Questions de cours

- ☐ Comme d'habitude, savoir citer toute définition, toute proposition du cours.
- ☐ Savoir citer la définition de la différentielle, de la dérivée en un point selon un vecteur, des dérivées partielles, du gradient, de la matrice jacobienne et de la matrice hessienne.
- ☐ Soit $f : \Omega \to F$ une application différentiable en $a \in \Omega$. Montrer qu'il existe une unique application linéaire $u : E \to F$ telle que, pour tout h au voisinage de 0, $f(a+h) = f(a) + u(h) + o(h)$.
- ☐ Soit $f : \Omega \to F$ une fonction de classe $\mathcal{C}^1$. Soit L une application linéaire de F dans G. Justifier que $L \circ f$ est de classe $\mathcal{C}^1$ et calculer sa différentielle *(de deux façons différentes)*.
- ☐ Déterminer la différentielle de f dans les cas particuliers suivantes : f est constante, f est linéaire, f est une fonction d'une variable réelle.
- ☐ On définit sur $\mathbb{R}^2$ la fonction f par :

$$f(x,y) = \begin{cases} \dfrac{x^4+y^4}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon.} \end{cases}$$

  a) Montrer que f est différentiable en $(0,0)$ et calculer sa différentielle.
  b) Montrer que f est de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

- ☐ Soit $f : \Omega \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ sur un ouvert Ω de $\mathbb{R}^n$. Montrer que si f admet un minimum local en $a \in \Omega$, alors a est un point critique de f et $H_f(a) \in S_n^+(\mathbb{R})$.
- ☐ Soit $f : \Omega \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ sur un ouvert Ω de $\mathbb{R}^n$. Montrer que si f admet un point critique en a et si $H_f(a) \in S_n^{++}(\mathbb{R})$, alors f admet un minimum local strict en a.

Chapitre 25 : Calcul différentiel

Dans ce chapitre, on étudie des fonctions $f : \Omega \to F$, où E et F sont des espaces vectoriels normés de dimension finie et Ω est un ouvert (non vide) de E.

I Différentielle en un point

I.1 Dérivée selon un vecteur

- définition de la dérivée selon un vecteur. En cas d'existence de la limite : $\mathrm{D}_v f(a) = \lim\limits_{t \to 0} \dfrac{f(a+tv) - f(a)}{t}$.
- dans le cas où f est une fonction d'<u>une</u> variable réelle, lien entre $\mathrm{D}_v f(a)$ et $f'(a)$.
- définition des dérivées partielles de f en a : lorsqu'on fixe une base $(e_1, \ldots, e_n)$ de E, alors $\dfrac{\partial f}{\partial x_i}(a)$ (aussi notée $\partial_i f(a)$) est égale à $\mathrm{D}_{e_i} f(a)$ c'est-à-dire à $\lim\limits_{t \to 0} \dfrac{f(a+te_i) - f(a)}{t}$.

I.2 Différentielle en un point

- notations o et O.
- on dit que f est différentiable en a s'il existe $u \in \mathcal{L}(E,F)$ telle que $f(a+h) = f(a) + u(h) + o(h)$ lorsque h tend vers 0_E. Cela revient à écrire $f(a+h) = f(a) + u(h) + \|h\|\varepsilon(h)$, où ε est une fonction qui tend vers 0_F en 0_E.
- si f est différentiable, l'application linéaire u définie ci-dessus est unique. Elle est appelée différentielle de f en a et notée $\mathrm{d}f(a)$. Ainsi, lorsque h tend vers 0_E, $f(a+h) = f(a) + \mathrm{d}f(a) \cdot h + o(h)$.
- si f est différentiable en a, alors f est continue en a.
- f est différentiable en a si et seulement si toutes ses applications coordonnées f_k sont différentiables en a.
- dans le cas où f est une fonction d'<u>une</u> variable réelle, f est différentiable en a si et seulement si f est dérivable en a. Dans ce cas, pour tout $h \in \mathbb{R}$, $\mathrm{d}f(a) \cdot h = hf'(a)$.
- une fonction constante est différentiable en tout point a et $\mathrm{d}f(a) : h \mapsto 0$.
- si $f : E \to F$ est linéaire, alors f est différentiable en tout point a et $\mathrm{d}f(a) : h \mapsto f(h)$.

I.3 Différentielle et dérivée partielle

- si f est différentiable en a, alors pour tout vecteur v, $\mathrm{D}_v f(a) = \mathrm{d}f(a) \cdot v$.
- si f est différentiable en a, alors toutes les dérivées partielles de f en a existent et $\mathrm{d}f(a) : h \mapsto \sum\limits_{k=1}^{n} h_k \dfrac{\partial f}{\partial x_k}(a)$.
- définition de la matrice jacobienne : $J_f(a) = \left(\dfrac{\partial f_i}{\partial x_j}(a)\right)$.

