# Résumé 18 – Calcul différentiel

## Applications de classe $\mathscr{C}^1$

Soit $f : \mathscr{U} \subset E \to F$ où $E$ et $F$ désignent deux e.v.n. sur $\mathbb{R}$ de dimensions respectives $p$ et $n$ et $\mathscr{U}$ un ouvert de $E$.

### → Différentielle

**Définition : Différentielle en un point**

L'application $f$ est dite différentiable en $a \in \mathscr{U}$ s'il existe $\varphi \in \mathscr{L}(E,F)$ tel que :

$$f(a+h) \underset{h\to 0}{=} f(a)+\varphi(h)+o(h)$$

L'application est alors unique, on l'appelle *différentielle* de $f$ au point $a$. On la note $\mathrm{d}f_a$ ou $\mathrm{d}f(a)$.

Notation : $o(h) = \|h\|\varepsilon(h)$ où $\varepsilon : E \to F$ et $\varepsilon(h) \xrightarrow[h\to 0_E]{} 0_F$.

$$f(a+h) \underset{h\to 0}{=} f(a)+\mathrm{d}f_a(h)+o(h)$$

**Proposition**

Si $f$ est différentiable en $a$, $f$ est continue en $a$.

- Si $f$ et $g$ sont différentiables en $a$, $\lambda g+\mu g$ aussi et :

$$\mathrm{d}(\lambda f+\mu g)_a = \lambda \mathrm{d}f_a + \mu \mathrm{d}g_a$$

- Si $f$ est différentiable en $a$ et $g$ en $f(a)$, alors $g\circ f$ est différentiable en $a$ et :

$$\mathrm{d}(g\circ f)_a = \mathrm{d}g_{f(a)} \circ \mathrm{d}f_a$$

**Définition : Différentielle, application de classe $\mathscr{C}^1$**

- $f$ est dite différentiable sur $\mathscr{U}$ si $f$ est différentiable en tout point de $\mathscr{U}$. On appelle alors différentielle de $f$ l'application :

$$\mathrm{d}f : \left| \begin{array}{rcl} \mathscr{U} \subset E & \longrightarrow & \mathscr{L}(E,F) \\ a & \longmapsto & \mathrm{d}f(a) = \mathrm{d}f_a \end{array} \right.$$

- L'application $f : \mathscr{U} \subset E \to F$ est dite de classe $\mathscr{C}^1$ sur $\mathscr{U}$ si $f$ est différentiable sur $\mathscr{U}$ et si sa différentielle $\mathrm{d}f$ est continue sur $\mathscr{U}$.

Si $f$ est de classe $\mathscr{C}^1$ sur $\mathscr{U}$, on dira aussi que $f$ est continûment différentiable sur $\mathscr{U}$.

### → Dérivée selon un vecteur et dérivées partielles

**Définition : Dérivée selon un vecteur**

Soit $u \in E$. L'application $f$ est dite dérivable en $a$ selon le vecteur $u$ si la fonction $t \mapsto f(a+tu)$ est dérivable en 0. On pose dans ce cas :

$$D_u(f)(a) = \lim_{t\to 0} \frac{f(a+tu)-f(a)}{t}$$

Quand une fonction est différentiable, elle est dérivable dans toutes les directions.

**Proposition**

Si $f$ est différentiable en $a$ alors $f$ est dérivable en $a$ selon $u$ pour tout vecteur $u \in E$ et $D_u(f)(a) = \mathrm{d}f_a(u)$.

On munit désormais $E$ d'une b.o.n. $(e_1,\dots,e_p)$.

