Module A3
Éléments de topologie

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.

1 Semi-continuité inférieure

1.1 Définition et exemples

Définition 1 (Fonction semi-continue inférieurement)

Soit $f : \mathcal{X} \to \mathbb{R}\cup\{+\infty,-\infty\}$ une fonction. Soit $x^0 \in \mathcal{X}$. On dit que f est *semi-continue inférieurement* en x^0, abrégé en *s.c.i.* en x^0, si pour toute suite $(x_k)_{k\in\mathbb{N}} \in \mathcal{X}^{\mathbb{N}}$,

$$\lim_{k\to+\infty} x_k = x^0 \qquad \Longrightarrow \qquad \liminf_{k\to+\infty} f(x_k) \geq f(x^0)$$

On dit que f est s.c.i. si elle est s.c.i. en tout point $x^0 \in \mathcal{X}$.

Cette définition est connue sous le nom de *caractérisation séquentielle* de la semi-continuité inférieure. On rappelle que

$$\liminf_{k\to+\infty} u_k = \sup_{k\in\mathbb{N}} \left\{ \inf_{n\geq k} u_n \right\}$$

(lorsque les u_k ne sont pas finis, on se reportera au module B1 pour une définition plus rigoureuse de ces notions). Une autre manière équivalente d'écrire cette définition est d'utiliser la notion de limite inférieure d'une fonction en un point :

$$\liminf_{x\to x^0} f(x) \geq f(x^0)$$

Exemple

Exemple d'une fonction s.c.i. Considérons la fonction réelle définie par

$$\forall\, x\in\mathbb{R}, \qquad f(x) = \begin{cases} x+1 & \text{si } x>0 \\ x^2 & \text{si } x \leq 0 \end{cases}$$

Son graphe est représenté en figure 1. Cette fonction est continue, donc s.c.i., en tout point $\mathbb{R}\setminus\{0\}$. En 0, on a

$$f(0) = 0 \qquad \text{et} \qquad \forall\, x\neq 0, \qquad f(x)\in\{0,1\}$$

En particulier, on a $f(x)\geq 0$. Il s'ensuit que

$$\liminf_{x\to 0} f(x) \geq 0 = f(0)$$

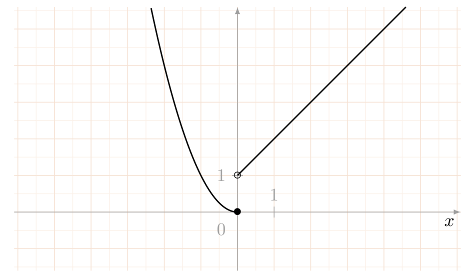

Figure 1 – Exemple de fonction s.c.i. non continue.

Comme on l'a vu à l'exemple précédent, une fonction s.c.i. n'est pas nécessairement continue. On a cependant immédiatement le résultat suivant :

Proposition 1

Soit $f : \mathcal{X} \to \mathbb{R}\cup\{+\infty\}$ une fonction. Soit $x^0 \in \operatorname{dom} f$. On suppose que f est continue en x^0. Alors f est s.c.i. en x^0.

Démonstration : Soit $(x_k)_{k\in\mathbb{N}}$ une suite de points de $\mathcal{X}$ convergeant vers x^0. Puisque $\mathcal{X} = \operatorname{dom} f \cup {}^{c}\operatorname{dom} f$, deux cas de figure sont possibles.