I.4 Gradient

- dans le cas où E est un espace euclidien et où f est à valeurs dans $\mathbb{R}$, le gradient de f en a est l'unique vecteur $\nabla f(a) \in E$ tel que, pour tout $h \in E$, $\mathrm{d}f(a) \cdot h = \langle \nabla f(a), h \rangle$.
- si $\nabla f(a) \neq 0_E$, alors $\nabla f(a)$ est positivement colinéaire au vecteur unitaire v selon lequel la valeur de $\mathrm{D}_v f(a)$ est maximale.
- dans une base orthonormée $(e_1, \dots, e_n)$, $\nabla f(a) = \sum\limits_{k=1}^{n} \dfrac{\partial f}{\partial x_k}(a) e_k$.

II Fonctions différentiables

II.1 Définitions et opérations usuelles

- définition d'une fonction différentiable sur un ouvert Ω. La différentielle $\mathrm{d}f$ est une application : $\mathrm{d}f : \begin{array}{ccc} \Omega & \to & \mathcal{L}(E,F) \\ a & \mapsto & \mathrm{d}f(a) \end{array}$.
- si f et g sont différentiables, alors $\lambda f + \mu g$ aussi et $\mathrm{d}(\lambda f + \mu g)(a) = \lambda \mathrm{d}f(a) + \mu \mathrm{d}g(a)$.
- si f et g sont différentiables et si B est bilinéaire, alors $B(f,g)$ est différentiable et $\mathrm{d}B(f,g)(a) \cdot h = B(\mathrm{d}f(a) \cdot h, g(a)) + B(f(a), \mathrm{d}g(a) \cdot h)$.
- en particulier, si f et g sont différentiables et à valeurs dans $\mathbb{R}$, alors fg est différentiable et $\mathrm{d}(fg)(a) \cdot h = g(a)\ \mathrm{d}f(a) \cdot h + f(a)\ \mathrm{d}g(a) \cdot h$.
- si $f_1, \dots, f_p$ sont différentiables et si M est multilinéaire, alors $M(f_1, \dots, f_p)$ est différentiable en a et $\mathrm{d}M(f_1, \dots, f_p)(a) \cdot h = \sum\limits_{k=1}^{p} M(f_1(a), \dots, \mathrm{d}f_k(a) \cdot h, \dots, f_p(a))$.

II.2 Différentielle d'une composée

- si f est différentiable en a et si g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et $\mathrm{d}(g \circ f)(a) = \mathrm{d}g(f(a)) \circ \mathrm{d}f(a)$.
- règle de la chaîne : si f est différentiable en a et si g est différentiable en $f(a)$, alors $\dfrac{\partial(g \circ f)}{\partial x_k}(a) = \sum\limits_{i=1}^{n} \dfrac{\partial f_i}{\partial x_k}(a)\dfrac{\partial g}{\partial y_i}(f(a))$.
- dérivation selon un arc : si $\gamma : I \subset \mathbb{R} \to \Omega$ est dérivable en t_0 et si f est différentiable en $\gamma(t_0)$, alors $f \circ \gamma$ est dérivable en t_0 et $(f \circ \gamma)'(t_0) = \mathrm{d}f(\gamma(t_0)) \cdot \gamma'(t_0) = \sum\limits_{k=1}^{n} \dfrac{\partial f}{\partial x_k}(\gamma(t_0))\gamma_k'(t_0)$.
- en tout point, le gradient est orthogonal aux lignes de niveau.

II.3 Fonctions de classe $\mathcal{C}^1$

- $f : \Omega \to F$ est de classe $\mathcal{C}^1$ sur Ω si $\mathrm{d}f : \begin{array}{ccc} \Omega & \to & \mathcal{L}(E,F) \\ a & \mapsto & \mathrm{d}f(a) \end{array}$ est continue sur Ω.
- stabilité par combinaison linéaire, par produit, par composition, ...
- une fonction f est de classe $\mathcal{C}^1$ si et seulement si toutes les dérivées partielles de f existent en tout point de Ω et sont continues sur Ω.
- si f est de classe $\mathcal{C}^1$ sur Ω et $\gamma : [0,1] \to \Omega$ est de classe $\mathcal{C}^1$, alors $f(\gamma(1)) - f(\gamma(0)) = \displaystyle\int_0^1 \mathrm{d}f(\gamma(t)) \cdot \gamma'(t)\mathrm{d}t$.
- si Ω est connexe par arcs, alors f est constante si et seulement si f est différentiable de différentielle nulle.

III Fonctions de classe $\mathcal{C}^k$

III.1 Dérivées partielles d'ordres supérieurs

- si $\dfrac{\partial f}{\partial x_i}$ est différentiable, alors sa dérivée partielle par rapport à la j-ème variable est notée $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ ou $\partial_j \partial_i f$ ou $\partial_{i,j} f$: c'est une dérivée partielle d'ordre 2 de f.
- les dérivées partielles d'ordre k de f, si elles existent, sont notées $\dfrac{\partial^k f}{\partial x_{j_k} \dots \partial x_{j_1}}$ ou $\partial_{j_k} \dots \partial_{j_1} f$ ou $\partial_{j_1,\dots,j_k} f$
- une fonction f est de classe $\mathcal{C}^k$ si et seulement si toutes les dérivées partielles d'ordre k de f existent en tout point de Ω et sont continues sur Ω.
- f est de classe $\mathcal{C}^\infty$ si f est de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$.
- stabilité par combinaison linéaire, par produit, par composition, ... Les fonctions polynomiales sont de classe $\mathcal{C}^\infty$.

