Nonsmooth analysis: Weak stationarity

Alexander Y. Kruger
Centre for Informatics and Applied Optimization
School of Information Technology and Mathematical Sciences
University of Ballarat
POB 663, Ballarat, Vic 3350, Australia
E-mail address: a.kruger@ballarat.edu.au
URL: http://uob-community.ballarat.edu.au/~akruger

Introduction

The article considers different stationarity and regularity concepts for extended real-valued functions on metric spaces.

All the properties can be characterized in terms of certain local constants. A function is said to be stationary at a point (in some sense) if the corresponding constant is zero (a critical point). Otherwise the function is said to be regular at this point (in the same sense) and the constant provides a quantitative estimate of regularity.

Traditionally the stationary behavior of a function at a point (*stationary point*) means that it is arbitrarily close to a constant near this point, that is the rate of change of the function is infinitely small compared to the increment of the variable. Of course, this is equivalent to the derivative being equal to zero (dual characterization of stationarity). The classical examples of the stationary behavior of a function, found in the textbooks, are given by the functions $y = x^2$, $y = -x^2$, and $y = x^3$. These three examples characterize the three possible types of the stationary behavior in the differentiable case.

Stationarity arises naturally in optimization theory: a point of minimum or maximum is necessarily a stationary point. At the same time it is easy to see that weaker stationarity concepts are applicable to optimization problems. For instance, when dealing with minimizing a real-valued function, only its decrement must be infinitely small at a stationary point, and the function itself does not need to be arbitrarily close to a constant or even differentiable at the point. Of course, if it is differentiable, one has the same classical stationarity concept. In the nondifferentiable case more types of the stationary behavior are possible. The examples are: $y = |x|$ and $y = \max(x, -x^2)$ (both functions are considered near the point $x = 0$.) One can speak about *inf-stationarity* (the term suggested by Vladimir F. Demianov.) From the point of view of maximization a concept of *sup-stationarity* can be considered in a similar way. Thus in the nondifferentiable case the stationarity splits into two "semi-stationarity" concepts.

Another pair of stationarity concepts can be of interest in optimization theory: the point itself may not be stationary (inf or sup), but in any its neighborhood there exists another point, in which the behavior of the function is arbitrarily close to stationary (a "fuzzy" condition). We will speak about *weak stationarity* (inf or sup). The exact definitions will be given below.

An examples of this type of stationary behavior is given by the function $y = x\sin(1/x)$ if $x \neq 0$, and $y = 0$ if $x = 0$. One can easily see that this function is differentiable on $\mathbb{R}\backslash\{0\}$ and there exists a sequence $\{x_k\}$ such that $x_k \to 0$ and x_k is a point of local minimum, $k = 1, 2, \ldots$. Another example: $y = x + x^2\sin(1/x)$ if $x \neq 0$, and $y = 0$ if $x = 0$. This function is everywhere differentiable, $y'(0) = 1$ and there exists a sequence $\{x_k\}$ such that $x_k \to 0$ and $y'(x_k) = 0$, $k = 1, 2, \ldots$.

Stationarity concepts can also be defined in terms of dual space elements (subdifferentials). The relations between primal and dual definitions provide dual characterizations of (primal space concepts of) stationarity.

The article contains no proofs. More detailed description of the stationarity and regularity concepts for real-valued functions can be found in [11].

Mostly standard notations are used throughout the article. X denotes a metric space with the distance d. $B_\rho(x)$ stands for a closed ball with center x and radius ρ.

Definitions

Inf-θ-stationarity and Inf-θ-regularity

Let f be a function on a metric space X with values in the extended real line $\mathbb{R}_\infty = \mathbb{R} \cup \{+\infty\}$. It is assumed to be finite at some point $x^\circ \in X$.

For $\rho > 0$ define the constant

$$\theta_\rho[f](x^\circ) = \inf_{x \in B_\rho(x^\circ)} f(x) - f(x^\circ). \tag{1}$$

Note that $\theta_\rho[f](x^\circ) \leq 0$ and the equality $\theta_\rho[f](x^\circ) = 0$ for some $\rho > 0$ (for all $\rho > 0$) means that x° is a point of local (global) minimum of f.

