Inégalités de Łojasiewicz : le point de vue de l'analyse

Jérôme Bolte

INRIA Saclay / CMAP & Equipe Combinatoire et Optimisation de Paris 6

travail commun avec A. Daniilidis, O. Ley, L. Mazet

Mai 2008, Journées Franco-chiliennes de Toulon

Préliminaires

Un cadre :

- H espace de Hilbert , $f : H \to \mathbb{R} \cup \{+\infty\}$ propre semicontinue inférieurement,
- $f(\cdot) + \alpha || \cdot ||^2$ convexe pour un certain $\alpha > 0$,
- f à sous-niveaux compacts
- 0 est une valeur critique isolée de f

Calcul à l'ordre 1 Soit x dans dom f
Sous-différentiel : $x^* \in \partial f(x)$ ssi il existe

$$f(u) \geq f(x) + \langle x^*, u - x \rangle + o(||u - x||).$$

Propriété KL

Tranches de niveaux : $\boxed{[0 < f < r_0] = \{x \in \mathbb{R}^n : 0 < f(x) < r\}\,.}$

On dit que f **a la propriété de KL sur** $[0 < f < r_0]$ si il existe une fonction $g : [0 < f < r_0] \to \mathbb{R} \cup \{+\infty\}$ ayant les même sous-niveaux $[f \leq r]_{r\in(0,r_0)}$ que f et telle que

$$||\partial g(x)||_- \geq 1 \text{ pour tout } x \text{ dans } [0 < f < r_0].$$

$\boxed{\textbf{EX}}$ $f(x) = \frac{1}{2}||x||^2$ **et** $g(x) = ||x||$

Propriété KL formelle

Formellement :

Fonctions désingularisantes sur $(0, r_0)$:
$\varphi \in C([0, r_0), \mathbb{R}_+)$**,** $\varphi \in C^1(0, r_0)$**,** $\varphi' > 0$ **et** $\varphi(0) = 0$**.**

Definition

f a la propriété KL sur $[0 < f < r_0]$ si il existe φ désingularisante telle que

$$||\partial(\varphi \circ f)(x)|| \geq 1, \ \forall x \in [0 < f < r_0].$$

La propriété est fondamentale pour l'obtention de propriété de convergence dans une très large gamme de dynamiques de gradients : méthodes proximales, de gradient explicites, équations de type chaleur/ondes amorties, algorithme alternés...

Les résultats fondamentaux

Travail pionnier **Łojasiewicz (63≃65)** $f : \mathbb{R}^n \to \mathbb{R}$ **analytique réelle:**
Autour de tout point $\bar{x}$ **il existe** θ **dans** $[0.5, 1)$ **tel que l'on ait**

$$||\nabla [f]^{1-\theta}(x)|| \geq 1$$

dès que x **est non critique.**

Théorème (Kurdyka 98)

Si f est définissable C^1 dans une structure o-minimale alors f a la propriété KL.

$\boxed{\text{Ex}}$: Extensions C^1 des fonctions de la forme $f(x) = \exp[-Q(x)]$ où Q est une fraction rationnelle...

D'autres résultats

Cas non-lisse sous-analytique/"o-minimal" : B-Daniilidis-Lewis/-Shiota

Cas convexe : voir Olivier Ley

Fonctions de Morse : Génériquement KL est vérifiée et la fonction $\sqrt{\ }$ est désingularisante

Régularité métrique et propriété KL : $F : \mathbb{R}^n \to \mathbb{R}^m$ C^1 dont le jacobien est surjectif sur $C \subset \mathbb{R}^n$,

$$\textbf{Pb : Résoudre } F(x) = 0, \quad x \in C$$

On résout

$$\min_{x \in C} f(x) := ||F(x)||^2.$$

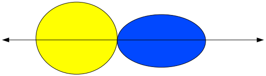
f a la propriété KL et admet $\sqrt{\ }$ pour fonction désingularisante.

