Calcul Différentiel Avancé

par

Emmanuel Hebey

Université de Cergy-Pontoise

Année 2017-2018

1. Applications continues

L'espace $\mathbb{R}^n$ est un ensemble sur lequel on peut placer une norme (la norme euclidienne, associée à sa structure d'espace vectoriel) et une distance (la distance euclidienne, associée à la norme euclidienne). Si $x = (x_1, \dots, x_n)$ est un point de $\mathbb{R}^n$ la norme euclidienne $\|x\|$ de x est donnée par

$$\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2} \; .$$

La distance d associée est donnée par : si $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ sont deux points de $\mathbb{R}^n$, alors

$$d(x, y) = \|y - x\| = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2} \; .$$

Dès lors qu'on a une distance on a une topologie associée (des ouverts, des fermés, des compacts etc.). Les ouverts de $\mathbb{R}^n$ sont les $\Omega \subset \mathbb{R}^n$ qui vérifient la propriété suivante : *pour tout point de* Ω *il existe une boule centrée en ce point qui est entièrement contenue dans* Ω. Donc :

$$\forall x \in \Omega, \exists r_x > 0 \text{ t.q. } B_x(r_x) \subset \Omega \ ,$$

où $B_x(r_x) = \{y \in \mathbb{R}^n \ / \ d(x, y) < r_x\}$. Les fermés de $\mathbb{R}^n$ sont les complémentaires des ouverts. Donc un sous ensemble $F \subset \mathbb{R}^n$ est un fermé si et seulement si $\Omega = \mathbb{R}^n \backslash F$ est un ouvert.

Intuitivement : les ouverts sont les patatoïdes qui ne contiennent aucun point de leur frontière et les fermés sont les patatoïdes qui contiennent tous les points de leur frontière). Un ensemble peut très bien n'être ni ouvert ni fermé. Par convention l'ensemble vide $\emptyset$ et $\mathbb{R}^n$ sont à la fois ouverts et fermés.

Un sous ensemble $K \subset \mathbb{R}^n$ est compact si de tout recouvrement $(\Omega_i)_I$ de K par des ouverts (à savoir $K \subset \bigcup_{i \in I} \Omega_i$) on peut extraire un sous recouvrement fini (à savoir $K \subset \bigcup_{i \in J} \Omega_i$, où $J \subset I$ est fini). On a alors

Théorème (Compacts de $\mathbb{R}^n$)

Un sous ensemble $K \subset \mathbb{R}^n$ est compact si et seulement si toute suite de points de K possède une sous suite convergente dans K. Les compacts de $\mathbb{R}^n$ sont précisément les fermés bornés de $\mathbb{R}^n$.

On rappelle qu'une suite $(x_n)_n$ d'un espace $\mathbb{R}^p$ converge vers un point x, on écrit $\lim\limits_{n \to +\infty} x_n = x$, si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \ / \ \forall n \geq N, \|x_n - x\| < \varepsilon \ ,$$

où $\| \cdot \|$ est la norme euclidienne de $\mathbb{R}^p$, et donc $(x_n)_n$ converge vers x si x_n est aussi proche que l'on veut de x pourvu que n soit suffisamment grand.

Pour finir avec ses rappels de topologie, un sous ensemble $X \subset \mathbb{R}^n$ est <u>connexe</u> si, par définition première, il ne peut pas s'écrire comme réunion disjointe de deux ouverts non vides. On montre que dans $\mathbb{R}^n$ un sous ensemble ouvert est connexe si et seulement si il est <u>connexe par arcs</u>, et donc si et seulement si deux points quelconques de Ω peuvent être joints par un chemin continu entièrement contenu dans Ω. A savoir :

$$\forall x, y \in \Omega, \exists \gamma : [0,1] \to \Omega, \gamma \text{ continue, avec } \gamma(0) = x \text{ et } \gamma(1) = y$$

Ici la continuité de γ sur $[0,1]$ se traduit par le fait que pour tout $t_0 \in [0,1]$, $\lim_{t \to t_0} \gamma(t) = \gamma(t_0)$ où les t qui tendent vers t_0 dans cette limite restent dans $[0,1]$.

Une première remarque simple : soit $D \subset \mathbb{R}^p$ un sous ensemble de $\mathbb{R}^p$ et soit $f : D \to \mathbb{R}^q$ une application de D dans $\mathbb{R}^q$. En écrivant que $f = (f_1, \ldots, f_q)$, l'application f est entièrement représentée par ses q fonctions coordonnées f_i, $i = 1, \ldots, q$. Les

$$f_i : D \to \mathbb{R}$$

sont des fonctions réelles définies sur D.

Une autre remarque simple : soit $(X_n)_n$ une suite de points d'un espace $\mathbb{R}^p$. En écrivant $X_n = (X_n^1, \ldots, X_n^p)$ pour tout n, la suite $(X_n)_n$ est entièrement représentée par ses p suites coordonnées $(X_n^i)_n$, $i = 1, \ldots, p$. Et bien sûr, si $X = (X_1, \ldots, X_p)$ est un point de $\mathbb{R}^p$ alors

$$\lim_{n\to\infty} X_n = X \text{ si et seulement si } \forall i = 1, \ldots, p, \lim_{n\to+\infty} X_n^i = X^i \ .$$

Ce dernier point est intuitif mais découle mathématiquement du double encadrement

$$\max_{i=1,\ldots,p} |x_i| \leq \|x\| \leq \sqrt{p} \max_{i=1,\ldots,p} |x_i|$$

pour tout $x = (x_1, \ldots, x_p)$ dans $\mathbb{R}^p$ qui entraîne que $\|X_n - X\|$ est petit si et seulement si tous les $|X_n^i - X^i|$ le sont.

Définition

Soient $D \subset \mathbb{R}^p$ un sous ensemble de $\mathbb{R}^p$, $a \in D$ un point de D et $f : D \to \mathbb{R}^q$ une application de D dans $\mathbb{R}^q$. On dit que f est continue au point a si l'une des trois propositions équivalentes suivantes est vérifiée :

(i) $\lim\limits_{x \to a} f(x) = f(a)$,

(ii) $\forall \varepsilon > 0, \exists \eta > 0 \ / \ \forall x \in D, \|x - a\| < \eta \ \Rightarrow \ \|f(x) - f(a)\| < \varepsilon$,

(iii) pour toute suite $(x_n)_n$ de points de D vérifiant que $\lim\limits_{n \to +\infty} x_n = a$, on a forcément que $\lim\limits_{n \to +\infty} f(x_n) = f(a)$,

(iv) les q fonctions coordonnées f_i de f sont toutes continues au point a.