Of course, the infimum in (1) can be limited to the set $\{x \in B_\rho(x^\circ) : f(x) \leq f(x^\circ)\}$, or even to the set $\{x \in B_\rho(x^\circ) : \ f(x) < f(x^\circ)\}$ under the additional agreement that the infimum over the empty subset of $\mathbb{R}_-$ is 0.

The function $\rho \to \theta_\rho[f](x^\circ)$ is nonincreasing on $\mathbb{R}_+$ and $\lim\limits_{\rho\to+0} \theta_\rho[f](x^\circ) \leq 0$. The equality $\lim\limits_{\rho\to+0} \theta_\rho[f](x^\circ) = 0$ means that f is lower semicontinuous at x°.

In the latter case it can be important to know how quickly $\theta_\rho[f](x^\circ)$ approaches 0 compared to ρ.

Define two more "derivative-like" constants based on (1):

$$\theta[f](x^\circ) = \limsup_{\rho\to+0} \frac{\theta_\rho[f](x^\circ)}{\rho}, \tag{2}$$

$$\hat{\theta}[f](x^\circ) = \limsup_{x\xrightarrow{f}x^\circ,\ \rho\to+0} \frac{\theta_\rho[f](x)}{\rho}, \tag{3}$$

where $x \xrightarrow{f} x^\circ$ means that $x \to x^\circ$ with $f(x) \to f(x^\circ)$. Due to the variations of x in (3), $\hat{\theta}[f](x^\circ)$ gains some properties of the strict derivative.

The constants (2) and (3) are nonpositive too and "zero cases" correspond to certain kinds of stationary behavior of f near x°. If a constant is strictly negative, this can be considered as a kind of regularity.

Definition 1. f is

(i) *inf-θ-stationary* at x° if $\theta[f](x^\circ) = 0$;

(ii) *weakly inf-θ-stationary* at x° if $\hat{\theta}[f](x^\circ) = 0$;

(iii) *inf-θ-regular* at x° if $\theta[f](x^\circ) < 0$;

(iv) *strongly inf-θ-regular* at x° if $\hat{\theta}[f](x^\circ) < 0$.

The purpose of the "inf" prefix in this definition is to emphasize that minimization problems are addressed here. Unlike the classical case, stationarity-regularity properties of nondifferentiable functions "from below" and "from above" can be essentially different.

Inf-τ-stationarity and Inf-τ-regularity

Another way of defining stationarity-regularity is based on using slightly modified versions of (2), (3):

$$\tau[f](x^\circ) = \liminf_{x\to x^\circ} \frac{[f(x) - f(x^\circ)]_-}{d(x, x^\circ)}, \tag{4}$$

$$\hat{\tau}[f](x^\circ) = \limsup_{x\xrightarrow{f}x^\circ,\ \rho\to+0} \inf_{u\in B_\rho(x)\setminus\{x\}} \frac{[f(u) - f(x)]_-}{d(u, x)}. \tag{5}$$

The notation $[\alpha]_- = \min(\alpha, 0)$ is used here. Again, only the points $x \in B_\rho(x^\circ)$ with $f(x) < f(x^\circ)$ and $u \in B_\rho(x)$ with $f(u) < f(x)$ are of interest

in (4) and (5) respectively. The role of the notation is to handle the case when the set of such points is empty. Similarly to (2), (3) these constants are nonpositive.

Remark 1. $\tau[f](x^\circ)$ coincides up to a sign with the *strong slope* $|\nabla f|(x^\circ)$ of f at x° [1] (see also [5]).

Definition 2. f is

(i) *inf-τ-stationary* at x° if $\tau[f](x^\circ) = 0$;

(ii) *weakly inf-τ-stationary* at x° if $\hat{\tau}[f](x^\circ) = 0$;

(iii) *inf-τ-regular* at x° if $\tau[f](x^\circ) < 0$;

(iv) *strongly inf-τ-regular* at x° if $\hat{\tau}[f](x^\circ) < 0$.

The relations between the constants (2), (3) and (4), (5), as well as between the corresponding stationarity and regularity concepts will be discussed in the next section.

