Leçon 215: Applications différentiables sur un ouvert de $\mathbb{R}^n$ Exemples et applications.

Adrien Fontaine

2 décembre 2012

Table des matières

Objectif : Il s'agit de généraliser à des applications de $\mathbb{R}^n$ dans $\mathbb{R}^p$ la dérivée usuelle des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. On approche l'accroissement d'une fonction "quelconque" par une application linéaire. C'est la "grande idée" du calcul différentiel.
Dans toute la leçon, U désigne un ouvert de $\mathbb{R}^n$.

1 Généralités

1.1 Différentielle

Définition 1

Une application $f : U \to \mathbb{R}^p$ est dite différentiable en $a \in U$ s'il existe $\varphi \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ telle que :

$$f(a+h) = f(a) + \varphi(h) + \circ(\|h\|) \text{ lorsque } h \to 0$$

Si une telle application φ existe, elle est unique et s'appelle la différentielle de f en a, notée $Df(a)$.
Si f est différentiable en tout point de U, on dit que f est différentiable sur U, et l'application $Df : a \mapsto Df(a)$ est appelée application différentielle de f.
Si Df est continue, on dit que f est de classe C^1.

Exemple 1

- *$f : I \to \mathbb{R}$ où I est un intervalle de $\mathbb{R}$ est différentiable en $a \in I$, si et seulement si, elle est dérivable en a, et $Df(a) : h \mapsto f'(a)h$. Ainsi, la notion de différentielle prolonge bien la notion de dérivée.*
- *Si $f : \mathbb{R}^n \to \mathbb{R}^p$ est linéaire, alors elle est différentiable et $\forall a, Df(a) = f$.*
- *Si f désigne l'application inverse sur l'espace des matrices $\mathbb{R}^{n\times n}$, alors pour tout X inversible, et $H \in \mathbb{R}^{n\times n}$,*
$$Df(X)H = -X^{-1}HX^{-1}$$
- *Si det désigne la fonction déterminant sur l'espace $\mathbb{R}^{n\times n}$ des matrices de taille n et tr la fonction trace, alors on a pour tout $X, H \in \mathbb{R}^{n\times n}$,*
$$Ddet(X)H = tr({}^t\tilde{X}H)$$
où $\tilde{X}$ est la comatrice de X. [1]

Proposition 1

Soit $f : U \to \mathbb{R}^n$ et $g : V \to \mathbb{R}^p$ deux applications telles que $f(U) \subset V$. On suppose que f est différentiable en $a \in U$ et g différentiable en $f(a)$. Alors $g \circ f$ est différentiable en a et :

$$Dg \circ f(a) = Dg(f(a)) \circ Df(a)$$

Démonstration : Gourdon, Proposition 3p304 ■

1.2 Dérivées partielles

Définition 2

Soit $f : U \to \mathbb{R}^p$, $a \in U$ et $v \in \mathbb{R}^n$. f est dite dérivable selon le vecteur v si

$$\lim_{t \to 0, t \neq 0} \frac{f(a+tv) - f(a)}{t} \text{ existe}$$

On note alors cette limite $f'_v(a)$.

Proposition 2

Si f est différentiable en a, alors f admet une dérivée selon tout vecteur et on a $f'_v(a) = Df(a)v$ pour tout $v \in \mathbb{R}^n$.

Démonstration : Gourdon, proposition 4p304 ■

Remarque 1

La réciproque est fausse. La dérivabilité selon tout vecteur en a n'entraîne pas forcément la différentiabilité de f en a. En effet, la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

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

est dérivable selon tout vecteur au point $(0,0)$ mais n'est même pas continue en $(0,0)$.

Démonstration : Gourdon, exercice 1p309 ■

Définition 3

Soit $(e_1, ..., e_n)$ la base canonique de $\mathbb{R}^n$. On appelle (lorsqu'elle existe) dérivée partielle d'indice i de f en a, le vecteur :

$$f'_{e_i} = \frac{\partial f}{\partial x_i}(a)$$

Définition 4

Si $f : U \to \mathbb{R}$ est une application différentiable en $a \in U$, alors $Df(a)$ est une forme linéaire continue sur $\mathbb{R}^n$, qui est donc caractérisée par un unique, vecteur appelé gradient de f en a et noté $\nabla f(a)$. On a alors :

$$\forall h \in \mathbb{R}^n, Df(a)h =< \nabla f(a), h >$$

où $<,>$ est le produit scalaire usuel sur $\mathbb{R}^n$. Les composantes de $\nabla f(a)$ dans la base canonique sont les dérivées partielles $\frac{\partial f}{\partial x_1}(a), ..., \frac{\partial f}{\partial x_n}(a)$.