**Définition : Dérivées partielles**

Pour $j \in [\![1,p]\!]$, on appelle dérivée partielle en $a$ d'indice $j$ la dérivée de $f$ en $a$ suivant $e_j$, c'est-à-dire :

$$\frac{\partial f}{\partial x_j}(a) = \lim_{t\to 0} \frac{f(a+te_j)-f(a)}{t}$$

Si $f$ est différentiable en $a$, alors les dérivées partielles existent et :

$$\forall j \in [\![1,p]\!], \quad \frac{\partial f}{\partial x_j}(a) = \mathrm{d}f_a(e_j)$$

$$\mathrm{d}f_a(h) = \mathrm{d}f_a\left(\sum_{j=1}^p h_j e_j\right) = \sum_{j=1}^p h_j \mathrm{d}f_a(e_j) = \sum_{j=1}^p h_j \frac{\partial f}{\partial x_j}(a).$$

En notant $\mathrm{d}x_j$ les applications $h \mapsto h_j$,

$$\mathrm{d}f_a = \frac{\partial f}{\partial x_1}(a)\mathrm{d}x_1 + \cdots + \frac{\partial f}{\partial x_p}(a)\mathrm{d}x_p = \sum_{j=1}^p \frac{\partial f}{\partial x_j}(a)\mathrm{d}x_j$$

**Théorème : Caractérisation**

Soit $f : \mathscr{U} \subset E \to F$. $f$ est de classe $\mathscr{C}^1$ sur $\mathscr{U}$ si et seulement si les dérivées partielles de $f$ existent et sont continues en tout point de $\mathscr{U}$.

Si $f$ est de classe $\mathscr{C}^1$ sur $\mathscr{U}$ et $a \in \mathscr{U}$, alors :

$$f(a+h) \underset{h\to 0}{=} f(a) + \sum_{j=1}^p h_j \frac{\partial f}{\partial x_j}(a) + o(h)$$

Pour calculer la différentielle en un point, on peut revenir à la définition ou bien calculer les dérivées partielles.

### → Jacobienne

On munit désormais $F$ d'une b.o.n. $(e'_1,\dots,e'_n)$. On appelle jacobienne de $f$ au point $x=(x_1,\dots,x_p)$ la matrice représentative de $\mathrm{d}f_x$ dans les bases $(e_j)_{1\leqslant j\leqslant p}$ et $(e'_i)_{1\leqslant i\leqslant n}$ :

$$J_f(x) = \left(\frac{\partial f_i}{\partial x_j}(x)\right)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_p}(x) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(x) & \dots & \frac{\partial f_n}{\partial x_p}(x) \end{bmatrix}$$

Si $f$ est de classe $\mathscr{C}^1$ sur $\mathscr{U}$, pour tout $x \in \mathscr{U}$,

$$J_{g\circ f}(x) = J_g(f(x)) \times J_f(x) \quad \text{car} \quad \mathrm{d}(g\circ f)_x = \mathrm{d}g_{f(x)} \circ \mathrm{d}f_x$$

- Soient $\mathscr{U}$ un ouvert de $\mathbb{R}^2$ et $I$ un intervalle de $\mathbb{R}$. On considère les deux applications de classe $\mathscr{C}^1$ :

$$\varphi : \left| \begin{array}{ccc} I & \longrightarrow & \mathscr{U} \\ t & \longmapsto & (x(t), y(t)) \end{array} \right. \quad \text{et} \quad f : \left| \begin{array}{ccc} \mathscr{U} & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(x,y) \end{array} \right.$$

L'application $t \mapsto f(x(t), y(t))$ est de classe $\mathscr{C}^1$ sur $I$ et pour tout $t \in I$,

$$(f \circ \varphi)'(t) = x'(t)\frac{\partial f}{\partial x}(x(t), y(t)) + y'(t)\frac{\partial f}{\partial y}(x(t), y(t))$$

- On considère les applications $f : \mathbb{R}^2 \to \mathbb{R}$, $\varphi : \mathbb{R}^2 \to \mathbb{R}$ et $\psi : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et :

$$F : \left| \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(\varphi(x,y), \psi(x,y)) \end{array} \right.$$