- **On peut extraire de la suite $(x_k)_{k\in\mathbb{N}}$ une sous-suite dans $\operatorname{dom} f$.** Dans ce cas, on a pour tout $n\in\mathbb{N}$

$$\liminf_{k\to+\infty} f(x_k) = \inf_{n\geq k} f(x_n) = \inf_{\substack{n\geq k \\ x_n\in\operatorname{dom} f}} f(x_n) = \lim_{\substack{k\to+\infty \\ x_k\in\operatorname{dom} f}} f(x_k) = f(x^0)$$

  où la limite est une conséquence de la continuité de f en x^0 ; en passant à la borne supérieure sur les $n\in\mathbb{N}$, on obtient

$$\sup_{n\in\mathbb{N}} \inf_{n\geq k} f(x_n) = f(x^0)$$

- **On ne peut pas extraire de la suite $(x_k)_{k\in\mathbb{N}}$ une sous-suite dans $\operatorname{dom} f$.** Dans ce cas, il existe un rang $k^0\in\mathbb{N}$ à partir duquel les x_k n'appartiennent pas à $\operatorname{dom} f$. Il s'ensuit que

$$\liminf_{k\to+\infty} f(x_k) = \lim_{k\to+\infty} f(x_k) = +\infty > f(x^0)$$

Dans les deux cas, on a démontré que f est s.c.i. en x^0. ■

Exercice

Fermeture d'une fonction convexe. Soit $f : \mathcal{X} \to \mathbb{R}\cup\{+\infty\}$ une fonction convexe. On définit

$$\underline{f} : \left\{ \begin{array}{ccc} \mathcal{X} & \to & \mathbb{R}\cup\{+\infty\} \\ x & \mapsto & \liminf_{z\to x} f(z) \end{array} \right.$$

la fermeture de f. Montrer que $\underline{f}$ est une fonction (convexe) s.c.i.

Un exemple important est celui des indicatrices d'ensembles fermés.

Proposition 2

Soit $\mathcal{E} \subset \mathcal{X}$ un ensemble fermé non vide. Alors la fonction $\chi_{\mathcal{E}} : \mathcal{X} \to \{0, +\infty\}$ définie par

$$\forall\, x \in \mathcal{X}, \qquad \chi_{\mathcal{E}}(x) = \begin{cases} 0 & \text{si } x \in \mathcal{E} \\ +\infty & \text{sinon} \end{cases}$$

est s.c.i.

Démonstration : Soit $x^0 \in \mathcal{X}$ et soit $(x_k)_{k\in\mathbb{N}}$ une suite de points de $\mathcal{X}$ convergeant vers x^0. Puisque $\mathcal{X} = \mathcal{E} \cup {}^c\mathcal{E}$, on peut considérer deux cas de figure :

- Si $x^0 \in \mathcal{E}$, alors on a
$$\chi_{\mathcal{E}}(x^0) = 0 \qquad \text{et} \qquad \forall\, k \in \mathbb{N}, \quad \chi_{\mathcal{E}}(x_k) \in \{0; +\infty\}$$
ce qui implique en particulier que
$$0 = \chi_{\mathcal{E}}(x^0) \leq \liminf_{n\to+\infty} \chi_{\mathcal{E}}(x_k) \in \{0; +\infty\}$$
La fonction $\chi_{\mathcal{E}}$ est donc semi-continue inférieurement en tout $x^0 \in \mathcal{E}$.
- Si $x^0 \notin \mathcal{E}$, alors on peut vérifier que l'ensemble
$$\Big\{ k \in \mathbb{N} \mid x_k \in \mathcal{E} \Big\}$$
est fini. En effet, dans le cas contraire, on pourrait extraire une sous-suite convergente de $(x_k)_{k\in\mathbb{N}}$ dont tous les termes appartiennent à $\mathcal{E}$. Puisque $\mathcal{E}$ est supposé fermé, il s'ensuivrait que la limite de cette sous-suite appartient également à $\mathcal{E}$; par unicité de la limite, cela contredit l'hypothèse considérée dans ce cas de figure. Ainsi, il existe un indice $k_0 \in \mathbb{N}$ tel que
$$\forall\, k \geq k_0, \qquad x_k \notin \mathcal{E}$$
de sorte que, pour tout $k \geq k_0$, on a $\chi_{\mathcal{E}}(x_k) = +\infty$. Par conséquent,
$$\liminf_{n\to+\infty} \chi_{\mathcal{E}}(x_k) = +\infty = \chi_{\mathcal{E}}(x^0)$$
ce qui assure la semi-continuité inférieure de $\chi_{\mathcal{E}}$ en tout point $x^0 \notin \mathcal{E}$. ■

On peut généraliser ce résultat aux fonctions continues sur leur domaine fermé :

Proposition 3

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction. On suppose que $\operatorname{dom} f$ est fermé et que f est continue sur son domaine. Alors f est s.c.i.

