RÉSUMÉ DE COURS : FONCTIONS DE PLUSIEURS VARIABLES

Dans ce chapitre, on va s'intéresser aux fonctions de plusieurs variables définies sur une partie de $\mathbb{R}^n$. Dans tout ce qui suit, on supposera l'espace vectoriel $\mathbb{R}^n$ muni du produit scalaire canonique noté $\langle\ ,\ \rangle$ et de la norme euclidienne associée notée $\|\ \|$, lesquels sont définis pour tout $x = (x_1, ..., x_n) \in \mathbb{R}^n$ et tout $y = (y_1, ..., y_n) \in \mathbb{R}^n$ par $\langle x, y\rangle = x_1y_1 + ... + x_ny_n$ et $\|x\| = \sqrt{x_1^2 + ... + x_n^2}$.

1 Eléments de topologie

Dans ce paragraphe, nous allons introduire les notions d'ouvert et de fermé, qui seront les ensembles de base sur lesquels on étudiera les fonctions de plusieurs variables. Commençons par la :

Définition 1.1 *Soient $A, B \in \mathbb{R}^n$. La distance de A à B est le réel $d(A, B) = \left\|\overrightarrow{AB}\right\| = \|B - A\|$.*

A noter que la distance sur $\mathbb{R}^n$ étend la notion de distance bien connue dans le plan ou l'espace. De plus, on voit facilement que $d(A, B) = 0$ si et seulement si $A = B$. En outre, on montre (et nous admettrons) que la distance sur $\mathbb{R}^n$ vérifie *l'inégalité triangulaire*, c'est-à-dire :

$$\boxed{\forall A, B, C \in \mathbb{R}^n, \quad d(A, C) \leq d(A, B) + d(B, C).}$$

Définition 1.2 *Soit A un point de $\mathbb{R}^n$. La boule ouverte de centre A et de rayon $r > 0$ est l'ensemble $B(A, r) = \{M \in \mathbb{R}^n | \ \|A - M\| < r\}$. De même, la boule fermée de centre A et de rayon $r \geq 0$ est l'ensemble $B_f(A, r) = \{M \in \mathbb{R}^n | \ \|A - M\| \leq r\}$.*

En d'autres termes, la boule ouverte (resp. fermée) de centre A et de rayon r est l'ensemble des points M de $\mathbb{R}^n$ dont la distance à A est $< r$ (resp. $\leq r$). En particulier, si $n = 1$, on voit que $B(A, r)$ est l'intervalle $]A - r, A + r[$. De même, si $n = 2$, alors $B(A, r)$ est le disque ouvert de centre A et de rayon r.

Définition 1.3 *Une partie U de $\mathbb{R}^n$ est appelée un ouvert de $\mathbb{R}^n$ si $U = \emptyset$ ou si, pour tout point $A \in U$, il existe un réel $r > 0$ tel que la boule ouverte $B(A, r)$ soit contenue dans U.*

Dans ce cas, on dit aussi que U est une partie ouverte de $\mathbb{R}^n$. A titre d'exemples, l'ensemble vide, l'espace $\mathbb{R}^n$ et toute boule ouverte de $\mathbb{R}^n$ sont des ouverts de $\mathbb{R}^n$, et tout intervalle ouvert est un ouvert de $\mathbb{R}$. De plus, on montre (et nous admettrons) que *tout produit cartésien de n ouverts de $\mathbb{R}$ est un ouvert de $\mathbb{R}^n$. En particulier, toute partie de la forme $]a, b[^n$, avec $-\infty \leq a < b \leq +\infty$, est un ouvert de $\mathbb{R}^n$.*

Théorème 1.4 *Toute réunion quelconque d'ouverts de $\mathbb{R}^n$ est un ouvert de $\mathbb{R}^n$. De plus, toute intersection <u>finie</u> d'ouverts de $\mathbb{R}^n$ est un ouvert de $\mathbb{R}^n$.*

A noter qu'une intersection quelconque d'ouverts n'est pas nécessairement un ouvert. Par exemple, l'ensemble $\{0\} = \bigcap_{n\in\mathbb{N}^*}] - \frac{1}{n}, \frac{1}{n}[$ est bien une intersection dénombrable d'ouverts, mais n'est pas un ouvert, vu qu'il ne contient aucune boule de centre 0 et de rayon $r > 0$.

Définition 1.5 *Une partie U de $\mathbb{R}^n$ est un fermé de $\mathbb{R}^n$ si son complémentaire est un ouvert de $\mathbb{R}^n$.*

Dans ce cas, on dit aussi que U est une partie fermée de $\mathbb{R}^n$. A titre d'exemples, l'ensemble vide, l'espace $\mathbb{R}^n$, toute boule fermée de $\mathbb{R}^n$ sont des fermés de $\mathbb{R}^n$, et tout intervalle fermé est un fermé de $\mathbb{R}$. En particulier, tout singleton de $\mathbb{R}^n$ est un fermé de $\mathbb{R}^n$ (puisqu'un singleton de $\mathbb{R}^n$ est un point, et donc une boule fermée de rayon nul). Par passage au complémentaire à partir du théorème 1.4, on montre le :

Théorème 1.6 *Toute réunion <u>finie</u> de fermés de $\mathbb{R}^n$ est un fermé de $\mathbb{R}^n$. De plus, toute intersection quelconque de fermés de $\mathbb{R}^n$ est un fermé de $\mathbb{R}^n$.*

