Методы оптимизации. Семинар 5. Субдифференциал.

Александр Катруца

Московский физико-технический институт

11 октября 2021 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 $\mathbf{a} = f'(\mathbf{x})$.
- Что делать, если f недифференцируема?

Определение

Субградиент

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

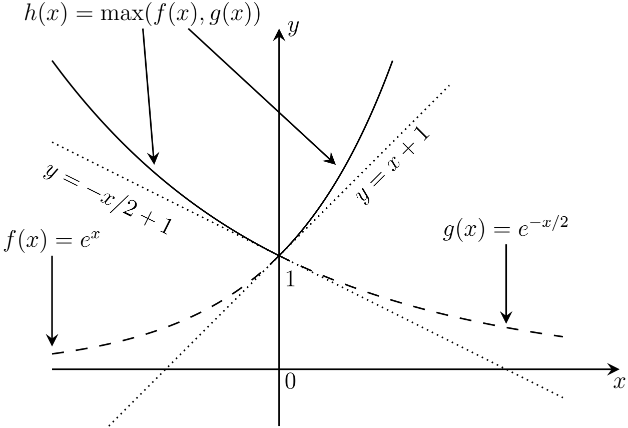
$$f(\mathbf{y}) - f(\mathbf{x}) \geq \langle \mathbf{a}, \mathbf{y} - \mathbf{x} \rangle$$

для всех $\mathbf{y} \in \mathcal{X}$.

Субдифференциал

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

Геометрическая интерпретация



Существование

Q: когда субдифференциал непустое множество?

Существование

Q: когда субдифференциал непустое множество?

Теорема

Если f выпуклая функция, то в любой точке $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\ f)$ выполнено $\partial f(\mathbf{x}) \neq \varnothing$

Полезные факты

Теорема Моро-Рокафеллара

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

Если функция — максимум

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

Полезные факты

Теорема Моро-Рокафеллара

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

Если функция — максимум

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

$\partial_{\mathcal{X}} f(\mathbf{x}) = \operatorname{conv}\left(\bigcup\limits_{i \in \mathcal{J}(\mathbf{x})} \partial_{\mathcal{X}} f_i(\mathbf{x}) \right)$, где

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

Примеры

Найдите субдифференциалы следующих функций.

- Модуль: $f(x) = |x|$
- Норма: $f(\mathbf{x}) = \|\mathbf{x}\|$
- 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 $\{\mathbf{a} \mid f(\mathbf{x}) - f(\mathbf{x}_0) \geq \langle \mathbf{a}, \mathbf{x} - \mathbf{x}_0 \rangle, \ \forall \mathbf{x} \in \mathcal{X}\}$ называется субдифференциалом f в $\mathbf{x}_0$ на множестве $\mathcal{X}$ и обозначается $\partial_{\mathcal{X}} f(\mathbf{x}_0)$.

От безусловного субдифференциала к условному

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

$$\partial g(\mathbf{x}_0) = \partial_{\mathcal{X}} f(\mathbf{x}_0) = \partial f(\mathbf{x}_0) + \partial \delta(\mathbf{x}_0 \mid \mathcal{X}).$$

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

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

Нормальный конус

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

Тогда $\partial_{\mathcal{X}} f(\mathbf{x}_0) = \partial f(\mathbf{x}_0) + N(\mathbf{x}_0 \mid \mathcal{X})$

Примеры

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

Резюме

- Субградиент
- Субдифференциал
- Условный субдифференциал
- Методы вычислений