Методы оптимизации
Лекция 7: Введение в условия оптимальности и сопряжённые функции

Александр Катруца

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

26 октября 2020 г.

На прошлой лекции

- Субградиент и субдифференциал

На прошлой лекции

- Субградиент и субдифференциал
- Субдифференциал и производная по направлению

На прошлой лекции

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

На прошлой лекции

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

План на эту лекцию

- Условия оптимальности

План на эту лекцию

- Условия оптимальности
- Необходимое условие первого порядка

План на эту лекцию

- Условия оптимальности
- Необходимое условие первого порядка
- Сопряжённые функции

План на эту лекцию

- Условия оптимальности
- Необходимое условие первого порядка
- Сопряжённые функции
- Примеры и свойства

Существование решения задачи оптимизации

Теорема

Пусть f полунепрерывная снизу функция на компактном множестве. Тогда эта функция достигает своего минимального значения на этом компакте.

Существование решения задачи оптимизации

Теорема

Пусть f полунепрерывная снизу функция на компактном множестве. Тогда эта функция достигает своего минимального значения на этом компакте.

Доказательство

Существование решения задачи оптимизации

Теорема

Пусть f полунепрерывная снизу функция на компактном множестве. Тогда эта функция достигает своего минимального значения на этом компакте.

Доказательство

- По определению инфимума найдётся последовательность $\{\mathbf{x}_n\} \subset \mathcal{C}$ такая что $f(\mathbf{x}_n) \to f^* = \inf_{\mathbf{x} \in \mathcal{C}} f(\mathbf{x})$

Существование решения задачи оптимизации

Теорема

Пусть f полунепрерывная снизу функция на компактном множестве. Тогда эта функция достигает своего минимального значения на этом компакте.

Доказательство

- По определению инфимума найдётся последовательность $\{\mathbf{x}_n\} \subset \mathcal{C}$ такая что $f(\mathbf{x}_n) \to f^* = \inf_{\mathbf{x} \in \mathcal{C}} f(\mathbf{x})$
- Так как $\mathcal{C}$ — компакт, то можно выделить сходящуюся подпоследовательность $\mathbf{x}_{n_k} \to \mathbf{x}^* \in \mathcal{C}$ в силу компактности множества $\mathcal{C}$

Существование решения задачи оптимизации

Теорема

Пусть f полунепрерывная снизу функция на компактном множестве. Тогда эта функция достигает своего минимального значения на этом компакте.

Доказательство

- По определению инфимума найдётся последовательность $\{\mathbf{x}_n\} \subset \mathcal{C}$ такая что $f(\mathbf{x}_n) \to f^* = \inf_{\mathbf{x} \in \mathcal{C}} f(\mathbf{x})$
- Так как $\mathcal{C}$ — компакт, то можно выделить сходящуюся подпоследовательность $\mathbf{x}_{n_k} \to \mathbf{x}^* \in \mathcal{C}$ в силу компактности множества $\mathcal{C}$
- В силу полунепрерывности снизу $f(\mathbf{x}^*) \leq \lim\limits_{k \to \infty} f(\mathbf{x}_{n_k})$

Существование решения задачи оптимизации

Теорема

Пусть f полунепрерывная снизу функция на компактном множестве. Тогда эта функция достигает своего минимального значения на этом компакте.

Доказательство

- По определению инфимума найдётся последовательность $\{\mathbf{x}_n\} \subset \mathcal{C}$ такая что $f(\mathbf{x}_n) \to f^* = \inf_{\mathbf{x} \in \mathcal{C}} f(\mathbf{x})$
- Так как $\mathcal{C}$ — компакт, то можно выделить сходящуюся подпоследовательность $\mathbf{x}_{n_k} \to \mathbf{x}^* \in \mathcal{C}$ в силу компактности множества $\mathcal{C}$
- В силу полунепрерывности снизу $f(\mathbf{x}^*) \leq \lim\limits_{k \to \infty} f(\mathbf{x}_{n_k})$
- Но мы знаем, что $f(\mathbf{x}_n) \to f^*$, а значит и $\lim\limits_{k \to \infty} f(\mathbf{x}_{n_k}) = f^*$

Существование решения задачи оптимизации

Теорема

Пусть f полунепрерывная снизу функция на компактном множестве. Тогда эта функция достигает своего минимального значения на этом компакте.

Доказательство

- По определению инфимума найдётся последовательность $\{\mathbf{x}_n\} \subset \mathcal{C}$ такая что $f(\mathbf{x}_n) \to f^* = \inf_{\mathbf{x} \in \mathcal{C}} f(\mathbf{x})$
- Так как $\mathcal{C}$ — компакт, то можно выделить сходящуюся подпоследовательность $\mathbf{x}_{n_k} \to \mathbf{x}^* \in \mathcal{C}$ в силу компактности множества $\mathcal{C}$
- В силу полунепрерывности снизу $f(\mathbf{x}^*) \leq \lim\limits_{k \to \infty} f(\mathbf{x}_{n_k})$
- Но мы знаем, что $f(\mathbf{x}_n) \to f^*$, а значит и $\lim\limits_{k \to \infty} f(\mathbf{x}_{n_k}) = f^*$
- Значит $f(\mathbf{x}^*) \leq f^*$ и следовательно инфимум достигается в точке $\mathbf{x}^*$

Как это нам поможет?

- Замкнутая выпуклая функция полунепрерывна снизу

Как это нам поможет?

- Замкнутая выпуклая функция полунепрерывна снизу
- Задача минимизации замкнутой выпуклой функции на компакте всегда имеет решение

Как это нам поможет?

- Замкнутая выпуклая функция полунепрерывна снизу
- Задача минимизации замкнутой выпуклой функции на компакте всегда имеет решение
- Таким образом, решение большинства практически важных задач существует

Общее условие оптимальности

Задача оптимизации в общем виде

$$f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathcal{X}}$$

Общее условие оптимальности

Задача оптимизации в общем виде

$$f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathcal{X}}$$

Условие оптимальности

Пусть f определена на множестве $\mathcal{X} \subset \mathbb{R}^n$. Тогда

Общее условие оптимальности

Задача оптимизации в общем виде

$$f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathcal{X}}$$

Условие оптимальности

Пусть f определена на множестве $\mathcal{X} \subset \mathbb{R}^n$. Тогда

1. Если $\mathbf{x}^*$ точка минимума функции f на $\mathcal{X}$, тогда
   $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$

Общее условие оптимальности

Задача оптимизации в общем виде

$$f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathcal{X}}$$

Условие оптимальности

Пусть f определена на множестве $\mathcal{X} \subset \mathbb{R}^n$. Тогда

1. Если $\mathbf{x}^*$ точка минимума функции f на $\mathcal{X}$, тогда $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$
2. Если в некоторой точке $\mathbf{x}^* \in \mathcal{X}$ существует субдифференциал $\partial_{\mathcal{X}} f(\mathbf{x}^*)$ и $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$, тогда $\mathbf{x}^*$ точка минимума функции f на множестве $\mathcal{X}$.

Общее условие оптимальности

Задача оптимизации в общем виде

$$
f(\mathbf{x}) \to \min_{\mathbf{x} \in \mathcal{X}}
$$

Условие оптимальности

Пусть f определена на множестве $\mathcal{X} \subset \mathbb{R}^n$. Тогда

1. Если $\mathbf{x}^*$ точка минимума функции f на $\mathcal{X}$, тогда $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$
2. Если в некоторой точке $\mathbf{x}^* \in \mathcal{X}$ существует субдифференциал $\partial_{\mathcal{X}} f(\mathbf{x}^*)$ и $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$, тогда $\mathbf{x}^*$ точка минимума функции f на множестве $\mathcal{X}$.

Q: какие недостатки у этого критерия?

Доказательство

1. Если $\mathbf{x}^*$ точка минимума, тогда

Доказательство

1. Если $\mathbf{x}^*$ точка минимума, тогда
   - для любого $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$

Доказательство

1. Если $\mathbf{x}^*$ точка минимума, тогда
   - для любого $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
   - Это неравенство можно переписать в виде: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$

Доказательство

1. Если $\mathbf{x}^*$ точка минимума, тогда
   - для любого $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
   - Это неравенство можно переписать в виде: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$
   - Это означает, что $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ по определению

Доказательство

1. Если $\mathbf{x}^*$ точка минимума, тогда
    - для любого $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
    - Это неравенство можно переписать в виде: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$
    - Это означает, что $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ по определению
2. Если $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ в некоторрой точке $\mathbf{x}^*$, тогда

Доказательство

1. Если $\mathbf{x}^*$ точка минимума, тогда
   - для любого $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
   - Это неравенство можно переписать в виде: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$
   - Это означает, что $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ по определению
2. Если $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ в некоторрой точке $\mathbf{x}^*$, тогда
   - по определению субдифференциала $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle \mathbf{a}, \mathbf{x} - \mathbf{x}^* \rangle$

Доказательство

1. Если $\mathbf{x}^*$ точка минимума, тогда
   - для любого $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
   - Это неравенство можно переписать в виде:
     $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$
   - Это означает, что $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ по определению
2. Если $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ в некоторрой точке $\mathbf{x}^*$, тогда
   - по определению субдифференциала
     $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle \mathbf{a}, \mathbf{x} - \mathbf{x}^* \rangle$
   - Если $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$, тогда получим неравенство
     $f(\mathbf{x}) \geq f(\mathbf{x}^*)$ для любого $\mathbf{x} \in \mathcal{X}$

Доказательство

1. Если $\mathbf{x}^*$ точка минимума, тогда
    - для любого $\mathbf{x} \in \mathcal{X}$: $f(\mathbf{x}) \geq f(\mathbf{x}^*)$
    - Это неравенство можно переписать в виде: $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle 0, \mathbf{x} - \mathbf{x}^* \rangle$
    - Это означает, что $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ по определению
2. Если $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$ в некоторрой точке $\mathbf{x}^*$, тогда
    - по определению субдифференциала $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \langle \mathbf{a}, \mathbf{x} - \mathbf{x}^* \rangle$
    - Если $0 \in \partial_{\mathcal{X}} f(\mathbf{x}^*)$, тогда получим неравенство $f(\mathbf{x}) \geq f(\mathbf{x}^*)$ для любого $\mathbf{x} \in \mathcal{X}$
    - Это означает, что $\mathbf{x}^*$ точка минимума

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \tag{1}$$

Теорема

Если $\mathbf{x}^*$ *решение задачи* (1), *то* $f'(\mathbf{x}^*) = 0$.

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \tag{1}$$

Теорема

Если $\mathbf{x}^*$ *решение задачи* (1)*, то* $f'(\mathbf{x}^*) = 0$*.*

Доказательство

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \tag{1}$$

Теорема

Если $\mathbf{x}^*$ *решение задачи* (1), *то* $f'(\mathbf{x}^*) = 0$.

