Chapitre 3

Théorèmes généraux du calcul différentiel

Sommaire

3.1 Dérivées partielles d'ordre supérieur

Définition 76

Soit U un ouvert non vide de $\mathbb{R}^n$, $a \in U$ et $f : U \to \mathbb{R}$ une fonction admettant sur U une dérivée partielle $\dfrac{\partial f}{\partial x_i}$. Si la fonction $\dfrac{\partial f}{\partial x_i} : U \to \mathbb{R}$ admet une dérivée partielle $\dfrac{\partial f}{\partial x_j}\left(\dfrac{\partial f}{\partial x_i}\right)$ par rapport à la j-ième variable au point a, on dit que $\dfrac{\partial f}{\partial x_j}\left(\dfrac{\partial f}{\partial x_i}\right)(a)$ est une dérivée partielle d'ordre 2 au point a par rapport à la i-ième et j-ième variables prises dans cet ordre.

Notation 2

La dérivée partielle $\frac{\partial f}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(a)$ *est généralement noté* $\frac{\partial^2 f}{\partial x_j \partial x_i}(a)$ *ou* $\partial^2_{x_j x_i} f(a)$.

Exemple 19

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par

$$f(x,y) = \begin{cases} \frac{x^3 y}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}.$$

Calculons $\frac{\partial^2 f}{\partial x \partial y}(0,0)$ et $\frac{\partial^2 f}{\partial y \partial x}(0,0)$. Il est facile de vérifier que

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{x^4 y + 3x^2 y^3}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases} \quad \text{et} \quad \frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{x^5 - x^3 y^2}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}.$$

Donc $\lim\limits_{x\to 0} \frac{\frac{\partial f}{\partial y}(x,0) - \frac{\partial f}{\partial y}(0,0)}{x} = \lim\limits_{x \to 0} \frac{\frac{x^5}{x^4}}{x} = 1$ et $\lim\limits_{y \to 0} \frac{\frac{\partial f}{\partial x}(0,y) - \frac{\partial f}{\partial x}(0,0)}{y} = 0$.
D'où $\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1$ et $\frac{\partial^2 f}{\partial y \partial x}(0,0) = 0$. ▪

À partir des dérivées partielles d'ordre 2, on définit les dérivées partielles d'ordre 3 lorsqu'elles existent. De proche en proche, on définit les dérivées partielles d'ordre quelconque lorsqu'elles existent. La dérivée partielle d'ordre k de $f : U \to \mathbb{R}$ au point a par rapport aux variables $x_{i_k}, \ldots, x_{i_2}, x_{i_1}$ prises dans cet ordre est notée $\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}}(a)$, qu'est par définition la dérivée partielle au point a par rapport à la variable d'indice x_{i_1} de la fonction $\frac{\partial^{k-1} f}{\partial x_{i_2} \ldots \partial x_{i_k}}$.

Exemple 20

Cherchons les dérivées partielles d'ordre 3 de la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x + y - x^2 y^3$.

1. Les dérivées partielles d'ordre 1 sont

$$\frac{\partial f}{\partial x}(x,y) = 1 - 2xy^3, \ \frac{\partial f}{\partial y}(x,y) = 1 - 3x^2 y^2.$$

2. Les dérivées partielles d'ordre 2 sont

$$\frac{\partial^2 f}{\partial x^2}(x,y) = -2y^3, \ \frac{\partial^2 f}{\partial y^2}(x,y) = -6x^2 y, \ \frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = -6xy^2.$$

3. Les dérivées partielles d'ordre 3 sont

$$\frac{\partial^3 f}{\partial x^3}(x,y) = 0, \ \frac{\partial^3 f}{\partial y^3}(x,y) = -6x^2, \ \frac{\partial^3 f}{\partial x \partial y^2}(x,y) = \frac{\partial^2 f}{\partial y^2 \partial x}(x,y) = -12xy,$$

$$\frac{\partial^3 f}{\partial x^2 \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x^2}(x,y) = -6y^2.$$

▪

Définition 77

Soit f une fonction de $U \subset \mathbb{R}^n$ dans $\mathbb{R}$, $k \in \mathbb{N}^*$. On dit que f est de *classe* $\mathcal{C}^k$ sur U et on écrit $f \in \mathcal{C}^k(U)$ si toutes ses dérivées partielles jusqu'à l'ordre k existent et sont continues sur U.

Si $f \in \mathcal{C}^k(U)$ pour tout $k \in \mathbb{N}^*$, on dit f est de classe $\mathcal{C}^\infty$ sur U et on écrit $f \in \mathcal{C}^\infty(U)$.

3.1.1 Théorème de Schwarz

Théorème 78

Soit V un ouvert non vide de $\mathbb{R}^2$, $a = (a_1, a_2) \in V$ et $g : V \to \mathbb{R}$ une fonction telle que $\frac{\partial^2 g}{\partial x \partial y}$ et $\frac{\partial^2 g}{\partial y \partial x}$ existent sur V et sont continues au point a. Alors $\frac{\partial^2 g}{\partial x \partial y}(a) = \frac{\partial^2 g}{\partial y \partial x}(a)$.