Remarque 2

L'hypothèse d'existence de dérivées partielles ne suffit pas à assurer la différentiabilité : plus de régularité sur la fonction est nécessaire.

Théorème 1

Soit $f : U \to \mathbb{R}^p$. Si toutes les dérivées partielles de f sur U existent et si elles sont continues *en un point a de U, alors f est différentiable en a et on a :*

$$Df(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) dx_i$$

Démonstration : Gourdon, théorème 1p305 ■

Remarque 3

La réciproque est fausse. Par exemple, l'application $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2 \sin(\frac{1}{x})$ si $x \neq 0$ et $f(0) = 0$, est différentiable en 0 mais f' n'est pas continue en 0.

Définition 5

Soit $f : U \to \mathbb{R}^m$ différentiable en $a \in U$. Soient $(f_1, ..., f_n)$ les composantes de f dans la base canonique. Alors, la matrice de $Df(a)$ dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^m$ est :

$$J_a = \left[\frac{\partial f_i}{\partial x_j}(a) \right]_{1 \leq i \leq m, 1 \leq j \leq n} \in M_{m,n}(\mathbb{R})$$

On l'appelle matrice jacobienne de f en a. Lorsque $m = n$, J_a est une matrice carrée et son déterminant est appelé jacobien de f en a.

Proposition 3

Soient deux applications $f : U \to \mathbb{R}$ et $\varphi : V \subset \mathbb{R}^m \to \mathbb{R}^n$ telle que $\varphi(V) \subset U$. On écrit φ sous la forme $\varphi = (\varphi_1, ..., \varphi_n)$ où $\varphi_i : V \to \mathbb{R}$. Soit $a \in V$ tel que φ est différentiable en a et f est différentiable en $\varphi(a)$. Alors l'application $F = f \circ \varphi : V \to \mathbb{R}$ est différentiable en a et :

$$\forall j \in \{1, ..., m\}, \frac{\partial F}{\partial u_j}(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\varphi(a)).\frac{\partial \varphi_i}{\partial u_j}(a)$$

Application : Calcul du Laplacien en coordonnées polaires

$$\triangle f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 F}{\partial r^2} + \frac{1}{r}\frac{\partial F}{\partial r} + \frac{1}{r^2}\frac{\partial^2 F}{\partial \theta^2}$$

1.3 Plan tangent à une surface

[2] voir Rouvière, exemple 5p52 et exemple 6p53

2. Cette partie est optionnelle... Cependant, il est important d'en connaître son contenu.

2 Le théorème des accroissements finis

Théorème 2 (***Théorème des accroissements finis***)

Soit $f : U \to \mathbb{R}^p$ une application différentiable en tout point de U. Soit $[a,b] \subset U$, et k une constante positive. On suppose que pour tout $x \in [a,b]$, la norme d'application linéaire de la différentielle $Df(x)$ vérifie l'inégalité

$$\|Df(x)\| \leq k$$

Alors, on a l'inégalité de la moyenne :

$$\|f(b) - f(a)\| \leq k\|b - a\|$$

Démonstration : Rouvière, théorème 3.1p104 ■

Corollaire 1

Soit $f : U \to \mathbb{R}^p$ différentiable sur U.

1. *Si U est un ouvert convexe et si $\|Df(x)\| \leq k$ pour tout $x \in U$ (où k est une constante positive), alors l'application f est k-lipschitzienne sur U, i.e :*

$$\|f(x) - f(y)\| \leq k\|x - y\|$$

   pour tous $x, y \in U$.

2. *Si U est un ouvert connexe et si $Df(x) = 0$ pour tout $x \in U$, alors f est constante sur U.*

Démonstration : Rouvière, corollaire 3.1p106 ■

3 Théorèmes d'inversion

Théorème 3 (***Théorème d'inversion locale***)

Soit $a \in U$ et $f : U \to \mathbb{R}^n$ de classe C^1. On suppose que la matrice jacobienne $Df(a)$ est inversible, i.e $det(Df(a)) \neq 0$.
Il existe alors un ouvert V contenant a (et contenu dans U) et un ouvert W contenant $b = f(a)$, tels que f (restreinte à V) soit un difféomorphisme de classe C^1 de V sur $W = f(V)$

Démonstration : Rouvière, théorème 5.1p188 pour l'énoncé et exercice 71p222 pour une preuve ■

Théorème 4 (***Théorème d'inversion globale***)

Soit $f : U \to \mathbb{R}^n$ de classe C^1. On suppose que f est injective sur U et que pour tout $x \in U$, la matrice jacobienne $Df(x)$ est inversible, i.e $det(Df(x)) \neq 0$.
Alors l'ensemble image $f(U)$ est un ouvert de $\mathbb{R}^n$ et f est un difféomorphisme de classe C^1 de U sur $f(U)$.