Доказательство

- $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y})$ и $\lim\limits_{\mathbf{y} \to \mathbf{x}^*} \frac{r(\mathbf{x}^*, \mathbf{y})}{\|\mathbf{x}^* - \mathbf{y}\|_2} = 0 \ (*)$

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \qquad (1)$$

Теорема

Если $\mathbf{x}^*$ *решение задачи* (1)*, то* $f'(\mathbf{x}^*) = 0$*.*

Доказательство

- $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle + r(\mathbf{x}^*, \mathbf{y})$ и $\lim\limits_{\mathbf{y} \to \mathbf{x}^*} \frac{r(\mathbf{x}^*, \mathbf{y})}{\|\mathbf{x}^* - \mathbf{y}\|_2} = 0 \ (*)$
- Если $f'(\mathbf{x}^*) \neq 0$, тогда рассмотрим $\mathbf{y}(\tau) = \mathbf{x}^* - \tau f'(\mathbf{x}^*)$, $\tau > 0$

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \qquad (1)$$

Теорема

Если $\mathbf{x}^*$ *решение задачи* (1), *то* $f'(\mathbf{x}^*) = 0$.

Доказательство

- $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^*\rangle + r(\mathbf{x}^*, \mathbf{y})$ и $\lim\limits_{\mathbf{y}\to\mathbf{x}^*} \frac{r(\mathbf{x}^*,\mathbf{y})}{\|\mathbf{x}^*-\mathbf{y}\|_2} = 0 \ (*)$
- Если $f'(\mathbf{x}^*) \neq 0$, тогда рассмотрим $\mathbf{y}(\tau) = \mathbf{x}^* - \tau f'(\mathbf{x}^*)$, $\tau > 0$
- $f(\mathbf{y}(\tau)) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y}(\tau) - \mathbf{x}^*\rangle + r(\mathbf{x}^*, \mathbf{y}(\tau)) = f(\mathbf{x}^*) - \tau\|f'(\mathbf{x}^*)\|_2^2 + r(\mathbf{x}^*, \mathbf{y}(\tau))$

Условие оптимальности для задачи без ограничений

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \tag{1}$$

Теорема

Если $\mathbf{x}^*$ *решение задачи* (1), *то* $f'(\mathbf{x}^*) = 0$.

Доказательство

- $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^*\rangle + r(\mathbf{x}^*, \mathbf{y})$ и $\lim\limits_{\mathbf{y}\to\mathbf{x}^*} \frac{r(\mathbf{x}^*,\mathbf{y})}{\|\mathbf{x}^*-\mathbf{y}\|_2} = 0$ $(*)$
- Если $f'(\mathbf{x}^*) \neq 0$, тогда рассмотрим $\mathbf{y}(\tau) = \mathbf{x}^* - \tau f'(\mathbf{x}^*)$, $\tau > 0$
- $f(\mathbf{y}(\tau)) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y}(\tau) - \mathbf{x}^*\rangle + r(\mathbf{x}^*, \mathbf{y}(\tau)) = f(\mathbf{x}^*) - \tau\|f'(\mathbf{x}^*)\|_2^2 + r(\mathbf{x}^*, \mathbf{y}(\tau))$
- В силу $(*)$ найдётся $\bar{\tau}$ такое что для всех $\tau \in (0, \bar{\tau})$ выполнено
$$f(\mathbf{y}(\tau)) - f(\mathbf{x}^*) \leq -\frac{\tau}{2}\|f'(\mathbf{x}^*)\|_2^2 < 0$$
Значит $\mathbf{x}^*$ не минимум, противоречие.

Замечания к предыдущей теореме

- Эта теорема частный случай первой теоремы для общей задачи оптимизации

Замечания к предыдущей теореме

- Эта теорема частный случай первой теоремы для общей задачи оптимизации
- Дополнительное доказательство приведено для демонстрации свойства убывания у направления $-f'(\mathbf{x})$

Замечания к предыдущей теореме

- Эта теорема частный случай первой теоремы для общей задачи оптимизации
- Дополнительное доказательство приведено для демонстрации свойства убывания у направления $-f'(\mathbf{x})$
- Обратите внимание, что субградиент таким свойством НЕ обладает!

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) *функция* f *выпукла, то* $\mathbf{x}^*$ *глобальный минимум тогда и только тогда* $f'(\mathbf{x}^*) = 0$

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) *функция* f *выпукла, то* $\mathbf{x}^*$ *глобальный минимум тогда и только тогда* $f'(\mathbf{x}^*) = 0$

Доказательство

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) *функция* f *выпукла, то* $\mathbf{x}^*$ *глобальный минимум тогда и только тогда* $f'(\mathbf{x}^*) = 0$

Доказательство

- Если $\mathbf{x}^*$ глобальный минимум, то $\mathbf{x}^*$ локальный минимум

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) *функция* f *выпукла, то* $\mathbf{x}^*$ *глобальный минимум тогда и только тогда* $f'(\mathbf{x}^*) = 0$

Доказательство

- Если $\mathbf{x}^*$ глобальный минимум, то $\mathbf{x}^*$ локальный минимум
- Значит $f'(\mathbf{x}^*) = 0$ по предыдущей теореме

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) *функция* f *выпукла, то* $\mathbf{x}^*$ *глобальный минимум тогда и только тогда* $f'(\mathbf{x}^*) = 0$

Доказательство

- Если $\mathbf{x}^*$ глобальный минимум, то $\mathbf{x}^*$ локальный минимум
- Значит $f'(\mathbf{x}^*) = 0$ по предыдущей теореме
- Пусть $\mathbf{x}^*$ такая точка, что $f'(\mathbf{x}^*) = 0$ и функция выпукла

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) *функция* f *выпукла, то* $\mathbf{x}^*$ *глобальный минимум тогда и только тогда* $f'(\mathbf{x}^*) = 0$

Доказательство

- Если $\mathbf{x}^*$ глобальный минимум, то $\mathbf{x}^*$ локальный минимум
- Значит $f'(\mathbf{x}^*) = 0$ по предыдущей теореме
- Пусть $\mathbf{x}^*$ такая точка, что $f'(\mathbf{x}^*) = 0$ и функция выпукла
- Тогда по критерию выпуклости

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle = f(\mathbf{x}^*)$$

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) *функция* f *выпукла, то* $\mathbf{x}^*$ *глобальный минимум тогда и только тогда* $f'(\mathbf{x}^*) = 0$

Доказательство

- Если $\mathbf{x}^*$ глобальный минимум, то $\mathbf{x}^*$ локальный минимум
- Значит $f'(\mathbf{x}^*) = 0$ по предыдущей теореме
- Пусть $\mathbf{x}^*$ такая точка, что $f'(\mathbf{x}^*) = 0$ и функция выпукла
- Тогда по критерию выпуклости

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle = f(\mathbf{x}^*)$$

- Значит $\mathbf{x}^*$ – глобальный минимум.

Критерий для выпуклой задачи без ограничений

Теорема

Если в задаче (1) *функция* f *выпукла, то* $\mathbf{x}^*$ *глобальный минимум тогда и только тогда* $f'(\mathbf{x}^*) = 0$

Доказательство

- Если $\mathbf{x}^*$ глобальный минимум, то $\mathbf{x}^*$ локальный минимум
- Значит $f'(\mathbf{x}^*) = 0$ по предыдущей теореме
- Пусть $\mathbf{x}^*$ такая точка, что $f'(\mathbf{x}^*) = 0$ и функция выпукла
- Тогда по критерию выпуклости

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle = f(\mathbf{x}^*)$$

- Значит $\mathbf{x}^*$ – глобальный минимум.

Замечание

Также можно воспользоваться свойством существования субдифференциала для выпуклой функции и общим критерием оптимальности.

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \tag{2}$$

Теорема

Точка $\mathbf{x}^*$ *– решение задачи* (2), *где* f *– выпуклая функция, тогда и только тогда, когда* $\mathbf{x}^* \in \mathcal{X}$ *и* $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ *для всех* $\mathbf{y} \in \mathcal{X}$.

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \tag{2}$$

Теорема

Точка $\mathbf{x}^*$ – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in \mathcal{X}$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ для всех $\mathbf{y} \in \mathcal{X}$.

Доказательство

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \tag{2}$$

Теорема

Точка $\mathbf{x}^*$ *– решение задачи* (2)*, где* f *– выпуклая функция, тогда и только тогда, когда* $\mathbf{x}^* \in \mathcal{X}$ *и* $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ *для всех* $\mathbf{y} \in \mathcal{X}$*.*

Доказательство

- Пусть $\mathbf{x}^* \in \mathcal{X}$ и выполнено неравенство. Тогда по критерию первого порядка для выпуклой функции $f(\mathbf{y}) \geq f(\mathbf{x}^*)$

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \qquad (2)$$

Теорема

Точка $\mathbf{x}^*$ – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in \mathcal{X}$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ для всех $\mathbf{y} \in \mathcal{X}$.

Доказательство

- Пусть $\mathbf{x}^* \in \mathcal{X}$ и выполнено неравенство. Тогда по критерию первого порядка для выпуклой функции $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- Пусть $\mathbf{x}^*$ решение задачи (2), но найдётся $\tilde{\mathbf{y}}$ такой что $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \qquad (2)$$

Теорема

Точка $\mathbf{x}^*$ – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in \mathcal{X}$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ для всех $\mathbf{y} \in \mathcal{X}$.

Доказательство

- Пусть $\mathbf{x}^* \in \mathcal{X}$ и выполнено неравенство. Тогда по критерию первого порядка для выпуклой функции $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- Пусть $\mathbf{x}^*$ решение задачи (2), но найдётся $\tilde{\mathbf{y}}$ такой что $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$
- Рассмотрим точку $\mathbf{z}(t) = t\tilde{\mathbf{y}} + (1 - t)\mathbf{x}^*$, $t \in [0, 1]$

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \qquad (2)$$

Теорема

Точка $\mathbf{x}^*$ – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in \mathcal{X}$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^* \rangle \geq 0$ для всех $\mathbf{y} \in \mathcal{X}$.

Доказательство

- Пусть $\mathbf{x}^* \in \mathcal{X}$ и выполнено неравенство. Тогда по критерию первого порядка для выпуклой функции $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- Пусть $\mathbf{x}^*$ решение задачи (2), но найдётся $\tilde{\mathbf{y}}$ такой что $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$
- Рассмотрим точку $\mathbf{z}(t) = t\tilde{\mathbf{y}} + (1 - t)\mathbf{x}^*$, $t \in [0, 1]$
- Тогда в силу $\frac{d}{dt} f(\mathbf{z}(t))\Big|_{t=0} = \langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^* \rangle < 0$

Дифференциальное условие оптимальности для задачи с ограничениями

$$\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) \qquad (2)$$

Теорема

Точка $\mathbf{x}^*$ – решение задачи (2), где f – выпуклая функция, тогда и только тогда, когда $\mathbf{x}^* \in \mathcal{X}$ и $\langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}^*\rangle \geq 0$ для всех $\mathbf{y} \in \mathcal{X}$.