Démonstration. Puisque V est un ouvert, alors il existe $r > 0$ tel $B(a, r) \subset V$. On définit alors la fonction ϕ sur $B\left(0, \frac{r}{2}\right)$ par

$$\phi(x, y) = g(x, y) - g(x, a_2) - g(a_1, y) + g(a_1, a_2).$$

On va montrer que $\frac{\phi(x,y)}{(x-a_1)(y-a_2)}$ admet une limite en $a = (a_1, a_2)$ et que celle-ci peut s'exprimer de deux façons. Par application du théorème des accroissements finis à la fonction $x \mapsto \phi(x, y)$ entre a_1 et x, il existe c_x entre a_1 et x tel que-0.4cm

$$\frac{\phi(x, y)}{(x - a_1)} = \frac{\partial \phi}{\partial x}(c_x, y) = \frac{\partial g}{\partial x}(c_x, y) - \frac{\partial g}{\partial x}(c_x, a_2).$$

Par réapplication du théorème des accroissements finis, cette fois à la fonction $y \mapsto \frac{\phi(x,y)}{(x-a_1)}$ entre a_2 et y, il existe c_y entre a_2 et y tel que-0.4cm

$$\frac{\phi(x, y)}{(x - a_1)(y - a_2)} = \frac{\partial^2 \phi}{\partial y \partial x}(c_x, c_y) = \frac{\partial^2 g}{\partial y \partial x}(c_x, c_y).$$

Comme $\frac{\partial^2 g}{\partial y \partial x}$ est continue, on en déduit alors que

$$\lim_{(x,y)\to(a_1,a_2)} \frac{\phi(x, y)}{(x - a_1)(y - a_2)} = \frac{\partial^2 g}{\partial y \partial x}(a_1, a_2).$$

Si on reprend ce raisonnement en intervertissant l'ordre des variables, on montre de la même façon que

$$\lim_{(x,y)\to(a_1,a_2)} \frac{\phi(x, y)}{(x - a_1)(y - a_2)} = \frac{\partial^2 g}{\partial x \partial y}(a_1, a_2).$$

D'où l'égalité des dérivées secondes croisées. ■

Les corollaires suivants peuvent être déduits de ce théorème.

Corollaire 79

Soit U un ouvert non vide de $\mathbb{R}^n$, $a \in U$ et f une fonction de U dans $\mathbb{R}$. Si $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ et $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ existent au voisinage de a et sont continues en a, alors $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a) = \dfrac{\partial^2 f}{\partial x_j \partial x_i}(a)$.

Corollaire 80

Si f est de classe $\mathcal{C}^2$ sur un ouvert $U \subset \mathbb{R}^n$, on a alors en tout point de U

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}, \quad 1 \leq i, j \leq n.$$

Ce résultat nous montre que la matrice des dérivées partielles d'ordre 2, $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{1 \leq i,j \leq n}$ est une matrice symétrique. On l'appelle la matrice hessienne de f.

L'exemple 19 montre que l'existence des dérivées partielles croisées $\frac{\partial^2 f}{\partial x \partial y}(x, y)$ et $\frac{\partial^2 f}{\partial y \partial x}(x, y)$ d'une fonction f n'implique pas nécessairement l'égalité de ces dérivées. Le théorème de Schwarz précise que la continuité de ces dérivées suffit à assurer leur égalité.

Corollaire 81

Si f est de classe $\mathcal{C}^k$ sur un ouvert $U \subset \mathbb{R}^n$, on a alors en tout point de U

$$\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} ... \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{\sigma(i_1)} \partial x_{\sigma(i_2)} ... \partial x_{\sigma(i_k)}}, \quad 1 \leq i_1, ..., i_k \leq n,$$

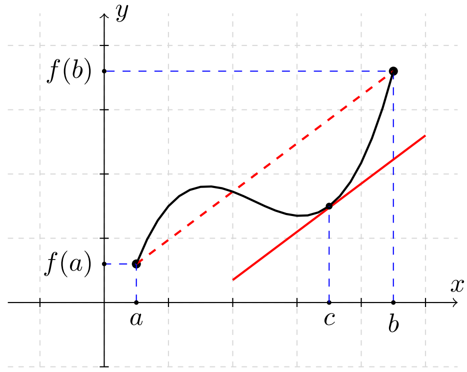
pour toute permutation σ de $\{i_1, ..., i_k\}$.

3.2 Théorème des accroissements finis

Rappelons le résultat vu au lycée et en première année.

