Module A3
Propriété de Kurdyka-Łojasiewicz

Sauf mention contraire, $\mathcal{X}$ est un espace vectoriel réel de dimension finie, muni du produit scalaire noté $\langle \cdot, \cdot \rangle$ et on note $\|\cdot\|$ la norme euclidienne associée. Pour toutes les autres notations et définitions, le lecteur est invité à se reporter aux modules précédents.

Dans ce module, on va introduire une notion qui prend une place grandissante en optimisation numérique non convexe du premier ordre, à savoir la propriété de Kurdyka-Łojasiewicz. Elle permet en effet dans de nombreuses preuves de convergence d'algorithmes de remplacer l'hypothèse de convexité.

1 Propriété de Kurdyka-Łojasiewicz

1.1 Définition

Commençons par introduire la définition principale de ce module :

Définition 1 (Propriété de Kurdyka-Łojasiewicz)

Soit $f : \mathcal{X} \to \overline{\mathbb{R}}$ une fonction. Soit $x^* \in \operatorname{dom} f$. On dit que f satisfait la propriété de Kurdyka-Łoajasiewicz en x^*, en abrégé la propriété KŁ en x^*, s'il existe $\eta \in \,]\,0\,;+\infty\,[$, un voisinage $\mathcal{V}$ de x^* et $\kappa : [\,0\,;\eta\,] \to [\,0\,;+\infty\,[$ une fonction concave tels que

(i) κ est continue en 0 et $\kappa(0) = 0$;

(ii) κ est continûment dérivable sur $]\,0\,;\eta\,[$;

(iii) $\kappa'(t) > 0$ pour tout $t \in \,]\,0\,;\eta\,[$;

(iv) pour tout $x \in \mathcal{V} \setminus \{x^*\}$ tel que $f(x^*) \leq f(x) \leq f(x^*) + \eta$, on a

$$\kappa'\big(f(x) - f(x^*)\big)\, \operatorname{dist}\big(0, \partial f(x)\big) \geq 1$$

On dit que f est une fonction KŁ si elle satisfait la propriété KŁ en tout point $x^* \in \operatorname{dom} f$ tel que $\partial f(x^*) \neq \emptyset$.

Dans cette définition, la distance est définie par

$$\operatorname{dist}\big(0, \partial f(x)\big) = \inf_{p \in \partial f(x)} \|p\| \in [\,0\,;+\infty\,]$$

la valeur infinie étant atteinte si et seulement si le sous-différentiel de f en x est vide.

Notons d'ores-et-déjà que la condition (iv) de la définition 1 s'applique à des points x voisins de x^* prenant une valeur plus grande que $f(x^*)$. Si de tels points n'existent pas (parce que x^* est un maximiseur local de f), alors n'importe quel fonction κ est acceptable et les conditions (i)–(iv) sont trivialement vérifiées.

EXEMPLE

Exemple de fonction κ. La fonction racine carrée satisfait pour tout $\eta > 0$ les conditions (i)–(iii) apparaissant dans la définition de la propriété KŁ.

1.2 Propriété KŁ et points critiques

Observons que toute fonction est KŁ en tout point non critique :

Proposition 1

Soit $f : \mathcal{X} \to \overline{\mathbb{R}}$ une fonction. Soit $x^0 \in \operatorname{dom} f$. On suppose que $0 \notin \partial f(x^0)$. Alors f satisfait la propriété KŁ en x^0.

DÉMONSTRATION : Scindons cette preuve en deux étapes.

- Commençons par montrer par l'absurde l'existence d'un réel $c > 0$ tel que, pour tout $x \in \operatorname{dom} f$,

$$\|x - x^0\| + \|f(x) - f(x^0)\| < c \qquad \Longrightarrow \qquad \operatorname{dist}\big(0, \partial f(x)\big) \geq c$$

  Supposons que ce n'est pas le cas. Dans ce cas, pour tout $k \in \mathbb{N}$, il existe un point $x_k \in \operatorname{dom} f$ tel que

$$\|x_k - x^0\| + \|f(x_k) - f(x^0)\| < \frac{1}{k+1} \qquad \text{et} \qquad \operatorname{dist}\big(0, \partial f(x_k)\big) < \frac{1}{k+1}$$

