Subdifferential and Subgradient

Mathematical Optimization

Georgiy Kormakov

CMC MSU

29 March 2024

Motivation

Subgradient and subdifferential

Subgradient

Let $f : E \to \mathbb{R}$ be a function defined on the set E in the Euclidean space V, and let $x_0 \in E$. The vector $g \in V$ is called the *subgradient* of the function f at the point x_0 if

$$f(x) \geqslant f(x_0) + \langle g, x - x_0 \rangle$$

for all $x \in E$.

Subdifferential at the point

The set of all possible subgradients of the function f at the point x_0 is called the *subdifferential of the function f at the point x_0* and is denoted by $\partial f(x_0)$.

Subdifferentiability

Subdifferential

The function $\partial f : E \to 2^V$ is called the *subdifferential* of the function f.

Subdifferentiability

If $\partial f(x_0) \neq \varnothing$, then f is called *subdifferentiable at the point* x_0;
If f is subdifferentiable at any point in *dom* f, then f is simply called *subdifferentiable*.

Properties

1. If $x_0 \in relint\ E$, then $\partial f(x_0)$ — convex compact set.

Properties

1. If $x_0 \in relint\ E$, then $\partial f(x_0)$ — convex compact set.
2. **Convex** function f is differentiable at $x_0 \Leftrightarrow \partial f(x_0) = \{\nabla f(x_0)\}$

Properties

1. If $x_0 \in \textit{relint}\, E$, then $\partial f(x_0)$ — convex compact set.
2. **Convex** function f is differentiable at $x_0 \Leftrightarrow \partial f(x_0) = \{\nabla f(x_0)\}$
3. If f is subdifferentiable $\forall x_0 \in E$ then $f(x)$ is convex on E

Properties

1. If $x_0 \in relint\ E$, then $\partial f(x_0)$ — convex compact set.
2. **Convex** function f is differentiable at $x_0 \Leftrightarrow \partial f(x_0) = \{\nabla f(x_0)\}$
3. If f is subdifferentiable $\forall x_0 \in E$ then $f(x)$ is convex on E
4. Show that x_0 is the minimum of $f \Leftrightarrow 0 \in \partial f\,(x_0)$.

Counterexample

Moreau-Rockafellar

Theorem

Consider f_i — convex functions on E_i.
If $\bigcap\limits_{i=1}^{n} \textit{relint}\, E_i \neq \varnothing$ and $f(x) = \sum\limits_{i=1}^{n} \alpha_i f_i(x)$, $\alpha_i > 0$,
then exists subdifferential $\partial_E f(x)$ on $E = \bigcap\limits_{i=1}^{n} E_i$ and

$$\partial_E f(x) = \sum_{i=1}^{n} \alpha_i \partial f_i(x)$$

Dubovitsky-Milutin

Theorem

Consider f_i — convex functions on the open convex set $E \subseteq \mathbb{R}^n$.
Let $f(x) = \max_i f_i(x)$, $x_0 \in E$, then

$$\partial f(x_0) = \mathrm{Cl}\left(\mathrm{Conv}\left\{\bigcup_{i \in I(x_0)} \partial f_i(x_0)\right\}\right).$$

$$I(x_0) := \{1 \leqslant i \leqslant m : f_i(x_0) = f(x_0)\}$$

Composition

Chain rule

Consider $g_i, i = \overline{1, m}$ — **convex** functions on the open convex set $E \subseteq \mathbb{R}^n$.
ϕ — a monotonically non-decreasing convex function on an open convex set $U \subseteq \mathbb{R}^m$ and $g(E) \subseteq U$.
Then the subdifferential of the function $f(x) = \phi(g(x))$ is equal to

$$\partial f(x_0) = \bigcup_{p \in \partial \phi(u), u = g(x)} \left(\sum_{i=1}^{m} p_i \partial g_i(x_0) \right)$$

Problems?

1. Let $c > 0$, and let $x_0 \in E$. Show that $\boxed{\partial(cf)(x_0) = c\partial f(x_0).}$

Problems?