Soit f une fonction définie sur le segment $[a, b]$ de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. Si f est continue sur $[a, b]$ et dérivable sur $]a, b[$, alors il existe $c \in]a, b[$ tel que

$$f(b) - f(a) = f'(c)(b - a).$$



Géométriquement, il existe un point c entre a et b où la tangente à la courbe de f a même pente que la droite qui passe par les points $(a, f(a))$ et $(b, f(b))$. Notez que ce théorème ne garantit pas l'unicité de c.

Définition 82

Soit a et b deux points de $\mathbb{R}^n$. On appelle segment de $\mathbb{R}^n$ d'extrémités a et b, l'ensemble de $\mathbb{R}^n$ noté $[a,b]$ et défini par

$$[a,b] = \{a + t\,(b-a)\,,\quad t \in [0,1]\}\,.$$

Cette définition généralise à $\mathbb{R}^n$ la notion bien connue de segment de $\mathbb{R}$.

Théorème 83 (Théorème des accroissements finis)

Soit U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ une fonction définie et continue U. Soit $a = (a_1, ..., a_n)$ et $b = (b_1, ..., b_n)$ dans U tels que le segment $[a,b]$ soit contenu dans U. Si f est différentiable en chaque point du segment ouvert $]a,b[$, alors il existe un point c de $]a,b[$ tel que

$$f(b) - f(a) = \nabla f(c) \cdot (b-a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(c)\,(b_i - a_i)\,.$$

Démonstration. Il suffit de se ramener au cas connu. Soit la fonction

$$\begin{array}{rccl} \varphi : & [0,1] & \longrightarrow & \mathbb{R} \\ & t & \longmapsto & \varphi(t) = f(a + t\,(b-a)) \end{array}.$$

La fonction φ est continue sur $[0,1]$ et dérivable sur $]0,1[$, donc il existe $t_0 \in]0,1[$ tel que

$$\varphi(1) - \varphi(0) = (1-0)\,\varphi'(t_0)\,.$$

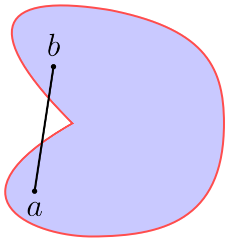
Or $\varphi(1) = b$, $\varphi(0) = a$, et par la règle de dérivation en chaîne

$$\varphi'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a + t\,(b-a))\,(b_i - a_i)\,.$$

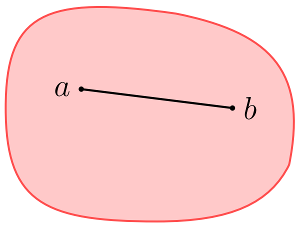
Le résultat suit en posant $c = a + t_0\,(b-a)$. ■

Définition 84

On dit qu'un ensemble U de $\mathbb{R}^n$ est *convexe* si pour tout $a, b \in U$, le segment $[a,b]$ est inclus dans U i.e. $\forall a, b \in U,\ [a,b] \subset U$.



Ensemble non convexe



Ensemble convexe

Exemple 21

- Les boules ouvertes ou fermées de $\mathbb{R}^n$ sont des ensembles convexes.
- Les sphères de $\mathbb{R}^n$ ne sont pas convexes. ■

Corollaire 85

Soit U un ouvert convexe de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ une fonction différentiable sur U. Si le gradient de f est nulle en tout point de U, alors f est constante sur U.

Démonstration. Soit x_0 un point fixé dans U et x un point quelconque de U. Il existe donc $c_x \in]x_0, x[$ tel que

$$f(x) - f(x_0) = \nabla f(c_x) \cdot (x - x_0).$$

Comme $\nabla f(x)$ pour tout $x \in U$, il vient $f(x) = f(x_0)$ pour tout $x \in U$. La fonction f est donc constante sur U. ■

Remarque 86

Soit U un ouvert convexe de $\mathbb{R}^n$ et $F : U \to \mathbb{R}^m$ différentiable sur U. Alors

$$\forall a, b \in U, \ \|F(a) - F(b)\| \leq \|b - a\| \sup_{x \in U} \|J_F(x)\|,$$

où $J_F(x)$ est la matrice jacobienne de F.

3.3 Formule de Taylor

Ici on généralise à une fonction numérique de n variables réelles la formule de Taylor dans le cas des fonctions numériques d'une variable réelle et qu'on rappelle ci-dessous.
Si g est une fonction numérique de classe $\mathcal{C}^p$ sur l'intervalle $[a, a+h]$, on a

$$g(a+h) = g(a) + \sum_{k=1}^{p} \tfrac{g^{(k)}(a)}{k!} h^k + h^p \varepsilon(h) \ \text{ avec } \lim_{h \to 0} \varepsilon(h) = 0.$$

Soit maintenant, U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^p$ sur U, $a, a+h \in U$. On note

$$D^{(k)} f(a)(h)^{(k)} = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \dots \sum_{i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}(a) \, h_{i_1} h_{i_2} \dots h_{i_k}, \ k = 1, \dots, p.$$

On a par exemple

$$k=1, \quad Df(a)(h)=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a)\,h_i=h\cdot\nabla f(a)\,,$$
$$k=2, \quad D^{(2)}f(a)(h)^{(2)}=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(a)\,h_ih_j=h^{\mathsf{T}}H(a)\,h,$$

où $H=\left(\frac{\partial^2 f}{\partial x_i\partial x_j}\right)_{1\le i,j\le n}$ est la matrice hessienne de f.