Sup-stationarity and Sup-regularity

Similarly to (1)–(5) corresponding "maximization" constants can be defined. To do this one has to replace "inf", "lim inf", "lim sup", and $[\cdot]_-$ by "sup", "lim sup", "lim inf", and $[\cdot]_+$ respectively in the corresponding definitions. The resulting constants are nonnegative. They are related to (1)–(5) by the following equalities:

$$\begin{gathered}\theta_\rho^+[f](x^\circ) = -\theta_\rho[-f](x^\circ),\\ \theta^+[f](x^\circ) = -\theta[-f](x^\circ), \quad \hat{\theta}^+[f](x^\circ) = -\hat{\theta}[-f](x^\circ),\\ \tau^+[f](x^\circ) = -\tau[-f](x^\circ), \quad \hat{\tau}^+[f](x^\circ) = -\hat{\tau}[-f](x^\circ)\end{gathered}$$

and lead to similar *sup-stationarity* and *sup-regularity* concepts.

Of course, for a function f the set of sup-stationary (sup-regular) points is different in general from that of inf-stationary (inf-regular) points.

The "combined" concepts can also be of interest. It is natural to say that a function is stationary (in some sense) at a point if it is either inf-stationary or sup-stationary at this point. On the contrary, the regularity property for a function is satisfied when this function is both inf-regular and sup-regular at the point.

Definition 3. f is

(i) *θ-stationary* at x° if $\max(\theta[f](x^\circ), \theta[-f](x^\circ)) = 0$;

(ii) *weakly θ-stationary* at x° if $\max(\hat{\theta}[f](x^\circ), \hat{\theta}[-f](x^\circ)) = 0$;

(iii) *θ-regular* at x° if $\max(\theta[f](x^\circ), \theta[-f](x^\circ)) < 0$;

(iv) *strongly θ-regular* at x° if $\max(\hat{\theta}[f](x^\circ), \hat{\theta}[-f](x^\circ)) < 0$;

(v) *τ-stationary* at x° if $\max(\tau[f](x^\circ), \tau[-f](x^\circ)) = 0$;

(vi) *weakly τ-stationary* at x° if $\max(\hat{\tau}[f](x^\circ), \hat{\tau}[-f](x^\circ)) = 0$;

(vii) *τ-regular* at x° if $\max(\tau[f](x^\circ), \tau[-f](x^\circ)) < 0$;

(viii) *strongly τ-regular* at x° if $\max(\hat{\tau}[f](x^\circ), \hat{\tau}[-f](x^\circ)) < 0$.

Strong inf-regularity can be interpreted in the following way: all points in a neighborhood of a given point have "descent sequences", and the rate of descent is uniform. In contrast to that, strong regularity is equivalent to the existence of both descent and ascent sequences with the uniformity property.

Dual Stationarity and Regularity

All definitions in the preceding subsections are primal space definitions. As in the classical analysis, dual characterizations of stationarity and regularity concepts are important. In the case of a normed linear space such characterizations can be formulated in terms of *Fréchet subdifferentials*.

Let X be a normed linear space. Its (topological) dual is denoted X^*. $\langle \cdot, \cdot \rangle$ is the bilinear form defining the duality pairing. Recall that the Fréchet subdifferential of f at x° is defined as

$$\partial f(x^\circ) = \left\{ x^* \in X^* : \liminf_{x \to x^\circ} \frac{f(x) - f(x^\circ) - \langle x^*, x - x^\circ \rangle}{\|x - x^\circ\|} \geq 0 \right\}. \quad (6)$$

Definition 4. f is

(i) *inf-d-stationary* at x° if $0 \in \partial f(x^\circ)$;

(ii) *inf-d-regular* at x° if $0 \notin \partial f(x^\circ)$.

It follows immediately from the definitions that in the normed space setting inf-τ-stationarity (inf-τ-regularity) is equivalent to inf-d-stationarity (inf-d-regularity).

A bit more complicated constructions are needed for characterization of weak stationarity and strong regularity.

Let us assume for simplicity that f is lower semicontinuous near x°.