Définition 1.7 *Une partie U de $\mathbb{R}^n$ est dite bornée si elle est contenue dans une boule (ouverte ou fermée).*

Théorème 1.8 *Une partie U de $\mathbb{R}^n$ est bornée si et seulement si : $\exists R > 0$, $\forall M \in U$, $\|M\| \leq R$.*

Partant de là, on peut montrer que toute réunion *finie* de parties bornées de $\mathbb{R}^n$ est une partie bornée de $\mathbb{R}^n$. A titre d'exemples, toute boule (ouverte ou fermée) de $\mathbb{R}^n$ et tout singleton de $\mathbb{R}^n$ sont des parties bornées de $\mathbb{R}^n$. De plus, tout ensemble de la forme $[a, b]^n$, avec $-\infty < a < b < +\infty$, est une partie bornée de $\mathbb{R}^n$. En général, pour montrer qu'une partie U de $\mathbb{R}^n$ est bornée, on cherche à déterminer une constante $R > 0$ telle que : $\forall M \in U$, $\|M\| \leq R$, ce qui revient à montrer que la norme $\|A\|$, lorsque A parcourt U, est majorée par une constante indépendante de A. A contrario, pour montrer qu'une partie U de $\mathbb{R}^n$ n'est pas bornée, on pourra construire une suite de points $(M_k)_{k\geq 0}$ de U dont la norme tend vers $+\infty$.

2 Généralités sur les fonctions de plusieurs variables

2.1 Graphe et ligne de niveau d'une fonction de plusieurs variables

Définition 2.1 *Soit U une partie non vide de $\mathbb{R}^n$. Le graphe de l'application $f : U \longrightarrow \mathbb{R}$ est l'ensemble :*

$$\Gamma_f = \{(x_1, ..., x_n, y) \in U \times \mathbb{R} |\ y = f(x_1, ..., x_n)\}.$$

En d'autres termes, le graphe de f n'est ni plus ni moins que la généralisation de la notion de graphe pour une fonction d'une partie de $\mathbb{R}$ dans $\mathbb{R}$. En particulier, si $n = 2$, le graphe de f correspond à une surface de $\mathbb{R}^3$ (tout comme le graphe d'une fonction d'une partie de $\mathbb{R}$ dans $\mathbb{R}$ est une courbe de $\mathbb{R}^2$). A titre d'exemple, pour tout $(a_0, a_1, a_2) \in \mathbb{R}^3$, le graphe de l'application $f : (x_1, x_2) \longmapsto a_0 + a_1x_1 + a_2x_2$ est le plan de $\mathbb{R}^3$ d'équation cartésienne $a_0 + a_1x_1 + a_2x_2 - y = 0$.

Définition 2.2 *Soit U une partie non vide de $\mathbb{R}^n$ et soit $\lambda \in \mathbb{R}$. La ligne de niveau λ d'une application $f : U \longrightarrow \mathbb{R}$ est l'ensemble $f^{-1}(\{\lambda\}) = \{(x_1, ..., x_n) \in U |\ f(x_1, ..., x_n) = \lambda\}$.*

A noter que les lignes de niveau d'une fonction d'une partie U de $\mathbb{R}^2$ (resp. de $\mathbb{R}^3$) dans $\mathbb{R}$ sont en général des courbes (resp. des surfaces), et qu'elle peuvent être éventuellement vides. De plus, de nombreux ensembles géométriques peuvent se représenter comme des lignes de niveau d'une fonction. Par exemple, tout cercle de rayon $r > 0$ centré à l'origine dans $\mathbb{R}^2$ est la ligne de niveau r de l'application $f : (x_1, x_2) \longmapsto \sqrt{x_1^2 + x_2^2}$. De même, toute droite du plan admet une équation cartésienne de la forme $a_0 + a_1x_1 + a_2x_2 = 0$ avec $(a_1, a_2) \neq (0, 0)$, et elle est donc la ligne de niveau 0 de l'application $g : (x_1, x_2) \longmapsto a_0 + a_1x_1 + a_2x_2$.

2.2 Notion de continuité d'une fonction de plusieurs variables

Définition 2.3 *Soit U une partie non vide de $\mathbb{R}^n$. Une application $f : U \longrightarrow \mathbb{R}$ est dite continue en $a \in U$ si :*

$$\forall \varepsilon > 0, \quad \exists \eta > 0, \quad \forall x \in U, \quad \|x - a\| \leq \eta \implies |f(x) - f(a)| \leq \varepsilon.$$

En d'autres termes, dire que la fonction f est continue en a signifie que la valeur $f(x)$ peut être rendue arbitrairement proche de $f(a)$ si l'on impose au point x d'être suffisamment proche de a. *A noter qu'il existe des applications f d'une partie U de $\mathbb{R}^n$ dans $\mathbb{R}$ qui ne sont pas continues en un point donné.*

Définition 2.4 *Soit U une partie non vide de $\mathbb{R}^n$. On dit que l'application f est continue sur U si f est continue en tout point $a \in U$.*

Comme pour les fonctions d'une partie de $\mathbb{R}$ dans $\mathbb{R}$, on montre (et nous admettrons) les :

Théorème 2.5 *Soit U une partie non vide de $\mathbb{R}^n$, soit $\lambda \in \mathbb{R}$ et soient $f, g : U \longrightarrow \mathbb{R}$ deux applications.*