Formellement, si on pose $\nabla=\left(\frac{\partial}{\partial x_1},...,\frac{\partial}{\partial x_2}\right)$, alors par le théorème de Schwarz, on peut calculer $D^{(k)}f(a)(h)^{(k)}$ comme $(h\cdot\nabla)^k f(a)$. Par exemple si $n=2$ et $k=2$ on a

$$\begin{aligned} D^{(2)}f(a)(h)^{(2)} &= \left((h_1,h_2)\cdot\left(\tfrac{\partial}{\partial x_1},\tfrac{\partial}{\partial x_2}\right)\right)^2 f(a)=\left(h_1\tfrac{\partial}{\partial x_1}+h_2\tfrac{\partial}{\partial x_2}\right)^2 f(a) \\ &= \left(h_1^2\tfrac{\partial^2}{\partial x_1^2}+2h_1h_2\tfrac{\partial^2}{\partial x_1\partial x_2}+h_2^2\tfrac{\partial^2}{\partial x_2^2}\right)f(a)\,. \end{aligned}$$

Théorème 87

Sous les conditions précédentes et en supposant que le segment $[a,a+h]$ est inclus dans U on a alors le développement de Taylor de f au voisinage de a

$$f(a+h)=f(a)+\sum_{k=1}^{p}\tfrac{1}{k!}D^{(k)}f(a)(h)^{(k)}+\|h\|^p\varepsilon(h) \text{ avec } \lim_{h\to 0}\varepsilon(h)=0.$$

Démonstration. Soit la fonction

$$\begin{array}{llll} \varphi: & [0,1] & \longrightarrow & \mathbb{R} \\ & t & \longmapsto & \varphi(t)=f(a+t\,h) \end{array}.$$

La fonction φ est dérivable sur $[0,1]$, donc pour avoir le résultat, il suffit d'appliquer la formule de Taylor pour les fonctions d'une variable. ■

Si $f:U\subset\mathbb{R}^n\to\mathbb{R}$ une fonction de classe $\mathcal{C}^2$ sur U, alors le développement de Taylor d'ordre 2 est

$$f(a+h)=f(a)+h\cdot\nabla f(a)+\tfrac{1}{2}h^{\mathsf{T}}H(a)\,h+\|h\|^2\varepsilon(h) \text{ avec } \lim_{h\to 0}\varepsilon(h)=0,$$

et $H=\left(\frac{\partial^2 f}{\partial x_i\partial x_j}\right)_{1\le i,j\le n}$ est la matrice hessienne de f.

Dans le cas particulier des fonctions de deux variables i.e. $n=2$, $a=(a_1,a_2)$ et $h=(h_1,h_2)$, le développement de Taylor d'ordre 2 est donné par la formule

$$f(a+h)=f(a)+\tfrac{\partial f}{\partial x}(a)\,h_1+\tfrac{\partial f}{\partial y}(a)\,h_2+\tfrac{1}{2}\left(\tfrac{\partial^2 f}{\partial x^2}(a)\,h_1^2+2\tfrac{\partial^2 f}{\partial x\partial y}(a)\,h_1h_2+\tfrac{\partial^2 f}{\partial y^2}(a)\,h_2^2\right)+\|h\|^2\varepsilon(h)\,,$$

avec $\lim_{h\to 0}\varepsilon(h)=0$.

Exemple 22

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par $f(x,y) = e^{x+y}$ qui est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$.

Le développement d'ordre 2 en $(0,0)$ est

$$e^{x+y} \simeq 1 + h_1 + h_2 + \frac{1}{2}\left(h_1^2 + 2h_1h_2 + h_2^2\right). \qquad \blacksquare$$

3.3.1 Points critiques et extrema libres

Dans ce paragraphe U est un ouvert de $\mathbb{R}^n$, f une fonction de U dans $\mathbb{R}$ et $a \in U$. On s'intéresse aux extrema de f.

Définition 88 (Points critiques)

Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ sur U. On dit que $a \in U$ est un point *critique*, ou singulier, ou stationnaire, de f si toutes les dérivées partielles de f sont nulles en a, c'est-à-dire

$$\nabla f(a) = 0.$$

Cette définition s'étend aux fonctions à valeurs dans $\mathbb{R}^m$ en remplaçant $\nabla f(a)$ par $J_F(a)$.