In the general nonconvex setting the subdifferential mapping $\partial f(\cdot)$ fails to posses good (semi-)continuity properties. In fact, the set $\partial f(x)$ can be empty rather often. Based on (6) one can define a more robust derivative-like object:

$$\hat{\partial}_\delta f(x^\circ) = \bigcup_{\substack{x\in B_\delta(x^\circ)\\ |f(x)-f(x^\circ)|\le\delta}} \partial f(x). \tag{7}$$

It depends on a positive parameter δ and accumulates information on "differential" properties of f at nearby points, thus attaining some properties of the strict derivative. The set (7) is called the *strict δ-subdifferential* of f at x° (see [7, 8, 9]). In contrast to (6), the set (7) can be nonconvex. However, it possesses certain subdifferential calculus.

Using (7) one more constant can be defined for characterizing stationarity/regularity properties of f:

$$\eta[f](x^\circ) = \lim_{\delta\to 0} \inf\{\|x^*\| :\ x^* \in \hat{\partial}_\delta f(x^\circ)\}. \tag{8}$$

Unlike the constants considered in the preceding subsections, this constant is nonnegative.

Definition 5. f is

(i) *inf-η-stationary* at x° if $\eta[f](x^\circ) = 0$;

(ii) *inf-η-regular* at x° if $\eta[f](x^\circ) > 0$.

Note that the inf-η-stationary condition $\eta[f](x^\circ) = 0$ does not imply the inclusion $0 \in \hat{\partial}_\delta f(x^\circ)$.

Example 1. Take $f(x) = x$, if $x < 0$, and $f(x) = x^2$ otherwise. One has $\partial f(0) = \emptyset$, $0 \notin \hat{\partial}_\delta f(0)$ for any $\delta > 0$ while $\eta[f](0) = 0$.

Fortunately (8) happens to be closely related to (3) and (5).

Sup-d-stationarity and sup-η-stationarity as well as the corresponding regularity concepts can be defined in a similar way.

Formulation

Relations between the "Elementary" Constants

Proposition 1. *The following assertions hold true:*

(i) $\tau[f](x^\circ) \le \theta[f](x^\circ)$;

(ii) *if* $\theta_\rho[f](x^\circ) = 0$ *for some* $\rho > 0$ *then* $\tau[f](x^\circ) = \theta[f](x^\circ) = 0$.

Proposition 1 (i) implies the relations between the corresponding stationarity and regularity concepts:

- inf-τ-stationarity $\Rightarrow$ inf-θ-stationarity;
- inf-θ-regularity $\Rightarrow$ inf-τ-regularity.

Proposition 1 (ii) means that at a point of local minimum a function is both inf-τ-stationary and inf-θ-stationary.

Inequality (i) in Proposition 1 can be strict even for functions from $\mathbb{R}$ to $\mathbb{R}$.

Example 2. Take $f(x) = -|x|$, if $|x| = 1/2^n$, $n = 1, 2, \ldots$, and $f(x) = 0$ otherwise. Obviously $\tau[f](0) = -1$. At the same time, for any $\rho \in \Xi_n = \{\rho : 1/2^n \le \rho < 1/2^{n-1}\}$ one has $\theta_\rho[f](0) = -1/2^n$ and

$$\sup_{\rho \in \Xi_n} \frac{\theta_\rho[f](0)}{\rho} = \frac{-1/2^n}{1/2^{n-1}} = -\frac{1}{2}.$$

Thus, $\theta[f](0) = -1/2$.

It is possible to modify the above example to make $\theta[f](0)$ equal zero.

Example 3. Take $f(x) = -|x|$, if $|x| = 1/n^n$, $n = 1, 2, \ldots$, and $f(x) = 0$ otherwise. One still has $\tau[f](0) = -1$ while $\theta[f](0) = 0$.

Thus, in the above example f is inf-τ-regular at 0 while being inf-θ-stationary at this point.

It is possible to modify the example further to make f continuous and even differentiable near 0 (but not strictly differentiable!) while keeping the inequality (i) in Proposition 1 strict.

Relations between the "Strict" Constants

The relations between the elementary constants and their "strict" counterparts, as well as between the two "strict" constants are given by the next theorem.