Démonstration : Commençons par remarquer que, pour tout $x^0 \in \operatorname{dom} f$, la fonction f est continue en x^0, donc s.c.i. en x^0 (proposition 1). On s'intéresse donc aux points $x^0 \notin \operatorname{dom} f$. Par hypothèse sur le domaine de f, son complémentaire est ouvert, donc il existe $\varepsilon > 0$ tel que

$$\forall\, x \in \mathcal{X}, \qquad \|x - x^0\| \leq \varepsilon \qquad \Longrightarrow \qquad x \notin \operatorname{dom} f$$

Autrement dit, pour tout tel x, on a $f(x) = +\infty$. Ainsi, pour toute suite $(x_k)_{k\in\mathbb{N}}$

une suite de points de $\mathcal{X}$ convergeant vers x^0, il existe un rang $k^0 \in \mathbb{N}$ tel que

$$\forall\, k \geq k^0, \qquad f(x_k) = +\infty \qquad \text{donc} \qquad \liminf_{k\to+\infty} f(x_k) = +\infty = f(x^0)$$

et f est s.c.i. en x^0. ■

1.2 Propriétés

Proposition 4

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty, -\infty\}$ une fonction. Alors on a l'équivalence suivante :

(i) f est s.c.i. ;

(ii) pour tout $t \in \mathbb{R}$, l'ensemble de niveau inférieur
$$\operatorname{niv}_{\leq t} f = \Big\{ x \in \mathcal{X} \mid f(x) \leq t \Big\}$$
est fermé.

Démonstration : On démontre séparément les deux sens de l'équivalence.

- **Sens direct.** Soit $t \in \mathbb{R}$ et $(x_k)_{k\in\mathbb{N}}$ une suite de $\operatorname{niv}_{\leq t} f$, c'est-à-dire que
$$\forall\, k \in \mathbb{N}, \qquad f(x_k) \leq t$$
On suppose que cette suite est convergente, de limite x^*. Par comparaison, $f(x^*) \leq t$, de sorte que $x^* \in \operatorname{niv}_{\leq t} f$. Ainsi, on prouve que cet ensemble est fermé.
- **Réciproque.** Soit $(x_k)_{k\in\mathbb{N}}$ une suite convergente de limite x^0. Par définition de la limite inférieure, on a
$$\liminf_{k\to+\infty} f(x_k) = \sup_{k\in\mathbb{N}} \left\{ \inf_{n\geq k} f(x_n) \right\}$$
Par l'absurde, supposons que cette quantité est strictement inférieure à $f(x^0)$. Il s'ensuit qu'il existe $\varepsilon > 0$ tel que, pour tout $k \in \mathbb{N}$,
$$\inf_{n\geq k} f(x_n) \leq f(x^0) - \varepsilon$$
Soit $k \in \mathbb{N}$. Par définition de la borne inférieure, il existe $n_k \in \mathbb{N}$ tel que
$$f(x_{n_k}) \leq f(x^0) - \frac{\varepsilon}{2}$$
L'ensemble de niveau inférieur $f(x^0) - \varepsilon/2$ étant fermé, on a
$$\lim_{k\to+\infty} x_{n_k} = x^0 \in \operatorname{niv}_{\leq f(x^0)-\varepsilon/2} f \qquad \text{soit} \qquad f(x^0) \leq f(x^0) - \frac{\varepsilon}{2}$$
ce qui est absurde. ■

On a en particulier le résultat (équivalent) suivant :

Corollaire 1

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty, -\infty\}$ une fonction. Alors on a l'équivalence suivante :

(i) f est s.c.i. ;

(ii) son épigraphe $\operatorname{epi} f = \Big\{ (x,t) \in \mathcal{X} \times \mathbb{R} \mid f(x) \leq t \Big\}$ est fermé.