Définition 89

On dit que f admet un *minimum local* en a s'il existe un voisinage V de a tel que

$$f(a) \leq f(x) \quad \forall x \in V.$$

On dit que f admet un *minimum global* ou *absolu* en a si $f(a) \leq f(x) \quad \forall x \in U$.

On définit de la même manière les notions de maximum local et global en remplaçant $\leq$ par $\geq$.
On utilise le nom *extremum* pour désigner sans distinction un maximum ou minimum.
On dit qu'un extremum est strict si les inégalités précédentes sont strictes.

Condition nécessaire d'existence d'un extremum

Théorème 90 (Condition nécessaire d'existence d'un extremum)

Si f est différentiable sur U et qu'elle présente en $a \in U$ un extremum local ou global alors a est un point critique de f, i.e. $\nabla f(a) = 0$.

Démonstration. Supposons que cet extremum est un minimum. L'hypothèse faite sur f implique qu'il existe une boule ouverte $B(a,r)$ telle que $f(a) \leq f(x)\ \forall x \in B(a,r)$.

Pour h fixé dans $\mathbb{R}^n$ tel que $a+h \in B(a,r)$, on définit la fonction φ de classe $\mathcal{C}^1$ sur $]-1,1[$ par $\varphi(t) = f(a+t\,h)$.

On a $\varphi(0) = f(a) \leq f(a+t\,h) = \varphi(t)\ \ \forall t \in]-1,1[$. Ce qui implique que $\varphi'(0) = 0$.

Comme $\varphi'(t) = \nabla f(a+t\,h) \cdot h$, alors $\nabla f(a) \cdot h = 0$ pour tout h fixé tel que $\|h\| < r$.

Si $k \neq 0$ un point quelconque de $\mathbb{R}^n$, alors en choisissant $h = \frac{rk}{2\|k\|}$, on obtient $\nabla f(a) \cdot \frac{rk}{2\|k\|} = 0$, ce qui implique que $\nabla f(a) \cdot k = 0$.

Comme k est quelconque dans $\mathbb{R}^n \setminus \{0\}$, on conclut que $\nabla f(a) = 0$. ■

Remarquons que la nullité de la différentielle constitue une condition nécessaire mais non suffisante d'existence d'un extremum. En effet, soit la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^3 + y^3$. Un calcul simple montre que $\nabla f(0,0) = 0$.

D'autre part, on a

$$-h^3 = f(-h,0) < f(0,0) < f(h,0) = h^3\ \forall h \in \mathbb{R}_+^*.$$

Ce qui montre que f n'admet pas d'extremum en $(0,0)$.

Le théorème précédent dit alors que les extrema d'une fonction différentiable, quand ils existent, sont à chercher parmi ses points critiques, c'est-à-dire parmi les solutions de l'équation

$$\nabla f(x) = 0, x \in U.$$

Condition suffisante d'existence d'un extremum

Soit f une fonction d'un ouvert U de $\mathbb{R}^n$ dans $\mathbb{R}$ de classe $\mathcal{C}^2$ et $a \in U$ un point critique de f. Pour tout $h \in \mathbb{R}^n$ tel que $\|h\|$ est suffisamment petit, on a

$$f(a+h) = f(a) + \tfrac{1}{2} h^{\mathsf{T}} H(a)\, h + \|h\|^2 \varepsilon(h),$$

où $\lim\limits_{h \to 0} \varepsilon(h) = 0$ et $H = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{1 \leq i,j \leq n}$ est la matrice hessienne de f qui est symétrique.

Le terme $\frac{1}{2} h^{\mathsf{T}} H(a)\, h = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, h_i h_j$ c'est la forme quadratique associée à la matrice hessienne $H(a)$ de f en a. On rappelle qu'une matrice symétrique se diagonalise dans une base orthonormée et que toutes ses valeurs propres sont réelles.

En notant $(v_1, ..., v_n)$ et $(\lambda_1, ..., \lambda_n)$ la base orthonormée et les valeurs propres pour la matrice $H(a)$, et en décomposant h suivant cette base $h = \sum\limits_{i=1}^{n} \alpha_i v_i$, on obtient l'expression plus simple $h^{\mathsf{T}} H(a)\, h = \sum\limits_{i=1}^{n} \alpha_i^2 \lambda_i$. On peut à partir de cette remarque obtenir une condition suffisante pour avoir un extremum au point a.

Théorème 91

Soit f une fonction d'un ouvert U de $\mathbb{R}^n$ dans $\mathbb{R}$ de classe $\mathcal{C}^2$ et $a \in U$ un point critique de f.

- Si les λ_i sont tous strictement positifs alors f admet un minimum local au point a.
- Si les λ_i sont tous strictement négatifs alors f admet un maximum local au point a.

Dans le cas $n = 2$ la matrice hessienne est de taille 2×2 et il est particulièrement facile de vérifier le signe des λ_1 et λ_2 en remarquant que le déterminant $H(a)$ vaut $\lambda_1 \lambda_2$ et que la trace $tr(H(a))$ vaut $\lambda_1 + \lambda_2$.

