

# SOMMAIRE

---

<b>Chapitre 1</b> Optimisation : généralités	p01
1.1- Programmation mathématique -Définitions	p01
1.2- Conditions suffisantes d'existence d'un minimum global	p02
 <b>Chapitre 2</b> Rappels de différentiabilité	p04
2.1- Condition nécessaire pour que $f$ soit différentiable en $a$	p04
2.2- Condition suffisante pour que $f$ soit différentiable en $a$	p05
2.3- Les différentes formules de Taylor	p05
 <b>Chapitre 3</b> Optimisation locale des fonctions différentiables	p07
3.1- Conditions nécessaires de minimum local	p07
3.2- Conditions suffisantes de minimum local	p08
 <b>Chapitre 4</b> Optimisation des fonctions convexes	p10
4.1- Définitions-Propriétés	p10
4.2- Caractérisation des fonctions convexes différentiables	p11
4.3- Minimum global d'une fonction convexe	p12
 <b>Chapitre 5</b> Algorithmes pour l'optimisation sans contraintes	p13
5.1- Choix du pas de descente	p13
5.2- Choix de la direction de descente	p15
 <b>Chapitre 6</b> Optimisation avec contraintes	p17
6.1- Deux résultats auxiliaires	p17
6.2- Les relations de Kuhn-Tcker	p18
6.3- Extension à des problèmes avec contraintes égalités	p20
6.4- Cas des fonctions convexes	p22
 <b>Chapitre 7</b> Points selles et fonctions de Lagrange	p23
7.1- Fonction de Lagrange	p23
7.2- Cas des fonctions convexes	p24

<b><u>Chapitre 8</u></b>	Algorithmes pour l'optimisation avec contraintes	p25
8.1-	Méthodes de projection et de pénalisation	p25
8.2-	Méthodes duales	p27
<b><u>Annexe 1</u></b>	Rappels de topologie	p29
<b><u>Annexe 2</u></b>	Rappels de différentiabilité	p35

---

---

## Chapitre 1

# Optimisation : Généralités

### 1.1- Programmation mathématique-Définitions

D'une façon très générale, un programme mathématique dans  $\mathbb{R}^n$  est un problème d'optimisation sous contraintes de la forme :

$$(\mathcal{P}) \quad \begin{cases} \text{Minimiser } f(x) \\ \text{sous les contraintes :} \\ g_i(x) \leq 0 \quad i = 1, \dots, m \\ x \in S \subset \mathbb{R}^n \end{cases}$$

Le vecteur  $x \in \mathbb{R}^n$  a pour composantes  $x_1, \dots, x_n$  qui sont les inconnues du problème.

La fonction  $f$  est appelée fonction objectif (on dit aussi parfois fonction économique) et l'ensemble des conditions :  $g_i(x) \leq 0$  ( $i = 1, \dots, m$ ) et  $x \in S$  sont les contraintes du problème.

On appelle solution du problème  $(\mathcal{P})$  tout vecteur  $x$  vérifiant les contraintes (c'est-à-dire  $g_i(x) \leq 0$  pour  $1 \leq i \leq m$  et  $x \in S$ ).

On appelle solution optimale de  $(\mathcal{P})$  toute solution qui minimise  $f(x)$  sur l'ensemble de toutes les solutions.

#### ◇ Exemples vus en Recherche Opérationnelle :

##### 1) Programmation linéaire :

$$(\mathcal{P}) \quad \begin{cases} \text{Minimiser } f(x) \\ \text{sous les contraintes :} \\ g_i(x) \leq 0 \quad i = 1, \dots, m \end{cases}$$

avec  $f$  et  $g_i$  linéaires ( $S = \mathbb{R}^n$ ).

**Remarque :** Une contrainte égalité  $g_i(x) = 0$  peut s'exprimer avec deux contraintes inégalités  $g_i(x) \leq 0$  et  $-g_i(x) \leq 0$ .

##### 2) Programmation en nombres entiers :

$$(\mathcal{P}) \quad \begin{cases} \text{Minimiser } f(x) = \langle c, x \rangle \\ \text{sous les contraintes :} \\ A.x = b \\ x_i \geq 0, \quad x_i \text{ entiers, } 1 \leq i \leq n \end{cases}$$

Ici,  $S = \mathbb{Z}^n$ . Par exemple, une compagnie aérienne cherche à réaliser un certain programme annuel de vols, tout en minimisant le coût total de ses avions. Les variables représentent alors le nombre d'appareils de chaque type à acheter ou à louer. Il n'est alors pas admissible d'avoir une solution optimale fractionnaire.  $\diamond$

Dans, ce cours, à l'exception du chapitre 6, on aura toujours  $S = \mathbb{R}^n$ , c'est-à-dire seulement des contraintes du type  $g_i(x) \leq 0$ .

De manière plus générale, soit  $f : C \subset \mathbb{R}^n \rightarrow \mathbb{R}$  et  $a \in C$ .

**Définition 1.1 :** On dit que  $a$  est un minimum global de  $f$  sur  $C$  si, pour tout  $x \in C$ ,  $f(a) \leq f(x)$ . L'ensemble des minimums de  $f$  sur  $C$  est noté  $\text{Arg}_C \min f$ .

**Définition 1.2 :** On dit que  $a$  est un minimum local de  $f$  sur  $C$  s'il existe une boule ouverte  $B$  contenant  $a$  telle que  $a$  soit minimum global de  $f$  sur  $B \cap C$ .

**Remarques :**

1) Il est, en général, beaucoup plus facile de déterminer un minimum local qu'un minimum global.

2) Si  $f(a) \geq f(x)$  pour tout  $x \in C$ , alors  $-f(a) \leq -f(x)$  pour tout  $x \in C$ . Donc  $a$  maximum de  $f$  équivaut à  $a$  minimum de  $-f$ . Par conséquent, on se limitera, dans le cours, mais pas dans les exercices, à l'étude des minimums.

## 1.2- Conditions suffisantes d'existence d'un minimum global

Cette section nécessite quelques connaissances minimales de topologie. Pour un léger rafraîchissement, notamment en ce qui concerne les définitions, voir l'annexe 1.

**Théorème 1.1 :** Soit  $f$  une application continue de  $C \subset \mathbb{R}^n$  dans  $\mathbb{R}$ .

1) Si  $C$  est compact, alors  $\text{Arg}_C \min f$  est un compact non vide.

2) Si  $C$  est fermé et  $f$  coercive (c'est-à-dire  $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$ ), alors  $\text{Arg}_C \min f$  est un compact non vide.

*Preuve :*

1) Soit  $m = \inf f(C)$  où  $f(C) = \{f(x) ; x \in C\}$ .

Par définition de la borne inférieure, il existe une suite  $(y_k)$  de  $f(C)$  telle que  $m = \lim_{k \rightarrow +\infty} y_k$ . Comme  $y_k \in f(C)$ , il existe  $x^{(k)} \in C$  tel que  $y_k = f(x^{(k)})$ . La suite  $(x^{(k)})$  étant dans  $C$  compact, il existe une sous-suite  $(x^{(k')})$  de  $(x^{(k)})$  qui converge dans  $C$ . Soit  $a = \lim_{k \rightarrow +\infty} x^{(k')}$ . Comme  $f$  est continue, on a  $f(a) = \lim_{k \rightarrow +\infty} f(x^{(k')})$ .

Or  $m = \lim_{k \rightarrow +\infty} f(x^{(k)})$  et  $(f(x^{(k')}))$  est une sous-suite de  $(f(x^{(k)}))$  car  $(x^{(k')})$  est une sous-suite de  $(x^{(k)})$  ; donc  $\lim_{k \rightarrow +\infty} f(x^{(k')}) = m$  et comme  $f(a) = \lim_{k \rightarrow +\infty} f(x^{(k')})$ , on a alors  $m = f(a)$ .  
 Finalement,  $m = f(a)$  avec  $a \in C$ , et  $a \in \text{Arg}_C \min f$ . D'où  $\text{Arg}_C \min f \neq \emptyset$ .  
 On a  $\text{Arg}_C \min f = \{x \in C ; f(x) = m\} = f^{-1}(\{m\})$ . Comme  $\{m\}$  est un fermé et que  $f$  est continue,  $f^{-1}(\{m\})$  est un fermé. De plus,  $f^{-1}(\{m\})$  est contenu dans  $C$  qui est un compact donc borné et donc  $f^{-1}(\{m\}) = \text{Arg}_C \min f$  est borné, puis c'est un compact car il est aussi fermé.

**2)** Soit  $x^{(0)} \in C$  quelconque. Comme  $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$ , il existe  $M > 0$  tel que, pour tout  $x \in C$ , si  $\|x\| \geq M$ , alors  $f(x) \geq f(x^{(0)}) + 1$ .  
 L'ensemble  $C \cap B_F(0, M)$  est un fermé, comme intersection de 2 fermés ; c'est donc un fermé borné de  $\mathbb{R}^n$ , c'est-à-dire un compact de  $\mathbb{R}^n$ .  
 D'après le 1),  $\text{Arg}_{C \cap B_F(0, M)} \min f$  est non vide.  
 Soit  $a \in \text{Arg}_{C \cap B_F(0, M)} \min f$ . Pour tout  $x \in C \cap B_F(0, M)$ , on a  $f(a) \leq f(x)$ .  
 D'autre part,  $x^{(0)} \in C \cap B_F(0, M)$ , car si on avait  $\|x^{(0)}\| > M$ , on aurait  $f(x^{(0)}) \geq f(x^{(0)}) + 1$ , ce qui est absurde. D'où  $f(x^{(0)}) \geq f(a)$ .  
 On en déduit que, si  $\|x\| > M$ , alors  $f(x) \geq f(x^{(0)}) + 1 \geq f(a) + 1 > f(a)$  : c'est donc que finalement,  $a \in \text{Arg}_C \min f$  et que  $\text{Arg}_C \min f = \text{Arg}_{C \cap B_F(0, M)} \min f$ .  
 D'après le 1),  $\text{Arg}_C \min f$  est donc compact.

□

**Attention!** : La coercivité ne peut être utilisée que pour l'existence des minimums, pas pour celle des maximums.

**Remarque** : Lorsque les conditions de ce théorème sont satisfaites, les minimums globaux d'une fonction  $f$  sur  $C$  fermé sont, soit des minimums locaux à l'intérieur de  $C$ , soit des points de la frontière de  $C$ .

---

## Chapitre 2

### Rappels de différentiabilité

Pour plus de détails, voir l'annexe 2.

Soit  $U$  un ouvert de  $\mathbb{R}^n$ ,  $a$  un élément de  $U$  et  $f$  une application de  $U$  dans  $\mathbb{R}$ .

**Définition 2.1:** On dit que  $f$  est *différentiable en  $a$*  s'il existe une forme linéaire (que l'on notera  $d_a f$ ) telle que, pour tout  $h$  pour lequel  $a + h \in U$  :

$$f(a + h) - f(a) = d_a f(h) + \|h\|\varepsilon(h) \text{ avec } \varepsilon(h) \rightarrow 0 \text{ si } \|h\| \rightarrow 0.$$

On dit que  $f$  est *différentiable sur  $U$*  si elle est différentiable en tout point de  $U$ .

On notera  $(e_1, \dots, e_n)$  la base canonique de  $\mathbb{R}^n$  ( $e_1 = (1, 0, \dots, 0)$ ,  $e_2 = (0, 1, 0, \dots, 0)$ ...)  $\mathbb{R}^n$  est un espace euclidien doté du produit scalaire  $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$  si  $x = (x_1, \dots, x_n)$  et  $y = (y_1, \dots, y_n)$ .

**Remarque :**  $d_a f$  étant une forme linéaire, il existe un unique élément de  $\mathbb{R}^n$ , noté  $\nabla f(a)$  et appelé *gradient* de  $f$  en  $a$  tel que, pour tout  $h \in \mathbb{R}^n$ ,  $d_a f(h) = \langle \nabla f(a), h \rangle$ .

**Définition 2.2:** On dit qu'une fonction  $f : U \rightarrow \mathbb{R}^p$  a une *dérivée en  $a$  suivant un vecteur  $h \in \mathbb{R}^n$*  si la fonction de la variable réelle  $t \mapsto \varphi_h(t) = f(a + th) \in \mathbb{R}^p$  est dérivable en 0.

On appelle alors *dérivée de  $f$  en  $a$  suivant  $h$*  le vecteur  $D_h f(a) = \varphi'_h(0)$  :

$$D_h f(a) = \lim_{t \rightarrow 0} \frac{f(a + th) - f(a)}{t}$$

Si  $h = e_j$ , on note  $\varphi_{e_j}$  par  $\varphi_j$  :  *$j$ -ème application partielle de  $f$  en  $a$* , et on note  $D_{e_j} f(a)$  par  $D_j f(a)$  :  *$j$ -ème dérivée partielle de  $f$  en  $a$* .

### 2.1- Condition nécessaire pour que $f$ soit différentiable en $a$

**Proposition 2.1 :** Si  $f$  est différentiable en  $a$ , alors elle admet une dérivée au point  $a$  suivant tout vecteur  $h$ , et celle-ci est égale à  $D_h f(a) = df_a(h)$ .

En particulier, les dérivées partielles  $D_i f$  existent en  $a$  pour tout  $i \in \{1, \dots, n\}$ .

**Attention :** Cette condition n'est pas suffisante ! (voir exemple A2.5)

**Définition 2.3 :**  $f$  est dite de classe  $C^1$  sur  $U$  si, pour tout  $1 \leq j \leq n$ ,  $D_j f$  existe et est continue sur  $U$ .  $f$  est de classe  $C^1$  en  $a$  si les dérivées partielles  $D_i f$  existent dans un voisinage  $V$  de  $a$  et sont continues en  $a$ .

## 2.2- Condition suffisante pour que $f$ soit différentiable en $a$

**Proposition 2.2 :** Si  $f$  est de classe  $C^1$  en  $a$  (resp. sur  $U$ ), alors elle est différentiable (et donc continue) en  $a$  (resp. sur  $U$ ).

**Remarque :** La condition est suffisante, mais pas nécessaire !

On peut définir, par récurrence, la classe  $\mathcal{C}^k$  d'une fonction  $f$  de la façon suivante :

**Définition 2.4:**  $f$  est dite de classe  $C^k$  sur l'ouvert  $U$  si  $f$  est de classe  $C^1$  sur  $U$  et si  $D_j f$  est de classe  $C^{k-1}$  sur  $U$  pour tout  $j$ . Enfin, on dit que  $f$  est de classe  $C^\infty$  si elle est de classe  $C^k$  pour tout  $k \in \mathbb{N}$ .

Dans ce cours,  $f$  sera, en général, de classe  $C^1$  au moins, sauf éventuellement en quelques points.

**Proposition 2.3 (Théorème de Schwarz)**  
Si  $f$  est de classe  $C^2$ , on a  $D_i D_j f = D_j D_i f$  pour tout  $1 \leq i, j \leq n$ .

On appelle différentielle d'ordre 2 de  $f$  en  $a$ , la forme bilinéaire  $d_a^2 f$  définie sur  $(\mathbb{R}^n)^2$  par :

$$d_a^2 f(h, k) = \sum_{1 \leq i, j \leq n} D_i D_j f(a) h_i k_j.$$

On note  $\nabla^2 f(a)$  et on appelle Hessienne de  $f$  en  $a$ , la matrice  $(D_i D_j f(a))_{1 \leq i, j \leq n}$ .

D'après le théorème de Schwarz,  $D_i D_j f(a) = D_j D_i f(a)$  donc  $d_a^2 f$  est une forme bilinéaire symétrique (et  $\nabla^2 f(a)$  une matrice symétrique).

$$d_a^2 f(h, k) = \langle \nabla^2 f(a) h, k \rangle = {}^t k \nabla^2 f(a) h = {}^t h \nabla^2 f(a) k.$$

## 2.3- Les différentes formules de Taylor

**Formule de Taylor Mac-Laurin à l'ordre 1 :**

Si  $f$  est de classe  $C^1$  sur  $[a, a + h[ \subset U$ , alors il existe  $\theta \in ]0, 1[$  tel que :

$$f(a + h) = f(a) + d_{a+\theta h} f(h) = f(a) + \langle \nabla f(a + \theta h), h \rangle.$$

**Formule de Taylor-Mac-Laurin à l'ordre 2 :**

Si  $f$  est de classe  $C^2$  sur  $[a, a + h] \subset U$ , il existe  $\theta \in ]0, 1[$  tel que :

$$f(a + h) = f(a) + d_a f(h) + \frac{1}{2} d_{a+\theta h}^2 f(h, h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle \nabla^2 f(a + \theta h) h, h \rangle.$$

**Formule de Taylor-Young à l'ordre 1 :**

Si  $f$  est différentiable en  $a$ , on a, au voisinage de  $a \in U$  :

$$f(a+h) = f(a) + d_a f(h) + \|h\| \varepsilon(h)$$

avec  $\varepsilon(h) \rightarrow 0$  quand  $\|h\| \rightarrow 0$ , si  $a+h \in U$  (c'est la définition de la différentiabilité).

**Formule de Taylor-Young à l'ordre 2 :**

Si  $f$  est de classe  $C^2$  en  $a$ , on a, au voisinage de  $a \in U$  :

$$\begin{aligned} f(a+h) &= f(a) + d_a f(h) + \frac{1}{2} d_a^2 f(h, h) + \|h\|^2 \varepsilon(h) \\ &= f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} \langle \nabla^2 f(a) h, h \rangle + \|h\|^2 \varepsilon(h) \end{aligned}$$

avec  $\varepsilon(h) \rightarrow 0$  quand  $\|h\| \rightarrow 0$ , si  $a+h \in U$ .

---



---



## Chapitre 3

# Optimisation locale des fonctions différentiables

Soit  $U$  un ouvert de  $\mathbb{R}^n$ ,  $a$  un élément de  $U$  et  $f$  une application de  $U$  dans  $\mathbb{R}$ .

### 3.1- Conditions nécessaires de minimum local

**Proposition 3.1 :** Soit  $a \in U$ . Si  $f : U \rightarrow \mathbb{R}$  est différentiable en  $a$  et si  $f$  admet un minimum (ou un maximum) local en  $a$ , alors

$$d_a f = 0.$$

Les points  $a$  solutions de  $d_a f = 0$  (ou  $\nabla f(a) = 0$ ) sont appelés points critiques de  $f$  (ou points stationnaires de  $f$ ).

*Preuve :*

Soit  $h \in \mathbb{R}^n$  et soit  $\varphi_h : t \mapsto f(a + th)$  où  $t$  est réel.

On a  $a \in U$ , donc  $\varphi_h(0)$  existe et, comme  $U$  est ouvert,  $\varphi_h$  est aussi définie sur un intervalle réel  $I_h$  contenant 0. Comme  $a$  est un minimum local de  $f$ , il existe un intervalle  $I'$  contenu dans  $I_h$  et contenant 0 tel que  $f(a + th) - f(a) \geq 0$  pour tout  $t \in I'$ .

On a supposé  $f$  différentiable en  $a$ , donc :

$$f(a + th) - f(a) = d_a f(th) + \|th\|\varepsilon(th).$$

Or, par linéarité,  $d_a f(th) = t d_a f(h)$ , donc

$$f(a + th) - f(a) = t \left[ d_a f(h) + \frac{|t|}{t} \|h\| \varepsilon(th) \right] \geq 0.$$

Ceci donne, pour  $t > 0$ ,  $d_a f(h) + \frac{|t|}{t} \|h\| \varepsilon(th) \geq 0$  et pour  $t < 0$ ,  $d_a f(h) + \frac{|t|}{t} \|h\| \varepsilon(th) \leq 0$ , puis, par passage à la limite, quand  $t \rightarrow 0^+$ ,  $d_a f(h) \geq 0$  et quand  $t \rightarrow 0^-$ ,  $d_a f(h) \leq 0$ .

Finalement,  $d_a f(h) = 0$  et, comme  $h \in \mathbb{R}^n$  était pris quelconque,  $d_a f = 0$ .

□

◇**Exemple 3.1 :** Si  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  est définie par  $f(x, y) = x^2 + 3y^2$ , on a  $D_1 f(x, y) = 2x$ ,  $D_2 f(x, y) = 6y$  et  $\nabla f(x, y) = 0$  si et seulement si  $x = y = 0$  :  $(0, 0)$  est le seul point critique de  $f$ . ◇

**Proposition 3.2 :** Si  $f$  admet un minimum local en  $a$  et si  $f$  est de classe  $\mathcal{C}^2$  en  $a$ , alors  $d_a^2 f(h, h) \geq 0$  pour tout  $h \in \mathbb{R}^n$ , c'est-à-dire que la matrice  $\nabla^2 f(a)$  est positive, ou encore que  $d_a^2 f$  est une forme bilinéaire positive.

*Preuve :*

On utilise la formule de Taylor-Young à l'ordre 2 :

$$\begin{aligned} f(a+th) &= f(a) + d_a f(th) + \frac{1}{2} d_a^2 f(th, th) + \|th\|^2 \varepsilon(th) \\ &= f(a) + t d_a f(h) + \frac{1}{2} t^2 [d_a^2 f(h, h) + 2\|h\|^2 \varepsilon(th)] \end{aligned}$$

Or  $f(a+th) - f(a) \geq 0$  et, d'après la proposition 3.1,  $d_a f(h) = 0$ , donc, pour tout  $t \in I' \setminus \{0\}$ ,  $d_a^2 f(h, h) + 2\|h\|^2 \varepsilon(th) \geq 0$ . On obtient alors le résultat en faisant  $t \rightarrow 0$ . □

**Attention !** La réciproque est fausse comme le montre l'exemple ci-dessous.

◇ **Exemple 3.2 :**  $f : \mathbb{R} \rightarrow \mathbb{R}$ , définie par  $f(x) = x^3$ . On a  $f'(x) = 3x^2$  et  $f''(x) = 6x$  donc 0 seul point critique de  $f$  ;  $f''(0) = 0 \geq 0$  et pourtant 0 n'est pas minimum local de  $f$ . ◇

### 3.2- Conditions suffisantes de minimum local

**Proposition 3.3 :** Soit  $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$  différentiable en  $a$  stationnaire. Dans les 2 cas suivants,  $f$  admet un minimum local en  $a$  :

- 1)  $f$  est de classe  $\mathcal{C}^2$  en  $a$  et il existe  $\alpha > 0$  tel que  $d_a^2 f(h, h) \geq \alpha \|h\|^2$  pour tout  $h \in \mathbb{R}^n$ .
- 2) Il existe  $r > 0$  tel que  $f$  soit de classe  $\mathcal{C}^2$  sur la boule  $B(a, r)$  et tel que, pour tout  $h \in \mathbb{R}^n$ ,  $d_x^2 f(h, h) \geq 0$  pour  $x \in B(a, r)$ .