> Cette dernière propriété justifie l'appellation *fonction convexe fermée* que l'on rencontre parfois pour désigner les fonctions convexes s.c.i., ainsi que la dénomination de fermeture pour $\underline{f}$.

Proposition 5

Soit $(f_k)_{k\in\mathcal{I}}$ une famille de fonctions $\mathcal{X}\to\mathbb{R}\cup\{+\infty,-\infty\}$. Soit $x^0\in\mathcal{X}$. On suppose que, pour tout $k\in\mathcal{I}$, la fonction f_k est s.c.i. en x^0. Alors l'enveloppe supérieure des f_k

$$f:\left\{\begin{array}{ccc}\mathcal{X} & \to & \mathbb{R}\cup\{+\infty,-\infty\}\\ x & \mapsto & \sup\limits_{k\in\mathcal{I}} f_k(x)\end{array}\right.$$

est s.c.i. en x^0.

Démonstration : Il suffit de remarquer que

$$\mathrm{niv}_{\leq t} f = X\setminus\Big\{x\in\mathcal{X}\mid f(x)>t\Big\} = X\setminus\Big\{x\in\mathcal{X}\mid \sup_{k\in\mathcal{I}} f_k(x)>t\Big\}$$

Or, on a

$$\Big\{x\in\mathcal{X}\mid \sup_{k\in\mathcal{I}} f_k(x)>t\Big\} = \bigcup_{k\in\mathcal{I}}\Big\{x\in\mathcal{X}\mid f_k(x)>t\Big\} = \bigcup_{k\in\mathcal{I}}\Big(X\setminus \mathrm{niv}_{\leq t} f_k\Big)$$

où chaque ensemble de niveau inférieur est fermé d'après la proposition 4, on en déduit, puisque l'union d'ensembles ouverts est ouvert, que l'ensemble de niveau inférieur de f est fermé. ■

2 Propriété de Kurdyka-Łojasiewicz

2.1 Définition

Définition 2 (Propriété de Kurdyka-Łoajasiewicz)

Soit $f:\mathcal{X}\to\mathbb{R}\cup\{+\infty,-\infty\}$ 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 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 $f(x^*)<f(x)<f(x^*)+\eta$, on a

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

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

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

la valeur infinie étant atteinte si et seulement si le sous-différentiel de f en x est vide. *A contrario*, cette distance est nulle si et seulement si x est un point critique. Ainsi, une fonction satisfaisant la propriété KŁ en un point x^* est nécessairement telle qu'il existe $\eta>0$ vérifiant

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

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.

Proposition 6

Soit $f:\mathcal{X}\to\mathbb{R}\cup\{+\infty,-\infty\}$ une fonction s.c.i. 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 $x\in\operatorname{dom} f$,

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

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

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

  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

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

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

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

  ce qui prouve le résultat désiré. ■

Ainsi, on voit que les seuls points où la propriété KŁ est intéressante sont les points critiques, c'est-à-dire les points $x^*\in\operatorname{dom} f$ tels que $0\in\partial f(x^*)$.

Pour comprendre l'intérêt de la propriété KŁ, on va considérer le cas particulier où f est une fonction continûment différentiable sur $\mathcal{V}$ et x^* un point critique. Le quatrième point de la définition devient alors : pour tout $x \in \mathcal{V}$, on a

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

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 (autour de x^*) près, la fonction f n'est pas *plate* autour 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").

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 $x \neq 0$,

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

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.

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

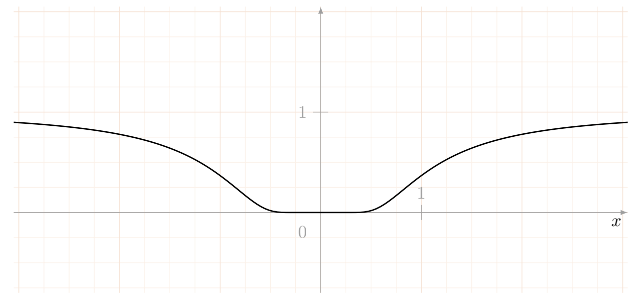

Figure 2 – Exemple de fonction non KŁ.

