Методы оптимизации
Лекция 6: Субдифференциал, его свойства и способы вычисления

Александр Катруца

Физтех-школа прикладной математики и информатики
Московский физико-технический институт

19 октября 2020 г.

На прошлой лекции

- Непрерывность выпуклых функций

На прошлой лекции

- Непрерывность выпуклых функций
- Производная по направлению

На прошлой лекции

- Непрерывность выпуклых функций
- Производная по направлению
- L-гладкость выпуклых функций

План на эту лекцию

- Субградиент

План на эту лекцию

- Субградиент
- Субдифференциал

План на эту лекцию

- Субградиент
- Субдифференциал
- Свойства

План на эту лекцию

- Субградиент
- Субдифференциал
- Свойства
- Способы вычисления

Негладкие выпуклые функции

- На прошлой лекции показали Липшицевость выпуклой функции f в любой точки из $\mathrm{int}\,(\mathrm{dom}\,(f))$

Негладкие выпуклые функции

- На прошлой лекции показали Липшицевость выпуклой функции f в любой точки из $\mathrm{int}\,(\mathrm{dom}\,(f))$
- Что можно сказать про количество точек, в которых она может быть недифференцируема?

Негладкие выпуклые функции

- На прошлой лекции показали Липшицевость выпуклой функции f в любой точки из $\mathrm{int}\,(\mathrm{dom}\,(f))$
- Что можно сказать про количество точек, в которых она может быть недифференцируема?

Теорема (Радемахер)

Пусть задано открытое множество и Липшицева функция на нём. Тогда она дифференцируема почти всюду, то есть мера Лебега точек, в которых она недифференцируема, равна нулю[1].

[1] Доказательство http://www.math.jyu.fi/research/reports/rep100.pdf, стр. 18

Негладкие выпуклые функции

- На прошлой лекции показали Липшицевость выпуклой функции f в любой точки из $\mathrm{int}\,(\mathrm{dom}\,(f))$
- Что можно сказать про количество точек, в которых она может быть недифференцируема?

Теорема (Радемахер)

Пусть задано открытое множество и Липшицева функция на нём. Тогда она дифференцируема почти всюду, то есть мера Лебега точек, в которых она недифференцируема, равна нулю[1].

Типичные источники недифференцируемых выпуклых функций

- Операция взятия максимума
- В частности, использование модуля
- И норм $\|\cdot\|_1$ и $\|\cdot\|_\infty$

[1] Доказательство http://www.math.jyu.fi/research/reports/rep100.pdf, стр. 18

Субградиент и субдифференциал

Определение субградиента

Вектор $\mathbf{g}$ называется *субградиентом* функции f в точке $\mathbf{x}$ если $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle$ для всех точек $\mathbf{y} \in \mathrm{dom}\,(f)$.

Субградиент и субдифференциал

Определение субградиента

Вектор $\mathbf{g}$ называется *субградиентом* функции f в точке $\mathbf{x}$ если $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle$ для всех точек $\mathbf{y} \in \mathrm{dom}\,(f)$.

Определение субдифференциала

Множество всех субградиентов функции f в точке $\mathbf{x}$ называется *субдифференциалом* функции f в точке $\mathbf{x}$ и обозначается $\partial f(\mathbf{x}) = \{\mathbf{g} \mid f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle,$ для всех $\mathbf{y} \in \mathrm{dom}\,(f)\}$.

Субградиент и субдифференциал

Определение субградиента

Вектор $\mathbf{g}$ называется *субградиентом* функции f в точке $\mathbf{x}$ если $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle$ для всех точек $\mathbf{y} \in \mathrm{dom}\,(f)$.

Определение субдифференциала

Множество всех субградиентов функции f в точке $\mathbf{x}$ называется *субдифференциалом* функции f в точке $\mathbf{x}$ и обозначается $\partial f(\mathbf{x}) = \{\mathbf{g} \mid f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle, \text{для всех } \mathbf{y} \in \mathrm{dom}\,(f)\}$.

Замечания

Субградиент и субдифференциал

Определение субградиента

Вектор $\mathbf{g}$ называется *субградиентом* функции f в точке $\mathbf{x}$ если $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle$ для всех точек $\mathbf{y} \in \mathrm{dom}\,(f)$.

Определение субдифференциала

Множество всех субградиентов функции f в точке $\mathbf{x}$ называется *субдифференциалом* функции f в точке $\mathbf{x}$ и обозначается $\partial f(\mathbf{x}) = \{\mathbf{g} \mid f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle,$ для всех $\mathbf{y} \in \mathrm{dom}\,(f)\}$.

Замечания

- В определении не требуется выпуклость функции f

Субградиент и субдифференциал

Определение субградиента

Вектор $\mathbf{g}$ называется *субградиентом* функции f в точке $\mathbf{x}$ если $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle$ для всех точек $\mathbf{y} \in \mathrm{dom}\,(f)$.

Определение субдифференциала

Множество всех субградиентов функции f в точке $\mathbf{x}$ называется *субдифференциалом* функции f в точке $\mathbf{x}$ и обозначается $\partial f(\mathbf{x}) = \{\mathbf{g} \mid f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle,$ для всех $\mathbf{y} \in \mathrm{dom}\,(f)\}$.

Замечания

- В определении не требуется выпуклость функции f
- Геометрически субградиент — вектор нормали гиперплоскости, которая касается графика f и ограничивает её снизу

Субградиент и субдифференциал

Определение субградиента

Вектор $\mathbf{g}$ называется *субградиентом* функции f в точке $\mathbf{x}$ если $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle$ для всех точек $\mathbf{y} \in \mathrm{dom}\,(f)$.

Определение субдифференциала

Множество всех субградиентов функции f в точке $\mathbf{x}$ называется *субдифференциалом* функции f в точке $\mathbf{x}$ и обозначается $\partial f(\mathbf{x}) = \{\mathbf{g} \mid f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle,$ для всех $\mathbf{y} \in \mathrm{dom}\,(f)\}$.

Замечания

- В определении не требуется выпуклость функции f
- Геометрически субградиент — вектор нормали гиперплоскости, которая касается графика f и ограничивает её снизу
- Вопросы непустоты субдифференциала и его вид для дифференцируемых выпуклых функций предстоит обсудить...

Простой пример

1. $f(x) = |x|$

Простой пример

1. $f(x) = |x|$
    - Рассмотрим $\partial f(0)$

Простой пример

1. $f(x) = |x|$
   - Рассмотрим $\partial f(0)$
   - По определению $|y| \geq |0| + a(y - 0)$ для всех y

Простой пример

1. $f(x) = |x|$
    - Рассмотрим $\partial f(0)$
    - По определению $|y| \geq |0| + a(y - 0)$ для всех y
    - Если $y > 0$, получим $y \geq ay$ и $a \leq 1$

Простой пример

1. $f(x) = |x|$
   - Рассмотрим $\partial f(0)$
   - По определению $|y| \geq |0| + a(y - 0)$ для всех y
   - Если $y > 0$, получим $y \geq ay$ и $a \leq 1$
   - Если $y < 0$, получим $-y \geq ay$ и $a \geq -1$

Простой пример

1. $f(x) = |x|$
   - Рассмотрим $\partial f(0)$
   - По определению $|y| \geq |0| + a(y - 0)$ для всех y
   - Если $y > 0$, получим $y \geq ay$ и $a \leq 1$
   - Если $y < 0$, получим $-y \geq ay$ и $a \geq -1$
   - В итоге ответом будет $\partial f(0) = [-1, 1]$

Простой пример

1. $f(x) = |x|$
   - Рассмотрим $\partial f(0)$
   - По определению $|y| \geq |0| + a(y - 0)$ для всех y
   - Если $y > 0$, получим $y \geq ay$ и $a \leq 1$
   - Если $y < 0$, получим $-y \geq ay$ и $a \geq -1$
   - В итоге ответом будет $\partial f(0) = [-1, 1]$
2. Индикаторная функция множества

$$\delta_{\mathcal{X}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \mathcal{X} \\ +\infty, & \mathbf{x} \notin \mathcal{X} \end{cases}$$

Простой пример

1. $f(x) = |x|$
   - Рассмотрим $\partial f(0)$
   - По определению $|y| \geq |0| + a(y - 0)$ для всех y
   - Если $y > 0$, получим $y \geq ay$ и $a \leq 1$
   - Если $y < 0$, получим $-y \geq ay$ и $a \geq -1$
   - В итоге ответом будет $\partial f(0) = [-1, 1]$
2. Индикаторная функция множества
   $$\delta_{\mathcal{X}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \mathcal{X} \\ +\infty, & \mathbf{x} \notin \mathcal{X} \end{cases}$$
   - Рассмотрим произвольную точку $\mathbf{x} \in \mathcal{X}$, тогда $\mathbf{z} \in \partial\delta_{\mathcal{X}}(\mathbf{x})$ iff $\delta_{\mathcal{X}}(\mathbf{y}) \geq \delta_{\mathcal{X}}(\mathbf{x}) + \langle \mathbf{z}, \mathbf{y} - \mathbf{x} \rangle$ для всех $\mathbf{y} \in \mathcal{X}$

Простой пример

1. $f(x) = |x|$
   - Рассмотрим $\partial f(0)$
   - По определению $|y| \geq |0| + a(y - 0)$ для всех y
   - Если $y > 0$, получим $y \geq ay$ и $a \leq 1$
   - Если $y < 0$, получим $-y \geq ay$ и $a \geq -1$
   - В итоге ответом будет $\partial f(0) = [-1, 1]$
2. Индикаторная функция множества
   $$\delta_{\mathcal{X}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \mathcal{X} \\ +\infty, & \mathbf{x} \notin \mathcal{X} \end{cases}$$
   - Рассмотрим произвольную точку $\mathbf{x} \in \mathcal{X}$, тогда $\mathbf{z} \in \partial \delta_{\mathcal{X}}(\mathbf{x})$ iff $\delta_{\mathcal{X}}(\mathbf{y}) \geq \delta_{\mathcal{X}}(\mathbf{x}) + \langle \mathbf{z}, \mathbf{y} - \mathbf{x} \rangle$ для всех $\mathbf{y} \in \mathcal{X}$
   - Что значит $\langle \mathbf{z}, \mathbf{y} - \mathbf{x} \rangle \leq 0$ для всех $\mathbf{y} \in \mathcal{X}$

Простой пример

1. $f(x) = |x|$
   - Рассмотрим $\partial f(0)$
   - По определению $|y| \geq |0| + a(y - 0)$ для всех y
   - Если $y > 0$, получим $y \geq ay$ и $a \leq 1$
   - Если $y < 0$, получим $-y \geq ay$ и $a \geq -1$
   - В итоге ответом будет $\partial f(0) = [-1, 1]$
2. Индикаторная функция множества
   $$\delta_{\mathcal{X}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \mathcal{X} \\ +\infty, & \mathbf{x} \notin \mathcal{X} \end{cases}$$
   - Рассмотрим произвольную точку $\mathbf{x} \in \mathcal{X}$, тогда $\mathbf{z} \in \partial \delta_{\mathcal{X}}(\mathbf{x})$ iff $\delta_{\mathcal{X}}(\mathbf{y}) \geq \delta_{\mathcal{X}}(\mathbf{x}) + \langle \mathbf{z}, \mathbf{y} - \mathbf{x} \rangle$ для всех $\mathbf{y} \in \mathcal{X}$
   - Что значит $\langle \mathbf{z}, \mathbf{y} - \mathbf{x} \rangle \leq 0$ для всех $\mathbf{y} \in \mathcal{X}$
   - $\mathcal{N}_{\mathcal{X}}(\mathbf{x}) = \{\mathbf{z} \mid \langle \mathbf{z}, \mathbf{y} - \mathbf{x} \rangle \leq 0$ для всех $\mathbf{y} \in \mathcal{X}\}$ — нормальный конус к множеству $\mathcal{X}$ в точке $\mathbf{x}$

Простой пример

1. $f(x) = |x|$
   - Рассмотрим $\partial f(0)$
   - По определению $|y| \geq |0| + a(y - 0)$ для всех y
   - Если $y > 0$, получим $y \geq ay$ и $a \leq 1$
   - Если $y < 0$, получим $-y \geq ay$ и $a \geq -1$
   - В итоге ответом будет $\partial f(0) = [-1, 1]$
2. Индикаторная функция множества

   $$\delta_{\mathcal{X}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \mathcal{X} \\ +\infty, & \mathbf{x} \notin \mathcal{X} \end{cases}$$

   - Рассмотрим произвольную точку $\mathbf{x} \in \mathcal{X}$, тогда $\mathbf{z} \in \partial \delta_{\mathcal{X}}(\mathbf{x})$ iff $\delta_{\mathcal{X}}(\mathbf{y}) \geq \delta_{\mathcal{X}}(\mathbf{x}) + \langle \mathbf{z}, \mathbf{y} - \mathbf{x} \rangle$ для всех $\mathbf{y} \in \mathcal{X}$
   - Что значит $\langle \mathbf{z}, \mathbf{y} - \mathbf{x} \rangle \leq 0$ для всех $\mathbf{y} \in \mathcal{X}$
   - $\mathcal{N}_{\mathcal{X}}(\mathbf{x}) = \{\mathbf{z} \mid \langle \mathbf{z}, \mathbf{y} - \mathbf{x} \rangle \leq 0$ для всех $\mathbf{y} \in \mathcal{X}\}$ — нормальный конус к множеству $\mathcal{X}$ в точке $\mathbf{x}$
   - Это конус, он выпуклый и замкнутый (проверьте почему!)

Свойства субдифференциала

Теорема о выпуклости и замкнутости

Субдифференциал является выпуклым и замкнутым множеством.

Свойства субдифференциала

Теорема о выпуклости и замкнутости

Субдифференциал является выпуклым и замкнутым множеством.

Доказательство

Свойства субдифференциала

Теорема о выпуклости и замкнутости

Субдифференциал является выпуклым и замкнутым множеством.

Доказательство

- Для произвольного $\mathbf{x}$ субдифференциал $\partial f(\mathbf{x}) = \bigcap\limits_{\mathbf{y} \in \mathrm{dom}(f)} \{\mathbf{g} \mid \langle \mathbf{y} - \mathbf{x}, \mathbf{g} \rangle \leq f(\mathbf{y}) - f(\mathbf{x})\}$

Свойства субдифференциала

Теорема о выпуклости и замкнутости

Субдифференциал является выпуклым и замкнутым множеством.

Доказательство

- Для произвольного $\mathbf{x}$ субдифференциал $\partial f(\mathbf{x}) = \bigcap\limits_{\mathbf{y} \in \mathrm{dom}(f)} \{\mathbf{g} \mid \langle \mathbf{y} - \mathbf{x}, \mathbf{g} \rangle \leq f(\mathbf{y}) - f(\mathbf{x})\}$
- Множества, которые пересекаются, — полупространства

Свойства субдифференциала

Теорема о выпуклости и замкнутости

Субдифференциал является выпуклым и замкнутым множеством.