Доказательство

- Пусть $\mathbf{x}^* \in \mathcal{X}$ и выполнено неравенство. Тогда по критерию первого порядка для выпуклой функции $f(\mathbf{y}) \geq f(\mathbf{x}^*)$
- Пусть $\mathbf{x}^*$ решение задачи (2), но найдётся $\tilde{\mathbf{y}}$ такой что $\langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^*\rangle < 0$
- Рассмотрим точку $\mathbf{z}(t) = t\tilde{\mathbf{y}} + (1-t)\mathbf{x}^*$, $t \in [0, 1]$
- Тогда в силу $\frac{d}{dt} f(\mathbf{z}(t))\Big|_{t=0} = \langle f'(\mathbf{x}^*), \tilde{\mathbf{y}} - \mathbf{x}^*\rangle < 0$
- Значит для малого t выполнено $f(\mathbf{z}(t)) < f(\mathbf{x}^*)$. Противоречие.

Достаточное условие второго порядка

Теорема

Пусть f дважды непрерывно дифференцируема. Точка $\mathbf{x}^*$ удовлетворяет уравнению $f'(\mathbf{x}^*) = 0$. Если $\mathbf{s}^\top f''(\mathbf{x}^*)\mathbf{s} > 0$ для всех $\mathbf{s} \neq 0$, тогда $\mathbf{x}^*$ точка локального минимума.

Достаточное условие второго порядка

Теорема

Пусть f дважды непрерывно дифференцируема. Точка $\mathbf{x}^*$ удовлетворяет уравнению $f'(\mathbf{x}^*) = 0$. Если $\mathbf{s}^\top f''(\mathbf{x}^*)\mathbf{s} > 0$ для всех $\mathbf{s} \neq 0$, тогда $\mathbf{x}^*$ точка локального минимума.

Доказательство от противного

Достаточное условие второго порядка

Теорема

Пусть f дважды непрерывно дифференцируема. Точка $\mathbf{x}^*$ удовлетворяет уравнению $f'(\mathbf{x}^*) = 0$. Если $\mathbf{s}^\top f''(\mathbf{x}^*)\mathbf{s} > 0$ для всех $\mathbf{s} \neq 0$, тогда $\mathbf{x}^*$ точка локального минимума.

Доказательство от противного

- Пусть найдётся точка $\mathbf{y}$ близкая к $\mathbf{x}^*$, такая что $f(\mathbf{y}) < f(\mathbf{x}^*)$

Достаточное условие второго порядка

Теорема

Пусть f дважды непрерывно дифференцируема. Точка $\mathbf{x}^*$ удовлетворяет уравнению $f'(\mathbf{x}^*) = 0$. Если $\mathbf{s}^\top f''(\mathbf{x}^*)\mathbf{s} > 0$ для всех $\mathbf{s} \neq 0$, тогда $\mathbf{x}^*$ точка локального минимума.

Доказательство от противного

- Пусть найдётся точка $\mathbf{y}$ близкая к $\mathbf{x}^*$, такая что $f(\mathbf{y}) < f(\mathbf{x}^*)$
- Тогда рассмотрим разложение $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2}\langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}) \rangle + o(\|\mathbf{y} - \mathbf{x}\|_2^2)$

Достаточное условие второго порядка

Теорема

Пусть f дважды непрерывно дифференцируема. Точка $\mathbf{x}^*$ удовлетворяет уравнению $f'(\mathbf{x}^*) = 0$. Если $\mathbf{s}^\top f''(\mathbf{x}^*)\mathbf{s} > 0$ для всех $\mathbf{s} \neq 0$, тогда $\mathbf{x}^*$ точка локального минимума.

Доказательство от противного

- Пусть найдётся точка $\mathbf{y}$ близкая к $\mathbf{x}^*$, такая что $f(\mathbf{y}) < f(\mathbf{x}^*)$
- Тогда рассмотрим разложение $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}\rangle + \frac{1}{2}\langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}^*)(\mathbf{y} - \mathbf{x})\rangle + o(\|\mathbf{y} - \mathbf{x}\|_2^2)$
- Из условия стационарности следует, что $f(\mathbf{y}) - f(\mathbf{x}^*) = \frac{1}{2}\langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}^*)(\mathbf{y} - \mathbf{x})\rangle + o(\|\mathbf{y} - \mathbf{x}\|_2^2)$

Достаточное условие второго порядка

Теорема

Пусть f дважды непрерывно дифференцируема. Точка $\mathbf{x}^*$ удовлетворяет уравнению $f'(\mathbf{x}^*) = 0$. Если $\mathbf{s}^\top f''(\mathbf{x}^*)\mathbf{s} > 0$ для всех $\mathbf{s} \neq 0$, тогда $\mathbf{x}^*$ точка локального минимума.

Доказательство от противного

- Пусть найдётся точка $\mathbf{y}$ близкая к $\mathbf{x}^*$, такая что $f(\mathbf{y}) < f(\mathbf{x}^*)$
- Тогда рассмотрим разложение $f(\mathbf{y}) = f(\mathbf{x}^*) + \langle f'(\mathbf{x}^*), \mathbf{y} - \mathbf{x}\rangle + \frac{1}{2}\langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}^*)(\mathbf{y} - \mathbf{x})\rangle + o(\|\mathbf{y} - \mathbf{x}\|_2^2)$
- Из условия стационарности следует, что $f(\mathbf{y}) - f(\mathbf{x}^*) = \frac{1}{2}\langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}^*)(\mathbf{y} - \mathbf{x})\rangle + o(\|\mathbf{y} - \mathbf{x}\|_2^2)$
- Если $\mathbf{y} \to \mathbf{x}^*$, тогда у нас есть направление $\mathbf{z} \neq 0$ такое что $\mathbf{z}^\top f''(\mathbf{x}^*)\mathbf{z} \leq 0$, получили противоречие.

Сопряжённая функция

Определение

Пусть $f : \mathbb{R}^n \to \mathbb{R}$. Функция $f^* : \mathbb{R}^n \to \mathbb{R}$ называется сопряжённой функцией к функции f и определена как

$$f^*(\mathbf{y}) = \sup_{\mathbf{x} \in \mathrm{dom}\, f} (\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x})).$$

Область определения f^* — это множество таких $\mathbf{y}$, что супремум конечен.

Сопряжённая функция

Определение

Пусть $f : \mathbb{R}^n \to \mathbb{R}$. Функция $f^* : \mathbb{R}^n \to \mathbb{R}$ называется сопряжённой функцией к функции f и определена как

$$f^*(\mathbf{y}) = \sup_{\mathbf{x} \in \mathrm{dom}\, f} (\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x})).$$

Область определения f^* — это множество таких $\mathbf{y}$, что супремум конечен.

Свойства

Сопряжённая функция

Определение

Пусть $f : \mathbb{R}^n \to \mathbb{R}$. Функция $f^* : \mathbb{R}^n \to \mathbb{R}$ называется сопряжённой функцией к функции f и определена как

$$f^*(\mathbf{y}) = \sup_{\mathbf{x} \in \mathrm{dom}\, f} (\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x})).$$

Область определения f^* — это множество таких $\mathbf{y}$, что супремум конечен.

Свойства

- Выпуклая и замкнутая как супремум линейных функций, которые выпуклы и замкнуты

Сопряжённая функция

Определение

Пусть $f : \mathbb{R}^n \to \mathbb{R}$. Функция $f^* : \mathbb{R}^n \to \mathbb{R}$ называется сопряжённой функцией к функции f и определена как

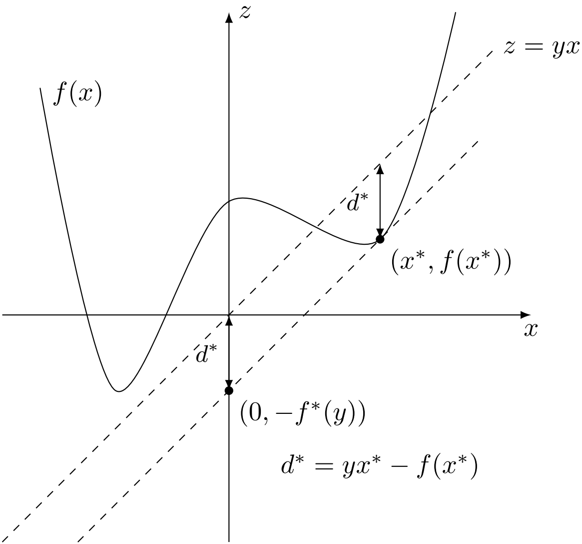
$$f^*(\mathbf{y}) = \sup_{\mathbf{x} \in \mathrm{dom}\ f} (\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x})).$$

Область определения f^* — это множество таких $\mathbf{y}$, что супремум конечен.

Свойства

- Выпуклая и замкнутая как супремум линейных функций, которые выпуклы и замкнуты
- Неравенство Юнга-Фенхеля: $\langle \mathbf{y}, \mathbf{x} \rangle \leq f^*(\mathbf{y}) + f(\mathbf{x})$

Геометрический смысл



Индикаторная и опорная функции

Пример

- Рассмотрим индикаторную функцию
$\delta_{\mathcal{X}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \mathcal{X} \\ +\infty, & \mathbf{x} \notin \mathcal{X} \end{cases}$
- Для неё сопряжённая функция
$\delta^*_{\mathcal{X}}(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - \delta_{\mathcal{X}}(\mathbf{x}) \right) = \sup_{\mathbf{x} \in \mathcal{X}} \langle \mathbf{y}, \mathbf{x} \rangle = \sigma_{\mathcal{X}}(\mathbf{y})$

Инфимальная конволюция

Определение

Пусть f_1 и f_2 выпуклые функции. Тогда функция $f(\mathbf{x}) = (f_1 \square f_2)(\mathbf{x}) = \inf\limits_{\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}} (f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2))$ называется инфимальной конволюцией функций f_1 и f_2

Инфимальная конволюция

Определение

Пусть f_1 и f_2 выпуклые функции. Тогда функция $f(\mathbf{x}) = (f_1 \square f_2)(\mathbf{x}) = \inf\limits_{\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}} (f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2))$ называется инфимальной конволюцией функций f_1 и f_2

Теорема

Инфимальная конволюция выпуклых функций есть выпуклая функция

Инфимальная конволюция

Определение

Пусть f_1 и f_2 выпуклые функции. Тогда функция $f(\mathbf{x}) = (f_1 \square f_2)(\mathbf{x}) = \inf\limits_{\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}} (f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2))$ называется инфимальной конволюцией функций f_1 и f_2

Теорема

Инфимальная конволюция выпуклых функций есть выпуклая функция

Доказательство

Инфимальная конволюция

Определение

Пусть f_1 и f_2 выпуклые функции. Тогда функция $f(\mathbf{x}) = (f_1 \Box f_2)(\mathbf{x}) = \inf\limits_{\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}} (f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2))$ называется инфимальной конволюцией функций f_1 и f_2