Démonstration : Rouvière, théorème 5.3p190 ■

Application :

1. Développement : Surjectivité de l'exponentielle de $M_n(\mathbb{C})$ dans $GL_n(\mathbb{C})$

Théorème 5 (*Surjectivité de l'exponentielle*)

Soit $A \in GL_n(\mathbb{C})$, alors il existe $P \in \mathbb{C}[X]$ tel que $A = exp(P(A))$. En particulier, l'application exponentielle de $M_n(\mathbb{C})$ dans $Gl_n(\mathbb{C})$ est surjective.

Démonstration : On considère la sous-algèbre $\mathbb{C}[A]$ de $M_n(\mathbb{C})$ engendrée par A. C'est une algèbre commutative de dimension finie sur $\mathbb{C}$, isomorphe à $\mathbb{C}[X]/(\Pi)$ avec Π le polynôme minimal de A. Comme $\mathbb{C}[A]$ est un sous-espace vectoriel de dimension finie, il est fermé dans $M_n(\mathbb{C})$ et donc pour $M \in \mathbb{C}[A]$, $exp(M)$ étant la limite d'une suite de polynômes en M (et donc en A), on $M \in \mathbb{C}[A]$. De plus, on a également $exp(M)^{-1} \in \mathbb{C}[A]$ car $\exp(M)^{-1} = exp(-M)$.
Comme $\mathbb{C}[A]$ est une algèbre commutative, on en déduit que exp réalise un morphisme de groupes de $(\mathbb{C}[A], +)$ dans $(\mathbb{C}[A]^{\times}, .)$.
On va maintenant montrer que l'image $exp(\mathbb{C}[A])$ est un ouvert fermé de $\mathbb{C}[A]^{\times}$. Pour $H \in \mathbb{C}[A]$, on a :

$$exp(0 + H) = I_n + H + o(H^2)$$

Donc, exp est différentiable en $0 \in \mathbb{C}[A]$ et sa différentielle en 0 est l'identité qui est inversible. On peut donc appliquer le théorème d'inversion locale : il existe un voisinage ouvert V_0 de 0 et V un voisinage ouvert de l'identité, tels que exp réalise un difféomorphisme de V_0 dans V.

(a) $exp(\mathbb{C}[A])$ **est ouvert dans** $(\mathbb{C}[A]^{\times}, .)$ **:** Soit $M \in \mathbb{C}[A]^{\times}$, comme $In \in V$, on a $M \in MV \subset exp(\mathbb{C}[A]) \subset \mathbb{C}[A]$. Par continuité de $(M, M') \mapsto MM'$, il vient que MV est un voisinage ouvert de M, inclus dans $exp(\mathbb{C}[A])$. Donc, $exp(\mathbb{C}[A])$ est un voisinage ouvert de chacun de ses points, donc c'est un ouvert.

(b) $exp(\mathbb{C}[A])$ **est fermé dans** $(\mathbb{C}[A]^{\times}, .)$ **:** On peut écrire

$$^{c}exp(\mathbb{C}[A]) = \bigcup_{M \notin exp(\mathbb{C}[A])} MV$$

Comme MV est ouvert pour tout M, $^{c}exp(\mathbb{C}[A])$ est ouvert, et le résultat est montré.