Доказательство

- Для произвольного $\mathbf{x}$ субдифференциал $\partial f(\mathbf{x}) = \bigcap\limits_{\mathbf{y} \in \mathrm{dom}(f)} \{\mathbf{g} \mid \langle \mathbf{y} - \mathbf{x}, \mathbf{g} \rangle \leq f(\mathbf{y}) - f(\mathbf{x})\}$
- Множества, которые пересекаются, — полупространства
- Они выпуклы и замкнуты

Свойства субдифференциала

Теорема о выпуклости и замкнутости

Субдифференциал является выпуклым и замкнутым множеством.

Доказательство

- Для произвольного $\mathbf{x}$ субдифференциал $\partial f(\mathbf{x}) = \bigcap\limits_{\mathbf{y} \in \mathrm{dom}(f)} \{\mathbf{g} \mid \langle \mathbf{y} - \mathbf{x}, \mathbf{g} \rangle \leq f(\mathbf{y}) - f(\mathbf{x})\}$
- Множества, которые пересекаются, — полупространства
- Они выпуклы и замкнуты
- Значит их пересечение также выпукло и замкнуто

Свойства субдифференциала

Теорема о выпуклости и замкнутости

Субдифференциал является выпуклым и замкнутым множеством.

Доказательство

- Для произвольного $\mathbf{x}$ субдифференциал $\partial f(\mathbf{x}) = \bigcap\limits_{\mathbf{y} \in \mathrm{dom}(f)} \{\mathbf{g} \mid \langle \mathbf{y} - \mathbf{x}, \mathbf{g} \rangle \leq f(\mathbf{y}) - f(\mathbf{x})\}$
- Множества, которые пересекаются, — полупространства
- Они выпуклы и замкнуты
- Значит их пересечение также выпукло и замкнуто

Замечания

- Субдифференциал может быть пустым множеством
- Обозначим точки, в которых субдифференциал непустое множество $\mathrm{dom}\,(\partial f)$

Признак выпуклости

Теорема

Если у функции f выпуклая область определения и в каждой точке из $\mathrm{dom}\,(f)$ субдифференциал непуст, тогда f выпукла.

Признак выпуклости

Теорема

Если у функции f выпуклая область определения и в каждой точке из $\mathrm{dom}\,(f)$ субдифференциал непуст, тогда f выпукла.

Доказательство

Признак выпуклости

Теорема

Если у функции f выпуклая область определения и в каждой точке из $\mathrm{dom}\,(f)$ субдифференциал непуст, тогда f выпукла.

Доказательство

- Возьмём произвольные $\mathbf{x}, \mathbf{y} \in \mathrm{dom}\,(f)$ и $\alpha \in [0, 1]$

Признак выпуклости

Теорема

Если у функции f выпуклая область определения и в каждой точке из $\mathrm{dom}\,(f)$ субдифференциал непуст, тогда f выпукла.

Доказательство

- Возьмём произвольные $\mathbf{x}, \mathbf{y} \in \mathrm{dom}\,(f)$ и $\alpha \in [0, 1]$
- Рассмотрим $\mathbf{z}_\alpha = \alpha\mathbf{y} + (1 - \alpha)\mathbf{x} \in \mathrm{dom}\,(f)$ в силу выпуклости области определения

Признак выпуклости

Теорема

Если у функции f выпуклая область определения и в каждой точке из $\mathrm{dom}\,(f)$ субдифференциал непуст, тогда f выпукла.

Доказательство

- Возьмём произвольные $\mathbf{x}, \mathbf{y} \in \mathrm{dom}\,(f)$ и $\alpha \in [0, 1]$
- Рассмотрим $\mathbf{z}_\alpha = \alpha\mathbf{y} + (1 - \alpha)\mathbf{x} \in \mathrm{dom}\,(f)$ в силу выпуклости области определения
- Значит $\mathbf{g} \in \partial f(\mathbf{z}_\alpha)$ существует

Признак выпуклости

Теорема

Если у функции f выпуклая область определения и в каждой точке из $\mathrm{dom}\,(f)$ субдифференциал непуст, тогда f выпукла.

Доказательство

- Возьмём произвольные $\mathbf{x}, \mathbf{y} \in \mathrm{dom}\,(f)$ и $\alpha \in [0, 1]$
- Рассмотрим $\mathbf{z}_\alpha = \alpha\mathbf{y} + (1-\alpha)\mathbf{x} \in \mathrm{dom}\,(f)$ в силу выпуклости области определения
- Значит $\mathbf{g} \in \partial f(\mathbf{z}_\alpha)$ существует
- Запишем два неравенства

$$f(\mathbf{y}) \geq f(\mathbf{z}_\alpha) + \langle \mathbf{g}, \mathbf{y} - \mathbf{z}_\alpha \rangle = f(\mathbf{z}_\alpha) + (1-\alpha)\langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle \mid \cdot \alpha$$
$$f(\mathbf{x}) \geq f(\mathbf{z}_\alpha) + \langle \mathbf{g}, \mathbf{x} - \mathbf{z}_\alpha \rangle = f(\mathbf{z}_\alpha) - \alpha\langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle \mid \cdot (1-\alpha)$$

Признак выпуклости

Теорема

Если у функции f выпуклая область определения и в каждой точке из $\mathrm{dom}\,(f)$ субдифференциал непуст, тогда f выпукла.

Доказательство

- Возьмём произвольные $\mathbf{x}, \mathbf{y} \in \mathrm{dom}\,(f)$ и $\alpha \in [0, 1]$
- Рассмотрим $\mathbf{z}_\alpha = \alpha\mathbf{y} + (1-\alpha)\mathbf{x} \in \mathrm{dom}\,(f)$ в силу выпуклости области определения
- Значит $\mathbf{g} \in \partial f(\mathbf{z}_\alpha)$ существует
- Запишем два неравенства

$$f(\mathbf{y}) \geq f(\mathbf{z}_\alpha) + \langle \mathbf{g}, \mathbf{y} - \mathbf{z}_\alpha \rangle = f(\mathbf{z}_\alpha) + (1-\alpha)\langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle \mid \cdot\, \alpha$$
$$f(\mathbf{x}) \geq f(\mathbf{z}_\alpha) + \langle \mathbf{g}, \mathbf{x} - \mathbf{z}_\alpha \rangle = f(\mathbf{z}_\alpha) - \alpha\langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle \mid \cdot\, (1-\alpha)$$

- Сложим их и получим, что $f(\mathbf{z}_\alpha) \leq \alpha f(\mathbf{y}) + (1-\alpha) f(\mathbf{x})$

Признак выпуклости

Теорема

Если у функции f выпуклая область определения и в каждой точке из $\mathrm{dom}\,(f)$ субдифференциал непуст, тогда f выпукла.

Доказательство

- Возьмём произвольные $\mathbf{x}, \mathbf{y} \in \mathrm{dom}\,(f)$ и $\alpha \in [0, 1]$
- Рассмотрим $\mathbf{z}_\alpha = \alpha\mathbf{y} + (1-\alpha)\mathbf{x} \in \mathrm{dom}\,(f)$ в силу выпуклости области определения
- Значит $\mathbf{g} \in \partial f(\mathbf{z}_\alpha)$ существует
- Запишем два неравенства

$$
\begin{aligned}
f(\mathbf{y}) &\geq f(\mathbf{z}_\alpha) + \langle \mathbf{g}, \mathbf{y} - \mathbf{z}_\alpha \rangle = f(\mathbf{z}_\alpha) + (1-\alpha)\langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle \mid \cdot\, \alpha \\
f(\mathbf{x}) &\geq f(\mathbf{z}_\alpha) + \langle \mathbf{g}, \mathbf{x} - \mathbf{z}_\alpha \rangle = f(\mathbf{z}_\alpha) - \alpha\langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle \mid \cdot\, (1-\alpha)
\end{aligned}
$$

- Сложим их и получим, что $f(\mathbf{z}_\alpha) \leq \alpha f(\mathbf{y}) + (1-\alpha) f(\mathbf{x})$
- Это выполнено для любых $\mathbf{x}, \mathbf{y}$ из выпуклой области определения, значит f выпукла

Существование

- Если субдифференциал не пуст в любой точки выпуклой области определения, то функция выпукла

Существование

- Если субдифференциал не пуст в любой точки выпуклой области определения, то функция выпукла
- Обратное утверждение неверно!

Существование

- Если субдифференциал не пуст в любой точки выпуклой области определения, то функция выпукла
- Обратное утверждение неверно!
- Пример $f(x) = -\sqrt{x}, x \geq 0$

Существование

- Если субдифференциал не пуст в любой точки выпуклой области определения, то функция выпукла
- Обратное утверждение неверно!
- Пример $f(x) = -\sqrt{x}, x \geq 0$
- При $x = 0$ субдифференциал — пустое множество (покажите это!), хотя функция выпукла.

Существование

- Если субдифференциал не пуст в любой точки выпуклой области определения, то функция выпукла
- Обратное утверждение неверно!
- Пример $f(x) = -\sqrt{x}, x \geq 0$
- При $x = 0$ субдифференциал — пустое множество (покажите это!), хотя функция выпукла.

Теорема о существовании

Пусть f выпуклая функция и $\hat{\mathbf{x}} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Тогда $\partial f(\hat{\mathbf{x}}) \neq \varnothing$ и ограниченное множество.

Существование

- Если субдифференциал не пуст в любой точки выпуклой области определения, то функция выпукла
- Обратное утверждение неверно!
- Пример $f(x) = -\sqrt{x}, x \geq 0$
- При $x = 0$ субдифференциал — пустое множество (покажите это!), хотя функция выпукла.

Теорема о существовании

Пусть f выпуклая функция и $\hat{\mathbf{x}} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Тогда $\partial f(\hat{\mathbf{x}}) \neq \varnothing$ и ограниченное множество.

Доказательство

Существование

- Если субдифференциал не пуст в любой точки выпуклой области определения, то функция выпукла
- Обратное утверждение неверно!
- Пример $f(x) = -\sqrt{x}, x \geq 0$
- При $x = 0$ субдифференциал — пустое множество (покажите это!), хотя функция выпукла.

Теорема о существовании

Пусть f выпуклая функция и $\hat{\mathbf{x}} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Тогда $\partial f(\hat{\mathbf{x}}) \neq \varnothing$ и ограниченное множество.

Доказательство

- Рассмотрим точку $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}})) \in \partial \mathrm{epi} f$

Существование

- Если субдифференциал не пуст в любой точки выпуклой области определения, то функция выпукла
- Обратное утверждение неверно!
- Пример $f(x) = -\sqrt{x}, x \geq 0$
- При $x = 0$ субдифференциал — пустое множество (покажите это!), хотя функция выпукла.

Теорема о существовании

Пусть f выпуклая функция и $\hat{\mathbf{x}} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Тогда $\partial f(\hat{\mathbf{x}}) \neq \varnothing$ и ограниченное множество.

Доказательство

- Рассмотрим точку $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}})) \in \partial \mathrm{epi} f$
- $\mathrm{epi} f$ выпуклое множество в силу выпуклости f

Существование

- Если субдифференциал не пуст в любой точки выпуклой области определения, то функция выпукла
- Обратное утверждение неверно!
- Пример $f(x) = -\sqrt{x}, x \geq 0$
- При $x = 0$ субдифференциал — пустое множество (покажите это!), хотя функция выпукла.

Теорема о существовании

Пусть f выпуклая функция и $\hat{\mathbf{x}} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Тогда $\partial f(\hat{\mathbf{x}}) \neq \varnothing$ и ограниченное множество.

Доказательство

- Рассмотрим точку $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}})) \in \partial \mathrm{epi} f$
- $\mathrm{epi} f$ выпуклое множество в силу выпуклости f
- Значит через $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}}))$ можно провести опорную гиперплоскость

- То есть $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \mathbf{x} \rangle - \alpha t$ для всех $(\mathbf{x}, t) \in \mathrm{epi} f$

- То есть $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \mathbf{x} \rangle - \alpha t$ для всех $(\mathbf{x}, t) \in \mathrm{epi} f$
- Покажем, что $\alpha \geq 0$

- То есть $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \mathbf{x} \rangle - \alpha t$ для всех $(\mathbf{x}, t) \in \mathrm{epi} f$
- Покажем, что $\alpha \geq 0$
    - Точка $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}}) + 1) \in \mathrm{epi} f$

- То есть $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \mathbf{x} \rangle - \alpha t$ для всех $(\mathbf{x}, t) \in \mathrm{epi} f$
- Покажем, что $\alpha \geq 0$
  - Точка $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}}) + 1) \in \mathrm{epi} f$
  - Подставим её в неравенство для опорной гиперплоскости: $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha(f(\hat{\mathbf{x}}) + 1)$

- То есть $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \mathbf{x} \rangle - \alpha t$ для всех $(\mathbf{x}, t) \in \mathrm{epi} f$
- Покажем, что $\alpha \geq 0$
  - Точка $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}}) + 1) \in \mathrm{epi} f$
  - Подставим её в неравенство для опорной гиперплоскости: $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha(f(\hat{\mathbf{x}}) + 1)$
  - Получаем $\alpha \geq 0$

- То есть $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \mathbf{x} \rangle - \alpha t$ для всех $(\mathbf{x}, t) \in \mathrm{epi} f$
- Покажем, что $\alpha \geq 0$
  - Точка $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}}) + 1) \in \mathrm{epi} f$
  - Подставим её в неравенство для опорной гиперплоскости: $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha(f(\hat{\mathbf{x}}) + 1)$
  - Получаем $\alpha \geq 0$
- Покажем, что $\alpha > 0$

- То есть $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \mathbf{x} \rangle - \alpha t$ для всех $(\mathbf{x}, t) \in \mathrm{epi} f$
- Покажем, что $\alpha \geq 0$
  - Точка $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}}) + 1) \in \mathrm{epi} f$
  - Подставим её в неравенство для опорной гиперплоскости: $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha(f(\hat{\mathbf{x}}) + 1)$
  - Получаем $\alpha \geq 0$
- Покажем, что $\alpha > 0$
  - Так как $\hat{\mathbf{x}} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ то f Липшицева в этой точке, то есть найдётся $\varepsilon > 0$ и $L > 0$ что для всех $\mathbf{x} \in B_2(\varepsilon, \hat{\mathbf{x}}) \subset \mathrm{dom}\,(f)$ выполнено $|f(\mathbf{x}) - f(\hat{\mathbf{x}})| \leq L\|\mathbf{x} - \hat{\mathbf{x}}\|_2$

