Sous gradients des fonctions convexes

Définition : Soient $f : \mathbb{R}^n \longrightarrow]-\infty, +\infty]$ une fonction convexe et $\bar{x} \in \operatorname{dom} f$ (i.e., $f(\bar{x}) < +\infty$). Un élément $v \in \mathbb{R}^n$ est appelé sous-gradient de f au point $\bar{x}$ si

$$\langle v, x - \bar{x} \rangle \leqslant f(x) - f(\bar{x}) \quad \forall x \in \mathbb{R}^n$$

La collection de tous les sous-gradients de f au point $\bar{x}$ est appelée sous-différentiel de f au point $\bar{x}$.

On la note $\partial f(\bar{x})$

Exemple 1 : Considérons la fonction $f(x) = x$, $x \in \mathbb{R}$.

Alors $\partial f(0) = [-1, 1]$.

En effet, soit $v \in \partial f(0)$, alors

$$v(x - 0) \leqslant f(x) - f(0) \quad \forall x \in \mathbb{R}$$

$$\Rightarrow \quad vx \leqslant |x| \quad \forall x \in \mathbb{R}$$

- $\underline{x = 1}$
  $$v . 1 \leqslant 1$$
- $\underline{x = -1 :}$ $\quad v(-1) \leqslant |-1| = 1$
  $$\Rightarrow \quad v \geqslant -1$$

Ainsi $v \in [-1, 1]$ et par suite

$$\partial f(0) \subset [-1, 1]$$

On va montrer l'inclusion inverse :

Soit $v \in [-1, 1]$. Alors $|v| \leqslant 1$

$\forall x \in \mathbb{R}$

$$v(x - 0) = vx \leqslant |vx| = |v|\,|x| \leqslant |x| = |x| - |0| = f(x) - f(0)$$

$$\Rightarrow v \in \partial f(0)$$

Il s'ensuit que $[-1, 1] \subset \partial f(0)$

A. KHELOUFI

Exemple ② : Considérons la fonction

$$f(x) = \|x\|, \quad x \in \mathbb{R}^n$$

Alors $\partial f(0) = \mathbb{B}(0,1) = \{ v \in \mathbb{R}^n : \|v\| \leq 1 \}$.

En effet, soit $v \in \partial f(0)$. Alors

$$\langle v, x - 0 \rangle \leq f(x) - f(0) \quad \forall x \in \mathbb{R}^n$$

$$\Rightarrow \quad \langle v, x \rangle \leq \|x\| \qquad \forall x \in \mathbb{R}^n$$

Pour $x = v$

$$\langle v, v \rangle \leq \|v\|$$

$$\Rightarrow \quad \|v\|^2 \leq \|v\|$$

Ce qui implique que

$$\|v\| \leq 1.$$

Donc $v \in \mathbb{B}(0,1)$

Ainsi $\partial f(0) \subset \mathbb{B}(0,1)$.

* Soit $v \in \mathbb{R}^n$: $\|v\| \leq 1$ $\quad (v \in \mathbb{B}(0,1))$.

Soit $x \in \mathbb{R}^n$,

$$\langle v, x - 0 \rangle = \langle v, x \rangle \leq \|v\| \, \|x\|$$

$$\leq \|x\| = f(x) - f(0).$$

Par conséquent, $v \in \partial f(0)$

Ainsi $\mathbb{B}(0,1) \subset \partial f(0)$.

A. KHELOUFI
Convexité et optimisation

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction. On dit que f est différentiable en $\bar{x}$ s'il existe $v \in \mathbb{R}^n$:

$$\lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} = 0.$$

Le vecteur v est appelé gradient de f en $\bar{x}$.
On le note $\nabla f(\bar{x})$.