2.2 Fonctions semi-algébriques

Définition 3 (Ensemble semi-algébrique)

Soit $\mathcal{A} \subset \mathcal{X}$ un ensemble. On dit que $\mathcal{A}$ est *semi-algébrique* s'il existe deux familles de polynômes $\{P_{i,j}\}_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ et $\{Q_{i,j}\}_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ telles que

$$\mathcal{A} = \bigcup_{i=1}^{p} \Big\{ x \in \mathcal{X} \mid \forall\, j \in [\![\, 1\,;\, n \,]\!]\,, P_{i,j}(x) = 0 \text{ et } Q_{i,j}(x) > 0 \Big\}$$

Remarque : Dans la définition précédente, les polynômes peuvent être nuls ou constants.

Proposition 7 (Complémentaire)

Soit $\mathcal{A} \subset \mathcal{X}$ un ensemble semi-algébrique. Alors $\mathcal{X} \setminus \mathcal{A}$ est semi-algébrique.

Démonstration : Laissé au lecteur.

Proposition 8 (Intersection, union et produit cartésien)

Soit $\{\mathcal{A}_\ell\}_{1 \leq \ell \leq L}$ une famille finie d'ensembles de $\mathcal{X}$. On suppose que chacun des $\mathcal{A}_\ell$ est semi-algébrique. Alors

$$\bigcup_{\ell=1}^{L} \mathcal{A}_\ell, \qquad \bigcap_{\ell=1}^{L} \mathcal{A}_\ell \qquad \text{et} \qquad \mathcal{A}_1 \times \cdots \times \mathcal{A}_L$$

sont semi-algébriques.

Démonstration : Laissé au lecteur.

Exemple

Quelques exemples d'ensembles semi-algébriques. Les ensembles suivants sont semi-algébriques :

- tout singleton $\{a\}$ (il suffit de choisir $P_{1,1}(x) = x - a$ et $Q_{1,1}(x) = 1$) ;
- tout sous-espace vectoriel U de $\mathcal{X}$ (avec $P_{1,1}$ linéaire tel que $\ker P_{1,1} = U$) ;
- tout demi-espace $\mathcal{H} = \{x \in \mathcal{X} \mid \langle x, a\rangle \leq b\}$ pour $a \in \mathcal{X}$ et $b \in \mathbb{R}$;
- tout intervalle (fermé, ouvert ou semi-ouvert) dans $\mathbb{R}$ (par intersection).

Définition 4 (Fonction semi-algébrique)

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty, -\infty\}$ une fonction. On dit que f est *semi-algébrique* si $\operatorname{dom} f$ est semi-algébrique et que, pour tout ensemble semi-algébrique $\mathcal{A} \subset \mathcal{Y} \times \mathbb{R}$, l'ensemble

$$\widetilde{\mathcal{A}} = \Big\{(y, x) \in \mathcal{Y} \times \mathcal{X} \mid x \in \operatorname{dom} f \text{ et } (y, f(x)) \in \mathcal{A}\Big\}$$

est semi-algébrique.

Exemple

Fonctions polynomiales. Soit $P : \mathcal{X} \to \mathbb{R}$ une fonction polynomiale. Soit $p \in \mathbb{N}$ et $\mathcal{A} \subset \mathbb{R}^{p+1}$ un ensemble semi-algébrique. Pour tout $x \in \mathcal{X}$ et pour tout $y \in \mathcal{R}^p$, le point $(y, P(x))$ appartient à $\mathcal{A}$ si et seulement si

$$\exists\, i \in [\![1\,;\, m]\!], \forall\, j \in [\![1\,;\, n]\!], \quad P_{i,j}(y, P(x)) = 0 \quad \text{et} \quad Q_{i,j}(y, P(x)) > 0$$