1. *Si f et g sont continues sur U, alors $\lambda f, f + g, fg$ sont continues sur U.*
2. *Si f et g sont continues en U et si g ne s'annule pas sur U, alors f/g est continue sur U.*

Théorème 2.6 *Soit U une partie non vide de $\mathbb{R}^n$, soit I une partie non vide de $\mathbb{R}$ et soient $f : U \longrightarrow \mathbb{R}$ et $g : I \longrightarrow \mathbb{R}$ deux applications continues sur U et sur I respectivement, telles que $f(U) \subset I$. Alors $g \circ f$ est continue sur U.*

En particulier, ces résultats permettent de montrer qu'une fonction est continue sur une partie U de $\mathbb{R}^n$, en utilisant le fait qu'elle s'obtient par sommes, produits, quotients, compositions de fonctions continues.

Théorème 2.7 *Pour tout $i \in \{1, ..., n\}$, l'application $p_i : (x_1, ..., x_n) \longmapsto x_i$ est continue sur $\mathbb{R}^n$.*

Définition 2.8 *Une application $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est dite polynomiale si elle est combinaison linéaire d'un nombre fini d'applications de la forme $(x_1, ..., x_n) \longmapsto x_1^{i_1} ... x_n^{i_n}$, où $(i_1, ..., i_n) \in \mathbb{N}^n$.*

A titre d'exemples, les fonctions $f : (x_1, x_2) \longmapsto x_1^2x_2 + x_2^2x_1 + x_1$ et $g : (x_1, x_2, x_3) \longmapsto x_1x_2x_3$ sont polynomiales (resp. sur $\mathbb{R}^2$ et $\mathbb{R}^3$). De façon générale, comme toute fonction polynomiale sur $\mathbb{R}^n$ s'obtient à partir des fonctions p_i par une succession de sommes et de produits, le théorème 3.11 entraine le :

Corollaire 2.9 *Toute fonction polynomiale de $\mathbb{R}^n$ dans $\mathbb{R}$ est continue sur $\mathbb{R}^n$.*

De même, comme la norme sur $\mathbb{R}^n$ est la composée de la fonction polynomiale $(x_1, ..., x_n) \longmapsto x_1^2 + ... + x_n^2$ et de la racine carrée (qui est continue sur $\mathbb{R}_+$), les théorèmes 2.6 et 2.9 impliquent le :

Corollaire 2.10 *La norme sur $\mathbb{R}^n$, c'est-à-dire l'application $x \longmapsto \|x\|$, est continue sur $\mathbb{R}^n$.*

A noter que *toute fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, qui est continue sur $\mathbb{R}^n$, est aussi continue sur toute partie non vide U de $\mathbb{R}^n$*. En outre, on montre (et nous admettrons) le :

Théorème 2.11 *Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une application continue sur $\mathbb{R}^n$, et soit $a \in \mathbb{R}$. Alors l'ensemble $\{x \in \mathbb{R}^n | \ f(x) < a\}$ est un ouvert de $\mathbb{R}^n$. De même, l'ensemble $\{x \in \mathbb{R}^n | \ f(x) \leq a\}$ est un fermé de $\mathbb{R}^n$.*

Ce résultat permet dans la majorité des cas de montrer qu'une partie U de $\mathbb{R}^n$ est ouverte. Pour ce faire, il suffit de vérifier que U s'obtient par réunions quelconques et intersections finies d'ensembles de la forme $\{x \in \mathbb{R}^n | \ f(x) < a\}$, où f est une fonction continue de $\mathbb{R}^n$ dans $\mathbb{R}$. En particulier, si $f_1, ..., f_p$ sont des fonctions continues de $\mathbb{R}^n$ dans $\mathbb{R}$ et si $a_1, ..., a_p \in \mathbb{R}$, alors tout ensemble de la forme :

$$U = \{x \in \mathbb{R}^n | \ f_1(x) < a_1, ..., f_p(x) < a_p\}$$

est un ouvert de $\mathbb{R}^n$, vu que $U = \bigcap_{i=1}^p \{x \in \mathbb{R}^n | \ f_i(x) < a_i\}$. De même, si f est continue de $\mathbb{R}^n$ dans $\mathbb{R}$ et si $a \in \mathbb{R}$, alors tout ensemble de la forme :

$$U = \{x \in \mathbb{R}^n | \ f(x) \neq a\}$$

est un ouvert de $\mathbb{R}^n$, vu que $U = \{x \in \mathbb{R}^n | \ f(x) < a\} \bigcup \{x \in \mathbb{R}^n | \ -f(x) < a\}$. Bien entendu, on peut utiliser le même procédé pour montrer qu'une partie de $\mathbb{R}^n$ est fermée. Par exemple, si f est continue de $\mathbb{R}^n$ dans $\mathbb{R}$ et si $a \in \mathbb{R}$, alors l'ensemble :

$$U = f^{-1}(\{a\}) = \{x \in \mathbb{R}^n | \ f(x) = a\}$$

est un fermé de $\mathbb{R}^n$, vu que $U = \{x \in \mathbb{R}^n | \ -f(x) \leq -a\} \bigcap \{x \in \mathbb{R}^n | \ f(x) \leq a\}$. Bien entendu, on doit savoir refaire ce type de raisonnement pour montrer qu'une partie de $\mathbb{R}^n$ est ouverte (resp. fermée).