*Preuve :*

- 1) On applique la formule de Taylor-Young à l'ordre 2 :

$$f(a+h) = f(a) + d_a f(h) + \frac{1}{2} d_a^2 f(h, h) + \|h\|^2 \varepsilon(h) \text{ et } d_a f(h) = 0 \text{ donc :}$$

$$f(a+h) - f(a) \geq \left( \frac{1}{2} \alpha + \varepsilon(h) \right) \|h\|^2.$$

Comme  $\varepsilon(h) \rightarrow 0$  quand  $\|h\| \rightarrow 0$ , il existe  $\gamma$  tel que, pour  $\|h\| \leq \gamma$ , on ait  $\frac{1}{2} \alpha + \varepsilon(h) \geq 0$  et donc  $f(a+h) - f(a) \geq 0$ .

- 2) On applique la formule de Taylor-Mac-Laurin à l'ordre 2 : il existe  $\theta \in ]0, 1[$  tel que :

$$f(a+h) = f(a) + d_a f(h) + \frac{1}{2} d_{a+\theta h}^2 f(h, h)$$

et comme  $d_a f(h) = 0$ , on a  $f(a+h) - f(a) = \frac{1}{2} d_{a+\theta h}^2 f(h, h)$ . Or, pour  $\|h\| < r$ , on a alors  $a + \theta h \in B(a, r)$  et donc  $d_{a+\theta h}^2 f(h, h) \geq 0$ . □

**Remarque :** Comme  $d_a^2 f$  est une forme bilinéaire symétrique,  $\nabla^2 f(a)$  est une matrice symétrique réelle, donc diagonalisable dans une base orthonormée et on sait qu'il existe une matrice orthogonale  $P$  telle que  $D = {}^t P \nabla^2 f(a) P$  où  $D$  désigne donc la matrice dont les éléments diagonaux sont les valeurs propres  $\lambda_1, \lambda_2, \dots, \lambda_n$ .

Quitte à poser  $y = {}^t P h$ , on obtient alors :

$$\frac{{}^t h \nabla^2 f(a) h}{{}^t h h} = \frac{{}^t h P D {}^t P h}{{}^t h h} = \frac{{}^t y D y}{{}^t y y} = \frac{\lambda_1 y_1^2 + \dots + \lambda_n y_n^2}{y_1^2 + \dots + y_n^2}.$$

En supposant  $\lambda_1 \leq \dots \leq \lambda_n$ , on a alors :

- le rapport précédent est compris entre  $\lambda_1$  et  $\lambda_n$
- la valeur  $\lambda_1$  est atteinte pour  $y_1 = 1$  et  $y_2 = \dots = y_n = 0$ .
- la valeur  $\lambda_n$  est atteinte pour  $y_n = 1$  et  $y_1 = \dots = y_{n-1} = 0$ .

Ainsi, la connaissance de la plus petite et de la plus grande valeurs propres  $\lambda_1$  et  $\lambda_n$  de la matrice  $\nabla^2 f(a)$  peut permettre de connaître le signe de  ${}^t h \nabla^2 f(a) h$  lorsque  $h$  décrit  $\mathbb{R}^n$ .

$d_a^2 f(h, h) \geq 0$  pour tout  $h \in \mathbb{R}^n$  si et seulement si les valeurs propres de  $d_a^2 f$  sont toutes positives (on dit alors que  $f$  est positive). De plus, il existe  $\alpha > 0$  tel que  $d_a^2 f(h, h) \geq \alpha \|h\|^2$  si et seulement si les valeurs propres de  $d_a^2 f$  sont strictement positives (on dit alors que  $d_a^2 f$  est définie positive).

- Si  $\lambda_1 \lambda_n > 0$ , il y a extrémum local en  $a$  (maximum si  $\lambda_n < 0$ , minimum si  $\lambda_1 > 0$ ) ; en particulier, on retrouve le 1) de la proposition 3.3 avec  $\alpha = \lambda_1$  si  $\lambda_1 > 0$ ,
- Si  $\lambda_1 \lambda_n < 0$ , il n'y a pas d'extrémum en  $a$  ( $a$  est dit col ou point-selle),
- Si  $\lambda_1 \lambda_n = 0$ , on ne peut pas conclure directement.

◇ **Exemple 3.3 :** Si  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  est définie par  $f(x, y) = x^2 + 3y^2$ ,  $(0, 0)$  est le seul point critique,  $\nabla^2 f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix}$  est définie positive (valeurs propres 2 et 6) donc  $(0, 0)$  est l'unique minimum local de  $f$ . ◇

◇ **Exemple 3.4 :** Extrémums locaux de  $f : (x, y) \mapsto x^2 + x^2 y + y^4$ .

$f$  est de classe  $\mathcal{C}^1$  sur  $\mathbb{R}^2$  et  $D_1 f(x, y) = 2x(1 + y)$  et  $D_2 f(x, y) = x^2 + 4y^3$ . Un point critique vérifie donc  $x = 0$  et  $y = 0$  ou  $y = -1$  et  $x = 2\varepsilon$ . Comme  $f(-x, y) = f(x, y)$ , il suffit d'étudier  $(0, 0)$  et  $(2, -1)$ .

Si  $y \geq -1$ ,  $f(x, y) - f(0, 0) = f(x, y) \geq y^4 \geq 0$ , donc  $(0, 0)$  présente un minimum local. Il n'est pas global, car  $f(x, -2) = -x^2 + 16$  donc  $f(5, -2) < 0$ .

$f(2 + h, -1 + k) = k(2 + h)^2 + (-1 + k)^4 = 1 + 4hk + 6k^2 + kh^2 - 4k^3 + k^4$ . Donc,  $f(x, y) = f(2, -1) + 4hk + 6k^2 + o(h^2 + k^2)$  car avec  $\|(h, k)\| = \max(|h|, |k|)$ ,  $|kh^2| \leq \|(h, k)\|^3 = o(\|(h, k)\|^2)$ ,  $|k^3| \leq \|(h, k)\|^3$  et  $k^4 \leq \|(h, k)\|^4$ . Localement, le signe de  $f(x, y) - f(2, -1)$  est donné par celui de  $4hk + 6k^2$ . Pour  $h = 0$ ,  $f(2, -1 + k) - f(2, -1) = 4k + (-1 + k)^4 - 1 \sim 6k^2 \geq 0$  et pour  $h = -2k$ ,  $f(2 - 2k, -1 + k) - f(2, -1) \sim -2k^2$ . On n'a donc pas d'extrémum local. ◇

## Chapitre 4

### Optimisation des fonctions convexes

#### 4.1- Définitions, propriétés

**Définition 4.1:** Un sous-ensemble  $C$  de  $\mathbb{R}^n$  est dit convexe si, pour tout  $(a, b) \in C^2$ ,  $[a, b] \subset C$  (c'est-à-dire pour tout  $\lambda \in [0, 1]$ ,  $\lambda a + (1 - \lambda)b \in C$ ).

◇ **Exemple 4.1:** Une boule est convexe.

*Preuve :* On part de  $B(a, r)$ , par exemple, avec  $x, y$  dans cette boule, et  $t \in [0, 1]$ .

$$\|a - tx - (1 - t)y\| \leq t\|x - a\| + (1 - t)\|y - a\| < tr + (1 - t)r = r$$

en écrivant  $a - tx - (1 - t)y = t(a - x) + (1 - t)(a - y)$ . ◇

**Définition 4.2:** : Une fonction réelle  $f$  définie sur  $C$  convexe est dite convexe si, pour tout  $(a, b) \in C^2$  et pour tout  $\lambda \in [0, 1]$ ,  $f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b)$ .

**Propriétés 4.1: 1)** Si  $C$  est un convexe de  $\mathbb{R}^n$  et  $(f_i)_{i \in I}$ , une famille quelconque de fonctions convexes alors :

a)  $\sup_{i \in I} f_i$  est convexe ;

b) si  $I$  est fini et si  $(\lambda_i)_{i \in I}$  est une famille de réels positifs, alors  $\sum_{i \in I} \lambda_i f_i$  est convexe.

**2)** Si  $C$  est un convexe de  $\mathbb{R}^n$ , si  $f$  est une fonction convexe de  $C$  sur  $\mathbb{R}$  et si  $\varphi$  est une fonction convexe croissante sur  $\mathbb{R}$ , alors  $\varphi \circ f$  est une fonction convexe.

*Preuve :*

**1)a)**  $f_i(\lambda a + (1 - \lambda)b) \leq \lambda f_i(a) + (1 - \lambda)f_i(b) \leq \lambda(\sup f_i)(a) + (1 - \lambda)(\sup f_i)(b)$ . Ceci est vérifié pour tout  $i \in I$  donc  $(\sup f_i)(\lambda a + (1 - \lambda)b) \leq \lambda(\sup f_i)(a) + (1 - \lambda)(\sup f_i)(b)$ .

**1)b)**  $\lambda_i \geq 0$  donc  $\lambda_i f_i(\lambda a + (1 - \lambda)b) \leq \lambda_i \lambda f_i(a) + \lambda_i (1 - \lambda) f_i(b)$ . En faisant la somme, on a :

$$\left(\sum_i \lambda_i f_i\right)(\lambda a + (1 - \lambda)b) \leq \lambda \left(\sum_i \lambda_i f_i\right)(a) + (1 - \lambda) \left(\sum_i \lambda_i f_i\right)(b).$$

**2)**  $f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b)$  et, en utilisant la croissance, puis la convexité de  $\varphi$ , on a :

$$\begin{aligned} \varphi(f(\lambda a + (1 - \lambda)b)) &\leq \varphi(\lambda f(a) + (1 - \lambda)f(b)) \\ &\leq \lambda \varphi(f(a)) + (1 - \lambda) \varphi(f(b)). \quad \square \end{aligned}$$

## 4.2- Caractérisation des fonctions convexes différentiables

**Théorème 4.1:** Soit  $U$  un ouvert convexe de  $\mathbb{R}^n$  et  $f$  une fonction de  $U$  sur  $\mathbb{R}$ .

**1)** Si  $f$  est différentiable sur  $U$ , alors  $f$  est convexe si et seulement si, pour tout  $(x, y) \in U^2$ ,

$$f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle.$$

**2)** Si  $f$  est de classe  $\mathcal{C}^2$  sur  $U$ , alors  $f$  est convexe sur  $U$  si et seulement si, pour tout  $x \in U$ , la matrice  $\nabla^2 f(x)$  est positive.

*Preuve :*

**1)** Si  $f$  est convexe, pour  $\theta \in ]0, 1[$ , on a :

$$f(x + \theta(y - x)) = f((1 - \theta)x + \theta y) \leq (1 - \theta)f(x) + \theta f(y)$$

d'où  $f(x + \theta(y - x)) - f(x) \leq \theta(f(y) - f(x))$ .

Or  $f(x + \theta(y - x)) = f(x) + \theta d_x f(y - x) + \theta \|y - x\| \varepsilon(\theta(y - x))$ .

En simplifiant par  $\theta > 0$ , on obtient :

$$d_x f(y - x) + \|y - x\| \varepsilon(\theta(y - x)) \leq f(y) - f(x)$$

et en faisant tendre  $\theta$  vers 0, on a alors :

$$d_x f(y - x) \leq f(y) - f(x).$$

Réciproquement, si  $f(b) \geq f(a) + d_a f(b - a)$  pour tout  $(a, b) \in U^2$ , alors, avec  $b = y$  et  $a = x + \theta(y - x)$ , puis avec  $b = x$  et  $a = x + \theta(y - x)$ ,

$$f(y) \geq f(x + \theta(y - x)) + d_{x + \theta(y - x)} f(y - x) \times (1 - \theta)$$

$$f(x) \geq f(x + \theta(y - x)) + d_{x + \theta(y - x)} f(y - x) \times (-\theta).$$

On multiplie la première inégalité par  $\theta$ , la deuxième par  $(1 - \theta)$  et on fait la somme :

$$\theta f(y) + (1 - \theta)f(x) \geq (\theta + 1 - \theta)f(x + \theta(y - x))$$

soit  $f((1 - \theta)x + \theta y) \leq (1 - \theta)f(x) + \theta f(y)$ .

**2)** On applique la formule de Taylor-Mac-Laurin à l'ordre 2 en  $x$  :  
il existe  $\theta \in ]0, 1[$  tel que :

$$f(y) = f(x) + d_x f(y - x) + \frac{1}{2} d_{x + \theta(y - x)}^2 f(y - x, y - x).$$

Or  $d_{x + \theta(y - x)}^2 f(y - x, y - x) \geq 0$  donc  $f(y) \geq f(x) + d_x f(y - x)$  et  $f$  est convexe.

Réciproquement, si  $f$  est convexe, alors, pour tout  $h \in \mathbb{R}^n$  et pour tout  $t$  tel que  $x + th \in U$ ,

$$f(x + th) \geq f(x) + d_x f(th).$$

Or  $f(x + th) = f(x) + d_x f(th) + \frac{1}{2}t^2 d_x^2 f(h, h) + t^2 \|h\|^2 \varepsilon(th)$ . D'où :

$$d_x^2 f(h, h) + 2\|h\|^2 \varepsilon(th) \geq 0$$

et en faisant  $t \rightarrow 0$ , on obtient finalement  $d_x^2 f(h, h) \geq 0$ .

□

### 4.3- Minimum global d'une fonction convexe.

**Théorème 4.2:** Soit  $C$  un convexe de  $\mathbb{R}^n$ ,  $f$  une fonction convexe de  $C$  sur  $\mathbb{R}$  et  $a \in C$ , alors

- 1) un minimum local est un minimum global ;
- 2) si  $f$  est de classe  $\mathcal{C}^1$  sur  $C$  et si  $C$  est ouvert, alors  $a \in \text{Arg}_C \min f$  si et seulement si  $\nabla f(a) = 0$ .

*Preuve :*

1)  $f(x + \theta(y - x)) \leq (1 - \theta)f(x) + \theta f(y)$  donc  $f(x + \theta(y - x)) - f(x) \leq \theta(f(y) - f(x))$ .

Si  $x$  est un minimum local, alors, pour  $\theta$  assez petit, on a  $f(x + \theta(y - x)) - f(x) \geq 0$ . Donc  $f(y) - f(x) \geq 0$  et ceci pour  $y$  quelconque. Donc  $x$  est un minimum global.

2) Si  $a \in \text{Arg}_C \min f$ , alors  $f(y) \geq f(a)$  pour tout  $y \in C$ . Or  $C$  est ouvert, donc  $a$  est aussi minimum local et  $\nabla f(a) = 0$ .

Réciproquement, si  $\nabla f(a) = 0$ , i.e.  $d_a f = 0$ , comme  $f(y) \geq f(a) + d_a f(y - a)$  pour tout  $y \in C$ , on a donc, pour tout  $y \in C$ ,  $f(y) \geq f(a)$  et  $a$  est un minimum global de  $f$  sur  $C$ .

□

## Chapitre 5

# Algorithmes pour l'optimisation sans contraintes

On s'intéresse ici à la résolution pratique de problèmes d'optimisation, c'est-à-dire à la mise en œuvre d'algorithmes pour calculer une approximation de la solution optimale.

La plupart des algorithmes considérés ont la forme générale suivante :

**Proposition 1 :** (Forme générale des méthodes d'optimisation considérées)

$$x_0 \text{ étant donné, on calcule } x_{k+1} = x_k + \rho_k d_k.$$

Le vecteur  $d_k$  s'appelle direction de descente et  $\rho_k$  le pas associé.

On cherche alors, en général, à avoir  $f(x_{k+1}) \leq f(x_k)$ .

De tels algorithmes sont ainsi souvent appelés méthodes de descente.

Essentiellement, la différence entre les méthodes de descente réside dans le choix de la direction  $d_k$ . Une fois cette direction choisie, le calcul du pas  $\rho_k$  est ensuite un problème unidimensionnel.

### 5.1- Choix du pas de descente

On pourrait choisir  $\rho_k$  en cherchant à minimiser la fonction :

$$q(\rho) = f(x_k + \rho d_k).$$

Quelle que soit la direction de descente, on aurait alors nécessairement

$$f(x_{k+1}) \leq f(x_k).$$

Cependant, même si ce n'est plus qu'en une seule dimension, il faudrait donc résoudre un problème de minimisation à chaque étape de l'algorithme de descente.

Le pas optimal  $\rho^*$  est alors cherché comme l'annulateur de la dérivée de la fonction  $q$  qui s'écrit

$$q'(\rho) = \langle \nabla f(x_k + \rho d_k), d_k \rangle.$$

Ceci est en général très peu rentable en termes de temps de calcul par rapport à la qualité du résultat obtenu et n'est donc en pratique jamais utilisé.

En pratique, on préfère le compromis qui consiste à choisir une approximation “suffisamment satisfaisante” de ce pas optimal, mais pour un effort en temps de calcul qui sera réduit.

Il existe de nombreuses méthodes pour effectuer ce compromis et on vous en propose ici un exemple particulièrement simple ainsi qu’un exemple assez efficace très souvent utilisé dans la pratique. Pour en savoir plus sur le sujet, le nom anglais de la recherche du pas de descente est “line search”.

### Méthode de dichotomie

Une première méthode de calcul du pas de descente  $\rho_k$  est la méthode de dichotomie qui est basée sur le théorème des valeurs intermédiaires. L’approximation du pas optimal obtenue en temps raisonnable est plutôt grossière mais a le mérite d’être particulièrement simple à implémenter :

1. On détermine un intervalle  $[a, b]$  qui contient un zéro de  $q'$
2. On pose  $c = \frac{a+b}{2}$
3. Si  $q'(a) \cdot q'(c) \leq 0$  on pose  $b = c$   
Sinon, on pose  $a = c$
4. On boucle les étapes 2 et 3 autant de fois que voulu
5. On choisit  $\rho_k = c$

### Remarques :

1. Comment terminer les boucles d’algorithmes itératifs de recherche d’une valeur *a priori* inconnue ? Ici, on sait que le pas optimal  $\rho^*$  est par construction dans l’intervalle  $[a, b]$  qui est de longueur  $\alpha/2^N$ , où  $\alpha$  est la longueur de l’intervalle initial et  $N$  est le nombre de découpes successives. On peut donc fixer à l’avance un nombre d’étapes de dichotomie et connaître la précision avec laquelle  $c$  sera une approximation de  $\rho^*$ .

2. Quand on ne dispose pas précisément de cette vitesse de convergence, une autre méthode est de s’arrêter quand la variation de la variable recherchée d’une étape de la boucle sur l’autre devient petite. Cependant, “petite” est un terme subjectif et il convient alors d’utiliser une variation relative : on s’arrête quand  $\|\rho_{k+1} - \rho_k\|/\|\rho_k\|$  devient plus petit qu’un certain seuil fixé.

### Règles de Wolfe

On choisit deux nombres  $m_1$  et  $m_2$  avec  $0 < m_1 < m_2 < 1$  (par exemple  $m_1 = 0.1$  et  $m_2 = 0.7$ ) et on recherche  $\rho$  qui vérifie

$$\begin{cases} q(\rho) \leq q(0) = m_1 \rho q'(0) \\ q'(\rho) \geq m_2 q'(0) \end{cases}$$



1. - On choisit  $m_1$  et  $m_2$  avec  $0 < m_1 < m_2 < 1$   
 - On pose  $\rho = 1$ ,  $\rho_- = 0$ ,  $\rho_+ = 0$
2. - Si  $q(\rho) \leq q(0) + m_1 \rho q'(0)$  et  $q'(\rho) \geq m_2 q'(0)$  alors  
 - Choisir  $\rho_k = \rho$  et quitter la boucle  
 - Sinon  
 - Si  $q(\rho) > q(0) + m_1 \rho q'(0)$  alors  
 - Poser  $\rho_+ = \rho$   
 - Sinon, Si  $q(\rho) \leq q(0) + m_1 \rho q'(0)$  et  $q'(\rho) < m_2 q'(0)$  alors  
 - Poser  $\rho_- = \rho$   
 - Fin si  
 - Fin si
3. - Si  $\rho_+ = 0$  alors  
 - Choisir  $\rho_- > \rho$   
 - Sinon, Si  $\rho_+ > 0$  alors  
 - Choisir  $\rho \in ]\rho_-; \rho_+[$   
 - Fin si
4. On boucle les étapes 2 et 3

## 5.2-Choix de la direction de descente

### Définition :

On dit que  $d \in \mathbb{R}^n$  est une direction de descente pour  $f$  en un point  $a \in \mathbb{R}^n$  si  $\langle \nabla f(a), d \rangle < 0$

### Proposition 2 :

Si  $d$  est une direction de descente pour  $f$  en  $a$ , alors on peut trouver  $x$  au voisinage de  $a$  tel que  $f(x) < f(a)$ , plus précisément, on peut trouver un réel  $\alpha > 0$  tel que  $f(a + \rho d) < f(a)$  pour tout  $\rho \in ]0, \alpha[$ .

*Preuve :*

Pour  $\rho > 0$ , on peut écrire,  $d$  étant fixé,  $f(a + \rho d) = f(a) + \rho[\langle \nabla f(a), d \rangle + \varepsilon(\rho)]$  avec  $\lim_{\rho \rightarrow 0} \varepsilon(\rho) = 0$ . Comme  $\langle \nabla f(a), d \rangle < 0$ , pour  $\rho > 0$  assez petit (c'est-à-dire  $\rho \in ]0, \alpha[$ ), on a  $\langle \nabla f(a), d \rangle + \varepsilon(\rho) < 0$ , donc  $f(a + \rho d) < f(a)$ .

□

## Méthode de plus profonde descente

Une direction de descente naturelle est alors

$$d_k = -\nabla f(x_k).$$

Cependant, ce choix de direction se révèle en pratique souvent lent à converger et on lui préfère une méthode de gradient conjugué ou de Newton.

## Méthode de Newton

La méthode de Newton pour trouver le minimum de  $f$  est en fait la méthode de Newton pour trouver le zéro de  $\nabla f$ . La direction de descente est alors donnée par

$$d_k = -\left(Hf(x_k)\right)^{-1}\nabla f(x_k).$$

**Proposition :** On suppose que  $f \in \mathcal{C}^2$  et que  $x^*$  est un minimum local strict de  $f$  ( $Hf(x^*)$  est définie positive). Pour  $x$  au voisinage de  $x^*$ , le vecteur  $d = -\left(Hf(x)\right)^{-1}\nabla f(x)$  est alors une direction de descente.

*Preuve :*

On a  $\langle \nabla f(x), d \rangle = -\langle Hf(x) \cdot d, d \rangle < 0$  car  $Hf(x)$  est une matrice définie positive. En effet,  $Hf(x^*)$  est définie positive et  $Hf$  est continue. Donc, on peut trouver un voisinage de  $x^*$  sur lequel  $Hf(x)$  est symétrique et définie positive. □

Quand le point de départ  $x_0$  est proche de la solution optimale  $x^*$ , la convergence avec cette direction de descente est généralement très bonne. Cependant, il faut calculer la dérivée seconde de  $f$  et l'inverser à chaque étape. Ceci peut être très difficile, voire simplement inaccessible quand la fonction coût qu'on minimise n'est connue qu'à travers des mesures ponctuelles d'un système complexe.