On dit que f est continue sur D lorsque f est continue en tout point de D.

(1) L'équivalence (i) $\Leftrightarrow$ (ii) suit du fait que (ii) n'est rien d'autre que la définition mathématique de (i). L'équivalence (i) $\Leftrightarrow$ (iii) se démontre comme dans le cas réel en remplaçant les valeurs absolues de $\mathbb{R}$ par les normes euclidiennes des espaces $\mathbb{R}^p$ et $\mathbb{R}^q$. L'équivalence (i) $\Leftrightarrow$ (iv) suit de la remarque précédant la définition.
(2) Si f et g sont continues en a, alors $f+g$ est continue en a. Dans le cas $q=1$, fg et $\frac{f}{g}$ si $g(a)\neq 0$ sont aussi continues en a.
(3) Soient $D\subset\mathbb{R}^{n_1}$ et $D'\subset\mathbb{R}^{n_2}$ deux sous ensembles de $\mathbb{R}^{n_1}$ et $\mathbb{R}^{n_2}$, $a\in D$ un point de D, $f: D\to\mathbb{R}^{n_2}$ une application et $g: D'\to\mathbb{R}^{n_3}$ une autre application. On suppose que $Im(f)\subset D'$, où $Im(f)$ est l'ensemble des valeurs prises par f sur D, de sorte que $g\circ f$ est définie sur D. Si f est continue en a et si g est continue en $f(a)$, alors $g\circ f$ est aussi continue en a.

Proposition (Extension d'une des propositions du Chapitre 1)

Soient $D \subset \mathbb{R}^p$ un sous ensemble connexe de $\mathbb{R}^p$, $f : D \to \mathbb{R}$ un fonction réelle continue sur D et $a, b \in D$ deux points de D. On suppose que f change de signe en a et b, et donc que $f(a)f(b) < 0$. Il existe alors $c \in D$ tel que $f(c) = 0$.

Preuve 1 : Comme connexité et connexité par arcs coïncident dans les espaces $\mathbb{R}^p$, D est aussi connexe par arcs. Soit $\gamma : [0, 1] \to D$ un chemin entièrement contenu dans D joignant a à b. On définit $g : [0, 1] \to \mathbb{R}$ par $g(t) = f(\gamma(t))$. Par composition, g est continue sur $[0, 1]$. Par hypothèse g change de signe en 0 et 1. Le Chapitre 1 donne alors l'existence d'un point $t_0 \in]0, 1[$ tel que $g(t_0) = 0$. En posant $c = \gamma(t_0)$ on récupère que $f(c) = 0$. La proposition est démontrée. CQFD

Preuve 2 : On raisonne par l'absurde et on suppose que f ne s'annule jamais sur D. On pose

$$\Omega_1 = \{x \in D \ / \ f(x) < 0\} \quad \text{et} \quad \Omega_2 = \{x \in D \ / \ f(x) > 0\} \ .$$

Alors $\Omega_1 = f^{-1}\left(]-\infty, 0[\right)$ et $\Omega_2 = f^{-1}\left(]0, +\infty[\right)$ de sorte que Ω_1 et Ω_2 sont deux ouverts de D (l'image réciproque d'un ouvert par une application continue est un ouvert). Par hypothèse $\Omega_1 \neq \emptyset$ et $\Omega_2 \neq \emptyset$ puisque $a \in \Omega_1$ et $b \in \Omega_2$ (ou réciproquement). Par hypothèse d'absurde $\Omega_1 \bigcup \Omega_2 = D$. Clairement $\Omega_1 \bigcap \Omega_2 = \emptyset$. Une contradiction avec la connexité de D. CQFD

Théorème

Soient $K \subset \mathbb{R}^p$ un sous ensemble compact de $\mathbb{R}^p$ et $f : K \to \mathbb{R}$ une fonction continue sur K. Alors f est bornée sur K et elle y atteint ses bornes.

Preuve : Elle est identique à la preuve du théorème correspondant du Chapitre 1. On prend une suite $(x_n)_n$ de points de K telle que $f(x_n) \to m$ lorsque $n \to +\infty$, où $m = \inf_K f$ (ou alors aussi $m = \sup_K f$). Par compacité cette suite possède une sous suite convergente dans K. Si $\overline{x}$ est la limite de la sous suite, et par continuité de f, on récupère $f(\overline{x}) = m$, ce qui clôt la preuve du théorème. CQFD.

Même remarque qu'au Chapitre 1, un autre résultat fréquemment associé à la compacité est que toute fonction continue sur un compact y est uniformément continue.

On cherche à définir la notion de "dérivabilité" d'une fonction $f : \mathbb{R}^p \to \mathbb{R}^q$. Comme on ne peut plus diviser par des vecteurs on perd la notion même du quotient

$$?? \ \frac{f(x) - f(a)}{x - a} \ ??$$

qui ne fait plus sens. Soient Ω un ouvert de $\mathbb{R}^p$, $a = (a_1, \ldots, a_p)$ un point de Ω, et $f : \Omega \to \mathbb{R}$ une fonction réelle définie sur Ω. On considère les p fonctions réelles f_a^i, $i = 1, \ldots, p$, définies par

$$f_a^i(t) = f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_p) \ .$$

En d'autres termes,

$$\begin{aligned} f_a^1(t) &= f(t, a_2, \ldots, a_p), \\ f_a^2(t) &= f(a_1, t, a_3, \ldots, a_p), \\ &\ldots, \\ f_a^p(t) &= f(a_1, \ldots, a_{p-1}, t) \ . \end{aligned}$$

Sachant que Ω est un ouvert de $\mathbb{R}^p$, il existe un $\epsilon > 0$ pour lequel

$B_a(\epsilon) \subset \Omega$. Il s'ensuit que pour tout $i = 1, \ldots, p$, f_a^i est définie sur au moins l'intervalle ouvert $]a_i - \varepsilon, a_i + \varepsilon[$.

Définition

Soient Ω un ouvert de $\mathbb{R}^p$, $a = (a_1, \ldots, a_p)$ un point de Ω, et $f : \Omega \to \mathbb{R}$ une fonction réelle définie sur Ω. On dit que f admet une dérivée partielle par rapport à x_i au point a si la fonction f_a^i est dérivable au point a_i. On note alors $\frac{\partial f}{\partial x_i}(a)$ la dérivée de f_a^i en a_i, et

$$\frac{\partial f}{\partial x_i}(a) = (f_a^i)'(a_i)$$

s'appelle la dérivée partielle de f par rapport à x_i au point a.