Proposition : Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction convexe.
Si f est différentiable en $\bar{x}$, alors

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \quad \forall x \in \mathbb{R}^n.$$

Démonstration : Soit $x \in \mathbb{R}^n$ et $t \in]0,1[$.

$$\begin{aligned} f(\bar{x} + t(x - \bar{x})) &= f(tx + (1-t)\bar{x}) \\ &\leq t f(x) + (1-t) f(\bar{x}) \\ &= f(\bar{x}) + t(f(x) - f(\bar{x})). \end{aligned}$$

Alors

$$\frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x}) - \langle \nabla f(\bar{x}), t(x - \bar{x}) \rangle}{\|t(x - \bar{x})\|}$$

$$\leq \frac{t(f(x) - f(\bar{x})) - t \langle \nabla f(\bar{x}), x - \bar{x} \rangle}{t \|x - \bar{x}\|}$$

$$= \frac{f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|}$$

En prenant la limite lorsque $t \longrightarrow 0^+$, on a

$$0 \leq \frac{f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|}$$

$x \neq \bar{x}$
$\|x - \bar{x}\| > 0$

Donc $0 \leq f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle$

Par conséquent, $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq f(x) - f(\bar{x})$

A. KHELOUF
Convexité et Optimisation

Proposition : Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction convexe.
Si f est différentiable en $\bar{x}$, alors $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$.

Lemme : Supposons que $\langle v, h \rangle \leqslant \varepsilon \|h\|$ lorsque $\|h\| < \delta$ $(\varepsilon, \delta > $
Alors $\|v\| \leqslant \varepsilon$.

Preuve : $h = \dfrac{\delta v}{2\|v\|}$, $v \neq 0$

$$\langle v, \frac{\delta}{2}\frac{v}{\|v\|} \rangle \leqslant \varepsilon \frac{\delta}{2}. \text{ Alors}$$

$$\langle v, \frac{v}{\|v\|} \rangle \leqslant \varepsilon \Rightarrow \frac{\|v\|^2}{\|v\|} \leqslant \varepsilon$$

$$\Rightarrow \|v\| \leqslant \varepsilon.$$

Démonstration de la proposition : On a

$$\forall x \in \mathbb{R}^n \quad \langle \nabla f(\bar{x}), x - \bar{x} \rangle \leqslant f(x) - f(\bar{x})$$

Donc $\nabla f(\bar{x}) \in \partial f(\bar{x})$.

Soit $v \in \partial f(\bar{x})$. Montrons que $v = \nabla f(\bar{x})$

Par définition, on a

$$\langle v, x - \bar{x} \rangle \leqslant f(x) - f(\bar{x}) \quad \forall x \in \mathbb{R}^n$$

Puisque f est différentiable en $\bar{x}$, alors

$$\lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|}$$

Ainsi $\forall \varepsilon > 0$, $\exists \delta > 0$, tels que

$$\left| \frac{f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle}{\|x - \bar{x}\|} \right| < \varepsilon \quad \text{lorsque } \|x - \bar{x}\| < \delta \quad x \neq \bar{x}$$

Alors

$$f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle < \varepsilon \|x - \bar{x}\| \quad \text{lorsque } \|x - \bar{x}\| < \delta$$

Par conséquent,

$$f(x) - f(\bar{x}) < \langle \nabla f(\bar{x}), x - \bar{x} \rangle + \varepsilon \|x - \bar{x}\| \quad \text{lorsque } \|x - \bar{x}\| < \delta$$

On a

$$\langle v, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \quad \forall x \in \mathbb{R}^n$$

Par suite,

$$\langle v, x - \bar{x} \rangle < \langle \nabla f(\bar{x}), x - \bar{x} \rangle + \varepsilon \|x - \bar{x}\| \quad \text{lorsque } \|x - \bar{x}\| < \delta$$

$$\Rightarrow \quad \langle v - \nabla f(\bar{x}), x - \bar{x} \rangle < \varepsilon \|x - \bar{x}\| \quad \text{lorsque } \|x - \bar{x}\| < \delta$$

Il s'ensuit

$$\langle v - \nabla f(\bar{x}), h \rangle < \varepsilon \|h\| \quad \text{lorsque } \|h\| < \delta$$