  La première inégalité implique que $x_k \to x^0$ et que $f(x_k) \to f(x^0)$; la seconde entraîne que, pour tout $k \in \mathbb{N}$, il existe un vecteur $p_k \in \mathcal{X}$ tel que

$$p_k \in \partial f(x_k) \qquad \text{et} \qquad \|p_k\| < \frac{1}{k+1}$$

  Par comparaison, $p_k \to 0$. Ainsi, la fermeture du sous-différentiel implique que $0 \in \partial f(x^0)$, ce qui contredit l'hypothèse sur x^0.
- Définissons la fonction concave (car linéaire)

$$\kappa : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ t & \mapsto & \dfrac{t}{c} \end{array} \right. \qquad \text{avec} \qquad \forall\, t \in \mathbb{R}, \quad \kappa'(t) = \frac{1}{c}$$

  Cette fonction est continûment dérivable sur $\mathbb{R}$, avec $\kappa(0) = 0$ et κ' est strictement positif sur $\mathbb{R}$. Considérons le voisinage $\mathcal{V} = \mathcal{B}(x^0, c/2)$ de x^0 et posons $\eta = c/2$. Alors pour tout $x \in \mathcal{X}$ tel que $\|x - x^0\| < c/2$, on a

$$f(x^0) \leq f(x) \leq f(x^0) + \frac{c}{2} \qquad \Longrightarrow \qquad 0 \leq f(x) - f(x^0) \leq \frac{c}{2}$$

  de sorte que le point précédent montre que, pour tout $x \in \mathcal{V}$,

$$f(x^*) \leq f(x) \leq f(x^*) + \eta \qquad \Longrightarrow \qquad \kappa'\big(f(x) - f(x^*)\big) \operatorname{dist}\big(0, \partial f(x)\big) \geq 1$$

  ce qui prouve le résultat désiré. ■

Ainsi, on voit que les seuls points où la propriété KŁ présente un intérêt sont les points critiques, c'est-à-dire les points $x^* \in \operatorname{dom} f$ tels que $0 \in \partial f(x^*)$, et qui, de plus, ne sont pas maximiseurs locaux (attention : dans le cas non différentiable, un maximiseur local n'est pas nécessairement point critique !).

EXEMPLE

Exemple de fonction KŁ différentiable. Soit $f = \|\cdot\|^2$ et $x^* = 0$ son unique point critique. En choisissant κ la fonction racine carrée, on montre que f est une fonction KŁ puisque pour tout $x \neq 0$,

$$\kappa'\big(f(x) - f(0)\big)\, \|\nabla f(x)\| = \frac{1}{2\sqrt{\|x\|^2}} \times 2\,\|x\| = 1$$

Revenons à la définition de la propriété KŁ, en particulier la propriété (iv). Remarquons que la distance qui apparaît dans cette propriété est nulle si x est un point critique ; dans ce cas,

$$\kappa'\big(f(x) - f(x^*)\big)\, \mathrm{dist}\big(0, \partial f(x)\big) = 0 < 1$$

Aussi, une fonction satisfaisant la propriété KŁ en un point (critique) x^* est nécessairement telle qu'il existe $\eta > 0$ vérifiant pour x **voisin** de x^*

$$f(x^*) \leq f(x) \leq f(x^*) + \eta \qquad \Longrightarrow \qquad 0 \notin \partial f(x)$$

Supposons que x^* est un minimiseur local de $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$. Par définition du minimiseur local, pour tout point $x \in \mathcal{X}$ voisin de x^*, on a $f(x^*) \leq f(x)$. La propriété KŁ assure donc qu'il existe un voisinage $\mathcal{V}$ de x^* tel que $0 \notin \partial f(x)$ pour tout $x \in \mathcal{V}$. Autrement dit, aucun point de $\mathcal{V} \setminus \{x^*\}$ n'est critique (et, en particulier, n'est minimiseur local). Ainsi, si f est KŁ en x^* un minimiseur local, alors ce minimiseur local est nécessairement *isolé*.

CONTRE-EXEMPLE

Fonction constante Soit $c \in \mathbb{R}$. Considérons la fonction constante définie par

$$\forall\, x \in \mathcal{X}, \qquad f(x) = c$$

Soit $x^* \in \mathcal{X}$. Pour tout x voisin de x^*, on a $f(x) = f(x^*)$ et $\nabla f(x) = 0$, de sorte que, pour tout voisinage $\mathcal{V}$ de x^*, tout $\eta > 0$ et toute fonction concave $\kappa : [\,0\,;\eta\,]$ satisfaisant les hypothèses de la définition 1, on a

$$\forall\, x \in \mathcal{V} \setminus \{x^*\}, \qquad \|\nabla(\kappa \circ f)(x)\| = 0 < 1$$

Ainsi, les fonctions constantes ne sont KŁ en aucun point de leur domaine.