Dire que $f : \Omega \to \mathbb{R}^q$ est différentiable en a c'est dire que les dérivées partielles des f_j en a existent toutes et que

$$f(x) = f(a) + Df(a).(x - a) + o(x - a)$$

où $o(x - a)$ est une application à valeurs $\mathbb{R}^q$ telle que $\frac{o(x-a)}{\|x-a\|} \to 0$ lorsque $x \to a$, et où $Df(a) \in L(\mathbb{R}^p, \mathbb{R}^q)$ est comme ci-après.

Définition

Soient Ω *un ouvert de* $\mathbb{R}^p$, $a = (a_1, \ldots, a_p)$ *un point de* Ω, *et* $f : \Omega \to \mathbb{R}^q$, $f = (f_1, \ldots, f_q)$, *une application définie sur* Ω. *On dit que* f *est différentiable au point* a *si*

(i) Pour tout $i = 1, \ldots, p$ *et tout* $j = 1, \ldots, q$ *les* $\frac{\partial f_j}{\partial x_i}(a)$ *existent ,*

(ii) $\forall \varepsilon > 0$, $\exists \eta > 0 \,/\, \forall x \in \Omega \setminus \{a\}$, $\|x - a\| < \eta \Rightarrow$

$$\|f(x) - f(a) - Df(a).(x - a)\| < \varepsilon \|x - a\| \;,$$

où $Df(a) \in L(\mathbb{R}^p, \mathbb{R}^q)$ *est l'application linéaire de* $\mathbb{R}^p$ *dans* $\mathbb{R}^q$ *qui est donnée par*

$$Df(a)^j.(X) = \sum_{i=1}^{p} \frac{\partial f_j}{\partial x_i}(a) X^i$$

pour tout $X \in \mathbb{R}^p$ *et tout* $j = 1, \ldots, q$. *On dit que* $Df(a)$ *est la différentielle de* f *en* a. *L'application est dite différentiable sur* Ω *si elle est différentiable en tout point de* Ω.

On peut très bien avoir (i) sans pour autant avoir (ii).

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

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

a toutes ses dérivées partielles en $(0,0)$ mais elle n'est pas différentiable en $(0,0)$. En effet, comme $f(x,0) = 0$ et $f(0,y) = 0$ pour tout x et tout y, les dérivées partielles $\frac{\partial f}{\partial x}(0,0)$ et $\frac{\partial f}{\partial y}(0,0)$ existent et valent 0. Donc (i) est vérifiée par f en $a = (0,0)$. Par contre (ii) équivaut à

$$\forall \varepsilon > 0, \exists \eta > 0 \ / \ \forall (x,y) \in \mathbb{R}^2 \backslash \{a\}, \|(x,y)\| < \eta \\ \Rightarrow \|f(x,y)\| < \varepsilon \|(x,y)\| \ .$$

En écrivant $(x,y) = (r\cos\theta, r\sin\theta)$ on devrait donc avoir que $\frac{1}{r} f(r\cos\theta, r\sin\theta) \to 0$ uniformément par rapport à θ lorsque $r \to 0$. Or $f(r\cos\frac{\pi}{4}, r\sin\frac{\pi}{4}) = \frac{1}{2}$ pour tout $r > 0$.

Définition

La matrice de représentation de $Df(a)$ dans les bases canoniques de $\mathbb{R}^p$ et $\mathbb{R}^q$ est la matrice

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_p}(a) \\ \vdots & & \vdots \\ \frac{\partial f_q}{\partial x_1}(a) & \dots & \frac{\partial f_q}{\partial x_p}(a) \end{pmatrix}$$

Elle est appelée matrice jacobienne de f en a. Son déterminant (lorsque $p = q$) est appelé jacobien de f en a.

Pour mémoire la matrice de représentation d'une application linéaire dans deux bases (une au départ et une à l'arrivée) est la matrice dont les colonnes sont constituées des coordonnées dans la base d'arrivée de l'image par l'application linéaire des vecteurs de la base de départ.

Si $p = q = 1$ on a $Df(a).(x) = f'(a) \times x$ et $(f'(a))$ se regarde comme la matrice jacobienne de f en a.

Proposition

Soient Ω un ouvert de $\mathbb{R}^p$, $a = (a_1, \ldots, a_p)$ un point de Ω, et $f : \Omega \to \mathbb{R}^q$, $f = (f_1, \ldots, f_q)$, une application définie sur Ω. Alors f est différentiable en a si et seulement si les f_j sont différentiables en a pour tout $j = 1, \ldots, q$.

Preuve : L'équivalence des points (i) ne présente pas de difficulté. Pour ce qui est de l'équivalence des points (ii) on remarque que la $j$ème coordonnée de $Y = f(x) - f(a) - Df(a).(x - a)$ vaut $Y_j = f_j(x) - f_j(a) - \sum_{i=1}^{p} \frac{\partial f_j}{\partial x_i}(a)(x_i - a_i)$ et on sait que

$$\max_{j=1,\ldots,q} |Y_j| \leq \|Y\| \leq \sqrt{q} \max_{j=1,\ldots,q} |Y_j| \ .$$

Par suite (ii) est bien équivalent à

$$\forall j = 1, \ldots, q, \forall \varepsilon > 0, \exists \eta > 0 \ / \ \forall x \in \Omega \backslash \{a\}, \|x - a\| < \eta$$
$$\Rightarrow \left| f_j(x) - f_j(a) - \sum_{i=1}^{p} \frac{\partial f_j}{\partial x_i}(a)(x_i - a_i) \right| < \varepsilon \|x - a\| \ .$$

CQFD

Dans la preuve précédente on voit que $Df(a)_j = Df_j(a)$ pour tout $j = 1, \ldots, q$ (reconstruction de la différentielle).

Retour sur la formule

$$Df(a).(x) = f'(a)x$$

pour les $f : \mathbb{R} \to \mathbb{R}$: on véfifie qu'en effet, pour ces fonctions,

$$\begin{aligned}
& f(x) - f(a) - Df(a).(x-a) = o(|x-a|) \\
& \Leftrightarrow f(x) - f(a) - f'(a)(x-a) = o(x-a) \\
& \Leftrightarrow \frac{f(x) - f(a) - f'(a)(x-a)}{x-a} = o(1) \\
& \Leftrightarrow \frac{f(x) - f(a)}{x-a} - f'(a) = o(1) \\
& \Leftrightarrow \lim_{x \to a, x \neq a} \frac{f(x) - f(a)}{x-a} = f'(a)\ .
\end{aligned}$$

Théorème

Une application différentiable en un point est continue en ce point.