- То есть $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \mathbf{x} \rangle - \alpha t$ для всех $(\mathbf{x}, t) \in \mathrm{epi} f$
- Покажем, что $\alpha \geq 0$
  - Точка $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}}) + 1) \in \mathrm{epi} f$
  - Подставим её в неравенство для опорной гиперплоскости: $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha(f(\hat{\mathbf{x}}) + 1)$
  - Получаем $\alpha \geq 0$
- Покажем, что $\alpha > 0$
  - Так как $\hat{\mathbf{x}} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ то f Липшицева в этой точке, то есть найдётся $\varepsilon > 0$ и $L > 0$ что для всех $\mathbf{x} \in B_2(\varepsilon, \hat{\mathbf{x}}) \subset \mathrm{dom}\,(f)$ выполнено $|f(\mathbf{x}) - f(\hat{\mathbf{x}})| \leq L\|\mathbf{x} - \hat{\mathbf{x}}\|_2$
  - $(\mathbf{x}, f(\mathbf{x})) \in \mathrm{epi} f$ для всех $\mathbf{x} \in B_2(\varepsilon, \hat{\mathbf{x}})$

- То есть $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \mathbf{x} \rangle - \alpha t$ для всех $(\mathbf{x}, t) \in \mathrm{epi} f$
- Покажем, что $\alpha \geq 0$
    - Точка $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}}) + 1) \in \mathrm{epi} f$
    - Подставим её в неравенство для опорной гиперплоскости: $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha(f(\hat{\mathbf{x}}) + 1)$
    - Получаем $\alpha \geq 0$
- Покажем, что $\alpha > 0$
    - Так как $\hat{\mathbf{x}} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ то f Липшицева в этой точке, то есть найдётся $\varepsilon > 0$ и $L > 0$ что для всех $\mathbf{x} \in B_2(\varepsilon, \hat{\mathbf{x}}) \subset \mathrm{dom}\,(f)$ выполнено $|f(\mathbf{x}) - f(\hat{\mathbf{x}})| \leq L\|\mathbf{x} - \hat{\mathbf{x}}\|_2$
    - $(\mathbf{x}, f(\mathbf{x})) \in \mathrm{epi} f$ для всех $\mathbf{x} \in B_2(\varepsilon, \hat{\mathbf{x}})$
    - Подставим эту точку в неравенство для опорной гиперплоскости: $\langle \mathbf{p}, \mathbf{x} - \hat{\mathbf{x}} \rangle \leq \alpha(f(\mathbf{x}) - f(\hat{\mathbf{x}})) \leq \alpha L\|\mathbf{x} - \hat{\mathbf{x}}\|_2$

- То есть $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \mathbf{x} \rangle - \alpha t$ для всех $(\mathbf{x}, t) \in \mathrm{epi} f$
- Покажем, что $\alpha \geq 0$
  - Точка $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}}) + 1) \in \mathrm{epi} f$
  - Подставим её в неравенство для опорной гиперплоскости: $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha(f(\hat{\mathbf{x}}) + 1)$
  - Получаем $\alpha \geq 0$
- Покажем, что $\alpha > 0$
  - Так как $\hat{\mathbf{x}} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ то f Липшицева в этой точке, то есть найдётся $\varepsilon > 0$ и $L > 0$ что для всех $\mathbf{x} \in B_2(\varepsilon, \hat{\mathbf{x}}) \subset \mathrm{dom}\,(f)$ выполнено $|f(\mathbf{x}) - f(\hat{\mathbf{x}})| \leq L\|\mathbf{x} - \hat{\mathbf{x}}\|_2$
  - $(\mathbf{x}, f(\mathbf{x})) \in \mathrm{epi} f$ для всех $\mathbf{x} \in B_2(\varepsilon, \hat{\mathbf{x}})$
  - Подставим эту точку в неравенство для опорной гиперплоскости: $\langle \mathbf{p}, \mathbf{x} - \hat{\mathbf{x}} \rangle \leq \alpha(f(\mathbf{x}) - f(\hat{\mathbf{x}})) \leq \alpha L\|\mathbf{x} - \hat{\mathbf{x}}\|_2$
  - Возьмём $\mathbf{d}$ такой что $\|\mathbf{d}\|_2 = 1$ и $\langle \mathbf{p}, \mathbf{d} \rangle = \|\mathbf{p}\|_2$, и представим $\mathbf{x} \in B_2(\varepsilon, \hat{\mathbf{x}})$ в виде $\hat{\mathbf{x}} + \varepsilon\mathbf{d}$

- То есть $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \mathbf{x} \rangle - \alpha t$ для всех $(\mathbf{x}, t) \in \mathrm{epi} f$
- Покажем, что $\alpha \geq 0$
  - Точка $(\hat{\mathbf{x}}, f(\hat{\mathbf{x}}) + 1) \in \mathrm{epi} f$
  - Подставим её в неравенство для опорной гиперплоскости: $\langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha f(\hat{\mathbf{x}}) \geq \langle \mathbf{p}, \hat{\mathbf{x}} \rangle - \alpha(f(\hat{\mathbf{x}}) + 1)$
  - Получаем $\alpha \geq 0$
- Покажем, что $\alpha > 0$
  - Так как $\hat{\mathbf{x}} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ то f Липшицева в этой точке, то есть найдётся $\varepsilon > 0$ и $L > 0$ что для всех $\mathbf{x} \in B_2(\varepsilon, \hat{\mathbf{x}}) \subset \mathrm{dom}\,(f)$ выполнено $|f(\mathbf{x}) - f(\hat{\mathbf{x}})| \leq L \|\mathbf{x} - \hat{\mathbf{x}}\|_2$
  - $(\mathbf{x}, f(\mathbf{x})) \in \mathrm{epi} f$ для всех $\mathbf{x} \in B_2(\varepsilon, \hat{\mathbf{x}})$
  - Подставим эту точку в неравенство для опорной гиперплоскости: $\langle \mathbf{p}, \mathbf{x} - \hat{\mathbf{x}} \rangle \leq \alpha(f(\mathbf{x}) - f(\hat{\mathbf{x}})) \leq \alpha L \|\mathbf{x} - \hat{\mathbf{x}}\|_2$
  - Возьмём $\mathbf{d}$ такой что $\|\mathbf{d}\|_2 = 1$ и $\langle \mathbf{p}, \mathbf{d} \rangle = \|\mathbf{p}\|_2$, и представим $\mathbf{x} \in B_2(\varepsilon, \hat{\mathbf{x}})$ в виде $\hat{\mathbf{x}} + \varepsilon \mathbf{d}$
  - Тогда $\varepsilon \|\mathbf{p}\|_2 \leq \alpha L \varepsilon \|\mathbf{d}\|_2$ и если $\alpha = 0$, то и $\mathbf{p} = 0$, что противоречит существованию опорной гиперплоскости

- Возьмём $t = f(\mathbf{x})$ в неравенстве для опорной гиперплоскости и получим $\alpha f(\mathbf{x}) \geq \alpha f(\hat{\mathbf{x}}) + \langle \mathbf{p}, \mathbf{x} - \hat{\mathbf{x}} \rangle$

- Возьмём $t = f(\mathbf{x})$ в неравенстве для опорной гиперплоскости и получим $\alpha f(\mathbf{x}) \geq \alpha f(\hat{\mathbf{x}}) + \langle \mathbf{p}, \mathbf{x} - \hat{\mathbf{x}} \rangle$
- Поделив обе части на α, получим, что субградиент $\mathbf{g} = \mathbf{p}/\alpha$ в точке $\hat{\mathbf{x}}$

- Возьмём $t = f(\mathbf{x})$ в неравенстве для опорной гиперплоскости и получим $\alpha f(\mathbf{x}) \geq \alpha f(\hat{\mathbf{x}}) + \langle \mathbf{p}, \mathbf{x} - \hat{\mathbf{x}} \rangle$
- Поделив обе части на α, получим, что субградиент $\mathbf{g} = \mathbf{p}/\alpha$ в точке $\hat{\mathbf{x}}$
- Значит $\partial f(\hat{\mathbf{x}}) \neq \varnothing$

- Возьмём $t = f(\mathbf{x})$ в неравенстве для опорной гиперплоскости и получим $\alpha f(\mathbf{x}) \geq \alpha f(\hat{\mathbf{x}}) + \langle \mathbf{p}, \mathbf{x} - \hat{\mathbf{x}} \rangle$
- Поделив обе части на α, получим, что субградиент $\mathbf{g} = \mathbf{p}/\alpha$ в точке $\hat{\mathbf{x}}$
- Значит $\partial f(\hat{\mathbf{x}}) \neq \varnothing$
- Покажем ограниченность. Пусть $\mathbf{g} \in \partial f(\hat{\mathbf{x}})$, значит $f(\mathbf{x}) \geq f(\hat{\mathbf{x}}) + \langle \mathbf{g}, \mathbf{x} - \hat{\mathbf{x}} \rangle$ для всех $\mathbf{x} \in \operatorname{dom}(f)$

- Возьмём $t = f(\mathbf{x})$ в неравенстве для опорной гиперплоскости и получим $\alpha f(\mathbf{x}) \geq \alpha f(\hat{\mathbf{x}}) + \langle \mathbf{p}, \mathbf{x} - \hat{\mathbf{x}} \rangle$
- Поделив обе части на α, получим, что субградиент $\mathbf{g} = \mathbf{p}/\alpha$ в точке $\hat{\mathbf{x}}$
- Значит $\partial f(\hat{\mathbf{x}}) \neq \varnothing$
- Покажем ограниченность. Пусть $\mathbf{g} \in \partial f(\hat{\mathbf{x}})$, значит $f(\mathbf{x}) \geq f(\hat{\mathbf{x}}) + \langle \mathbf{g}, \mathbf{x} - \hat{\mathbf{x}} \rangle$ для всех $\mathbf{x} \in \mathrm{dom}\,(f)$
- Возьмём $\mathbf{x} = \hat{\mathbf{x}} + \varepsilon \mathbf{d}$, где $\mathbf{d}$ такой что $\|\mathbf{d}\|_2 = 1$ и $\langle \mathbf{g}, \mathbf{d} \rangle = \|\mathbf{g}\|_2$

- Возьмём $t = f(\mathbf{x})$ в неравенстве для опорной гиперплоскости и получим $\alpha f(\mathbf{x}) \geq \alpha f(\hat{\mathbf{x}}) + \langle \mathbf{p}, \mathbf{x} - \hat{\mathbf{x}} \rangle$
- Поделив обе части на α, получим, что субградиент $\mathbf{g} = \mathbf{p}/\alpha$ в точке $\hat{\mathbf{x}}$
- Значит $\partial f(\hat{\mathbf{x}}) \neq \varnothing$
- Покажем ограниченность. Пусть $\mathbf{g} \in \partial f(\hat{\mathbf{x}})$, значит $f(\mathbf{x}) \geq f(\hat{\mathbf{x}}) + \langle \mathbf{g}, \mathbf{x} - \hat{\mathbf{x}} \rangle$ для всех $\mathbf{x} \in \mathrm{dom}\,(f)$
- Возьмём $\mathbf{x} = \hat{\mathbf{x}} + \varepsilon \mathbf{d}$, где $\mathbf{d}$ такой что $\|\mathbf{d}\|_2 = 1$ и $\langle \mathbf{g}, \mathbf{d} \rangle = \|\mathbf{g}\|_2$
- Тогда $\varepsilon \|\mathbf{g}\|_2 \leq f(\mathbf{x}) - f(\hat{\mathbf{x}}) \leq L \|\mathbf{x} - \hat{\mathbf{x}}\|_2 = L\varepsilon$

- Возьмём $t = f(\mathbf{x})$ в неравенстве для опорной гиперплоскости и получим $\alpha f(\mathbf{x}) \geq \alpha f(\hat{\mathbf{x}}) + \langle \mathbf{p}, \mathbf{x} - \hat{\mathbf{x}} \rangle$
- Поделив обе части на α, получим, что субградиент $\mathbf{g} = \mathbf{p}/\alpha$ в точке $\hat{\mathbf{x}}$
- Значит $\partial f(\hat{\mathbf{x}}) \neq \varnothing$
- Покажем ограниченность. Пусть $\mathbf{g} \in \partial f(\hat{\mathbf{x}})$, значит $f(\mathbf{x}) \geq f(\hat{\mathbf{x}}) + \langle \mathbf{g}, \mathbf{x} - \hat{\mathbf{x}} \rangle$ для всех $\mathbf{x} \in \mathrm{dom}\,(f)$
- Возьмём $\mathbf{x} = \hat{\mathbf{x}} + \varepsilon \mathbf{d}$, где $\mathbf{d}$ такой что $\|\mathbf{d}\|_2 = 1$ и $\langle \mathbf{g}, \mathbf{d} \rangle = \|\mathbf{g}\|_2$
- Тогда $\varepsilon \|\mathbf{g}\|_2 \leq f(\mathbf{x}) - f(\hat{\mathbf{x}}) \leq L\|\mathbf{x} - \hat{\mathbf{x}}\|_2 = L\varepsilon$
- Значит $\mathbf{g} \in B_2(L, 0)$, то есть субдифференциал является ограниченным множеством

- Возьмём $t = f(\mathbf{x})$ в неравенстве для опорной гиперплоскости и получим $\alpha f(\mathbf{x}) \geq \alpha f(\hat{\mathbf{x}}) + \langle \mathbf{p}, \mathbf{x} - \hat{\mathbf{x}} \rangle$
- Поделив обе части на α, получим, что субградиент $\mathbf{g} = \mathbf{p}/\alpha$ в точке $\hat{\mathbf{x}}$
- Значит $\partial f(\hat{\mathbf{x}}) \neq \varnothing$
- Покажем ограниченность. Пусть $\mathbf{g} \in \partial f(\hat{\mathbf{x}})$, значит $f(\mathbf{x}) \geq f(\hat{\mathbf{x}}) + \langle \mathbf{g}, \mathbf{x} - \hat{\mathbf{x}} \rangle$ для всех $\mathbf{x} \in \mathrm{dom}\,(f)$
- Возьмём $\mathbf{x} = \hat{\mathbf{x}} + \varepsilon \mathbf{d}$, где $\mathbf{d}$ такой что $\|\mathbf{d}\|_2 = 1$ и $\langle \mathbf{g}, \mathbf{d} \rangle = \|\mathbf{g}\|_2$
- Тогда $\varepsilon \|\mathbf{g}\|_2 \leq f(\mathbf{x}) - f(\hat{\mathbf{x}}) \leq L\|\mathbf{x} - \hat{\mathbf{x}}\|_2 = L\varepsilon$
- Значит $\mathbf{g} \in B_2(L, 0)$, то есть субдифференциал является ограниченным множеством

Замечание

Теорему существования можно обобщить с внутренности на относительную внутренность области определения f.