3 Calcul différentiel à l'ordre 1

3.1 Dérivées partielles - Gradient - Fonctions de classe $\mathcal{C}^1$

Définition 3.1 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $a = (a_1, ..., a_n) \in U$ et soit $f : U \longrightarrow \mathbb{R}$ une application. La i-ème application partielle de f en a est l'application $f_i : x \longmapsto f(a_1, ..., a_{i-1}, x, a_{i+1}, ..., a_n)$.*

Définition 3.2 *Soit U un ouvert de $\mathbb{R}^n$, soit $a = (a_1, ..., a_n) \in U$ et soit $f : U \longrightarrow \mathbb{R}$ une application. Si la i-ème application partielle f_i de f en a est dérivable en a_i, alors la dérivée de f_i en a_i est appelée la dérivée partielle d'ordre 1 de f en a par rapport à x_i, et notée $\partial_i f(a)$ ou $\partial_i(f)(a)$.*

A noter que, sous réserve d'existence, on a :

$$\boxed{\partial_i f(a) = \lim_{h \to 0} \frac{f(a_1, .., a_{i-1}, a_i + h, a_{i+1}, .., a_n) - f(a_1, ..., a_n)}{h}.}$$

Définition 3.3 *Soit U un ouvert non vide de $\mathbb{R}^n$ et soit $f : U \longrightarrow \mathbb{R}$ une application. Si f admet une dérivée partielle d'ordre 1 par rapport à x_i en tout point $a \in U$, alors l'application $\partial_i f : U \longrightarrow \mathbb{R}$, $a \longmapsto \partial_i f(a)$ est appelée la dérivée partielle d'ordre 1 de f par rapport à x_i.*

A noter que, pour calculer la dérivée partielle d'ordre 1 d'une fonction f par rapport à x_i, il suffit de dériver la fonction f comme si elle ne dépendait que de la variable x_i, c'est-à-dire comme si les autres variables $x_1, .., x_{i-1}, x_{i+1}, .., x_n$ étaient des constantes.

Définition 3.4 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $a \in U$ et soit $f : U \longrightarrow \mathbb{R}$ une application. Si f admet une dérivée partielle d'ordre 1 par rapport à x_i en a pour tout $i \in \{1, ..., n\}$, alors le gradient de f en a est défini comme le vecteur :*

$$\nabla f(a) = (\partial_1 f(a), ..., \partial_n f(a)) \, .$$

Définition 3.5 *Soit U un ouvert non vide de $\mathbb{R}^n$ et soit $f : U \longrightarrow \mathbb{R}$ une application. Si f admet une dérivée partielle d'ordre 1 par rapport à x_i en tout point $a \in U$ et pour tout $i \in \{1, ..., n\}$, alors le gradient de f est l'application $\nabla f : U \longrightarrow \mathbb{R}^n$, $a \longmapsto \nabla f(a)$.*

Définition 3.6 *Soit U un ouvert non vide de $\mathbb{R}^n$. Une application $f : U \longrightarrow \mathbb{R}$ est dite de classe $\mathcal{C}^1$ sur U si f admet une dérivée partielle d'ordre 1 par rapport à x_i sur U pour tout $i \in \{1, ..., n\}$, et si toutes ses dérivées partielles d'ordre 1 sont continues sur U.*

En d'autres termes, la notion de fonction de classe $\mathcal{C}^1$ sur un ouvert U de $\mathbb{R}^n$ n'est ni plus ni moins que la généralisation de la même notion pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$. On montre (et nous admettrons) les :

Théorème 3.7 *Toute application f de classe $\mathcal{C}^1$ sur un ouvert U non vide de $\mathbb{R}^n$ est continue sur U.*

Théorème 3.8 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $\lambda \in \mathbb{R}$ et soient $f : U \longrightarrow \mathbb{R}$ et $g : U \longrightarrow \mathbb{R}$ deux applications de classe $\mathcal{C}^1$ sur U. Alors :*

1. *λf et $f + g$ sont de classe $\mathcal{C}^1$ sur U et de plus : $\nabla(\lambda f) = \lambda\nabla(f)$ et $\nabla(f + g) = \nabla(f) + \nabla(g)$.*
2. *fg est de classe $\mathcal{C}^1$ sur U et de plus : $\nabla(fg) = f\nabla(g) + g\nabla(f)$ (Identité de Leibniz).*
3. *Si g ne s'annule pas sur U, alors f/g est de classe $\mathcal{C}^1$ sur U et de plus : $\nabla\left(\frac{f}{g}\right) = \frac{\nabla(f)g - \nabla(g)f}{g^2}$.*

A noter que toutes les relations du théorème 3.8 sur le calcul du gradient sont encore valables pour le calcul des dérivées partielles d'ordre 1, c'est-à-dire si l'on remplace ∇ par ∂_i pour tout $i \in \{1, ..., n\}$.