Pour mieux comprendre l'intérêt de la propriété KŁ, réécrivons le quatrième point de la définition : pour tout $x \in \mathcal{V} \setminus \{x^*\}$, on a

$$f(x^*) \leq f(x) \leq f(x^*) + \eta \qquad \Longrightarrow \qquad \|\nabla f(x)\|\, \kappa' \circ f(x) \geq 1$$

On peut vérifier que cette dernière inégalité s'écrit

$$\|\nabla(\kappa \circ f)(x)\| \geq 1$$

Si x^* est un minimiseur local isolé, alors la quatrième condition apparaissant dans la définition de la propriété KŁ se lit

$$\forall\, x \in \mathcal{V} \setminus \{x^*\}, \qquad \|\nabla(\kappa \circ f)(x)\| \geq 1$$

La fonction κ pouvant être vue comme une fonction de reparamétrisation, la définition de la propriété KŁ pour un point critique peut être interprétée de la manière suivante : à une reparamétrisation locale de f (au voisinage de x^*) près, la fonction f n'est pas *plate* au voisinage x^*.

La fonction de reparamétrisation κ peut être vue comme une fonction *désingularisante*, en ce sens où elle transforme les gradients de f arbitrairement proches de zéro (donc "singulières") en des éléments à l'extérieur de la boule unité (donc "régulières").

CONTRE-EXEMPLE

Exemple de fonction non KŁ. Considérons la fonction définie par

$$\forall\, x \in \mathbb{R}, \qquad f(x) = \begin{cases} \exp\left(-\dfrac{1}{x^2}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Le graphe de cette fonction est représenté en figure 2. La fonction f est dérivable (et même de classe $\mathcal{C}^\infty$) sur $\mathbb{R}$. Sa dérivée vaut

$$\forall\, x \in \mathbb{R}, \qquad f'(x) = \begin{cases} \dfrac{2}{x^3} \exp\left(-\dfrac{1}{x^2}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Alors on peut montrer que f ne satisfait pas la propriété KŁ en 0, qui est pourtant un minimiseur isolé (puisque $f(x) > 0$ pour tout $x \neq 0$.

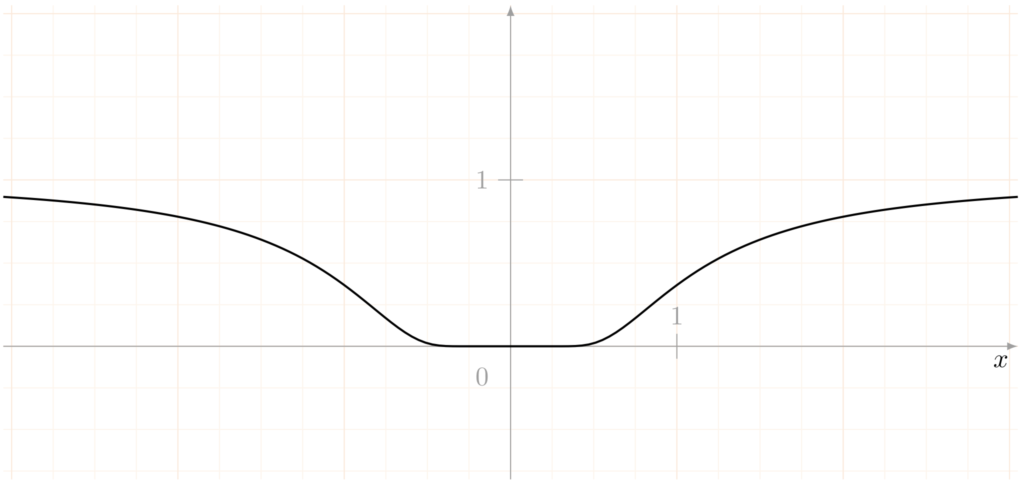

FIGURE 1 – Exemple de fonction non KŁ.

De même, on peut trouver des exemples de fonctions convexes qui ne sont pas KŁ.

2 Classes de fonctions KŁ

La propriété KŁ, parce qu'elle repose sur l'**existence** d'une fonction désingularisante κ difficile à déterminer, est en général difficile à vérifier. Dans cette section, on va donner quelques résultats qui permettent de vérifier, sans expliciter cette fonction κ, qu'une fonction est KŁ.

