# 24

# Calcul différentiel et optimisation : cas de $\mathbb{R}^n$

Dans tout le chapitre, $E = \mathbb{R}^n$, $F = \mathbb{R}^m$, $\mathscr{U}$ désigne un ouvert de $E = \mathbb{R}^n$.

## I Dérivées partielles

Dans cette partie, $\mathscr{U}$ est un ouvert non vide de $\mathbb{R}^n$, où $n \in \mathbb{N}^*$.

### 1 Compléments sur la continuité des fonctions de variable vectorielle

**Définition 1 : Application partielle**

Soit $f : \mathscr{U} \to \mathbb{R}$ une fonction, $a = (a_1, \dots, a_n) \in \mathscr{U}$, $j \in [\![1, n]\!]$. On appelle $j^e$ **application partielle de $f$ en** $a$ l'application $t \mapsto f(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n)$.

**Lemme 1 : « de partition »**

*Soit $A$ une partie de $E$, $f : A \to \mathbb{R}^m$, $B_1, B_2$ deux parties de $A$ telles que $B_1 \cup B_2 = A$, $a \in \overline{B_1} \cap \overline{B_2}$, $\ell \in F$. Si $f_{|B_1}(x) \xrightarrow[x \to a]{} \ell$ et $f_{|B_2}(x) \xrightarrow[x \to a]{} \ell$, alors $f(x) \xrightarrow[x \to a]{} \ell$.*

*En particulier, si $a \in A$, $f$ est continue en $a$.*

### 2 Dérivées partielles

$\mathscr{U}$ est toujours un ouvert de $\mathbb{R}^n$.

**Définition 2 : Dérivées partielles**

Soit $f : \mathscr{U} \to \mathbb{R}$, $a = (a_1, \dots, a_n) \in \mathscr{U}$, $j \in [\![1, n]\!]$. On appelle $j^e$ **dérivée partielle de $f$ en** $a$, lorsqu'elle existe, la dérivée de la $j^e$ application partielle de $f$ en $a$. On note $\partial_j f(a) = \dfrac{\partial f}{\partial x_j}(a)$ le nombre dérivé en ce point.

On appelle $j^e$ **dérivée partielle** la fonction définie sur $\mathscr{U}$ par $a \mapsto \dfrac{\partial f}{\partial x_j}(a)$.

**Définition 3 : Vecteur gradient**

Soit $f : \mathscr{U} \to \mathbb{R}$ admettant des dérivées partielles en $a \in \mathscr{U}$. On appelle **gradient** de $f$ en $a$ le vecteur

$$\nabla f(a) = \overrightarrow{\operatorname{grad}} f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)\right) \in \mathbb{R}^n.$$

**Définition 4 : Extension aux fonctions à valeurs vectorielles**

Soit $\mathscr{U}$ ouvert de $\mathbb{R}^n$ et $f : \mathscr{U} \to \mathbb{R}^m$. On note $f = (f_1, \dots, f_m)$ : $f_i$ est la $i^e$ composante de $f$.

On dit que $f$ admet des dérivées partielles en $a \in \mathscr{U}$ si chacune des $f_i$ admet une dérivée partielle en $a$.

On appelle alors $j^e$ **dérivée partielle de $f$ en** $a$ le vecteur

$$\frac{\partial f}{\partial x_j}(a) = \left(\frac{\partial f_1}{\partial x_j}(a), \dots, \frac{\partial f_m}{\partial x_j}(a)\right).$$

**Définition 5 : Matrice jacobienne**

Soit $\mathscr{U}$ ouvert de $\mathbb{R}^n$ et $f : \mathscr{U} \to \mathbb{R}^m$. On note $f = (f_1, \dots, f_m)$. On appelle, lorsque existe, **matrice jacobienne de $f$ en** $a$ la matrice

$$J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} \in \mathscr{M}_{m,n}(\mathbb{R})$$

### 3 Fonctions de classe $\mathscr{C}^1$

**Définition 6 : Classe $\mathscr{C}^1$**

Soit $f : \mathscr{U} \to \mathbb{R}^m$. On dit que $f$ est de classe $\mathscr{C}^1$ sur $\mathscr{U}$ lorsqu'en tout point de $\mathscr{U}$, les dérivées partielles de $f$ existent, et que ces dérivées partielles sont continues.

On note $\mathscr{C}^1(\mathscr{U}, \mathbb{R}^m)$ l'ensemble de ces fonctions.