Preuve : Il existe $C > 0$ tel que

$$\begin{aligned}\|f(x) - f(a) - Df(a).(x-a)\| &\geq \|f(x) - f(a)\| - \|Df(a).(x-a)\| \\ &\geq \|f(x) - f(a)\| - C\|x-a\|\end{aligned}$$

pour tout $x \in \mathbb{R}^p$. Par suite il est clair que (ii) entraîne que

$$\forall \varepsilon > 0, \exists 0 < \eta < \frac{1}{2}\min(1, \frac{\varepsilon}{C}) \ / \ \forall x \in \Omega, \|x-a\| < \eta$$
$$\Rightarrow \ \|f(x) - f(a)\| < \varepsilon$$

et ceci prouve la continuité de f en a. CQFD

Si f et g sont différentiables en a, alors $f+g$ est différentiable en a et de plus

$$D(f+g)(a) \equiv Df(a) + Dg(a)$$

(égalité entre applications linéaires de $\mathbb{R}^p$ dans $\mathbb{R}^q$). Dans le cas $q=1$, fg et $\frac{f}{g}$ si $g(a) \neq 0$ sont aussi différentiable en a et de plus

$$D(fg)(a) \equiv f(a)Dg(a) + g(a)Df(a)$$

et

$$D\left(\frac{f}{g}\right)(a) \equiv \frac{1}{g(a)^2}\left(g(a)Df(a) - f(a)Dg(a)\right)$$

(là encore il s'agit de deux égalités entre applications linéaires de $\mathbb{R}^p$ dans $\mathbb{R}^q$, $q=1$ ici).

Soient $\Omega \subset \mathbb{R}^{n_1}$ *et* $\Omega' \subset \mathbb{R}^{n_2}$ *deux ouverts de* $\mathbb{R}^{n_1}$ *et* $\mathbb{R}^{n_2}$, $a \in \Omega$ *un point de* Ω, $f : \Omega \to \mathbb{R}^{n_2}$ *une application et* $g : \Omega' \to \mathbb{R}^{n_3}$ *une autre application. On suppose que* $Im(f) \subset \Omega'$, *où* $Im(f)$ *est l'ensemble des valeurs prises par* f *sur* D, *de sorte que* $g \circ f$ *est définie sur* Ω. *Si* f *est différentiable en* a *et si* g *est différentiable en* $f(a)$, *alors* $g \circ f$ *est aussi différentiable en* a *et*

$$D\left(g \circ f\right)(a) \equiv Dg\left(f(a)\right) \circ Df(a)$$

(composition d'applications linéaires). Les matrices jacobiennes se multiplient.

Le théorème des accroissements finis en égalité, sous la forme du Chapitre 1, cesse d'être vraie en général dans le cas vectoriel. Considérons par exemple l'application $f : \mathbb{R} \to \mathbb{R}^2$ définie par $f(x) = (x^2, x^3)$. Soient aussi $a = -1$ et $b = +1$. Clairement f est différentiable sur $\mathbb{R}$ et sa différentielle est donnée par

$$Df(c).h = h(2c, 3c^2)$$

pour tout $h \in \mathbb{R}$. Notons f_1 et f_2 les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ définies par $f_1(x) = x^2$ et $f_2(x) = x^3$. L'égalité $f(b) - f(a) = Df(c).(b - a)$ pour un $c \in]-1, +1[$, si elle était vraie, entraînerait que

$$f_1(+1) - f_1(-1) = 4c \text{ et } \; f_2(+1) - f_2(-1) = 6c^2 \; .$$

Comme $f_1(-1) = f_1(+1)$, la première de ses deux équations donne $c = 0$. La seconde est alors impossible. D'où l'affirmation.

Il faut corriger l'énoncé du théorème (par exemple) sous la forme suivante (nous reviendrons plus en détail sur les accroissements finis lorsque nous traiterons du cas des espaces de Banach). On rappelle qu'un sous ensemble A d'un espace euclidien est convexe si pour tous $x, y \in A$ le segment $[x, y]$ est entièrement contenu dans A.

Théorème

Soient Ω un ouvert convexe de $\mathbb{R}^p$ et $f : \Omega \to \mathbb{R}^q$ une application différentiable sur Ω. On suppose qu'il existe $C > 0$ tel que pour tout $i = 1, \ldots, p$, tout $j = 1, \ldots, q$ et tout $x \in \Omega$, $\left|\frac{\partial f_j}{\partial x_i}(x)\right| \leq C$ (donc que les dérivées partielles de f sont toutes bornées dans Ω). Alors il existe $C' > 0$ telle que pour tous $x, y \in \Omega$,

$$\|f(y) - f(x)\| \leq C'\|y - x\| \ ,$$

et $C' = \sup_{x \in \Omega} \|Df(x)\|$ n'est rien d'autre que le sup sur les $x \in \Omega$ des normes des applications linéaires $Df(x)$.

4. Applications de classe C^1

Les “ennuis” avec la théorie $\mathbb{R}^p \to \mathbb{R}^q$ commencent quand on veut définir la notion d'application C^1 ou encore la notion de dérivée d'ordre supérieur. Le problème (non assumé, contrairement à la théorie “Banach”) est que, contrairement à la théorie $\mathbb{R} \to \mathbb{R}$, Df n'est plus de même nature que f puisque si f est différentiable sur Ω alors

$$Df : \Omega \to L(\mathbb{R}^p, \mathbb{R}^q) \ .$$

Pour contourner cette difficulté on“ triche” un peu en “sortant des définitions du chapeau”.

Définition

Soient Ω un ouvert de $\mathbb{R}^p$ et $f : \Omega \to \mathbb{R}^q$, $f = (f_1, \ldots, f_q)$ une application définie sur Ω. On dit que f est de classe C^1 sur Ω si pour tout $i = 1, \ldots, p$ et tout $j = 1, \ldots, q$ les $\frac{\partial f_j}{\partial x_i}$ existent en tout point de Ω et, en tant que fonctions $\frac{\partial f_j}{\partial x_i} : \Omega \to \mathbb{R}$, sont continues sur Ω.

Cette définition élégante pose d'entrée un problème : il faut démontrer que "$C^1 \Rightarrow$ différentiable", cette affirmation n'allant pas de soit à ce stade.