Theorem 1. *The following assertions hold true:*

(i) $\hat{\theta}[f](x^\circ) \ge \limsup_{x \xrightarrow{f} x^\circ} \theta[f](x)$,

(ii) $\hat{\tau}[f](x^\circ) = \limsup\limits_{x \xrightarrow{f} x^\circ} \tau[f](x)$;

(iii) $\hat{\tau}[f](x^\circ) \le \hat{\theta}[f](x^\circ)$;

(iv) *If X is complete and f is lower semicontinuous near x°, then* $\hat{\tau}[f](x^\circ) = \hat{\theta}[f](x^\circ)$.

Parts (i) and (ii) of Theorem 1 imply the inequalities

$$\theta[f](x^\circ) \le \hat{\theta}[f](x^\circ), \quad \tau[f](x^\circ) \le \hat{\tau}[f](x^\circ),$$

and both of them can be strict.

Example 4. Take the function f from Example 1. Evidently, f attains a local minimum at $x_n = 1/2^n$ for any $n = 1, 2, \ldots$, and consequently, $\theta_\rho[f](x_n) = 0$ for some $\rho > 0$. It follows from Proposition 1 (ii) that $\tau[f](x_n) = \theta[f](x_n) = 0$. Consequently, $\hat{\tau}[f](0) = \hat{\theta}[f](0) = 0$. Recall that $\tau[f](0) = -1$ and $\theta[f](0) = -1/2$.

Inequalities (i) and (iii) in Theorem 1 can be strict too.

Example 5. Define the function $f : \mathbb{R} \to \mathbb{R}$ in the following way: $f(x) = x$ if $x \le 0$, $f(x) = x - 1/n$ if $1/n < x \le 1/(n-1)$, $n = 2, 3, \ldots$, $f(x) = x - 1/2$ if $x > 1/2$. It is easy to see that $\theta[f](x) = \tau[f](x) = -1$ for any $x \in \mathbb{R}$. Then $\hat{\tau}[f](0) = -1$. On the other hand, take $x_n = 1/n + 1/n^2$, $\rho_n = 1/n$, $n = 1, 2, \ldots$. Then $f(x_n) = 1/n^2$ and consequently, $\theta_{\rho_n}[f](x_n) \ge -1/n^2$. It follows immediately that $\hat{\theta}[f](0) = 0$.

Due to part (iv) of Theorem 1 in the case of a lower semicontinuous function on a complete metric space two weak stationarity concepts as well as two strong regularity concepts coincide and the prefixes θ and τ can be omitted.

Corollary 1.1. *The following assertions hold true:*

(i) *inf-θ-stationarity $\Rightarrow$ weak inf-θ-stationarity;*
strong inf-θ-regularity $\Rightarrow$ inf-θ-regularity;

(ii) *inf-τ-stationarity $\Rightarrow$ weak inf-τ-stationarity;*
strong inf-τ-regularity $\Rightarrow$ inf-τ-regularity;

(iii) *weak inf-τ-stationarity $\Rightarrow$ weak inf-θ-stationarity;*
strong inf-θ-regularity $\Rightarrow$ strong inf-τ-regularity;

(iv) If X is complete and f is lower semicontinuous near x°, then

weak inf-τ-stationarity $\Leftrightarrow$ weak inf-θ-stationarity;

strong inf-θ-regularity $\Leftrightarrow$ strong inf-τ-regularity.

The next "fuzzy" characterization of weak inf-τ-stationarity can be convenient for applications. It follows directly from the definition (5).

Proposition 2. *f is weakly inf-τ-stationary at x° if and only if for any $\varepsilon > 0$ there exists an $x \in B_\varepsilon(x^\circ)$ such that $|f(x) - f(x^\circ)| \leq \varepsilon$ and*

$$f(u) + \varepsilon d(u, x) \geq f(x) \text{ for all } u \text{ near } x. \tag{9}$$

Remark 2. A point x satisfying (9) is referred to in [12] (see also [6]) as a *local Ekeland point* of f (with factor ε). If all the conditions in Proposition 2 are satisfied, then x° is said to be a *stationary point* of f *with respect to minimization* [12]. Thus, stationarity with respect to minimization is equivalent to weak inf-τ-stationarity and, in the case of a lower semicontinuous function on a complete metric space, also to weak inf-θ-stationarity.