**Propriété 1 : Équivalence avec les fonctions coordonnées**

*On note $f = (f_1, \dots, f_m)$. Alors $f$ est de classe $\mathscr{C}^1$ si et seulement si chacune des $f_i$ l'est.*

**Propriété 2 : Structure d'algèbre**

*Pour tout $m \in \mathbb{N}$, $\mathscr{C}^1(\mathscr{U}, \mathbb{R}^m)$ est un $\mathbb{R}$-espace vectoriel et $\mathscr{C}^1(\mathscr{U}, \mathbb{R})$ est une $\mathbb{R}$-algèbre.*

**Théorème 1 : $\mathrm{DL}_1$**

*Soit $f : \mathscr{U} \to \mathbb{R}$ de classe $\mathscr{C}^1$ et $a \in \mathscr{U}$. Pour tout $h \in \mathbb{R}^n$ tel que $a + h \in \mathscr{U}$,*

$$\begin{aligned} f(a+h) &= f(a) + h_1 \frac{\partial f}{\partial x_1}(a) + \cdots + h_n \frac{\partial f}{\partial x_n}(a) + \underset{h \to 0}{\mathrm{o}}(\|h\|) \\ &= f(a) + \left(\nabla f(a) \middle| h\right) + \underset{h \to 0}{\mathrm{o}}(\|h\|). \end{aligned}$$

*en utilisant le produit scalaire canonique sur $\mathbb{R}^n$.*

**Corollaire 1 : $\mathscr{C}^1 \Longrightarrow$ continue**

*Une fonction de classe $\mathscr{C}^1$ est continue.*

**Propriété 3 : Règle de la chaîne**

*Soient $\mathscr{U}, \mathscr{V}$ deux ouverts respectifs de $\mathbb{R}^n$ et $\mathbb{R}^m$ et $a \in \mathscr{U}$. Soient*

$$f : \left| \begin{array}{ccc} \mathscr{U} & \longrightarrow & \mathscr{V} \\ (x_1,\dots,x_n) & \longmapsto & (f_1(x_1,\dots,x_n),\dots,f_m(x_1,\dots,x_n)) \end{array} \right.$$

*et*

$$g : \left| \begin{array}{ccc} \mathscr{V} & \longrightarrow & \mathbb{R} \\ (y_1,\dots,y_m) & \longmapsto & g(y_1,\dots,y_m) \end{array} \right.$$

*deux fonctions de classe $\mathscr{C}^1$. Alors $g \circ f = g \circ (f_1,\dots,f_m)$ est de classe $\mathscr{C}^1$ et pour $j \in [\![1,n]\!]$,*

$$\frac{\partial (g\circ f)}{\partial x_j}(a) = \sum_{k=1}^{m} \frac{\partial g}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a)$$

**Cas particulier 1 : important – dérivée le long d'un arc**

*Si $\gamma : \mathbb{R} \to \mathbb{R}^3$ et $f : \mathbb{R}^3 \to \mathbb{R}$ sont de classe $\mathscr{C}^1$, en notant, pour $t \in \mathbb{R}$, $\gamma(t) = (x(t), y(t), z(t))$,*

$$(f\circ\gamma)'(t) = x'(t)\frac{\partial f}{\partial x}(\gamma(t)) + y'(t)\frac{\partial f}{\partial y}(\gamma(t)) + z'(t)\frac{\partial f}{\partial z}(\gamma(t)).$$

**Corollaire 2 : Règle de la chaîne vectorielle**

*Soient $\mathscr{U}, \mathscr{V}$ deux ouverts respectifs de $\mathbb{R}^n$ et $\mathbb{R}^m$ et $a \in \mathscr{U}$. Soient $f : \mathscr{U} \to \mathscr{V}$ et $g : \mathscr{V} \to \mathbb{R}^p$ deux fonctions de classe $\mathscr{C}^1$. Alors $g \circ f$ est de classe $\mathscr{C}^1$ et pour $j \in [\![1,n]\!]$ et $\ell \in [\![1,p]\!]$,*

$$\frac{\partial (g_\ell \circ f)}{\partial x_j}(a) = \frac{\partial (g\circ f)_\ell}{\partial x_j}(a) = \sum_{k=1}^{m} \frac{\partial g_\ell}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a)$$