Réalisabilité et propriété KL

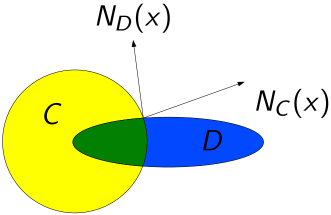
$C, D \subset H$ ss-ensembles fermés. C et D transverses en $x \in C \cap D$ si

$$N_x C \cap [-N_x D] = \{0\}.$$

Non transverse



Transverse



Réalisabilité et propriété KL 2

Pb : trouver $x \in C \cap D$

Remarquer que la bifonction

$$H^2 \ni (x, y) \longrightarrow \begin{cases} \frac{1}{2}||x - y||^2 & \text{si } (x, y) \in C \times D \\ +\infty & \text{sinon.} \end{cases}$$

satisfait KL (si H de DF) et en déduire que

$$\begin{cases} x_{k+1} \in P_C\left(\dfrac{x_k + y_k}{2}\right) \\ y_{k+1} \in P_D\left(\dfrac{y_k + x_{k+1}}{2}\right). \end{cases}$$

converge [B.-Attouch-Redont-Soubeyran].

Pourquoi caractériser ?

Motivations

– Nouveaux outils pour l'analyse des problèmes impliquant des points critiques (cf réalisabilité/algo. alternés/KL)

– Nouveaux outils pour établir de nouvelles inégalités

– Comprendre les limites de validité de KL : a-t-on par exemple "longueur finie des trajectoires de gradient $\Rightarrow$ propriété KL" ?

– Compréhension approfondie des fonctions désingularisantes (impact sur les vitesses de convergence et la complexité des méthodes de gradient)

Aspect métriques : bornes d'erreurs

Azé-Corvellec, Ioffe....

- $d_\varphi(r,s) = |\varphi(r) - \varphi(s)|, \quad r, s \in (0, r_0).$
- **Distance de Hausdorff**
  $D(C,D) = \max\{\sup_{x\in D} \textbf{dist}\ (x,C), \sup_{y\in C} \textbf{dist}\ (y,D)\}$
- **La (multi)application/correspondance sous-niveaux**

$$(0, r_0) \ni r \longrightarrow [f \leq r] \in \mathcal{K} \subset \mathcal{P}(H)$$

Théorème

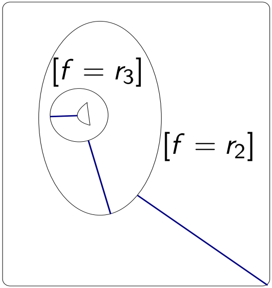
(i) f satisfait KL
(ii) **L'application sous-niveaux** *est lipschitzienne de* $[(0, r_0), d_\varphi]$ *vers* $[\mathcal{K}, D]$.

Talweg : "chemin dans la vallée"

$$\rho > 0, \quad V(r) := \left\{ x \in [f = r] : ||\partial f(x)|| \leq (1+\rho) \inf_{y \in [f=r]} ||\partial f(y)||. \right\}$$

$\boxed{\text{Illustration}}$ Soit $r_1 > r_2 > r_3 \ldots$

$[f = r_1]$



Caractérisation en terme de Talweg

Théorème

(i) f a la propriété KL sur $[0 < f < r_0]$.

(ii) Il existe un **talweg de longueur finie**, i.e. $\exists \theta : (0, r) \to H$ *absolument continue par morceaux telle que*

$$\theta(s) \in V(s), \forall s \in (0, r)$$

De plus

$$\psi(s) = \int_0^s ||\dot{\theta}(\tau)|| d\tau$$

est désingularisante pour f.

$\boxed{\text{Remarque}}$ θ mesure le pire chemin qu'une trajectoire de gradient pourrait suivre

Courbes de sous-gradient

$$(1) \qquad \dot{x}(t) + \partial f(x(t)) \ni 0, \text{pp sur } (0, +\infty) \;\text{ avec } x(0) \in \mathrm{dom} f$$

- Si x est absolument continue et satisfait (1) $\longrightarrow$ c'est une courbe de sous-gradient
- Existence-unicité (Marcellin-Thibault, Brézis, Degiovanni-Marino-Tosques)
- Les valeurs d'adhérence faibles de $x(\cdot)$ sont des points critiques de f
- SI γ est absolument continue par morceaux, satisfait (1) avec $f \circ \gamma$ DECROISSANTE $\longrightarrow$ γ *est une courbe de sous-gradient par morceaux.*

Troisième caractérisation : bornitude uniforme des courbes de gradients par morceaux

Théorème

(i) Propriété K-Ł de f sur $[0 < f < r_0]$.

(ii) Soit Γ l'ensemble des courbes de sous gradient de $[0 < f < r_0]$. Alors

$$\sup_{\gamma \in \Gamma} \textit{longueur } \gamma < +\infty.$$