Relations between the Primal and Dual Constants

From now on X is assumed to be a normed linear space. The next assertion is straightforward and has already been mentioned in the previous section.

Proposition 3. *(i) inf-τ-stationarity $\Leftrightarrow$ weak inf-d-stationarity;*

(ii) inf-τ-regularity $\Leftrightarrow$ inf-d-regularity.

Remark 3. Due to Propositions 1 and 3 the inclusion $0 \in \partial f(x^\circ)$ is sufficient for inf-θ-stationarity of f at x°. The opposite implication is not true in general (see Examples 2 and 3).

In what follows f is assumed to be lower semicontinuous near x°.

Theorem 2. *(i)* $\hat{\theta}[f](x^\circ) + \eta[f](x^\circ) \geq 0$.

(ii) If X is Asplund then

$$\frac{\hat{\theta}[f](x^\circ)}{[1 + \hat{\theta}[f](x^\circ)]_+} + \eta[f](x^\circ) \leq 0.$$

This theorem follows from [10], Theorem 2. The first part of the theorem is elementary. The proof of the second part is based on the application of the two fundamental results of variational analysis: the *Ekeland variational principle* [2] and the *fuzzy sum rule* due to Fabian [3].

Thus, in an Asplund space the constants $\hat{\theta}[f](x^\circ)$ and $\eta[f](x^\circ)$ can be zero or nonzero only simultaneously. Recall that a Banach space is called *Asplund* (see [4, 13, 14]) if any continuous convex function on it is Fréchet differentiable on a dense G_δ subset. Note that in a Banach space $\hat{\theta}[f](x^\circ) = \hat{\tau}[f](x^\circ)$ due to Theorem 1.

Corollary 2.1. *(i) inf-η-stationarity $\Rightarrow$ weak inf-θ-stationarity;*

(ii) strong inf-θ-regularity $\Rightarrow$ inf-η-regularity;

(iii) If X is Asplund then

weak inf-θ-stationarity $\Leftrightarrow$ weak inf-τ-stationarity $\Leftrightarrow$ inf-η-stationarity;

strong inf-θ-regularity $\Leftrightarrow$ strong inf-τ-regularity $\Leftrightarrow$ inf-η-regularity.

Differentiable Functions

The constants and corresponding stationarity/regularity concepts defined above take quite a traditional form when the function is assumed differentiable or convex. Fortunately, the number of different constants and concepts reduces significantly.

Theorem 3. *If f is Fréchet differentiable at x° with the derivative $\nabla f(x^\circ)$ then*

$$\theta[f](x^\circ) = \tau[f](x^\circ) = -\theta^+[f](x^\circ) = -\tau^+[f](x^\circ) = -\left\|\nabla f(x^\circ)\right\|.$$

If, additionally, the derivative is strict then

$$\hat{\theta}[f](x^\circ) = \hat{\tau}[f](x^\circ) = -\hat{\theta}^+[f](x^\circ) = -\hat{\tau}^+[f](x^\circ) = -\left\|\nabla f(x^\circ)\right\|.$$

Recall that f is called *strictly differentiable* [13, 15] at x° (with the derivative $\nabla f(x^\circ)$) if

$$\lim_{x \to x^\circ,\ u \to x^\circ} \frac{f(u) - f(x) - \langle \nabla f(x^\circ), u - x \rangle}{\|u - x\|} = 0.$$

This condition is stronger than the traditional Fréchet differentiability. Thus, condition $\nabla f(x^\circ) \neq 0$ does not guarantee strong regularity in the sense of Definition 1 (or Definition 2) unless f is strictly differentiable at x°.

Example 6. Take $f(x) = x + x^2 \sin(1/x)$, if $x \neq 0$, and $f(0) = 0$. This function is everywhere Fréchet differentiable and $\nabla f(0) = 1$. Thus, f is regular at zero. At the same time $\hat{\tau}[f](0) = \hat{\tau}^+[f](0) = 0$: there exists a sequence $x_k \to 0$ such that $\nabla f(x_k) \to 0$, and the assertion follows from Theorem 1, part (ii). Consequently, f is both weakly inf-stationary and weakly sup-stationary at zero.

