Dérivabilité des fonctions convexes

Théorème. *Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction convexe. Alors f est continue, et dérivable de dérivée continue sur un ensemble codénombrable.*

Théorème. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe. Alors f est continue, dérivable presque partout.*

On admettra le lemme suivant :

Lemme. *Soit $F : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe qui s'annule en 0, alors pour tout $a > 0$,*

$$\sup_{\|h\|_1 \leq a} |F(h)| \leq \max_{i=1..n, s=\pm 1} F(sae_i)$$

Démonstration. (du premier théorème) On utilise la caractérisation de la convexité suivante : pour tout $u < v < w$,

$$\frac{f(v)-f(u)}{v-u} \leq \frac{f(w)-f(v)}{w-v} \leq \frac{f(w)-f(u)}{w-u}$$

Une fonction convexe est localement Lipschitzienne ; en effet, soit $[a,b] \subset]c,d[$, $f_{|[a,b]}$ est $\max\left(|\frac{f(d)-f(b)}{d-b}|, |\frac{f(a)-f(c)}{a-c}|\right)$-Lipschitzienne (faire un dessin, et comparer les pentes).

Avec $(u,v,w) = (x-h, x, x+k)$, en gardant le premier et le dernier terme ;

$$\frac{f(x)-f(x-h)}{h} \leq \frac{f(x+k)-f(x)}{k}$$

Soient $h_1 < h_2$, $k_1 < k_2$, avec $(u,v,w) = (x-h_2, x-h_1, x)$ et $(x, x+k_1, x+k_2)$, on obtient que $h > 0 \mapsto \frac{f(x)-f(x-h)}{h}$ est décroissante et $k > 0 \mapsto \frac{f(x+k)-f(x)}{k}$ est croissante. Ainsi, $\begin{cases} \mathbb{R}^* \to \mathbb{R} \\ k \mapsto \frac{f(x+k)-f(x)}{k} \end{cases}$ est une fonction croissante et f est dérivable en x si et seulement si la fonction précédente se prolonge par continuité en 0, c'est à dire si et seulement si

$$g(x) := \sup_{h>0} \frac{f(x)-f(x-h)}{h} = \inf_{k>0} \mapsto \frac{f(x+k)-f(x)}{k} =: d(x)$$

On a toujours $g(x) \leq d(x)$. Soit $y > x$, alors $d(x) \leq \frac{f(y)-f(x)}{y-x} \leq g(y)$. Ainsi, g et d sont des fonctions croissantes "entrelacées". Elles sont donc continues sauf sur un ensemble dénombrable D. Soit x un point de continuité de g et l, pour tout $y > x$, $g(x) \leq d(x) \leq g(y) \underset{y \to x}{\longrightarrow} d(x)$, d'où $d(x) = g(x)$, ce qui achève la preuve. □

Démonstration. f est continue. En effet, soit $x \in \mathbb{R}^n$, $h \mapsto f(x+h) - f(x)$ vérifie les hypothèses du lemme, donc

$$\sup_{\|h\|_1 \leq a} |f(x+h) - f(x)| \leq \max_{i=1..n, s=\pm 1} f(x + sae_i) - f(x) \to 0$$

Soit $i \in \{1..n\}$, et $A_i = \{x \in \mathbb{R}^n$ tels que $\frac{\partial f}{\partial x_i}(x)$ existe $\}$. On a montré que pour tout $\widehat{x_i} = (x_1, .., x_{i-1}, x_{i+1}, .., x_n) \in \mathbb{R}^{n-1}$, $A_i \cap (\widehat{x_i} \times \mathbb{R})$ est codénombrable ; selon le théorème de Fubini-Tonelli, A_i est de mesure pleine.
Soit $A = \bigcap\limits_{i=1..n} A_i$, A est de mesure pleine. C'est l'ensemble de définition de ∇f (le vecteur de composantes $\frac{\partial f}{\partial x_i}$), reste à montrer que c'est bien l'ensemble de dérivabilité de f.
Soit $x \in V$, la fonction $h \mapsto f(x+h) - f(x) - \nabla f(x) \cdot h$ vérifie les hypothèses du lemme (somme d'une fonction convexe et d'une forme linéaire) ; on l'applique sur une boule de rayon $a > 0$:

$$\sup_{\|h\|_1 \leq a} |f(x+h) - f(x) - \nabla f(x) \cdot h| \leq \max_{i=1..n, s=\pm 1} f(x + sae_i) - f(x) - sa\frac{\partial f}{\partial x_i}(x) = o(a)$$

$h \mapsto \nabla f(x) \cdot h$ est donc bien la différentielle de f en x, d'où le résultat. □

Voici la preuve du lemme admis :

Démonstration. Notons que pour tout h tel que $\|h\|_1 \leq a$, $0 = F(0) \leq \frac{1}{2}(F(h) + F(-h))$, donc $\sup\limits_{\|h\|_1 \leq a} |F(h)| \leq \sup\limits_{\|h\|_1 \leq a} F(h)$.
De plus, $F_{|h\mathbb{R}}$ est convexe, d'où $\sup\limits_{\|h\|_1 \leq a} F(h) \leq \sup\limits_{\|h\|_1 = a} F(h)$.
Si $\|h\|_1 = a$, tel que $h_i \neq 0$ et $h_j \neq 0$ (où $i \neq j$). On note s le signe de h_i/h_j, alors $h + (e_i - se_j)\mathbb{R}$ est une droite telle que pour tout $t \in [-h_i, h_j]$, $h + (e_i - se_j)t$ est de norme a, et sur les bords, une des coordonnées (i ou j) s'annule sans toucher aux coordonnées $k \neq i, j$; de proche en proche, on se ramène à $h = \pm ae_i$ pour un certain $i \in \{1, .., n\}$; on en déduit la dernière inégalité. □

Corollaire. *Soit $\|.\|$ une norme sur $\mathbb{R}^n$, alors $\|.\|$ est presque partout différentiable.*