## Méthodes de Quasi-Newton

Le principe des méthodes dites de "Quasi-Newton", est de calculer rapidement un remplaçant satisfaisant à  $\left(Hf(x)\right)^{-1}$ , l'inverse de la Hessienne. Généralement on aura en fait une suite de matrices  $W_k$ , avec  $W_0 = Id$ , qui seront corrigées à chaque étape ( $W_{k+1} = W_k + B_k$ ) pour se rapprocher itérativement de la matrice  $\left(Hf(x_k)\right)^{-1}$ . On approche ensuite  $x^*$  par la suite

$$x_{k+1} = x_k - \rho_k W_k \nabla f(x_k).$$

Il existe bien sur de nombreuses stratégies pour choisir le terme correcteur  $B_k$ . Parmi celles-ci, la méthode BFGS, où la matrice  $W_k$  est donnée par

$$W_k = W_{k-1} - \frac{sy^T W_k + W_k y s^T}{y^T s} + \left(1 + \frac{y^T W_k y}{y^T s}\right) \frac{ss^T}{y^T s},$$

où  $s = x_k - x_{k-1}$  et  $y = \nabla f(x_k) - \nabla f(x_{k-1})$ , est considérée comme une des plus robustes.

---

## Chapitre 6

### Optimisation avec contraintes

Problème à résoudre : On considère  $f$  et  $(\varphi_i)_{i \in I}$ , fonctions de classe  $\mathcal{C}^1$  de  $\mathbb{R}^n$  sur  $\mathbb{R}$ . Soit  $\mathcal{P}$  le problème : minimiser  $f$  sur  $X = \{x \in \mathbb{R}^n ; \varphi_i(x) \leq 0 \text{ pour tout } i \in I\}$ .

#### 6.1- Deux résultats auxiliaires

**Théorème 6.1** (*Théorème de projection sur  $\mathbb{R}^n$* )

Soit  $C$  un convexe non vide fermé de  $\mathbb{R}^n$ .

Pour tout  $b \in \mathbb{R}^n$ , il existe un et un seul élément  $p(b)$  de  $\mathbb{R}^n$  tel que  $p(b) \in C$  et  $\|b - p(b)\| = \inf_{c \in C} \|b - c\|$ . Cet élément vérifie  $\langle p(b) - b, p(b) - c \rangle \leq 0$  pour tout  $c \in C$  et réciproquement, si  $\langle u - b, u - c \rangle \leq 0$  pour tout  $c \in C$ , alors  $u = p(b)$ .

*Preuve* : voir l'exercice 11.

**Théorème 6.2** (*lemme de Farkas-Minkowski*)

Soit  $I$  un ensemble fini,  $(a_i)_{i \in I}$  et  $b$  des éléments de  $\mathbb{R}^n$ . Alors

$$\{w \in \mathbb{R}^n ; \langle a_i, w \rangle \geq 0 \text{ pour tout } i \in I\} \subset \{w \in \mathbb{R}^n ; \langle b, w \rangle \geq 0\}$$

si et seulement si il existe une famille  $(\lambda_i)_{i \in I}$  de réels positifs, tels que  $b = \sum_{i \in I} \lambda_i a_i$ .

*Preuve* :

• Si  $b = \sum_{i \in I} \lambda_i a_i$  et si  $\langle a_i, w \rangle \geq 0$  pour tout  $i \in I$  alors  $\lambda_i \langle a_i, w \rangle \geq 0$  pour tout  $i \in I$  car  $\lambda_i \geq 0$  et  $\sum_{i \in I} \lambda_i \langle a_i, w \rangle = \langle \sum_{i \in I} \lambda_i a_i, w \rangle = \langle b, w \rangle \geq 0$ .

• La réciproque est beaucoup moins évidente et se fait en plusieurs étapes.

Soit  $C = \{\sum_{i \in I} \lambda_i a_i ; \lambda_i \geq 0 \text{ pour tout } i \in I\}$ .

i) Montrons que  $C$  est un cône convexe fermé de sommet 0.

→ Si  $x \in C$ ,  $x = \sum_{i \in I} \lambda_i a_i$  et  $\lambda x = \sum_{i \in I} (\lambda \lambda_i) a_i$  avec  $\lambda \lambda_i \geq 0$  si  $\lambda \geq 0$ , donc si  $x \in C$  et si  $\lambda \geq 0$ , alors  $\lambda x \in C$  :  $C$  est bien un cône de sommet 0.

→ Si  $(x, y) \in C^2$  et si  $\theta \in [0, 1]$ ,  $x = \sum_{i \in I} \lambda_i a_i$ ,  $y = \sum_{i \in I} \mu_i a_i$ ,  $\lambda_i, \mu_i \geq 0$  et  $\theta x + (1 - \theta)y = \sum_{i \in I} (\theta \lambda_i + (1 - \theta)\mu_i) a_i$  avec  $\theta \lambda_i + (1 - \theta)\mu_i \geq 0$  donc  $\theta x + (1 - \theta)y \in C$  et  $C$  est convexe.

→ Si les  $a_i$  sont linéairement indépendants,  $C$  est isomorphe à  $(\mathbb{R}_+)^{\text{card } I}$  et c'est donc un fermé.

→  $C$  fermé est plus difficile lorsque les  $a_i$  ne sont pas linéairement indépendants. Il existe alors des réels  $\mu_i$  tels que  $\sum_{i \in I} \mu_i a_i = 0$  avec  $J = \{i ; \mu_i < 0\} \neq \emptyset$ . Dans ces conditions, tout vecteur  $v = \sum_{i \in I} \lambda_i a_i$  s'écrit aussi  $v = \sum_{i \in I} (\lambda_i + t\mu_i) a_i$  où  $t$  est quelconque.

Posons  $t_0 = \min_{j \in J} \left( -\frac{\lambda_j}{\mu_j} \right)$ . Alors l'un au moins des  $\lambda_i + t_0 \mu_i$  est nul et, pour tout  $i \in I$ ,  $\lambda_i + t_0 \mu_i \geq 0$ . Donc  $C = \bigcup_{j \in J} \left\{ \sum_{i \in I \setminus \{j\}} \lambda_i a_i ; \lambda_i \geq 0 \text{ pour tout } i \in I \setminus \{j\} \right\}$ .

On recommence ce raisonnement jusqu'à ce que  $C$  soit décrit comme une réunion finie de cônes associés à des vecteurs linéairement indépendants et donc  $C$  est bien fermé.

ii) On va appliquer le Théorème 6.1 à  $C$  : on va montrer que si  $b \notin C$ , alors l'inclusion est fautive, c'est-à-dire qu'il existe  $u \in \mathbb{R}^n$  tel que  $\langle a_i, u \rangle \geq 0$  pour tout  $i \in I$  et  $\langle b, u \rangle < 0$ .

D'après le Théorème 6.1, on a, pour tout  $v \in C$ ,  $\langle p(b) - b, p(b) - v \rangle \leq 0$ .

Par suite,  $\langle v, p(b) - b \rangle \geq \langle p(b) - b, p(b) \rangle = \|p(b) - b\|^2 + \langle p(b) - b, b \rangle$ .

On a  $\|p(b) - b\|^2 > 0$  car  $p(b) \neq b$  ( $b \notin C$ ).

Soit  $\alpha \in ]\langle p(b) - b, b \rangle, \langle p(b) - p, p(b) \rangle[$  et  $u = p(b) - b$ . On a donc  $\langle v, u \rangle > \alpha$  pour tout  $v \in C$  et  $\langle b, u \rangle < \alpha$ .

En particulier, comme  $0 \in C$ ,  $\alpha < 0$  et  $\langle b, u \rangle < 0$ . D'autre part, pour tout  $i \in I$  et pour tout  $\lambda > 0$ ,  $\lambda a_i \in C$  donc  $\langle a_i, u \rangle > \frac{\alpha}{\lambda}$  et en faisant  $\lambda \rightarrow +\infty$ , on obtient  $\langle a_i, u \rangle \geq 0$ . D'où le résultat.  $\square$

## 6.2- Les relations de Kuhn-Tucker

**Définition 6.1:** Soit  $X = \{x \in \mathbb{R}^n ; \varphi_i(x) \leq 0 \text{ pour tout } i \in I\}$ . Un arc de courbe  $\Gamma$  défini par  $\gamma : \mathbb{R}_+ \rightarrow \mathbb{R}^n$  de classe  $\mathcal{C}^1$  est dit admissible en  $\bar{x} \in X$  si  $\gamma(0) = \bar{x}$  et si  $\gamma(\theta) \in X$  pour  $\theta$  assez petit. On appelle alors direction admissible en  $\bar{x}$  le vecteur  $v = \gamma'(0)$ .

On note  $C(\bar{x})$  la fermeture du cône formé par l'ensemble des directions admissibles en  $\bar{x}$ .

**Théorème 6.3 :** Si  $f$  admet en  $\bar{x}$  un minimum local sur  $X$ , alors  $\langle \nabla f(\bar{x}), w \rangle \geq 0$  pour tout  $w \in C(\bar{x})$ .

*Preuve :*

Si  $w$  est direction admissible en  $\bar{x}$ , il existe  $\gamma : \mathbb{R}_+ \rightarrow \mathbb{R}^n$  de classe  $\mathcal{C}^1$  telle que  $\gamma(\theta) = \bar{x} + \theta w + \theta \varepsilon(\theta)$  et  $f(\gamma(\theta)) = f(\gamma(0)) + (f \circ \gamma)'(0)\theta + \theta \varepsilon(\theta)$  avec  $f(\gamma(\theta)) = f(\gamma_1(\theta), \dots, \gamma_n(\theta))$  d'où  $(f \circ \gamma)'(\theta) = \sum_{i=1}^n D_i f(\gamma(\theta)) \gamma'_i(\theta)$  donc, pour  $\theta = 0$ ,  $(f \circ \gamma)'(0) = \langle \nabla f(\gamma(0)), \gamma'(0) \rangle = \langle \nabla f(\bar{x}), w \rangle$ .

Par suite,  $f(\gamma(\theta)) = f(\bar{x}) + \theta(\langle \nabla f(\bar{x}), w \rangle + \varepsilon(\theta))$ .

Comme  $\gamma(\theta) \in X$  et que  $\bar{x}$  est minimum local sur  $X$ , on a  $f(\gamma(\theta)) - f(\bar{x}) \geq 0$  et donc  $\langle \nabla f(\bar{x}), w \rangle + \varepsilon(\theta) \geq 0$  pour tout  $\theta$  assez petit. En faisant  $\theta \rightarrow 0$ , on obtient  $\langle \nabla f(\bar{x}), w \rangle \geq 0$ .

Si  $w$  n'est pas direction admissible,  $w = \lim_n w_n$  avec  $w_n$  direction admissible, donc  $\langle \nabla f(\bar{x}), w \rangle = \lim_n \langle \nabla f(\bar{x}), w_n \rangle \geq 0$ .

□

**Définition 6.2:** Si  $x \in X$ , on note  $I(x) = \{i \in I ; \varphi_i(x) = 0\}$  et on appelle  $I(x)$  l'ensemble des contraintes saturées en  $x$ .  
On pose  $C^*(x) = \{w \in \mathbb{R}^n ; \langle \nabla \varphi_i(x), w \rangle \leq 0 \text{ pour tout } i \in I(x)\}$ .

**Théorème 6.4:** On a  $C(x) \subset C^*(x)$  mais la réciproque est fausse en général.

*Preuve :*

En remplaçant  $f$  par  $\varphi_i$  dans le Théorème 6.3, on a :

$$\varphi_i(\gamma(\theta)) = \varphi_i(x) + \theta(\langle \nabla \varphi_i(x), w \rangle + \varepsilon(\theta)).$$

Si  $i \in I(x)$ , alors  $\varphi_i(x) = 0$  et  $\varphi_i(\gamma(\theta)) \leq 0$  car  $\gamma(\theta) \in X$ .

En faisant  $\theta \rightarrow 0$ , on obtient  $\langle \nabla \varphi_i(x), w \rangle \leq 0$ .

(Ceci pour  $w$  direction admissible, sinon,  $w = \lim_n w_n$  et on procède de même que précédemment.)

Donc, si  $w \in C(x)$ , alors  $w \in C^*(x)$ .

La réciproque n'est pas toujours vraie.

□

◇ **Exemple sur  $\mathbb{R}^2$  :**  $\varphi_1(x) = -x_1$ ,  $\varphi_2(x) = -x_2$ ,  $\varphi_3(x) = -(1 - x_1)^3 + x_2$

$$X = \{x \in \mathbb{R}^2 ; x_1 \geq 0, x_2 \geq 0, x_2 \leq (1 - x_1)^3\}$$

Pour  $x = (1, 0)$ ,  $I(x) = \{2, 3\}$ ,  $\nabla \varphi_2(x) : \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ ,  $\nabla \varphi_3(x) : \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ .

On a donc  $C^*(x) = \{w = (w_1, 0) ; w_1 \in \mathbb{R}\} : (1, 0) \in C^*(x)$  et  $(1, 0) \notin C(x)$  car sinon on aurait  $(1 + \theta, 0) \in X$ , ce qui est faux. ◇

**Définition 6.3 :** On dit que les contraintes sont qualifiées en  $x$  si  $C(x) = C^*(x)$ .

La condition précédente étant très peu aisée à vérifier, on donnera ici des conditions suffisantes (mais non nécessaires) de qualification des contraintes (QC) :

**Lemme admis :** Pour que (QC) soit vérifiée en tout point  $x$  de  $X$ , il suffit que les  $\varphi_i$  soient linéaires, ou convexes si  $X$  est d'intérieur non vide.  
Pour que (QC) soit vérifiée en  $\bar{x}$ , il suffit que la famille  $(\nabla \varphi_i(\bar{x}))_{i \in I(\bar{x})}$  soit libre.

**Remarque :** Les conditions du lemme précédent peuvent être combinées. Ainsi, par exemple,

- 1) (QC) est vérifiée en tout  $x \in X$  si les  $\varphi_i$  sont, soit linéaires, soit non linéaires convexes et s'il existe  $\bar{x} \in X$  tel que  $\varphi_i(\bar{x}) < 0$  pour tout  $\varphi_i$  non linéaire.  
 2) (QC) est vérifiée en  $\bar{x}$  si les  $\varphi_i$  sont, soit linéaires, soit non linéaires avec  $(\nabla \varphi_i(\bar{x}))_{i \in I(\bar{x}) \setminus \{j; \varphi_j \text{ linéaire}\}}$  libre.

**Théorème 6.5** (Kuhn-Tucker)

Si  $f$  admet un minimum en  $\bar{x} \in X = \{x; \varphi_i(x) \leq 0 \text{ pour tout } i \in I\}$ , et si les contraintes sont qualifiées en  $\bar{x}$ , alors il existe une famille  $(\lambda_i(\bar{x}))_{i \in I(\bar{x})}$  de réels positifs ou nuls tels que :

$$\nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i(\bar{x}) \nabla \varphi_i(\bar{x}) = 0.$$

*Preuve :*

D'après le Théorème 6.3,  $\langle \nabla f(\bar{x}), w \rangle \geq 0$  pour tout  $w \in C(\bar{x})$ . Or

$$\begin{aligned} C(\bar{x}) = C^*(\bar{x}) &= \{w; \langle \nabla \varphi_i(\bar{x}), w \rangle \leq 0 \text{ pour tout } i \in I(\bar{x})\} \\ &= \{w; \langle -\nabla \varphi_i(\bar{x}), w \rangle \geq 0 \text{ pour tout } i \in I(\bar{x})\} \end{aligned}$$

On a donc  $\{w; \langle -\nabla \varphi_i(\bar{x}), w \rangle \geq 0 \text{ pour tout } i \in I(\bar{x})\} \subset \{w; \langle \nabla f(\bar{x}), w \rangle \geq 0\}$  donc, d'après le Théorème 6.2, il existe une famille  $(\lambda_i(\bar{x}))_{i \in I(\bar{x})}$  de réels positifs ou nuls tels que  $\nabla f(\bar{x}) = - \sum_{i \in I(\bar{x})} \lambda_i(\bar{x}) \nabla \varphi_i(\bar{x})$ . □

**Remarques :**

- 1) En convenant que  $\lambda_i(\bar{x}) = 0$  si  $i \notin I(\bar{x})$ , le vecteur  $\lambda(\bar{x})$  est appelé vecteur de Kuhn-Tucker en  $\bar{x}$ .  
 2) Si  $(\nabla \varphi_i(\bar{x}))_{i \in I(\bar{x})}$  est libre, alors  $\lambda(\bar{x})$  est unique.

### 6.3- Extension à des problèmes avec contraintes égalités

Problème : Minimiser  $f$  sous les contraintes :

- $\rightarrow \varphi_i(x) \leq 0$  pour tout  $i \in I$  ;  
 $\rightarrow \psi_j(x) = 0$  pour tout  $j \in J$  ;

les fonctions  $f, \varphi_i, \psi_j$  étant de classe  $\mathcal{C}^1$ .

On note  $X = \{x \in \mathbb{R}^n; \varphi_i(x) \leq 0 \text{ pour tout } i \in I \text{ et } \psi_j(x) = 0 \text{ pour tout } j \in J\}$  et on pose  $I(\bar{x}) = \{i \in I; \varphi_i(\bar{x}) = 0\}$  et

$$C^*(\bar{x}) = \{w \in \mathbb{R}^n; \langle \nabla \varphi_i(\bar{x}), w \rangle \leq 0 \text{ pour tout } i \in I \text{ et } \langle \nabla \psi_j(\bar{x}), w \rangle = 0 \text{ pour tout } j \in J\}.$$

On définit de même les directions admissibles et  $C(\bar{x})$  et on a l'analogie du théorème de Kuhn-Tucker :

**Théorème 6.6:** Si  $f$  admet sur  $X$  un minimum en  $\bar{x}$  en lequel les contraintes sont qualifiées, alors il existe des nombres positifs ou nuls  $\lambda_i(\bar{x})$  pour  $i \in I$  et des nombres  $\mu_j(\bar{x})$  (non nécessairement positifs) pour  $j \in J$  tels que :

$$\nabla f(\bar{x}) + \sum_{i \in I} \lambda_i(\bar{x}) \nabla \varphi_i(\bar{x}) + \sum_{j \in J} \mu_j(\bar{x}) \nabla \psi_j(\bar{x}) = 0$$

et  $\lambda_i(\bar{x}) \varphi_i(\bar{x}) = 0$  pour tout  $i \in I$ .

Autre problème : Minimiser  $f$  sous les contraintes  $\psi_j(x) = 0$  pour tout  $j \in J$  ; les fonctions  $f, \psi_j$  étant de classe  $\mathcal{C}^1$  (contraintes égalités seulement).

On note  $X = \{x \in \mathbb{R}^n ; \psi_j(x) = 0 \text{ pour tout } j \in J\}$ .

**Théorème 6.7:** Si  $f$  admet sur  $X$  un minimum en  $\bar{x}$  en lequel les contraintes sont qualifiées, alors il existe une famille de nombres  $(\mu_j(\bar{x}))_{j \in J}$  (non nécessairement positifs) telle que :

$$\nabla f(\bar{x}) + \sum_{j \in J} \mu_j(\bar{x}) \nabla \psi_j(\bar{x}) = 0.$$

**Remarque :** on retrouve les multiplicateurs de Lagrange et le théorème des extrema liés.

◇ **Exemples :**

1) Extrémums de  $f : (x, y) \mapsto ax + by + c$  sous la contrainte  $\varphi(x, y) = x^2 + y^2 - 1 = 0$ .

$$\nabla f(x, y) : \begin{pmatrix} a \\ b \end{pmatrix} \quad ; \quad \nabla \varphi(x, y) : \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

Condition nécessaire :  $\nabla f(\bar{x}, \bar{y}) + \mu \nabla \varphi(\bar{x}, \bar{y}) = 0$  soit  $\begin{cases} a + 2\mu\bar{x} = 0 \\ b + 2\mu\bar{y} = 0 \end{cases}$ .

On trouve  $\bar{x} = -\frac{a}{2\mu}$  et  $\bar{y} = -\frac{b}{2\mu}$ . On a aussi  $\varphi(\bar{x}, \bar{y}) = 0$  donc  $a^2 + b^2 = 4\mu^2$  et  $\mu = \pm \frac{\sqrt{a^2 + b^2}}{2}$ .

• Si  $\mu = \frac{\sqrt{a^2 + b^2}}{2}$ , alors  $\bar{x} = -\frac{a}{\sqrt{a^2 + b^2}}$ ,  $\bar{y} = -\frac{b}{\sqrt{a^2 + b^2}}$  et  $f(\bar{x}, \bar{y}) = c - \sqrt{a^2 + b^2}$ .

• Si  $\mu = -\frac{\sqrt{a^2 + b^2}}{2}$ , alors  $\bar{x} = \frac{a}{\sqrt{a^2 + b^2}}$ ,  $\bar{y} = \frac{b}{\sqrt{a^2 + b^2}}$  et  $f(\bar{x}, \bar{y}) = c + \sqrt{a^2 + b^2}$ .

$\varphi$  étant convexe, la contrainte est qualifiée. De plus,  $X$  est fermé borné (cercle unité) et  $f$  est continue donc les extrémums existent.

On a donc  $\min_X f = c - \sqrt{a^2 + b^2}$ ,  $\max_X f = c + \sqrt{a^2 + b^2}$ , avec

$$\text{Arg}_X \min f = \left\{ \left( -\frac{a}{\sqrt{a^2 + b^2}}, -\frac{b}{\sqrt{a^2 + b^2}} \right) \right\} \text{ et } \text{Arg}_X \max f = \left\{ \left( \frac{a}{\sqrt{a^2 + b^2}}, \frac{b}{\sqrt{a^2 + b^2}} \right) \right\}.$$

2) Extrémums de  $f : (x, y, z) \mapsto x + y + z$  sous les contraintes  $\varphi_1(x, y, z) = x^2 + z^2 - 1 \leq 0$  et  $\varphi_2(x, y, z) = y^2 + z^2 - 1 \leq 0$ .

$f$  est continue,  $X = \varphi_1^{-1}(]-\infty, 0]) \cap \varphi_2^{-1}(]-\infty, 0])$  est fermé comme intersection de 2 fermés (images réciproques de  $]-\infty, 0]$  fermé par les  $\varphi_i$  qui sont continues), et  $X$  est borné, car si

$(x, y, z) \in X$ ,  $x^2 + y^2 + z^2 \leq 2$ .