où les $P_{i,j}$ et $Q_{i,j}$ sont des polynômes pour tout $(i, j) \in [\![1\,;\, m]\!] \times [\![1\,;\, n]\!]$. Or, puisque les fonctions $\widetilde{P}_{i,j} : (y, x) \mapsto P_{i,j}(y, P(x))$ et $\widetilde{Q}_{i,j} : (y, x) \mapsto Q_{i,j}(y, P(x))$ sont également des polynômes, on en déduit que $(y, x) \in \widetilde{\mathcal{A}}$ si et seulement si

$$\exists\, i \in [\![1\,;\, m]\!], \forall\, j \in [\![1\,;\, n]\!], \quad \widetilde{P}_{i,j}(y, x) = 0 \quad \text{et} \quad \widetilde{Q}_{i,j}(y, x) > 0$$

ce qui prouve que $\widetilde{A}$ est semi-algébrique.

On prouve de manière similaire que les indicatrices d'ensembles semi-algébriques sont semi-algébriques.

Il est possible de caractériser les fonctions semi-algébriques à l'aide de leur graphe.

Proposition 9

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty, -\infty\}$ une fonction. Alors on a l'équivalence entre les deux affirmations suivantes :

(i) f est semi-algébrique ;

(ii) son graphe $\quad \operatorname{gr} f = \Big\{(x, f(x)) \in \mathcal{X} \times \mathbb{R} \mid x \in \mathcal{X}\Big\}$ est semi-algébrique.

Démonstration : Admis.

Proposition 10 (Somme, différence et produit)

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ et $g : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ deux fonctions semi-algébriques telles que $\operatorname{dom} f = \operatorname{dom} g$. Alors

$$f + g, \qquad f - g \qquad \text{et} \qquad f \cdot g$$

sont semi-algébriques.

Autrement dit, l'ensemble des fonctions semi-algébriques définies sur un ensemble $\mathcal{A}$ semi-algébrique forme un anneau.

Démonstration : Laissé au lecteur.

Exercice

Fonctions réelles définies par morceaux. Soit $a_1 \leq \cdots \leq a_n$ des réels et $\{f_i\}_{1 \leq i \leq n+1}$ une famille de fonctions semi-algébriques. Considérons la fonction $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ définie par

$$\forall\, x \in \mathbb{R}, \qquad f(x) = \begin{cases} f_1(x) & \text{si } x \leq a_1 \\ f_i(x) & \text{si } a_i < x \leq a_{i+1} \\ f_{n+1}(x) & \text{si } a_{n+1} < x \end{cases}$$

Montrer que f est semi-algébrique.

Exercice

Valeur absolue et racine carrée.

- Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ une fonction semi-algébrique. Montrer que $|f|$ est semi-algébrique.
- Soit $f : \mathcal{X} \to [\,0\,; +\infty\,]$ une fonction semi-algébrique. Montrer que $\sqrt{f}$ est semi-algébrique.

Exemple

Distance à un ensemble semi-algébrique. Soit $\mathcal{A} \subset \mathcal{X}$ un ensemble semi-algébrique non vide. Alors la fonction

$$f : \left\{ \begin{array}{ccl} \mathcal{X} & \to & \mathbb{R} \\ x & \mapsto & \operatorname{dist}(x, \mathcal{A}) = \inf\limits_{a \in \mathcal{A}} \|x - a\| \end{array} \right.$$

est semi-algébrique.

Définition 5 (Application semi-algébrique)

Soit $f : \mathcal{X} \to (\mathbb{R} \cup \{+\infty, -\infty\})^n$ une application. On dit que f est *semi-algébrique* si chacune de ses composantes f_i pour $i \in [\![1\,;\, n]\!]$ est semi-algébrique.

Proposition 11 (Composition)

Soit $f : \mathcal{Y} \to (\mathbb{R} \cup \{+\infty, -\infty\})^n$ et $g : \mathcal{X} \to \mathcal{Y}$ deux applications semi-algébriques. Alors $f \circ g$ est semi-algébrique.