Corollary 3.1. *If f is Fréchet differentiable at x° with the derivative $\nabla f(x^\circ)$ then the following conditions are equivalent.*

(i) f is inf-θ-stationary at x°;

(ii) f is inf-τ-stationary at x°;

(iii) f is θ-stationary at x°;

(iv) f is τ-stationary at x°;

(v) $\nabla f(x^\circ) = 0$.

If, additionally, the derivative is strict then the above conditions are also equivalent to the following ones:

(vi) f is weakly inf-θ-stationary at x°;

(vii) f is weakly inf-τ-stationary at x°;

(viii) f is weakly θ-stationary at x°;

(ix) f is weakly τ-stationary at x°.

Remark 4. Stationary and weak stationary in the above corollary can be replaced with regularity and strong regularity respectively if one replaces the equality in (v) with the inequality $\nabla f(x^\circ) \neq 0$.

Convex Functions

In the convex case, as one could expect, all versions of inf-stationarity coincide and appear to be equivalent to just (local and global) minimality.

Theorem 4. *Let f be convex.*

(i) If $\theta_\rho[f](x^\circ) < 0$ for some $\rho > 0$ then $\theta_\rho[f](x^\circ) < 0$ for all $\rho > 0$.

(ii) The functions $\rho \to \theta_\rho[f](x^\circ)/\rho$ and $\rho \to \theta^+_\rho[f](x^\circ)/\rho$ are nondecreasing on $\mathbb{R}_+ \backslash \{0\}$.

(iii) *The following equalities hold true:*

$$\hat{\theta}[f](x^\circ) = \hat{\tau}[f](x^\circ) = \theta[f](x^\circ) = \tau[f](x^\circ) = \inf_{\rho>0} \frac{\theta_\rho[f](x^\circ)}{\rho} = \inf_{x\neq x^\circ} \frac{[f(x) - f(x^\circ)]_-}{\|x - x^\circ\|},$$

$$\theta^+[f](x^\circ) = \tau^+[f](x^\circ) = \inf_{\rho>0} \frac{\theta^+_\rho[f](x^\circ)}{\rho} = \inf_{\rho>0} \sup_{\|x-x^\circ\|=\rho} \frac{[f(x) - f(x^\circ)]_+}{\rho}.$$

(iv) $\tau[f](x^\circ) + \tau^+[f](x^\circ) \geq 0$.

(v) $\hat{\tau}[f](x^\circ) + \hat{\tau}^+[f](x^\circ) \geq 0$.

(vi) *If* $\tau[f](x^\circ) + \tau^+[f](x^\circ) = 0$ *and* $\{x_k\} \subset X$ *is a sequence defining* $\tau[f](x^\circ)$, *that is* $x_k \to 0$ *and*

$$\tau[f](x^\circ) = \lim_{k\to\infty} \frac{f(x^\circ + x_k) - f(x^\circ)}{\|x_k\|}$$

then $\{-x_k\}$ *is a sequence defining* $\tau^+[f](x^\circ)$*:*

$$\tau^+[f](x^\circ) = \lim_{k\to\infty} \frac{f(x^\circ - x_k) - f(x^\circ)}{\|x_k\|}.$$

Corollary 4.1. *If* f *is convex then the following conditions are equivalent:*

(i) f *attains a global minimum at* x°*;*

(ii) f *attains a local minimum at* x°*;*

(iii) f *is inf-*θ*-stationary at* x°*;*

(iv) f *is inf-*τ*-stationary at* x°*;*

(v) f *is* θ*-stationary at* x°*;*

(vi) f *is* τ*-stationary at* x°*;*

(vii) f *is weakly inf-*θ*-stationary at* x°*;*

(viii) f *is weakly inf-*τ*-stationary at* x°*;*

(ix) f *is weakly stationary at* x°*.*

Remark 5. The conditions $\tau[f](x^\circ) = \tau^+[f](x^\circ) = 0$ imply Fréchet differentiability of f at x° (with the derivative equal to zero). The weaker condition $\tau[f](x^\circ)+\tau^+[f](x^\circ) = 0$ in Theorem 4, (vi) implies linearity of the directional derivative of f along the direction of steepest descent (if the latter exists) with the opposite direction being automatically the direction of steepest ascent. This condition is not sufficient for differentiability of f at x° unless $X = \mathbb{R}$. Note also that the direction opposite to the direction of steepest ascent does not need to be a direction of steepest descent.

