Optimization methods. Seminar 7. Subdifferential

Alexandr Katrutsa

Moscow Institute of Physics and Technology
Department of Control and Applied Mathematics

October 22, 2017

Reminder

- Convex function
- Epigraph and sublevel set
- Criteria of convex function
- Jensen inequality

Motivation

For what?

The important property of any convex function f is that for any point $\mathbf{x}$ for all $\mathbf{y} \in \mathrm{dom}\, f$ the following inequality holds:

$$f(\mathbf{y}) - f(\mathbf{x}) \geq \langle \mathbf{a}, \mathbf{y} - \mathbf{x} \rangle$$

for some vector $\mathbf{a}$, namely tangent hyperplane to the function at the point $\mathbf{x}$ is a global lower bound for the function.

- If the function f is differentiable, then $\mathbf{a} = \nabla f(\mathbf{y})$.
- What if the function f is not differentiable?

Definition

Subgradient

A vector $\mathbf{a}$ is called *subgradient* of a function $f : X \to \mathbb{R}^n$ in a point $\mathbf{x}$, if

$$f(\mathbf{y}) - f(\mathbf{x}) \geq \langle \mathbf{a}, \mathbf{y} - \mathbf{x} \rangle$$

for all $\mathbf{y} \in X$.

Subdifferential

A set of subgradients of the function f in the point $\mathbf{x}$ is called *subdifferential* of the function f in the point $\mathbf{x}$ and is denoted as $\partial f(\mathbf{x})$.

Helpful facts

Moreau-Rockafellar theorem

Let $f_i(\mathbf{x})$ be convex functions defined over convex sets G_i, $i = 1, \dots, n$. Then, if $\bigcap\limits_{i=1}^{n} \mathrm{relint}(G_i) \neq \varnothing$, then a function $f(\mathbf{x}) = \sum\limits_{i=1}^{n} a_i f_i(\mathbf{x})$, $a_i > 0$ has subdifferential $\partial_G f(\mathbf{x})$ on the set $G = \bigcap\limits_{i=1}^{n} G_i$ and $\partial_G f(\mathbf{x}) = \sum\limits_{i=1}^{n} a_i \partial_{G_i} f_i(\mathbf{x})$.

Subdifferential of maximum

If $f(\mathbf{x}) = \max\limits_{i=1,\dots,m} (f_i(\mathbf{x}))$, then

Helpful facts

Moreau-Rockafellar theorem

Let $f_i(\mathbf{x})$ be convex functions defined over convex sets $G_i, \ i = 1, \dots, n$. Then, if $\bigcap\limits_{i=1}^{n} \mathrm{relint}(G_i) \neq \varnothing$, then a function $f(\mathbf{x}) = \sum\limits_{i=1}^{n} a_i f_i(\mathbf{x}), \ a_i > 0$ has subdifferential $\partial_G f(\mathbf{x})$ on the set $G = \bigcap\limits_{i=1}^{n} G_i$ and $\partial_G f(\mathbf{x}) = \sum\limits_{i=1}^{n} a_i \partial_{G_i} f_i(\mathbf{x})$.

Subdifferential of maximum

If $f(\mathbf{x}) = \max\limits_{i=1,\dots,m} (f_i(\mathbf{x}))$, then $\partial_G f(\mathbf{x}) = \mathrm{Conv}\left(\bigcup\limits_{i \in \mathcal{J}(\mathbf{x})} \partial_G f_i(\mathbf{x}) \right)$,
где $\mathcal{J}(\mathbf{x}) = \{ i = 1, \dots, m | f_i(\mathbf{x}) = f(\mathbf{x}) \}$

Examples

Find subdifferential fr the following functions

- Absolute value: $f(x) = |x|$
- ℓ_2 norm: $f(\mathbf{x}) = \|\mathbf{x}\|_2$
- Scalar maximum: $f(x) = \max(e^x, 1 - x)$
- Multivariate maximum: $f(\mathbf{x}) = |\mathbf{c}^\mathsf{T}\mathbf{x}|$
- $f(\mathbf{x}) = |\mathbf{c}_1^\mathsf{T}\mathbf{x}| + |\mathbf{c}_2^\mathsf{T}\mathbf{x}|$

Conditional subdifferential

Definition

A set $\{\mathbf{a} | f(\mathbf{x}) - f(\mathbf{x}_0) \geq \langle \mathbf{a}, \mathbf{x} - \mathbf{x}_0 \rangle, \ \forall \mathbf{x} \in X\}$ is called *subdifferential* of function f in a point $\mathbf{x}_0$ on a set X and denoted as $\partial_X f(\mathbf{x}_0)$.

From conditional subdifferential to unconditional one

If the function f is convex, then consider a function $g(\mathbf{x}) = f(\mathbf{x}) + \delta(\mathbf{x}|X)$, which is also convex. Thus

$$\partial g(\mathbf{x}_0) = \partial_X f(\mathbf{x}_0) = \partial f(\mathbf{x}_0) + \partial \delta(\mathbf{x}_0|X).$$

Find $\partial \delta(\mathbf{x}_0|X)$:

$$\delta(\mathbf{x}|X) - \delta(\mathbf{x}_0|X) \overset{\mathbf{x} \in X}{=} 0 \geq \langle \mathbf{a}, \mathbf{x} - \mathbf{x}_0 \rangle$$

Normal cone

A set $N(\mathbf{x}_0|X) = \{\mathbf{a} | \langle \mathbf{a}, \mathbf{x} - \mathbf{x}_0 \rangle \leq 0, \ \forall \mathbf{x} \in X\}$ is called normal cone to the set X in a point $\mathbf{x}_0$.

Then $\partial_X f(\mathbf{x}_0) = \partial f(\mathbf{x}_0) + N(\mathbf{x}_0|X)$

Examples

- $f(x) = |x|$, $X = \{-1 \leq x \leq 1\}$
- $f(\mathbf{x}) = |x_1 - x_2|$, $X = \{\mathbf{x} \,|\, \|\mathbf{x}\|_2^2 \leq 2\}$

Recap

- Subgradient
- Subdifferential
- Conditional subdifferential
- How to compute them