Démonstration : Admis.

2.3 Quelques classes de fonctions KŁ

On a mentionné plus haut que les fonctions convexes **ne sont pas nécessairement KŁ**. Cependant, le résultat suivant montre que les fonctions fortement convexes le sont :

Proposition 12 (Condition de Polyak–Łojasiewicz)

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty, -\infty\}$ une fonction s.c.i. On suppose que f est fortement convexe, de module α. 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, 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. ■

On va maintenant établir que les fonctions analytiques réelles sont KŁ.

Proposition 13 (Lojasiewicz)

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

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

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

Démonstration : Admis.

De manière générale, lorsque la fonction désingularisante κ apparaissant dans la définition de la propriété KŁ est de la forme

$$\kappa : t \mapsto c\,t^{1-\theta} \qquad \text{avec} \qquad c > 0 \qquad \text{et} \qquad \theta \in [\,0\,;1\,[$$

alors le réel θ est appelé *exposant de* Łojasiewicz. Cette quantité a une influence sur la vitesse de convergence de certains algorithmes impliquant de telles fonctions.

Une autre classe importante de fonctions KŁ sont les fonctions semi-algébriques.

Théorème 1 (Bolte, Daniilidis & Lewis)

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty, -\infty\}$ une fonction s.c.i. et semi-algébrique. Alors f est KŁ.

Démonstration : Admis.

Néanmoins, en pratique, ces deux résultats ne sont en général pas suffisants car les fonctions rencontrées dans les applications sont rarement analytiques ou semi-algébriques. Il s'agit plus couramment de somme / composition de telles fonctions.

Proposition 14 (Somme)

Soit $f : \mathcal{D} \to \mathbb{R}$ une fonction analytique réelle et $g : \mathbb{R} \to \mathbb{R} \cup \{+\infty, -\infty\}$ une fonction semi-algébrique continue sur son domaine. On suppose que f et g sont minorées et que $\operatorname{dom} g$ est fermé. Alors $f + g$ est KŁ.

Démonstration : Admis.

Proposition 15 (Composition)

Soit $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty, -\infty\}$ et $g : \mathcal{Y} \to \mathcal{X}$ deux fonctions analytiques et / ou semi-algébriques. On suppose que l'une des deux hypothèses suivantes est satisfaite :

(i) pour tout $\mathcal{E} \subset \mathcal{Y}$ borné, $g(\mathcal{E})$ est borné ;
(ii) pour tout $\mathcal{E} \subset \mathcal{Y}$ borné, $f^{-1}(\mathcal{E})$ est borné.

Alors $f \circ g$ est KŁ.

Démonstration : Admis.

> 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.

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

Théorème 2 (Attouch, Bolte & Svaiter)

Soit $J : \mathcal{X} \to \mathbb{R} \cup \{+\infty, -\infty\}$ 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 $k \in \mathbb{N}$,

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

(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

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

(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

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

- **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 $J(x^*) < J(x) < J(x^*) + \eta$, on a

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

- **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

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

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

- **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 $x_{k+1} \neq x_k$. 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

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

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

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

- **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 $2\sqrt{\alpha\,\beta} \le \alpha + \beta$ valable pour tous $\alpha, \beta \ge 0$,

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

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

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

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

Il s'ensuit en particulier que

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

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. ■

Pour aller plus loin

Semi-continuité inférieure. Cette propriété sera beaucoup utilisée dans les modules qui viennent, notamment dans certaines preuves de convergence, car elle autorise des passages à la limite intéressants dans le cadre de l'optimisation. Elle peut évidemment être avantageusement remplacée par une hypothèse de continuité, ou du moins, la continuité sur le domaine, lorsque celui-ci est fermé.

Propriété KŁ. Cette propriété permet de remplacer les hypothèses de régularité et/ou de convexité dans certains algorithmes d'optimisation, comme ceux étudiés dans le module B5.