Preuve : Pour simplifier on suppose que $p = 2$ et $q = 1$. Soit (a, b) un point de $\mathbb{R}^2$ et soit (h, k) suffisamment proche de $(0, 0)$ pour que pour tous $t_1, t_2 \in [0, 1]$, $(a + t_1 h, b + t_2 k) \in \Omega$. Avec la formule des accroissements finis pour les fonctions réelles définies sur $\mathbb{R}$, vue au Chapitre 1, on obtient qu'il existe $\theta_1 \in]0, 1[$ et $\theta_2 \in]0, 1[$ tels que

$$
\begin{aligned}
& f(a + h, b + k) - f(a, b) \\
& = (f(a + h, b + k) - f(a, b + k)) + (f(a, b + k) - f(a, b)) \\
& = h\frac{\partial f}{\partial x}(a + \theta_1 h, b + k) + k\frac{\partial f}{\partial y}(a, b + \theta_2 k)
\end{aligned}
$$

Preuve suite : On en déduit que

$$\left|f(a+h,b+k)-f(a,b)-\frac{\partial f}{\partial x}(a,b)h-\frac{\partial f}{\partial y}(a,b)k\right|$$
$$\leq \left|\frac{\partial f}{\partial x}(a+\theta_1 h,b+k)-\frac{\partial f}{\partial x}(a,b)\right|\times|h|$$
$$+\left|\frac{\partial f}{\partial y}(a,b+\theta_2 k)-\frac{\partial f}{\partial y}(a,b)\right|\times|k|\ .$$

Or les fonctions $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont par hypothèse continues sur Ω. Pour tout $\epsilon > 0$ il existe ainsi $\eta > 0$ tel que si $\|(h,k)\| < \eta$, alors

$$\left|\frac{\partial f}{\partial x}(a+\theta_1 h,b+k)-\frac{\partial f}{\partial x}(a,b)\right| < \frac{\epsilon}{2}\ ,$$
$$\left|\frac{\partial f}{\partial y}(a,b+\theta_2 k)-\frac{\partial f}{\partial y}(a,b)\right| < \frac{\epsilon}{2}\ .$$

Preuve suite et fin : Comme par ailleurs $|h| \leq \|(h,k)\|$ et $|k| \leq \|(h,k)\|$, on obtient que

$$\forall \epsilon > 0,\ \exists \eta > 0 \ / \ 0 < \|(h,k)\| < \eta \ \Rightarrow$$
$$\left| f(a+h,b+k) - f(a,b) - \frac{\partial f}{\partial x}(a,b)h - \frac{\partial f}{\partial y}(a,b)k \right| < \epsilon \|(h,k)\| \ .$$

Il s'ensuit que f est différentiable au point (a,b). CQFD

Si f et g sont C^1 alors $f+g$ est C^1. Dans le cas $q=1$, fg et $\frac{f}{g}$ si $g(a) \neq 0$ sont aussi C^1 dès que f et g le sont. Enfin, si $\Omega \subset \mathbb{R}^{n_1}$ et $\Omega' \subset \mathbb{R}^{n_2}$ sont deux ouverts de $\mathbb{R}^{n_1}$ et $\mathbb{R}^{n_2}$, $f : \Omega \to \mathbb{R}^{n_2}$ est une application de classe C^1 sur Ω, $g : \Omega' \to \mathbb{R}^{n_3}$ est une application de classe C^1 sur Ω' et $Im(f) \subset \Omega'$, alors $g \circ f$ est de classe C^1 sur Ω.

5. Le théorème d'inversion locale

Définition

Soient U un ouvert de $\mathbb{R}^p$, V un ouvert de $\mathbb{R}^q$, et $f : U \to \mathbb{R}^q$ une application définie sur U. On dit que f est un difféomorphisme de classe C^1 de U sur V si

(i) f est de classe C^1 sur U ,

(ii) f réalise une bijection de U sur V ,

(iii) f^{-1}, l'inverse de f, est de classe C^1 sur V .

On cherche ici des conditions simples sur f pour que f soit un difféomorphisme de classe C^1 de U sur V. De la formule de composition pour la différentielle et des formules $f^{-1} \circ f = \mathrm{Id}_{\mathbb{R}^p}$ et $f \circ f^{-1} = \mathrm{Id}_{\mathbb{R}^q}$ on tire par différenciation de ces formules que pour tout $a \in U$,

$$\begin{cases} Df^{-1}(f(a)) \circ Df(a) = \mathrm{Id}_{\mathbb{R}^p} \ , \\ Df(a) \circ Df^{-1}(f(a)) = \mathrm{Id}_{\mathbb{R}^q} \ . \end{cases}$$

Par suite si f est un C^1-difféomorphisme de U sur V alors $Df(a)$ est un isomorphisme de $\mathbb{R}^p$ sur $\mathbb{R}^q$, d'inverse $Df^{-1}(f(a))$:

$$Df^{-1}(f(a)) = Df(a)^{-1} \ .$$

En particulier, comme un isomorphisme présèrve les dimensions (l'image d'une base par un isomorphisme est encore une base), s'il existe un C^1-difféomorphisme de U sur V alors nécessairement $p = q$.

Le théorème qui répond à la question que nous avons posée plus haut est le théorème d'inversion locale. Si U est un ouvert de $\mathbb{R}^n$, si $a \in U$ et si $f : U \to \mathbb{R}^n$ est une application définie sur U et différentiable en a, on rappelle que le *jacobien* $Jf(a)$ de f en a est le déterminant de la matrice jacobienne de f en a. On a (c'est de l'algèbre linéaire) que $Df(a)$ est un isomorphisme de $\mathbb{R}^n$ si et seulement si le jacobien de f en a est non nul. Donc :

$$Df(a) \in \mathrm{Isom}(\mathbb{R}^n) \quad \Leftrightarrow \quad Jf(a) \neq 0 \ .$$

Le théorème d'inversion locale s'énonce de la façon suivante. Nous reviendrons plus en détail sur ce théorème lorsque nous traiterons du cas des espaces de Banach.

Théorème (Le théorème d'inversion locale)

Soient U un ouvert de $\mathbb{R}^n$, a un point de U, et $f : U \to \mathbb{R}^n$ une application de classe C^1 sur U. On suppose que $Jf(a) \neq 0$. Il existe alors $\eta > 0$ tel que

(i) $B_a(\eta) \subset U$ et $f\big(B_a(\eta)\big)$ est un ouvert de $\mathbb{R}^n$

(ii) f est un difféomorphisme de classe C^1 de $B_a(\eta)$ sur $f\big(B_a(\eta)\big)$,

où $B_a(\eta)$ est la boule ouverte de centre a et de rayon η dans $\mathbb{R}^n$. Si maintenant f est injective sur U, et si en tout point x de U, $Jf(x) \neq 0$, alors $f(U)$ est un ouvert de $\mathbb{R}^n$ et f est un difféomorphisme de classe C^1 de U sur $f(U)$.