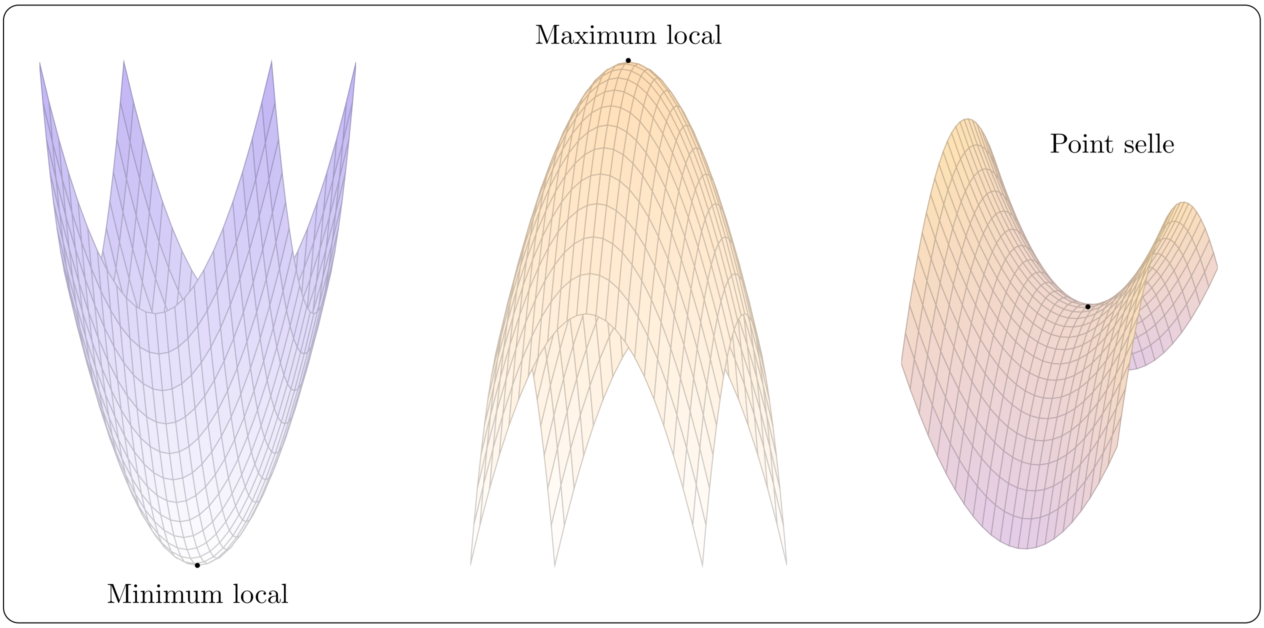
Théorème 92

Soit U un ouvert U de $\mathbb{R}^2$, $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ et $a \in U$ un point critique de f, i.e. $\frac{\partial f}{\partial x}(a) = \frac{\partial f}{\partial y}(a) = 0$. Alors, avec les notations de Monge

$$p = \frac{\partial^2 f}{\partial x^2}(a), \quad q = \frac{\partial^2 f}{\partial x \partial y}(a), \quad r = \frac{\partial^2 f}{\partial y^2}(a),$$

on a

1. Si $pr - q^2 > 0$ et $p > 0$: f admet en a un minimum local.
2. Si $pr - q^2 > 0$ et $p < 0$: f admet en a un maximum local.
3. Si $pr - q^2 < 0$: f n'admet en a ni maximum ni minimum local, mais un point selle.
4. Si $pr - q^2 = 0$: on ne peut conclure a priori.



Exemple 23

Étudier l'existence des extrema de la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = 3x^3+3y^3-x-y$.

Il est clair que f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$. Un simple calcul donne

$$\frac{\partial f}{\partial x}(x,y) = 9x^2-1,\ \frac{\partial f}{\partial y}(x,y) = 9y^2-1,$$

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 18x,\ \frac{\partial^2 f}{\partial y^2}(x,y) = 18y,\ \frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = 0.$$

Les points critiques de f sont les solutions du système $\{9x^2-1=0,\ 9y^2-1=0\}$.

Le système admet donc 4 solutions qui sont

$$s_1 = \left(\tfrac{1}{3},\tfrac{1}{3}\right),\ s_2 = \left(\tfrac{1}{3},-\tfrac{1}{3}\right),\ s_3 = \left(-\tfrac{1}{3},\tfrac{1}{3}\right),\ s_4 = \left(-\tfrac{1}{3},-\tfrac{1}{3}\right).$$

- En s_1 on a $pr-q^2 = 36 > 0$ et $p = 6 > 0$: f présente donc en s_1 un minimum local.
- En s_2 et s_3 on a $pr-q^2 = -36 < 0$: f n'admet d'extremum en aucun de ces deux points.
- En s_4 on a $pr-q^2 = 36 > 0$ et $p = -6 < 0$: f présente donc en s_4 un maximum local. ∎

3.4 Théorème des fonctions implicites

Un résultat classique pour les fonctions d'une variable affirme que si f est de classe $\mathcal{C}^1$ sur un intervalle I et $f'(x_0) \neq 0,\ x_0 \in I$, alors f est strictement monotone sur un un voisinage $I_0 \subset I$ de x_0 et est une bijection entre les intervalles I_0 et $f(I_0)$. De plus la bijection réciproque est aussi de classe $\mathcal{C}^1$. Ce résultat se généralise en dimension supérieure à un théorème appelé le théorème d'inversion locale.