Alors $F$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et pour tout $(x, y) \in \mathbb{R}^2$,

$$\begin{aligned} \frac{\partial F}{\partial x}(x,y) &= \frac{\partial \varphi}{\partial x}(x,y) \cdot \frac{\partial f}{\partial x}(\varphi(x,y), \psi(x,y)) \\ &\quad + \frac{\partial \psi}{\partial x}(x,y) \cdot \frac{\partial f}{\partial y}(\varphi(x,y), \psi(x,y)) \\ \frac{\partial F}{\partial y}(x,y) &= \frac{\partial \varphi}{\partial y}(x,y) \cdot \frac{\partial f}{\partial x}(\varphi(x,y), \psi(x,y)) \\ &\quad + \frac{\partial \psi}{\partial y}(x,y) \cdot \frac{\partial f}{\partial y}(\varphi(x,y), \psi(x,y)) \end{aligned}$$

### → Gradient associé à une fonction numérique

Soit $f : \mathscr{U} \subset E \to \mathbb{R}$ une fonction **numérique**, supposée différentiable en $a$. On peut alors définir le gradient de $f$ au point $a$ par ses coordonnées dans une b.o.n. $(e_1, \ldots, e_p)$ :

$$\nabla f(a) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_p}(a) \end{bmatrix}$$

Le théorème de Riesz fournit une définition intrinsèque :

**Définition : Gradient**

On appelle gradient de $f$ en $a$ et on note $\nabla f(a)$ le vecteur associé à la forme linéaire $\mathrm{d}f_a$. Pour tout $h \in E$,

$$\mathrm{d}f_a(h) = \nabla f(a)^\top h = \nabla f(a) \cdot h$$

Si $f : \mathscr{U} \subset E \to \mathbb{R}$ est de classe $\mathscr{C}^1$,

$$\forall x \in \mathscr{U}, \quad f(x+h) \underset{h \to 0}{=} f(x) + \nabla f(x)^\top h + \mathrm{o}(\|h\|)$$

### → Dérivées le long d'un arc et vecteurs tangents

**Proposition : Dérivation le long d'un arc**

Si $f : \mathscr{U} \subset E \to F$ et $\gamma : I \subset \mathbb{R} \to E$ sont de classe $\mathscr{C}^1$, alors $f \circ \gamma$ est de classe $\mathscr{C}^1$ sur $I$ et :

$$\forall t \in I, \quad (f \circ \gamma)'(t) = \mathrm{d}f_{\gamma(t)}(\gamma'(t))$$

**Proposition : Intégration le long d'un arc**

Si $f : \mathscr{U} \subset E \to F$ et $\gamma : [0,1] \subset \mathbb{R} \to E$ sont de classe $\mathscr{C}^1$ et si $\gamma(0) = a$ et $\gamma(1) = b$, alors :

$$f(b) - f(a) = \int_0^1 \mathrm{d}f_{\gamma(t)}(\gamma'(t))\,\mathrm{d}t$$

Si $\mathscr{U}$ est un ouvert connexe par arcs et $f : \mathscr{U} \subset E \to F$, $f$ est constante sur $\mathscr{U}$ ssi pour tout $a \in \mathscr{U}$, $\mathrm{d}f_a = 0_{\mathscr{L}(E,F)}$.

**Définition : Vecteur tangent**

Si $X$ est une partie de $E$ et $x \in X$, un vecteur $v$ de $E$ est tangent à $X$ en $x$ s'il existe $\varepsilon > 0$ et un arc $\gamma$ défini sur $]-\varepsilon, \varepsilon[$ dérivable en 0 à valeurs dans $X$, tels que $\gamma(0) = x$ et $\gamma'(0) = v$.

On note $T_x X$ l'ensemble des vecteurs tangents à $X$ en $x$.

**Définition : Hyperplan tangent**

Soient $g$ est une fonction numérique de classe $\mathscr{C}^1$ sur l'ouvert $\mathscr{U}$, $X$ l'ensemble des zéros de $g$, $x \in X$. Si $\mathrm{d}g_x$ est non nulle, $T_x X = \mathrm{Ker}(\mathrm{d}g_x) = \nabla f(x)^\perp$.