(1) Ce théorème généralise ce que vous saviez déjà pour les fonctions d'une variable réelle. Si I est un intervalle de $\mathbb{R}$, si $a \in I$, et si $f : I \to \mathbb{R}$ est une fonction de classe C^1 telle que $f'(a) \neq 0$ alors soit $f'(a) > 0$ soit $f'(a) < 0$. Par suite, par continuité de f', pour η petit, et pour tout x dans $]a-\eta, a+\eta[$, $f'(x) > 0$ (ou $f'(x) < 0$). En particulier, f est strictement croissante sur $]a-\eta, a+\eta[$ ou strictement décroissante sur $]a-\eta, a+\eta[$, et f réalise ainsi un difféomorphisme de classe C^1 de $]a-\eta, a+\eta[$ sur l'intervalle ouvert $f\big(]a-\eta, a+\eta[\big)$.
(2) La dernière partie du théorème tel qu'énoncé est une version globale du théorème. Elle se déduit facilement de la partie locale en remarquant que si f est injective sur U alors f réalise une bijection de U sur $f(U)$. La partie locale impose ensuite que $f(U)$ est ouvert et que f^{-1} est C^1 sur $f(U)$.

6. Le théorème des fonctions implicites

Soient U un ouvert de $\mathbb{R}^{n+p}$, et $f : U \to \mathbb{R}^p$ une application de classe C^1 définie sur U. On écrit $\mathbb{R}^{n+p}$ sous la forme $\mathbb{R}^{n+p} = \mathbb{R}^n \times \mathbb{R}^p$, et on écrit les points de $\mathbb{R}^{n+p}$ sous la forme (x, y) avec $x \in \mathbb{R}^n$ et $y \in \mathbb{R}^p$. Etant donné (a, b) un point de U, on note f_a l'application de $\mathbb{R}^p$ dans $\mathbb{R}^p$ définie par

$$f_a(y) = f(a, y) \ .$$

On vérifie facilement que le domaine de définition de f_a est un ouvert de $\mathbb{R}^p$. De même, on note f_b l'application de $\mathbb{R}^n$ dans $\mathbb{R}^p$ définie par

$$f_b(x) = f(x, b) \ .$$

Là encore, on vérifie facilement que le domaine de définition de f_b est un ouvert de $\mathbb{R}^n$. Puisque f est de classe C^1 sur U, il est clair que f_a et f_b sont aussi de classe C^1 sur leurs domaines de définition respectifs. On note alors $\frac{\partial f}{\partial y}(a, b)$ la différentielle de f_a au point b, et $\frac{\partial f}{\partial x}(a, b)$ la différentielle de f_b au point a.

On a donc

$$\begin{cases} \frac{\partial f}{\partial y}(a,b) = Df_a(b) \ , \\ \frac{\partial f}{\partial x}(a,b) = Df_b(a) \end{cases}$$

et $\frac{\partial f}{\partial y}(a,b)$ et $\frac{\partial f}{\partial x}(a,b)$ sont respectivement des applications linéaires de $\mathbb{R}^p$ dans $\mathbb{R}^p$ et de $\mathbb{R}^n$ dans $\mathbb{R}^p$.

Le théorème des fonctions implicites, objet de ce paragraphe, s'énonce de la façon suivante. Là encore nous reviendrons plus en détail sur ce théorème lorsque nous traiterons du cas des espaces de Banach.

Théorème (Théorème des fonctions implicites)

Soient U un ouvert de $\mathbb{R}^{n+p}$, (a,b) un point de U, et $f : U \to \mathbb{R}^p$ une fonction de classe C^1 sur U. On suppose que $\frac{\partial f}{\partial y}(a,b) \in Isom(\mathbb{R}^p)$. Il existe alors un ouvert V de $\mathbb{R}^{n+p}$ contenant le point (a,b) et contenu dans U, il existe un ouvert W de $\mathbb{R}^n$ contenant le point a, et il existe une fonction $g : W \to \mathbb{R}^p$ de classe C^1 sur W tels que

(i) $\forall (x,y) \in V$, *si* $f(x,y) = f(a,b)$ *alors* $x \in W$ *et* $y = g(x)$

(ii) $\forall x \in W$, $\big(x, g(x)\big) \in V$ *et* $f\big(x, g(x)\big) = f(a,b)$.

En particulier, $g(a) = b$.

L'application g est unique (au moins au voisinage de a).

7. Différentielles d'ordre supérieures et formule de Taylor à l'ordre deux

On est là encore confronté à une difficulté similaire à celle rencontrée pour parler de classe C^1 : l'application différentielle Df n'est pas de même nature que f et on ne peut se borner à dire que f est deux fois différentiable si Df est une fois différentiable puisque l'on ne sait pas (encore) définir la différentiabilité d'une application à valeur $L(\mathbb{R}^p, \mathbb{R}^q)$, sauf à assimiler de façon artificielle $L(\mathbb{R}^p, \mathbb{R}^q)$ à $\mathbb{R}^{pq}$.

Pour simplifier la présentation et éviter l'abondance d'indices on va ici supposer que $q = 1$ (on se restreint donc à parler des fonctions). Mais bien sûr on retiendra que pour $f = (f_1, \ldots, f_q)$, la différentiabilité de f, le caractère C^1 de f, la différentiabilité seconde de f, le caractère C^2 de f, etc. équivalent à ce que pour tout $j = 1, \ldots, q$ les f_j sont différentiables, de classe C^1, deux fois différentiables, de classe C^2, etc. On passe ainsi "facilement de la théorie $\mathbb{R}^p \to \mathbb{R}$ à la théorie $\mathbb{R}^p \to \mathbb{R}^q$.

Définition

Soient Ω un ouvert de $\mathbb{R}^n$, a un point de Ω, et $f : \Omega \to \mathbb{R}$ une fonction réelle définie sur Ω. On suppose que pour tout $i = 1, \ldots, n$, et tout $x \in \Omega$, $\frac{\partial f}{\partial x_i}(x)$ existe. Si elle existe, on appelle dérivée partielle seconde de f par rapport à x_i et x_j au point a, et on note $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$, la dérivée partielle par rapport à x_i au point a de la fonction $\frac{\partial f}{\partial x_j}$. Pour tous $i, j = 1, \ldots, n$, et tout $a \in \Omega$, on a donc

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

Lorsque $i = j$, on note $\frac{\partial^2 f}{\partial x_i^2}(a)$ au lieu de $\frac{\partial^2 f}{\partial x_i \partial x_i}(a)$.