III.2 Théorème de Schwarz

- si f est de classe $\mathcal{C}^2$, alors pour tout $(i,j) \in [\![1,n]\!]^2$, $\dfrac{\partial^2 f}{\partial x_j \partial x_i} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}$.

III.3 Matrice hessienne

- si $f : \Omega \to \mathbb{R}$ est de classe $\mathcal{C}^2$, on définit la matrice hessienne de f en a par : $H_f(a) = \left(\dfrac{\partial^2 f}{\partial x_j \partial x_i}(a)\right)$.
  D'après le théorème de Schwarz, la matrice hessienne est symétrique réelle.
- formule de Taylor-Young à l'ordre 2 : pour h au voisinage de 0,
  $f(a+h) = f(a) + \mathrm{d}f(a) \cdot h + \dfrac{1}{2} h^T H_f(a) h + o(\|h\|^2)$.
  En coordonnées, cela s'écrit : $f(a+h) = f(a) + \sum\limits_{k=1}^{n} \dfrac{\partial f}{\partial x_k}(a) h_k + \dfrac{1}{2} \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} \dfrac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + o(\|h\|^2)$.

IV Optimisation

IV.1 Extrema locaux et globaux

- définition d'un maximum global, d'un maximum local, d'un maximum local strict. Idem pour minimum et extremum.
- étude du signe de $f(a+h) - f(a)$ pour h au voisinage de 0.

IV.2 Étude au premier ordre

- si $f : A \to \mathbb{R}$ et si a est un point intérieur de A, on dit que f admet un point critique en a si $\mathrm{d}f(a) = 0$. Cela revient à $\nabla f(a) = 0$ ou à : $\forall k \in [\![1, n]\!]$, $\dfrac{\partial f}{\partial x_k}(a) = 0$.
- supposons f différentiable en un point intérieur a de A. Si f admet un extremum local en a, alors $\mathrm{d}f(a) = 0$.
- conséquence : les extrema locaux de f sont à rechercher parmi les points critiques de f et parmi les points de $A \setminus \overset{\circ}{A}$.

IV.3 Étude au second ordre

- si f est de classe $\mathcal{C}^2$ et admet un minimum local (resp. maximum local) en un point intérieur a de A, alors a est un point critique de f et $H_f(a) \in S_n^+(\mathbb{R})$ (resp $-H_f(a) \in S_n^+(\mathbb{R})$).
- si f est de classe $\mathcal{C}^2$ sur un ouvert, si a est un point critique de f et si $H_f(a) \in S_n^{++}(\mathbb{R})$ (resp. $-H_f(a) \in S_n^{++}(\mathbb{R})$), alors f admet un minimum local strict (resp. maximum local strict) en a.
- étude complète du cas de $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$, selon les signes du déterminant et de la trace de la matrice hessienne en chaque point critique.

V Optimisation sous une contrainte

V.1 Espace tangent

- si X est une partie de E et $a \in X$, on dit que v est un vecteur tangent à X en a s'il existe $\varepsilon > 0$ et $\gamma :]-\varepsilon, \varepsilon[\to E$ une application à valeurs dans X, dérivable en 0 et telle que $\gamma(0) = a$ et $\gamma'(0) = v$.
- l'ensemble des vecteurs tangents en a à X est l'espace tangent, noté $T_a X$.
- si $X = \{x \in \Omega, g(x) = 0\}$, où g est de classe $\mathcal{C}^1$, et si $\mathrm{d}g(a) \neq 0$, alors $T_a X = \mathrm{Ker}(\mathrm{d}g(a))$. Cela revient à : $T_a X = \{\nabla g(a)\}^\perp$.

V.2 Optimisation sous une contrainte

- si f est différentiable sur un ouvert Ω et si la restriction $f_{|X}$ admet un extremum local en $a \in X$, alors $T_a X \subset \operatorname{Ker}(\mathrm{d}f(a))$.
- théorème d'optimisation sous une contrainte : si f est de classe $\mathcal{C}^1$ sur un ouvert Ω, si $X = \{x \in \Omega, g(x) = 0\}$ où g est de classe $\mathcal{C}^1$, si $\mathrm{d}g(a) \neq 0$ et si la restriction $f_{|X}$ admet un extremum local en $a \in X$, alors il existe $\lambda \in \mathbb{R}$ tel que $\mathrm{d}f(a) = \lambda \mathrm{d}g(a)$. Cela revient à $\nabla f(a) = \lambda \nabla g(a)$.