Методы оптимизации. Семинар 7. Субдифференциал.

Александр Катруца

Московский физико-технический институт, Факультет Управления и Прикладной Математики

16 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 $\mathbf{y} \in \mathsf{dom}\ f$ выполнено неравенство:

$$f(y) - f(x) \ge \langle a, y - x \rangle$$

для некоторого вектора **a**, то есть касательная к графику функции является <mark>глобальной</mark> оценкой снизу для функции.

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

Определение

Субградиент

Вектор **a** называется субградиентов функции $f: X \to \mathbb{R}^n$ в точке **x**, если $f(\mathbf{y}) - f(\mathbf{x}) > \langle \mathbf{a}, \mathbf{y} - \mathbf{x} \rangle$

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

$$G_i,\;i=1,\ldots,n$$
. Тогда, если $igcap_{i=1}^n$ ri $G_i
eq arnothing$ то функция

$$f(\mathbf{x}) = \sum_{i=1}^n a_i f_i(\mathbf{x}), \ a_i > 0$$
 имеет субдифференциал $\partial_G f(\mathbf{x})$

на множестве
$$G = \bigcap_{i=1}^n G_i$$
 и $\partial_G f(\mathbf{x}) = \sum_{i=1}^n a_i \partial_{G_i} f_i(\mathbf{x})$.

Если функция — максимум

Если
$$f(\mathbf{x}) = \max_{i=1,\dots,m} (f_i(\mathbf{x}))$$
, тогда

Полезные факты

Теорема Моро-Рокафеллара

Пусть $f_i(\mathbf{x})$ — выпуклые функции на выпуклых множествах

$$G_i,\;i=1,\ldots,n$$
. Тогда, если $igcap_{i=1}^n$ ri $G_i
eq\varnothing$ то функция

$$f(\mathbf{x}) = \sum_{i=1}^n a_i f_i(\mathbf{x}), \ a_i > 0$$
 имеет субдифференциал $\partial_G f(\mathbf{x})$

на множестве
$$G = \bigcap_{i=1}^n G_i$$
 и $\partial_G f(\mathbf{x}) = \sum_{i=1}^n a_i \partial_{G_i} f_i(\mathbf{x})$.

Если функция — максимум

Если
$$f(\mathbf{x}) = \max_{i=1,\dots,m} (f_i(\mathbf{x}))$$
, тогда

$$\partial_G f(\mathsf{x}) = \mathsf{Conv}(\bigcup_{i \in \mathcal{J}(\mathsf{x})} \partial_G f_i(\mathsf{x}))$$
, где

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

Примеры

Найдите субдифференциал для следующих функций.

- Модуль: f(x) = |x|
- Норма ℓ_2 : $f(\mathbf{x}) = \|\mathbf{x}\|_2$
- ullet Скалярный максимум: $f(x) = \max(e^x, 1-x)$
- ullet Векторный максимум: $f(\mathbf{x}) = |\mathbf{c}^\mathsf{T} \mathbf{x}|$
- $f(x) = |c_1^T x| + |c_2^T x|$

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

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

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

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