Связь между субградиентом и производной по направлению

Теорема

Пусть f выпуклая функция. Тогда для любой точки $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$ и любого направления $\mathbf{d}$ выполнено $f'(\mathbf{x}, \mathbf{d}) = \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$

Связь между субградиентом и производной по направлению

Теорема

Пусть f выпуклая функция. Тогда для любой точки $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ и любого направления $\mathbf{d}$ выполнено $f'(\mathbf{x}, \mathbf{d}) = \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$

Доказательство

Связь между субградиентом и производной по направлению

Теорема

Пусть f выпуклая функция. Тогда для любой точки $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$ и любого направления $\mathbf{d}$ выполнено $f'(\mathbf{x}, \mathbf{d}) = \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$

Доказательство

- Пусть $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$, тогда по определению субдифференциала верна следующая цепочка неравенств для некоторого $\mathbf{g} \in \partial f(\mathbf{x})$

$$f'(\mathbf{x}, \mathbf{d}) = \lim_{\alpha \to +0} \frac{1}{\alpha}(f(\mathbf{x} + \alpha \mathbf{d}) - f(\mathbf{x})) \geq \langle \mathbf{g}, \mathbf{d} \rangle$$

Связь между субградиентом и производной по направлению

Теорема

Пусть f выпуклая функция. Тогда для любой точки $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ и любого направления $\mathbf{d}$ выполнено $f'(\mathbf{x}, \mathbf{d}) = \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$

Доказательство

- Пусть $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, тогда по определению субдифференциала верна следующая цепочка неравенств для некоторого $\mathbf{g} \in \partial f(\mathbf{x})$

$$f'(\mathbf{x}, \mathbf{d}) = \lim_{\alpha \to +0} \frac{1}{\alpha}(f(\mathbf{x} + \alpha \mathbf{d}) - f(\mathbf{x})) \geq \langle \mathbf{g}, \mathbf{d} \rangle$$

- Поскольку это выполнено для произвольного $\mathbf{g} \in \partial f(\mathbf{x})$, то $f'(\mathbf{x}, \mathbf{d}) \geq \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$

- Покажем, что $f'(\mathbf{x}, \mathbf{d}) \leq \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$

- Покажем, что $f'(\mathbf{x}, \mathbf{d}) \leq \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$
- Рассмотрим функцию $h(\mathbf{w}) = f'(\mathbf{x}, \mathbf{w})$

- Покажем, что $f'(\mathbf{x}, \mathbf{d}) \leq \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$
- Рассмотрим функцию $h(\mathbf{w}) = f'(\mathbf{x}, \mathbf{w})$
- Эта функция выпукла и определена на $\mathbb{R}^n$ (см. прошлую лекцию)

- Покажем, что $f'(\mathbf{x}, \mathbf{d}) \leq \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$
- Рассмотрим функцию $h(\mathbf{w}) = f'(\mathbf{x}, \mathbf{w})$
- Эта функция выпукла и определена на $\mathbb{R}^n$ (см. прошлую лекцию)
- Значит в любой точке $\mathbf{d} \in \mathbb{R}^n$ субдифференциал не пуст

- Покажем, что $f'(\mathbf{x}, \mathbf{d}) \leq \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$
- Рассмотрим функцию $h(\mathbf{w}) = f'(\mathbf{x}, \mathbf{w})$
- Эта функция выпукла и определена на $\mathbb{R}^n$ (см. прошлую лекцию)
- Значит в любой точке $\mathbf{d} \in \mathbb{R}^n$ субдифференциал не пуст
- Пусть $\hat{\mathbf{g}} \in \partial h(\mathbf{d})$, тогда для любого $\mathbf{v}$ и $\alpha \geq 0$

$$\alpha f'(\mathbf{x}, \mathbf{v}) = f'(\mathbf{x}, \alpha\mathbf{v}) = h(\alpha\mathbf{v}) \geq h(\mathbf{d}) + \langle \hat{\mathbf{g}}, \alpha\mathbf{v} - \mathbf{d} \rangle$$

- Покажем, что $f'(\mathbf{x}, \mathbf{d}) \le \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$
- Рассмотрим функцию $h(\mathbf{w}) = f'(\mathbf{x}, \mathbf{w})$
- Эта функция выпукла и определена на $\mathbb{R}^n$ (см. прошлую лекцию)
- Значит в любой точке $\mathbf{d} \in \mathbb{R}^n$ субдифференциал не пуст
- Пусть $\hat{\mathbf{g}} \in \partial h(\mathbf{d})$, тогда для любого $\mathbf{v}$ и $\alpha \ge 0$

$$\alpha f'(\mathbf{x}, \mathbf{v}) = f'(\mathbf{x}, \alpha \mathbf{v}) = h(\alpha \mathbf{v}) \ge h(\mathbf{d}) + \langle \hat{\mathbf{g}}, \alpha \mathbf{v} - \mathbf{d} \rangle$$

- Тогда $\alpha(f'(\mathbf{x}, \mathbf{v}) - \langle \hat{\mathbf{g}}, \mathbf{v} \rangle) \ge f'(\mathbf{x}, \mathbf{d}) - \langle \hat{\mathbf{g}}, \mathbf{d} \rangle$

- Покажем, что $f'(\mathbf{x}, \mathbf{d}) \leq \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$
- Рассмотрим функцию $h(\mathbf{w}) = f'(\mathbf{x}, \mathbf{w})$
- Эта функция выпукла и определена на $\mathbb{R}^n$ (см. прошлую лекцию)
- Значит в любой точке $\mathbf{d} \in \mathbb{R}^n$ субдифференциал не пуст
- Пусть $\hat{\mathbf{g}} \in \partial h(\mathbf{d})$, тогда для любого $\mathbf{v}$ и $\alpha \geq 0$

$$\alpha f'(\mathbf{x}, \mathbf{v}) = f'(\mathbf{x}, \alpha\mathbf{v}) = h(\alpha\mathbf{v}) \geq h(\mathbf{d}) + \langle \hat{\mathbf{g}}, \alpha\mathbf{v} - \mathbf{d} \rangle$$

- Тогда $\alpha(f'(\mathbf{x}, \mathbf{v}) - \langle \hat{\mathbf{g}}, \mathbf{v} \rangle) \geq f'(\mathbf{x}, \mathbf{d}) - \langle \hat{\mathbf{g}}, \mathbf{d} \rangle$
- Так как это выполнено для любого $\alpha > 0$, то выражение слева неотрицательно. Иначе для достаточно большого α неравенство бы нарушалось. В итоге $f'(\mathbf{x}, \mathbf{v}) \geq \langle \hat{\mathbf{g}}, \mathbf{v} \rangle$

- Покажем, что $f'(\mathbf{x}, \mathbf{d}) \leq \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$
- Рассмотрим функцию $h(\mathbf{w}) = f'(\mathbf{x}, \mathbf{w})$
- Эта функция выпукла и определена на $\mathbb{R}^n$ (см. прошлую лекцию)
- Значит в любой точке $\mathbf{d} \in \mathbb{R}^n$ субдифференциал не пуст
- Пусть $\hat{\mathbf{g}} \in \partial h(\mathbf{d})$, тогда для любого $\mathbf{v}$ и $\alpha \geq 0$

$$\alpha f'(\mathbf{x}, \mathbf{v}) = f'(\mathbf{x}, \alpha \mathbf{v}) = h(\alpha \mathbf{v}) \geq h(\mathbf{d}) + \langle \hat{\mathbf{g}}, \alpha \mathbf{v} - \mathbf{d} \rangle$$

- Тогда $\alpha(f'(\mathbf{x}, \mathbf{v}) - \langle \hat{\mathbf{g}}, \mathbf{v} \rangle) \geq f'(\mathbf{x}, \mathbf{d}) - \langle \hat{\mathbf{g}}, \mathbf{d} \rangle$
- Так как это выполнено для любого $\alpha > 0$, то выражение слева неотрицательно. Иначе для достаточно большого α неравенство бы нарушалось. В итоге $f'(\mathbf{x}, \mathbf{v}) \geq \langle \hat{\mathbf{g}}, \mathbf{v} \rangle$
- Используем утверждение с прошлой лекции (слайд 15):

$$f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x}, \mathbf{y} - \mathbf{x}) \geq f(\mathbf{x}) + \langle \hat{\mathbf{g}}, \mathbf{y} - \mathbf{x} \rangle$$

- Покажем, что $f'(\mathbf{x}, \mathbf{d}) \leq \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$
- Рассмотрим функцию $h(\mathbf{w}) = f'(\mathbf{x}, \mathbf{w})$
- Эта функция выпукла и определена на $\mathbb{R}^n$ (см. прошлую лекцию)
- Значит в любой точке $\mathbf{d} \in \mathbb{R}^n$ субдифференциал не пуст
- Пусть $\hat{\mathbf{g}} \in \partial h(\mathbf{d})$, тогда для любого $\mathbf{v}$ и $\alpha \geq 0$

$$\alpha f'(\mathbf{x}, \mathbf{v}) = f'(\mathbf{x}, \alpha \mathbf{v}) = h(\alpha \mathbf{v}) \geq h(\mathbf{d}) + \langle \hat{\mathbf{g}}, \alpha \mathbf{v} - \mathbf{d} \rangle$$

- Тогда $\alpha(f'(\mathbf{x}, \mathbf{v}) - \langle \hat{\mathbf{g}}, \mathbf{v} \rangle) \geq f'(\mathbf{x}, \mathbf{d}) - \langle \hat{\mathbf{g}}, \mathbf{d} \rangle$
- Так как это выполнено для любого $\alpha > 0$, то выражение слева неотрицательно. Иначе для достаточно большого α неравенство бы нарушалось. В итоге $f'(\mathbf{x}, \mathbf{v}) \geq \langle \hat{\mathbf{g}}, \mathbf{v} \rangle$
- Используем утверждение с прошлой лекции (слайд 15):

$$f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x}, \mathbf{y} - \mathbf{x}) \geq f(\mathbf{x}) + \langle \hat{\mathbf{g}}, \mathbf{y} - \mathbf{x} \rangle$$

- Значит $\hat{\mathbf{g}} \in \partial f(\mathbf{x})$

- Покажем, что $f'(\mathbf{x}, \mathbf{d}) \le \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$
- Рассмотрим функцию $h(\mathbf{w}) = f'(\mathbf{x}, \mathbf{w})$
- Эта функция выпукла и определена на $\mathbb{R}^n$ (см. прошлую лекцию)
- Значит в любой точке $\mathbf{d} \in \mathbb{R}^n$ субдифференциал не пуст
- Пусть $\hat{\mathbf{g}} \in \partial h(\mathbf{d})$, тогда для любого $\mathbf{v}$ и $\alpha \ge 0$

$$\alpha f'(\mathbf{x}, \mathbf{v}) = f'(\mathbf{x}, \alpha \mathbf{v}) = h(\alpha \mathbf{v}) \ge h(\mathbf{d}) + \langle \hat{\mathbf{g}}, \alpha \mathbf{v} - \mathbf{d} \rangle$$

- Тогда $\alpha(f'(\mathbf{x}, \mathbf{v}) - \langle \hat{\mathbf{g}}, \mathbf{v} \rangle) \ge f'(\mathbf{x}, \mathbf{d}) - \langle \hat{\mathbf{g}}, \mathbf{d} \rangle$
- Так как это выполнено для любого $\alpha > 0$, то выражение слева неотрицательно. Иначе для достаточно большого α неравенство бы нарушалось. В итоге $f'(\mathbf{x}, \mathbf{v}) \ge \langle \hat{\mathbf{g}}, \mathbf{v} \rangle$
- Используем утверждение с прошлой лекции (слайд 15):

$$f(\mathbf{y}) \ge f(\mathbf{x}) + f'(\mathbf{x}, \mathbf{y} - \mathbf{x}) \ge f(\mathbf{x}) + \langle \hat{\mathbf{g}}, \mathbf{y} - \mathbf{x} \rangle$$

- Значит $\hat{\mathbf{g}} \in \partial f(\mathbf{x})$
- При $\alpha = 0$ получим $f'(\mathbf{x}, \mathbf{d}) \le \langle \hat{\mathbf{g}}, \mathbf{d} \rangle \le \max_{\mathbf{g} \in \partial f(\mathbf{x})} \langle \mathbf{g}, \mathbf{d} \rangle$

Опорная функция

Определение

Пусть $\mathcal{X}$ некоторое непустое множество. Тогда опорной функцией для этого множества называется функция

$$\sigma_{\mathcal{X}}(\mathbf{y}) = \max_{\mathbf{x} \in \mathcal{X}} \langle \mathbf{y}, \mathbf{x} \rangle.$$

Опорная функция

Определение

Пусть $\mathcal{X}$ некоторое непустое множество. Тогда опорной функцией для этого множества называется функция

$$\sigma_{\mathcal{X}}(\mathbf{y}) = \max_{\mathbf{x} \in \mathcal{X}} \langle \mathbf{y}, \mathbf{x} \rangle.$$

Свойства

Опорная функция

Определение

Пусть $\mathcal{X}$ некоторое непустое множество. Тогда опорной функцией для этого множества называется функция

$$\sigma_{\mathcal{X}}(\mathbf{y}) = \max_{\mathbf{x} \in \mathcal{X}} \langle \mathbf{y}, \mathbf{x} \rangle.$$

Свойства

- Опорная функция *любого* множества $\mathcal{X}$ является выпуклой и замкнутой (почему?)

Опорная функция

Определение

Пусть $\mathcal{X}$ некоторое непустое множество. Тогда опорной функцией для этого множества называется функция

$$\sigma_{\mathcal{X}}(\mathbf{y}) = \max_{\mathbf{x}\in\mathcal{X}}\langle \mathbf{y}, \mathbf{x}\rangle.$$

Свойства

- Опорная функция *любого* множества $\mathcal{X}$ является выпуклой и замкнутой (почему?)

Замечание

Предыдущая теорема может быть сформулирована в виде

$$f'(\mathbf{x}, \mathbf{d}) = \sigma_{\partial f(\mathbf{x})}(\mathbf{d})$$

Опорная функция полностью описывает выпуклое замкнутое множество

Теорема

Пусть $\mathcal{A}$ и $\mathcal{B}$ выпуклые замкнутые множества. Тогда $\mathcal{A} = \mathcal{B}$ iff $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$.

Опорная функция полностью описывает выпуклое замкнутое множество

Теорема

Пусть $\mathcal{A}$ и $\mathcal{B}$ выпуклые замкнутые множества. Тогда $\mathcal{A} = \mathcal{B}$ iff $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$.

Доказательство

Опорная функция полностью описывает выпуклое замкнутое множество

Теорема

Пусть $\mathcal{A}$ и $\mathcal{B}$ выпуклые замкнутые множества. Тогда $\mathcal{A} = \mathcal{B}$ iff $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$.

Доказательство

1. Пусть $\mathcal{A} = \mathcal{B}$, тогда очевидно $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$ так как максимум берётся по одному и тому же множеству

Опорная функция полностью описывает выпуклое замкнутое множество

Теорема

Пусть $\mathcal{A}$ и $\mathcal{B}$ выпуклые замкнутые множества. Тогда $\mathcal{A} = \mathcal{B}$ iff $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$.

Доказательство

1. Пусть $\mathcal{A} = \mathcal{B}$, тогда очевидно $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$ так как максимум берётся по одному и тому же множеству
2. Пусть $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$

Опорная функция полностью описывает выпуклое замкнутое множество

Теорема

Пусть $\mathcal{A}$ и $\mathcal{B}$ выпуклые замкнутые множества. Тогда $\mathcal{A} = \mathcal{B}$ iff $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$.