Il nous reste à montrer que $\mathbb{C}[A]^{\times}$ est connexe pour conclure.
Soient M et N dans $\mathbb{C}[A]^{\times}$. Soit $P(X) = det(XM + (1-N)X)$, on note $\mathbb{C}_{M,N} = \mathbb{C} \setminus \{\text{zéros de } P\}$. L'ensemble des zéros de P étant en nombre fini, $\mathbb{C}_{M,N}$ est un ensemble connexe. L'image de $\mathbb{C}_{M,N}$ par l'application continue $z \mapsto zM + (1-z)N$ est donc une connexe de $\mathbb{C}[A]^{\times}$ contenant M et N, donc $\mathbb{C}[A]^{\times}$ est connexe.
En conclusion, $exp(\mathbb{C}[A])$ est un ouvert-fermé de $\mathbb{C}[A]^{\times}$ qui est connexe, donc $exp(\mathbb{C}[A]) = \mathbb{C}[A]^{\times}$. ■

2. Développement : Théorème de Hadamard-Lévy

Théorème 6 (*Hadamard-Lévy*)

Soit f une application de classe C^1 de $\mathbb{R}^n$ dans lui même. Les propriétés suivantes sont équivalentes :

(a) f est un difféomorphisme de $\mathbb{R}^n$ sur $\mathbb{R}^n$.

(b) f est propre et, pour tout $x \in \mathbb{R}^n$, la matrice jacobienne $Df(x)$ est inversible.

Une application est dite propre si l'image réciproque de tout compact est compacte. Lorsque $\mathbb{R}^n$ est l'espace de départ et d'arrivée, cela revient à dire que $\| f(x) \|$ tend vers l'infini quand $\| x \|$ tend vers l'infini.

Démonstration : La preuve de ce résultat n'est pas simple, cependant si f est de classe C^2, il y a une preuve abordable que l'on présente ici. On suppose donc désormais que f est de classe C^2.

- 2) $\Rightarrow$ 1) : f^{-1} étant continue, elle envoie un compact sur un compact donc f est propre. Ensuite, comme $f^{-1} \circ f = Id$, on a $d(f^{-1})(f(x)) \circ df(x) = Id$, donc df est inversible en tout point.
- (a) Montrons que f est surjective.
  Pour cela on montre que $f(\mathbb{R}^n)$ est un ouvert/fermé de $\mathbb{R}^n$ et donc que $f(\mathbb{R}^n) = \mathbb{R}^n$ par connexité de $\mathbb{R}^n$.
  Soit $y_0 \in f(\mathbb{R}^n)$ et $x_0 \in \mathbb{R}^n$ tel que $f(x_0) = y_0$. D'après le théorème d'inversion locale, f est un difféomorphisme d'un voisinage V de x_0 sur un voisinage W de y_0, alors $W = f(V) \subset f(\mathbb{R}^n)$, donc $f(\mathbb{R}^n$ est ouvert.
  Soit $(y_k)_{k\in\mathbb{N}} \in f(\mathbb{R}^n)^{\mathbb{N}}$ telle que $y_k \to y \in \mathbb{R}^n$ quand $k \to \infty$. Posons $K = \{y_k, k \in \mathbb{N}\} \cup \{y\}$ qui est un compact de $\mathbb{R}^n$. Soit $(x_k)_{k\in\mathbb{N}}$ tel que $f(x_k) = y_k$ pour tout k, on a $x_k \in f^{-1}(K)$ qui est un compact puisque f est propre. Il existe donc une sous-suite convergente $x_{k_i} \to x$ dans $\mathbb{R}^n$ quand $i \to \infty$. Alors par continuité de f, $y_{k_i} = f(x_{k_i}) \to f(x) \in f(\mathbb{R}^n)$ d'où par unicité de la limite $y = f(x) \in f(\mathbb{R}^n)$, ce qui montre que $f(\mathbb{R}^n)$ est fermé.

  (b) Montrons que f est injective.
  C'est ici que l'on va utiliser l'hypothèse $f \in C^2$.
  Soit $x_0 \in \mathbb{R}^n$. On considère la fonction $g(x) = f(x) - f(x_0)$. Elle vérifie les deux conditions de 2) à savoir g propre et le déterminant du jacobien est partout non nul. Il nous faut montrer que l'ensemble $S = \{x/g(x) = 0\}$ est réduit à un point $\{x_0\}$.

  i. S a un nombre fini d'éléments.
  En effet, S est compact car g est propre ($S = g^{-1}(\{0\})$). Si S avait un nombre infini d'éléments (x_k), il y aurait dans S un point d'accumulation q. Comme $\mid Jac_q(g) \mid \neq 0$, il existe un voisinage V de q tel que g soit un difféomorphisme de V sur $h(V)$. En particulier, g est injective. Or, dans V, il y a au moins un $x_k \neq q$ et on a $g(x_k) = g(q) = 0$ ce qui contredit l'injectivité de g.