2.1 Fonctions fortement convexes

On a signalé plus haut que certaines fonctions convexes ne sont pas KŁ. En revanche, toutes les fonctions fortement convexes le sont :

Proposition 2 (Condition de POLYAK–ŁOJASIEWICZ)

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction. On suppose que f est fortement convexe, de module α, et propre. Soit x^* le minimiseur de f. Alors on a

$$\forall\, x \in \operatorname{dom} f, \forall\, p \in \partial f(x), \qquad \|p\|^2 \geq 2\,\alpha\,\big(f(x) - f(x^*)\big)$$

Ainsi, si $x \neq x^*$, on a $f(x^*) < f(x)$ (le minimiseur d'une fonction fortement convexe étant unique), de sorte que la condition de POLYAK–ŁOJASIEWICZ devient

$$\forall\, x \in \operatorname{dom} f, \qquad \frac{1}{\sqrt{2\,\alpha\,\big(f(x) - f(x^*)\big)}}\operatorname{dist}\big(0, \partial f(x)\big) \geq 1$$

ce qui est la définition de la propriété KŁ en l'unique point critique x^* de f (il suffit de prendre $\kappa(t) = \sqrt{2\,t/\alpha}$).

DÉMONSTRATION : Soit $x \in \operatorname{dom} f$ et $p \in \partial f(x)$ (le sous-différentiel est supposé non vide). D'après le corollaire 1 du module **A2 : Sous-différentiabilité**, on a

$$f(x) - f(x^*) \leq \langle p, x - x^* \rangle - \frac{\alpha}{2}\,\|x - x^*\|^2 = \frac{1}{2\,\alpha}\,\|p\|^2 - \frac{1}{2}\,\left\|\frac{p}{\sqrt{\alpha}} - \sqrt{\alpha}\,(x - x^*)\right\|^2$$

où l'on a utilisé une identité remarquable. On conclut en remarquant que le dernier terme de cet encadrement est négatif. ■

> En réalité, dans ce cours, la propriété KŁ étant utilisée pour remplacer l'hypothèse de convexité dans les preuves de convergence, ce résultat ne nous sera pas utile.

2.2 Fonctions polynomiales et polynomiales par morceaux

Un premier exemple important de fonctions KŁ est celui des fonctions polynomiales.

Proposition 3

On suppose que $\mathcal{X}$ est de dimension finie. Soit $P : \mathcal{X} \to \mathbb{R}$ une fonction polynomiale. Alors P est KŁ.

DÉMONSTRATION : Admis.

REMARQUE : Il s'agit d'un cas particulier des fonctions analytiques réelles, dont on va voir qu'elles sont aussi KŁ dans le paragraphe suivant.

Dans de nombreuses applications, on utilise également souvent des fonctions polynomiales par morceaux.

Définition 2 (Fonction polynomiale par morceaux)

Soit $n \in \mathbb{N}^*$. Soit $a_1 \leq \cdots \leq a_n$ des réels et $\{P_i\}_{1 \leq i \leq n+1}$ une famille de fonctions polynomiales réelles. On appelle *fonction polynomiale par morceaux* la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par

$$\forall\, x \in \mathbb{R}, \qquad f(x) = \begin{cases} P_1(x) & \text{si } x \leq a_1 \\ P_i(x) & \text{si } a_i < x \leq a_{i+1} \\ P_{n+1}(x) & \text{si } a_{n+1} < x \end{cases}$$

.

Proposition 4

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction polynomiale par morceaux. Alors f est KŁ.

Démonstration : Admis. La preuve utilise la notion de *fonction semi-algébrique.*

En particulier, la fonction valeur absolue est KŁ. Le caractère KŁ est par ailleurs préservé par passage à la racine carrée dans ce cas :

Proposition 5

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction polynomiale par morceaux. Alors $\sqrt{f}$ est KŁ.

Démonstration : Admis. La preuve utilise la notion de *fonction semi-algébrique.*

> En réalité, les fonctions polynomiales par morceaux appartiennent à une famille de fonctions appelées *semi-algébriques*. Celle-ci est en particulier stable par passage à la valeur absolue et à la racine carrée, mais également par somme, multiplication par un scalaire, produit et composition. De plus, les fonctions semi-algébriques sont KŁ.

> Dans ce cours, on ne considère que les fonctions polynomiales par morceaux **réelles**. Il est en fait possible de généraliser les deux propositions précédentes à des fonctions polynomiales par morceaux définies sur un espace euclidien. Il faut cependant pour cela introduire la notion d'*ensembles semi-algébriques* pour partitionner le domaine de la fonction. Signalons que, parmi eux, se trouvent les sous-espaces vectoriels ou les demi-espaces, et que les ensembles semi-algébriques sont stables par union finie, intersection finie et produit cartésien.

2.3 Fonctions analytiques réelles

Dans ce paragraphe, on introduit une autre classe de fonctions KŁ, les fonctions *analytiques* réelles.