Théorème 3.9 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit I un intervalle de $\mathbb{R}$, et soient $f : U \longrightarrow \mathbb{R}$ et $g : I \longrightarrow \mathbb{R}$ deux applications telles que $f(U) \subset I$. Si f est de classe $\mathcal{C}^1$ sur U et si g est de classe $\mathcal{C}^1$ sur I, alors $g \circ f$ est de classe $\mathcal{C}^1$ sur U, et de plus : $\nabla(g \circ f) = g'(f)\nabla(f)$.*

Théorème 3.10 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit I un intervalle, soit $f : U \longrightarrow \mathbb{R}$ et $\gamma_1, ..., \gamma_n : I \longrightarrow \mathbb{R}$ des applications de classe $\mathcal{C}^1$, et posons $\gamma : I \longrightarrow \mathbb{R}^n$, $t \longmapsto (\gamma_1(t), ..., \gamma_n(t))$. Si $\gamma(I) \subset U$, alors l'application $f \circ \gamma : I \longrightarrow \mathbb{R}$, $t \longmapsto f(\gamma_1(t), ..., \gamma_n(t))$ est de classe $\mathcal{C}^1$ sur I. De plus, si l'on pose $\gamma' = (\gamma_1', ..., \gamma_n')$, alors on a pour tout $t \in I$:*

$$(f \circ \gamma)'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle = \sum_{i=1}^{n} \partial_i f(\gamma(t))\gamma_i'(t).$$

A noter que, comme pour la continuité, les résultats ci-dessus permettent de montrer qu'une fonction est de classe $\mathcal{C}^1$ sur un ouvert U non vide de $\mathbb{R}^n$, en utilisant le fait qu'elle s'obtient par sommes, produits, quotients et compositions successives de fonctions de classe $\mathcal{C}^1$.

Théorème 3.11 *Pour tout $i \in \{1, ..., n\}$, l'application $p_i : (x_1, ..., x_n) \longmapsto x_i$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$.*

Corollaire 3.12 *Toute fonction polynomiale de $\mathbb{R}^n$ dans $\mathbb{R}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$.*

A noter enfin que *toute fonction de classe $\mathcal{C}^1$ de $\mathbb{R}^n$ dans $\mathbb{R}$ est aussi de classe $\mathcal{C}^1$ sur tout ouvert U non vide de $\mathbb{R}^n$*. En particulier, toute application polynomiale de $\mathbb{R}^n$ dans $\mathbb{R}$ est de classe $\mathcal{C}^1$ sur tout ouvert U non vide de $\mathbb{R}^n$.

3.2 Notion de développement limité à l'ordre 1

Définition 3.13 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $a \in U$ et soit $f : U \longrightarrow \mathbb{R}$ une application. On dit que f admet un développement limité à l'ordre 1 en a s'il existe une application linéaire L de $\mathbb{R}^n$ dans $\mathbb{R}$ et une fonction ε qui tend vers 0 en $(0, ..., 0)$ telles que, pour tout vecteur h situé au voisinage de $(0, ..., 0)$:*

$$f(a + h) = f(a) + L(h) + \|h\|\varepsilon(h).$$

Précisons que l'expression "h situé au voisinage de $(0, ..., 0)$" signifie simplement que h doit appartenir à une boule ouverte centrée en $(0, ..., 0)$ et de rayon $r > 0$ suffisamment petit. En d'autres termes, une fonction f admet un développement limité à l'ordre 1 en $a \in U$ si f peut s'écrire comme somme d'une constante, d'une forme linéaire sur $\mathbb{R}^n$ et d'un terme (que l'on appelle le *reste*) qui est négligeable devant cette forme linéaire. On montre alors (et nous admettrons) le résultat suivant :

Théorème 3.14 (Formule de Taylor-Young à l'ordre 1) *Soit U un ouvert non vide de $\mathbb{R}^n$ et soit f une application de U dans $\mathbb{R}$. Si f est de classe $\mathcal{C}^1$ sur U, alors f admet un développement limité à l'ordre 1 en tout point $a \in U$, et ce dernier est donné par :*

$$f(a + h) = f(a) + \langle \nabla f(a), h \rangle + \|h\|\varepsilon(h).$$

En particulier, le développement limité de f à l'ordre 1 au voisinage de $a=(a_1,...,a_n)$ s'écrit sous la forme suivante, pour tout vecteur $h=(h_1,...,h_n)$ situé au voisinage de $(0,...,0)$:

$$\boxed{f(a_1+h_1,...,a_n+h_n)=f(a)+\sum_{i=1}^{n}\partial_i f(a)h_i+\sqrt{h_1^2+...+h_n^2}\,\varepsilon(h_1,...,h_n).}$$

A noter que, si f est de classe $\mathcal{C}^1$ sur un ouvert U contenant a, on définit *l'hyperplan tangent au graphe de f en a comme l'ensemble des points $(x_1,...,x_n,y)\in\mathbb{R}^{n+1}$ tels que $y=f(a)+\sum_{i=1}^{n}\partial_i f(a)(x_i-a_i)$.*

3.3 Dérivée directionnelle d'ordre 1

Définition 3.15 *Soit a un point de $\mathbb{R}^n$ et soit u un vecteur non nul de $\mathbb{R}^n$. La droite passant par a et de vecteur directeur u, notée $\mathcal{D}_{a,u}$, est l'ensemble défini par : $\mathcal{D}_{a,u}=\{M=a+tu|\ t\in\mathbb{R}\}$.*

Par définition, l'application $\gamma:\mathbb{R}\longrightarrow\mathbb{R}^n$, $t\longmapsto a+tu$ est appelée une paramétrisation de cette droite.