  ii. Notons $S = \{p_1, ..., p_N\}$. On va montrer que $N = 1$.
  On considère la fonction
  $$F(x) = [dg(x)]^{-1} g(x)$$
  Cette fonction est bien définie et de classe C^1 sur $\mathbb{R}^n$ par hypothèse sur f. On introduit alors le système différentiel suivant :
  $$\begin{cases} \dot{x}(t) & = -F(x(t)) \\ x(0) & = q \end{cases}$$
  où $q \in \mathbb{R}^n$. On appelle alors trajectoire issue de q une solution du problème de Cauchy ci-dessus. Puisque $F \in C^1$, ce problème admet une solution maximale définie sur $[0, T^*[$. Soit x une telle solution.
  Assertion 1 : $T^* = +\infty$.
  En effet, pour $t \in [0, T^*($, on a :
  $$\begin{aligned} \frac{d}{dt}[g(x(t))] &= dg(x(t))\dot{x}(t) \\ &= -dg(x(t))(dg(x(t)))^{-1} g(x(t)) \\ &= -g(x(t)) \end{aligned}$$
  Donc, $g(x(t)) = e^{-t} g(q) \leq g(q)$ car $t \geq 0$. Donc, $x(t) \in g^{-1}([-g(q), g(q)])$ qui est un compact car g est propre. Donc, T^* d'après le lemme de sortie de tout compact (ou théorème de majoration a priori).

Assertion 2 : Chaque p_i est un point asymptotiquement stable. Autrement dit, $F(p_i) = 0$ et $\exists\delta > 0$ tel que si x est une trajectoire issue de q telle que $\mid q - p_i \mid \delta$ alors $\lim_{t\to+\inf} x(t) = p_i$.
On a tout d'abord $F(p_i) = [dg(p_i)]^{-1}g(p_i) = 0$ car $g(p_i) = 0$.
Ensuite, g est un difféomorphisme d'une boule $B(p_i, \delta)$ sur un voisinage V de 0. Soit $\varepsilon > 0$ tel que $B(0, \varepsilon) \subset V$, $y \in B(0, \varepsilon)$ et posons $q = g^{-1}(y) \in B(p_i, \delta)$. La fonction $x(t) = g^{-1}(e^{-t}q)$ est la trajectoire issue de q. En effet, $x(0) = g^{-1}(y) = q$ et

$$\begin{aligned}\dot{x}(t) &= d(g^{-1})(e^{-t}y)(-e^{-t}y) \\ &= -d(g^{-1})(g(x(t))g(x(t)) \\ &= -[dg(x(t))]^{-1}g(x(t)) \\ &= -F(x(t))\end{aligned}$$

C'est donc par unicité la trajectoire issue de q. Mais alors, $\lim_{t\to+\infty} x(t) = \lim_{t\to+\infty} g^{-1}(e^{-t}y) = g^{-1}(0) = p_i$.
Soit W_i l'ensemble des points $q \in \mathbb{R}^n$ tel que la trajectoire issue de q converge vers p_i quand $t \to +\infty$.
Assertion 3 : $\mathbb{R}^n = \bigcup_{i=1}^N W_i$
Soit $q \in \mathbb{R}^n$ et $x(t)$ la trajectoire issue de q. On a vu dans l'assertion 1, que $x(t)$ reste dans un compact pour tout $t \geq 0$. Il existe donc $t_k \to +\infty$ telle que $x(t_k)$ converge vers l. Mais donc $g(l) = 0$ (car $g(x(t)) = e^{-t}g(q)$), donc $l = p_i$ pour un certain i.
Fixons k_0 assez grand pour que $x(t_k) \in B(p_i, \delta)$ où δ est défini dans l'assertion 2. Alors la trajectoire $y(t)$ issue de $x(t_{k_0})$ converge vers p_i. Or, la fonction $z(t) = x(t + t_{k_0})$ est aussi trajectoire issue de $x(t_{k_0})$. Par unicité globale, on a $x(t + t_{k_0}) = y(t)$ et donc $x(t) \to p_i$ lorsque $t \to +\infty$.
Assertion 4 : Chaque W_i est ouvert.
Soit $q \in W_i$ et $x(t)$ la trajectoire issue de q. Il existe par définition de W_i, $T > 0$ tel que $\mid x(T) - p_i \mid\leq \delta/2$. Soit y la trajectoire issue de $q' \in \mathbb{R}^n$. Par continuité des solutions par rapport aux conditions initiales, il existe $\varepsilon > 0$ tel que $\mid q - q' \mid\leq \varepsilon$ implique $\mid x(T) - y(T) \mid\leq \delta/2$. Alors

$$\mid y(T) - p_i \mid\leq\mid y(T) - x(T) \mid + \mid x(T) - p_i \mid\leq \delta$$