Example 7 ([1]). Take the function $f(x, y) = \max(x, y)$ on $\mathbb{R}^2$ and assume that $\mathbb{R}^2$ is equipped with the max type norm: $\|x, y\| = \max(|x|, |y|)$. f is obviously not differentiable at 0. At the same time $\tau[f](0) = -1$, $\tau^+[f](0) = 1$. The vector $(-1, -1)$ defines the (unique) direction of steepest descent. The opposite vector $(1, 1)$ defines the direction of steepest ascent and f is linear along the line defined by these vectors. Note that the direction of steepest ascent is not unique. For instance, the vector $(1, 0)$ also defines the direction of steepest ascent, while the opposite vector does not define the direction of steepest descent and f is not linear along this line.

Remark 6. Stationary and weak stationary in the assertions (iii) – (ix) of the above corollary can be replaced with regularity and strong regularity respectively if one replaces (i) and (ii) with the opposite assertions: x° is not a point of (local or global) minimum of f.

See also: **Nonsmooth analysis: Fréchet subdifferentials**

References

[1] De Giorgi, E., Marino, A., and Tosques, M. Problemi di evoluzione in spazi metrici e curve di massima pendenza. *Atti Accad. Nat. Lincei, Cl. Sci Fiz. Mat. Natur. 68* (1980), 180–187.

[2] Ekeland, I. On the variational principle. *J. Math. Anal. Appl. 47* (1974), 324–353.

[3] Fabian, M. Subdifferentiability and trustworthiness in the light of a new variational principle of Borwein and Preiss. *Acta Univ. Carolinae 30* (1989), 51–56.

[4] Fabian, M. *Gâteaux Differentiability of Convex Functions and Topology. Weak Asplund Spaces*. Canadian Mathematical Society Series of

[1] The example was suggested by Alexander Rubinov (personal communication).

Monographs and Advanced Texts. John Wiley & Sons Inc., New York, 1997.

[5] Ioffe, A. D. Metric regularity and subdifferential calculus. *Russian Math. Surveys 55* (2000), 501–558.

[6] Klatte, D., and Kummer, B. *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications*, vol. 60 of *Nonconvex Optimization and Its Applications*. Kluwer Academic Publishers, Dordrecht, 2002.

[7] Kruger, A. Y. On calculus of strict ε-semidifferentials. *Dokl. Akad. Nauk Belarusi 40*, 4 (1996), 34–39. in Russian.

[8] Kruger, A. Y. Strict (ε, δ)-semidifferentials and extremality conditions. *Optimization 51* (2002), 539–554.

[9] Kruger, A. Y. On Fréchet subdifferentials. *J. Math. Sci. (N. Y.) 116*, 3 (2003), 3325–3358. Optimization and related topics, 3.

[10] Kruger, A. Y. Weak stationarity: eliminating the gap between necessary and sufficient conditions. *Optimization 53*, 2 (2004), 147–164.

[11] Kruger, A. Y. Stationarity and regularity of real-valued functions. *Appl. Comput. Math. 5*, 1 (2006), 79–93.

[12] Kummer, B. Inverse functions of pseudo regular mappings and regularity conditions. *Math. Program., Ser. B 88* (2000), 313–339.

[13] Mordukhovich, B. S. *Variational Analysis and Generalized Differentiation. I*, vol. 330 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006. Basic theory.

[14] Phelps, R. R. *Convex Functions, Monotone Operators and Differentiability, 2nd edition*. Lecture Notes in Mathematics, Vol. 1364. Springer-Verlag, New York, 1993.

[15] Rockafellar, R. T., and Wets, R. J.-B. *Variational Analysis*. Springer-Verlag, New York, 1998.

MSC 2000: 90C46, 90C48, 58C20, 58E30
Key words and phrases: variational analysis, subdifferential, normal cone, optimality, stationarity, regularity, slope, multifunction, Asplund space.