COURS COMPLET
CALCUL DIFFÉRENTIEL ET OPTIMISATION

Différentiabilité, Gradient, Hessienne et Extrema

Table des matières

1 Dérivées Partielles et Directionnelles

Soit $f : U \subset E \to F$ définie sur un ouvert U d'un e.v.n de dimension finie E.

1.1 Dérivée selon un vecteur

DÉFINITION : Dérivée Directionnelle

Soit $a \in U$ et $v \in E$. La **dérivée de f en a selon le vecteur** v est, si elle existe :

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

C'est la dérivée en 0 de la fonction d'une variable $\phi(t) = f(a+tv)$.

1.2 Dérivées Partielles

Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E, on note $f(x) = f(x_1, \dots, x_n)$.

DÉFINITION : Dérivée Partielle

La i-ème dérivée partielle est la dérivée selon le vecteur e_i de la base :

$$\frac{\partial f}{\partial x_i}(a) = D_{e_i} f(a) = \lim_{t \to 0} \frac{f(a_1, \dots, a_i + t, \dots, a_n) - f(a)}{t}$$

2 Différentiabilité

2.1 Définition et Différentielle

DÉFINITION : Différentiabilité (DL1)

f est **différentiable** en a s'il existe une application linéaire $L \in \mathcal{L}(E, F)$ telle que :

$$f(a+h) = f(a) + L(h) + o(\|h\|) \quad \text{quand } h \to 0$$

L est unique, on l'appelle la **différentielle** de f en a, notée $df(a)$.

THÉORÈME : Propriétés

Si f est différentiable en a, alors :

1. f est continue en a.
2. f admet des dérivées selon tout vecteur v, et $D_v f(a) = df(a) \cdot v$.
3. En particulier, $df(a) \cdot e_i = \frac{\partial f}{\partial x_i}(a)$.

Attention : L'existence des dérivées partielles n'implique PAS la différentiabilité (ni même la continuité).

2.2 Matrice Jacobienne et Gradient

DÉFINITION : Matrice Jacobienne

La matrice de l'application linéaire $df(a)$ dans les bases canoniques est la **Matrice Jacobienne** $J_f(a) \in \mathcal{M}_{p,n}(\mathbb{R})$:

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

DÉFINITION : Gradient (Cas Euclidien)

Si E est euclidien et f à valeurs réelles ($p = 1$), il existe un unique vecteur $\nabla f(a) \in E$ tel que :

$$\forall h \in E, \quad df(a) \cdot h = \langle \nabla f(a), h \rangle$$

Dans une base orthonormée : $\nabla f(a) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix}$.

2.3 Opérations et Règle de la Chaîne

THÉORÈME : Différentielle d'une composée

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

$$d(g \circ f)(a) = dg(f(a)) \circ df(a)$$

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

3 Fonctions de Classe $\mathcal{C}^1$

THÉORÈME : Caractérisation

f est de classe $\mathcal{C}^1$ sur U ssi ses dérivées partielles $\frac{\partial f}{\partial x_i}$ existent et sont **continues** sur U. Dans ce cas, f est différentiable.

PROPOSITION : Théorème fondamental (sur un arc)

Si f est $\mathcal{C}^1$ et $\gamma : [0, 1] \to U$ est un arc $\mathcal{C}^1$ reliant a à b :

$$f(b) - f(a) = \int_0^1 df(\gamma(t)) \cdot \gamma'(t)\, dt$$

4 Dérivées d'ordre supérieur

DÉFINITION : Classe $\mathcal{C}^k$

f est $\mathcal{C}^k$ si toutes ses dérivées partielles d'ordre k existent et sont continues.

THÉORÈME : Théorème de Schwarz

Si f est de classe $\mathcal{C}^2$, l'ordre de dérivation n'importe pas :

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

5 Géométrie et Optimisation

5.1 Vecteurs Tangents

Soit X une partie de E. Un vecteur v est tangent à X en x s'il est la vitesse en 0 d'une courbe tracée sur X passant par x. **Exemple :** Si $X = \{x \in E \mid g(x) = 0\}$ (Ligne de niveau), l'espace tangent est $\ker dg(x)$.

5.2 Optimisation Sans Contrainte (Ordre 1)

THÉORÈME : Condition Nécessaire (Point Critique)

Si f admet un extremum local en a (et U ouvert), alors a est un **point critique** :

$$df(a) = 0_{\mathcal{L}(E,F)} \quad (\text{ou } \nabla f(a) = 0)$$

5.3 Optimisation Sous Contrainte (Lagrange)

THÉORÈME : Théorème des Multiplicateurs de Lagrange

Soit f la fonction à optimiser et $g(x) = 0$ la contrainte (f, g de classe $\mathcal{C}^1$). Si f admet un extremum en a sous la contrainte $g(x) = 0$, et si $dg(a) \neq 0$, alors :

$$\exists \lambda \in \mathbb{R}, \quad df(a) = \lambda dg(a)$$

Les gradients $\nabla f(a)$ et $\nabla g(a)$ sont colinéaires.

5.4 Optimisation : Étude au Second Ordre

Si a est un point critique ($df(a) = 0$) et f est $\mathcal{C}^2$:

DÉFINITION : Matrice Hessienne

La Hessienne $H_f(a)$ est la matrice symétrique des dérivées secondes :

$$H_f(a) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right]_{1 \leq i,j \leq n}$$

THÉORÈME : Taylor-Young ordre 2

$$f(a+h) = f(a) + \underbrace{df(a) \cdot h}_{0} + \frac{1}{2}\,{}^t h H_f(a) h + o(\|h\|^2)$$

PROPOSITION : Nature du point critique

On regarde les valeurs propres de la matrice symétrique réelle $H_f(a)$:

- **Toutes** > 0 **:** $H_f(a)$ définie positive $\Longrightarrow$ **Minimum Local Strict**.
- **Toutes** < 0 **:** $H_f(a)$ définie négative $\Longrightarrow$ **Maximum Local Strict**.
- **Valeurs propres de signes opposés : Point Col** (pas d'extremum).