Théorème 3.16 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $a\in U$, soit $u=(u_1,...,u_n)$ un vecteur non nul de $\mathbb{R}^n$ et soit $f:U\longrightarrow\mathbb{R}$ une application de classe $\mathcal{C}^1$ sur U. Alors la fonction $g:t\longmapsto f(a+tu)$ est dérivable en tout réel t tel que $a+tu\in U$, et de plus on a :*

$$g'(t)=\langle\nabla f(a+tu),u\rangle=\sum_{i=1}^{n}u_i\partial_i f(a+tu).$$

En particulier, on a : $g'(0)=\langle\nabla f(a),u\rangle=\displaystyle\sum_{i=1}^{n}u_i\partial_i f(a)$.

Définition 3.17 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $a\in U$, soit $u=(u_1,...,u_n)$ un vecteur non nul de $\mathbb{R}^n$ et soit $f:U\longrightarrow\mathbb{R}$ une application. On dit que f admet une dérivée directionnelle d'ordre 1 au point a dans la direction u si la fonction $g:t\longmapsto f(a+tu)$ est dérivable en 0. Dans ce cas, le nombre dérivé $g'(0)$ est appelée la dérivée directionnelle d'ordre 1 de f en a dans la direction u, et se note $f'_u(a)$.*

En particulier, les dérivées partielles de f au point $a\in U$ (si elles existent) sont les dérivées directionnelles d'ordre 1 de f en a dans les directions données par les vecteurs $e_1,...,e_n$ de la base canonique de $\mathbb{R}^n$. A noter que, sous réserve d'existence, on a :

$$\boxed{f'_u(a)=\lim_{t\to 0}\frac{f(a+tu)-f(a)}{t}.}$$

Corollaire 3.18 *Soit U un ouvert non vide de $\mathbb{R}^n$. Alors toute application $f:U\longrightarrow\mathbb{R}$ de classe $\mathcal{C}^1$ sur U admet une dérivée directionnelle d'ordre 1 en tout point $a\in U$ et dans toute direction $u\neq 0$, laquelle est donnée pour tout $a\in U$ et tout $u\in\mathbb{R}^n\setminus\{0\}$ par : $f'_u(a)=\langle\nabla f(a),u\rangle=\sum_{i=1}^{n}\partial_i f(a)u_i$.*

D'après l'inégalité de Cauchy-Schwarz, on voit que, si $\|u\|=1$, alors $\langle\nabla(f)(a),u\rangle$ est maximal si et seulement si $\nabla(f)(a)$ et u sont colinéaires et de même sens. En d'autres termes, le gradient de f en a (s'il est non nul) nous donne la direction dans laquelle la variation de f au voisinage de a est la plus forte. A contrario, la dérivée directionnelle en a suivant une direction u orthogonale au gradient est nulle. Dans ce cas, la variation de f au voisinage de a suivant u est la plus faible, et on dit que *u est tangent à la ligne de niveau $f^{-1}(\{f(a)\})$*. En d'autres termes, *le gradient de f est orthogonal aux lignes de niveau de f.*

4 Recherche d'extrema : condition d'ordre 1

Définition 4.1 *Soit U une partie non vide de $\mathbb{R}^n$, soit $f:U\longrightarrow\mathbb{R}$ une application et soit $a\in U$. Alors :*

1. *f admet un maximum local en a si : $\exists r>0,\ \forall x\in U,\ \|x-a\|\leq r\Longrightarrow f(x)\leq f(a)$.*
2. *f admet un minimum local en a si : $\exists r>0,\ \forall x\in U,\ \|x-a\|\leq r\Longrightarrow f(x)\geq f(a)$.*
3. *f admet un extremum local en a si elle admet un maximum ou un minimum local en a.*
4. *f admet un maximum global en a si : $\forall x\in U,\ f(x)\leq f(a)$.*
5. *f admet un minimum global en a si : $\forall x\in U,\ f(x)\geq f(a)$.*
6. *f admet un extremum global en a si elle admet un maximum ou un minimum global en a.*

En particulier, on voit qu'une fonction de U dans $\mathbb{R}$ admet un extremum local en a si elle admet un extremum global sur un ensemble de la forme $U \cap B_f(a,r)$, et ce pour un certain réel $r > 0$. A noter qu'un extremum global en a est aussi un extremum local en a, mais que la réciproque n'est pas vraie.

Définition 4.2 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $a \in U$ et soit $f : U \longrightarrow \mathbb{R}$ une application de classe $\mathcal{C}^1$. On dit que a est un point critique de f si $\nabla f(a) = 0$.*

A noter qu'en un point critique a de f, toutes les dérivées directionnelles de f en a à l'ordre 1 sont nulles d'après le corollaire 3.18. A présent, nous allons donner une condition nécessaire pour l'existence d'un extremum local au point a.

Théorème 4.3 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $a \in U$ et soit $f : U \longrightarrow \mathbb{R}$ une application de classe $\mathcal{C}^1$. Si f admet un extremum local en a, alors a est un point critique de f.*

En d'autres termes et au vu de l'expression du gradient, on retiendra que :

$$\boxed{\text{si } f \text{ admet un extremum local en } a\text{, alors : } \forall i \in \{1, ..., n\}, \ \ \partial_i f(a) = 0.}$$

A noter que ce résultat n'est ni plus ni moins que la généralisation de la condition "$f'(a) = 0$" pour la recherche d'extrema des fonctions d'une seule variable. A noter aussi que ce résultat n'est pas une équivalence, mais seulement une condition nécessaire d'existence d'un extremum local. En effet, il existe des fonctions admettant des points critiques qui ne sont pas des extrema locaux.

