INU Champollion

Licence Info 3ème Année

Maths pour l'I.A.

Thierry Montaut

Plan du cours

- Compléments d'algèbre linéaire
- **Compléments d'analyse**
  - Fonctions de plusieurs variables
  - **Dérivées partielles, gradient**
  - **formules de Taylor et plan tangent**
  - Dérivées partielles d'une composée
  - Optimisation des fonctions de plusieurs variables
  - Méthodes numériques d'optimisation
- Compléments de probabilité et de statistiques

Dérivées partielles

Soit $f : \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}$ et $a = (a_1, ..., a_n) \in \mathcal{D}$.

Définition

$\forall i \in \llbracket 1, n \rrbracket$ *on définit la* **ième application partielle de** f *en* a *par :*

$$f_i : \left(\begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ t & \mapsto & f(a_1, ..a_{i-1}, t, a_{i+1}..a_n) \end{array} \right.$$

Définition

On dit que f admet une $i$ème dérivée partielle en a (ou une dérivée partielle par rapport à x_i en a) ssi f_i est dérivable en a_i, et on note alors :

$$\frac{\partial f}{\partial x_i}(a) = f_i'(a_i).$$

Un calcul de dérivée partielle se ramène donc à un calcul de dérivée, donc dans les cas les plus délicats à un calcul de limite du taux d'accroissement.

Le calcul des dérivées partielles relevant de la dérivabilité des fonctions réelles, à vous de prévoir les révisions qui s'imposent...

Exercice 1 : Etudier l'existence de dérivées partielles en tout point de $\mathbb{R}^2$ pour les fonctions suivantes :

1. $f(x,y) = x^2 + xy - y^3$.
2. $f(x,y) = xy\cos(x+y)$
3. $f(x,y) = \ln(1 + x^2y^2)$
4. $f(x,y) = \sqrt{x^2 + y^2}$

Opérations et dérivées partielles

Les dérivées partielles héritent des propriétés des dérivées des fonctions réelles :

Propriété

(Opérations et dérivées partielles)

Si f, et g admettent des dérivées partielles en a par rapport à x_i alors

- *$f + g$ aussi et*

$$\frac{\partial(f+g)}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a) + \frac{\partial g}{\partial x_i}(a).$$

- *$f.g$ aussi et*

$$\frac{\partial f.g}{\partial x_i}(a) = f(a).\frac{\partial g}{\partial x_i}(a) + \frac{\partial f}{\partial x_i}(a).g(a).$$

Propriété

- *Si* $g(a) \neq 0$ *alors* $\dfrac{1}{g}$ *admet une dérivée partielle en a et*

$$\frac{\partial \frac{1}{g}}{\partial x_i}(a) = -\frac{1}{g(a)^2}\frac{\partial g}{\partial x_i}(a).$$

- *On déduit des deux dernières propriétés que :*

Si $g(a) \neq 0$ *alors* $\dfrac{f}{g}$ *admet une dérivée partielle en a et*

$$\frac{\partial \frac{f}{g}}{\partial x_i}(a) = \frac{1}{g(a)^2}\left(\frac{\partial f}{\partial x_i}(a).g(a) - f(a).\frac{\partial g}{\partial x_i}(a).\right).$$

Propriété

- *Si* $g(a) \neq 0$ *alors* $\frac{1}{g}$ *admet une dérivée partielle en a et*

$$\frac{\partial \frac{1}{g}}{\partial x_i}(a) = -\frac{1}{g(a)^2}\frac{\partial g}{\partial x_i}(a).$$

- *On déduit des deux dernières propriétés que :*

Si $g(a) \neq 0$ *alors* $\frac{f}{g}$ *admet une dérivée partielle en a et*

$$\frac{\partial \frac{f}{g}}{\partial x_i}(a) = \frac{1}{g(a)^2}\left(\frac{\partial f}{\partial x_i}(a).g(a) - f(a).\frac{\partial g}{\partial x_i}(a).\right).$$

dérivées partielles d'une composée (1)

Théorème

Si f : $\mathbb{R}^n \to \mathbb{R}$ *admet une dérivée partielle en a par rapport à* x_i *et que h est une fonction réelle dérivable en f(a), alors* $h \circ f : \mathbb{R}^n \to \mathbb{R}$ *admet une dérivée partielle en a par rapport à* x_i *et*

$$\frac{\partial h \circ f}{\partial x_i}(a) = h'(f(a)) \frac{\partial f}{\partial x_i}(a).$$

Calculer les dérivées partielles de $g(x, y) = e^{2x+y^2}$.

dérivées partielles d'une composée (2)

Théorème

Si $f : \mathbb{R}^n \to \mathbb{R}$ *admet une dérivée partielle en a par rapport à* x_i *et que* $u_1(t), \dots u_n(t)$ *sont n fonctions réelles dérivables en* t_0 *et telles que* $a = (u_1(t_0), \dots u_n(t_0))$ *alors* $g : \mathbb{R} \to \mathbb{R}$ *définie par* $g(t) = f((u_1(t), \dots u_n(t))$*est dérivable en* t_0 *et*

$$g'(t_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \times u_i'(t_0).$$

dérivées partielles d'ordre supérieur

Définition

soit $f : \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}$,
On définit les dérivées partielles secondes de f, *notées*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial \left(\frac{\partial f}{\partial x_j} \right)}{\partial x_i},$$

comme les dérivées partielles (si elles existent) des dérivées partielles de f.
On définit alors par récurrence les dérivées partielles d'ordre p *de* f, *notées* $\frac{\partial^p f}{\partial x_{i_1} \partial x_{i_2} .. \partial x_{i_p}}$, *comme les dérivées partielles des dérivées partielles d'ordre* $p-1$ *de* f.