Теорема

Инфимальная конволюция выпуклых функций есть выпуклая функция

Доказательство

- Запишем инфимальную конволюцию в виде $(f_1 \Box f_2)(\mathbf{x}) = \inf\limits_{\mathbf{x}_2} (f_1(\mathbf{x} - \mathbf{x}_2) + f_2(\mathbf{x}_2))$

Инфимальная конволюция

Определение

Пусть f_1 и f_2 выпуклые функции. Тогда функция $f(\mathbf{x}) = (f_1 \square f_2)(\mathbf{x}) = \inf\limits_{\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}} (f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2))$ называется инфимальной конволюцией функций f_1 и f_2

Теорема

Инфимальная конволюция выпуклых функций есть выпуклая функция

Доказательство

- Запишем инфимальную конволюцию в виде $(f_1 \square f_2)(\mathbf{x}) = \inf\limits_{\mathbf{x}_2} (f_1(\mathbf{x} - \mathbf{x}_2) + f_2(\mathbf{x}_2))$
- Рассмотрим функцию $g(\mathbf{x}, \mathbf{x}_2) = f_1(\mathbf{x} - \mathbf{x}_2) + f_2(\mathbf{x}_2)$

Инфимальная конволюция

Определение

Пусть f_1 и f_2 выпуклые функции. Тогда функция $f(\mathbf{x}) = (f_1 \square f_2)(\mathbf{x}) = \inf\limits_{\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}} (f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2))$ называется инфимальной конволюцией функций f_1 и f_2

Теорема

Инфимальная конволюция выпуклых функций есть выпуклая функция

Доказательство

- Запишем инфимальную конволюцию в виде $(f_1 \square f_2)(\mathbf{x}) = \inf\limits_{\mathbf{x}_2} (f_1(\mathbf{x} - \mathbf{x}_2) + f_2(\mathbf{x}_2))$
- Рассмотрим функцию $g(\mathbf{x}, \mathbf{x}_2) = f_1(\mathbf{x} - \mathbf{x}_2) + f_2(\mathbf{x}_2)$
- Так как f_1 и f_2 выпуклы, то и g выпукла по обоим аргументам

Инфимальная конволюция

Определение

Пусть f_1 и f_2 выпуклые функции. Тогда функция $f(\mathbf{x}) = (f_1 \square f_2)(\mathbf{x}) = \inf\limits_{\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}} (f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2))$ называется инфимальной конволюцией функций f_1 и f_2

Теорема

Инфимальная конволюция выпуклых функций есть выпуклая функция

Доказательство

- Запишем инфимальную конволюцию в виде $(f_1 \square f_2)(\mathbf{x}) = \inf\limits_{\mathbf{x}_2} (f_1(\mathbf{x} - \mathbf{x}_2) + f_2(\mathbf{x}_2))$
- Рассмотрим функцию $g(\mathbf{x}, \mathbf{x}_2) = f_1(\mathbf{x} - \mathbf{x}_2) + f_2(\mathbf{x}_2)$
- Так как f_1 и f_2 выпуклы, то и g выпукла по обоим аргументам
- Значит инфимальная конволюция — это операция частичной минимизации, которая сохраняет выпуклость

Сопряжённая функция к инфимальной конволюции

$$
\begin{aligned}
&f^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x}\rangle - \inf_{\mathbf{x}_1}(f_1(\mathbf{x}_1) + f_2(\mathbf{x} - \mathbf{x}_1))) = \\
&\sup_{\mathbf{x}} \sup_{\mathbf{x}_1}(\langle \mathbf{y}, \mathbf{x}\rangle - f_1(\mathbf{x}_1) - f_2(\mathbf{x} - \mathbf{x}_1)) = \\
&\sup_{\mathbf{x}} \sup_{\mathbf{x}_1}(\langle \mathbf{y}, \mathbf{x} - \mathbf{x}_1\rangle + \langle \mathbf{y}, \mathbf{x}_1\rangle - f_1(\mathbf{x}_1) - f_2(\mathbf{x} - \mathbf{x}_1)) = \\
&\sup_{\mathbf{x}_1} \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x} - \mathbf{x}_1\rangle + \langle \mathbf{y}, \mathbf{x}_1\rangle - f_1(\mathbf{x}_1) - f_2(\mathbf{x} - \mathbf{x}_1)) = \\
&\sup_{\mathbf{x}_1}(f_2^*(\mathbf{y}) + \langle \mathbf{y}, \mathbf{x}_1\rangle - f_1(\mathbf{x}_1)) = f_2^*(\mathbf{y}) + f_1^*(\mathbf{y})
\end{aligned}
$$

Сопряжённая функция к квадрату нормы

- Рассмотрим функцию $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$

Сопряжённая функция к квадрату нормы

- Рассмотрим функцию $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$
- Тогда $f^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x} \rangle - \frac{1}{2}\|\mathbf{x}\|^2)$

Сопряжённая функция к квадрату нормы

- Рассмотрим функцию $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$
- Тогда $f^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x}\rangle - \frac{1}{2}\|\mathbf{x}\|^2)$
- Вспомним определение сопряжённой нормы $\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\|\leq 1} \mathbf{z}^\top \mathbf{x}$

Сопряжённая функция к квадрату нормы

- Рассмотрим функцию $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$
- Тогда $f^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x}\rangle - \frac{1}{2}\|\mathbf{x}\|^2)$
- Вспомним определение сопряжённой нормы $\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\|\leq 1} \mathbf{z}^\top \mathbf{x}$
- $\langle \mathbf{z}, \mathbf{x}\rangle = \|\mathbf{x}\| \left\langle \mathbf{z}, \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\rangle \leq \|\mathbf{x}\| \|\mathbf{z}\|_*$

Сопряжённая функция к квадрату нормы

- Рассмотрим функцию $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$
- Тогда $f^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x} \rangle - \frac{1}{2}\|\mathbf{x}\|^2)$
- Вспомним определение сопряжённой нормы $\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \leq 1} \mathbf{z}^\top \mathbf{x}$
- $\langle \mathbf{z}, \mathbf{x} \rangle = \|\mathbf{x}\| \left\langle \mathbf{z}, \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\rangle \leq \|\mathbf{x}\| \|\mathbf{z}\|_*$
- Тогда $f^*(\mathbf{y}) \leq \sup_{\mathbf{x}} \left(\|\mathbf{x}\| \|\mathbf{y}\|_* - \frac{1}{2}\|\mathbf{x}\|^2 \right) = \sup_x \left(x\|\mathbf{y}\|_* - \frac{1}{2}x^2 \right)$

Сопряжённая функция к квадрату нормы

- Рассмотрим функцию $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$
- Тогда $f^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x}\rangle - \frac{1}{2}\|\mathbf{x}\|^2)$
- Вспомним определение сопряжённой нормы $\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\|\leq 1} \mathbf{z}^\top \mathbf{x}$
- $\langle \mathbf{z}, \mathbf{x}\rangle = \|\mathbf{x}\| \left\langle \mathbf{z}, \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\rangle \leq \|\mathbf{x}\|\|\mathbf{z}\|_*$
- Тогда $f^*(\mathbf{y}) \leq \sup_{\mathbf{x}} \left(\|\mathbf{x}\|\|\mathbf{y}\|_* - \frac{1}{2}\|\mathbf{x}\|^2\right) = \sup_x \left(x\|\mathbf{y}\|_* - \frac{1}{2}x^2\right)$
- Супремум достигается в точке $x^* = \|\mathbf{y}\|_*$

Сопряжённая функция к квадрату нормы

- Рассмотрим функцию $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$
- Тогда $f^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x}\rangle - \frac{1}{2}\|\mathbf{x}\|^2)$
- Вспомним определение сопряжённой нормы $\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \leq 1} \mathbf{z}^\top \mathbf{x}$
- $\langle \mathbf{z}, \mathbf{x}\rangle = \|\mathbf{x}\| \left\langle \mathbf{z}, \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\rangle \leq \|\mathbf{x}\|\|\mathbf{z}\|_*$
- Тогда $f^*(\mathbf{y}) \leq \sup_{\mathbf{x}} \left(\|\mathbf{x}\|\|\mathbf{y}\|_* - \frac{1}{2}\|\mathbf{x}\|^2\right) = \sup_x \left(x\|\mathbf{y}\|_* - \frac{1}{2}x^2\right)$
- Супремум достигается в точке $x^* = \|\mathbf{y}\|_*$
- $f^*(\mathbf{y}) \leq \frac{1}{2}\|\mathbf{y}\|_*^2$

Сопряжённая функция к квадрату нормы

- Рассмотрим функцию $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$
- Тогда $f^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x} \rangle - \frac{1}{2}\|\mathbf{x}\|^2)$
- Вспомним определение сопряжённой нормы $\|\mathbf{z}\|_* = \sup_{\|\mathbf{x}\| \leq 1} \mathbf{z}^\top \mathbf{x}$
- $\langle \mathbf{z}, \mathbf{x} \rangle = \|\mathbf{x}\| \left\langle \mathbf{z}, \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\rangle \leq \|\mathbf{x}\| \|\mathbf{z}\|_*$
- Тогда $f^*(\mathbf{y}) \leq \sup_{\mathbf{x}} \left(\|\mathbf{x}\| \|\mathbf{y}\|_* - \frac{1}{2}\|\mathbf{x}\|^2 \right) = \sup_x \left(x\|\mathbf{y}\|_* - \frac{1}{2}x^2 \right)$
- Супремум достигается в точке $x^* = \|\mathbf{y}\|_*$
- $f^*(\mathbf{y}) \leq \frac{1}{2}\|\mathbf{y}\|_*^2$
- Осталось предъявить вектор, на котором достигается равенство. Этот этап оставим в качестве упражнения

Сопряжённая функция для преобразований аргумента

Сдвиг аргумента

Сопряжённая функция для преобразований аргумента

Сдвиг аргумента

- Пусть $g(\mathbf{x}) = f(\mathbf{x} - \mathbf{a})$. Найдём связь между f^* и g^*

Сопряжённая функция для преобразований аргумента

Сдвиг аргумента

- Пусть $g(\mathbf{x}) = f(\mathbf{x} - \mathbf{a})$. Найдём связь между f^* и g^*
- $g^*(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x} - \mathbf{a}) \right) \overset{\mathbf{z}=\mathbf{x}-\mathbf{a}}{=}$
  $\langle \mathbf{y}, \mathbf{a} \rangle + \sup_{\mathbf{z}} \left(\langle \mathbf{y}, \mathbf{z} \rangle - f(\mathbf{z}) \right) = \langle \mathbf{y}, \mathbf{a} \rangle + f^*(\mathbf{y})$