On a donc existence des extrémums de  $f$  sur  $X$ .

$$\nabla f(x, y, z) : \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad ; \quad \nabla \varphi_1(x, y, z) : \begin{pmatrix} 2x \\ 0 \\ 2z \end{pmatrix} \quad ; \quad \nabla \varphi_2(x, y, z) : \begin{pmatrix} 0 \\ 2y \\ 2z \end{pmatrix}.$$

$$\text{Conditions de Kuhn-Tucker : } \begin{cases} 1 + 2\lambda_1 \bar{x} = 0 \\ 1 + 2\lambda_2 \bar{y} = 0 \\ 1 + 2(\lambda_1 + \lambda_2) \bar{z} = 0 \\ \lambda_1(\bar{x}^2 + \bar{z}^2 - 1) = 0 \\ \lambda_2(\bar{y}^2 + \bar{z}^2 - 1) = 0 \end{cases}.$$

On trouve  $\bar{x} = -\frac{1}{2\lambda_1}$  et  $\bar{y} = -\frac{1}{2\lambda_2}$ .

On a nécessairement  $\lambda_1 \neq 0$  et  $\lambda_2 \neq 0$  donc  $\varphi_1(\bar{x}, \bar{y}, \bar{z}) = 0$  et  $\varphi_2(\bar{x}, \bar{y}, \bar{z}) = 0$  soit

$$\begin{cases} \bar{x}^2 + \bar{z}^2 - 1 = 0 \\ \bar{y}^2 + \bar{z}^2 - 1 = 0 \end{cases}.$$

En faisant la différence, on obtient  $\bar{x}^2 = \bar{y}^2$  et, comme  $\bar{x}$  et  $\bar{y}$  sont de même signe, on a alors  $\bar{x} = \bar{y}$ , donc  $\lambda_1 = \lambda_2 = \lambda$ , puis  $\bar{x} = \bar{y} = -\frac{1}{2\lambda}$  et  $\bar{z} = -\frac{1}{4\lambda} = \frac{\bar{x}}{2}$ . Enfin  $\bar{x}^2(1 + \frac{1}{4}) - 1 = 0$  d'où  $\bar{x} = \pm \frac{2}{\sqrt{5}}$ .

On a donc  $\min_X f = -\sqrt{5}$ ,  $\max_X f = \sqrt{5}$ , avec

$$\text{Arg}_X \min f = \left\{ \left( -\frac{2}{\sqrt{5}}, -\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}} \right) \right\} \quad \text{et} \quad \text{Arg}_X \max f = \left\{ \left( \frac{2}{\sqrt{5}}, \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right) \right\}. \quad \diamond$$

## 6.4- Cas des fonctions convexes

**Théorème 6.8:** Si  $X$  est convexe, si  $f$  et les  $\varphi_i$  sont convexes et s'il existe des nombres  $\lambda_i$  tels que les relations de Kuhn-Tucker soient satisfaites en  $\bar{x}$ , alors  $f$  admet sur  $X$  un minimum en  $\bar{x}$ .

*Preuve :*

Pour tout  $x \in X$ ,  $\varphi_i(x) \leq 0$  et  $\lambda_i \geq 0$  pour  $i \in I(\bar{x})$  donc  $\lambda_i \varphi_i(x) \leq 0$  et

$$f(\bar{x}) \leq f(x) - \sum_{i \in I(\bar{x})} \lambda_i \varphi_i(x) = f(\bar{x}) - \sum_{i \in I(\bar{x})} \lambda_i (\varphi_i(x) - \varphi_i(\bar{x}))$$

car  $\varphi_i(\bar{x}) = 0$  pour  $i \in I(\bar{x})$ . Comme les  $\varphi_i$  sont convexes,  $\varphi_i(x) - \varphi_i(\bar{x}) \geq \langle \nabla \varphi_i(\bar{x}), x - \bar{x} \rangle$  donc

$$f(\bar{x}) \leq f(x) - \left\langle \sum_{i \in I(\bar{x})} \lambda_i \nabla \varphi_i(\bar{x}), x - \bar{x} \right\rangle = f(\bar{x}) + \langle \nabla f(\bar{x}), x - \bar{x} \rangle$$

d'après le Théorème 6.5 (Kuhn-Tucker). Mais  $f$  étant convexe,  $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq f(x) - f(\bar{x})$ . D'où  $f(\bar{x}) \leq f(x)$  si  $x \in X$  et  $\bar{x}$  est bien minimum de  $f$  sur  $X$ . □



## Chapitre 7

### Points selles et fonctions de Lagrange

Le but est d'étudier des conditions suffisantes d'optimalité pour des problèmes du type  $\mathcal{P}$  : minimiser  $f$  sur  $S \cap \{x ; \varphi_i(x) \leq 0 \text{ pour } i \in I\}$ , où  $S \subset \mathbb{R}^n$ .

Si  $S = \mathbb{R}^n$ , le problème est le même que précédemment, mais ce qui suit s'applique dans le cas général. Par exemple, si  $S$  est l'ensemble des points à coordonnées entières, le problème devient un problème de programmation en nombres entiers.

#### 7.1- Fonction de Lagrange

**Définition 7.1:** La fonction de Lagrange associée à  $\mathcal{P}$  est définie par :

$$L(x, \lambda) = f(x) + \sum_{i \in I} \lambda_i \varphi_i(x).$$

Si  $\bar{x} \in S$  et  $\bar{\lambda} \geq 0$ , on dit que  $(\bar{x}, \bar{\lambda})$  est un point-selle (ou col) de  $L$  si, pour tout  $x \in S$  et pour tout  $\lambda \geq 0$  :

$$L(\bar{x}, \lambda) \leq L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda})$$

On peut traduire la dernière double inégalité en disant que  $\bar{x}$  minimise  $x \mapsto L(x, \bar{\lambda})$  et  $\bar{\lambda}$  maximise  $\lambda \mapsto L(\bar{x}, \lambda)$ .

**Théorème 7.1:** Soit  $\bar{x} \in S$  et  $\bar{\lambda} \geq 0$ . Alors  $(\bar{x}, \bar{\lambda})$  est un point-selle de  $L$  si et seulement si :

- i)  $L(\bar{x}, \bar{\lambda}) = \min_{x \in S} L(x, \bar{\lambda})$ ,
- ii)  $\varphi_i(\bar{x}) \leq 0$  et  $\bar{\lambda}_i \varphi_i(\bar{x}) = 0$  pour tout  $i \in I$ .

*Preuve :*

• Si  $(\bar{x}, \bar{\lambda})$  est un point-selle de  $L$ , alors i) est vérifié et  $L(\bar{x}, \lambda) \leq L(\bar{x}, \bar{\lambda})$ . On a donc  $\sum_{i \in I} (\lambda_i - \bar{\lambda}_i) \varphi_i(\bar{x}) \leq 0$  pour tout  $\lambda \geq 0$ . Si  $\varphi_{i_0}(\bar{x}) > 0$ , on pourra toujours trouver un  $\lambda_{i_0}$  suffisamment grand pour que  $\sum_{i \in I} (\lambda_i - \bar{\lambda}_i) \varphi_i(\bar{x}) > 0$ .

D'autre part, pour  $\lambda = 0$ ,  $\sum_{i \in I} \bar{\lambda}_i \varphi_i(\bar{x}) \geq 0$  mais  $\bar{\lambda}_i \geq 0$  et  $\varphi_i(\bar{x}) \leq 0$  donc  $\sum_{i \in I} \bar{\lambda}_i \varphi_i(\bar{x}) \leq 0$  et  $\sum_{i \in I} \bar{\lambda}_i \varphi_i(\bar{x}) = 0$ . Comme, pour tout  $i \in I$ ,  $\bar{\lambda}_i \varphi_i(\bar{x}) \leq 0$ , on a donc, pour tout  $i \in I$ ,  $\bar{\lambda}_i \varphi_i(\bar{x}) = 0$  et on a bien ii).

• Réciproquement, on a déjà  $L(\bar{x}, \bar{\lambda}) \leq L(x, \bar{\lambda})$  pour tout  $x \in S$ .  
 D'autre part,  $\sum_{i \in I} \bar{\lambda}_i \varphi_i(\bar{x}) = 0$  donc  $L(\bar{x}, \bar{\lambda}) = f(\bar{x})$ .  
 Mais  $L(\bar{x}, \lambda) = f(\bar{x}) + \sum_{i \in I} \lambda_i \varphi_i(\bar{x}) \leq f(\bar{x}) = L(\bar{x}, \bar{\lambda})$ , d'où le résultat.  $\square$

**Théorème 7.2:** Si  $(\bar{x}, \bar{\lambda})$  est un point-selle de  $L$ , alors  $\bar{x}$  est un optimum global de  $\mathcal{P}$ .

*Preuve :*

D'après i)  $f(\bar{x}) + \sum_{i \in I} \bar{\lambda}_i \varphi_i(\bar{x}) \leq f(x) + \sum_{i \in I} \bar{\lambda}_i \varphi_i(x)$ . Mais, d'après ii),  $\sum_{i \in I} \bar{\lambda}_i \varphi_i(\bar{x}) = 0$  et  $\sum_{i \in I} \bar{\lambda}_i \varphi_i(x) \leq 0$ , donc  $f(\bar{x}) \leq f(x)$ .  $\square$

◇ **Exemple de problème sans point-selle :** minimiser  $f$  définie par  $f(x) = -x^2$  sous les contraintes  $2x - 1 \leq 0$  et  $0 \leq x \leq 1$  où l'on pose  $S = [0, 1]$  et  $\varphi(x) = 2x - 1$ .  $L(x, \lambda) = -x^2 + \lambda(2x - 1)$ .

Pour  $\lambda$  fixé,  $x \mapsto L(x, \lambda)$  est concave, donc le minimum est atteint, soit pour  $x = 0$ , soit pour  $x = 1$  et les seuls points-selles possibles sont  $(0, \bar{\lambda})$  ou  $(1, \bar{\lambda})$ . Mais le problème est en fait de minimiser  $f$  sur  $[0, \frac{1}{2}]$  et le minimum est atteint pour  $\bar{x} = \frac{1}{2}$ , alors qu'il n'existe pas de points-selles de la forme  $(\frac{1}{2}, \bar{\lambda})$ . ◇

## 7.2- Cas des fonctions convexes

**Théorème 7.3 :** Si  $f$  et les  $\varphi_i$  sont convexes, si  $S = \mathbb{R}^n$  et si  $\bar{x}$  est solution de  $\mathcal{P}$ , alors, si les contraintes sont qualifiées en  $\bar{x}$ , il existe  $\lambda \in (\mathbb{R}_+)^n$  tel que  $(\bar{x}, \lambda)$  soit point-selle de  $L$ .

*Preuve :*

D'après le Théorème de Kuhn-Tucker, il existe  $\lambda(\bar{x}) \in (\mathbb{R}_+)^n$  tel que :

$$\nabla f(\bar{x}) + \sum_{i \in I} \lambda_i(\bar{x}) \nabla \varphi_i(\bar{x}) = 0$$

avec  $\varphi_i(\bar{x}) \leq 0$  et  $\lambda_i(\bar{x}) \varphi_i(\bar{x}) = 0$ , soit ii) du Théorème 6.1.

$g : x \mapsto L(x, \lambda(\bar{x})) = f(x) + \sum_{i \in I} \lambda_i(\bar{x}) \varphi_i(x)$  est convexe comme somme de fonctions convexes (et parce que  $\lambda_i(\bar{x}) \geq 0$ ) donc  $g$  admet un minimum en  $\bar{x}$  si et seulement  $\nabla g(\bar{x}) = 0$ . Mais  $\nabla g(\bar{x}) = \nabla f(\bar{x}) + \sum_{i \in I} \lambda_i(\bar{x}) \nabla \varphi_i(\bar{x})$  qui vaut bien 0.  $\square$

## Chapitre 8

# Algorithmes pour l'optimisation sous contraintes

On présente ici des méthodes de résolution pratique du problème qu'est la minimisation d'une fonctionnelle  $J$  sur un ensemble fermé non vide  $C$

$$(\mathcal{P}) \begin{cases} \min J(x), \\ x \in C. \end{cases} \quad (1)$$

### 8.1- Méthodes de projection et de pénalisation

On commence par présenter deux méthodes dont le principe est de se ramener à la minimisation standard de  $J$ , mais en prenant compte de la contrainte. On verra par la suite une autre approche.

#### Méthodes de projection

On a vu dans le Chapitre 5 des méthodes pour construire une suite  $x_n$  avec

$$x_{k+1} = x_k + \rho_k d_k,$$

où on se donne un point de départ  $x_0$  et une méthode pour choisir la direction de descente  $d_k$  ainsi que le pas de descente  $\rho_k$ .

Dans les bonnes conditions, cette suite peut converger plus ou moins rapidement vers la solution optimale  $x^*$  du problème sans contraintes.

Par contre, si on veut minimiser  $J$  sur l'ensemble  $C$ , il n'y a aucune garantie que la suite  $x_n$  reste dans  $C$  même en prenant  $x_0 \in C$ .

On construit alors naturellement l'algorithme du gradient projeté en intercalant une étape de projection sur  $C$  entre chaque calcul du terme suivant dans la suite  $x_k$  :

$$\begin{aligned} \tilde{x} &= x_k + \rho_k d_k, \\ x_{k+1} &= \Pi_C(\tilde{x}), \end{aligned}$$

où  $\Pi_C : \mathbb{R}^n \rightarrow C$  est l'opérateur de projection. Cette idée très simple peut néanmoins être très difficile à mettre en œuvre du fait justement qu'il faut savoir projeter sur  $C$ .

Cependant, il existe des cas particuliers pour lesquels on peut expliciter l'opérateur de projection.

Par exemple pour des contraintes de type  $x_i \geq a_i$  données sur les composantes du vecteur  $x$ , on voit que la  $i$ -ième composante de  $\Pi_C(x)$  sera  $x_i$  si  $x_i \geq a_i$  et  $a_i$  sinon.

On peut résumer ceci par  $(\Pi_C(x))_i = \max(x_i, a_i)$ .

Autre exemple : si l'ensemble des contraintes  $C = \overline{B(a, r)}$  est une boule fermée, centrée au point  $a$  et de rayon  $r$ , la projection  $\Pi_C$  s'écrit alors

$$\Pi_C(x) = \begin{cases} x & \text{si } x \in C, \\ a + r \frac{x-a}{\|x-a\|} & \text{si } x \notin C. \end{cases}$$

Il existe bien entendu des variantes “projetées” de toutes les méthodes de minimisation sans contrainte.

## Méthodes de pénalisation

Le principe des méthodes de pénalisation est de remplacer le problème de minimisation avec contraintes par un problème sans contraintes.

$$\boxed{(\mathcal{P}_\varepsilon) \begin{cases} \min J(x) + \frac{1}{\varepsilon} R(x), \\ x \in \mathbb{R}^n, \end{cases}} \quad (2)$$

où  $R : \mathbb{R}^n \rightarrow \mathbb{R}$  est une fonction de pénalisation et  $\varepsilon > 0$ .

Le but est de trouver des fonctions  $R$  telles que les problèmes  $(\mathcal{P})$  et  $(\mathcal{P}_\varepsilon)$  soient équivalents (qu'ils aient les mêmes solutions).

On dit dans ce cas que la pénalisation est exacte. C'est par exemple le cas avec

$$R(x) = \begin{cases} 0 & \text{si } x \in C, \\ +\infty & \text{si } x \notin C. \end{cases}$$

Cependant, cette fonction n'a pas de bonnes propriétés mathématiques (continuité, dérivabilité,...) nécessaires pour pouvoir utiliser les méthodes d'optimisation sans contrainte classiques.

On utilise alors en général une pénalisation inexacte et on espère que lorsque  $\varepsilon$  devient petit la solution optimale vérifie les contraintes de manière suffisamment satisfaisante.

Néanmoins, on voit déjà que cette méthode générale demandera de faire un compromis entre minimiser  $J$  et satisfaire les contraintes.

Pour être une pénalisation valable, on demande à  $R$  de vérifier au moins les propriétés suivantes :

- $R$  continue sur  $\mathbb{R}^n$ .
- $R(x) \geq 0$  sur  $\mathbb{R}^n$ .
- $R(x) = 0 \iff x \in C$ .

Voici quelques exemples de pénalisations pour les contraintes classiques :

- Contrainte  $x \leq 0$  :  $R(x) = \|x^+\|^2$ .
- Contrainte  $h(x) = 0$  :  $R(x) = \|h(x)\|^2$ .
- Contrainte  $g(x) \leq 0$  :  $R(x) = \|g(x)^+\|^2$ .

L'alternance des étapes “trouver une approximation de la solution optimale sur  $\mathbb{R}^n$  de  $J(x) + \frac{1}{\varepsilon}R(x)$ ” et “réduire la valeur de  $\varepsilon$ ” donne alors ce qu'on appelle l'algorithme de pénalisation extérieure (car on approche la solution optimale sous contrainte par des points *a priori* à l'extérieur de la contrainte).

## 8.2- Méthodes duales

Un autre grande classe de méthodes pour l'optimisation sous contraintes est celle des méthodes dites duales.

Le principe de ces méthodes est de gérer la difficulté apportée par les contraintes en cherchant les points critiques du Lagrangien.

On se ramène ainsi à de l'optimisation vraiment sans contraintes, mais pour une nouvelle fonctionnelle.

### La méthode d'Uzawa

On a vu précédemment que les points-selle  $(\bar{x}, \bar{\lambda}, \bar{\mu})$  du Lagrangien permettent de localiser les extrema sous contraintes de  $J$ .

La méthode d'Uzawa est la recherche de ces points-selle :

1. Pour des multiplicateurs  $(\bar{\lambda}, \bar{\mu})$  fixés, on cherche le minimum (sans contrainte) de  $x \mapsto L(x, \bar{\lambda}, \bar{\mu})$ .
2. Pour un point  $\bar{x}$  fixé, on cherche le maximum de  $(\lambda, \mu) \mapsto L(\bar{x}, \lambda, \mu)$ .

L'algorithme est alors le suivant :

- Choisir  $x_0, \lambda_0, \mu_0$ .
- Calculer  $x_{k+1}$  en cherchant un minimiseur de  $L(x, \lambda_k, \mu_k)$  (optimisation sans contrainte classique).
- Calculer  $\mu_{k+1}$  avec  $\mu_{k+1} = \mu_k + \rho\psi_j(x_{k+1})$ .
- Calculer  $\lambda_{k+1}$  avec  $\lambda_{k+1} = \lambda_k + \rho\varphi_j(x_{k+1})$ .

## La méthode SQP

Le principe de la méthode SQP (Sequential Quadratic Programming) consiste à simplifier les contraintes et le Lagrangien par une approximation de Taylor pour trouver une direction de descente  $d$ . On aura ensuite comme avant

$$x_{k+1} = x_k + \rho_k d_k.$$

Pour ce faire, les contraintes sont approximées par

$$\begin{aligned}\varphi_j(x_k + d) &\approx \varphi(x_k) + D\varphi(x_k) \cdot d, \\ \psi_j(x_k + d) &\approx \psi_j(x_k) + D\psi(x_k) \cdot d,\end{aligned}$$

et le Lagrangien par

$$L(x_k + d, \lambda, \mu) \approx L(x_k, \lambda, \mu) + \partial_x L(x_k, \lambda, \mu) \cdot d + \frac{1}{2} \partial_{xx} L(x_k, \lambda, \mu) \cdot (d, d).$$

Comme  $L(x_k, \lambda, \mu)$  est fixé, et que

$$\partial_x L(x_k, \lambda, \mu) \cdot d = DJ(x_k) \cdot d + \langle D\varphi(x_k) \cdot d, \lambda \rangle + \langle D\psi(x_k) \cdot d, \mu \rangle,$$

et que  $\langle D\varphi(x_k) \cdot d, \lambda \rangle$  et  $\langle D\psi(x_k) \cdot d, \mu \rangle$  sont fixés aussi, on cherche donc une solution  $d$  optimale au système

$$(QP) \begin{cases} \min DJ(x_k) \cdot d + \frac{1}{2} \partial_{xx} L(x_k, \lambda, \mu) \cdot (d, d), \\ \varphi(x_k) + D\varphi(x_k) \cdot d \leq 0, \\ \psi(x_k) + D\psi(x_k) \cdot d = 0. \end{cases}$$

La boucle de l'algorithme est alors la suivante :

- Trouver  $d_k$  en résolvant  $(QP)$ .
- Mettre à jour  $x_{k+1} = x_k + \rho_k d_k$ .
- Mettre à jour  $\lambda$  et  $\mu$  pour qu'ils soient les multiplicateurs associés à  $x_{k+1}$ .

Comme pour les autres méthodes, la question de la convergence n'a pas été abordée mais elle existe ... sous certaines conditions.

---



---

## Rappels de topologie

### 1- Normes, parties bornées, boules

Soit  $E$  un  $\mathbb{R}$ -espace vectoriel.

#### 1.1 Normes

**Définition A1.1:** Une norme sur  $E$  est une application  $\| \cdot \| : E \rightarrow \mathbb{R}, x \mapsto \|x\|$  (ou  $x \mapsto N(x)$ ) vérifiant :

- i)  $\|x\| \geq 0$  pour tout  $x \in E$  et  $\|x\| = 0$  équivaut à  $x = 0$
  - ii) pour tout  $(\lambda, x) \in \mathbb{K} \times E$ ,  $\|\lambda x\| = |\lambda| \|x\|$
  - iii) pour tout  $(x, y) \in E^2$ ,  $\|x + y\| \leq \|x\| + \|y\|$ .
- $E$  muni d'une norme  $\| \cdot \|$  s'appelle un espace vectoriel normé.

◇ **Exemple A1.1:** Supposons  $E$  de dimension  $n \geq 1$ , muni d'une base  $\mathcal{B} = (e_1, \dots, e_n)$ . En notant  $x = \sum_{i=1}^n x_i e_i$ , on a trois normes sur  $E$  en posant :

$$\|x\|_1 = \sum_{i=1}^n |x_i| \quad ; \quad \|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} \quad ; \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \quad \diamond$$

#### 1.2 Parties bornées

**Définition A1.2:** Soit  $E$  un espace vectoriel normé par  $\| \cdot \|$ . Une partie  $P$  de  $E$  est bornée (pour  $\| \cdot \|$ ) s'il existe  $M \geq 0$  tel que, pour tout  $x \in P$ ,  $\|x\| \leq M$  (il existe alors  $\sup_{x \in P} \|x\|$ ).

#### 1.3 Boules

**Définition A1.3:** Dans  $E$  normé par  $\| \cdot \|$ , pour  $a \in E$ , on définit :

- a) si  $r > 0$ ,  $B(a, r) = \{x \in E ; \|x - a\| < r\}$  : boule ouverte de centre  $a$  et de rayon  $r$
- b) si  $r \geq 0$ ,  $B'(a, r) = \{x \in E ; \|x - a\| \leq r\}$  : boule fermée de centre  $a$  et de rayon  $r$
- c) si  $r \geq 0$ ,  $S(a, r) = \{x \in E ; \|x - a\| = r\}$  : sphère de centre  $a$  et de rayon  $r$ .