Classe $\mathcal{C}^k$

Définition

On dit que f est de **classe** $\mathcal{C}^1$ *sur $\mathcal{D}$ ssi f admet sur $\mathcal{D}$ des dérivées partielles par rapport à toutes ses variables et si elles sont toutes continues sur $\mathcal{D}$.*
On note alors $f \in \mathcal{C}^1(\mathcal{D})$.

Définition

On dit que f est de **classe** $\mathcal{C}^k$ *sur $\mathcal{D}$ ssi f admet sur $\mathcal{D}$ toutes ses dérivées partielles d'ordre k et si elles sont toutes continues sur $\mathcal{D}$.*
On note alors $f \in \mathcal{C}^k(\mathcal{D})$.

Définition

On dit enfin que f est de **classe** $\mathcal{C}^\infty$ *sur $\mathcal{D}$ ssi f est de* **classe** $\mathcal{C}^k$ *pour tout $k \in \mathbb{N}$. On note alors $f \in \mathcal{C}^\infty(\mathcal{D})$.*

Théorème

(de Schwarz)
Si f est de classe $\mathcal{C}^k$, les dérivées partielles d'ordre k de f sont indépendantes de l'ordre de dérivation :
Ce qui pour $n = 2$ et $p = 2$ s'écrit simplement : Si f est de classe $\mathcal{C}^2$, alors

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

Proposition

1. *Toute application polynomiale est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^n$.*
2. *Toute fraction rationnelle est de classe $\mathcal{C}^\infty$ sur son ensemble de définition.*

Attention aux liens entre ces notions qui différent de ceux des fonctions réelles :

Proposition

1. *Si f est de classe $\mathcal{C}^1$ alors toute ses dérivées partielles existent et f est continue.*
2. *Mais f peut admettre toutes ses dérivées partielles sans pour autant être continue (cf TD).*
3. *Comme dans $\mathbb{R}$ un fonction peut être continue sans admettre toutes ses dérivées partielles.*

Vecteur gradient

Définition

Si f admet toutes ses dérivées partielles en a on appelle vecteur gradient de f au point a le vecteur de $\mathbb{R}^n$:

$$grad_a(f) = \left(\frac{\partial f}{\partial x_1}(a), .., \frac{\partial f}{\partial x_n}(a) \right).$$

On le note également à la physicienne $\bigtriangledown f(a)$.
Si f admet toutes ses dérivées partielles sur tout $\mathcal{D}$, on peut considérer la fonction :

$$\bigtriangledown f : \left(\begin{array}{ccc} \mathcal{D} \subset \mathbb{R}^n & \rightarrow & \mathbb{R}^n \\ a & \mapsto & \bigtriangledown f(a) \end{array} \right.$$

Formule de Taylor à l'ordre 1

La formule de Taylor à l'ordre 1 permet d'obtenir une approximation locale de la variation de f entre deux points voisins :

Théorème

(Développement de Taylor à l'ordre 1) Soit f de classe $\mathcal{C}^1$ en a. Il existe une application $\varepsilon : \mathbb{R} \to \mathbb{R}$ tq

$$\forall x \in \mathbb{R}^n, \quad f(x) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a).(x_i - a_i) + \|(x_i - a_i)\| \varepsilon(x_i - a_i)$$

avec

$$\varepsilon(h) \underset{h \to 0}{\longrightarrow} 0.$$

Ou en notant $x = a + h$:

$$f(x) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a).h_i + \|h_i\| \varepsilon(h_i)$$

Cette formule permet donc d'estimer la variation de f entre les points a et x, en fonction des variations h_i des composantes de $x - a$
Les physiciens la notent en général :

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a).dx_i = (\bigtriangledown f(a) | dx)$$

en oubliant le reste (quitte à considérer que $(x - a)$ est assez petit pour se le permettre), où df représente la variation de f entre a et x et où dx_i note $x_i - a_i$.

Retour sur le vecteur gradient

De $df = (\bigtriangledown f(a) | dx)$ on déduit que

Propriété

1. *La direction du vecteur gradient est la direction de plus grande variation de f,*
2. *la direction orthogonale au vecteur gradient est la direction suivant laquelle f ne varie pas.*
3. *les lignes de niveau sont en tout point orthogonales au vecteur gradient*

Plan tangent

Une application classique de cette formule est l'étude en géométrie du plan tangent d'une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$:
Si f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ en $a = (a_1, a_2) \in U$, alors on rappelle que le graphe de f,

$$\mathcal{S} = \{(x, y, z) \in \mathbb{R}^3, \ \ (x, y) \in U, \ \ z = f(x, y)\}$$

est la surface de $\mathbb{R}^3$ d'équation cartésienne $z = f(x, y)$.

D'après la formule de Taylor à l'ordre 1,

$$z = f(x,y) = f(a) + \frac{\partial f}{\partial x}(a).(x - a_1) + \frac{\partial f}{\partial y}(a).(y - a_2) + ||(x,y) - a||\varepsilon((x,y)$$

Définition

On appelle plan tangent à la surface $\mathcal{S}$ d'équation cartésienne $z = f(x,y)$, le plan de $\mathbb{R}^3$ d'équation

$$z = f(a) + \frac{\partial f}{\partial x}(a).(x - a_1) + \frac{\partial f}{\partial y}(a).(y - a_2).$$

Exercice 2 Soit la fonction

$$f \ : \begin{pmatrix} \mathbb{R}^2 & \rightarrow & \mathbb{R} \\ (x,y) & \mapsto & x^2 y \end{pmatrix}$$

Déterminer le plan tangent au graphe de f aux points $A = (0,0)$ et $B = (1,1)$.