Il résulte de l'assertion 2 que $y(t)$ converge vers p_i lorsque $t \to +\infty$, *i.e* $q' \in W_i$ et $B(q, \varepsilon) \subset W_i$ donc W_i est ouvert.
Assertion 5 : $N = 1$
Chaque W_i est un ouvert non vide et $W_i \cap W_j = \emptyset$ si $i \neq j$ (si $N \geq 2$), et $\mathbb{R}^n = \cup_{i=1}^N W_i$. Si $N \geq 2$, cela contredit la connexité de $\mathbb{R}^n$. Donc, $N = 1$! ■

Théorème 7 (***Théorème des fonctions implicites***)

Soient U un ouvert de $\mathbb{R}^n \times \mathbb{R}^p$, (a, b) un point de U, et $f : (x, y) \mapsto f(x, y)$ une application de classe C^1 de U dans $\mathbb{R}^p$. On suppose que $f(a, b) = 0$ et que la matrice jacobienne $D_y f(a, b)$, formée des dérivées partielles par rapport à y, est inversible, i.e $det(D_y f(a, b)) \neq 0$.
Alors l'équation $f(x, y) = 0$ peut être résolue localement par rapport aux variables y : il existe V (voisinage ouvert de a dans $\mathbb{R}^n$, W (voisinage ouvert de b dans $\mathbb{R}^p$), avec $V \times W \subset U$ et $D_y f(x, y)$ inversible pour tout $x, y \in V \times W$, et une application $\varphi : V \to W$ telle que :

$$(x \in V, y \in W, \text{ et } f(x, y) = 0) \Leftrightarrow (x \in V, \text{ et } y = \varphi(x))$$

De plus, φ est de classe C^1 sur V.

Démonstration : Rouvière, théorème 5.7p192 pour l'énoncé et exercice 86p259 pour une preuve ■

Remarque 4

On peut remplacer l'hypothèse fC^1 par fC^k, et on obtient des conclusions C^k (C^k difféomorphisme dans les théorèmes d'inversion, et φ C^k dans le théorème des fonctions implicites).

4 Dérivées d'ordre supérieur

4.1 Définitions

Définition 6

Soit $f : U \to \mathbb{R}^p$ une fonction diffférentiable sur U et $a \in U$. On dit que f est deux fois différentiable en a si l'application Df est différentiable en a. On note alors $D^2f(a) = D(Df)(a)$ la différentielle seconde ainsi obtenue.

Définition 7

Si $f : U\mathbb{R} \to \mathbb{R}$ est deux fois différentiable en $a \in U$, la matrice de $D^2f(a)$ est appelée matrice Hessienne de f en a.

Cette matrice est symétrique :

Théorème 8 (*Théorème de Schwarz*)

Soit $f : U \to \mathbb{R}^p$, supposée deux fois différentiable en $a \in U$. Alors :

$$\forall 1 \leq i, j \leq n, \frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$$

Autrement dit, $D^2f(a)$ est une application bilinéaire symétrique.

Démonstration : Rouvière, théorème 6.1p294 pour l'énoncé et Gourdon, théorème 2p306 pour une preuve sous des hypothèses légèrement differentes. ■

Remarque 5

L'existence des $\frac{\partial^2 f}{\partial x_i \partial x_j}$ ne suffirait pas. La fonction

$$f(x, y) = xy\frac{x^2 - y^2}{x^2 + y^2}, f(0, 0) = 0$$

est de classe C^1 sur $\mathbb{R}^2$ et admet des dérivées $\frac{\partial^2 f}{\partial x_i \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_i}$ en tout point, mais distinctes à l'origine.

Démonstration : Rouvière, p296 ■

Définition 8

Pour $k \geq 1$, l'application f est dite de classe C^k sur l'ouvert U de $\mathbb{R}^n$ si ses dérivées partielles premières sont de classe C^{k-1} sur U.
f est dite de classe C^∞ sur U si elle est de classe C^k pour tout k.