Définition 3 (Fonction analytique)

Soit $U \subset \mathbb{R}$ un ouvert. Soit $f : U \to \mathbb{R}$ une fonction. On dit que f est *analytique* sur U si pour tout $a \in U$, il existe une suite réelle $(a_k)_{k\in\mathbb{N}}$ et $r > 0$ tel que

$$\forall\, x \in \,]\, a - r\,; a + r\,[\,, \qquad f(x) = \sum_{k=1}^{+\infty} a_k\,(x-a)^k$$

Autrement dit, f est analytique si et seulement si f est développement en série entière au voisinage de tout point de son domaine U.

Proposition 6 (Lojasiewicz)

Soit $f : U \to \mathbb{R}$ une fonction analytique. Alors il existe $\theta \in [\,1/2\,; 1\,[$ tel que pour tout $x \in U$,

$$f(x^*) \leq f(x) \leq f(x^*) + \eta \qquad \Longrightarrow \qquad \|\nabla f(x)\| \geq \big(f(x) - f(x^*)\big)^\theta$$

Autrement dit, toute fonction analytique réelle est KŁ ; en effet, il suffit de considérer la fonction désingularisante $\kappa(t) = t^{1-\theta}$.

Démonstration : Admis.

Néanmoins, en pratique, les fonctions rencontrées dans les applications sont rarement analytiques ou polynomiales par morceaux. Il s'agit plus couramment de somme / composition de telles fonctions.

Proposition 7 (Somme)

Soit $f : U \to \mathbb{R}$ une fonction analytique et $g : \mathbb{R} \to \mathbb{R}$ une fonction polynomiale par morceaux. On suppose que f et g sont minorées et que $\operatorname{dom} g$ est fermé. Alors $f + g$ et $f + \sqrt{g}$ sont KŁ.

Démonstration : Admis. La preuve utilise la notion de *fonction sous-analytique.*

Les fonctions analytiques sont stables par somme et produit. Aussi, la composition avec une fonction polynomiale reste analytique. On a par ailleurs le résultat plus général suivant :

Proposition 8 (Composition I)

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction polynomiale et $g_i : U \to \mathbb{R}$ une fonction analytique pour tout $i \in [\![\, 1\,; n \,]\!]$. On suppose par ailleurs que pour tout $E \subset U$ borné, les $g_i(E)$ sont bornés. Alors $f \circ g$ et $\sqrt{f} \circ g$ sont KŁ.

Démonstration : Admis. La preuve utilise la notion de *fonction sous-analytique.*

Proposition 9 (Composition II)

Soit $f : U \to \mathbb{R}$ une fonction analytique et $g : \mathbb{R} \to \mathbb{R}$ une fonction polynomiale par morceaux telle que $g(\mathbb{R}) \subset U$ (resp. $\sqrt{g}(\mathbb{R}) \subset U$). On suppose par ailleurs que pour tout $\mathcal{E} \subset \mathcal{Y}$ borné, $f^{-1}(\mathcal{E})$ est borné. Alors $f \circ g$ (resp. $f \circ \sqrt{g}$) est KŁ.

Démonstration : Admis. La preuve utilise la notion de *fonction sous-analytique*.

> On peut démontrer ces résultats en passant par la notion de fonctions *sous-analytiques*, dont font partie les fonctions analytiques réelles et les fonctions semi-algébriques. Les fonctions analytiques sont KŁ et l'ensemble de ces fonctions est stable (sous certaines conditions) par somme et composition.

3 Théorème d'Attouch–Bolte–Svaiter

On termine ce module par un résultat central qui démontre l'utilité de la propriété KŁ en optimisation.

Théorème 1 (Attouch, Bolte & Svaiter)

Soit $J : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction s.c.i. Soit $(x_k)_{k\in\mathbb{N}}$ une suite de $\mathcal{X}$ vérifiant les trois conditions suivantes :

(C1) (condition de décroissance) il existe $a > 0$ tel que, pour tout $k \in \mathbb{N}$,

$$J(x_{k+1}) + a\,\|x_{k+1} - x_k\|^2 \leq J(x_k)$$

(C2) (condition d'erreur relative) il existe $b > 0$ tel que, pour tout $k \in \mathbb{N}$, il existe $p_{k+1} \in \partial J(x_{k+1})$ vérifiant

$$\|p_{k+1}\| \leq b\,\|x_{k+1} - x_k\|$$

(C3) (condition de continuité) il existe une sous-suite $(x_{k_j})_{j\in\mathbb{N}}$ et $x^* \in \operatorname{dom} J$ tels que

$$\lim_{j\to+\infty} x_{k_j} = x^* \qquad \text{et} \qquad \lim_{j\to+\infty} J(x_{k_j}) = J(x^*)$$

On suppose que J satisfait la propriété KŁ en x^*. Alors la suite $(x_k)_{k\in\mathbb{N}}$ converge vers x^*.