Définition

Soient Ω *un ouvert de* $\mathbb{R}^n$, *a un point de* Ω, *et* $f : \Omega \to \mathbb{R}$ *une fonction différentiable sur* Ω. *On dit que* f *est deux fois différentiable au point* a *si les dérivées partielles premières de* f *sont différentiables au point* a, *autrement dit si pour tout* $i = 1, \ldots, n$, *les fonctions* $\frac{\partial f}{\partial x_i} : \Omega \to \mathbb{R}$ *sont différentiables au point* a. *Si* f *est deux fois différentiable au point* a, *la différentielle seconde* $D^2 f(a)$ *de* f *au point* a *est alors (par définition) la forme bilinéaire sur* $\mathbb{R}^n$ *définie en tous* $h = (h_1, \ldots, h_n), k = (k_1, \ldots, k_n) \in \mathbb{R}^n$ *par*

$$D^2 f(a).(h, k) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i k_j$$

où les $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$ *sont les dérivées partielles secondes de* f *au point* a. *Par extension, on dit que* f *est deux fois différentiable sur* Ω *si* f *est deux fois différentiable en tout point de* Ω.

Définition

On dit que f est de classe C^2 sur Ω si les dérivées partielles premières de f sont de classe C^1 sur Ω, autrement dit si pour tout $i = 1, \ldots, n$, les fonctions $\frac{\partial f}{\partial x_i} : \Omega \to \mathbb{R}$ sont de classe C^1 sur Ω.

Comme toujours, si f et g sont deux fois différentiable (resp. de classe C^2) alors $f + g$, fg et $\frac{f}{g}$ si $g \neq 0$ sont aussi deux fois différentiables (resp. de classe C^2). De même, si $\Omega \subset \mathbb{R}^{n_1}$ et $\Omega' \subset \mathbb{R}^{n_2}$ sont deux ouverts de $\mathbb{R}^{n_1}$ et $\mathbb{R}^{n_2}$, $f : \Omega \to \mathbb{R}^{n_2}$ est une application deux fois différentiable (resp. de classe C^2) sur Ω, $g : \Omega' \to \mathbb{R}^{n_3}$ est une application deux fois différentiable (resp. de classe C^2) sur Ω' et $Im(f) \subset \Omega'$, alors $g \circ f$ est deux fois différentiable (resp. de classe C^2) sur Ω.

A priori, étant donnée f une fonction réelle différentiable sur un ouvert Ω de $\mathbb{R}^n$ et deux fois différentiable en un point a de Ω, f possède n^2 dérivées partielles secondes. En fait, pour des raisons de symétrie, il n'y en a que $n(n+1)/2$. C'est l'objet du théorème de Schwarz. On renvoie à l'ouvrage d'Henri Cartan (*Cours de calcul différentiel*, Hermann) pour sa preuve (elle n'est pas triviale).

Théorème (Théorème de Schwarz)

Soient Ω un ouvert de $\mathbb{R}^n$, a un point de Ω, et $f : \Omega \to \mathbb{R}$ une fonction différentiable sur Ω. Si f est deux fois différentiable au point a, alors pour tous $i, j = 1, \ldots, n$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a) \ .$$

Une autre façon de dire les choses est encore que la différentielle seconde $D^2 f(a)$ de f au point a est symétrique au sens où pour tous $h, k \in \mathbb{R}^n$, $D^2 f(a).(h, k) = D^2 f(a).(k, h)$.

Théorème (Formule de Taylor à l'ordre 2)

Soient Ω un ouvert de $\mathbb{R}^n$, a un point de Ω, et $f : \Omega \to \mathbb{R}$ une fonction réelle différentiable sur Ω et deux fois différentiable au point a. Pour tout $h \in \mathbb{R}^n$, avec $\|h\|$ petit,

$$f(a+h) = f(a) + Df(a).(h) + \frac{1}{2}D^2f(a).(h,h) + \|h\|^2\epsilon(h)$$

où $\epsilon : \mathbb{R}^n \to \mathbb{R}$ est une fonction qui vaut 0 en 0 et qui tend vers 0 en 0.

Preuve : Supposons pour simplifier la preuve que f est C^2. La formule à démontrer équivaut à

$$\lim_{h\to 0} \frac{1}{\|h\|^2}\left(f(a+h) - f(a) - Df(a).(h) - \frac{1}{2}D^2f(a).(h,h)\right) = 0 \ .$$

Considérons $g : [0,1] \to \mathbb{R}$ la fonction $g(t) = f(a+th)$. On a $g'(t) = Df(a+th).(h)$ et $g''(t) = D^2f(a+th).(h,h)$. La formule de Taylor-Lagrange dans le cas $\mathbb{R} \to \mathbb{R}$ (Chapitre 1) donne pour tout h petit l'existence d'un $t_h \in]0,1[$ tel que

Preuve suite :

$$g(1) = g(0) + g'(0) + \frac{1}{2} g''(t_h) \ .$$

Par suite,

$$\begin{aligned} &\frac{1}{\|h\|^2} \left| f(a+h) - f(a) - Df(a).(h) - \frac{1}{2} D^2 f(a).(h,h) \right| \\ &= \frac{1}{2\|h\|^2} \left| \left(D^2 f(a + t_h h) - D^2 f(a) \right) (h,h) \right| \\ &\leq C \sum_{i,j} \left| \frac{\partial^2 f}{\partial x_i \partial x_j}(a + t_h h) - \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right| \ , \end{aligned}$$

où $C > 0$ est indépendante de h, et puisque les $\frac{\partial^2 f}{\partial x_i \partial x_j}$ sont supposées continues, et $a + t_h h \to a$ lorsque $h \to 0$ ($|t_h| \leq 1$ et donc $t_h h \to 0$ lorsque $h \to 0$), on voit que

$$\frac{1}{\|h\|^2} \left| f(a+h) - f(a) - Df(a).(h) - \frac{1}{2} D^2 f(a).(h,h) \right| \to 0$$

lorsque $h \to 0$. Le théorème est démontré. CQFD

Théorème (Condition nécessaire 1)

Soient $\Omega \subset \mathbb{R}^n$ un ouvert de $\mathbb{R}^n$, $a \in \Omega$ un point de Ω et $f : \Omega \to \mathbb{R}$ une fonction. Si f admet un extremum local en a et si f est différentiable en a alors $\frac{\partial f}{\partial x_i}(a) = 0$ pour tout $i = 1, \dots, n$ (ce qui revient encore à dire que $Df(a) \equiv 0$ en tant qu'application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$). En d'autres termes, les extremums locaux d'une fonction différentiable sont des points critiques de cette fonction.

