_1 \neq 0$, car sinon il contredit l'équation $2\lambda_1 y = 0$, donc $x = 0$. Alors la dernière équation donne $z = 1$ et de 4-ième équation il découle $y = \pm\sqrt{2}$. comme en $(0, \sqrt{2}, 1)$ $f(0, \sqrt{2}, 1) = 1 + \sqrt{2}$ est positive et en $(0, -\sqrt{2}, 1)$, $f(0, -\sqrt{2}, 1) = 1 - \sqrt{2}$ est négative, alors en $(0, \sqrt{2}, 1)$ on a un maximum local et en $(0, -\sqrt{2}, 1)$ on a un minimum local.

Exemple 5.8.3

Une grande brasserie veut allouer un budget de 500 000 euros à la publicité au cours des six prochains mois. Les annonces publicitaires seront présentées dans deux médias : la télévision et les journaux. Les profits générés par cette campagne sont estimés par la fonction suivante :

$$p(x, y) = -x^2 - y^2 + 500x + 1000y,$$

sous condition $x + y = 500$ où $x =$ montant investi dans la publicité dans les journaux (en milliers de euros) et $y =$ montant investi dans la publicité à la télévision (en milliers de euros).

- (i) En considérant que le budget est totalement dépensé, déterminer l'allocation aux deux médias qui permette de maximiser les profits de la brasserie en utilisant la méthode du multiplicateur de Lagrange.
- (ii) Quel est le profit maximum obtenu sous cette contrainte ?
- (iii) Si on décide de doubler le total d'investissement, est-ce que la proportion du montant d'investissement dans les journaux sur le montant d'investissement dans la télévision pour un profit maximum augmentera ou diminuera ?

Solution.

- (i) Comme $p(x, y) = -x^2 - y^2 + 500x + 1000y$ et la fonction de contrainte est $g(x, y) = x + y - 150$. Donc la méthode du multiplicateur de Lagrange donne

$$\begin{aligned} \nabla p(x_0, y_0) &= \lambda \nabla g(x_0, y_0) \\ (-2x_0 + 500, -2y_0 + 1000) &= (\lambda, \lambda) \end{aligned}$$

D'où

$$\begin{aligned} -2x_0 + 500 &= \lambda \\ -2y_0 + 1000 &= \lambda, \end{aligned}$$

ce qui donne

$$\begin{cases} y_0 - x_0 = 250, \\ y_0 + x_0 = 500. \end{cases}$$

D'où

$$\lambda = 250 \quad x_0 = 125 \quad \text{et} \quad y_0 = 375.$$

(ii) Calculons la matrice hessienne de p

$$H_p(x_0, y_0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

Comme $\Delta_1 = -2 < 0$ et $\Delta_2 = 4 > 0$. La matrice H_p est définie négative et le point $(125, 375)$ est donc un maximum absolu du problème d'optimisation sous contrainte. Le profit maximum est alors

$$p(125, 375) = -(125)^2 - (375)^2 + (500)(125) + (1000)(375) = 281250 \text{ mille euros.}$$

(iii) Dans ce cas on aura

$$\begin{cases} y_0 - x_0 = 250, \\ y_0 + x_0 = 1000. \end{cases}$$

D'où

$$\lambda = -250 \quad x_0 = 375 \quad \text{et} \quad y_0 = 625.$$

Donc si dans le premier cas $\frac{x_0}{y_0} = \frac{125}{375} = \frac{1}{3}$, dans ce cas $\frac{x_0}{y_0} = \frac{375}{625} = \frac{3}{5}$, c'est-à-dire il faut augmenter d'investissement dans les journaux plus qu'à la télévision.