Ces définitions généralisent les situations classiques sur  $\mathbb{R}$  (intervalles avec  $| \cdot |$ ), sur  $\mathbb{C}$  (disques et cercles avec  $| \cdot |$ ) et en géométrie sur  $\mathbb{R}^2$  ou  $\mathbb{R}^3$  avec  $\| \cdot \|_2$  (cercles et sphères notamment).

## 1.4 Normes équivalentes

**Définition A1.4:**  $N$  et  $N'$ , normes sur  $E$ , sont équivalentes s'il existe  $\alpha > 0$  et  $\beta > 0$  tels que, pour tout  $x \in E$ ,  $\alpha N(x) \leq N'(x) \leq \beta N(x)$ .

Si  $N$  et  $N'$  sont équivalentes, et si  $a \in E$ , toute boule ouverte (resp. fermée) pour  $N$ , de centre  $a$  en contient une pour  $N'$ , et vice-versa.

**Théorème A1.1:** Toutes les normes sur un espace vectoriel de dimension FINIE sont deux à deux équivalentes.

## 2- Suites dans $\mathbb{R}^n$

**Définition A1.5:** Si  $\| \cdot \|$  est une norme sur  $\mathbb{R}^n$ , une suite  $(x^{(k)}) \in (\mathbb{R}^n)^{\mathbb{N}}$  est bornée pour  $\| \cdot \|$  s'il existe  $M \geq 0$  tel que  $\|x^{(k)}\| \leq M$  pour tout  $k$  (et alors  $\sup \|x^{(k)}\|$  existe). Une suite  $(x^{(k)})$  converge vers  $l \in \mathbb{R}^n$  (pour  $\| \cdot \|$ ) si  $\lim_k \|x^{(k)} - l\| = 0$  : on note  $l = \lim_k x^{(k)}$ . Sinon,  $(x^{(k)})$  diverge (pour  $\| \cdot \|$ ).

On peut montrer que, si une suite converge, sa limite est unique.

**Remarque :** Comme  $\mathbb{R}^n$  est de dimension finie, le choix de la norme n'a aucune importance. Une suite bornée (resp. convergente) pour une norme l'est aussi pour n'importe quelle autre norme.

**Définition A1.6:** On dit que  $x$  est un point d'accumulation de la suite  $(x^{(k)})$  si on peut extraire de cette suite une sous suite  $(x^{(k')})$  qui converge vers  $x$ .

Une suite convergente a donc la propriété de posséder un point d'accumulation unique, la limite de la suite. Mais, évidemment, une suite ayant un point d'accumulation unique n'est pas nécessairement convergente.

## 3- Sous ensembles ouverts et sous-ensembles fermés de $\mathbb{R}^n$

### 3.1 Intérieur d'un sous-ensemble

**Définition A1.7:** Si  $S$  est un sous-ensemble de  $\mathbb{R}^n$ , on dit que  $y \in S$  est un point intérieur de  $S$  s'il existe une boule centrée en  $y$  et contenue dans  $S$ . L'ensemble des points intérieurs à  $S$  est appelé l'intérieur de  $S$  et est noté  $\text{Int}(S)$ .

◇ **Exemple A1.2:** La boule unité fermée  $B'(0; 1)$  a pour intérieur la boule unité ouverte  $B(0; 1)$ . ◇



**Remarque :** L'intérieur d'un ensemble peut être vide ; c'est le cas par exemple d'un hyperplan dans  $\mathbb{R}^n$ .

### 3.2 Sous-ensembles ouverts

**Définition A1.8:**  $\mathcal{O} \subset \mathbb{R}^n$  est un ouvert si, pour tout  $x \in \mathbb{R}^n$ , il existe une boule ouverte de centre  $x$  incluse dans  $\mathcal{O}$ .

Ainsi, un ensemble est ouvert s'il coïncide avec son intérieur. L'ensemble vide  $\emptyset$  et l'ensemble  $\mathbb{R}^n$  tout entier sont des ouverts.

◇ **Exemple A1.3 :** Une boule ouverte de  $\mathbb{R}^n$  est un ouvert.

*Preuve :* On se donne donc  $N$  une norme sur  $\mathbb{R}^n$ , et  $B(a, r)$ .

Soit  $x$  dans cette boule, et  $\varepsilon = r - N(x - a)$ . Si  $y \in B(x, \varepsilon)$ ,

$$N(y - a) \leq N(y - x) + N(x - a) < \varepsilon + N(x - a) = r.$$

Donc,  $B(x, \varepsilon) \subset B(a, r)$ . ◇

On rappelle qu'il n'est pas utile de préciser la norme utilisée. En effet, si cette boule  $B(x, r)$  est issue de la norme  $N$ , et si  $N'$  est une autre norme,  $B_N(x, r)$  contient, par équivalence des normes sur  $\mathbb{R}^n$ , une boule  $B_{N'}(x, r')$  qui est à son tour contenue dans  $\mathcal{O}$ . Dans la suite, ce sera une attitude constante : si l'expression "boule" est utilisée sans référence à la norme, c'est parce que toute norme peut-être utilisée.

**Théorème A1.2:** a) Une réunion d'ouverts de  $\mathbb{R}^n$  est un ouvert de  $\mathbb{R}^n$ .  
b) Une intersection FINIE d'ouverts de  $\mathbb{R}^n$  est un ouvert de  $\mathbb{R}^n$ .

*Preuve :* a) Soit  $x \in \bigcup_{i \in I} \mathcal{O}_i$ , donc  $x \in \mathcal{O}_{i_0}$ , et il existe  $B(x, r_{i_0}) \subset \mathcal{O}_{i_0} \subset \bigcup_{i \in I} \mathcal{O}_i$ .

b) Soit  $x \in \bigcap_{i=1}^n \mathcal{O}_i$ . Pour tout  $i$ , il existe  $B(x, r_i) \subset \mathcal{O}_i$ . Soit  $\rho = \min r_i$ .  $\rho > 0$ , car la famille est finie, et donc  $\rho = r_{i_0}$ . Puis,  $B(x, \rho) \subset \bigcap_{i=1}^n \mathcal{O}_i$ .

□

**Remarque :**  $\bigcap_{n \in \mathbb{N}^*} B(a, \frac{1}{n}) = \{a\}$  car  $N(x - a) < \frac{1}{n}$  pour tout  $n$  donne, avec  $n \rightarrow +\infty$ ,  $x = a$ .

### 3.3 Voisinage

**Définition A1.9:** Étant donné  $x \in \mathbb{R}^n$ ,  $\mathcal{V}(x)$  est un voisinage de  $x$  s'il contient une boule ouverte de centre  $x$ .

### 3.4 Point adhérent à une partie

**Définition A1.10:** On dit que  $x \in \mathbb{R}^n$  est un point d'adhérence d'un sous-ensemble  $S$  de  $\mathbb{R}^n$  si toute boule ouverte de centre  $x$  rencontre  $S$ .

L'ensemble de tous les points d'adhérence de  $S$  est appelé l'adhérence de  $S$  (ou encore la clôture de  $S$ ). Il est noté  $\text{cl}(S)$ .

**Remarques :** 1) On a évidemment  $S \subset \text{cl}(S)$ .

2) On peut remplacer “toute boule ouverte” par “toute boule fermée de rayon  $> 0$ ” : si  $B(x, r) \cap S \neq \emptyset$ , alors  $B'(x, r) \cap S \neq \emptyset$  et, si  $B'(x, \frac{r}{2}) \cap S \neq \emptyset$ , alors  $B(x, r) \cap S \neq \emptyset$ .

◇ **Exemple A1.4:** L'adhérence de la boule unité  $B(0; 1)$  est l'ensemble  $\text{cl}(B(0; 1)) = B'(0; 1)$  (boule unité fermée). ◇

**Théorème A1.3:** Si  $a \in \mathbb{R}^n$  est adhérent à  $S \subset \mathbb{R}^n$ , il existe une suite  $(x^{(k)}) \in S^{\mathbb{N}}$  telle que  $\lim_k x^{(k)} = a$  (même si  $n = 1$  et  $a = +\infty$  ou  $a = -\infty$ ).

*Preuve :* Fixons une norme  $N$ . Si  $a$  est adhérent à  $A$ ,  $B(a, \frac{1}{k}) \cap A \neq \emptyset$ , donc on peut prendre  $x^{(k)}$  dans cet ensemble. Comme  $N(x^{(k)} - a) < \frac{1}{k}$ , on a bien  $\lim_k x^{(k)} = a$ .

Si  $S \subset \mathbb{R}$ , et si  $+\infty$  est adhérent à  $S$ , on prend  $x^{(k)} \in ]k, +\infty[ \cap S$ .

□

### 3.5 Fermés de $\mathbb{R}^n$

**Théorème A1.4 :** Le complémentaire (dans  $\mathbb{R}^n$ ) d'un ouvert est un fermé ; le complémentaire (dans  $\mathbb{R}^n$ ) d'un fermé est un ouvert.

*Preuve :*

• Soit  $S \subset \mathbb{R}^n$  un ouvert de  $S$  et  $\bar{S}$  son complémentaire dans  $\mathbb{R}^n$ .

Soit  $x \in S$ . Il existe une boule de centre  $x$  et contenue entièrement dans  $S$ . Cette boule ne rencontre pas  $\bar{S}$  donc  $x$  n'est pas un point d'adhérence de  $\bar{S}$ . Par suite, tous les points d'adhérence de  $\bar{S}$  sont dans  $\bar{S}$ , et  $\bar{S}$  est fermé.

• Soit maintenant  $T \subset \mathbb{R}^n$  un fermé,  $\bar{T}$  son complémentaire dans  $\mathbb{R}^n$ . Soit  $x \in \bar{T}$  quelconque ;  $x$  n'est pas un point d'adhérence de  $T$  (car  $T$  contient tous ses points d'adhérence). Donc il existe une boule de centre  $x$  et qui ne rencontre pas  $T$ .

Cette boule est entièrement contenue dans  $\bar{T}$  et, par suite  $\bar{T}$  est ouvert.

□

La propriété suivante se démontre alors à partir du théorème A1.2 par passage aux complémentaires :

**Théorème A1.5:** a) Une intersection de fermés de  $\mathbb{R}^n$  est un fermé de  $\mathbb{R}^n$ .  
b) Une réunion FINIE de fermés de  $\mathbb{R}^n$  est un fermé de  $\mathbb{R}^n$ .

◇**Exemple A1.5:** Une boule fermée de  $\mathbb{R}^n$  est un fermé.

*Preuve :* On se donne donc  $N$  une norme sur  $\mathbb{R}^n$ , et  $B'(a, r)$ . Soit  $x$  dans  $E \setminus B'(a, r)$ , et  $\varepsilon = N(x - a) - r$ . Si  $y \in B(x, \varepsilon)$ ,  $N(y - a) \geq -N(x - y) + N(x - a) > -\varepsilon + N(x - a) = r$ . Donc,  $B(x, \varepsilon) \subset \mathbb{R}^n \setminus B'(a, r)$ . ◇

**Remarque :**  $\bigcup_{k \in \mathbb{N}^*} B'(a, 1 - \frac{1}{k}) = B(a, 1)$ .

Une propriété fondamentale des sous-ensembles fermés est la suivante :

**Théorème A1.6:**  $S \subset \mathbb{R}^n$  est fermé si et seulement si toute suite convergente d'éléments de  $S$  a une limite dans  $S$ .

*Preuve :* Si  $x$  est la limite d'une suite  $(x^{(k)})$  d'éléments de  $S$ , alors, pour tout  $\varepsilon > 0$ , il existe des éléments de la suite dans la boule centrée en  $x$  et de rayon  $\varepsilon$ . Donc  $x$  est un point d'adhérence de  $S$ .

Par suite, si  $S$  est fermé, toute suite convergente d'éléments de  $S$  a une limite dans  $S$ .

Inversement, si  $S$  n'est pas fermé,  $S$  admet un point d'adhérence  $x \notin S$ . Comme  $x$  est un point d'adhérence, pour tout  $k$  entier  $\geq 1$ , il existe un point  $x^{(k)} \in S$  tel que  $N(x - x^{(k)}) < \frac{1}{k}$ .

On met ainsi en évidence une suite  $(x^{(k)})$  dont la limite  $x$  n'est pas dans  $S$ .

□

Autre propriété utile :

**Théorème A1.7:** L'image réciproque d'un fermé (resp. d'un ouvert) par une application continue est un fermé (resp. un ouvert).

## 4- Sous-ensembles compacts

**Définition A1.11:** Un sous-ensemble  $K \subset \mathbb{R}^n$  est dit compact si, de toute suite infinie  $(x^{(k)})$  d'éléments de  $K$ , on peut extraire une sous-suite infinie  $(x^{(k')})$  convergeant vers un élément de  $K$ .

La propriété suivante donne une caractérisation simple des ensembles compacts dans  $\mathbb{R}^n$  (et, plus généralement dans les espaces vectoriels normés de dimension finie) ;

**Théorème A1.8:** Dans  $\mathbb{R}^n$ , un sous-ensemble  $K$  est un compact si et seulement s'il est fermé et borné.

Cette propriété cesse d'être vraie en dimension infinie et c'est là l'origine de certaines difficultés pour l'optimisation en dimension infinie.

Les résultats suivants sont fondamentaux et concernent l'existence d'une solution optimale pour un problème d'optimisation.

**Théorème A1.9:** Soit  $f : K \rightarrow F$  continue, avec  $K$  compact de  $\mathbb{R}^n$ . Alors,  $f(K)$  est un compact de  $F$ .

**Théorème A1.10:**  $f : K \rightarrow F$  continue, avec  $K$  compact de  $\mathbb{R}^n$ . Alors,  $f$  est bornée et atteint ses bornes, c'est-à-dire qu'il existe  $\alpha$  et  $\beta$  dans  $K$  avec  $f(\alpha) = \min_K f$  et  $f(\beta) = \max_K f$ .

---

---

## Rappels de différentiabilité

Dans tout ce cours,  $f$  est une fonction définie sur un ouvert  $U$  d'un espace vectoriel normé  $E$  vers un espace vectoriel normé  $F$ .

L'objectif est de linéariser  $f$  au voisinage de  $a \in U$ , c'est-à-dire d'approcher au voisinage de  $a$  la différence  $f(a+h) - f(a)$  par une expression linéaire continue en la variable  $h$ .

### 1- Fonctions différentiables

#### 1.1 Définition de la différentiabilité

**Définition A2.1:** On dit qu'une fonction  $f : U \rightarrow F$  est différentiable en  $a \in U$  s'il existe une application linéaire continue  $L : E \rightarrow F$  telle que :

$$f(a+h) - f(a) = L(h) + o(\|h\|) \quad (\text{si } a+h \in U).$$

On dit que  $f$  est différentiable sur  $U$  si elle est différentiable en tout point de  $U$ .

La relation  $f(a+h) - f(a) = L(h) + o(\|h\|)$  peut aussi s'écrire comme suit :

- $f(a+h) - f(a) = L(h) + \|h\|\varepsilon(h)$  où  $\lim_{h \rightarrow 0_E} \varepsilon(h) = 0_F$ .
- Pour tout  $\varepsilon > 0$ , il existe  $\alpha > 0$ , tel que, pour tout  $h \in E$ ,

$$\text{si } \|h\| \leq \alpha, \text{ alors } \|f(a+h) - f(a) - L(h)\| \leq \varepsilon \|h\|.$$

- $\lim_{\substack{h \rightarrow 0_E \\ h \neq 0}} \frac{f(a+h) - f(a) - L(h)}{\|h\|} = 0_F$ .

**Remarque importante :** Dans tout ce cours, on prendra  $E = \mathbb{R}^n$  et  $F = \mathbb{R}^p$  où  $n$  et  $p$  sont dans  $\mathbb{N}^*$ .

On notera  $(e_1, \dots, e_n)$  la base canonique de  $\mathbb{R}^n$  ( $e_1 = (1, 0, \dots, 0)$ ,  $e_2 = (0, 1, 0, \dots, 0)$ ...)  $\mathbb{R}^n$  est un espace euclidien doté du produit scalaire  $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$  si  $x = (x_1, \dots, x_n)$  et  $y = (y_1, \dots, y_n)$ .

On rappelle que, lorsque  $E$  est de dimension finie :

- toutes les normes sur  $E$  sont équivalentes ;
- une application linéaire  $L : E \rightarrow F$  est continue et  $\|L\| = \sup_{h \neq 0_E} \frac{\|L(h)\|}{\|h\|}$  ;
- si  $E$  est euclidien, pour toute forme linéaire  $L : E \rightarrow \mathbb{R}$ , il existe un unique élément  $v \in E$  tel que  $L(h) = \langle v, h \rangle$  pour tout  $h \in E$  (*théorème de Riez*).

**Proposition A2.1** (*Unicité de la différentielle en  $a$* )

Si  $f : U \rightarrow \mathbb{R}^p$  est différentiable en  $a \in U$ , alors il existe une et une seule application linéaire  $L : \mathbb{R}^n \rightarrow \mathbb{R}^p$  telle que :

$$f(a+h) - f(a) = L(h) + o(\|h\|) \quad (\text{si } a+h \in U).$$

*Preuve :* L'existence de  $L$  est assurée par l'hypothèse de différentiabilité de  $f$  en  $a$ . Supposons alors qu'il existe deux applications linéaires continues  $L_1$  et  $L_2$  telles que :

$$f(a+h) - f(a) = L_1(h) + o(\|h\|) = L_2(h) + o(\|h\|).$$

Par différence, on en déduit alors  $\lim_{h \rightarrow 0} \frac{(L_1 - L_2)(h)}{\|h\|} = 0$ . Donc, si  $x \in \mathbb{R}^n \setminus \{0\}$  et  $h = tx$ , où  $t \in \mathbb{R}_+$ , on a  $\lim_{t \rightarrow 0} \frac{t(L_1 - L_2)(x)}{t\|x\|} = 0 = \frac{L_1(x) - L_2(x)}{\|x\|}$ , donc  $L_1(x) = L_2(x)$ .

□

L'unicité de cette application linéaire  $L$  légitime alors la définition suivante :

**Définition A2.2:** Si  $f : U \rightarrow F$  est différentiable en  $a \in U$ , on appelle différentielle de  $f$  en  $a$  l'unique application linéaire  $df_a \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$  telle que :

$$f(a+h) = f(a) + df_a(h) + o(\|h\|).$$

Comme précédemment, on pourra écrire  $f(a+h) = f(a) + df_a(h) + \|h\|\varepsilon(h)$  avec  $\lim_{\|h\| \rightarrow 0} \varepsilon(h) = 0$ .

**Remarque : Différentiabilité des fonctions d'une variable réelle**

Lorsque  $E = \mathbb{R}$  ( $n = 1$ ), on a, par linéarité, pour tout  $h = h \times 1 \in \mathbb{R}$ ,  $df_a(h) = h df_a(1)$  et on en déduit :

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = \lim_{h \rightarrow 0} \frac{h df_a(1) + o(h)}{h} = df_a(1).$$

Dans ce cas, on a donc  $df_a(h) = h df_a(1) = h f'(a)$  et on retrouve bien la relation classique :

$$f(a+h) = f(a) + h f'(a) + o(h).$$

**Proposition A2.2** (*Une application différentiable est continue*)

Si  $f : U \rightarrow \mathbb{R}^p$  est différentiable en  $a \in U$ , alors elle est continue en  $a$ .

*Preuve :* On a, par définition de la différentiabilité en  $a$  :

$$f(a+h) = f(a) + df_a(h) + o(\|h\|).$$

Comme  $df_a$  est continue (car linéaire), on a  $\lim_{\|h\| \rightarrow 0} df_a(h) = 0$  et donc  $\lim_{\|h\| \rightarrow 0} f(a+h) = f(a)$ .

□

**Attention :** la réciproque est fausse !

◇ **Exemple A2.1:** Montrer qu'une norme sur  $\mathbb{R}^n$  est continue et non différentiable en 0. On sait que  $\| \cdot \|$  est continue car  $\|a+h\| - \|a\| \leq \|h\|$ . Si  $\| \cdot \|$  était différentiable en 0, il existerait  $L$  telle que  $\|h\| = L(h) + o(\|h\|)$ . L'égalité  $\| -th \| = \|th\|$  donne, pour tout  $t \in \mathbb{R}_+$ , avec  $h$  fixé :

$$L(th) + o(\|th\|) = L(-th) + o(\|th\|).$$

On en tire  $2tL(h) = o(\|th\|)$  d'où  $2tL(h) = t\|h\|\varepsilon(th)$  et  $L(h) = 0$  en divisant par  $t$  puis en faisant  $t \rightarrow 0$ . Comme  $h$  est arbitraire, on aurait  $L = 0$  et la relation  $\|h\| = o(\|h\|)$ , soit  $1 = \varepsilon(h)$ , ce qui est absurde.  $\diamond$

## Exemples d'applications différentiables

### 1. Une application constante est différentiable

En effet, si  $f = c$ ,  $f(a+h) - f(a) = c - c = 0$  et  $\boxed{df_a = 0}$ .

### 2. Une application linéaire est différentiable

En effet, on a, par linéarité,  $L(a+h) = L(a) + L(h)$  et  $\boxed{dL_a = L}$ .

### 3. Une application bilinéaire est différentiable

On peut normer  $\mathbb{R}^n \times \mathbb{R}^p$  par  $\|(x, y)\| = \max(\|x\|, \|y\|)$ .

Si  $B : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^m$  est bilinéaire, on a alors :

$$B(x+h, y+k) = B(x, y) + (B(x, k) + B(h, y)) + B(h, k).$$

L'application  $(h, k) \in \mathbb{R}^n \times \mathbb{R}^p \mapsto B(x, k) + B(h, y) \in \mathbb{R}^m$  est linéaire.

Le dernier terme est  $o(\|(h, k)\|)$  puisque  $\|B(h, k)\| \leq K\|h\| \|k\| \leq K\|(h, k)\|^2$ .

Ainsi,  $\boxed{dB_{(x,y)}(h, k) = B(x, k) + B(h, y)}$ .

## 1.2 Opérations sur les fonctions différentiables

### **Proposition A2.3** (*combinaison linéaire de fonctions différentiables*)

Soit  $\lambda, \mu \in \mathbb{R}$ , et  $f, g : U \rightarrow \mathbb{R}^p$  différentiables en  $a \in U$  (resp. sur  $U$ ).

Alors  $\lambda f + \mu g$  est différentiable en  $a \in U$  (resp. sur  $U$ ) et on a :

$$d(\lambda f + \mu g)_a = \lambda df_a + \mu dg_a.$$

*Preuve :* Si  $f$  et  $g$  sont différentiables en  $a$ , on a par définition :

$$f(a+h) = f(a) + df_a(h) + o(\|h\|) \quad ; \quad g(a+h) = g(a) + dg_a(h) + o(\|h\|).$$

Par combinaison linéaire, on en déduit

$$(\lambda f + \mu g)(a+h) = (\lambda f + \mu g)(a) + (\lambda df_a + \mu dg_a)(h) + o(\|h\|).$$

De plus, comme  $df_a$  et  $dg_a$  sont linéaires,  $\lambda df_a + \mu dg_a$  l'est aussi.