Preuve : Sous les hypothèses du théorème les fonctions partielles f_a^i sont dérivables en a_i et admettent un extremum local en a_i. Donc (cf. Chapitre 1), $(f_a^i)'(a_i) = 0$, ce qui revient à dire que $\frac{\partial f}{\partial x_i}(a) = 0$. CQFD

Théorème (Condition nécessaire 2)

Soient $\Omega \subset \mathbb{R}^n$ un ouvert de $\mathbb{R}^n$, $a \in \Omega$ un point de Ω et $f : \Omega \to \mathbb{R}$ une fonction différentiable sur Ω et deux fois différentiables en a. Si f admet un minimum local en a alors $D^2 f(a) \geq 0$ au sens des formes bilinéaires et si f admet un maximum local en a alors $D^2 f(a) \leq 0$ au sens des formes bilinéaires.

Dire que $D^2 f(a) \geq 0$ au sens des formes bilinéaires signifie que $D^2 f(a).(h, h) \geq 0$ pour tout $h \in \mathbb{R}^n$, et dire que $D^2 f(a) \leq 0$ au sens des formes bilinéaires signifie que $D^2 f(a).(h, h) \leq 0$ pour tout $h \in \mathbb{R}^n$.

Preuve : Si f admet un extremum local au point a, on sait avec le Théorème précédent que a est un point critique de f. Il s'ensuit de la formule de Taylor à l'ordre 2 de f au point a que pour tout h,

$$f(a + h) = f(a) + \frac{1}{2} D^2 f(a).(h, h) + \|h\|^2 \epsilon(h) \ .$$

Preuve suite et fin : Soit alors h donné, et soit t un réel strictement positif. On écrit que

$$\begin{aligned} f(a+th)-f(a) &= \frac{1}{2}D^2 f(a).(th,th)+\|th\|^2\epsilon(th) \\ &= \Big(\frac{1}{2}D^2 f(a).(h,h)+\|h\|^2\epsilon(th)\Big)t^2 \ . \end{aligned}$$

Si maintenant f admet un minimum local au point a, on obtient que pour $t>0$, t petit,

$$\frac{1}{2}D^2 f(a).(h,h)+\|h\|^2\epsilon(th)\geq 0 \ .$$

En faisant tendre t vers 0 dans cette inégalité, et puisque par propriété de la fonction ϵ, $\lim_{t\to 0}\epsilon(th)=0$, on obtient que $D^2 f(a).(h,h)\geq 0$. Le raisonnement est bien sûr de même nature si f admet un maximum local au point a. On récupère dans ce cas que $D^2 f(a).(h,h)\leq 0$. Le théorème est démontré. CQFD

Lemma

Soit $Q : \mathbb{R}^n \to \mathbb{R}$ une fonction qui pour tout $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ s'écrit sous la forme

$$Q(h) = \sum_{i,j=1}^{n} a_{ij} h_i h_j$$

où les a_{ij} sont des réels. On suppose que pour tout $h \in \mathbb{R}^n \backslash \{0\}$, $Q(h) > 0$. Il existe alors un réel $K > 0$ tel que pour tout $h \in \mathbb{R}^n$, $Q(h) \geq K \|h\|^2$.

Preuve : Notons S la sphère unité de $\mathbb{R}^n$ définie par

$$S = \{h \in \mathbb{R}^n \ / \ \|h\| = 1\} \ .$$

Alors S est un fermé borné de $\mathbb{R}^n$, donc un compact de $\mathbb{R}^n$. Par ailleurs, la fonction Q est continue sur $\mathbb{R}^n$, et donc en particulier sur S.

Preuve suite et fin : Puisque toute fonction continue sur un compact est bornée est atteint ses bornes, il s'ensuit qu'il existe $h_0 \in S$ tel que pour tout $h \in S$, $Q(h) \geq Q(h_0)$. On pose $K = Q(h_0)$. Comme Q est supposée strictement positive en tout point différent de 0, K est strictement positif. On remarque par ailleurs que pour tout $t \in \mathbb{R}$, et tout $h \in \mathbb{R}^n$, $Q(th) = t^2 Q(h)$. Pour $h \in \mathbb{R}^n \backslash \{0\}$ on pourra ainsi écrire que

$$Q(h) = Q\left(\|h\|\left(\frac{1}{\|h\|}h\right)\right) = \|h\|^2 Q\left(\frac{1}{\|h\|}h\right) \geq K\|h\|^2$$

puisque $\frac{1}{\|h\|}h \in S$. Le résultat s'ensuit. CQFD

Note : On aurait aussi pu démontrer ce résultat en utilisant Sylvester et l'équivalence des normes dans $\mathbb{R}^n$.

Théorème (Condition suffisante)

Soient Ω un ouvert de $\mathbb{R}^n$, a un point de Ω, et $f : \Omega \to \mathbb{R}$ une fonction différentiable sur Ω et deux fois différentiable au point a. On suppose que a est un point critique de f. Si pour tout $h \in \mathbb{R}^n \backslash \{0\}$, $D^2 f(a).(h,h) > 0$ (resp. pour tout $h \in \mathbb{R}^n \backslash \{0\}$, $D^2 f(a).(h,h) < 0$) alors f admet un minimum local strict (resp. un maximum local strict) au point a.

Preuve : On traite du cas des minimums sachant que celui des maximums en découle en changeant f en $-f$. D'après le lemme précédent, il existe $K > 0$ tel que pour tout $h \in \mathbb{R}^n$,

$$D^2 f(a).(h,h) \geq K\|h\|^2 \ .$$

On écrit avec la formule de Taylor à l'ordre deux pour f au point a que

$$f(a+h)-f(a)=\frac{1}{2}D^2f(a).(h,h)+\|h\|^2\epsilon(h)$$
$$\geq \|h\|^2\left(\frac{K}{2}+\epsilon(h)\right)\ .$$

Par propriété de la fonction ϵ, il existe alors un $\eta>0$ tel que pour tout $h\in\mathbb{R}^n$, si $\|h\|<\eta$, $|\epsilon(h)|\leq K/4$. Il s'ensuit que pour h tel que $\|h\|<\eta$,

$$f(a+h)-f(a)\geq\frac{K}{4}\|h\|^2\ .$$

En écrivant $x=a+h$, on en déduit qu'il existe $\eta>0$ tel que pour $x\in B_a(\eta)\backslash\{a\}$, $f(x)>f(a)$. D'où le résultat. CQFD