## Applications de classe $\mathscr{C}^k$

### → Dérivées partielles d'ordres supérieurs

On définit par récurrence les dérivées partielles d'ordres supérieurs :

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

**Définition : Application de classe $\mathscr{C}^k$**

Une application est dite de classe $\mathscr{C}^k$ sur un ouvert $\mathscr{U}$ si toutes ses dérivées partielles d'ordre $k$ existent et sont continues.

**Théorème : Théorème de Schwarz**

Soit $f : \mathscr{U} \subset E \to F$ une application de classe $\mathscr{C}^2$ sur un ouvert $\mathscr{U}$ de $\mathbb{R}^2$. Alors,

$$\forall a \in \mathscr{U}, \quad \frac{\partial^2 f}{\partial x \partial y}(a) = \frac{\partial^2 f}{\partial y \partial x}(a)$$

### → Hessienne

La hessienne au point $a \in \mathscr{U}$ de la fonction numérique $f : \mathscr{U} \subset \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^2$ est la matrice symétrique :

$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial {x_1}^2}(a) & \ldots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \ldots & \dfrac{\partial^2 f}{\partial {x_n}^2}(a) \end{bmatrix}$$

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

$$\forall x \in \mathscr{U}, \; f(x+h) \underset{h \to 0}{=} f(x) + \nabla f(x)^\top h + \frac{1}{2} h^\top H_f(x) h + \mathrm{o}(\|h\|^2)$$

## Optimisation

Toutes les fonctions considérées sont des fonctions numériques.

### → Condition d'ordre 1

**Définition : Point critique**

Soit $f : \mathscr{U} \subset \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur l'ouvert $\mathscr{U}$.
On dit que $a \in \mathscr{U}$ est un point critique de $f$ si :

$$\mathrm{d}f_a = 0_{\mathscr{L}(\mathbb{R}^p,\mathbb{R})} \text{ c'est-à-dire } \nabla f(a) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a) \end{bmatrix} = \vec{0}$$

$f$ admet un maximum en $a \in \mathbb{R}^p$ si et seulement s'il existe un voisinage $\mathscr{U}$ de $a$ tel que :

$$\forall x \in \mathscr{U}, \quad f(x) \leqslant f(a)$$

La définition est analogue pour un minimum.

**Théorème : CN d'existence d'un extremum**

Si $f : \mathscr{U} \subset \mathbb{R}^p \to \mathbb{R}$ de classe $\mathscr{C}^1$ sur l'ouvert $\mathscr{U}$ admet un extremum au point $a \in \mathscr{U}$ alors $a$ est un point critique. Cela revient à dire que $\nabla f(a) = \vec{0}$.

La propriété est fausse ailleurs que sur un ouvert. De plus, tout point critique ne correspond pas nécessairement à un extremum (cas des points selles).

### → Condition d'ordre 2

**Théorème : CS d'existence d'un extremum**

Soit $f : \mathscr{U} \subset \mathbb{R}^n \to \mathbb{R}$ de classe $\mathscr{C}^2$ sur l'ouvert $\mathscr{U}$.
Si $a$ est un point critique de $f$ et $H_f(a) \in \mathscr{S}_n^{++}(\mathbb{R})$, alors $f$ atteint un minimum local strict en $a$.

Pour $n = 2$, le déterminant et la trace nous permettent de trouver facilement le signe des valeurs propres.

### → Optimisation sous contrainte

**Théorème : Multiplicateur de Lagrange**

Soient $f$ et $g$ deux fonctions numériques de classe $\mathscr{C}^1$ sur l'**ouvert** $\mathscr{U}$ de $E$ et $X = \{x \in \mathscr{U} \mid g(x) = 0\}$.
Si la restriction de $f$ à $X$ admet un extremum local en $a \in X$ et $\mathrm{d}g_a \neq 0$, alors $\mathrm{d}f_a$ est colinéaire à $\mathrm{d}g_a$.

L'existence de $\lambda \in \mathbb{R}^*$ tel que $\nabla f(a) = \lambda \nabla g(a)$ en un point critique $a$ est une simple condition nécessaire. On mène ensuite une étude locale ou on donne un argument de compacité pour justifier la présence d'un extremum.