Démonstration : Décomposons cette preuve en plusieurs étapes distinctes.

- **Convergence de $(J(x_k))_{k\in\mathbb{N}}$.** Commençons par noter que la condition (C1) implique la décroissance de la suite réelle $(J(x_k))_{k\in\mathbb{N}}$. Ainsi, il y a deux possibilités : soit elle diverge vers $-\infty$, soit elle est convergente. La condition (C3) assure que le second cas s'applique ; l'unicité de la limite implique alors que $(J(x_k))_{k\in\mathbb{N}}$ converge vers $J(x^*)$. Par ailleurs, on a

$$\forall\, k \in \mathbb{N}, \qquad J(x_k) \geq J(x^*)$$

- **Propriété KŁ.** Soit $\delta > 0$ tel que $\mathcal{B}(x^*, \delta) \subset \mathcal{V}$ et $\rho \in \,]\,0\,;\delta\,[$, où $\mathcal{V}$ est le voisinage de x^* apparaissant dans la définition de la propriété KŁ. Sans perte de généralité, on peut supposer qu'il existe $\eta < a\,(\delta-\rho)^2$ tel que une fonction concave $\kappa : [\,0\,;\eta\,] \to [\,0\,;+\infty\,[$ tels que

(i) $\kappa(0)=0$;
(ii) κ est continûment dérivable sur $]\,0\,;\eta\,[$;
(iii) $\kappa'(t)>0$ pour tout $t\in\,]\,0\,;\eta\,[$;
(iv) pour tout $x\in\mathcal{V}$ tel que $J(x^*)\le J(x)\le J(x^*)+\eta$, on a

$$\kappa'\big(J(x)-J(x^*)\big)\,\mathrm{dist}\big(0,\partial J(x)\big)\ge 1$$

- **Définition de $x_{k_{j_0}}$.** La convergence de $(J(x_k))_{k\in\mathbb{N}}$ vers $J(x^*)$ implique l'existence de k_0 tel que

$$\forall\, k\ge k_0,\qquad J(x^*)\le J(x_k)\le J(x^*)+\eta$$

Par ailleurs, puisque κ est continue au voisinage de 0 et que la sous-suite $(x_{k_j})_{j\in\mathbb{N}}$ tend vers x^*, il existe $j_0\in\mathbb{N}$ tel que

$$\|x_{k_{j_0}}-x^*\|+2\sqrt{\frac{J(x_{k_{j_0}})-J(x^*)}{a}}+\frac{b}{a}\,\kappa\big(J(x_{k_{j_0}})-J(x^*)\big)<\rho$$

Sans perte de généralité, on peut supposer que $k_0\le k_{j_0}$. Notons que, grâce à la condition (C1), $x_{k_{j_0}}$ vérifie également

$$\|x_{k_{j_0}}-x^*\|+2\,\|x_{k_{j_0}+1}-x_{k_{j_0}}\|+\frac{b}{a}\,\kappa\big(J(x_{k_{j_0}})-J(x^*)\big)<\rho \qquad (\star)$$

car $\qquad \|x_{k_{j_0}+1}-x_{k_{j_0}}\|\le\sqrt{\dfrac{J(x_{k_{j_0}})-J(x_{k_{j_0}+1})}{a}}\le\sqrt{\dfrac{J(x_{k_{j_0}})-J(x^*)}{a}}$

- **Propriété KŁ en x^*.** Soit $k>k_{j_0}$ tel que $\|x_k-x^*\|<\delta$. Par hypothèse sur δ, on a donc $x_k\in\mathcal{V}$. On suppose que $x_{k+1}\ne x_k$. Alors, d'après la condition (C1) et la définition de k_{j_0}, on a

$$J(x^*)\le J(x_{k+1})+a\,\|x_{k+1}-x_k\|^2<J(x_k)<J(x^*)+\eta$$

D'après la propriété KŁ en x^* appliqué au point x_k, on a alors

$$\kappa'\big(J(x_k)-J(x^*)\big)\,\mathrm{dist}\big(0,\partial J(x_k)\big)\ge 1$$

Il s'ensuit que $0\notin\partial J(x_k)$ (en particulier, $p_k\ne 0$). Or, puisque, par définition de la distance, $\mathrm{dist}\big(0,\partial J(x_k)\big)\le\|p_k\|$, on a d'après la condition (ii),

$$x_{k-1}\ne x_k\qquad\text{et}\qquad \kappa'\big(J(x_k)-J(x^*)\big)\ge\frac{1}{b\,\|x_k-x_{k-1}\|}$$