*En particulier,*

$$J_{g\circ f}(a) = J_g(f(a)) J_f(a).$$

## 4 Fonctions de classe $\mathscr{C}^k$

**Définition 7 : Dérivées partielles d'ordre supérieur**

Soit $\mathscr{U}$ ouvert de $\mathbb{R}^n$, $f : \mathscr{U} \to \mathbb{R}^m$. On appelle **dérivée partielle d'ordre** $k \in \mathbb{N}^*$, une dérivée $\dfrac{\partial \varphi}{\partial x_j}$ où $\varphi$ est une dérivée partielle d'ordre $k-1$ de $f$. Les dérivées partielles d'ordre $k$ sont de la forme

$$\frac{\partial}{\partial x_{i_k}}\left(\frac{\partial}{\partial x_{i_{k-1}}}\left(\cdots \frac{\partial}{\partial x_{i_2}}\left(\frac{\partial f}{\partial x_{i_1}}\right)\cdots\right)\right)$$

notées $\dfrac{\partial^k f}{\partial x_{i_k}\cdots\partial x_{i_1}} = \partial_{i_k,\dots,i_1} f$ où $i_1,\dots,i_k \in [\![1,n]\!]$.

**Définition 8 : Classe $\mathscr{C}^k$**

Soit $k \in \mathbb{N}^*$, $\mathscr{U}$ ouvert de $\mathbb{R}^n$, $f : \mathscr{U} \to \mathbb{R}^m$. On dit que $f$ **est de classe** $\mathscr{C}^k$ si toutes ses dérivées partielles d'ordre $k$ existent et sont continues.

On dit que $f$ **est de classe** $\mathscr{C}^\infty$ si elle est de classe $\mathscr{C}^k$ pour tout $k \in \mathbb{N}$.

**Propriété 4 : Caractérisation par les applications coordonnées**

*On note $f = (f_1,\dots,f_m)$, $k \in \mathbb{N} \cup \{+\infty\}$. Alors $f$ est de classe $\mathscr{C}^k$ si et seulement si chacune des $f_i$ l'est.*

**Théorème 2 : de Schwarz**

*Soit $\mathscr{U}$ ouvert de $\mathbb{R}^n$, $f : \mathscr{U} \to \mathbb{R}^m$ de classe $\mathscr{C}^2$. Alors, pour tout $i, j \in [\![1,n]\!]$ tel que $i \neq j$,*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

**Propriété 5 : Opérations sur les fonctions de classe $\mathscr{C}^k$**

*Soit $k \in \mathbb{N} \cup \{+\infty\}$.*

(i) *Toute combinaison linéaire, toute composée, tout produit, tout quotient de fonctions de classe $\mathscr{C}^k$ l'est encore.*

(ii) *Si $M : \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_p} \to \mathbb{R}^m$ est $p$-linéaire, $f_1 : \mathscr{U} \to \mathbb{R}^{n_1}, \dots, f_p : \mathscr{U} \to \mathbb{R}^{n_p}$ de classe $\mathscr{C}^k$ alors $M(f_1,\dots,f_p)$ est de classe $\mathscr{C}^k$.*

(iii) *Toute fonction polynomiale à $n$ variables est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}^n$.*

(iv) *Toute fonction rationnelle (quotient de fonctions polynomiales) est de classe $\mathscr{C}^\infty$ sur son domaine de définition.*

## 5 Matrice hessienne et $DL_2$

**Définition 9 : Matrice hessienne**

Soit $f : \mathscr{U} \to \mathbb{R}$ où $\mathscr{U}$ est un ouvert de $\mathbb{R}^n$. On appelle, lorsqu'elle existe, **matrice hessienne** de $f$ en $x \in \mathscr{U}$ la matrice

$$H_f(x) = \left(\partial_{i,j}\right)_{1\leqslant i,j\leqslant n} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{1\leqslant i,j\leqslant n} \in \mathscr{M}_n(\mathbb{R}).$$

**Théorème 3 : $DL_2$ : formule de Taylor-Young à l'ordre 2**

*Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^n$, $f : \mathscr{U} \to \mathbb{R}$ de classe $\mathscr{C}^2$, $a \in \mathscr{U}$ et $h \in \mathbb{R}^n$ tel que $a+h$ reste dans $\mathscr{U}$ lorsque $h \to 0$.*

*En confondant $\mathbb{R}^n$ et $\mathscr{M}_{n,1}(\mathbb{R})$, et en utilisant le produit scalaire canonique sur $\mathbb{R}^n$, on a*

$$f(a+h) = f(a) + \left(\nabla f(a) \middle| h\right) + \frac{1}{2}\left(H_f(a)h \middle| h\right) + \underset{h\to 0}{o}\left(\|h\|^2\right)$$

$$f(a+h) = f(a) + \nabla f(a)^{\mathsf{T}} h + \frac{1}{2} h^{\mathsf{T}} H_f(a) h + \underset{h\to 0}{o}\left(\|h\|^2\right)$$