Donc

$$\|v - \nabla f(\bar{x})\| < \varepsilon$$

Puisque $\varepsilon > 0$ est arbitraire

$$v = \nabla f(\bar{x})$$

donc $\partial f(\bar{x}) \subset \{\nabla f(\bar{x})\}$

Finalement,

$$\partial f(\bar{x}) = \{\nabla f(\bar{x})\}.$$

Sous-différentiel : règle de composition

Soit $B : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ une application affine

$$B(x) = Ax + b, \quad x \in \mathbb{R}^n$$

A est une $p \times n$ matrice et $b \in \mathbb{R}^p$.

Lemme : Soit $B : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ l'application affine définie précédemment. Alors pour tout $(\bar{x}, \bar{y}) \in gph(B)$, on a

$$N((\bar{x}, \bar{y}); gph(B)) = \{(u, v) \in \mathbb{R}^n \times \mathbb{R}^p ; u = -A^T v\}$$

Preuve : Fixons $(u, v) \in N((\bar{x}, \bar{y}); gph(B))$. Par définition

$$\langle u, x - \bar{x} \rangle + \langle v, y - \bar{y} \rangle \leqslant 0 \quad \forall (x, y) \in gph(B).$$

On a

$$\langle u, x - \bar{x} \rangle + \langle v, B(x) - B(\bar{x}) \rangle = \langle u, x - \bar{x} \rangle + \langle v, Ax - A\bar{x} \rangle$$

$$= \langle u, x - \bar{x} \rangle + \langle v, A(x - \bar{x}) \rangle$$

$$= \langle u, x - \bar{x} \rangle + \langle A^T v, x - \bar{x} \rangle$$

$$= \langle u + A^T v, x - \bar{x} \rangle.$$

Alors $(u, v) \in N((\bar{x}, \bar{y}), gph(B))$ si et seulement si

$$\langle u + A^T v, x - \bar{x} \rangle \leqslant 0 \quad \forall x \in \mathbb{R}^n$$

$$\iff \langle u + A^T v, y \rangle \leqslant 0 \quad \forall y \in \mathbb{R}^n$$

$$\iff \langle u + A^T v, y \rangle = 0 \quad \forall y \in \mathbb{R}^n \quad (\text{car c'est vrai pour } -y)$$

$$\iff u + A^T v = 0$$

$$\iff u = -A^T v$$

Proposition : Soit $B : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ l'application affine définie précédemment. Soit $f : \mathbb{R}^p \longrightarrow]-\infty, +\infty]$ une fonction convexe. Fixons $\bar{x} \in \mathbb{R}^n$ tel que $\bar{y} = B(\bar{x}) \in \operatorname{dom} f$.

Alors

$$\partial (f \circ B)(\bar{x}) \supset A^T(\partial f(\bar{y})) = \{A^T v \,/\, v \in \partial f(\bar{y})\} = \bigcup_{v \in \partial f(\bar{y})} \{A^T v\}$$

Preuve :

Fixons $u \in A^T(\partial f(\bar{y}))$. Alors

$$u = A^T v, \quad v \in \partial f(\bar{y})$$

$\forall x \in \mathbb{R}^n$, on a

$$\begin{aligned}\langle u, x - \bar{x} \rangle &= \langle A^T v, x - \bar{x} \rangle \\ &= \langle v, Ax - A\bar{x} \rangle \\ &= \langle v, B(x) - B(\bar{x}) \rangle \\ &\leq f(B(x)) - f(B(\bar{x})) \\ &= (f \circ B)(x) - (f \circ B)(\bar{x})\end{aligned}$$

Par suite

$$u \in \partial (f \circ B)(\bar{x}).$$

Théorème :

Considérons les hypothèses de la proposition précédente

Supposons que $ri(\operatorname{dom} f) \cap B(\mathbb{R}^n) \neq \phi$

Alors

$$\partial (f \circ B)(\bar{x}) = A^T(\partial f(\bar{y})), \quad \bar{y} = B(\bar{x}).$$