5 Calcul différentiel à l'ordre 2

5.1 Dérivées partielles d'ordre 2 - Fonctions de classe $\mathcal{C}^2$ - Hessienne

Définition 5.1 *Soit U un ouvert non vide de $\mathbb{R}^n$ et soit $f : U \longrightarrow \mathbb{R}$ une application. Si $\partial_j f$ est bien définie sur U et si elle admet une dérivée partielle par rapport à x_i au point $a \in U$, on définit la dérivée partielle d'ordre 2 de f (par rapport à x_j puis x_i) au point a comme le réel : $\partial^2_{i,j} f(a) = \partial_i(\partial_j f)(a)$.*

A noter que la notation "$\partial^2_{i,j}(f)(a)$" est aussi valable et tout à fait acceptée.

Définition 5.2 *Soit U un ouvert non vide de $\mathbb{R}^n$ et soit $f : U \longrightarrow \mathbb{R}$ une application. Si $\partial_j f$ est bien définie sur U et si elle admet une dérivée partielle par rapport à x_i en tout point $a \in U$, on définit la dérivée partielle d'ordre 2 de f (par rapport à x_j puis x_i) comme l'application :*

$$\partial^2_{i,j} f : \left\{ \begin{array}{ccc} U & \longrightarrow & \mathbb{R} \\ a & \longmapsto & \partial^2_{i,j} f(a) \end{array} \right. .$$

De façon générale, les fonctions $\partial^2_{i,j} f$ (si elles existent) sont appelées les dérivées partielles d'ordre 2 ou dérivées partielles secondes de f. Comme pour les dérivées partielles d'ordre 1, on a la :

Définition 5.3 *Soit U un ouvert non vide de $\mathbb{R}^n$ et soit $f : U \longrightarrow \mathbb{R}$ une application. On dit que f est de classe $\mathcal{C}^2$ sur U si, pour tous $i, j \in \{1, ..., n\}$, l'application f admet une dérivée partielle par rapport à x_i puis x_j sur U, et que $\partial^2_{i,j} f$ est continue sur U.*

En d'autres termes, la notion de fonction f de classe $\mathcal{C}^2$ d'un ouvert U de $\mathbb{R}^n$ dans $\mathbb{R}$ n'est ni plus ni moins que la généralisation de la même notion pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Tout comme pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$, on montre les :

Théorème 5.4 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $\lambda \in \mathbb{R}$ et soient $f, g : U \longrightarrow \mathbb{R}$ deux applications de classe $\mathcal{C}^2$ sur U. Alors les applications $\lambda f, f + g, fg$ sont de classe $\mathcal{C}^2$ sur U. Si de plus $g(a) \neq 0$ pour tout $a \in U$, alors l'application f/g est de classe $\mathcal{C}^2$ sur U.*

Théorème 5.5 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit I un intervalle de $\mathbb{R}$, soit $f : U \longrightarrow \mathbb{R}$ une application de classe $\mathcal{C}^2$ sur U et soit $g : I \longrightarrow \mathbb{R}$ une application de classe $\mathcal{C}^2$ sur I, telles que $f(U) \subset I$. Alors $g \circ f$ est de classe $\mathcal{C}^2$ sur U.*

Comme pour les fonctions de classe $\mathcal{C}^1$, ces résultats sont fondamentaux car ils permettent de montrer qu'une fonction est de classe $\mathcal{C}^2$ sur U, en utilisant le fait qu'elle s'obtient par sommes, produits, quotients et compositions successives de fonctions de classe $\mathcal{C}^2$ sur U. C'est ce que l'on fera en général pour vérifier qu'une fonction f est de classe $\mathcal{C}^2$ sur U.

Théorème 5.6 *Pour tout $i \in \{1, ..., n\}$, l'application $p_i : (x_1, ..., x_n) \longmapsto x_i$ est de classe $\mathcal{C}^2$ sur $\mathbb{R}^n$.*

De façon générale, comme toute fonction polynomiale sur $\mathbb{R}^n$ s'obtient à partir des fonctions p_i par une succession de sommes et de produits, le théorème 5.6 entraine que :

Corollaire 5.7 *Toute fonction polynomiale de $\mathbb{R}^n$ dans $\mathbb{R}$ est de classe $\mathcal{C}^2$ sur $\mathbb{R}^n$.*

A noter que *toute fonction de classe $\mathcal{C}^2$ de $\mathbb{R}^n$ dans $\mathbb{R}$ est de classe $\mathcal{C}^2$ sur tout ouvert non vide U de $\mathbb{R}^n$.* En particulier, toute fonction polynomiale de $\mathbb{R}^n$ dans $\mathbb{R}$ est $\mathcal{C}^2$ sur tout ouvert non vide U de $\mathbb{R}^n$.

Théorème 5.8 (Schwarz) *Soit U un ouvert non vide de $\mathbb{R}^n$ et soit $f : U \longrightarrow \mathbb{R}$ une application de classe $\mathcal{C}^2$ sur U. Alors, pour tout $a \in U$ et pour tous $i, j \in \{1, ..., n\}$, on a : $\partial^2_{i,j} f(a) = \partial^2_{j,i} f(a)$.*

En d'autres termes, le théorème de Schwarz nous assure que l'on peut intervertir les dérivées partielles lors du calcul des dérivées partielles secondes d'une fonction de classe $\mathcal{C}^2$. A noter qu'il existe des fonctions qui ne sont pas de classe $\mathcal{C}^2$ et pour lesquelles $\partial^2_{i,j} f \neq \partial^2_{i,j} f$.