Il en résulte que  $\lambda f + \mu g$  est donc différentiable en  $a$  et sa différentielle est

$$d(\lambda f + \mu g)_a = \lambda df_a + \mu dg_a.$$

□

**Proposition A2.4** (*produit d'une fonction scalaire et d'une fonction différentiable*)  
 Soit  $\lambda : U \rightarrow \mathbb{R}$  et  $f : U \rightarrow \mathbb{R}^p$  différentiables en  $a \in U$  (resp. sur  $U$ ).  
 Alors  $\lambda f$  est différentiable en  $a \in U$  (resp. sur  $U$ ) et on a :

$$d(\lambda f)_a = f(a)d\lambda_a + \lambda(a)df_a.$$

*Preuve :* Remarquons d'abord que :

$$(\lambda f)(a+h) - (\lambda f)(a) = (\lambda(a+h) - \lambda(a))f(a+h) + \lambda(a)(f(a+h) - f(a)).$$

Si  $\lambda$  et  $f$  sont différentiables en  $a$ , on a par définition :

$$\lambda(a+h) = \lambda(a) + d\lambda_a(h) + o(\|h\|) \quad ; \quad f(a+h) = f(a) + df_a(h) + o(\|h\|).$$

En reportant ci-dessus, on en déduit que

$$\begin{aligned} (\lambda f)(a+h) &= (\lambda f)(a) + [d\lambda_a(h) + o(\|h\|)]f(a+h) + \lambda(a)[(df_a(h) + o(\|h\|))] \\ &= (\lambda f)(a) + d\lambda_a(h)f(a) + \lambda(a)df_a(h) + \\ &\quad d\lambda_a(h)(f(a+h) - f(a)) + o(\|h\|)f(a+h) + \lambda(a)o(\|h\|). \end{aligned}$$

On observe que les trois termes de cette dernière ligne sont négligeables devant  $\|h\|$ .  
 C'est en effet évident pour les deux derniers et c'est vrai pour le premier car :

$$\|d\lambda_a(h)(f(a+h) - f(a))\| \leq \|d\lambda_a\| \|h\| \|f(a+h) - f(a)\| = o(\|h\|).$$

Il en résulte que  $\lambda f$  est donc différentiable en  $a$  et sa différentielle est :

$$d(\lambda f)_a = f(a)d\lambda_a + \lambda(a)df_a.$$

□

**Proposition A2.5** (*composée d'une application linéaire et d'une fonction différentiable*)  
 Soit  $f : U \rightarrow \mathbb{R}^p$  différentiable en  $a \in U$  (resp. sur  $U$ ) et  $L : \mathbb{R}^p \rightarrow \mathbb{R}^m$  linéaire.  
 Alors  $L \circ f$  est différentiable en  $a$  (resp. sur  $U$ ) et sa différentielle en  $a$  est :

$$d(L \circ f)_a = L \circ df_a.$$

*Preuve :* On a, par définition de la différentiabilité en  $a$  (avec  $\lim_{\|h\| \rightarrow 0} \varepsilon(h) = 0$ ) :

$$f(a+h) = f(a) + df_a(h) + \|h\|\varepsilon(h).$$

En composant par  $L$ , on en déduit par linéarité :

$$L \circ f(a+h) = L \circ f(a) + L \circ df_a(h) + \|h\|L \circ \varepsilon(h).$$

Comme  $L$  est continue, on a  $\lim_{\|h\| \rightarrow 0} L \circ \varepsilon(h) = L(\lim_{\|h\| \rightarrow 0} \varepsilon(h)) = L(0_p) = 0_m$ .

Ainsi donc  $\|h\|L \circ \varepsilon(h)$  est négligeable devant  $\|h\|$  et on peut écrire :

$$L \circ f(a+h) = L \circ f(a) + L \circ df_a(h) + o(\|h\|).$$

De plus, comme  $L$  et  $df_a$  sont linéaires continues,  $L \circ df_a$  l'est aussi.



Il en résulte que  $L \circ f$  est différentiable en  $a$  et sa différentielle est :

$$d(L \circ f)_a = L \circ df_a.$$

□

**Proposition A2.6** (*application à une simplification intéressante*).

On pose  $f = (f_1, \dots, f_p)$  où  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ . Il y a équivalence entre :

- la fonction  $f : U \rightarrow \mathbb{R}^p$  est différentiable en  $a \in U$  (resp. sur  $U$ ).
- les composantes  $f_1, \dots, f_p : U \rightarrow \mathbb{R}$  sont différentiables en  $a \in U$  (resp. sur  $U$ ).

*Preuve :*

• Supposons  $f$  différentiable en  $a$ . En notant  $v_i^* : \mathbb{R}^p \rightarrow \mathbb{R}$ ,  $y = (y_1, \dots, y_p) \mapsto y_i$ , on a donc  $f_i = v_i^* \circ f$ .

Comme  $v_i^*$  est linéaire, par la proposition A2.5,  $f_i = v_i^* \circ f$  est différentiable en  $a$ , de différentielle  $v_i^* \circ df_a$ .

- Supposons  $f_1, \dots, f_p$  différentiables en  $a$ , de différentielles  $(df_1)_a, \dots, (df_p)_a$  :

$$\text{pour tout } i \in \{1, \dots, p\}, f_i(a+h) = f_i(a) + (df_i)_a(h) + \|h\|\varepsilon_i(h).$$

Comme  $f = (f_1, \dots, f_p)$ , il vient avec  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_p)$  :

$$f(a+h) = f(a) + ((df_1)_a(h), \dots, (df_p)_a(h)) + \|h\|\varepsilon(h).$$

Ainsi,  $f$  est différentiable en  $a$  et sa différentielle est  $h \mapsto ((df_1)_a(h), \dots, (df_p)_a(h))$ . □

En résumé, il est équivalent d'étudier la différentiabilité de  $f$  ou de ses  $p$  composantes  $f_1, \dots, f_p$ .

Les différentielles des composantes de  $f$  sont les composantes de la différentielle de  $f$ .

**Proposition A2.7** (*Composition des fonctions différentiables*).

Soit une fonction  $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$  différentiable en  $a \in U$  (resp. sur  $U$ ).

Soit une fonction  $g : V \subset \mathbb{R}^p \rightarrow \mathbb{R}^m$  différentiable en  $f(a)$  (resp. sur  $V$ ).

Si  $f(U) \subset V$ , alors  $g \circ f$  est différentiable en  $a$  (resp. sur  $U$ ) et on a :

$$d(g \circ f)_a = dg_{f(a)} \circ df_a.$$

*Preuve :* Si  $f$  est différentiable en  $a$ , on a par définition, avec  $\lim_{\|h\| \rightarrow 0} \varepsilon_f(h) = 0$  :

$$f(a+h) = f(a) + df_a(h) + \|h\|\varepsilon_f(h).$$

Si  $g$  est différentiable en  $f(a)$ , on a par définition, avec  $\lim_{\|k\| \rightarrow 0} \varepsilon_g(k) = 0$  :

$$g(f(a) + k) = g(f(a)) + dg_{f(a)}(k) + \|k\|\varepsilon_g(k).$$

Faisant alors  $k = f(a+h) - f(a) = df_a(h) + \|h\|\varepsilon_f(h)$ , on obtient :

$$g \circ f(a+h) = g \circ f(a) + dg_{f(a)}(df_a(h) + \|h\|\varepsilon_f(h)) + \|df_a(h) + \|h\|\varepsilon_f(h)\|\varepsilon_g(f(a+h) - f(a)).$$

Utilisant la linéarité de  $df_a$  et de  $dg_{f(a)}$ , cette expression s'écrit aussi :

$$g \circ f(a+h) = g \circ f(a) + dg_{f(a)} \circ df_a(h) + \|h\| [dg_{f(a)}(\varepsilon_f(h)) + \|df_a(h/\|h\|) + \varepsilon_f(h)\| \varepsilon_g(f(a+h) - f(a))].$$

Ce dernier terme en facteur de  $\|h\|$  tend vers 0 lorsque  $h$  tend vers 0 car

$$\begin{aligned} &\rightarrow \lim_{h \rightarrow 0} dg_{f(a)}(\varepsilon_f(h)) = dg_{f(a)}(0) = 0. \\ &\rightarrow \|df_a(h/\|h\|) + \varepsilon_f(h)\| \leq \|df_a\| + \|\varepsilon_f(h)\| \leq \|df_a\| + 1 \text{ pour } \|h\| \text{ assez petit.} \\ &\rightarrow \lim_{h \rightarrow 0} \varepsilon_g(f(a+h) - f(a)) = 0. \end{aligned}$$

Il en résulte que  $g \circ f(a+h) = g \circ f(a) + dg_{f(a)} \circ df_a(h) + o(\|h\|)$ .

De plus, comme  $dg_{f(a)}$  et  $df_a$  sont linéaires,  $dg_{f(a)} \circ df_a$  l'est aussi.

Finalement,  $g \circ f$  est différentiable et sa différentielle en  $a$  est  $d(g \circ f)_a = dg_{f(a)} \circ df_a$ .

□

◇ **Exemple A2.2:** différentielle de  $1/\lambda$  où  $\lambda : U \rightarrow \mathbb{R}$  est différentiable en  $a$ .

Si  $\lambda$  ne s'annule pas et si  $\lambda$  est différentiable en  $a$ , alors  $1/\lambda$  est différentiable en  $a$  et :

$$d\left(\frac{1}{\lambda}\right)_a = -\frac{1}{\lambda^2(a)} d\lambda_a.$$

En effet, il suffit d'appliquer le résultat précédent avec les deux applications  $f$  et  $g$  définies par  $f : x \in U \rightarrow \lambda(x) \in \mathbb{R}$  et  $g : x \in \mathbb{R}^* \rightarrow 1/x \in \mathbb{R}$ . ◇

### 1-3 Dérivées suivant un vecteur, applications et dérivées partielles.

Soit  $a \in U$ ,  $h \in \mathbb{R}^n$ , et  $f : U \rightarrow \mathbb{R}^p$ . Comme  $U$  est ouvert, il existe  $r > 0$  tel que  $B(a, r) \subset U$ , et comme  $f$  est définie sur  $U$ , donc sur  $B(a, r)$ , alors  $t \in \mathbb{R} \rightarrow \varphi_h(t) = f(a + th) \in \mathbb{R}^p$  est définie pour  $\|th\| < r$ , donc sur  $] -\delta, \delta[$  où  $\delta = r/\|h\|$  si  $h \neq 0$  et sur  $\mathbb{R}$  si  $h = 0$ , ce qui légitime la définition suivante :

**Définition A2.3:** On dit qu'une fonction  $f : U \rightarrow \mathbb{R}^p$  a une dérivée en  $a$  suivant un vecteur  $h \in \mathbb{R}^n$  si la fonction de la variable réelle  $t \mapsto \varphi_h(t) = f(a + th) \in \mathbb{R}^p$  est dérivable en 0.

On appelle alors dérivée de  $f$  en  $a$  suivant  $h$  le vecteur  $D_h f(a) = \varphi'_h(0)$  :

$$D_h f(a) = \lim_{t \rightarrow 0} \frac{f(a + th) - f(a)}{t}$$

(Si  $h = 0$ ,  $D_0 f(a) = 0$ ).

Si  $h = e_j$ , on note  $\varphi_{e_j}$  par  $\varphi_j$  :  $j$ -ème application partielle de  $f$  en  $a$ , et on note  $D_{e_j} f(a)$  par  $D_j f(a)$  :  $j$ -ème dérivée partielle de  $f$  en  $a$ .

**Autre notation :**  $D_j f(a) = \lim_{t \rightarrow 0} \frac{f(a + te_j) - f(a)}{t}$  est parfois notée  $\frac{\partial f}{\partial x_j}(a)$  mais cette écriture est parfois ambiguë comme dans  $\frac{\partial f}{\partial x_1}(x_2, x_1)$ .

**Remarque :** Si  $f : U \rightarrow \mathbb{R}^p$  est continue en  $a$ , alors, pour tout  $h$ ,  $\varphi_h$  l'est en 0, avec  $\varphi_h(0) = f(a)$ .

En effet,  $\psi_h : t \mapsto a + th$  est continue sur  $] -\delta, \delta[$ , avec  $\psi_h(] -\delta, \delta[) \subset U$  et  $\varphi_h = f \circ \psi_h$  est continue par composition.

◇ **Exemple A2.3:** Continuité en  $(0, 0)$  de  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $\begin{cases} (x, y) \neq (0, 0) \mapsto \frac{xy}{x^2+y^2} \\ (0, 0) \mapsto l \end{cases}$ .

$\varphi_1(t) = f(t, 0) = 0$  si  $t \neq 0$ , donc il faut  $l = 0$  pour que  $f$  soit continue,  $\varphi_1$  et  $\varphi_2$  l'étant alors. Prenons  $h = (\alpha, \beta)$ .  $\varphi_h(t) = f(t\alpha, t\beta) = \frac{\alpha\beta}{\alpha^2+\beta^2}$  si  $t \neq 0$ . Par exemple,  $\varphi_{(1,1)}(t) = \frac{1}{2}$  donc, avec  $l = 0$ ,  $\varphi_{(1,1)}$  n'est pas continue en  $(0, 0)$  et  $f$  non plus. ◇

**Attention :** la réciproque est fausse !

◇ **Exemple A2.4:** Continuité en  $(0, 0)$  de  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $\begin{cases} (x, y) \neq (0, 0) \mapsto \frac{x^2y}{x^4+y^2} \\ (0, 0) \mapsto l \end{cases}$ .

Prenons  $h = (\alpha, \beta)$ .  $\varphi_h(t) = \frac{t\alpha^2\beta}{t^2\alpha^4+\beta^2}$  pour  $t \neq 0$ . Si  $\beta = 0$ ,  $\varphi_h(t) = 0$ . Sinon,  $\lim_{t \rightarrow 0} \varphi_h(t) = 0$ . Il faut que  $l = 0$  pour que  $f$  soit continue, toutes les  $\varphi_h$  l'étant alors. Mais  $f(x, x^2) = \frac{1}{2}$  si  $x \neq 0$ , donc, si  $l = 0$ ,  $f$  n'est pas continue en  $(0, 0)$ . ◇

### Expression des coordonnées de la dérivée suivant un vecteur

Si  $f = (f_1, \dots, f_p)$ , il y a équivalence entre :

- la fonction  $f$  admet une dérivée en  $a \in U$  suivant un vecteur  $h$ .
- les fonctions  $f_1, \dots, f_p$  admettent une dérivée en  $a \in U$  suivant un vecteur  $h$ .

La dérivée de  $f$  en  $a$  suivant ce vecteur  $h$  est alors  $D_h f(a) = (D_h f_1(a), \dots, D_h f_p(a))$ .

#### **Proposition A2.8** (Une propriété des fonctions différentiables)

Si une fonction  $f : U \rightarrow \mathbb{R}^p$  est différentiable en  $a \in U$ , alors elle admet une dérivée au point  $a$  suivant tout vecteur  $h$ , et celle-ci est égale à  $D_h f(a) = df_a(h)$ .

*Preuve :* Supposons  $f$  différentiable en  $a$ , on a alors pour  $h$  fixé et  $t$  assez petit :

$$f(a + th) - f(a) = df_a(th) + o(\|th\|) = tdf_a(h) + |t| \|h\| \varepsilon(th).$$

On en déduit que  $f$  est dérivable en  $a$  suivant  $h$ , et la dérivée de  $f$  en  $a$  suivant  $h$  est :

$$D_h f(a) = \lim_{t \rightarrow 0} \frac{f(a + th) - f(a)}{t} = \lim_{t \rightarrow 0} \frac{tdf_a(h) + |t| \|h\| \varepsilon(th)}{t} = df_a(h).$$

□

#### **Proposition A2.9** (Expression de la différentielle à l'aide des dérivées partielles)

Si  $f : U \rightarrow \mathbb{R}^p$  est différentiable en  $a \in U$ , elle a des dérivées partielles en  $a$  qui sont égales à  $D_j f(a) = df_a(e_j)$ , et on a :

$$\text{pour tout } h = (h_1, \dots, h_n) \in \mathbb{R}^n, \quad df_a(h) = h_1 D_1 f(a) + \dots + h_n D_n f(a).$$

*Preuve :* Si  $f$  est différentiable en  $a \in U$ , elle a des dérivées en  $a$  suivant tout vecteur comme on vient de le voir, et elle a donc des dérivées partielles en  $a$  :

$$D_j f(a) = \lim_{t \rightarrow 0} \frac{f(a + te_j) - f(a)}{t} = \lim_{t \rightarrow 0} \frac{tdf_a(e_j) + o(\|te_j\|)}{t} = df_a(e_j).$$

On en déduit, par linéarité de la différentielle  $df_a$  pour  $h = (h_1, \dots, h_n) = h_1 e_1 + \dots + h_n e_n \in \mathbb{R}^n$  :

$$df_a(h) = df_a \left( \sum_{j=1}^n h_j e_j \right) = \sum_{j=1}^n h_j df_a(e_j) = \sum_{j=1}^n h_j D_j f(a).$$

□

### Remarque : calcul pratique des dérivées partielles

→ On dérive en 0 la fonction  $t \in \mathbb{R} \mapsto f(a_1, \dots, a_j + t, \dots, a_n)$ .

→ On dérive en  $a_j$  la fonction  $t \in \mathbb{R} \mapsto f(a_1, \dots, t, \dots, a_n)$ .

**Attention !** La réciproque de la proposition précédente est fausse : une fonction peut avoir des dérivées partielles sans être différentiable.

◇ **Exemple A2.5 :** Dérivées partielles de  $f : \mathbb{R}^2 \rightarrow \mathbb{R}, \begin{cases} (x, y) \neq (0, 0) \mapsto \frac{xy}{x^2+y^2} \\ (0, 0) \mapsto 0 \end{cases}$ .

• La fonction  $x \mapsto \frac{xy}{x^2+y^2}$  est dérivable en tout  $x$  si  $y \neq 0$ , et en tout  $x \neq 0$  si  $y = 0$ , de dérivée  $\frac{y(y^2-x^2)}{(x^2+y^2)^2} = D_1 f(x, y)$ .

• De même,  $D_2 f(x, y)$  existe en tout  $(x, y) \neq (0, 0)$  avec  $\frac{x(x^2-y^2)}{(x^2+y^2)^2} = D_2 f(x, y)$ .

• Si  $x = y = 0$ ,  $f(t, 0) = 0$  et  $t \mapsto f(t, 0) = 0$  est dérivable en tout  $t$ , de dérivée 0, donc  $D_1 f(0, 0) = 0$  et  $f(0, t) = 0$  donc  $D_2 f(0, 0) = 0$ .

Par contre, si  $a = (0, 0)$ ,  $\varphi_h$  n'est pas continue en 0 si  $h = (\alpha, \beta)$  où  $\alpha\beta \neq 0$ , donc  $D_h f(0, 0)$  n'existe pas et  $f$  n'est pas différentiable en  $(0, 0)$ . ◇

◇ **Exemple A2.6 :** Dérivées suivant  $h$  en  $(0, 0)$  de  $f : \mathbb{R}^2 \rightarrow \mathbb{R}, \begin{cases} (x, y) \neq (0, 0) \mapsto \frac{x^2 y}{x^4 + y^2} \\ (0, 0) \mapsto 0 \end{cases}$ .

Si  $h = (\alpha, \beta)$ ,  $\varphi_h(t) = f(t\alpha, t\beta) = \frac{t\alpha^2\beta}{t\alpha^4+\beta^2}$  et  $\frac{\varphi_h(t)}{t} = \frac{\alpha^2\beta}{t\alpha^4+\beta^2}$ .

• Si  $\alpha = 0$  ou  $\beta = 0$ ,  $\frac{\varphi_h(t)}{t} = 0$  et  $\lim_{t \rightarrow 0} \frac{\varphi_h(t)}{t} = 0$  donc  $D_{(\alpha, 0)} f(0, 0) = 0 = D_{(0, \beta)} f(0, 0)$ .

• Si  $\alpha\beta \neq 0$ ,  $\lim_{t \rightarrow 0} \frac{\varphi_h(t)}{t} = \frac{\alpha^2}{\beta}$  donc  $D_{(\alpha, \beta)} f(0, 0) = \frac{\alpha^2}{\beta} \neq 0$  et

$$D_h f(0, 0) \neq \alpha D_1 f(0, 0) + \beta D_2 f(0, 0). \quad \diamond$$

## 1-4 Interprétation matricielle de la différentiabilité (jacobienne, gradient)

On note  $f_1, \dots, f_p$  les fonctions composantes de  $f : f = (f_1, \dots, f_p)$ .

**Définition A2.4:** Soit un ouvert  $U$  de  $\mathbb{R}^n$  et une application  $f : U \rightarrow \mathbb{R}^p$ , différentiable en  $x \in U$ .

On appelle alors *matrice jacobienne de  $f$  en  $x$*  la matrice  $Jf_x$  de  $df_x$  relativement aux bases canoniques de  $\mathbb{R}^n$  et  $\mathbb{R}^p$ , c'est-à-dire, puisque  $df_x(e_j) = D_j f(x)$ , la matrice dont le coefficient  $(i, j)$  est  $D_j f_i(x)$  :

$$Jf_x = \begin{bmatrix} D_1 f_1(x) & D_2 f_1(x) & \cdots & D_n f_1(x) \\ D_1 f_2(x) & D_2 f_2(x) & \cdots & D_n f_2(x) \\ \vdots & \vdots & & \vdots \\ D_1 f_p(x) & D_2 f_p(x) & \cdots & D_n f_p(x) \end{bmatrix}.$$

Notons que  $Jf_x \in \mathcal{M}_{p,n}(\mathbb{R})$  et que si  $f$  est fonction que d'une seule variable, sa matrice jacobienne est une matrice colonne (et cette colonne est alors constituée des composantes du vecteur dérivée  $f'(x)$ ).

Enfin, la relation de différentiabilité  $f(x+h) = f(x) + df_x(h) + \|h\|\varepsilon(h)$  s'écrit :

$$\begin{bmatrix} f_1(x+h) \\ f_2(x+h) \\ \vdots \\ f_p(x+h) \end{bmatrix} = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_p(x) \end{bmatrix} + \begin{bmatrix} D_1 f_1(x) & D_2 f_1(x) & \cdots & D_n f_1(x) \\ D_1 f_2(x) & D_2 f_2(x) & \cdots & D_n f_2(x) \\ \vdots & \vdots & & \vdots \\ D_1 f_p(x) & D_2 f_p(x) & \cdots & D_n f_p(x) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} + \|h\| \begin{bmatrix} \varepsilon_1(h) \\ \varepsilon_2(h) \\ \vdots \\ \varepsilon_p(h) \end{bmatrix}.$$

Si  $p = 1$ , on peut caractériser vectoriellement la différentielle d'une fonction.