Preuve :

Fixons $u \in \partial (f \circ B)(\bar{x})$. Par définition

$$\begin{aligned}\langle u, x - \bar{x} \rangle &\leq (f \circ B)(x) - (f \circ B)(\bar{x}) \\ &= f(B(x)) - f(B(\bar{x})) \qquad \forall x \in \mathbb{R}^n\end{aligned}$$

Définissons

$$\Omega_1 = gph(B) \times \mathbb{R} \subset \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}$$

$$\Omega_2 = \mathbb{R}^n \times \operatorname{epi} f \subset \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}$$

Première étape : $(u, 0, -1) \in N((\bar{x}, \bar{y}, \bar{z}); \Omega_1 \cap \Omega_2)$, $\bar{y} = B(\bar{x})$, $\bar{z} = f(\bar{y})$

Fixons $(x, y, \lambda) \in \Omega_1 \cap \Omega_2$. Alors

$$y = B(x) \text{ et } f(y) \leq \lambda$$

Ainsi

$$\lambda \geq f(y) = f(B(x)).$$

On a

$$\langle (u, 0, -1), (x, y, \lambda) - (\bar{x}, \bar{y}, \bar{z}) \rangle$$

$$= \langle u, x - \bar{x} \rangle + \langle 0, y - \bar{y} \rangle + (-1)(\lambda - \bar{z})$$

$$= \langle u, x - \bar{x} \rangle - (\lambda - f(B(\bar{x})))$$

$$\leq \langle u, x - \bar{x} \rangle - [f(B(x)) - f(B(\bar{x}))]$$

≤ 0 car $u \in \partial(f \circ B)(\bar{x})$.

Deuxième étape : Application de la règle de l'intersection d'un cône normal

Par les hypothèses faites : $ri(\Omega_1) \cap ri(\Omega_2) \neq \phi$ (voir la 3ème étape)

Alors

$$(u, 0, -1) \in N((\bar{x}, \bar{y}, \bar{z}); \Omega_1 \cap \Omega_2)$$

$$= N((\bar{x}, \bar{y}, \bar{z}); \Omega_1) + N((\bar{x}, \bar{y}, \bar{z}); \Omega_2).$$

Alors

$$(u, 0, -1) = (u, -v_2, 0) + (0, v_2, -1)$$

où $(u, -v_2) \in N((\bar{x}, \bar{y}); gph\, B)$, $(v_2, -1) \in N(\bar{y}, \bar{z}; epi\, f)$

Alors

$$u = A^T v_2 \quad \text{et} \quad v_2 \in \partial f(\bar{y})$$

Ainsi $u \in A^T(\partial f(\bar{y}))$

Donc $\partial(f \circ B)(\bar{x}) \subset A^T(\partial f(\bar{y}))$

Troisième étape : Montrons que $ri(\Omega_1) \cap ri(\Omega_2) \neq \phi$.

Sous l'hypothèse $B(\mathbb{R}^n) \cap ri(dom\, f) \neq \phi$

On a

$$ri(\Omega_1) = \Omega_1 = gph(B) \times \mathbb{R}$$

$$ri(\Omega_2) = ri(\mathbb{R}^n) \times ri(epi\, f)$$

$$= \{(x, y, \lambda),\ x \in \mathbb{R}^n,\ y \in ri(dom\, f),\ \lambda > f(y)\}$$

Soit $\bar{y} \in B(\mathbb{R}^n)$ et $\bar{y} \in ri(\text{dom} f)$ (car $B(\mathbb{R}^n) \cap ri(\text{dom} f) \neq \phi$)

Choisissons $\bar{x} \in \mathbb{R}^n$, $\bar{y} = B(\bar{x})$

alors $(\bar{x}, \bar{y}, f(\bar{y}) + 1) \in ri(\Omega_1) \cap ri(\Omega_2)$.

A. KHELOUFI

Convexité et Optimisation