Théorème 93 (Théorème d'inversion locale)

Soit f de classe $\mathcal{C}^1$ sur un ouvert $U \subset \mathbb{R}^n$ et à valeurs dans $\mathbb{R}^n$ et soit $x \in U$ telle que la matrice jacobienne $J_F(x)$ est inversible. Alors il existe deux ouverts, $U' \subset U$ et V tels que $x \in U'$ et $f(x) \in V$ et tels que f est un $\mathcal{C}^1$ difféomorphisme de U' dans V.

Une application du théorème d'inversion locale concerne le problème suivant : pour f de classe $\mathcal{C}^1$ sur un ouvert $U \subset \mathbb{R}^2$, on considère l'équation $f(x,y) = 0$ et on cherche à comprendre si cette équation est en un certain sens équivalente à $y = g(x)$ où g est une fonction d'une variable. L'exemple de la fonction $f(x,y) = x^2+y^2-1$ nous montre que ceci n'est possible que localement : certaines portions du cercle unité s'identifient au graphe de la fonction $y = \sqrt{1-x^2}$, d'autres à celui de la fonction $y = -\sqrt{1-x^2}$. D'autre portions, telles qu'au voisinage du point $(1,0)$ ne peuvent s'identifier à un graphe. Le théorème des fonctions implicites donne un résultat général allant dans ce sens.

Théorème 94 (Théorème des fonctions implicites dans $\mathbb{R}^2$)

Soit U un ouvert de $\mathbb{R}^2$, (a,b) un élément de U, f une fonction de U dans $\mathbb{R}$. On suppose que
1) f est de classe $\mathcal{C}^1$ sur U.
2) $f(a,b)=0$ et $\frac{\partial f}{\partial y}(a,b)\neq 0$.
Alors il existe un intervalle ouvert I contenant a, un intervalle ouvert J contenant b et une fonction φ unique de I dans J de classe $\mathcal{C}^1$ sur I et vérifiant
$I\times J\subset U,\quad \varphi(a)=b,\quad f(x,\varphi(x))=0\ \forall x\in I$ et $\varphi'(x)=-\frac{\frac{\partial f}{\partial x}(x,\varphi(x))}{\frac{\partial f}{\partial y}(x,\varphi(x))}$.

Démonstration. La démonstration de ce théorème peut s'effectuer en considérant la fonction $h(x,y)=(x,f(x,y))$ et en lui appliquant le théorème d'inversion locale qui permet de définir $(x,g(x))$ comme l'antécédent de $(x,0)$. ■

Exemple 24

Soit $f:\mathbb{R}^2\to\mathbb{R}$ la fonction définie par $f(x,y)=e^{y-1}+y-x$. Montrer que la relation $f(x,y)=0$ définit implicitement y en fonction de x au voisinage du point $(2,1)$.
Il est clair que f est de classe $\mathcal{C}^1$ sur l'ouvert $\mathbb{R}^2$. D'autre part on a $f(2,1)=0$ et $\frac{\partial f}{\partial y}(x,y)=e^{y-1}+1$, $\frac{\partial f}{\partial y}(2,1)=2\neq 0$. Il existe donc un intervalle ouvert I contenant 2, un intervalle ouvert J contenant 1 et une fonction $\varphi:I\to J$ de classe $\mathcal{C}^1$ sur I et vérifiant $\varphi(2)=1$, $f(x,\varphi(x))=0\ \forall x\in I$ et $\varphi'(x)=-\frac{\frac{\partial f}{\partial x}(x,\varphi(x))}{\frac{\partial f}{\partial y}(x,\varphi(x))}=\frac{1}{e^{\varphi(x)-1}+1}$. ■

Le théorème des fonctions implicite se généralise aux fonctions de plus de deux variables.