**Définition A2.5:** Étant donné une application  $f : U \rightarrow \mathbb{R}$  différentiable en  $a \in U$ , on appelle *gradient de  $f$  en  $a$*  l'unique vecteur de  $\mathbb{R}^n$  et noté  $\nabla f(a) = \text{grad} f(a)$  tel que :

$$\text{pour tout } h \in \mathbb{R}^n, \quad df_a(h) = D_h f(a) = \langle \nabla f(a), h \rangle.$$

Le point  $a$  est alors dit *régulier* si  $\nabla f(a) \neq 0$ , *singulier* ou *critique* si  $\nabla f(a) = 0$ .

La différentielle de  $f$  en  $a$  est ainsi déterminée par la donnée du gradient de  $f$  en  $a$  et la différentiabilité de  $f$  en  $a$  s'écrit donc dans ce contexte :

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + o(\|h\|).$$

**Remarque :** L'expression  $D_h f(a) = \sum_{j=1}^n h_j D_j f(a)$  conduit à :

$$\nabla f(a) = (D_1 f(a), \dots, D_n f(a)).$$

## 1-5 Fonctions de classe $C^1$ sur un ouvert

**Définition A2.6:**  $f : U \rightarrow \mathbb{R}^p$  est dite *de classe  $C^1$  sur  $U$*  si, pour tout  $1 \leq j \leq n$ ,  $D_j f$  existe et est continue sur  $U$ .

**Proposition A2.10** (*Théorème fondamental*)

Si une fonction  $f : U \rightarrow \mathbb{R}^p$  est de classe  $C^1$  sur  $U$ , alors elle est différentiable (et donc continue) sur  $U$ .

*Preuve :* Pour simplifier les notations, on établit le résultat lorsque  $n = 2$ .  
On munit  $\mathbb{R}^2$  de la norme  $\| \cdot \| = \| \cdot \|_\infty$ .

Considérons  $a \in U$  et montrons que  $f$  est différentiable en  $a$ .  
Si elle existe, cette différentielle est nécessairement  $h \mapsto h_1 D_1 f(a) + h_2 D_2 f(a)$ .  
Il suffit donc d'établir que la différence  $f(a+h) - f(a) - h_1 D_1 f(a) - h_2 D_2 f(a)$  est négligeable devant  $\|h\|$  lorsque  $h \rightarrow 0_n$ , c'est-à-dire ici

$$f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) - h_1 D_1 f(a_1, a_2) - h_2 D_2 f(a_1, a_2) = o(\|h\|).$$

À cet effet, on écrit cette expression comme somme des deux suivantes :

$$A(h) = f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) - h_1 D_1 f(a_1, a_2)$$

$$B(h) = f(a_1, a_2 + h_2) - f(a_1, a_2) - h_2 D_2 f(a_1, a_2).$$

Nous majorons maintenant chacune de ces 2 expressions.

La première expression s'écrit  $A(h) = \alpha(h_1) - \alpha(0)$  avec :

$$\alpha(t) = f(a_1 + t, a_2 + h_2) - t D_1 f(a_1, a_2).$$

Comme  $\alpha'(t) = D_1 f(a_1 + t, a_2 + h_2) - D_1 f(a_1, a_2)$  on a, par continuité de  $D_1 f$  :

pour tout  $\varepsilon > 0$ , il existe  $\eta_1 > 0$  tel que, pour  $t \in [-|h_1|, |h_1|]$ , si  $\|h\| \leq \eta_1$ , alors  $\|\alpha'(t)\| \leq \varepsilon$

L'inégalité des accroissements finis (voir 1.6) permet alors d'en déduire  $\|\alpha(h_1) - \alpha(0)\| \leq \varepsilon |h_1|$ .  
Ainsi, on a  $\|A(h)\| \leq \varepsilon |h_1|$  pour  $\|h\| \leq \eta_1$ .

La seconde expression s'écrit  $B(h) = \beta(h_2) - \beta(0)$  avec :

$$\beta(t) = f(a_1, a_2 + t) - t D_2 f(a_1, a_2).$$

Comme  $\beta'(t) = D_2 f(a_1, a_2 + t) - D_2 f(a_1, a_2)$  on a, par continuité de  $D_2 f$  :

pour tout  $\varepsilon > 0$ , il existe  $\eta_2 > 0$  tel que, pour  $t \in [-|h_2|, |h_2|]$ , si  $\|h\| \leq \eta_2$ , alors  $\|\beta'(t)\| \leq \varepsilon$

L'inégalité des accroissements finis permet alors d'en déduire  $\|\beta(h_2) - \beta(0)\| \leq \varepsilon |h_2|$ . Ainsi, on a  $\|B(h)\| \leq \varepsilon |h_2|$  pour  $\|h\| \leq \eta_2$ .

Par addition, il en résulte que  $\|A(h) + B(h)\| \leq 2\varepsilon \|h\|$  pour  $\|h\| \leq \min(\eta_1, \eta_2)$ .

Cela signifie exactement que  $f$  est différentiable en  $a$ , puisque :

$$f(a+h) - f(a) - h_1 D_1 f(a) - h_2 D_2 f(a) = o(\|h\|).$$

□

**Proposition A2.11** (*Définition équivalente des fonctions de classe  $C^1$* )

Pour une fonction  $f : U \rightarrow \mathbb{R}^p$ , il y a équivalence entre :

- l'application  $x \mapsto D_h f(x) = df_x(h)$  est continue sur  $U$  pour tout  $h \in \mathbb{R}^n$ .
- $f$  est de classe  $C^1$  sur  $U$ .

*Preuve :*

- Supposons l'application  $x \mapsto D_h f(x)$  continue sur  $U$  pour tout vecteur  $h \in \mathbb{R}^n$ .

En faisant  $h = e_j$ , on en déduit la continuité de  $x \mapsto D_j f(x)$  pour  $1 \leq j \leq n$ .

• Supposons les dérivées partielles  $D_1f, \dots, D_nf$  de  $f$  continues sur  $U$ . Dans ce cas,  $f$  est différentiable sur  $U$  d'après la proposition A2.10 et on a :

$$D_hf(x) = df_x(h) = h_1D_1f(x) + \dots + h_nD_nf(x).$$

On en déduit la continuité de  $D_hf$  pour tout vecteur  $h \in \mathbb{R}^n$ .

□

**Remarque :** On établit l'équivalence, pour  $f : U \rightarrow \mathbb{R}^p$ , des 2 propriétés suivantes :

- $f$  est de classe  $C^1$  sur  $U$
- $f$  est différentiable sur  $U$  et l'application  $x \in U \mapsto df_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$  est continue.

En effet,

• Si  $f$  est différentiable sur  $U$  et si l'application  $x \in U \mapsto df_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$  est continue, comme  $D_jf(x) = df_x(e_j)$ , on a, en tout  $a \in U$  :

$$\|D_jf(x) - D_jf(a)\| = \|df_x(e_j) - df_a(e_j)\| \leq \|df_x - df_a\| \|e_j\|.$$

La continuité de  $x \mapsto df_x$  implique donc celle des dérivées partielles  $x \mapsto D_jf(x)$ .

• Si  $f$  est de classe  $C^1$  sur  $U$ , comme  $df_x(h) = h_1D_1f(x) + \dots + h_nD_nf(x)$ , on a en tout  $a \in U$ , en munissant  $\mathbb{R}^n$  de la norme  $\| \cdot \| = \| \cdot \|_\infty$  :

$$\|(df_x - df_a)(h)\| = \left\| \sum_{j=1}^n h_j(D_jf(x) - D_jf(a)) \right\| \leq \|h\| \sum_{j=1}^n \|D_jf(x) - D_jf(a)\|.$$

On en déduit, par définition de la norme de  $df_x - df_a$  ( $\|L\| = \sup_{h \neq 0} \frac{\|L(h)\|}{\|h\|}$ ) :

$$\|df_x - df_a\| \leq \sum_{j=1}^n \|D_jf(x) - D_jf(a)\|.$$

La continuité des dérivées partielles  $x \mapsto D_jf(x)$  implique donc celle de  $x \mapsto df_x$ .

### Exemples de fonctions différentiables et de différentielles

En pratique, on montrera le plus souvent qu'une fonction est différentiable en établissant que ses dérivées partielles sont continues (en utilisant le résultat qu'on vient d'établir).

◇ **Exemple A2.7:** gradient et classe  $C^1$  de la norme euclidienne  $N : x \mapsto \|x - a\|_2$  où  $a \in \mathbb{R}^n$ .

$N(x) = \sqrt{\sum_{j=1}^n (x_j - a_j)^2}$ .  $N$  n'est pas différentiable en  $a$  mais, pour  $x \neq a$ ,

$$D_jN(x) = \frac{2(x_j - a_j)}{2\sqrt{\sum_{j=1}^n (x_j - a_j)^2}} = \frac{x_j - a_j}{N(x - a)}.$$

Comme les composantes de  $\nabla N(x)$  sont les  $D_j N(x)$ , on a donc :

$$\nabla N(x) = \frac{x - a}{\|x - a\|_2}.$$

Enfin,  $N$  est de classe  $C^1$  sur  $\mathbb{R}^n \setminus \{a\}$  car les  $x \mapsto D_j N(x)$  y sont continues.  $\diamond$

### Opérations sur les fonctions de classe $C^1$

#### • Combinaison linéaire de fonctions de classe $C^1$

Soit deux réels  $\lambda, \mu$  et deux fonctions  $f, g : U \rightarrow \mathbb{R}^p$  de classe  $C^1$  sur  $U$ .

Les fonctions  $f$  et  $g$  étant différentiables en  $x \in U$ , par la proposition A2.3,  $\lambda f + \mu g$  est différentiable en  $x$  et on a :

$$\text{pour tout } j = 1, \dots, n, \quad d(\lambda f + \mu g)_x(e_j) = \lambda df_x(e_j) + \mu dg_x(e_j).$$

Cette égalité donne, pour  $1 \leq j \leq n$ ,  $D_j(\lambda f + \mu g)(x) = \lambda D_j f(x) + \mu D_j g(x)$ .

Ainsi, la continuité des dérivées partielles de  $f$  et  $g$  sur l'ouvert  $U$  implique celle des dérivées partielles de  $\lambda f + \mu g$ , et on en déduit que  $\lambda f + \mu g$  est de classe  $C^1$  sur  $U$ .

L'ensemble  $C^1(U, \mathbb{R}^p)$  des fonctions  $C^1$  de  $U$  dans  $\mathbb{R}^p$  forme un espace vectoriel.

#### • Produit d'une fonction scalaire de classe $C^1$ par une fonction de classe $C^1$

Soit deux fonctions  $\lambda : U \rightarrow \mathbb{R}$  et  $f : U \rightarrow \mathbb{R}^p$  de classe  $C^1$  sur  $U$ .

Les fonctions  $\lambda$  et  $f$  étant différentiables en  $x \in U$ , par la proposition A2.4,  $\lambda f$  est différentiable en  $x$  et on a :

$$\text{pour tout } j = 1, \dots, n, \quad d(\lambda f)_x(e_j) = d\lambda_x(e_j)f(x) + \lambda(x)df_x(e_j).$$

Cette égalité donne, pour  $1 \leq j \leq n$  :  $D_j(\lambda f)(x) = (D_j \lambda)(x)f(x) + \lambda(x)D_j f(x)$ .

Ainsi, la continuité des dérivées partielles des fonctions  $\lambda$  et  $f$  sur l'ouvert  $U$  implique celle des dérivées partielles de  $\lambda f$ , et on en déduit que  $\lambda f$  est de classe  $C^1$  sur  $U$ .

#### • Composée d'une application linéaire et d'une fonction de classe $C^1$

Soit une fonction  $f : U \rightarrow \mathbb{R}^p$  de classe  $C^1$  et une application linéaire  $L : \mathbb{R}^p \rightarrow \mathbb{R}^m$ .

La fonction  $f$  étant différentiable en  $x \in U$ , par la proposition A2.5,  $L \circ f$  est différentiable en  $x$  et on a :

$$\text{pour tout } j = 1, \dots, n, \quad d(L \circ f)_x(e_j) = L \circ df_x(e_j).$$

Cette égalité donne, pour  $1 \leq j \leq n$ ,  $D_j(L \circ f)(x) = L(D_j f(x))$ .

Ainsi, la continuité des dérivées partielles de la fonction  $f$  sur  $U$  et celle de  $L$  impliquent celle des dérivées partielles de  $L \circ f$ , et on en déduit que  $L \circ f$  est de classe  $C^1$  sur  $U$ .

#### • Composée de deux fonctions quelconques de classe $C^1$

Soit  $f : U \rightarrow \mathbb{R}^p$  de composantes  $f_1, \dots, f_p$ , de classe  $C^1$  sur  $U$  et  $g : V \rightarrow \mathbb{R}^m$ , de composantes  $g_1, \dots, g_m$ , de classe  $C^1$  sur l'ouvert  $V$ .



Si  $f(U) \subset V$ ,  $f$  étant différentiable en  $x \in U$  et  $g$  différentiable en  $f(x) \in V$ , par la proposition A2.7,  $g \circ f$  est différentiable en  $x$  et on a  $d(g \circ f)_x = dg_{f(x)} \circ df_x$ , c'est-à-dire matriciellement :

$$J(g \circ f)_x = Jg_{f(x)} \times Jf_x.$$

En effectuant le produit de ces deux matrices jacobienues et en examinant ses colonnes, on obtient les égalités suivantes, valables pour  $1 \leq i \leq p$  et  $1 \leq j \leq n$  :

$$D_j(g_i \circ f)(x) = \sum_{k=1}^p (D_k g_i)(f(x)) \times D_j f_k(x).$$

Les égalités obtenues ici pour  $1 \leq i \leq m$  peuvent être regroupées en une seule s'écrivant :

$$D_j(g \circ f)(x) = \sum_{k=1}^p (D_k g)(f(x)) \times D_j f_k(x).$$

Ainsi, la continuité des dérivées partielles des fonctions  $f$  sur  $U$  et  $g$  sur  $V$  implique celle des dérivées partielles de  $g \circ f$ , et on en déduit que  $\boxed{g \circ f \text{ est de classe } C^1 \text{ sur } U}$ .

### Autres formes de cette relation

Cette formule fondamentale peut s'écrire de façon plus développée comme suit :

$$D_j(g(f_1, \dots, f_p))(x) = \sum_{k=1}^p (D_k g)(f_1(x), \dots, f_p(x)) \times D_j f_k(x).$$

◇ **Exemple A2.8:** Classe  $C^1$  de  $F : (r, \theta) \in \mathbb{R}^2 \mapsto f(r \cos \theta, r \sin \theta)$  où  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  est supposée de classe  $C^1$ .

Par composition,  $(r, \theta) \mapsto F(r, \theta) = f(r \cos \theta, r \sin \theta)$  est de classe  $C^1$  sur  $\mathbb{R}^2$  et

$$\begin{aligned} D_1 F(r, \theta) &= \frac{\partial F}{\partial r}(r, \theta) = \cos \theta D_1 f(r \cos \theta, r \sin \theta) + \sin \theta D_2 f(r \cos \theta, r \sin \theta) \\ D_2 F(r, \theta) &= \frac{\partial F}{\partial \theta}(r, \theta) = -r \sin \theta D_1 f(r \cos \theta, r \sin \theta) + r \cos \theta D_2 f(r \cos \theta, r \sin \theta) \end{aligned}$$

**Remarque :** Dans le cas particulier où  $U$  désigne un intervalle de  $\mathbb{R}$  ( $n = 1$ ), alors  $g \circ f$  est une fonction d'une variable réelle à valeurs dans  $\mathbb{R}^m$  et la formule précédente devient :

$$\frac{d}{dx} g(f_1, \dots, f_p)(x) = \sum_{k=1}^p f'_k(x) (D_k g)(f_1(x), \dots, f_p(x)).$$

On observe que ce cas particulier équivaut au cas général, puisqu'on dérive toujours par rapport à une seule variable réelle.

### • Cas particulier des fonctions à valeurs réelles ( $p = 1$ )

→ L'ensemble des fonctions de classe  $C^1$  de  $U$  dans  $\mathbb{R}$  forme une algèbre :

Si  $\lambda, \mu$  sont deux réels, si  $f, g$  sont deux fonctions de classe  $C^1$  de  $U$  dans  $\mathbb{R}$ , on a :

$$\nabla(\lambda f + \mu g) = \lambda \nabla f + \mu \nabla g \quad ; \quad \nabla(f \times g) = g \nabla f + f \nabla g.$$

◇ **Exemple A2.9:** *gradient de la fonction  $x \mapsto \|x\|^k$  ( $k \geq 1$ )*

On vérifie par récurrence, si  $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$  est de classe  $C^1$ , que  $\nabla(f^k) = k f^{k-1} \nabla f$ .  
On en déduit le gradient en  $x$  de l'application  $F : x \mapsto \|x\|^k$  sur l'ouvert  $U = \mathbb{R}^n \setminus \{0\}$  :

$$\nabla F(x) = k \|x\|^{k-1} \frac{x}{\|x\|} = k \|x\|^{k-2} x. \quad \diamond$$

De plus, si  $g$  ne s'annule pas sur  $U$ , la fonction  $f/g$  est de classe  $C^1$  sur  $U$  et on a :

$$\nabla \left( \frac{f}{g} \right) = \frac{g \nabla f - f \nabla g}{g^2}.$$

→ Si  $u : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$  est de classe  $C^1$ , si  $f : U \rightarrow \mathbb{R}$  est de classe  $C^1$ , si  $f(I) \subset U$ , alors  $f \circ u$  est de classe  $C^1$  de  $I$  dans  $\mathbb{R}$  et on a,

$$\text{pour tout } x \in I, (f \circ u)'(x) = \langle \nabla f(u(x)), u'(x) \rangle.$$

C'est par exemple une conséquence de la formule déjà établie :

$$\frac{d}{dx} f(u_1(x), \dots, u_n(x)) = \sum_{j=1}^n D_j f(u(x)) \times u'_j(x).$$

## 1-6 Formules des accroissements finis

### a. Cas $p = 1$ .

**Proposition A2.12** (*Formule de Taylor Mac-Laurin à l'ordre 1*)

Soit  $f : U \rightarrow \mathbb{R}$  une fonction différentiable en tout point du segment  $[u, v]$  contenu dans l'ouvert  $U$  de  $\mathbb{R}^n$ . Alors il existe  $\theta \in ]0, 1[$  tel que :

$$f(u + h) - f(u) = df_{u+\theta h}(h).$$

*Preuve :* L'application  $\psi : t \in \mathbb{R} \mapsto u + th \in \mathbb{R}^n$  est différentiable sur  $\mathbb{R}$ , de dérivée  $\psi'(t) = h$  et donc de différentielle  $d\psi_t : \alpha \in \mathbb{R} \mapsto \alpha h$ . Il en résulte que la fonction :

$$\varphi = f \circ \psi : t \in [0, 1] \mapsto f(u + th) \in \mathbb{R}$$

est différentiable sur  $]0, 1[$  (proposition A2.7), de différentielle (d'après la remarque p.?) :

$$d\varphi_t : \alpha \in \mathbb{R} \mapsto df_{\psi(t)}(d\psi_t(\alpha)) = \alpha df_{\psi(t)}(h).$$

On a donc,

$$\text{pour tout } t \in ]0, 1[, \quad \varphi'(t) = df_{\psi(t)}(h).$$

Or, en utilisant la formule classique des accroissements finis pour  $\varphi$ , on sait qu'il existe  $\theta \in ]0, 1[$  tel que :

$$f(u+h) - f(u) = \varphi(1) - \varphi(0) = (1-0)\varphi'(\theta) = df_{\psi(\theta)}(h)$$

ce qui donne le résultat. □

On étend ainsi la formule classique des accroissements finis pour les fonctions réelles d'une ou plusieurs variables réelles. On peut donner une autre expression de cette formule :

**Proposition A2.12'** (*Formule de Taylor avec reste intégral*)

Soit  $f : U \rightarrow \mathbb{R}$  une fonction différentiable en tout point du segment  $[u, v]$  contenu dans l'ouvert  $U$  de  $\mathbb{R}^n$ . Alors on a :

$$f(u+h) = f(u) + \int_0^1 d_{u+th}f(h) dt.$$

*Preuve :* Reprenant les notations de la démonstration précédente, on peut écrire :

$$f(u+h) - f(u) = \varphi(1) - \varphi(0) = \int_0^1 \varphi'(t) dt.$$

Il suffit alors de remplacer  $\varphi'(t)$  par la valeur trouvée. □

## b. Cas $n = 1$

**Proposition A2.13** (*Inégalité des accroissements finis*)

Soit  $f : I \rightarrow \mathbb{R}^p$  une fonction de classe  $C^1$  en tout point de l'intervalle ouvert  $I$  de  $\mathbb{R}$  et vérifiant  $\|f'(t)\| \leq M$  pour tout  $t \in I$ . Alors, en tous  $a, b \in I$  où  $a < b$  :

$$\|f(b) - f(a)\| \leq M(b-a).$$

*Preuve :* Soit  $\varepsilon$  un réel strictement positif donné. On pose :

$$I(\varepsilon) = \{t \in [a, b] ; \|f(u) - f(a)\| \leq (\varepsilon + M)(u-a) \text{ pour tout } u \in [a, t]\}.$$

Il est clair que  $a$  appartient à  $I(\varepsilon)$  et que, puisque :

$$\text{si } t' < t \text{ et } t \in I(\varepsilon) \text{ alors } t' \in I(\varepsilon),$$

$I(\varepsilon)$  est un intervalle (contenu dans  $[a, b]$ ) dont une extrémité est  $a$ . Soit  $c$  l'autre extrémité de cet intervalle (*i.e.*  $c = \sup_{t \in I(\varepsilon)} t$ ). Comme :

$$\|f(u) - f(a)\| \leq (\varepsilon + M)(u-a) \text{ pour tout } u \in [a, c]$$

on déduit, en tenant compte de la continuité de  $f$ , que, si  $u$  tend vers  $c$  par valeurs inférieures, l'inégalité est encore vraie en  $u = c$  et  $I(\varepsilon)$  est un intervalle fermé.

Si  $c$  était strictement plus petit que  $b$ , par la classe  $C^1$  de  $f$ , il existerait  $h > 0$  tel que :

$$\left\| \frac{f(t) - f(c)}{t - c} - f'(c) \right\| \leq \varepsilon \text{ pour tout } t \in [c, c + h].$$

Pour les mêmes valeurs de  $t$ , on en déduirait, en utilisant l'inégalité triangulaire :

$$\left\| \frac{f(t) - f(c)}{t - c} \right\| \leq \left\| \frac{f(t) - f(c)}{t - c} - f'(c) \right\| + \|f'(c)\| \leq \varepsilon + M$$

soit encore  $\|f(t) - f(c)\| \leq (\varepsilon + M)(t - c)$ .