Доказательство

1. Пусть $\mathcal{A} = \mathcal{B}$, тогда очевидно $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$ так как максимум берётся по одному и тому же множеству
2. Пусть $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$
   - Предположим, что $\mathcal{A} \neq \mathcal{B}$, то есть найдётся $\mathbf{y} \in \mathcal{A}$ такой что $\mathbf{y} \notin \mathcal{B}$

Опорная функция полностью описывает выпуклое замкнутое множество

Теорема

Пусть $\mathcal{A}$ и $\mathcal{B}$ выпуклые замкнутые множества. Тогда $\mathcal{A} = \mathcal{B}$ iff $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$.

Доказательство

1. Пусть $\mathcal{A} = \mathcal{B}$, тогда очевидно $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$ так как максимум берётся по одному и тому же множеству
2. Пусть $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$
   - Предположим, что $\mathcal{A} \neq \mathcal{B}$, то есть найдётся $\mathbf{y} \in \mathcal{A}$ такой что $\mathbf{y} \notin \mathcal{B}$
   - Так как $\mathcal{B}$ выпуклое замкнутое множество и $\mathbf{y} \notin \mathcal{B}$, то они строго отделимы (лекция 2, слайд 20)

Опорная функция полностью описывает выпуклое замкнутое множество

Теорема

Пусть $\mathcal{A}$ и $\mathcal{B}$ выпуклые замкнутые множества. Тогда $\mathcal{A} = \mathcal{B}$ iff $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$.

Доказательство

1. Пусть $\mathcal{A} = \mathcal{B}$, тогда очевидно $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$ так как максимум берётся по одному и тому же множеству
2. Пусть $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$
   - Предположим, что $\mathcal{A} \neq \mathcal{B}$, то есть найдётся $\mathbf{y} \in \mathcal{A}$ такой что $\mathbf{y} \not\in \mathcal{B}$
   - Так как $\mathcal{B}$ выпуклое замкнутое множество и $\mathbf{y} \not\in \mathcal{B}$, то они строго отделимы (лекция 2, слайд 20)
   - Значит существует $\mathbf{p} \neq 0$ и $\alpha > 0$ такие то $\langle \mathbf{p}, \mathbf{x} \rangle \leq \alpha < \langle \mathbf{p}, \mathbf{y} \rangle$ для всех $\mathbf{x} \in \mathcal{B}$

Опорная функция полностью описывает выпуклое замкнутое множество

Теорема

Пусть $\mathcal{A}$ и $\mathcal{B}$ выпуклые замкнутые множества. Тогда $\mathcal{A} = \mathcal{B}$ iff $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$.

Доказательство

1. Пусть $\mathcal{A} = \mathcal{B}$, тогда очевидно $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$ так как максимум берётся по одному и тому же множеству
2. Пусть $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$
   - Предположим, что $\mathcal{A} \neq \mathcal{B}$, то есть найдётся $\mathbf{y} \in \mathcal{A}$ такой что $\mathbf{y} \notin \mathcal{B}$
   - Так как $\mathcal{B}$ выпуклое замкнутое множество и $\mathbf{y} \notin \mathcal{B}$, то они строго отделимы (лекция 2, слайд 20)
   - Значит существует $\mathbf{p} \neq 0$ и $\alpha > 0$ такие то $\langle \mathbf{p}, \mathbf{x} \rangle \leq \alpha < \langle \mathbf{p}, \mathbf{y} \rangle$ для всех $\mathbf{x} \in \mathcal{B}$
   - Возьмём максимум от обеих частей по $\mathbf{x} \in \mathcal{B}$, тогда $\sigma_{\mathcal{B}}(\mathbf{p}) \leq \alpha < \langle \mathbf{p}, \mathbf{y} \rangle \leq \sigma_{\mathcal{A}}(\mathbf{p})$

Опорная функция полностью описывает выпуклое замкнутое множество

Теорема

Пусть $\mathcal{A}$ и $\mathcal{B}$ выпуклые замкнутые множества. Тогда $\mathcal{A} = \mathcal{B}$ iff $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$.

Доказательство

1. Пусть $\mathcal{A} = \mathcal{B}$, тогда очевидно $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$ так как максимум берётся по одному и тому же множеству
2. Пусть $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$
   - Предположим, что $\mathcal{A} \neq \mathcal{B}$, то есть найдётся $\mathbf{y} \in \mathcal{A}$ такой что $\mathbf{y} \notin \mathcal{B}$
   - Так как $\mathcal{B}$ выпуклое замкнутое множество и $\mathbf{y} \notin \mathcal{B}$, то они строго отделимы (лекция 2, слайд 20)
   - Значит существует $\mathbf{p} \neq 0$ и $\alpha > 0$ такие то $\langle \mathbf{p}, \mathbf{x} \rangle \leq \alpha < \langle \mathbf{p}, \mathbf{y} \rangle$ для всех $\mathbf{x} \in \mathcal{B}$
   - Возьмём максимум от обеих частей по $\mathbf{x} \in \mathcal{B}$, тогда $\sigma_{\mathcal{B}}(\mathbf{p}) \leq \alpha < \langle \mathbf{p}, \mathbf{y} \rangle \leq \sigma_{\mathcal{A}}(\mathbf{p})$
   - Получили противоречие с тем, что $\sigma_{\mathcal{A}} = \sigma_{\mathcal{B}}$

Чему равен субдифференциал дифференцируемой функции?

Теорема

Пусть f выпуклая функция и $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Если f дифференцируема в $\mathbf{x}$, то $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$. Если f имеет единственный субградиент в $\mathbf{x}$, то она дифференцируема в $\mathbf{x}$ и $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$

Чему равен субдифференциал дифференцируемой функции?

Теорема

Пусть f выпуклая функция и $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Если f дифференцируема в $\mathbf{x}$, то $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$. Если f имеет единственный субградиент в $\mathbf{x}$, то она дифференцируема в $\mathbf{x}$ и $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$

Доказательство

Чему равен субдифференциал дифференцируемой функции?

Теорема

Пусть f выпуклая функция и $\mathbf{x} \in \text{int}\,(\text{dom}\,(f))$. Если f дифференцируема в $\mathbf{x}$, то $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$. Если f имеет единственный субградиент в $\mathbf{x}$, то она дифференцируема в $\mathbf{x}$ и $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$

Доказательство

- Пусть $\mathbf{x} \in \text{int}\,(\text{dom}\,(f))$ и f дифференцируема в $\mathbf{x}$, тогда для любого направления $\mathbf{d}$ выполнено $f'(\mathbf{x}, \mathbf{d}) = \langle f'(\mathbf{x}), \mathbf{d} \rangle$

Чему равен субдифференциал дифференцируемой функции?

Теорема

Пусть f выпуклая функция и $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Если f дифференцируема в $\mathbf{x}$, то $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$. Если f имеет единственный субградиент в $\mathbf{x}$, то она дифференцируема в $\mathbf{x}$ и $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$

Доказательство

- Пусть $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ и f дифференцируема в $\mathbf{x}$, тогда для любого направления $\mathbf{d}$ выполнено $f'(\mathbf{x}, \mathbf{d}) = \langle f'(\mathbf{x}), \mathbf{d} \rangle$
- Пусть $\mathbf{g} \in \partial f(\mathbf{x}) \neq \varnothing$ так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$

Чему равен субдифференциал дифференцируемой функции?

Теорема

Пусть f выпуклая функция и $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Если f дифференцируема в $\mathbf{x}$, то $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$. Если f имеет единственный субградиент в $\mathbf{x}$, то она дифференцируема в $\mathbf{x}$ и $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$

Доказательство

- Пусть $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ и f дифференцируема в $\mathbf{x}$, тогда для любого направления $\mathbf{d}$ выполнено $f'(\mathbf{x}, \mathbf{d}) = \langle f'(\mathbf{x}), \mathbf{d} \rangle$
- Пусть $\mathbf{g} \in \partial f(\mathbf{x}) \neq \varnothing$ так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$
- Используем связь между производной по направлению и субдифференциалом:
  $\langle f'(\mathbf{x}), \mathbf{d} \rangle = f'(\mathbf{x}, \mathbf{d}) = \max_{\hat{\mathbf{g}} \in \partial f(\mathbf{x})} \langle \hat{\mathbf{g}}, \mathbf{d} \rangle \geq \langle \mathbf{g}, \mathbf{d} \rangle$

Чему равен субдифференциал дифференцируемой функции?

Теорема

Пусть f выпуклая функция и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$. Если f дифференцируема в $\mathbf{x}$, то $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$. Если f имеет единственный субградиент в $\mathbf{x}$, то она дифференцируема в $\mathbf{x}$ и $\partial f(\mathbf{x}) = \{f'(\mathbf{x})\}$

Доказательство

- Пусть $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$ и f дифференцируема в $\mathbf{x}$, тогда для любого направления $\mathbf{d}$ выполнено $f'(\mathbf{x}, \mathbf{d}) = \langle f'(\mathbf{x}), \mathbf{d} \rangle$
- Пусть $\mathbf{g} \in \partial f(\mathbf{x}) \neq \varnothing$ так как $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$
- Используем связь между производной по направлению и субдифференциалом: $\langle f'(\mathbf{x}), \mathbf{d} \rangle = f'(\mathbf{x}, \mathbf{d}) = \max_{\hat{\mathbf{g}} \in \partial f(\mathbf{x})} \langle \hat{\mathbf{g}}, \mathbf{d} \rangle \geq \langle \mathbf{g}, \mathbf{d} \rangle$
- Тогда $\langle \mathbf{g} - f'(\mathbf{x}), \mathbf{d} \rangle \leq 0$ и, взяв максимум по $\mathbf{d}$ таким что $\|\mathbf{d}\|_2 \leq 1$, получим $\|\mathbf{g} - f'(\mathbf{x})\|_2 \leq 0$, значит $\mathbf{g} = f'(\mathbf{x})$.

Если в точке субдифференциал состоит из одного элемента $\mathbf{g}$...

- Пусть $\mathcal{U}_\alpha = \mathbf{x} + \alpha\mathcal{B}_2(1)$, где $\mathcal{B}_2(1)$ единичный шар в евклидовой норме с центром в нуле

Если в точке субдифференциал состоит из одного элемента $\mathbf{g}$...

- Пусть $\mathcal{U}_\alpha = \mathbf{x} + \alpha\mathcal{B}_2(1)$, где $\mathcal{B}_2(1)$ единичный шар в евклидовой норме с центром в нуле
- Так как $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$, то найдётся $0 < \bar{\alpha} \leq 1$ что $\mathcal{U}_{\bar{\alpha}} \subset \operatorname{int}(\operatorname{dom}(f))$

Если в точке субдифференциал состоит из одного элемента $\mathbf{g}$...

- Пусть $\mathcal{U}_\alpha = \mathbf{x} + \alpha\mathcal{B}_2(1)$, где $\mathcal{B}_2(1)$ единичный шар в евклидовой норме с центром в нуле
- Так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то найдётся $0 < \bar{\alpha} \leq 1$ что $\mathcal{U}_{\bar{\alpha}} \subset \mathrm{int}\,(\mathrm{dom}\,(f))$
- Рассмотрим функцию $\varphi(\alpha, \mathbf{s}) = \frac{f(\mathbf{x}+\alpha\mathbf{s})-f(\mathbf{x})}{\alpha} - \langle \mathbf{g}, \mathbf{s} \rangle$ для $\mathbf{s} \in \mathcal{B}_2(1)$ и $0 < \alpha \leq \bar{\alpha}$

Если в точке субдифференциал состоит из одного элемента $\mathbf{g}$...

- Пусть $\mathcal{U}_\alpha = \mathbf{x} + \alpha\mathcal{B}_2(1)$, где $\mathcal{B}_2(1)$ единичный шар в евклидовой норме с центром в нуле
- Так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то найдётся $0 < \bar{\alpha} \leq 1$ что $\mathcal{U}_{\bar{\alpha}} \subset \mathrm{int}\,(\mathrm{dom}\,(f))$
- Рассмотрим функцию $\varphi(\alpha, \mathbf{s}) = \frac{f(\mathbf{x}+\alpha\mathbf{s}) - f(\mathbf{x})}{\alpha} - \langle \mathbf{g}, \mathbf{s} \rangle$ для $\mathbf{s} \in \mathcal{B}_2(1)$ и $0 < \alpha \leq \bar{\alpha}$
- Так как $\mathbf{g}$ единственный элемент $\partial f(\mathbf{x})$, то $f'(\mathbf{x}, \mathbf{s}) = \langle \mathbf{g}, \mathbf{s} \rangle$

Если в точке субдифференциал состоит из одного элемента $\mathbf{g}$...

- Пусть $\mathcal{U}_\alpha = \mathbf{x} + \alpha\mathcal{B}_2(1)$, где $\mathcal{B}_2(1)$ единичный шар в евклидовой норме с центром в нуле
- Так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то найдётся $0 < \bar{\alpha} \le 1$ что $\mathcal{U}_{\bar{\alpha}} \subset \mathrm{int}\,(\mathrm{dom}\,(f))$
- Рассмотрим функцию $\varphi(\alpha, \mathbf{s}) = \frac{f(\mathbf{x}+\alpha\mathbf{s}) - f(\mathbf{x})}{\alpha} - \langle \mathbf{g}, \mathbf{s} \rangle$ для $\mathbf{s} \in \mathcal{B}_2(1)$ и $0 < \alpha \le \bar{\alpha}$
- Так как $\mathbf{g}$ единственный элемент $\partial f(\mathbf{x})$, то $f'(\mathbf{x}, \mathbf{s}) = \langle \mathbf{g}, \mathbf{s} \rangle$
- Значит при $\alpha \to +0$ существует предел $\varphi(\alpha, \mathbf{s}) \to 0$

Если в точке субдифференциал состоит из одного элемента $\mathbf{g}$...

- Пусть $\mathcal{U}_\alpha = \mathbf{x} + \alpha\mathcal{B}_2(1)$, где $\mathcal{B}_2(1)$ единичный шар в евклидовой норме с центром в нуле
- Так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то найдётся $0 < \bar{\alpha} \leq 1$ что $\mathcal{U}_{\bar{\alpha}} \subset \mathrm{int}\,(\mathrm{dom}\,(f))$
- Рассмотрим функцию $\varphi(\alpha, \mathbf{s}) = \frac{f(\mathbf{x}+\alpha\mathbf{s}) - f(\mathbf{x})}{\alpha} - \langle \mathbf{g}, \mathbf{s} \rangle$ для $\mathbf{s} \in \mathcal{B}_2(1)$ и $0 < \alpha \leq \bar{\alpha}$
- Так как $\mathbf{g}$ единственный элемент $\partial f(\mathbf{x})$, то $f'(\mathbf{x}, \mathbf{s}) = \langle \mathbf{g}, \mathbf{s} \rangle$
- Значит при $\alpha \to +0$ существует предел $\varphi(\alpha, \mathbf{s}) \to 0$
- Также $\varphi(\alpha, \mathbf{s})$ непрерывна по $\mathbf{s}$ на компакте $\mathcal{B}_2(1)$

Если в точке субдифференциал состоит из одного элемента $\mathbf{g}$...