Définition 5.9 *Soit U un ouvert non vide de $\mathbb{R}^n$ et soit $f : U \longrightarrow \mathbb{R}$ une application de classe $\mathcal{C}^2$ sur U. On définit la matrice hessienne de f en $a \in U$ comme la matrice :*

$$\nabla^2 f(a) = \left(\partial^2_{i,j} f(a)\right)_{1 \leq i,j \leq n}.$$

En particulier, le théorème de Schwarz implique que *la matrice hessienne est symétrique.* Dès lors, passons à des rappels sur les formes quadratiques :

Définition 5.10 *Soit A une matrice symétrique de $\mathcal{M}_n(\mathbb{R})$. Pour tout vecteur $h = (h_1, ..., h_n) \in \mathbb{R}^n$, on désigne par H le vecteur colonne de composantes $h_1, ..., h_n$. Alors la forme quadratique associée à A est l'application $q : \mathbb{R}^n \longrightarrow \mathbb{R}$, $h \longmapsto {}^tHAH$.*

A noter que, si $\lambda_1, ..., \lambda_n$ sont les valeurs propres de A, on sait d'après le cours sur les endomorphismes et matrices symétriques qu'il existe une base orthonormée $\mathcal{B}$ de $\mathbb{R}^n$ telle que, pour tout vecteur h de coordonnées $x_1, ..., x_n$ dans la base $\mathcal{B}$:

$$q(h) = \sum_{i=1}^{n} \lambda_i x_i^2.$$

Définition 5.11 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $a \in U$ et soit $f : U \longrightarrow \mathbb{R}$ une application de classe $\mathcal{C}^2$ sur U. On désigne par q_a la forme quadratique associée à la matrice hessienne de f en a, c'est-à-dire l'application définie pour tout $h = (h_1, ..., h_n) \in \mathbb{R}^n$ par :*

$$q_a(h) = {}^tH(\nabla^2 f(a))H = \sum_{i=1}^{n} \sum_{j=1}^{n} \partial^2_{i,j}(f)(a) h_i h_j.$$

On verra notamment des applications de la recherche des valeurs propres de la matrice hessienne $\nabla^2 f(a)$ dans le chapitre "Recherche d'extrema".

5.2 Dérivée directionnelle seconde - Formule de Taylor-Young à l'ordre 2

Théorème 5.12 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $a \in U$, soit u un vecteur non nul de $\mathbb{R}^n$ et soit f une application de classe $\mathcal{C}^2$ de U dans $\mathbb{R}$. Alors la fonction $g : t \longmapsto f(a + tu)$ est deux dérivable en tout réel t tel que $a + tu \in U$, et de plus $g''(t) = q_{a+tu}(u)$. En particulier, on a : $g''(0) = q_a(u)$.*

A noter que le théorème 5.12 est encore valable si u est le vecteur nul.

Définition 5.13 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $a \in U$, soit u un vecteur non nul de $\mathbb{R}^n$ et soit $f : U \longrightarrow \mathbb{R}$ une application. Si la fonction $g : t \longmapsto f(a + tu)$ est deux fois dérivable en 0, alors la dérivée seconde de g en 0 s'appelle la dérivée directionnelle seconde de f en a dans la direction u, et se note $f''_u(a)$.*

Dès lors, on en déduit à partir du théorème 5.12 le :

Corollaire 5.14 *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $u = (u_1, ..., u_n)$ un vecteur non nul de $\mathbb{R}^n$ et soit f une application de U dans $\mathbb{R}$. Si f est de classe $\mathcal{C}^2$ sur U, alors f admet une dérivée directionnelle seconde en tout point $a \in U$ dans la direction u, et de plus :*

$$f''_u(a) = q_a(u) = \sum_{i=1}^{n} \sum_{j=1}^{n} \partial^2_{i,j}(f)(a) u_i u_j.$$

En particulier, si f est de classe $\mathcal{C}^2$ sur U, alors la forme quadratique associée à la matrice hessienne de f permet de calculer la dérivée directionnelle seconde de f en tout point de U et dans toutes les directions.

Théorème 5.15 (Formule de Taylor-Young à l'ordre 2) *Soit U un ouvert non vide de $\mathbb{R}^n$, soit $a \in U$ et soit $f : U \longrightarrow \mathbb{R}$ une application de classe $\mathcal{C}^2$ sur U. Alors il existe une fonction ε qui tend vers 0 en $(0, ..., 0)$ telle que, pour tout vecteur h situé au voisinage de $(0, ..., 0)$:*

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} q_a(h) + \|h\|^2 \varepsilon(h).$$

On parlera alors de l'expression ci-dessus comme *du développement limité de f à l'ordre 2 en a.* A noter que le développement limité de f à l'ordre 2 au voisinage de $a = (a_1, ..., a_n)$ s'écrit explicitement sous la forme suivante, pour tout vecteur $h = (h_1, ..., h_n)$ situé au voisinage de $(0, ..., 0)$:

$$\boxed{f(a_1 + h_1, ..., a_n + h_n) = f(a) + \sum_{i=1}^{n} \partial_i(f)(a) h_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \partial^2_{i,j}(f)(a) h_i h_j + \left(h_1^2 + ... + h_n^2\right) \varepsilon(h_1, ..., h_n).}$$