Puisque  $c \in I(\varepsilon)$ , on aurait, pour tout  $t \in ]c, c + h[$ , en utilisant à nouveau l'inégalité triangulaire :

$$\begin{aligned} \|f(t) - f(a)\| &\leq \|f(t) - f(c)\| + \|f(c) - f(a)\| \\ &\leq (\varepsilon + M)(t - c) + (\varepsilon + M)(c - a) = (\varepsilon + M)(t - a) \end{aligned}$$

et  $c + h$  (avec  $h > 0$ ) appartiendrait à  $I(\varepsilon)$ , ce qui est absurde par définition de  $c$ . Il en résulte que l'on a  $I(\varepsilon) = [a, b]$  et donc :

$$\|f(t) - f(a)\| \leq (\varepsilon + M)(t - a) \text{ pour tout } \varepsilon > 0 \text{ et pour tout } t \in [a, b].$$

Il suffit alors de faire  $\varepsilon \rightarrow 0$  et  $t = b$  pour avoir le résultat.

□

### c. Cas général

**Proposition A2.13'** (*Inégalité des accroissements finis dans le cas général*)

Soit  $f : U \rightarrow \mathbb{R}^p$  une fonction de classe  $C^1$  en tout point de l'ouvert  $U$  de  $\mathbb{R}^n$  et vérifiant  $\|df_u\| \leq M$  pour tout  $u \in U$ . Alors, pour tout segment  $[u, v]$  contenu dans  $U$ , on a l'inégalité :

$$\|f(v) - f(u)\| \leq M\|v - u\|.$$

*Preuve :* Notant  $\psi$  l'application  $t \in \mathbb{R} \mapsto u + t(v - u) \in \mathbb{R}^n$  (qui est dérivable sur  $\mathbb{R}$  et de différentielle en  $t$  :  $d\psi_t : \alpha \in \mathbb{R} \mapsto \alpha(v - u)$ ), l'application  $\varphi = f \circ \psi$  est différentiable en tout  $t \in [0, 1]$  et on a :

$$d\varphi_t : \alpha \in \mathbb{R} \mapsto df_{\psi(t)}(d\psi_t(\alpha)) = \alpha df_{\psi(t)}(v - u) \in \mathbb{R}^p.$$

La fonction  $\varphi : t \in [0, 1] \mapsto \mathbb{R}^p$  est donc dérivable en tout  $t \in [0, 1]$  et on a :

$$\varphi'(t) = df_{\psi(t)}(v - u).$$

On a alors, pour tout  $t \in [0, 1]$ ,

$$\|\varphi'(t)\| \leq \|df_{\psi(t)}\| \|v - u\| \leq M\|v - u\|.$$

On peut alors appliquer le résultat de la proposition A2.15 à  $\varphi$  sur le segment  $[0, 1]$  et il vient :

$$\|f(v) - f(u)\| = \|\varphi(1) - \varphi(0)\| \leq M\|v - u\|(1 - 0) = M\|v - u\|. \quad \square$$

**Remarque :** Cette inégalité permet en particulier la propriété : pour qu'une fonction de classe  $C^1$  soit constante sur  $U$ , il faut et il suffit que sa différentielle soit nulle sur  $U$ .

## 2- Dérivées d'ordre supérieur

**Définition A2.7:** On définit par récurrence qu'une fonction  $f : U \rightarrow \mathbb{R}^p$  est de classe  $C^k$  sur l'ouvert  $U$  si  $f$  est de classe  $C^1$  et si  $D_j f$  est de classe  $C^{k-1}$  pour tout  $j$ .

Enfin, on dit que  $f$  est de classe  $C^\infty$  si elle est de classe  $C^k$  pour tout  $k \in \mathbb{N}$ .

Les dérivées partielles secondes de  $f$  se notent :

$$D_i D_j f = D_i(D_j f), \quad \text{avec } D_j^2 f = D_j(D_j f) \text{ si } i = j.$$

Une autre notation (moins conseillée) est la suivante :

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_j} \right), \quad \text{avec } \frac{\partial^2 f}{\partial x_j^2} = \frac{\partial}{\partial x_j} \left( \frac{\partial f}{\partial x_j} \right) \text{ si } i = j.$$

**Proposition A2.14** (*Définitions équivalentes des fonctions de classe  $C^2$* )

Pour une fonction  $f : U \rightarrow \mathbb{R}^p$  de classe  $C^1$ , il y a équivalence entre :

- pour tout vecteur  $h \in \mathbb{R}^n$ , la fonction  $D_h f : U \rightarrow \mathbb{R}^p$  est de classe  $C^1$  ;
- $f$  est de classe  $C^2$ .

*Preuve :*

- Supposons l'application  $x \mapsto D_h f(x)$  de classe  $C^1$  sur  $U$  pour tout vecteur  $h \in E$ .

En faisant  $h = e_j$ , on obtient la classe  $C^1$  de  $x \mapsto D_j f(x)$  pour  $1 \leq j \leq n$ .

- Supposons les dérivées partielles  $D_1 f, \dots, D_n f$  de classe  $C^1$  sur  $U$ .

Leur classe  $C^1$  implique celle de  $D_h f$  pour tout vecteur  $h$  d'après la relation :

$$D_h f(x) = df_x(h) = h_1 D_1 f(x) + \dots + h_n D_n f(x).$$

□

Ce résultat qui généralise celui obtenu à la proposition A2.11 pour les fonctions de classe  $C^1$ , peut se généraliser aux fonctions de classe  $C^k$  :

**Proposition A2.14'** (*Définition équivalentes des fonctions de classe  $C^k$* )

Pour une fonction  $f : U \rightarrow \mathbb{R}^p$  de classe  $C^1$ , il y a équivalence entre :

- pour tout vecteur  $h \in \mathbb{R}^n$ , la fonction  $D_h f : U \rightarrow \mathbb{R}^p$  est de classe  $C^{k-1}$  ;
- $f$  est de classe  $C^k$ .

### Caractérisation d'une fonction de classe $C^k$ par ses composantes

Une fonction  $f : U \rightarrow \mathbb{R}^p$  est de classe  $C^k$  si et seulement si ses composantes  $f_1, \dots, f_p$  sont de classe  $C^k$ .

En effet,

- On sait que  $f$  est de classe  $C^1$  si et seulement si ses composantes le sont, et on a :

$$\text{pour tout } j \in \{1, \dots, n\}, \quad D_j f(x) = (D_j f_1(x), \dots, D_j f_p(x)).$$

- Supposons qu'une fonction est de classe  $C^{k-1}$  si et seulement si ses composantes le sont. Par définition,  $f$  est de classe  $C^k$  si et seulement si  $D_j f$  est  $C^{k-1}$  pour  $1 \leq j \leq n$ .

Selon l'hypothèse de récurrence, les  $n$  fonctions  $D_j f$  sont de classe  $C^{k-1}$  si et seulement si leurs composantes  $D_j f_i$  sont de classe  $C^{k-1}$  ( $1 \leq i \leq p$ ), et donc, par définition, si et seulement si les composantes  $f_1, \dots, f_p$  sont de classe  $C^k$ , d'où le résultat.

### Opérations sur les fonctions de classe $C^k$

- **Combinaison linéaire de fonctions de classe  $C^k$**

Soit deux réels  $\lambda, \mu$  et deux fonctions  $f, g : U \rightarrow \mathbb{R}^p$ , de classe  $C^k$  sur  $U$ .

Comme  $f$  et  $g$  sont de classe  $C^1$ , on sait que  $\lambda f + \mu g$  est de classe  $C^1$  sur  $U$  et :

$$D_j(\lambda f + \mu g)(x) = \lambda D_j f(x) + \mu D_j g(x).$$

Supposons que si  $f$  et  $g$  sont de classe  $C^{k-1}$  sur  $U$ ,  $\lambda f + \mu g$  est de classe  $C^{k-1}$  sur  $U$ . Comme  $f$  et  $g$  sont de classe  $C^k$ , les dérivées partielles  $D_j f$  et  $D_j g$  sont de classe  $C^{k-1}$ . L'hypothèse de récurrence et la formule précédente montrent que les dérivées partielles de  $\lambda f + \mu g$  sont aussi de classe  $C^{k-1}$ , et donc que  $\lambda f + \mu g$  est de classe  $C^k$  sur  $U$ . L'ensemble  $C^k(U, \mathbb{R}^p)$  des fonctions  $C^k$  de  $U$  dans  $\mathbb{R}^p$  forme un espace vectoriel.

- **Produit d'une fonction scalaire et d'une fonction de classe  $C^k$**

En raisonnant sur  $D_j(\lambda f)(x) = D_j \lambda(x) f(x) + \lambda(x) D_j f(x)$ , on établit de même que :

$$\boxed{\lambda f \text{ est de classe } C^k \text{ sur } U}.$$

- **Composée de deux fonctions quelconques de classe  $C^k$**

Soit  $f : U \rightarrow \mathbb{R}^p$  de classe  $C^k$  sur  $U$  et  $g : V \rightarrow \mathbb{R}^m$  de classe  $C^k$  sur l'ouvert  $V$  avec  $f(U) \subset V$ .

Comme  $f$  et  $g$  sont de classe  $C^1$ , on sait que  $g \circ f$  est de classe  $C^1$  sur  $U$  et :

$$D_j(g \circ f)(x) = \sum_{q=1}^m (D_q g)(f(x)) \times D_j f_q(x).$$

Supposons que, si  $g$  et  $f$  sont de classe  $C^{k-1}$ , alors  $g \circ f$  est de classe  $C^{k-1}$  sur  $U$ . Comme  $f$  et  $g$  sont de classe  $C^k$ , les dérivées partielles  $D_j f_q$ ,  $D_q g$  sont de classe  $C^{k-1}$ . L'hypothèse de récurrence et la formule précédente montrent que les dérivées partielles

de  $g \circ f$  sont aussi de classe  $C^{k-1}$ , et donc  $\boxed{g \circ f \text{ est de classe } C^k \text{ sur } U}$ .

**Proposition A2.15** (*Théorème de Schwarz*)

Pour toute fonction  $f : U \rightarrow \mathbb{R}^p$  de classe  $C^2$ , on a, pour tout  $1 \leq i, j \leq p$  :

$$D_i D_j f = D_j D_i f \text{ ou } \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

*Preuve* : (à sauter à première lecture) Puisque seules interviennent les variables  $i$  et  $j$ , tout se passe comme s'il n'y avait que deux variables (qu'on notera  $x_1$  et  $x_2$ ), ce qui ramène au cas  $n = 2$ .

Cet espace  $\mathbb{R}^2$  est muni de la norme  $\| \cdot \| = \| \cdot \|_\infty$ .

Pour  $x = (x_1, x_2)$  et  $h = (h_1, h_2)$  de composantes supposées non nulles et assez petites pour avoir  $[x_1, x_1 + h_1] \times [x_2, x_2 + h_2] \subset U$ , on pose :

$$\Delta(h) = \frac{1}{h_1 h_2} (f(x_1 + h_1, x_2 + h_2) - f(x_1 + h_1, x_2) - f(x_1, x_2 + h_2) + f(x_1, x_2)).$$

En agencant de deux façons différentes cette expression, nous allons démontrer que sa limite quand  $h \rightarrow 0$  est d'une part  $D_1 D_2 f(x)$ , d'autre part  $D_2 D_1 f(x)$ .

L'unicité de la limite établira alors le théorème de Schwarz.

- D'une part,  $h_1 h_2 \Delta(h)$  est la différence des deux expressions suivantes :

$$f(x_1 + h_1, x_2 + h_2) - f(x_1 + h_1, x_2) = \int_{x_2}^{x_2 + h_2} D_2 f(x_1 + h_1, t_2) dt_2$$

$$f(x_1, x_2 + h_2) - f(x_1, x_2) = \int_{x_2}^{x_2 + h_2} D_2 f(x_1, t_2) dt_2.$$

On en déduit par différence :

$$h_1 h_2 \Delta(h) = \int_{x_2}^{x_2 + h_2} (D_2 f(x_1 + h_1, t_2) - D_2 f(x_1, t_2)) dt_2.$$

Il en résulte, en considérant l'intégrande comme une intégrale :

$$h_1 h_2 \Delta(h) = \int_{x_2}^{x_2 + h_2} \left( \int_{x_1}^{x_1 + h_1} D_1 D_2 f(t_1, t_2) dt_1 \right) dt_2.$$

La différence  $\Delta(h) - D_1 D_2 f(x)$  s'exprime alors comme suit :

$$\frac{1}{h_1 h_2} \int_{x_2}^{x_2 + h_2} \left( \int_{x_1}^{x_1 + h_1} (D_1 D_2 f(t_1, t_2) - D_1 D_2 f(x_1, x_2)) dt_1 \right) dt_2.$$

L'inégalité de la moyenne permet alors de majorer  $\|\Delta(h) - D_1 D_2 f(x)\|$  par :

$$\frac{1}{|h_1 h_2|} \left| \int_{x_2}^{x_2 + h_2} \left| \int_{x_1}^{x_1 + h_1} \|D_1 D_2 f(t_1, t_2) - D_1 D_2 f(x_1, x_2)\| dt_1 \right| dt_2 \right|.$$

Exploitions alors la continuité de  $D_1 D_2 f$  en  $x$  et donnons nous un réel  $\varepsilon > 0$  :

il existe  $\alpha_1 > 0$  tel que, pour tout  $t \in U$ , si  $\|t - x\| \leq \alpha_1$ , alors  $\|D_1 D_2 f(t) - D_1 D_2 f(x)\| \leq \varepsilon$ .

On en déduit alors que  $\|\Delta(h) - D_1 D_2 f(x)\| \leq \varepsilon$  pour  $\|h\| \leq \alpha_1$ .

Il en résulte que  $\lim_{h \rightarrow 0} \Delta(h) = D_1 D_2 f(x)$ .

- D'autre part,  $h_1 h_2 \Delta(h)$  est la différence des expressions suivantes :

$$f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2 + h_2) = \int_{x_1}^{x_1 + h_1} D_1 f(t_1, x_2 + h_2) dt_1$$

$$f(x_1 + h_1, x_2) - f(x_1, x_2) = \int_{x_1}^{x_1 + h_1} D_1 f(t_1, x_2) dt_1.$$

On en déduit par différence :

$$h_1 h_2 \Delta(h) = \int_{x_1}^{x_1 + h_1} (D_1 f(t_1, x_2 + h_2) - D_1 f(t_1, x_2)) dt_1.$$

Il en résulte, en exprimant à nouveau l'intégrande comme une intégrale :

$$h_1 h_2 \Delta(h) = \int_{x_1}^{x_1 + h_1} \left( \int_{x_2}^{x_2 + h_2} D_2 D_1 f(t_1, t_2) dt_2 \right) dt_1.$$

On remarque alors que la différence  $\Delta(h) - D_2 D_1 f(x)$  s'exprime comme suit :

$$\frac{1}{h_1 h_2} \int_{x_1}^{x_1 + h_1} \left( \int_{x_2}^{x_2 + h_2} (D_2 D_1 f(t_1, t_2) - D_2 D_1 f(x_1, x_2)) dt_2 \right) dt_1.$$

L'inégalité de la moyenne permet alors de majorer  $\|\Delta(h) - D_1 D_2 f(x)\|$  par :

$$\frac{1}{|h_1 h_2|} \left| \int_{x_2}^{x_2 + h_2} \left| \int_{x_1}^{x_1 + h_1} \|D_2 D_1 f(t_1, t_2) - D_2 D_1 f(x_1, x_2)\| dt_1 \right| dt_2 \right|.$$

En utilisant cette fois-ci la continuité de  $D_2 D_1 f$  en  $x$ , on obtient, par un raisonnement analogue  $\lim_{h \rightarrow 0} \Delta(h) = D_2 D_1 f(x)$ .

Par unicité de la limite, on en déduit le résultat désiré  $D_1 D_2 f(x) = D_2 D_1 f(x)$ .

□

**Attention !** Le théorème de Schwarz peut tomber en défaut lorsque les dérivées  $D_1 D_2 f$  et  $D_2 D_1 f$  ne sont pas continues.

◇ **Exemple A2.10:** Calculer  $D_1 D_2 f(0, 0)$  et  $D_2 D_1 f(0, 0)$  pour

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \begin{cases} (x, y) \neq (0, 0) \mapsto xy \frac{x^2 - y^2}{x^2 + y^2} \\ (0, 0) \mapsto 0 \end{cases}.$$

La fonction  $f$  est continue, y compris en  $(0, 0)$  car  $|f(x, y)| \leq |xy| \leq (x^2 + y^2)/2$ .

Elle est de classe  $C^1$ , y compris en  $(0, 0)$ , et ses dérivées partielles, nulles à l'origine, sont :

$$D_1 f(x, y) = y \frac{x^4 + 4x^2 y^2 - y^4}{(x^2 + y^2)^2} ; \quad D_2 f(x, y) = x \frac{x^4 - 4x^2 y^2 - y^4}{(x^2 + y^2)^2}.$$



Mais elle n'est pas de classe  $C^2$  en  $(0,0)$  car

$$D_2 D_1 f(0,0) = \lim_{t \rightarrow 0} \frac{D_1 f(0,t) - D_1 f(0,0)}{t} = -1 ;$$

$$D_1 D_2 f(0,0) = \lim_{t \rightarrow 0} \frac{D_2 f(t,0) - D_2 f(0,0)}{t} = 1. \diamond$$

**Proposition A2.16** (*formule de Taylor-Young*)

Pour toute fonction  $f : U \rightarrow \mathbb{R}^p$  de classe  $C^2$ , on a, au voisinage de  $a \in U$  :

$$f(a+h) = f(a) + \sum_{i=1}^n h_i D_i f(a) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n h_i h_j D_i D_j f(a) + o(\|h\|^2).$$

*Preuve* : Pour simplifier les notations, nous écrirons la démonstration pour  $n = 2$ .  
Considérons une boule  $B(a,r)$  incluse dans  $U$  et un vecteur  $h$  tel que  $\|h\| < r$ .  
Introduisons alors la fonction d'une variable réelle  $t$  définie sur  $[0,1]$  par :

$$\varphi_h(t) = f(a+th) = f(a_1+th_1, a_2+th_2).$$

Le calcul de dérivation donne

$$\varphi'_h(t) = h_1 D_1 f(a_1+th_1, a_2+th_2) + h_2 D_2 f(a_1+th_1, a_2+th_2)$$

$$\varphi''_h(t) = h_1^2 D_1^2 f(a+th) + 2h_1 h_2 D_1 D_2 f(a+th) + h_2^2 D_2^2 f(a+th).$$

Appliquons alors à  $\varphi_h$  la formule de Taylor avec reste intégral sur  $[0,1]$  :

$$\begin{aligned} \varphi_h(1) &= \varphi_h(0) + \varphi'_h(0) + \int_0^1 (1-t) \varphi''_h(t) dt \\ &= \varphi_h(0) + \varphi'_h(0) + \frac{1}{2} \varphi''_h(0) + \int_0^1 (1-t) (\varphi''_h(t) - \varphi''_h(0)) dt \end{aligned}$$

On établit maintenant que ce dernier terme est négligeable devant  $\|h\|^2$  en  $(0,0)$ .

Comme  $f$  est de classe  $C^2$  sur  $U$ , ses dérivées secondes sont continues sur  $U$ , donc au point  $a$ , et pour tout réel  $\varepsilon > 0$ , il existe un réel  $\alpha > 0$  tel que :

$$\text{pour tout } i, j \in \{1, 2\}, \text{ pour tout } x \in U, \text{ si } \|x\| \leq \alpha, \text{ alors } |D_i D_j f(x) - D_i D_j f(a)| \leq \frac{\varepsilon}{4}.$$

On en déduit la majoration suivante de  $\varphi''_h(t) - \varphi''_h(0)$  pour  $\|h\| \leq \alpha$  :

$$\text{pour tout } t \in [0, 1], |\varphi''_h(t) - \varphi''_h(0)| \leq h_1^2 \frac{\varepsilon}{4} + 2|h_1| |h_2| \frac{\varepsilon}{4} + h_2^2 \frac{\varepsilon}{4} \leq \varepsilon \|h\|^2.$$

Par intégration sur  $[0,1]$ , on obtient donc pour  $\|h\| \leq \alpha$  :

$$\left| \int_0^1 (1-t) (\varphi''_h(t) - \varphi''_h(0)) dt \right| \leq \int_0^1 |\varphi''_h(t) - \varphi''_h(0)| dt \leq \varepsilon \|h\|^2.$$

Cette intégrale est donc négligeable devant  $\|h\|^2$  lorsque  $h$  tend vers  $(0,0)$ .

En reportant ces expressions de  $\varphi_h(0)$ ,  $\varphi_h(1)$ ,  $\varphi'_h(0)$ ,  $\varphi''_h(0)$  dans la formule de Taylor avec reste intégral écrite pour  $\varphi$ , on obtient la formule de Taylor-Young pour  $f$  :

$$\begin{aligned} f(a+h) &= f(a) + h_1 D_1 f(a) + h_2 D_2 f(a) + \\ &\quad + \frac{1}{2} (h_1^2 D_1^2 f(a) + 2h_1 h_2 D_1 D_2 f(a) + h_2^2 D_2^2 f(a)) + \|h\|^2 \varepsilon(\|h\|) \end{aligned}$$

C'est bien la formule cherchée.

□

### Interprétation matricielle

Introduisons les matrices suivantes :

- $Jf_a = (D_1 f(a), \dots, D_n f(a))$  matrice jacobienne de  $f$  en  $a$  (c'est une matrice ligne)
- $Hf_a = (D_i D_j f(a))$  matrice hessienne de  $f$  en  $a$  (c'est une matrice symétrique réelle)
- $h$  qu'on identifie à la matrice colonne de ses composantes.

La formule précédente s'écrit alors comme suit :

$$f(a+h) = f(a) + (Jf_a)h + \frac{1}{2} h(Hf_a)h + o(\|h\|^2).$$

Lorsque  $n = 2$ , cette formule explicite  $f(a+h)$  comme suit :

$$f(a) + (D_1 f(a), D_2 f(a)) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + \frac{1}{2} (h_1, h_2) \begin{pmatrix} D_1^2 f(a) & D_1 D_2 f(a) \\ D_2 D_1 f(a) & D_2^2 f(a) \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + o(\|h\|^2).$$

On remarquera que ce développement limité donne  $f(a+h)$  comme la somme :

- d'un terme de degré 0 :  $f(a)$
- des termes de degré 1 :  $h_1 D_1 f(a) + h_2 D_2 f(a)$
- des termes de degré 2 :  $\frac{1}{2} (h_1^2 D_1^2 f(a) + 2h_1 h_2 D_1 D_2 f(a) + h_2^2 D_2^2 f(a))$
- du reste de young :  $o(\|h\|^2) = \|h\|^2 \varepsilon(h)$  avec  $\lim_{h \rightarrow 0} \varepsilon(h) = 0$ .