Théorème 95 (Théorème des fonctions implicites dans $\mathbb{R}^{n+1}$)

Soit U un ouvert de $\mathbb{R}^n\times\mathbb{R}$, $(a,b)\in U$, où $a=(a_1,a_2,...,a_n)\in\mathbb{R}^n, b\in\mathbb{R}$,

$$\begin{aligned} f:\ \ U\subset\mathbb{R}^n\times\mathbb{R}\ &\to\mathbb{R}\\ (x,y)\ &\mapsto f(x,y)\end{aligned}$$

une fonction de classe $\mathcal{C}^1$ sur U. On suppose $f(a,b)=0$ et $\frac{\partial f}{\partial y}(a,b)\neq 0$.
Alors il existe un pavé ouvert A contenant a, un intervalle ouvert J contenant b et une fonction unique $\varphi:A\to J$ de classe $\mathcal{C}^1$ sur A et vérifiant
$A\times J\subset U,\quad \varphi(a)=b,\quad f(x,\varphi(x))=0\ \forall x\in A$ et $\nabla\varphi(x)=-\frac{\left(\frac{\partial f}{\partial x_1}(x,\varphi(x)),...,\frac{\partial f}{\partial x_n}(x,\varphi(x))\right)}{\frac{\partial f}{\partial y}(x,\varphi(x))}$.

3.4.1 Extrema liés

On s'est déjà intéressé à l'étude des extrema d'une fonction f définie sur un ouvert de $\mathbb{R}^n$. Dans de nombreuses applications, il est fréquent que l'on cherche à minimiser ou maximiser f sur un sous-ensemble restreint ce qui correspond à imposer une contrainte sur la variable x.

Définition 96 (Lagrangien)

On cherche les extrema de la fonction $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^1$ sous la contrainte $g(x, y) = 0$ avec $g : U \subset \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^1$. Le Lagrangien associé au problème est la fonction de 3 variables et de classe $\mathcal{C}^1$

$$\begin{array}{rccl} L: & U \times \mathbb{R} & \to & \mathbb{R} \\ & (x, y, \lambda) & \mapsto & f(x, y) + \lambda g(x, y), \end{array}$$

La variable λ est appelée multiplicateur de Lagrange.

Dans le cas général, on a à optimiser

$$\begin{array}{rccl} f: & \mathbb{R}^n & \to & \mathbb{R} \\ & (x_1, ..., x_n) & \mapsto & f(x_1, ..., x_n), \end{array}$$

sous p contraintes $(p \leq n)$: $g_i(x_1, ..., x_n) = 0, \; i = 1, ..., p$.
Il y a alors p multiplicateurs de Lagrange $\lambda_1, ..., \lambda_p$ et le Lagrangien associé est la fonction de $(n + p)$ variables $L(x_1, ..., x_n, \lambda_1, ..., \lambda_p) = f(x_1, ..., x_n) + \sum\limits_{i=1}^{p} \lambda_i g_i(x_1, ..., x_n)$.

Théorème 97 (de Lagrange dans $\mathbb{R}^2$)

Pour que f admette en (x_0, y_0) un extremum local sous la contrainte $g(x, y) = 0$, le point (x_0, y_0) n'étant pas un point critique de g, il faut qu'il existe λ_0 tel que (x_0, y_0, λ_0) soit un point critique du Lagrangien L i.e. $\nabla f(x_0, y_0) = \lambda_0 \nabla g(x_0, y_0)$.

Le théorème de Lagrange n'est pas un théorème d'existence des extrema liés d'une fonction ; il dit simplement que, dans le cas où la fonction f admet un extremum lié, cet extremum est à chercher parmi les solutions du système $\nabla f = \lambda \nabla g$.

Exemple 25

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par $f(x, y) = x^4 + y^2 - 5x^2$. Déterminons les extrema de f sachant que x, y sont liées par la relation $g(x, y) = x + y = 0$.

Les points qui vérifient la condition nécessaire de Lagrange sont des solution du système

$$\left\{\frac{\partial f}{\partial x}=\lambda\frac{\partial g}{\partial x},\ \frac{\partial f}{\partial y}=\lambda\frac{\partial g}{\partial y}\right\},\ \text{i.e. } 4x^3-10x=\lambda \text{ et } 2y=\lambda.$$

D'où $(x_0, y_0, \lambda_0) = (0,0,0)\,, (2,-2,-4)$ ou $(-2,2,4)$.

Observons que $f(x,-x) = x^2(x^2-4)$, alors que le point $(0,0)$ qui est un minimum local lié. ∎

Ce théorème se généralise aux fonctions de plus de deux variables sous plusieurs contraintes.

Théorème 98 (de Lagrange dans $\mathbb{R}^n$)

Soit U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur U. Soit $x_0 \in U$ un extremum local de f sous les contraintes $g_i(x) = 0,\ i = 1, ..., p$. On suppose que les vecteurs $\nabla g_1(x_0), \nabla g_2(x_0), ..., \nabla g_p(x_0)$ sont linéairement indépendants. Alors, il existe des réels $\lambda_i,\ i = 1, ..., p$, multiplicateurs de Lagrange tels que

$$\nabla f(x_0) = \lambda_1 \nabla g_1(x_0) + \lambda_2 \nabla g_2(x_0) + ... + \lambda_p \nabla g_p(x_0)\,.$$