Сопряжённая функция для преобразований аргумента

Сдвиг аргумента

- Пусть $g(\mathbf{x}) = f(\mathbf{x} - \mathbf{a})$. Найдём связь между f^* и g^*
- $g^*(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x} - \mathbf{a}) \right) \overset{\mathbf{z}=\mathbf{x}-\mathbf{a}}{=}$
  $\langle \mathbf{y}, \mathbf{a} \rangle + \sup_{\mathbf{z}} \left(\langle \mathbf{y}, \mathbf{z} \rangle - f(\mathbf{z}) \right) = \langle \mathbf{y}, \mathbf{a} \rangle + f^*(\mathbf{y})$

Умножение на константу

Сопряжённая функция для преобразований аргумента

Сдвиг аргумента

- Пусть $g(\mathbf{x}) = f(\mathbf{x} - \mathbf{a})$. Найдём связь между f^* и g^*
- $g^*(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x} - \mathbf{a}) \right) \overset{\mathbf{z}=\mathbf{x}-\mathbf{a}}{=}$
  $\langle \mathbf{y}, \mathbf{a} \rangle + \sup_{\mathbf{z}} \left(\langle \mathbf{y}, \mathbf{z} \rangle - f(\mathbf{z}) \right) = \langle \mathbf{y}, \mathbf{a} \rangle + f^*(\mathbf{y})$

Умножение на константу

1. Пусть $g(\mathbf{x}) = \alpha f(\mathbf{x})$

Сопряжённая функция для преобразований аргумента

Сдвиг аргумента

- Пусть $g(\mathbf{x}) = f(\mathbf{x} - \mathbf{a})$. Найдём связь между f^* и g^*
- $g^*(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x} - \mathbf{a})\right) \overset{\mathbf{z}=\mathbf{x}-\mathbf{a}}{=}$
$\langle \mathbf{y}, \mathbf{a} \rangle + \sup_{\mathbf{z}} \left(\langle \mathbf{y}, \mathbf{z} \rangle - f(\mathbf{z})\right) = \langle \mathbf{y}, \mathbf{a} \rangle + f^*(\mathbf{y})$

Умножение на константу

1. Пусть $g(\mathbf{x}) = \alpha f(\mathbf{x})$
    - $g^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x} \rangle - \alpha f(\mathbf{x})) = \alpha \sup_{\mathbf{x}}(\langle \mathbf{y}/\alpha, \mathbf{x} \rangle - f(\mathbf{x})) =$
    $\alpha f^*(\mathbf{y}/\alpha)$

Сопряжённая функция для преобразований аргумента

Сдвиг аргумента

- Пусть $g(\mathbf{x}) = f(\mathbf{x} - \mathbf{a})$. Найдём связь между f^* и g^*
- $g^*(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x} - \mathbf{a}) \right) \overset{\mathbf{z}=\mathbf{x}-\mathbf{a}}{=}$ $\langle \mathbf{y}, \mathbf{a} \rangle + \sup_{\mathbf{z}} \left(\langle \mathbf{y}, \mathbf{z} \rangle - f(\mathbf{z}) \right) = \langle \mathbf{y}, \mathbf{a} \rangle + f^*(\mathbf{y})$

Умножение на константу

1. Пусть $g(\mathbf{x}) = \alpha f(\mathbf{x})$
   - $g^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x} \rangle - \alpha f(\mathbf{x})) = \alpha \sup_{\mathbf{x}}(\langle \mathbf{y}/\alpha, \mathbf{x} \rangle - f(\mathbf{x})) =$ $\alpha f^*(\mathbf{y}/\alpha)$
2. Пусть $g(\mathbf{x}) = \alpha f(\mathbf{x}/\alpha)$

Сопряжённая функция для преобразований аргумента

Сдвиг аргумента

- Пусть $g(\mathbf{x}) = f(\mathbf{x} - \mathbf{a})$. Найдём связь между f^* и g^*
- $g^*(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x} - \mathbf{a})\right) \overset{\mathbf{z}=\mathbf{x}-\mathbf{a}}{=}$
$\langle \mathbf{y}, \mathbf{a} \rangle + \sup_{\mathbf{z}} \left(\langle \mathbf{y}, \mathbf{z} \rangle - f(\mathbf{z})\right) = \langle \mathbf{y}, \mathbf{a} \rangle + f^*(\mathbf{y})$

Умножение на константу

1. Пусть $g(\mathbf{x}) = \alpha f(\mathbf{x})$
   - $g^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x} \rangle - \alpha f(\mathbf{x})) = \alpha \sup_{\mathbf{x}}(\langle \mathbf{y}/\alpha, \mathbf{x} \rangle - f(\mathbf{x})) =$
   $\alpha f^*(\mathbf{y}/\alpha)$
2. Пусть $g(\mathbf{x}) = \alpha f(\mathbf{x}/\alpha)$
   - $g^*(\mathbf{y}) = \sup_{\mathbf{x}}(\langle \mathbf{y}, \mathbf{x} \rangle - \alpha f(\mathbf{x}/\alpha)) \overset{\mathbf{z}=\mathbf{x}/\alpha}{=}$
   $\sup_{\mathbf{z}}(\alpha \langle \mathbf{y}, \mathbf{z} \rangle - \alpha f(\mathbf{z})) = \alpha f^*(\mathbf{y})$

Разделение переменных

Теорема

Если $g(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \sum_{i=1}^{k} f_i(\mathbf{x}_i)$, тогда
$g^*(\mathbf{y}_1, \ldots, \mathbf{y}_k) = \sum_{i=1}^{k} f_i^*(\mathbf{y}_i)$

Разделение переменных

Теорема

Если $g(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \sum_{i=1}^k f_i(\mathbf{x}_i)$, тогда
$g^*(\mathbf{y}_1, \ldots, \mathbf{y}_k) = \sum_{i=1}^k f_i^*(\mathbf{y}_i)$

Доказательство

Разделение переменных

Теорема

Если $g(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \sum_{i=1}^k f_i(\mathbf{x}_i)$, тогда
$g^*(\mathbf{y}_1, \ldots, \mathbf{y}_k) = \sum_{i=1}^k f_i^*(\mathbf{y}_i)$

Доказательство

- Если вставить вид функции g в определение сопряжённой функции, то взятие супремума по $\mathbf{x}$ распадётся на решение k независимых задач для каждого k

Разделение переменных

Теорема

Если $g(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \sum_{i=1}^{k} f_i(\mathbf{x}_i)$, тогда
$g^*(\mathbf{y}_1, \ldots, \mathbf{y}_k) = \sum_{i=1}^{k} f_i^*(\mathbf{y}_i)$

Доказательство

- Если вставить вид функции g в определение сопряжённой функции, то взятие супремума по $\mathbf{x}$ распадётся на решение k независимых задач для каждого k
- После разделения задач, результатом будет сумма супремумов, то есть сумма f_i^*

Разделение переменных

Теорема

Если $g(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \sum_{i=1}^{k} f_i(\mathbf{x}_i)$, тогда
$g^*(\mathbf{y}_1, \ldots, \mathbf{y}_k) = \sum_{i=1}^{k} f_i^*(\mathbf{y}_i)$

Доказательство

- Если вставить вид функции g в определение сопряжённой функции, то взятие супремума по $\mathbf{x}$ распадётся на решение k независимых задач для каждого k
- После разделения задач, результатом будет сумма супремумов, то есть сумма f_i^*
- Формальные выкладки оставлены в качестве упражнения

Дважды сопряжённая функция

Определение

Пусть $f : \mathbb{R}^n \to \mathbb{R}$. Функция $f^{**} : \mathbb{R}^n \to \mathbb{R}$ называется дважды сопряжённой функцией к функции f и определена как

$$f^{**}(\mathbf{y}) = \sup_{\mathbf{x} \in \mathrm{dom}\, f} (\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{x})).$$

Дважды сопряжённая функция

Определение

Пусть $f : \mathbb{R}^n \to \mathbb{R}$. Функция $f^{**} : \mathbb{R}^n \to \mathbb{R}$ называется дважды сопряжённой функцией к функции f и определена как

$$f^{**}(\mathbf{y}) = \sup_{\mathbf{x} \in \mathrm{dom}\ f} (\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{x})).$$

Теорема

Для любой функции f выполнено $f^{**} \leq f$

Дважды сопряжённая функция

Определение

Пусть $f : \mathbb{R}^n \to \mathbb{R}$. Функция $f^{**} : \mathbb{R}^n \to \mathbb{R}$ называется дважды сопряжённой функцией к функции f и определена как

$$f^{**}(\mathbf{y}) = \sup_{\mathbf{x} \in \mathrm{dom}\, f} (\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{x})).$$

Теорема

Для любой функции f выполнено $f^{**} \leq f$

Доказательство

Дважды сопряжённая функция

Определение

Пусть $f : \mathbb{R}^n \to \mathbb{R}$. Функция $f^{**} : \mathbb{R}^n \to \mathbb{R}$ называется дважды сопряжённой функцией к функции f и определена как

$$f^{**}(\mathbf{y}) = \sup_{\mathbf{x} \in \mathrm{dom}\ f} (\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{x})).$$

Теорема

Для любой функции f выполнено $f^{**} \leq f$

Доказательство

- Из неравенства Юнга-Фенхеля следует, что $\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{y}) \leq f(\mathbf{x})$

Дважды сопряжённая функция

Определение

Пусть $f : \mathbb{R}^n \to \mathbb{R}$. Функция $f^{**} : \mathbb{R}^n \to \mathbb{R}$ называется дважды сопряжённой функцией к функции f и определена как

$$f^{**}(\mathbf{y}) = \sup_{\mathbf{x} \in \text{dom } f} (\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{x})).$$

Теорема

Для любой функции f выполнено $f^{**} \leq f$

Доказательство

- Из неравенства Юнга-Фенхеля следует, что $\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{y}) \leq f(\mathbf{x})$
- Возьмём супремум по $\mathbf{y}$ и получим требуемое неравенство

Дважды сопряжённая функция для выпуклой функции

Теорема

Пусть f, f^*, f^{**} определены на $\mathbb{R}^n$. Тогда $f = f^{**}$ iff f выпуклая функция.

Дважды сопряжённая функция для выпуклой функции

Теорема

Пусть f, f^*, f^{**} определены на $\mathbb{R}^n$. Тогда $f = f^{**}$ iff f выпуклая функция.

Доказательство

Дважды сопряжённая функция для выпуклой функции

Теорема

Пусть f, f^*, f^{**} определены на $\mathbb{R}^n$. Тогда $f = f^{**}$ iff f выпуклая функция.

Доказательство

- Пусть $f = f^{**}$, но сопряжённая функция выпукла, а значит и f выпукла

Дважды сопряжённая функция для выпуклой функции

Теорема

Пусть f, f^*, f^{**} определены на $\mathbb{R}^n$. Тогда $f = f^{**}$ iff f выпуклая функция.

Доказательство

- Пусть $f = f^{**}$, но сопряжённая функция выпукла, а значит и f выпукла
- Пусть f выпукла, тогда

Дважды сопряжённая функция для выпуклой функции

Теорема

Пусть f, f^*, f^{**} определены на $\mathbb{R}^n$. Тогда $f = f^{**}$ iff f выпуклая функция.