- Пусть $\mathcal{U}_\alpha = \mathbf{x} + \alpha \mathcal{B}_2(1)$, где $\mathcal{B}_2(1)$ единичный шар в евклидовой норме с центром в нуле
- Так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то найдётся $0 < \bar{\alpha} \leq 1$ что $\mathcal{U}_{\bar{\alpha}} \subset \mathrm{int}\,(\mathrm{dom}\,(f))$
- Рассмотрим функцию $\varphi(\alpha, \mathbf{s}) = \frac{f(\mathbf{x}+\alpha\mathbf{s}) - f(\mathbf{x})}{\alpha} - \langle \mathbf{g}, \mathbf{s} \rangle$ для $\mathbf{s} \in \mathcal{B}_2(1)$ и $0 < \alpha \leq \bar{\alpha}$
- Так как $\mathbf{g}$ единственный элемент $\partial f(\mathbf{x})$, то $f'(\mathbf{x}, \mathbf{s}) = \langle \mathbf{g}, \mathbf{s} \rangle$
- Значит при $\alpha \to +0$ существует предел $\varphi(\alpha, \mathbf{s}) \to 0$
- Также $\varphi(\alpha, \mathbf{s})$ непрерывна по $\mathbf{s}$ на компакте $\mathcal{B}_2(1)$
- Значит сходимость при $\alpha \to +0$ равномерна по $\mathbf{s}$, то есть для любого $\varepsilon > 0$ найдётся $0 < \hat{\alpha} \leq \bar{\alpha}$ такое что $0 \leq \varphi(\alpha, \mathbf{s}) \leq \varepsilon$ для $0 \leq \alpha \leq \hat{\alpha}$ и всех $\mathbf{s} \in \mathcal{B}_2(1)$

Если в точке субдифференциал состоит из одного элемента $\mathbf{g}$...

- Пусть $\mathcal{U}_\alpha = \mathbf{x} + \alpha \mathcal{B}_2(1)$, где $\mathcal{B}_2(1)$ единичный шар в евклидовой норме с центром в нуле
- Так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то найдётся $0 < \bar{\alpha} \leq 1$ что $\mathcal{U}_{\bar{\alpha}} \subset \mathrm{int}\,(\mathrm{dom}\,(f))$
- Рассмотрим функцию $\varphi(\alpha, \mathbf{s}) = \frac{f(\mathbf{x}+\alpha\mathbf{s}) - f(\mathbf{x})}{\alpha} - \langle \mathbf{g}, \mathbf{s} \rangle$ для $\mathbf{s} \in \mathcal{B}_2(1)$ и $0 < \alpha \leq \bar{\alpha}$
- Так как $\mathbf{g}$ единственный элемент $\partial f(\mathbf{x})$, то $f'(\mathbf{x}, \mathbf{s}) = \langle \mathbf{g}, \mathbf{s} \rangle$
- Значит при $\alpha \to +0$ существует предел $\varphi(\alpha, \mathbf{s}) \to 0$
- Также $\varphi(\alpha, \mathbf{s})$ непрерывна по $\mathbf{s}$ на компакте $\mathcal{B}_2(1)$
- Значит сходимость при $\alpha \to +0$ равномерна по $\mathbf{s}$, то есть для любого $\varepsilon > 0$ найдётся $0 < \hat{\alpha} \leq \bar{\alpha}$ такое что $0 \leq \varphi(\alpha, \mathbf{s}) \leq \varepsilon$ для $0 \leq \alpha \leq \hat{\alpha}$ и всех $\mathbf{s} \in \mathcal{B}_2(1)$
- Выберем $\mathbf{s} \in \mathcal{B}_2(1)$ такой что $\|\mathbf{s}\|_2 \leq \hat{\alpha}^2$

Если в точке субдифференциал состоит из одного элемента $\mathbf{g}$...

- Пусть $\mathcal{U}_\alpha = \mathbf{x} + \alpha \mathcal{B}_2(1)$, где $\mathcal{B}_2(1)$ единичный шар в евклидовой норме с центром в нуле
- Так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то найдётся $0 < \bar{\alpha} \leq 1$ что $\mathcal{U}_{\bar{\alpha}} \subset \mathrm{int}\,(\mathrm{dom}\,(f))$
- Рассмотрим функцию $\varphi(\alpha, \mathbf{s}) = \frac{f(\mathbf{x}+\alpha\mathbf{s}) - f(\mathbf{x})}{\alpha} - \langle \mathbf{g}, \mathbf{s} \rangle$ для $\mathbf{s} \in \mathcal{B}_2(1)$ и $0 < \alpha \leq \bar{\alpha}$
- Так как $\mathbf{g}$ единственный элемент $\partial f(\mathbf{x})$, то $f'(\mathbf{x}, \mathbf{s}) = \langle \mathbf{g}, \mathbf{s} \rangle$
- Значит при $\alpha \to +0$ существует предел $\varphi(\alpha, \mathbf{s}) \to 0$
- Также $\varphi(\alpha, \mathbf{s})$ непрерывна по $\mathbf{s}$ на компакте $\mathcal{B}_2(1)$
- Значит сходимость при $\alpha \to +0$ равномерна по $\mathbf{s}$, то есть для любого $\varepsilon > 0$ найдётся $0 < \hat{\alpha} \leq \bar{\alpha}$ такое что $0 \leq \varphi(\alpha, \mathbf{s}) \leq \varepsilon$ для $0 \leq \alpha \leq \hat{\alpha}$ и всех $\mathbf{s} \in \mathcal{B}_2(1)$
- Выберем $\mathbf{s} \in \mathcal{B}_2(1)$ такой что $\|\mathbf{s}\|_2 \leq \hat{\alpha}^2$
- Обозначим $\beta = \|\mathbf{s}\|_2 / \hat{\alpha}$ и $0 \leq \beta \leq \hat{\alpha}$

Если в точке субдифференциал состоит из одного элемента $\mathbf{g}$...

- Пусть $\mathcal{U}_\alpha = \mathbf{x} + \alpha\mathcal{B}_2(1)$, где $\mathcal{B}_2(1)$ единичный шар в евклидовой норме с центром в нуле
- Так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то найдётся $0 < \bar{\alpha} \leq 1$ что $\mathcal{U}_{\bar{\alpha}} \subset \mathrm{int}\,(\mathrm{dom}\,(f))$
- Рассмотрим функцию $\varphi(\alpha, \mathbf{s}) = \frac{f(\mathbf{x}+\alpha\mathbf{s}) - f(\mathbf{x})}{\alpha} - \langle \mathbf{g}, \mathbf{s} \rangle$ для $\mathbf{s} \in \mathcal{B}_2(1)$ и $0 < \alpha \leq \bar{\alpha}$
- Так как $\mathbf{g}$ единственный элемент $\partial f(\mathbf{x})$, то $f'(\mathbf{x}, \mathbf{s}) = \langle \mathbf{g}, \mathbf{s} \rangle$
- Значит при $\alpha \to +0$ существует предел $\varphi(\alpha, \mathbf{s}) \to 0$
- Также $\varphi(\alpha, \mathbf{s})$ непрерывна по $\mathbf{s}$ на компакте $\mathcal{B}_2(1)$
- Значит сходимость при $\alpha \to +0$ равномерна по $\mathbf{s}$, то есть для любого $\varepsilon > 0$ найдётся $0 < \hat{\alpha} \leq \bar{\alpha}$ такое что $0 \leq \varphi(\alpha, \mathbf{s}) \leq \varepsilon$ для $0 \leq \alpha \leq \hat{\alpha}$ и всех $\mathbf{s} \in \mathcal{B}_2(1)$
- Выберем $\mathbf{s} \in \mathcal{B}_2(1)$ такой что $\|\mathbf{s}\|_2 \leq \hat{\alpha}^2$
- Обозначим $\beta = \|\mathbf{s}\|_2 / \hat{\alpha}$ и $0 \leq \beta \leq \hat{\alpha}$
- Также пусть $\mathbf{h} = \hat{\alpha}\frac{\mathbf{s}}{\|\mathbf{s}\|_2}$ и $\|\mathbf{h}\|_2 = \hat{\alpha} \leq 1$

Если в точке субдифференциал состоит из одного элемента $\mathbf{g}$...

- Пусть $\mathcal{U}_\alpha = \mathbf{x} + \alpha \mathcal{B}_2(1)$, где $\mathcal{B}_2(1)$ единичный шар в евклидовой норме с центром в нуле
- Так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то найдётся $0 < \bar{\alpha} \leq 1$ что $\mathcal{U}_{\bar{\alpha}} \subset \mathrm{int}\,(\mathrm{dom}\,(f))$
- Рассмотрим функцию $\varphi(\alpha, \mathbf{s}) = \frac{f(\mathbf{x}+\alpha\mathbf{s}) - f(\mathbf{x})}{\alpha} - \langle \mathbf{g}, \mathbf{s} \rangle$ для $\mathbf{s} \in \mathcal{B}_2(1)$ и $0 < \alpha \leq \bar{\alpha}$
- Так как $\mathbf{g}$ единственный элемент $\partial f(\mathbf{x})$, то $f'(\mathbf{x}, \mathbf{s}) = \langle \mathbf{g}, \mathbf{s} \rangle$
- Значит при $\alpha \to +0$ существует предел $\varphi(\alpha, \mathbf{s}) \to 0$
- Также $\varphi(\alpha, \mathbf{s})$ непрерывна по $\mathbf{s}$ на компакте $\mathcal{B}_2(1)$
- Значит сходимость при $\alpha \to +0$ равномерна по $\mathbf{s}$, то есть для любого $\varepsilon > 0$ найдётся $0 < \hat{\alpha} \leq \bar{\alpha}$ такое что $0 \leq \varphi(\alpha, \mathbf{s}) \leq \varepsilon$ для $0 \leq \alpha \leq \hat{\alpha}$ и всех $\mathbf{s} \in \mathcal{B}_2(1)$
- Выберем $\mathbf{s} \in \mathcal{B}_2(1)$ такой что $\|\mathbf{s}\|_2 \leq \hat{\alpha}^2$
- Обозначим $\beta = \|\mathbf{s}\|_2 / \hat{\alpha}$ и $0 \leq \beta \leq \hat{\alpha}$
- Также пусть $\mathbf{h} = \hat{\alpha}\frac{\mathbf{s}}{\|\mathbf{s}\|_2}$ и $\|\mathbf{h}\|_2 = \hat{\alpha} \leq 1$
- Тогда $0 \leq \varphi(\beta, \mathbf{h}) \leq \varepsilon$

- Подставим значения $\beta, \mathbf{h}$:
$\varphi(\beta, \mathbf{h}) = \frac{1}{\beta}(f(\mathbf{x} + \beta\mathbf{h}) - f(\mathbf{x})) - \langle \mathbf{g}, \mathbf{h} \rangle =$
$\frac{\hat{\alpha}}{\|\mathbf{s}\|_2}(f(\mathbf{x}+\mathbf{s}) - f(\mathbf{x})) - \langle \mathbf{g}, \mathbf{h} \rangle = \frac{\hat{\alpha}}{\|\mathbf{s}\|_2}(f(\mathbf{x}+\mathbf{s}) - f(\mathbf{x}) - \langle \mathbf{g}, \mathbf{s} \rangle)$

- Подставим значения $\beta, \mathbf{h}$:
  $\varphi(\beta, \mathbf{h}) = \frac{1}{\beta}(f(\mathbf{x} + \beta\mathbf{h}) - f(\mathbf{x})) - \langle \mathbf{g}, \mathbf{h} \rangle =$
  $\frac{\hat{\alpha}}{\|\mathbf{s}\|_2}(f(\mathbf{x}+\mathbf{s}) - f(\mathbf{x})) - \langle \mathbf{g}, \mathbf{h} \rangle = \frac{\hat{\alpha}}{\|\mathbf{s}\|_2}(f(\mathbf{x}+\mathbf{s}) - f(\mathbf{x}) - \langle \mathbf{g}, \mathbf{s} \rangle)$
- Значит $\frac{1}{\|\mathbf{s}\|_2}(f(\mathbf{x} + \mathbf{s}) - f(\mathbf{x}) - \langle \mathbf{g}, \mathbf{s} \rangle) \leq \frac{\varepsilon}{\hat{\alpha}}$

- Подставим значения $\beta, \mathbf{h}$:
  $\varphi(\beta, \mathbf{h}) = \frac{1}{\beta}(f(\mathbf{x} + \beta\mathbf{h}) - f(\mathbf{x})) - \langle \mathbf{g}, \mathbf{h} \rangle =$
  $\frac{\hat{\alpha}}{\|\mathbf{s}\|_2}(f(\mathbf{x}+\mathbf{s}) - f(\mathbf{x})) - \langle \mathbf{g}, \mathbf{h} \rangle = \frac{\hat{\alpha}}{\|\mathbf{s}\|_2}(f(\mathbf{x}+\mathbf{s}) - f(\mathbf{x}) - \langle \mathbf{g}, \mathbf{s} \rangle)$
- Значит $\frac{1}{\|\mathbf{s}\|_2}(f(\mathbf{x} + \mathbf{s}) - f(\mathbf{x}) - \langle \mathbf{g}, \mathbf{s} \rangle) \leq \frac{\varepsilon}{\hat{\alpha}}$
- Получаем, что существует предел

$$\lim_{\mathbf{s} \to 0} \frac{f(\mathbf{x} + \mathbf{s}) - f(\mathbf{x}) - \langle \mathbf{g}, \mathbf{s} \rangle}{\|\mathbf{s}\|_2} = 0$$

  и функция f дифференцируема в $\mathbf{x}$

Основные операции над функциями

- Умножение на число

Основные операции над функциями

- Умножение на число
- Сложение

Основные операции над функциями

- Умножение на число
- Сложение
- Взятие максимума

Основные операции над функциями

- Умножение на число
- Сложение
- Взятие максимума

Упражнение

Докажите, что $\partial(\alpha f)(\mathbf{x}) = \alpha \partial f(\mathbf{x})$

Субдифференциал суммы выпуклых функций

Теорема

Пусть f_1 и f_2 выпуклые функции и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f_1)) \cap \operatorname{int}(\operatorname{dom}(f_2))$. Тогда выполнено следующее равенство $\partial(f_1 + f_2)(\mathbf{x}) = \partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x})$.

Субдифференциал суммы выпуклых функций

Теорема

Пусть f_1 и f_2 выпуклые функции и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f_1)) \cap \operatorname{int}(\operatorname{dom}(f_2))$. Тогда выполнено следующее равенство $\partial(f_1 + f_2)(\mathbf{x}) = \partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x})$.

Доказательство

Субдифференциал суммы выпуклых функций

Теорема

Пусть f_1 и f_2 выпуклые функции и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f_1)) \cap \operatorname{int}(\operatorname{dom}(f_2))$. Тогда выполнено следующее равенство $\partial(f_1 + f_2)(\mathbf{x}) = \partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x})$.