4.2 Formules de Taylor

Théorème 9 (*Formule de Taylor-Young*)

Si f est k-fois différentiable en $a \in U$, on a :

$$f(a+h) = f(a) + Df(a)h + ... + \frac{1}{k!}D^k f(a)(h)^k + o(\|h\|^k)$$

Démonstration : Rouvière, théorème 6.2p297 ■

Théorème 10 (*Formule de Taylor avec reste intégral*)

Si f est de classe C^{k+1} sur U et si le segment $[a, a+h]$ est tout entier contenu dans U, on a :

$$f(a+h) = f(a) + Df(a)h + ... + \frac{1}{k!}D^k f(a)(h)^k + \int_0^1 \frac{(1-t)^k}{k!} D^{k+1} f(a+th)(h)^{k+1} dt$$

En particulier, pour tout compact convexe contenu dans U, il existe une constante $C > 0$ telle que, pour $a \in K$ et $a+h \in K$:

$$\|f(a+h) - f(a) - Df(a)h - ... - \frac{1}{k!}D^k f(a)(h)^k\| \leq C\|h\|^{k+1}$$

Démonstration : Rouvière, théorème 6.3p298 ■

Application :
DÉVELOPPEMENT : Lemme de Morse

Lemme 1 (*Lemme de Morse*)

Soit $f : U \to \mathbb{R}^3$ une fonction de classe C^3 sur un ouvert U de $\mathbb{R}^n$ contenant l'origine. On suppose que $Df(0) = 0$ et que la forme quadratique associée à la matrice hessienne en 0, $D^2 f(0)$, est non dégénérée, de signature $(p, n-p)$. Alors il existe un C^1-difféomorphisme φ entre deux voisinages de l'origine dans $\mathbb{R}^n$, de classe C^1, tel que $\varphi(0) = 0$ et, en posant $u = \varphi(x)$:

$$f(x) - f(0) = u_1^2 + ... + u_p^2 - u_{p+1}^2 - ... - u_n^2$$

On a d'abord besoin de la proposition suivante :

Proposition 4

On note E l'espace des matrices réelles de taille n à coefficients dans $\mathbb{R}$ et S le sous-espace des matrices symétriques. On fixe $A_0 \in S$ inversible. Soit $\varphi : E \to S$ l'application définie par

$$\varphi(M) = {}^t M A_0 M$$

Alors, il existe un voisinage V de A_0 dans S, et une application $A \mapsto= \psi(A) = M$ de V dans l'ensemble des matrices inversibles, de classe C^1, telle que,

$$A = {}^t\psi(A) A_0 \psi(A)$$

pour tout $a \in V$.

Démonstration : L'application φ est polynomiale donc de classe C^1 sur E. Pour $H \in E$, on a en munissant E et S d'une norme d'application linéaire :

$$\begin{aligned}\varphi(I+h) - \varphi(I) &= {}^tHA_0 + A_0H + {}^tHA_0H \\ &= {}^t(A_0)H + (A_0H) + \bigcirc(\|H\|^2)\end{aligned}$$

Par suite,

$$D\varphi(I)H = {}^t(A_0H) + A_0H$$

Le noyau de l'application linéaire $D\varphi(I) : E \to S$ est donc formé des matrices H telles que A_0H soit antisymétriques. Le noyau $D\varphi(I)$ a même dimension que l'espace des matrices antisymétriques de taille n, c'est à dire $n(n-1)/2$, la dimension de son image est donc $dim(E) - n(n-1)/2 = n(n+1)/2 = dim(S)$. Donc $D\varphi(I)$ est surjective.
Toute matrice étant de manière unique somme d'une symétrique et d'une antisymétrique, le sous-espace F de E formé des matrices M telles que $A_0M \in S$ est un supplémentaire du noyau de $D\varphi(I)$ dans E, de plus I appartient à F.
Soit $\Phi : F \to S$ la restriction de φ à F. La différentielle $D\Phi(I)$ restriction à F de $D\varphi(I)$ est bijective puisque $Ker(D\varphi(I) \cap F = \{0\}$. Par le théorème d'inversion locale, il existe un voisinage ouvert U de I dans F (que l'on peut supposer contenu dans l'ouvert des matrices inversibles) tel que Φ soit un difféomorphisme de classe C^1 de U sur $V = \Phi(U)$. Ainsi, V est un voisinage ouvert $A_0 = \Phi(I) = \varphi(I)$ dans S, et pour tout $A \in V$, il existe une unique matrice inversible $M \in U$ telle que

$$A = {}^tMA_0M$$

et $M = \phi^{-1}(A) = \psi(A)$ est fonction continument différentiable de A. ■

Le résultat obtenu est que toute forme quadratique suffisamment voisine d'une forme quadratique non dégénérée lui est équivalente, c'est à dire se ramène à celle-ci par changement de base.
On est maintenant en mesure de démontrer le lemme de Morse.