Доказательство

- Пусть $f = f^{**}$, но сопряжённая функция выпукла, а значит и f выпукла
- Пусть f выпукла, тогда
  - Запишем определение в виде $f^{**}(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{x})\right) = \sup_{\mathbf{x}} \inf_{\mathbf{z}} \left(\langle \mathbf{y}, \mathbf{x} \rangle + f(\mathbf{z}) - \langle \mathbf{z}, \mathbf{x} \rangle\right) = \sup_{\mathbf{x}} \inf_{\mathbf{z}} \left(f(\mathbf{z}) + \langle \mathbf{x}, \mathbf{y} - \mathbf{z} \rangle\right)$

Дважды сопряжённая функция для выпуклой функции

Теорема

Пусть f, f^*, f^{**} определены на $\mathbb{R}^n$. Тогда $f = f^{**}$ iff f выпуклая функция.

Доказательство

- Пусть $f = f^{**}$, но сопряжённая функция выпукла, а значит и f выпукла
- Пусть f выпукла, тогда
  - Запишем определение в виде $f^{**}(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{x})\right) = \sup_{\mathbf{x}} \inf_{\mathbf{z}} \left(\langle \mathbf{y}, \mathbf{x} \rangle + f(\mathbf{z}) - \langle \mathbf{z}, \mathbf{x} \rangle\right) = \sup_{\mathbf{x}} \inf_{\mathbf{z}} \left(f(\mathbf{z}) + \langle \mathbf{x}, \mathbf{y} - \mathbf{z} \rangle\right)$
  - В силу выпуклости выполнено $f(\mathbf{z}) \geq f(\mathbf{y}) + \langle \mathbf{a}, \mathbf{z} - \mathbf{y} \rangle$, где $\mathbf{a} \in \partial f(\mathbf{y})$

Дважды сопряжённая функция для выпуклой функции

Теорема

Пусть f, f^*, f^{**} определены на $\mathbb{R}^n$. Тогда $f = f^{**}$ iff f выпуклая функция.

Доказательство

- Пусть $f = f^{**}$, но сопряжённая функция выпукла, а значит и f выпукла
- Пусть f выпукла, тогда
  - Запишем определение в виде $f^{**}(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{x})\right) = \sup_{\mathbf{x}} \inf_{\mathbf{z}} \left(\langle \mathbf{y}, \mathbf{x} \rangle + f(\mathbf{z}) - \langle \mathbf{z}, \mathbf{x} \rangle\right) = \sup_{\mathbf{x}} \inf_{\mathbf{z}} \left(f(\mathbf{z}) + \langle \mathbf{x}, \mathbf{y} - \mathbf{z} \rangle\right)$
  - В силу выпуклости выполнено $f(\mathbf{z}) \geq f(\mathbf{y}) + \langle \mathbf{a}, \mathbf{z} - \mathbf{y} \rangle$, где $\mathbf{a} \in \partial f(\mathbf{y})$
  - Тогда получим $f^{**}(\mathbf{y}) \geq f(\mathbf{y}) + \sup_{\mathbf{x}} \inf_{\mathbf{z}} \langle \mathbf{a} - \mathbf{x}, \mathbf{z} - \mathbf{y} \rangle$

Дважды сопряжённая функция для выпуклой функции

Теорема

Пусть f, f^*, f^{**} определены на $\mathbb{R}^n$. Тогда $f = f^{**}$ iff f выпуклая функция.

Доказательство

- Пусть $f = f^{**}$, но сопряжённая функция выпукла, а значит и f выпукла
- Пусть f выпукла, тогда
  - Запишем определение в виде $f^{**}(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{x})\right) = \sup_{\mathbf{x}} \inf_{\mathbf{z}} \left(\langle \mathbf{y}, \mathbf{x} \rangle + f(\mathbf{z}) - \langle \mathbf{z}, \mathbf{x} \rangle\right) = \sup_{\mathbf{x}} \inf_{\mathbf{z}} \left(f(\mathbf{z}) + \langle \mathbf{x}, \mathbf{y} - \mathbf{z} \rangle\right)$
  - В силу выпуклости выполнено $f(\mathbf{z}) \geq f(\mathbf{y}) + \langle \mathbf{a}, \mathbf{z} - \mathbf{y} \rangle$, где $\mathbf{a} \in \partial f(\mathbf{y})$
  - Тогда получим $f^{**}(\mathbf{y}) \geq f(\mathbf{y}) + \sup_{\mathbf{x}} \inf_{\mathbf{z}} \langle \mathbf{a} - \mathbf{x}, \mathbf{z} - \mathbf{y} \rangle$
  - Поскольку $\inf_{\mathbf{z}} \langle \mathbf{a} - \mathbf{x}, \mathbf{z} - \mathbf{y} \rangle = \begin{cases} 0, & \mathbf{a} = \mathbf{x} \\ -\infty, & \text{иначе} \end{cases}$

Дважды сопряжённая функция для выпуклой функции

Теорема

Пусть f, f^*, f^{**} определены на $\mathbb{R}^n$. Тогда $f = f^{**}$ iff f выпуклая функция.

Доказательство

- Пусть $f = f^{**}$, но сопряжённая функция выпукла, а значит и f выпукла
- Пусть f выпукла, тогда
  - Запишем определение в виде $f^{**}(\mathbf{y}) = \sup_{\mathbf{x}} (\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{x})) = \sup_{\mathbf{x}} \inf_{\mathbf{z}} (\langle \mathbf{y}, \mathbf{x} \rangle + f(\mathbf{z}) - \langle \mathbf{z}, \mathbf{x} \rangle) = \sup_{\mathbf{x}} \inf_{\mathbf{z}} (f(\mathbf{z}) + \langle \mathbf{x}, \mathbf{y} - \mathbf{z} \rangle)$
  - В силу выпуклости выполнено $f(\mathbf{z}) \geq f(\mathbf{y}) + \langle \mathbf{a}, \mathbf{z} - \mathbf{y} \rangle$, где $\mathbf{a} \in \partial f(\mathbf{y})$
  - Тогда получим $f^{**}(\mathbf{y}) \geq f(\mathbf{y}) + \sup_{\mathbf{x}} \inf_{\mathbf{z}} \langle \mathbf{a} - \mathbf{x}, \mathbf{z} - \mathbf{y} \rangle$
  - Поскольку $\inf_{\mathbf{z}} \langle \mathbf{a} - \mathbf{x}, \mathbf{z} - \mathbf{y} \rangle = \begin{cases} 0, & \mathbf{a} = \mathbf{x} \\ -\infty, & \text{иначе} \end{cases}$
  - Значит $f^{**}(\mathbf{y}) \geq f(\mathbf{y})$, а также $f^{**}(\mathbf{y}) \leq f(\mathbf{y}) \to f^{**} = f$

Сопряжённая функция и субдифференциал

Теорема

Пусть f выпуклая функция на $\mathbb{R}^n$. Тогда следующие условия эквивалентны

- $\langle \mathbf{x}, \mathbf{y} \rangle = f(\mathbf{x}) + f^*(\mathbf{y})$
- $\mathbf{y} \in \partial f(\mathbf{x})$
- $\mathbf{x} \in \partial f^*(\mathbf{y})$

Сопряжённая функция и субдифференциал

Теорема

Пусть f выпуклая функция на $\mathbb{R}^n$. Тогда следующие условия эквивалентны

- $\langle \mathbf{x}, \mathbf{y} \rangle = f(\mathbf{x}) + f^*(\mathbf{y})$
- $\mathbf{y} \in \partial f(\mathbf{x})$
- $\mathbf{x} \in \partial f^*(\mathbf{y})$

Доказательство

Сопряжённая функция и субдифференциал

Теорема

Пусть f выпуклая функция на $\mathbb{R}^n$. Тогда следующие условия эквивалентны

- $\langle \mathbf{x}, \mathbf{y} \rangle = f(\mathbf{x}) + f^*(\mathbf{y})$
- $\mathbf{y} \in \partial f(\mathbf{x})$
- $\mathbf{x} \in \partial f^*(\mathbf{y})$

Доказательство

- $\mathbf{y} \in \partial f(\mathbf{x})$ iff $f(\mathbf{z}) \geq f(\mathbf{x}) + \langle \mathbf{y}, \mathbf{z} - \mathbf{x} \rangle$ для всех $\mathbf{z}$

Сопряжённая функция и субдифференциал

Теорема

Пусть f выпуклая функция на $\mathbb{R}^n$. Тогда следующие условия эквивалентны

- $\langle \mathbf{x}, \mathbf{y} \rangle = f(\mathbf{x}) + f^*(\mathbf{y})$
- $\mathbf{y} \in \partial f(\mathbf{x})$
- $\mathbf{x} \in \partial f^*(\mathbf{y})$

Доказательство

- $\mathbf{y} \in \partial f(\mathbf{x})$ iff $f(\mathbf{z}) \geq f(\mathbf{x}) + \langle \mathbf{y}, \mathbf{z} - \mathbf{x} \rangle$ для всех $\mathbf{z}$
- Перепишем в виде $f(\mathbf{z}) - \langle \mathbf{y}, \mathbf{z} \rangle \geq f(\mathbf{x}) - \langle \mathbf{y}, \mathbf{x} \rangle$

Сопряжённая функция и субдифференциал

Теорема

Пусть f выпуклая функция на $\mathbb{R}^n$. Тогда следующие условия эквивалентны

- $\langle \mathbf{x}, \mathbf{y} \rangle = f(\mathbf{x}) + f^*(\mathbf{y})$
- $\mathbf{y} \in \partial f(\mathbf{x})$
- $\mathbf{x} \in \partial f^*(\mathbf{y})$

Доказательство

- $\mathbf{y} \in \partial f(\mathbf{x})$ iff $f(\mathbf{z}) \geq f(\mathbf{x}) + \langle \mathbf{y}, \mathbf{z} - \mathbf{x} \rangle$ для всех $\mathbf{z}$
- Перепишем в виде $f(\mathbf{z}) - \langle \mathbf{y}, \mathbf{z} \rangle \geq f(\mathbf{x}) - \langle \mathbf{y}, \mathbf{x} \rangle$
- Взяв супремум по $\mathbf{z}$ получим $f^*(\mathbf{z}) \leq \langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x})$

Сопряжённая функция и субдифференциал

Теорема

Пусть f выпуклая функция на $\mathbb{R}^n$. Тогда следующие условия эквивалентны

- $\langle \mathbf{x}, \mathbf{y} \rangle = f(\mathbf{x}) + f^*(\mathbf{y})$
- $\mathbf{y} \in \partial f(\mathbf{x})$
- $\mathbf{x} \in \partial f^*(\mathbf{y})$