Доказательство

- Обозначим $f \equiv f_1 + f_2$

Субдифференциал суммы выпуклых функций

Теорема

Пусть f_1 и f_2 выпуклые функции и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f_1)) \cap \operatorname{int}(\operatorname{dom}(f_2))$. Тогда выполнено следующее равенство $\partial(f_1 + f_2)(\mathbf{x}) = \partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x})$.

Доказательство

- Обозначим $f \equiv f_1 + f_2$
- Так как $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$, то $\sigma_{\partial f(\mathbf{x})}(\mathbf{d}) = f'(\mathbf{x}, \mathbf{d})$

Субдифференциал суммы выпуклых функций

Теорема

Пусть f_1 и f_2 выпуклые функции и $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f_1)) \cap \mathrm{int}\,(\mathrm{dom}\,(f_2))$. Тогда выполнено следующее равенство $\partial(f_1 + f_2)(\mathbf{x}) = \partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x})$.

Доказательство

- Обозначим $f \equiv f_1 + f_2$
- Так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то $\sigma_{\partial f(\mathbf{x})}(\mathbf{d}) = f'(\mathbf{x}, \mathbf{d})$
- Тогда в силу аддитивности производной по направлению $\sigma_{\partial f(\mathbf{x})}(\mathbf{d}) = f'(\mathbf{x}, \mathbf{d}) = f_1'(\mathbf{x}, \mathbf{d}) + f_2'(\mathbf{x}, \mathbf{d}) = \max\limits_{\mathbf{g}_1 \in \partial f_1(\mathbf{x})} \langle \mathbf{g}_1, \mathbf{d} \rangle + \max\limits_{\mathbf{g}_2 \in \partial f_2(\mathbf{x})} \langle \mathbf{g}_2, \mathbf{d} \rangle = \max\limits_{\mathbf{g}_1 \in \partial f_1(\mathbf{x}), \mathbf{g}_2 \in \partial f_2(\mathbf{x})} \langle \mathbf{g}_1 + \mathbf{g}_2, \mathbf{d} \rangle = \sigma_{\partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x})}(\mathbf{d})$

Субдифференциал суммы выпуклых функций

Теорема

Пусть f_1 и f_2 выпуклые функции и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f_1)) \cap \operatorname{int}(\operatorname{dom}(f_2))$. Тогда выполнено следующее равенство $\partial(f_1 + f_2)(\mathbf{x}) = \partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x})$.

Доказательство

- Обозначим $f \equiv f_1 + f_2$
- Так как $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$, то $\sigma_{\partial f(\mathbf{x})}(\mathbf{d}) = f'(\mathbf{x}, \mathbf{d})$
- Тогда в силу аддитивности производной по направлению $\sigma_{\partial f(\mathbf{x})}(\mathbf{d}) = f'(\mathbf{x}, \mathbf{d}) = f_1'(\mathbf{x}, \mathbf{d}) + f_2'(\mathbf{x}, \mathbf{d}) = \max\limits_{\mathbf{g}_1 \in \partial f_1(\mathbf{x})} \langle \mathbf{g}_1, \mathbf{d} \rangle + \max\limits_{\mathbf{g}_2 \in \partial f_2(\mathbf{x})} \langle \mathbf{g}_2, \mathbf{d} \rangle = \max\limits_{\mathbf{g}_1 \in \partial f_1(\mathbf{x}), \mathbf{g}_2 \in \partial f_2(\mathbf{x})} \langle \mathbf{g}_1 + \mathbf{g}_2, \mathbf{d} \rangle = \sigma_{\partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x})}(\mathbf{d})$
- Так как $\partial f_1(\mathbf{x})$ и $\partial f_2(\mathbf{x})$ выпуклые компакты, то и их сумма аналогично (почему?)

Субдифференциал суммы выпуклых функций

Теорема

Пусть f_1 и f_2 выпуклые функции и $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f_1)) \cap \mathrm{int}\,(\mathrm{dom}\,(f_2))$. Тогда выполнено следующее равенство $\partial(f_1 + f_2)(\mathbf{x}) = \partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x})$.

Доказательство

- Обозначим $f \equiv f_1 + f_2$
- Так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то $\sigma_{\partial f(\mathbf{x})}(\mathbf{d}) = f'(\mathbf{x}, \mathbf{d})$
- Тогда в силу аддитивности производной по направлению $\sigma_{\partial f(\mathbf{x})}(\mathbf{d}) = f'(\mathbf{x}, \mathbf{d}) = f_1'(\mathbf{x}, \mathbf{d}) + f_2'(\mathbf{x}, \mathbf{d}) = \max\limits_{\mathbf{g}_1 \in \partial f_1(\mathbf{x})} \langle \mathbf{g}_1, \mathbf{d} \rangle + \max\limits_{\mathbf{g}_2 \in \partial f_2(\mathbf{x})} \langle \mathbf{g}_2, \mathbf{d} \rangle = \max\limits_{\mathbf{g}_1 \in \partial f_1(\mathbf{x}), \mathbf{g}_2 \in \partial f_2(\mathbf{x})} \langle \mathbf{g}_1 + \mathbf{g}_2, \mathbf{d} \rangle = \sigma_{\partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x})}(\mathbf{d})$
- Так как $\partial f_1(\mathbf{x})$ и $\partial f_2(\mathbf{x})$ выпуклые компакты, то и их сумма аналогично (почему?)
- Если равны опорные функции, то совпадают и множества

Производная по направлению для максимума

Теорема

Пусть $f(\mathbf{x}) = \max_{i=1,\dots,k} f_i(\mathbf{x})$, где f_i такие функции, что для них существует производная по направлению $\mathbf{d}$. Тогда для $\mathbf{x} \in \bigcap\limits_{i=1}^{k} \operatorname{int}(\operatorname{dom}(f_i))$ выполнено

$$f'(\mathbf{x}, \mathbf{d}) = \max_{i \in \mathcal{I}(\mathbf{x})} f_i'(\mathbf{x}, \mathbf{d}),$$

где $\mathcal{I}(\mathbf{x}) = \{i \mid f_i(\mathbf{x}) = f(\mathbf{x})\}$.

Производная по направлению для максимума

Теорема

Пусть $f(\mathbf{x}) = \max_{i=1,\dots,k} f_i(\mathbf{x})$, где f_i такие функции, что для них существует производная по направлению $\mathbf{d}$. Тогда для $\mathbf{x} \in \bigcap\limits_{i=1}^{k} \operatorname{int}(\operatorname{dom}(f_i))$ выполнено

$$f'(\mathbf{x}, \mathbf{d}) = \max_{i \in \mathcal{I}(\mathbf{x})} f_i'(\mathbf{x}, \mathbf{d}),$$

где $\mathcal{I}(\mathbf{x}) = \{i \mid f_i(\mathbf{x}) = f(\mathbf{x})\}$.

Доказательство

Производная по направлению для максимума

Теорема

Пусть $f(\mathbf{x}) = \max_{i=1,\dots,k} f_i(\mathbf{x})$, где f_i такие функции, что для них существует производная по направлению $\mathbf{d}$. Тогда для $\mathbf{x} \in \bigcap\limits_{i=1}^{k} \operatorname{int}\left(\operatorname{dom}\left(f_i\right)\right)$ выполнено

$$f'(\mathbf{x}, \mathbf{d}) = \max_{i \in \mathcal{I}(\mathbf{x})} f_i'(\mathbf{x}, \mathbf{d}),$$

где $\mathcal{I}(\mathbf{x}) = \{i \mid f_i(\mathbf{x}) = f(\mathbf{x})\}$.

Доказательство

- Из существования производной по направлению $f_i'(\mathbf{x}, \mathbf{d})$ следует, что $\lim_{t \to +0} f_i(\mathbf{x} + t\mathbf{d}) = f_i(\mathbf{x})$

Производная по направлению для максимума

Теорема

Пусть $f(\mathbf{x}) = \max_{i=1,\ldots,k} f_i(\mathbf{x})$, где f_i такие функции, что для них существует производная по направлению $\mathbf{d}$. Тогда для $\mathbf{x} \in \bigcap\limits_{i=1}^{k} \mathrm{int}\,(\mathrm{dom}\,(f_i))$ выполнено

$$f'(\mathbf{x}, \mathbf{d}) = \max_{i \in \mathcal{I}(\mathbf{x})} f_i'(\mathbf{x}, \mathbf{d}),$$

где $\mathcal{I}(\mathbf{x}) = \{i \mid f_i(\mathbf{x}) = f(\mathbf{x})\}$.

Доказательство

- Из существования производной по направлению $f_i'(\mathbf{x}, \mathbf{d})$ следует, что $\lim_{t \to +0} f_i(\mathbf{x} + t\mathbf{d}) = f_i(\mathbf{x})$
- Значит найдётся $\varepsilon > 0$ что для всех $t \in (0, \varepsilon]$ выполнено $f_i(\mathbf{x} + t\mathbf{d}) > f_j(\mathbf{x} + t\mathbf{d})$ для всех $i \in \mathcal{I}(\mathbf{x})$ и $j \notin \mathcal{I}(\mathbf{x})$

Производная по направлению для максимума

Теорема

Пусть $f(\mathbf{x}) = \max_{i=1,\ldots,k} f_i(\mathbf{x})$, где f_i такие функции, что для них существует производная по направлению $\mathbf{d}$. Тогда для $\mathbf{x} \in \bigcap\limits_{i=1}^{k} \mathrm{int}\,(\mathrm{dom}\,(f_i))$ выполнено

$$f'(\mathbf{x}, \mathbf{d}) = \max_{i \in \mathcal{I}(\mathbf{x})} f_i'(\mathbf{x}, \mathbf{d}),$$

где $\mathcal{I}(\mathbf{x}) = \{i \mid f_i(\mathbf{x}) = f(\mathbf{x})\}$.

Доказательство

- Из существования производной по направлению $f_i'(\mathbf{x}, \mathbf{d})$ следует, что $\lim_{t \to +0} f_i(\mathbf{x} + t\mathbf{d}) = f_i(\mathbf{x})$
- Значит найдётся $\varepsilon > 0$ что для всех $t \in (0, \varepsilon]$ выполнено $f_i(\mathbf{x} + t\mathbf{d}) > f_j(\mathbf{x} + t\mathbf{d})$ для всех $i \in \mathcal{I}(\mathbf{x})$ и $j \notin \mathcal{I}(\mathbf{x})$
- Поэтому для всех $t \in (0, \varepsilon]$ выполнено $f(\mathbf{x}) = \max_{i=1,\ldots,k} f_i(\mathbf{x}) = \max_{i \in \mathcal{I}(\mathbf{x})} f_i(\mathbf{x})$

Производная по направлению для максимума

Теорема

Пусть $f(\mathbf{x}) = \max_{i=1,\dots,k} f_i(\mathbf{x})$, где f_i такие функции, что для них существует производная по направлению $\mathbf{d}$. Тогда для $\mathbf{x} \in \bigcap\limits_{i=1}^{k} \mathrm{int}\,(\mathrm{dom}\,(f_i))$ выполнено

$$f'(\mathbf{x}, \mathbf{d}) = \max_{i \in \mathcal{I}(\mathbf{x})} f_i'(\mathbf{x}, \mathbf{d}),$$

где $\mathcal{I}(\mathbf{x}) = \{i \mid f_i(\mathbf{x}) = f(\mathbf{x})\}$.

Доказательство

- Из существования производной по направлению $f_i'(\mathbf{x}, \mathbf{d})$ следует, что $\lim_{t \to +0} f_i(\mathbf{x} + t\mathbf{d}) = f_i(\mathbf{x})$
- Значит найдётся $\varepsilon > 0$ что для всех $t \in (0, \varepsilon]$ выполнено $f_i(\mathbf{x} + t\mathbf{d}) > f_j(\mathbf{x} + t\mathbf{d})$ для всех $i \in \mathcal{I}(\mathbf{x})$ и $j \not\in \mathcal{I}(\mathbf{x})$
- Поэтому для всех $t \in (0, \varepsilon]$ выполнено $f(\mathbf{x}) = \max_{i=1,\dots,k} f_i(\mathbf{x}) = \max_{i \in \mathcal{I}(\mathbf{x})} f_i(\mathbf{x})$
-

- Рассмотрим отношение для всех $t \in (0, \varepsilon]$

$$\frac{f(\mathbf{x}+t\mathbf{d})-f(\mathbf{x})}{t} = \frac{\max\limits_{i \in \mathcal{I}(\mathbf{x})} f_i(\mathbf{x})-f(\mathbf{x})}{t} = \max_{i \in \mathcal{I}(\mathbf{x})} \frac{f_i(\mathbf{x}+t\mathbf{d})-f_i(\mathbf{x})}{t}$$

- Рассмотрим отношение для всех $t \in (0, \varepsilon]$
$$\frac{f(\mathbf{x}+t\mathbf{d})-f(\mathbf{x})}{t} = \frac{\max\limits_{i\in\mathcal{I}(\mathbf{x})} f_i(\mathbf{x})-f(\mathbf{x})}{t} = \max_{i\in\mathcal{I}(\mathbf{x})} \frac{f_i(\mathbf{x}+t\mathbf{d})-f_i(\mathbf{x})}{t}$$
- Тогда производная по направлению
$$f'(\mathbf{x}, \mathbf{d}) = \lim_{t\to+0} \max_{i\in\mathcal{I}(\mathbf{x})} \frac{f_i(\mathbf{x}+t\mathbf{d})-f_i(\mathbf{x})}{t} =$$
$$\max_{i\in\mathcal{I}(\mathbf{x})} \lim_{t\to+0} \frac{f_i(\mathbf{x}+t\mathbf{d})-f_i(\mathbf{x})}{t} = \max_{i\in\mathcal{I}(\mathbf{x})} f_i'(\mathbf{x}, \mathbf{d})$$

- Рассмотрим отношение для всех $t \in (0, \varepsilon]$
$\frac{f(\mathbf{x}+t\mathbf{d})-f(\mathbf{x})}{t} = \frac{\max\limits_{i \in \mathcal{I}(\mathbf{x})} f_i(\mathbf{x})-f(\mathbf{x})}{t} = \max\limits_{i \in \mathcal{I}(\mathbf{x})} \frac{f_i(\mathbf{x}+t\mathbf{d})-f_i(\mathbf{x})}{t}$
- Тогда производная по направлению
$f'(\mathbf{x}, \mathbf{d}) = \lim\limits_{t \to +0} \max\limits_{i \in \mathcal{I}(\mathbf{x})} \frac{f_i(\mathbf{x}+t\mathbf{d})-f_i(\mathbf{x})}{t} =$
$\max\limits_{i \in \mathcal{I}(\mathbf{x})} \lim\limits_{t \to +0} \frac{f_i(\mathbf{x}+t\mathbf{d})-f_i(\mathbf{x})}{t} = \max\limits_{i \in \mathcal{I}(\mathbf{x})} f_i'(\mathbf{x}, \mathbf{d})$

Следствие

Пусть $f(\mathbf{x}) = \max\limits_{i=1,\dots,k} f_i(\mathbf{x})$ и f_i выпуклые. Тогда для $\mathbf{x} \in \bigcap\limits_{i=1}^{k} \operatorname{int}(\operatorname{dom}(f_i))$ выполнено

$$f'(\mathbf{x}, \mathbf{d}) = \max_{i \in \mathcal{I}(\mathbf{x})} f_i'(\mathbf{x}, \mathbf{d}),$$

где $\mathcal{I}(\mathbf{x}) = \{i \mid f_i(\mathbf{x}) = f(\mathbf{x})\}$.