# II Applications

## 1 Équations aux dérivées partielles

**Méthode 1**

- Savoir résoudre les ÉDP fondamentales auxquelles on se ramène systématiquement :

$$\frac{\partial f}{\partial x}(x,y) = 0 \qquad \frac{\partial f}{\partial x}(x,y) = g(x) \qquad \frac{\partial f}{\partial x}(x,y) = g(y)$$

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 0 \qquad \frac{\partial^2 f}{\partial x \partial y}(x,y) = 0$$

- Dans la pratique, on s'y ramène via un changement de variables $(u,v) = \varphi(x,y)$, en écrivant $f(x,y) = g(u,v)$ et en remplaçant soit $(x,y)$ en fonction de $(u,v)$, soit $(u,v)$ en fonction de $(x,y)$.
- La changement de variable doit être bijectif, entre deux ouverts et suffisamment régulier (classe $\mathscr{C}^1$ ou $\mathscr{C}^2$ suivant l'ordre de l'équation.)
- S'il n'est pas donné, il doit être affine ou polaire.
- Appliquer la règle de la chaîne (ou utiliser des matrices jacobiennes) pour exprimer les dérivées de $f$ en fonction de celle de $g$ ou l'inverse, et simplifier l'ÉDP.

## 2 Optimisation : recherche d'extremums

Ici, toutes les fonctions sont à valeurs dans $\mathbb{R}$.

On notera plutôt $n$ la dimension au départ : $\mathscr{U}$ désigne un ouvert de $\mathbb{R}^n$.

### a Extremums libres

**Définition 10 : Extremum**

Soit $A$ une partie de $\mathbb{R}^n$, $a \in A$, $f : A \to \mathbb{R}$.

(i) On dit que $f$ présente en $a$ un **maximum** (respectivement **minimum**) local s'il existe $\mathscr{V}$ voisinage de $a$ dans $A$ tel que pour tout $x \in \mathscr{V}$, $f(x) \leqslant f(a)$ (respectivement $f(x) \geqslant f(a)$).

Il est **strict** lorsque, $\forall x \in \mathscr{V} \setminus \{a\}$, , $f(x) < f(a)$ (respectivement $f(x) > f(a)$).

(ii) On dit que $f$ présente un **maximum** (respectivement **minimum**) global si cette inégalité est en fait valable pour tout $x \in A$.

**Définition 11 : Point critique**

Soit $f : \mathscr{U} \to \mathbb{R}$ admettant des dérivées partielles en $a \in \mathscr{U}$.

Lorsque $\nabla f(a) = 0$, c'est-à-dire $\forall i \in [\![1,n]\!]$, $\dfrac{\partial f}{\partial x_i}(a) = 0$, on dit que $a$ est un **point critique** de $f$.

**Propriété 6 : Condition nécessaire d'extremum local à l'ordre 1**

*On suppose que*

**H1** *$\mathscr{U}$ **ouvert** (très important !) de $\mathbb{R}^n$.*

**H2** *$a \in \mathscr{U}$ tel que $f : \mathscr{U} \to \mathbb{R}$ admettant des dérivées partielles en $a$.*

**H3** *$f$ présente un extremum local en $a$.*

*Alors*

**C1** *$a$ est un point critique de $f$.*

*La réciproque est fausse, et un contre-exemple est appelé **point selle** ou **point col**.*

**Propriété 7 : Condition nécessaire de minimum local à l'ordre 2**

*On suppose que*

**H1** *$\mathscr{U}$ **ouvert** (très important !) de $\mathbb{R}^n$*

**H2** *$a \in \mathscr{U}$ et $f \in \mathscr{C}^2(\mathscr{U}, \mathbb{R})$.*

**H3** *$f$ présente un minimum local en $a$*

*alors*

**C1** *$a$ est un point critique de $f$.*

**C2** *$H_f(a) \in \mathscr{S}_n^+(\mathbb{R})$.*

*Si c'est un maximum local en $a$, alors $-H_f(a) \in \mathscr{S}_n^+(\mathbb{R})$ (ie $H_f(a)$ est symétrique « négative », ses valeurs propres sont toutes dans $\mathbb{R}^-$).*