Доказательство

- $\mathbf{y} \in \partial f(\mathbf{x})$ iff $f(\mathbf{z}) \geq f(\mathbf{x}) + \langle \mathbf{y}, \mathbf{z} - \mathbf{x} \rangle$ для всех $\mathbf{z}$
- Перепишем в виде $f(\mathbf{z}) - \langle \mathbf{y}, \mathbf{z} \rangle \geq f(\mathbf{x}) - \langle \mathbf{y}, \mathbf{x} \rangle$
- Взяв супремум по $\mathbf{z}$ получим $f^*(\mathbf{z}) \leq \langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x})$
- С учётом неравенства Юнга-Фенхеля получим, что $\langle \mathbf{x}, \mathbf{y} \rangle = f(\mathbf{x}) + f^*(\mathbf{y})$

Сопряжённая функция и субдифференциал

Теорема

Пусть f выпуклая функция на $\mathbb{R}^n$. Тогда следующие условия эквивалентны

- $\langle \mathbf{x}, \mathbf{y} \rangle = f(\mathbf{x}) + f^*(\mathbf{y})$
- $\mathbf{y} \in \partial f(\mathbf{x})$
- $\mathbf{x} \in \partial f^*(\mathbf{y})$

Доказательство

- $\mathbf{y} \in \partial f(\mathbf{x})$ iff $f(\mathbf{z}) \geq f(\mathbf{x}) + \langle \mathbf{y}, \mathbf{z} - \mathbf{x} \rangle$ для всех $\mathbf{z}$
- Перепишем в виде $f(\mathbf{z}) - \langle \mathbf{y}, \mathbf{z} \rangle \geq f(\mathbf{x}) - \langle \mathbf{y}, \mathbf{x} \rangle$
- Взяв супремум по $\mathbf{z}$ получим $f^*(\mathbf{z}) \leq \langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x})$
- С учётом неравенства Юнга-Фенхеля получим, что $\langle \mathbf{x}, \mathbf{y} \rangle = f(\mathbf{x}) + f^*(\mathbf{y})$
- Так как f выпукла, то $f^{**} = f$. Значит $\langle \mathbf{x}, \mathbf{y} \rangle = f^{**}(\mathbf{x}) + f^*(\mathbf{y})$ и $\mathbf{x} \in \partial f^*(\mathbf{y})$

Сопряжённая функция к сумме выпуклых функций

Теорема

Пусть f_1, f_2 выпуклые функции. Тогда $(f_1 + f_2)^* = f_1^* \square f_2^*$

Сопряжённая функция к сумме выпуклых функций

Теорема

Пусть f_1, f_2 выпуклые функции. Тогда $(f_1 + f_2)^* = f_1^* \Box f_2^*$

Доказательство

Сопряжённая функция к сумме выпуклых функций

Теорема

Пусть f_1, f_2 выпуклые функции. Тогда $(f_1 + f_2)^* = f_1^* \square f_2^*$

Доказательство

- Ранее показали, что $(f_1 \square f_2)^* = f_1^* + f_2^*$

Сопряжённая функция к сумме выпуклых функций

Теорема

Пусть f_1, f_2 выпуклые функции. Тогда $(f_1 + f_2)^* = f_1^* \Box f_2^*$

Доказательство

- Ранее показали, что $(f_1 \Box f_2)^* = f_1^* + f_2^*$
- Тогда выполнено $(f_1^* \Box f_2^*)^* = f_1^{**} + f_2^{**} = f_1 + f_2$

Сопряжённая функция к сумме выпуклых функций

Теорема

Пусть f_1, f_2 выпуклые функции. Тогда $(f_1 + f_2)^* = f_1^* \Box f_2^*$

Доказательство

- Ранее показали, что $(f_1 \Box f_2)^* = f_1^* + f_2^*$
- Тогда выполнено $(f_1^* \Box f_2^*)^* = f_1^{**} + f_2^{**} = f_1 + f_2$
- Так как инфимальная конволюция выпуклая функция, то $(f_1 \Box f_2)^{**} = f_1 \Box f_2$

Сопряжённая функция к сумме выпуклых функций

Теорема

Пусть f_1, f_2 выпуклые функции. Тогда $(f_1 + f_2)^* = f_1^* \square f_2^*$

Доказательство

- Ранее показали, что $(f_1 \square f_2)^* = f_1^* + f_2^*$
- Тогда выполнено $(f_1^* \square f_2^*)^* = f_1^{**} + f_2^{**} = f_1 + f_2$
- Так как инфимальная конволюция выпуклая функция, то $(f_1 \square f_2)^{**} = f_1 \square f_2$
- Тогда $f_1^* \square f_2^* = (f_1 + f_2)^*$

Свойство сильно выпуклой функции

Теорема

Если $\mathbf{x}^*$ точка минимума для сильно выпуклой функции f с константой $m > 0$, то $\mathbf{x}^*$ единственная такая точка и

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \frac{m}{2}\|\mathbf{x}^* - \mathbf{y}\|_2^2, \quad \mathbf{y} \in \mathrm{dom}(f)$$

Свойство сильно выпуклой функции

Теорема

Если $\mathbf{x}^*$ точка минимума для сильно выпуклой функции f с константой $m > 0$, то $\mathbf{x}^*$ единственная такая точка и

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \frac{m}{2}\|\mathbf{x}^* - \mathbf{y}\|_2^2, \quad \mathbf{y} \in \mathrm{dom}(f)$$

Доказательство

Свойство сильно выпуклой функции

Теорема

Если $\mathbf{x}^*$ точка минимума для сильно выпуклой функции f с константой $m > 0$, то $\mathbf{x}^*$ единственная такая точка и

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \frac{m}{2}\|\mathbf{x}^* - \mathbf{y}\|_2^2, \quad \mathbf{y} \in \mathrm{dom}(f)$$

Доказательство

- Воспользуемся фактом о том, что f сильно выпукла с константой m iff $g(\mathbf{x}) = f(\mathbf{x}) - \frac{m}{2}\|\mathbf{x}\|_2^2$ выпукла

Свойство сильно выпуклой функции

Теорема

Если $\mathbf{x}^*$ точка минимума для сильно выпуклой функции f с константой $m > 0$, то $\mathbf{x}^*$ единственная такая точка и

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \frac{m}{2}\|\mathbf{x}^* - \mathbf{y}\|_2^2, \quad \mathbf{y} \in \mathrm{dom}(f)$$

Доказательство

- Воспользуемся фактом о том, что f сильно выпукла с константой m iff $g(\mathbf{x}) = f(\mathbf{x}) - \frac{m}{2}\|\mathbf{x}\|_2^2$ выпукла
- Запишем критерий выпуклости первого порядка для g: $f(\mathbf{y}) - \frac{m}{2}\|\mathbf{y}\|_2^2 \geq f(\mathbf{x}) - \frac{m}{2}\|\mathbf{x}\|_2^2 + \langle f'(\mathbf{x}) - m\mathbf{x}, \mathbf{y} - \mathbf{x}\rangle$

Свойство сильно выпуклой функции

Теорема

Если $\mathbf{x}^*$ точка минимума для сильно выпуклой функции f с константой $m > 0$, то $\mathbf{x}^*$ единственная такая точка и

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \frac{m}{2}\|\mathbf{x}^* - \mathbf{y}\|_2^2, \quad \mathbf{y} \in \mathrm{dom}(f)$$

Доказательство

- Воспользуемся фактом о том, что f сильно выпукла с константой m iff $g(\mathbf{x}) = f(\mathbf{x}) - \frac{m}{2}\|\mathbf{x}\|_2^2$ выпукла
- Запишем критерий выпуклости первого порядка для g: $f(\mathbf{y}) - \frac{m}{2}\|\mathbf{y}\|_2^2 \geq f(\mathbf{x}) - \frac{m}{2}\|\mathbf{x}\|_2^2 + \langle f'(\mathbf{x}) - m\mathbf{x}, \mathbf{y} - \mathbf{x}\rangle$
- Неравенство перепишем в виде $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x}\rangle + \frac{m}{2}\left(\|\mathbf{y}\|_2^2 + \|\mathbf{x}\|_2^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle\right)$

Свойство сильно выпуклой функции

Теорема

Если $\mathbf{x}^*$ точка минимума для сильно выпуклой функции f с константой $m > 0$, то $\mathbf{x}^*$ единственная такая точка и

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \frac{m}{2}\|\mathbf{x}^* - \mathbf{y}\|_2^2, \quad \mathbf{y} \in \mathrm{dom}(f)$$

Доказательство

- Воспользуемся фактом о том, что f сильно выпукла с константой m iff $g(\mathbf{x}) = f(\mathbf{x}) - \frac{m}{2}\|\mathbf{x}\|_2^2$ выпукла
- Запишем критерий выпуклости первого порядка для g: $f(\mathbf{y}) - \frac{m}{2}\|\mathbf{y}\|_2^2 \geq f(\mathbf{x}) - \frac{m}{2}\|\mathbf{x}\|_2^2 + \langle f'(\mathbf{x}) - m\mathbf{x}, \mathbf{y} - \mathbf{x}\rangle$
- Неравенство перепишем в виде $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x}\rangle + \frac{m}{2}\left(\|\mathbf{y}\|_2^2 + \|\mathbf{x}\|_2^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle\right)$
- Или $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x}\rangle + \frac{m}{2}\|\mathbf{y} - \mathbf{x}\|_2^2$

Свойство сильно выпуклой функции

Теорема

Если $\mathbf{x}^*$ точка минимума для сильно выпуклой функции f с константой $m > 0$, то $\mathbf{x}^*$ единственная такая точка и

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \frac{m}{2}\|\mathbf{x}^* - \mathbf{y}\|_2^2, \quad \mathbf{y} \in \mathrm{dom}(f)$$

Доказательство

- Воспользуемся фактом о том, что f сильно выпукла с константой m iff $g(\mathbf{x}) = f(\mathbf{x}) - \frac{m}{2}\|\mathbf{x}\|_2^2$ выпукла
- Запишем критерий выпуклости первого порядка для g: $f(\mathbf{y}) - \frac{m}{2}\|\mathbf{y}\|_2^2 \geq f(\mathbf{x}) - \frac{m}{2}\|\mathbf{x}\|_2^2 + \langle f'(\mathbf{x}) - m\mathbf{x}, \mathbf{y} - \mathbf{x}\rangle$
- Неравенство перепишем в виде $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x}\rangle + \frac{m}{2}\left(\|\mathbf{y}\|_2^2 + \|\mathbf{x}\|_2^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle\right)$
- Или $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x}\rangle + \frac{m}{2}\|\mathbf{y} - \mathbf{x}\|_2^2$
- В точке минимума $f'(\mathbf{x}^*) = 0$, так что после подстановки получим требуемое неравенство

Сопряжённая функция от m-сильно выпуклой функции есть $\frac{1}{m}$-гладкая функция

Теорема

Если f сильно выпуклая функция с константой m, сопряжённая к которой

$$f^*(\mathbf{y}) = \sup_{\mathbf{x}} \left(\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})\right)$$