- **Concavité de κ.** La convexité de $-\kappa$ entraîne d'une part que

$$\begin{aligned}-\kappa\big(J(x_{k+1})-J(x^*)\big)\ge&-\kappa\big(J(x_k)-J(x^*)\big)\\&-\kappa'\big(J(x_k)-J(x^*)\big)\,\big(J(x_{k+1})-J(x_k)\big)\end{aligned}$$

D'autre part, en utilisant la positivité de κ', la condition (C1) et en réarrangeant les termes, on obtient que

$$\begin{aligned}&\kappa\big(J(x_k)-J(x^*)\big)-\kappa\big(J(x_{k+1})-J(x^*)\big)\\&\qquad\ge\kappa'\big(J(x_k)-J(x^*)\big)\,a\,\|x_{k+1}-x_k\|^2\end{aligned}$$

de sorte qu'en combinant cette inégalité avec l'inégalité KŁ écrite au point précédent, on obtient

$$\kappa\big(J(x_k)-J(x^*)\big)-\kappa\big(J(x_{k+1})-J(x^*)\big)\ge\frac{a\,\|x_{k+1}-x_k\|^2}{b\,\|x_k-x_{k-1}\|}$$

soit encore, en utilisant l'inégalité $2\sqrt{\alpha\,\beta} \leq \alpha + \beta$ valable pour tous $\alpha, \beta \geq 0$,

$$2\,\|x_{k+1} - x_k\| \leq \|x_k - x_{k-1}\| + \frac{b}{a}\left(\kappa\big(J(x_k) - J(x^*)\big) - \kappa\big(J(x_{k+1}) - J(x^*)\big)\right) \quad (\star\star)$$

L'inégalité $(\star\star)$ reste trivialement vraie lorsque $x_{k+1} = x_k$.

- $\|x_k - x^*\| < \rho \implies \|x_{k+1} - x^*\| < \delta$**.** Soit $k \in \mathbb{N}$ tel que $\|x_k - x^*\| < \rho$. Alors, d'après la condition (C1), on a

$$\|x_{k+1} - x_k\| \leq \sqrt{\frac{J(x_k) - J(x_{k+1})}{a}} \leq \sqrt{\frac{\eta}{a}} < \delta - \rho$$

Il s'ensuit en particulier que

$$\|x_{k+1} - x^*\| \leq \|x_{k+1} - x_k\| + \|x_k - x^*\| < \delta - \rho + \rho = \delta$$

soit $x_{k+1} \in \mathcal{B}(x^*, \delta) \subset \mathcal{V}$.

- **Pour tout** $k \geq k_{j_0} + 1$**,** $x_k \in \mathcal{B}(x^*, \rho)$ **: initialisation de la récurrence.** On va démontrer par récurrence que, pour tout $k \geq k_{j_0} + 1$, on a $\|x_k - x^*\| < \rho$ et que

$$\sum_{\ell=k_{j_0}+1}^{k-1} \|x_{\ell+1} - x_\ell\| + 2\,\|x_{k+1} - x_k\| \leq \|x_{k_{j_0}+1} - x_{k_{j_0}}\| + \frac{b}{a}\,\kappa\big(J(x_{k_{j_0}+1}) - J(x^*)\big) - \kappa\big(J(x_{k+1}) - J(x^*)\big)$$

Commençons par l'initialisation. L'inégalité triangulaire, combinée avec la relation $(\star)$, permet de montrer que

$$\|x_{k_{j_0}+1} - x^*\| \leq \|x_{k_{j_0}+1} - x_{k_{j_0}}\| + \|x_{k_{j_0}} - x^*\| < \rho$$

Par conséquent, $\|x_{k_{j_0}+1} - x^*\| < \delta$ et la propriété KŁ en x^* (équation $(\star\star)$) assure donc que

$$2\,\|x_{k_{j_0}+2} - x_{k_{j_0}+1}\| \leq \|x_{k_{j_0}+1} - x_{k_{j_0}}\| + \frac{b}{a}\left(\kappa\big(J(x_{k_{j_0}+1}) - J(x^*)\big) - \kappa\big(J(x_{k_{j_0}+2}) - J(x^*)\big)\right)$$