**Propriété 8 : Condition suffisante d'extremum local à l'ordre 2**

*On suppose que*

**H1** *$\mathscr{U}$ ouvert (très important !) de $\mathbb{R}^n$.*

**H2** *$a \in \mathscr{U}$ et $f \in \mathscr{C}^2(\mathscr{U}, \mathbb{R})$.*

**H3** *$a$ est un point critique de $f$.*

**H4** *$H_f(a) \in \mathscr{S}_n^{++}(\mathbb{R})$*
*(respectivement $-H_f(a) \in \mathscr{S}_n^{++}(\mathbb{R})$)*

*alors*

**C1** *$f$ atteint un minimum (respectivement maximum) local **strict** en $a$.*

**Corollaire 3 : Discussion sur le spectre**

*Si $\mathscr{U}$ **ouvert**, $f \in \mathscr{C}^2(\mathscr{U},\mathbb{R})$ et $a \in \mathscr{U}$ point critique de $f$.*

- *Si $\mathrm{Sp}\left(H_f(a)\right) \subset \mathbb{R}_+^*$, $f$ atteint en $a$ un minimum local.*
- *Si $\mathrm{Sp}\left(H_f(a)\right) \subset \mathbb{R}_-^*$, $f$ atteint en $a$ un maximum local.*
- *Si $H_f(a)$ possède des valeurs propres non nulles de signes opposés, $a$ est un point selle.*

Si $n=2$ (fonction de 2 variables), on note

$$H_f(a) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

où $r = \dfrac{\partial^2 f}{\partial x^2}(a)$, $s = \dfrac{\partial^2 f}{\partial x \partial y}(a) = \dfrac{\partial^2 f}{\partial y \partial x}(a)$ (théorème de Schwarz), $t = \dfrac{\partial^2 f}{\partial y^2}(a)$.

On suppose que $a$ est un point critique. On a $\det\left(H_f(a)\right) = rt - s^2$ (égal produit de ses deux valeurs propres réelles) et $\mathrm{tr}\left(H_f(a)\right) = r+t$ (égal à leur somme).

1. Si $rt - s^2 > 0$, les deux valeurs propres de $H_f(a)$ ont même signe et sont non nulles : **$f$ présente en $a$ un extremum local strict.**
   (a) Si $r+t > 0$ alors $H_f(a) \in \mathscr{S}_n^{++}(\mathbb{R})$ et $f$ présente un **minimum local strict** en $a$.
   (b) Si $r+t < 0$ alors $-H_f(a) \in \mathscr{S}_n^{++}(\mathbb{R})$ et $f$ présente un **maximum local strict** en $a$.
2. Si $rt - s^2 < 0$, les deux valeurs propres de $H_f(a)$ ont des signes opposés et sont non nulles.
   Dans ce cas, **$f$ présente en $a$ un point col** : dans les directions propres, on a respectivement un minimum et un maximum local.
3. Si $rt - s^2 = 0$, on ne peut rien conclure en général.

**Méthode 2 : Recherche d'extremum**

(i) Soit $f : \mathscr{U} \to \mathbb{R}$ admettant des dérivées partielles sur $\mathscr{U}$. Pour déterminer des extremums locaux de $f$ on cherche ses points critiques.
Si $f$ est de classe $\mathscr{C}^2$, on peut s'intéresser à la Hessienne aux points critiques qui permet de conclure lorsqu'elle est inversible (voir la remarque précédente pour $n=2$.)
Sinon, pour un point critique $a$, on étudie $f(a+h) - f(a)$ pour $h$ proche de $0$. Comme $a$ est un point critique, les termes obtenus dans la différence sont au moins d'ordre 2.

(ii) Si $\mathscr{U}$ n'est pas un ouvert, on le décompose en son intérieur (ouvert) et son bord. Sur le bord, on étudie « à la main », en général à l'aide d'un paramétrage du bord.

(iii) Si $f : K \to \mathbb{R}$ où $K$ est un compact, on est assuré de l'existence d'un minimum et d'un maximum globaux. On les cherche comme dans la méthode précédente.

## b Extremums liés

**Théorème 4 : d'optimisation sous contrainte**

*Soient $f, g \in \mathscr{C}^1(\mathscr{U},\mathbb{R})$ où $\mathscr{U}$ est un ouvert de $\mathbb{R}^n$, et $X = \{x \in \mathscr{U},\ g(x) = 0\}$.*
*Si $f_{|X}$ admet un extremum local en $a \in X$ et si $\nabla g(a) \neq 0_{\mathbb{R}^n}$, alors $\nabla f(a)$ est colinéaire à $\nabla g(a)$.*