Démonstration : La formule de Taylor avec reste intégral à l'ordre un s'écrit au voisinage de 0

$$f(x) - f(0) = {}^txQ(x)x$$

où $Q(x)$ est la matrice symétrique

$$Q(x) = \int_0^1 (1-t)D^2f(tx)dt$$

fonction C^1 de x. D'après la proposition précédente, il existe une matrice inversible $M(x)$, fonction C^1 de x au voisinage de 0 dans $\mathbb{R}^n$, telle que

$$Q(x) = {}^tM(x)Q(0)M(x)$$

d'où

$$f(x) - f(0) = {}^tyQ(0)y \text{ avec } y = M(x)x$$

Or, $Q(0) = (1/2)D^2f(0)$ est de signature $(p, n-p)$, et il existe donc une changement linéaire de coordonnées $y = Au$, où A est une matrice inversible, tel que

$${}^tyQ(0)y = {}^tu{}^tAQ(0)Au = u_1^2 + ... + u_p^2 - up + 1^2 - ... - u_n^2$$

ce qui donne à f l'expression voulue.
Enfin, l'application $x \mapsto u = A^{-1}M(x)$ a pour différentielle à l'origine $A^{-1}M(0)$ matrice inversible. D'après le théorème d'inversion locale, c'est un difféomorphisme de classe C^1 entre deux voisinages de l'origine dans $\mathbb{R}^n$. ■

5 Problèmes d'extremum

Définition 9

Soit X une partie de $\mathbb{R}^n$, a un point de X et $f : X \to \mathbb{R}$ une fonction numérique.

1. *On dit que f admet en a un maximum global si $f(x) \leq f(a)$ pour tout $x \in X$.*
2. *On dit que f admet en a un maximum local s'il existe un voisinage V de a dans $\mathbb{R}^n$ tel que $f(x) \leq f(a)$ pour tout $x \in V \cap X$.*
3. *On dit que f admet en a un maximum strict si les inégalités précédentes sont strictes pour $x \neq a$.*

Les définitions correspondantes d'un minimum global, resp. local, resp. strict, s'en déduisent en renversant le sens des inégalités. Le terme d'extremum signifie maximum ou minimum.

Théorème 11

Soit $f : U \to \mathbb{R}$ une fonction numérique.

1. Condition nécessaire (pas suffisante) du premier ordre : *Si f admet en a un extremum local et si $Df(a)$ existe , alors nécessairement a est un point critique de f, i.e $Df(a) = 0$.*
2. Condition nécessaire (pas suffisante) du second ordre : *Si f admet en a un minimum local et si $D^2f(a)$ existe, alors nécessairement $Df(a) = 0$ et $D^2f(a)$ est une forme quadratique positive, i.e $D^2f(a)(h,h) \geq 0$ pour tout vecteur $h \in \mathbb{R}^n$.*
3. Condition suffisante (pas nécessaire) du second ordre : *Si $Df(a) = 0$ et si $D^2f(a)$ est une forme quadratique définie positive (i.e $D^2f(a)(h,h) > 0$ pour tout vecteur $h \neq 0$), alors f admet en a un minimum local strict.*

Démonstration : Rouvière, théorème 7.1p371 ■

Théorème 12 (*Extrema liés*)

Soient $f, g_1, ..., g_r : U \subset \mathbb{R}^n \to \mathbb{R}$ des fonctions de classe C^1, où U est un ouvert de $\mathbb{R}^n$. On désigne par Γ l'ensemble $\{x \in U / g_1(x) = ... = g_r(x) = 0\} := \Gamma$. Si $f_{|\Gamma}$ (restriction de f à Γ) admet un extremum relatif en $a \in \Gamma$ et si les formes linéaires $dg_{1,a}, ..., dg_{r,a}$ sont linéairement indépendantes, alors il existe des réels $\lambda_1, ..., \lambda_r$ (appelés multiplicateurs de Lagrange) tels que

$$df_a = \lambda_1 dg_{1,a} + ... + \lambda_r dg_{r,a}$$

Démonstration : Gourdon, théorème 2p317 pour l'énoncé et p327 pour la preuve ■

On peut trouver dans Rouvière, p373, une interprétation géométrique de ce théorème.