- **Pour tout** $k > k_{j_0} + 1$**,** $x_k \in \mathcal{B}(x^*, \rho)$ **: démonstration par récurrence.** Soit $k > k_{j_0} + 1$. Supposons que $\|x_k - x^*\| < \rho$ et que

$$\sum_{\ell=k_{j_0}+1}^{k-1} \|x_{\ell+1} - x_\ell\| + 2\,\|x_{k+1} - x_k\| \leq \|x_{k_{j_0}+1} - x_{k_{j_0}}\| + \frac{b}{a}\left(\kappa\big(J(x_{k_{j_0}+1}) - J(x^*)\big) - \kappa\big(J(x_{k+1}) - J(x^*)\big)\right)$$

La première propriété implique $\|x_{k+1} - x^*\| < \delta$. Aussi, on peut appliquer la relation $(\star)$ à l'indice $k+1$, ce qui donne

$$2\,\|x_{k+2} - x_{k+1}\| \leq \|x_{k+1} - x_k\| + \frac{b}{a}\left(\kappa\big(J(x_{k+1}) - J(x^*)\big) - \kappa\big(J(x_{k+2}) - J(x^*)\big)\right)$$

Sommons ces deux inégalités ; on démontre ainsi que la seconde inégalité de l'hypothèse de récurrence est vraie au rang $k+1$. On termine en écrivant à l'aide de l'inégalité triangulaire que

$$\|x_{k+1} - x^*\| \leq \|x_{k+1} - x_k\| + \sum_{\ell=k_{j_0}+1}^{k-1} \|x_{\ell+1} - x_\ell\| + \|x_{k_{j_0}+1} - x_{k_{j_0}}\| + \|x_{k_{j_0}} - x^*\|$$

On majore la somme à l'aide de l'hypothèse de récurrence ; on obtient alors

$$\|x_{k+1} - x^*\| \leq \|x_{k_{j_0}} - x^*\| + 2\,\|x_{k_{j_0}+1} - x_{k_{j_0}}\| + \frac{b}{a}\,\kappa\big(J(x_{k_{j_0}+1}) - J(x^*)\big)$$

Il suffit alors d'utiliser la croissance de κ, la condition (C1) et la relation $(\star\star)$ pour obtenir le résultat désiré. ■

Faisons quelques commentaires sur ce théorème. On applique généralement ce théorème sur une suite $(x_k)_{k\in\mathbb{N}}$ générée par un algorithme $\mathcal{A}$ d'optimisation du premier ordre (cf. module **B1 : Méthodes d'optimisation du premier ordre**).

1. Lorsque la condition de décroissance est vérifiée, la suite $(J(x_k))_{k\in\mathbb{N}}$ est décroissante, puisque $a\,\|x_{k+1}-x_k\|^2$ est positif. Aussi, si J est minorée (par exemple si J admet un minimiseur), alors la suite $(J(x_k))_{k\in\mathbb{N}}$ est décroissante minorée, donc convergente, de sorte que la propriété de convergence en valeur / convergence du critère (B-I) est vérifiée.
2. Dans ce cas, l'inégalité de la condition de décroissance assure par comparaison que la suite $(\|x_{k+1}-x_k\|)_{k\in\mathbb{N}}$ converge vers 0. Ainsi, si de plus la condition d'erreur relative est vérifiée, la suite des sous-gradients p_k converge vers 0 et la propriété de convergence vers le critère d'optimalité (C-II) est satisfaite.
3. Enfin, la condition de continuité implique que la suite $(x_k)_{k\in\mathbb{N}}$ admet une sous-suite qui converge ; si par ailleurs les itérées x_k appartiennent au domaine de J, l'hypothèse de continuité de J sur son domaine, si celui-ci est fermé (ce qui implique la semi-continuité inférieure de J) suffit.
4. Les hypothèses du théorème d'Attouch–Bolte–Svaiter empêchent donc la suite des itérées $(x_k)_{k\in\mathbb{N}}$ d'avoir plusieurs points d'adhérence.

Pour aller plus loin

Propriété KŁ. Cette propriété permet de remplacer les hypothèses de régularité et/ou de convexité dans certains algorithmes d'optimisation, comme ceux étudiés dans les modules **A5 : Opérateur proximal de Moreau** et **B4 : Éclatement primal d'opérateurs**. Il faut noter que vérifier qu'une fonction est KŁ peut être une tâche relativement ardue, et que les hypothèses du théorème d'Attouch–Bolte–Svaiter peuvent être difficiles à vérifier. Néanmoins, on verra qu'il s'agit de propriétés qui apparaissent naturellement pour de nombreux algorithmes classiques et leurs variantes.

Insistons enfin sur le fait que les exemples de fonctions KŁ présentés dans ce module sont volontairement simples. La théorie des fonctions KŁ dévoile son intérêt lorsque l'on considère plus largement les notions de fonctions semi-algébriques et d'ensembles semi-algébriques.