Методы оптимизации. Семинар 7. Субдифференциал.

Александр Катруца

Московский физико-технический институт,
Факультет Управления и Прикладной Математики

22 октября 2017 г.

Напоминание

- Выпуклая функция
- Надграфик и множество подуровня функции
- Критерии выпуклости функции
- Неравенство Йенсена

Мотивация

Зачем?

Важным свойством непрерывной выпуклой функции f является то, что в выбранной точке $\mathbf{x}$ для всех $\mathbf{y} \in \mathrm{dom}\, f$ выполнено неравенство:

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

для некоторого вектора $\mathbf{a}$, то есть касательная к графику функции является глобальной оценкой снизу для функции.

- Если f дифференцируема, то $\mathbf{a} = \nabla f(\mathbf{y})$.
- Что делать, если f недифференцируема?

Определение

Субградиент

Вектор $\mathbf{a}$ называется субградиентом функции $f : X \to \mathbb{R}^n$ в точке $\mathbf{x}$, если

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

для всех $\mathbf{y} \in X$.

Субдифференциал

Множество субградиентов функции f в точке $\mathbf{x}$ называется субдифференциалом f в $\mathbf{x}$ и обозначается $\partial f(\mathbf{x})$.

Полезные факты

Теорема Моро-Рокафеллара

Пусть $f_i(\mathbf{x})$ — выпуклые функции на выпуклых множествах $G_i,\ i = 1, \dots, n$. Тогда, если $\bigcap\limits_{i=1}^{n} \text{relint}(G_i) \neq \varnothing$ то функция $f(\mathbf{x}) = \sum\limits_{i=1}^{n} a_i f_i(\mathbf{x}),\ a_i > 0$ имеет субдифференциал $\partial_G f(\mathbf{x})$ на множестве $G = \bigcap\limits_{i=1}^{n} G_i$ и $\partial_G f(\mathbf{x}) = \sum\limits_{i=1}^{n} a_i \partial_{G_i} f_i(\mathbf{x})$.

Если функция — максимум

Если $f(\mathbf{x}) = \max\limits_{i=1,\dots,m} (f_i(\mathbf{x}))$, где $f_i(\mathbf{x})$ выпуклы, тогда

Полезные факты

Теорема Моро-Рокафеллара

Пусть $f_i(\mathbf{x})$ — выпуклые функции на выпуклых множествах $G_i, \; i = 1, \ldots, n$. Тогда, если $\bigcap\limits_{i=1}^{n} \text{relint}(G_i) \neq \varnothing$ то функция $f(\mathbf{x}) = \sum\limits_{i=1}^{n} a_i f_i(\mathbf{x}), \; a_i > 0$ имеет субдифференциал $\partial_G f(\mathbf{x})$ на множестве $G = \bigcap\limits_{i=1}^{n} G_i$ и $\partial_G f(\mathbf{x}) = \sum\limits_{i=1}^{n} a_i \partial_{G_i} f_i(\mathbf{x})$.

Если функция — максимум

Если $f(\mathbf{x}) = \max\limits_{i=1,\ldots,m}(f_i(\mathbf{x}))$, где $f_i(\mathbf{x})$ выпуклы, тогда

$\partial_G f(\mathbf{x}) = \text{Conv}\left(\bigcup\limits_{i \in \mathcal{J}(\mathbf{x})} \partial_G f_i(\mathbf{x})\right)$, где

$\mathcal{J}(\mathbf{x}) = \{i = 1, \ldots, m | f_i(\mathbf{x}) = f(\mathbf{x})\}$

Примеры

Найдите субдифференциал для следующих функций.

- Модуль: $f(x) = |x|$
- Норма ℓ_2: $f(\mathbf{x}) = \|\mathbf{x}\|_2$
- Скалярный максимум: $f(x) = \max(e^x, 1 - x)$
- Векторный максимум: $f(\mathbf{x}) = |\mathbf{c}^\top \mathbf{x}|$
- $f(\mathbf{x}) = |\mathbf{c}_1^\top \mathbf{x}| + |\mathbf{c}_2^\top \mathbf{x}|$

Условный субдифференциал

Определение

Множество $\{\mathbf{a} | f(\mathbf{x}) - f(\mathbf{x}_0) \geq \langle \mathbf{a}, \mathbf{x} - \mathbf{x}_0 \rangle, \ \forall \mathbf{x} \in X\}$ называется субдифференциалом f в $\mathbf{x}_0$ на множестве X и обозначается $\partial_X f(\mathbf{x}_0)$.

Как перейти от безусловного субдифференциала к условному?

Если f — выпуклая функция, то рассмотрим функцию $g(\mathbf{x}) = f(\mathbf{x}) + \delta(\mathbf{x}|X)$, которая тоже выпуклая. Тогда

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

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

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

Нормальный конус

Множество $N(\mathbf{x}_0|X) = \{\mathbf{a} | \langle \mathbf{a}, \mathbf{x} - \mathbf{x}_0 \rangle \leq 0, \ \forall \mathbf{x} \in X\}$ называется нормальным конусом к множеству X в точке $\mathbf{x}_0$.

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

Примеры

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

Резюме

- Субградиент
- Субдифференциал
- Условный субдифференциал
- Методы вычислений