1. Let $c > 0$, and let $x_0 \in E$. Show that $\boxed{\partial(cf)(x_0) = c\partial f(x_0).}$
2. Let $m \geqslant 2$ be an integer, $f_i, i = \overline{1, m}$ defined on sets E_i, and let $x_0 \in \cap_{i=1}^m E_i$. Then $\boxed{\partial\left(\sum_{i=1}^m f_i\right)(x_0) \supseteq \sum_{i=1}^m \partial f_i(x_0)}$

   If f_i are convex, and $\bigcap_{i=1}^{m-1} \operatorname{relint}(E_i) \neq \varnothing$, then

   $$\boxed{\partial\left(\sum_{i=1}^m f_i\right)(x_0) = \sum_{i=1}^m \partial f_i(x_0)}$$

Problems?

1. Let $c > 0$, and let $x_0 \in E$. Show that $\boxed{\partial(cf)(x_0) = c\partial f(x_0).}$
2. Let $m \geqslant 2$ be an integer, $f_i, i = \overline{1, m}$ defined on sets E_i, and let $x_0 \in \cap_{i=1}^m E_i$. Then $\boxed{\partial\left(\sum\limits_{i=1}^m f_i\right)(x_0) \supseteq \sum\limits_{i=1}^m \partial f_i(x_0)}$

   If f_i are convex, and $\bigcap\limits_{i=1}^{m-1} \operatorname{relint}(E_i) \neq \varnothing$, then

   $$\boxed{\partial\left(\sum_{i=1}^m f_i\right)(x_0) = \sum_{i=1}^m \partial f_i(x_0)}$$
3. For $L : V \to W$, $L(x) := Ax + b$, $g : E \to \mathbb{R}$, $E \subseteq W$, and $x_0 \in L^{-1}(E)$ subgradient $\boxed{\partial(g \circ L)(x_0) \supseteq A^* \partial g(L(x_0))}$

   If g is convex, $L(V) \cap \operatorname{int}(E) \neq \varnothing$, then

   $\boxed{\partial(g \circ L)(x_0) = A^* \partial g(L(x_0)).}$

Modulus. Graphically

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

Modulus. Properties

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

Modulus 2. Properties

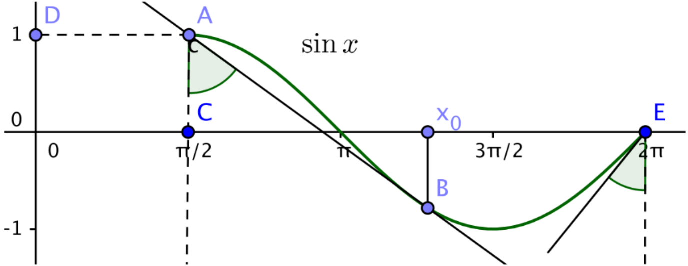
$f(x) = |x - 1| + |x + 1|, x \in \mathbb{R}$

Composition. Properties

$f(x) = (\max(0, f_0(x)))^q, q \geqslant 1$. f_0 — convex on the open convex set E.

Sine?!

$f(x) = \sin x, x \in [\frac{\pi}{2}, 2\pi]$



l_1

$f(x) = \|x\|_1, x \in \mathbb{R}^n$

Linear combination

$f(x) = \|Ax - b\|_1, x \in \mathbb{R}^n$

Let's try

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

Let's try

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

Danskin theorem (not general formulation)

For the function $g(x) = \sup\limits_{\|s\|_* \leqslant 1} \langle s, x \rangle$ the subdifferential is equal

$$\partial g(x) = \mathrm{Cl}\left(\mathrm{Conv}\left\{\bigcup_{\|s\|_*=1} \{s | \langle s, x \rangle = \|x\|\}\right\}\right)$$

Connection with the conjugate function

Theorem

Let $x_0 \in E$, $f : E \to \mathbb{R}, f^* : E^* \to \mathbb{R}$, then

$$\partial f(x_0) = \{s | \langle s, x_0 \rangle = f(x_0) + f^*(s)\}$$

Conclusions

The following equations are equivalent:

1. $\langle s, x_0 \rangle = f(x_0) + f^*(s)$
2. $s \in \partial f(x)$
3. $x \in \partial f^*(s)$

Norm again.