Субдифференциал для максимума выпуклых функций

Теорема

Пусть $f_1, \ldots, f_m$ выпуклые функции и $f(\mathbf{x}) = \max\limits_{i=1,\ldots,m} f_i(\mathbf{x})$.

Пусть $\mathbf{x} \in \bigcap\limits_{i=1}^{m} \mathrm{int}\,(\mathrm{dom}\,(f_i))$. Тогда

$$\partial f(\mathbf{x}) = \mathrm{conv}\left(\bigcup_{i \in \mathcal{I}(\mathbf{x})} \partial f_i(\mathbf{x}) \right),$$

где $\mathcal{I}(\mathbf{x}) = \{i \mid f_i(\mathbf{x}) = f(\mathbf{x})\}$.

Субдифференциал для максимума выпуклых функций

Теорема

Пусть $f_1, \ldots, f_m$ выпуклые функции и $f(\mathbf{x}) = \max\limits_{i=1,\ldots,m} f_i(\mathbf{x})$.

Пусть $\mathbf{x} \in \bigcap\limits_{i=1}^{m} \mathrm{int}\,(\mathrm{dom}\,(f_i))$. Тогда

$$\partial f(\mathbf{x}) = \mathrm{conv}\left(\bigcup_{i \in \mathcal{I}(\mathbf{x})} \partial f_i(\mathbf{x})\right),$$

где $\mathcal{I}(\mathbf{x}) = \{i \mid f_i(\mathbf{x}) = f(\mathbf{x})\}$.

Доказательство

Субдифференциал для максимума выпуклых функций

Теорема

Пусть $f_1, \ldots, f_m$ выпуклые функции и $f(\mathbf{x}) = \max\limits_{i=1,\ldots,m} f_i(\mathbf{x})$.

Пусть $\mathbf{x} \in \bigcap\limits_{i=1}^{m} \operatorname{int}(\operatorname{dom}(f_i))$. Тогда

$$\partial f(\mathbf{x}) = \operatorname{conv}\left(\bigcup_{i \in \mathcal{I}(\mathbf{x})} \partial f_i(\mathbf{x}) \right),$$

где $\mathcal{I}(\mathbf{x}) = \{ i \mid f_i(\mathbf{x}) = f(\mathbf{x}) \}$.

Доказательство

- f выпуклая функция как максимум выпуклых функций

Субдифференциал для максимума выпуклых функций

Теорема

Пусть $f_1, \ldots, f_m$ выпуклые функции и $f(\mathbf{x}) = \max\limits_{i=1,\ldots,m} f_i(\mathbf{x})$.

Пусть $\mathbf{x} \in \bigcap\limits_{i=1}^{m} \mathrm{int}\,(\mathrm{dom}\,(f_i))$. Тогда

$$\partial f(\mathbf{x}) = \mathrm{conv}\left(\bigcup_{i \in \mathcal{I}(\mathbf{x})} \partial f_i(\mathbf{x})\right),$$

где $\mathcal{I}(\mathbf{x}) = \{i \mid f_i(\mathbf{x}) = f(\mathbf{x})\}$.

Доказательство

- f выпуклая функция как максимум выпуклых функций
- По следствию из предыдущей теоремы для любого направления $\mathbf{d}$: $f'(\mathbf{x}, \mathbf{d}) = \max_{i \in \mathcal{I}(\mathbf{x})} f_i'(\mathbf{x}, \mathbf{d})$

Субдифференциал для максимума выпуклых функций

Теорема

Пусть $f_1, \ldots, f_m$ выпуклые функции и $f(\mathbf{x}) = \max\limits_{i=1,\ldots,m} f_i(\mathbf{x})$.

Пусть $\mathbf{x} \in \bigcap\limits_{i=1}^{m} \mathrm{int}\left(\mathrm{dom}\left(f_i\right)\right)$. Тогда

$$\partial f(\mathbf{x}) = \mathrm{conv}\left(\bigcup_{i \in \mathcal{I}(\mathbf{x})} \partial f_i(\mathbf{x})\right),$$

где $\mathcal{I}(\mathbf{x}) = \{i \mid f_i(\mathbf{x}) = f(\mathbf{x})\}$.

Доказательство

- f выпуклая функция как максимум выпуклых функций
- По следствию из предыдущей теоремы для любого направления $\mathbf{d}$: $f'(\mathbf{x}, \mathbf{d}) = \max_{i \in \mathcal{I}(\mathbf{x})} f'_i(\mathbf{x}, \mathbf{d})$
- Без ограничения общности будем считать, что $\mathcal{I}(\mathbf{x}) = \{1, 2, \ldots, k\}$ для некоторого $k \in \{1, \ldots, m\}$

- По формуле для связи производной по направлению и субдифференциала имеем $f'(\mathbf{x}, \mathbf{d}) = \max\limits_{i=1,\ldots,k} \max\limits_{\mathbf{g}_i \in \partial f_i(\mathbf{x})} \langle \mathbf{g}_i, \mathbf{d} \rangle$

- По формуле для связи производной по направлению и субдифференциала имеем $f'(\mathbf{x}, \mathbf{d}) = \max\limits_{i=1,\dots,k} \max\limits_{\mathbf{g}_i \in \partial f_i(\mathbf{x})} \langle \mathbf{g}_i, \mathbf{d} \rangle$
- Воспользуемся равенством $\max\{y_1, \dots, y_k\} = \max\limits_{\boldsymbol{\lambda} \in \Delta_k} \sum_{i=1}^k y_i \lambda_i$, его легко проверить из геометрических соображений

- По формуле для связи производной по направлению и субдифференциала имеем $f'(\mathbf{x}, \mathbf{d}) = \max\limits_{i=1,\dots,k} \max\limits_{\mathbf{g}_i \in \partial f_i(\mathbf{x})} \langle \mathbf{g}_i, \mathbf{d} \rangle$
- Воспользуемся равенством $\max\{y_1, \dots, y_k\} = \max\limits_{\boldsymbol{\lambda} \in \Delta_k} \sum_{i=1}^k y_i \lambda_i$, его легко проверить из геометрических соображений
- Тогда $f'(\mathbf{x}, \mathbf{d}) = \max\limits_{\boldsymbol{\lambda} \in \Delta_k} \left\{ \sum_{i=1}^k \lambda_i \max\limits_{\mathbf{g}_i \in \partial f_i(\mathbf{x})} \langle \mathbf{g}_i, \mathbf{d} \rangle \right\} = \max\limits_{\boldsymbol{\lambda} \in \Delta_k,\, \mathbf{g}_i \in \partial f_i(\mathbf{x})} \left\langle \sum_{i=1}^k \lambda_i \mathbf{g}_i, \mathbf{d} \right\rangle = \max\limits_{\mathbf{g} \in \mathrm{conv}\left(\cup_{i=1}^k \partial f_i(\mathbf{x})\right)} \langle \mathbf{g}, \mathbf{d} \rangle$

- По формуле для связи производной по направлению и субдифференциала имеем $f'(\mathbf{x}, \mathbf{d}) = \max\limits_{i=1,\ldots,k} \max\limits_{\mathbf{g}_i \in \partial f_i(\mathbf{x})} \langle \mathbf{g}_i, \mathbf{d} \rangle$
- Воспользуемся равенством $\max\{y_1, \ldots, y_k\} = \max\limits_{\boldsymbol{\lambda} \in \Delta_k} \sum_{i=1}^k y_i \lambda_i$, его легко проверить из геометрических соображений
- Тогда $f'(\mathbf{x}, \mathbf{d}) = \max\limits_{\boldsymbol{\lambda} \in \Delta_k} \left\{ \sum_{i=1}^k \lambda_i \max\limits_{\mathbf{g}_i \in \partial f_i(\mathbf{x})} \langle \mathbf{g}_i, \mathbf{d} \rangle \right\} = \max\limits_{\boldsymbol{\lambda} \in \Delta_k,\, \mathbf{g}_i \in \partial f_i(\mathbf{x})} \left\langle \sum_{i=1}^k \lambda_i \mathbf{g}_i, \mathbf{d} \right\rangle = \max\limits_{\mathbf{g} \in \mathrm{conv}\left(\cup_{i=1}^k \partial f_i(\mathbf{x})\right)} \langle \mathbf{g}, \mathbf{d} \rangle$
- Выразим результат через опорную функцию $f'(\mathbf{x}, \mathbf{d}) = \sigma_{\mathcal{A}}(\mathbf{d})$, где $\mathcal{A} = \mathrm{conv}\left(\cup_{i=1}^k \partial f_i(\mathbf{x})\right)$

- По формуле для связи производной по направлению и субдифференциала имеем $f'(\mathbf{x}, \mathbf{d}) = \max\limits_{i=1,\dots,k} \max\limits_{\mathbf{g}_i \in \partial f_i(\mathbf{x})} \langle \mathbf{g}_i, \mathbf{d} \rangle$
- Воспользуемся равенством $\max\{y_1, \dots, y_k\} = \max\limits_{\boldsymbol{\lambda} \in \Delta_k} \sum_{i=1}^k y_i \lambda_i$, его легко проверить из геометрических соображений
- Тогда $f'(\mathbf{x}, \mathbf{d}) = \max\limits_{\boldsymbol{\lambda} \in \Delta_k} \left\{ \sum_{i=1}^k \lambda_i \max\limits_{\mathbf{g}_i \in \partial f_i(\mathbf{x})} \langle \mathbf{g}_i, \mathbf{d} \rangle \right\} = \max\limits_{\boldsymbol{\lambda} \in \Delta_k,\, \mathbf{g}_i \in \partial f_i(\mathbf{x})} \left\langle \sum_{i=1}^k \lambda_i \mathbf{g}_i, \mathbf{d} \right\rangle = \max\limits_{\mathbf{g} \in \mathrm{conv}\left(\cup_{i=1}^k \partial f_i(\mathbf{x})\right)} \langle \mathbf{g}, \mathbf{d} \rangle$
- Выразим результат через опорную функцию $f'(\mathbf{x}, \mathbf{d}) = \sigma_{\mathcal{A}}(\mathbf{d})$, где $\mathcal{A} = \mathrm{conv}\left(\cup_{i=1}^k \partial f_i(\mathbf{x})\right)$
- С другой стороны $f'(\mathbf{x}, \mathbf{d})$ связана с $\partial f(\mathbf{x})$ в точках $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ как $f'(\mathbf{x}, \mathbf{d}) = \sigma_{\partial f(\mathbf{x})}(\mathbf{d})$

- По формуле для связи производной по направлению и субдифференциала имеем $f'(\mathbf{x}, \mathbf{d}) = \max\limits_{i=1,\dots,k} \max\limits_{\mathbf{g}_i \in \partial f_i(\mathbf{x})} \langle \mathbf{g}_i, \mathbf{d} \rangle$
- Воспользуемся равенством $\max\{y_1, \dots, y_k\} = \max\limits_{\boldsymbol{\lambda} \in \Delta_k} \sum_{i=1}^k y_i \lambda_i$, его легко проверить из геометрических соображений
- Тогда $f'(\mathbf{x}, \mathbf{d}) = \max\limits_{\boldsymbol{\lambda} \in \Delta_k} \left\{ \sum_{i=1}^k \lambda_i \max\limits_{\mathbf{g}_i \in \partial f_i(\mathbf{x})} \langle \mathbf{g}_i, \mathbf{d} \rangle \right\} = \max\limits_{\boldsymbol{\lambda} \in \Delta_k,\, \mathbf{g}_i \in \partial f_i(\mathbf{x})} \left\langle \sum_{i=1}^k \lambda_i \mathbf{g}_i, \mathbf{d} \right\rangle = \max\limits_{\mathbf{g} \in \mathrm{conv}\left(\cup_{i=1}^k \partial f_i(\mathbf{x})\right)} \langle \mathbf{g}, \mathbf{d} \rangle$
- Выразим результат через опорную функцию $f'(\mathbf{x}, \mathbf{d}) = \sigma_{\mathcal{A}}(\mathbf{d})$, где $\mathcal{A} = \mathrm{conv}\left(\cup_{i=1}^k \partial f_i(\mathbf{x})\right)$
- С другой стороны $f'(\mathbf{x}, \mathbf{d})$ связана с $\partial f(\mathbf{x})$ в точках $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ как $f'(\mathbf{x}, \mathbf{d}) = \sigma_{\partial f(\mathbf{x})}(\mathbf{d})$
- Таким образом, $\sigma_{\partial f(\mathbf{x})}(\mathbf{d}) = \sigma_{\mathcal{A}}(\mathbf{d})$, где $\mathcal{A} = \mathrm{conv}\left(\cup_{i=1}^k \partial f_i(\mathbf{x})\right)$

- Множество $\partial f(\mathbf{x})$ выпукло, замкнуто и так как $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$, то ограничено и непусто

- Множество $\partial f(\mathbf{x})$ выпукло, замкнуто и так как $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$, то ограничено и непусто
- Для $\partial f_i(\mathbf{x})$ выполнена компактность и непустота

- Множество $\partial f(\mathbf{x})$ выпукло, замкнуто и так как $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$, то ограничено и непусто
- Для $\partial f_i(\mathbf{x})$ выполнена компактность и непустота
- Значит для объединения выполнены те же свойства

- Множество $\partial f(\mathbf{x})$ выпукло, замкнуто и так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то ограничено и непусто
- Для $\partial f_i(\mathbf{x})$ выполнена компактность и непустота
- Значит для объединения выполнены те же свойства
- Тогда $\mathcal{A}$ также выпукло и замкнуто

- Множество $\partial f(\mathbf{x})$ выпукло, замкнуто и так как $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$, то ограничено и непусто
- Для $\partial f_i(\mathbf{x})$ выполнена компактность и непустота
- Значит для объединения выполнены те же свойства
- Тогда $\mathcal{A}$ также выпукло и замкнуто
- В итоге, из теоремы о связи равенства опорных функций и совпадения множеств следует, что $\mathcal{A} = \partial f(\mathbf{x})$

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 $\partial_{\mathcal{X}} f(\mathbf{x}_0)$.

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

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

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

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

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

Главное

- Субградиент и субдифференциал

Главное

- Субградиент и субдифференциал
- Субдифференциал и производная по направлению

Главное

- Субградиент и субдифференциал
- Субдифференциал и производная по направлению
- Опорная функция

Главное

- Субградиент и субдифференциал
- Субдифференциал и производная по направлению
- Опорная функция
- Способы вычисления субдифференциалов