тогда

Сопряжённая функция от m-сильно выпуклой функции есть $\frac{1}{m}$-гладкая функция

Теорема

Если f сильно выпуклая функция с константой m, сопряжённая к которой

$$f^*(\mathbf{y}) = \sup_{\mathbf{x}} \left(\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})\right)$$

тогда

- f^* определена и дифференцируема для всех $\mathbf{y}$ и при этом

$$\nabla f^*(\mathbf{y}) = \arg\max_{\mathbf{x}} \left(\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})\right)$$

Сопряжённая функция от m-сильно выпуклой функции есть $\frac{1}{m}$-гладкая функция

Теорема

Если f сильно выпуклая функция с константой m, сопряжённая к которой

$$f^*(\mathbf{y}) = \sup_{\mathbf{x}} \left(\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})\right)$$

тогда

- f^* определена и дифференцируема для всех $\mathbf{y}$ и при этом

$$\nabla f^*(\mathbf{y}) = \arg\max_{\mathbf{x}} \left(\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})\right)$$

- $\nabla f^*(\mathbf{y})$ удовлетворяет условию Липшица с константой $\frac{1}{m}$

Доказательство (часть 1)

- Так как $f(\mathbf{x})$ сильно выпуклая функция, то у функции $\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})$ единственная точка максимума для каждого $\mathbf{y}$. Обозначим её $\mathbf{x}_y$

Доказательство (часть 1)

- Так как $f(\mathbf{x})$ сильно выпуклая функция, то у функции $\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})$ единственная точка максимума для каждого $\mathbf{y}$. Обозначим её $\mathbf{x}_y$
- В силу условия оптимальности выполнено

$$\mathbf{y} = f'(\mathbf{x}_y) \quad f^*(\mathbf{y}) = \langle \mathbf{y}, \mathbf{x}_y \rangle - f(\mathbf{x}_y)$$

Доказательство (часть 1)

- Так как $f(\mathbf{x})$ сильно выпуклая функция, то у функции $\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})$ единственная точка максимума для каждого $\mathbf{y}$. Обозначим её $\mathbf{x}_y$
- В силу условия оптимальности выполнено

$$\mathbf{y} = f'(\mathbf{x}_y) \quad f^*(\mathbf{y}) = \langle \mathbf{y}, \mathbf{x}_y \rangle - f(\mathbf{x}_y)$$

- Тогда для произвольного $\mathbf{u}$ выполнено $f^*(\mathbf{u}) = \sup_{\mathbf{v}} (\mathbf{u}^\top \mathbf{v} - f(\mathbf{v})) \geq \mathbf{u}^\top \mathbf{x}_y - f(\mathbf{x}_y) =$ $\mathbf{x}_y^\top (\mathbf{u} - \mathbf{y}) + \mathbf{x}_y^\top \mathbf{y} - f(\mathbf{x}_y) = \mathbf{x}_y^\top (\mathbf{u} - \mathbf{y}) + f^*(\mathbf{y})$

Доказательство (часть 1)

- Так как $f(\mathbf{x})$ сильно выпуклая функция, то у функции $\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})$ единственная точка максимума для каждого $\mathbf{y}$. Обозначим её $\mathbf{x}_y$
- В силу условия оптимальности выполнено

$$\mathbf{y} = f'(\mathbf{x}_y) \quad f^*(\mathbf{y}) = \langle \mathbf{y}, \mathbf{x}_y \rangle - f(\mathbf{x}_y)$$

- Тогда для произвольного $\mathbf{u}$ выполнено $f^*(\mathbf{u}) = \sup_{\mathbf{v}} (\mathbf{u}^\top \mathbf{v} - f(\mathbf{v})) \geq \mathbf{u}^\top \mathbf{x}_y - f(\mathbf{x}_y) = \mathbf{x}_y^\top (\mathbf{u} - \mathbf{y}) + \mathbf{x}_y^\top \mathbf{y} - f(\mathbf{x}_y) = \mathbf{x}_y^\top (\mathbf{u} - \mathbf{y}) + f^*(\mathbf{y})$
- То есть по определению субдифференциала $\mathbf{x}_y \in \partial f^*(\mathbf{y})$, но так как $\mathbf{x}_y$ единственная точка максимума функции $\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})$, то $\partial f^*(\mathbf{y}) = \{\mathbf{x}_y\}$ и функция f^* дифференцируема

Доказательство (часть 1)

- Так как $f(\mathbf{x})$ сильно выпуклая функция, то у функции $\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})$ единственная точка максимума для каждого $\mathbf{y}$. Обозначим её $\mathbf{x}_y$
- В силу условия оптимальности выполнено

$$\mathbf{y} = f'(\mathbf{x}_y) \quad f^*(\mathbf{y}) = \langle \mathbf{y}, \mathbf{x}_y \rangle - f(\mathbf{x}_y)$$

- Тогда для произвольного $\mathbf{u}$ выполнено $f^*(\mathbf{u}) = \sup_{\mathbf{v}} (\mathbf{u}^\top \mathbf{v} - f(\mathbf{v})) \geq \mathbf{u}^\top \mathbf{x}_y - f(\mathbf{x}_y) = \mathbf{x}_y^\top (\mathbf{u} - \mathbf{y}) + \mathbf{x}_y^\top \mathbf{y} - f(\mathbf{x}_y) = \mathbf{x}_y^\top (\mathbf{u} - \mathbf{y}) + f^*(\mathbf{y})$
- То есть по определению субдифференциала $\mathbf{x}_y \in \partial f^*(\mathbf{y})$, но так как $\mathbf{x}_y$ единственная точка максимума функции $\mathbf{y}^\top \mathbf{x} - f(\mathbf{x})$, то $\partial f^*(\mathbf{y}) = \{\mathbf{x}_y\}$ и функция f^* дифференцируема
- В итоге $\nabla f^*(\mathbf{y}) = \mathbf{x}_y$

Доказательство (часть 2)

- Рассмотрим две точки $\mathbf{u}$ и $\mathbf{v}$, в которых соответственно

$$\mathbf{x}_u = \nabla f^*(\mathbf{u}) \quad \mathbf{x}_v = \nabla f^*(\mathbf{v})$$

Доказательство (часть 2)

- Рассмотрим две точки $\mathbf{u}$ и $\mathbf{v}$, в которых соответственно

$$\mathbf{x}_u = \nabla f^*(\mathbf{u}) \quad \mathbf{x}_v = \nabla f^*(\mathbf{v})$$

- Тогда по свойству сильно выпуклой функции, приведённому выше и применённому к сильно выпуклой функции $f(\mathbf{x}) - \mathbf{y}^\top \mathbf{x}$

$$f(\mathbf{x}_u) + \mathbf{v}^\top \mathbf{x}_u \geq f(\mathbf{x}_v) + \mathbf{v}^\top \mathbf{x}_v + \frac{m}{2}\|\mathbf{x}_v - \mathbf{x}_u\|_2^2$$
$$f(\mathbf{x}_v) + \mathbf{u}^\top \mathbf{x}_v \geq f(\mathbf{x}_u) + \mathbf{u}^\top \mathbf{x}_u + \frac{m}{2}\|\mathbf{x}_v - \mathbf{x}_u\|_2^2$$

Доказательство (часть 2)

- Рассмотрим две точки $\mathbf{u}$ и $\mathbf{v}$, в которых соответственно
$$\mathbf{x}_u = \nabla f^*(\mathbf{u}) \quad \mathbf{x}_v = \nabla f^*(\mathbf{v})$$
- Тогда по свойству сильно выпуклой функции, приведённому выше и применённому к сильно выпуклой функции $f(\mathbf{x}) - \mathbf{y}^\top \mathbf{x}$
$$f(\mathbf{x}_u) + \mathbf{v}^\top \mathbf{x}_u \geq f(\mathbf{x}_v) + \mathbf{v}^\top \mathbf{x}_v + \frac{m}{2}\|\mathbf{x}_v - \mathbf{x}_u\|_2^2$$
$$f(\mathbf{x}_v) + \mathbf{u}^\top \mathbf{x}_v \geq f(\mathbf{x}_u) + \mathbf{u}^\top \mathbf{x}_u + \frac{m}{2}\|\mathbf{x}_v - \mathbf{x}_u\|_2^2$$
- Сложим оба неравенства и получим
$m\|\mathbf{x}_v - \mathbf{x}_u\|_2^2 \leq (\mathbf{x}_u - \mathbf{x}_v)^\top(\mathbf{u} - \mathbf{v}) \leq \|\mathbf{x}_u - \mathbf{x}_v\|_2\|\mathbf{u} - \mathbf{v}\|_2$
или
$\|\nabla f^*(\mathbf{v}) - \nabla f^*(\mathbf{u})\|_2 \leq \frac{1}{m}\|\mathbf{u} - \mathbf{v}\|_2$

Доказательство (часть 2)

- Рассмотрим две точки $\mathbf{u}$ и $\mathbf{v}$, в которых соответственно

$$\mathbf{x}_u = \nabla f^*(\mathbf{u}) \quad \mathbf{x}_v = \nabla f^*(\mathbf{v})$$

- Тогда по свойству сильно выпуклой функции, приведённому выше и применённому к сильно выпуклой функции $f(\mathbf{x}) - \mathbf{y}^\top \mathbf{x}$

$$f(\mathbf{x}_u) + \mathbf{v}^\top \mathbf{x}_u \geq f(\mathbf{x}_v) + \mathbf{v}^\top \mathbf{x}_v + \frac{m}{2}\|\mathbf{x}_v - \mathbf{x}_u\|_2^2$$
$$f(\mathbf{x}_v) + \mathbf{u}^\top \mathbf{x}_v \geq f(\mathbf{x}_u) + \mathbf{u}^\top \mathbf{x}_u + \frac{m}{2}\|\mathbf{x}_v - \mathbf{x}_u\|_2^2$$

- Сложим оба неравенства и получим $m\|\mathbf{x}_v - \mathbf{x}_u\|_2^2 \leq (\mathbf{x}_u - \mathbf{x}_v)^\top(\mathbf{u} - \mathbf{v}) \leq \|\mathbf{x}_u - \mathbf{x}_v\|_2\|\mathbf{u} - \mathbf{v}\|_2$ или $\|\nabla f^*(\mathbf{v}) - \nabla f^*(\mathbf{u})\|_2 \leq \frac{1}{m}\|\mathbf{u} - \mathbf{v}\|_2$
- Таким образом, $\nabla f^*(\mathbf{u})$ Липшицев с константой Липшица $\frac{1}{m}$

Почему важна эта теорема?

- $f(\mathbf{x})$ выпуклая, но *негладкая*
- Moreau-Yosida envelope $(\lambda > 0)$

$$M_{\lambda f}(\mathbf{x}) = \inf_{\mathbf{u}} (f(\mathbf{u}) + \frac{1}{2\lambda}\|\mathbf{x} - \mathbf{u}\|_2^2) = \left(f \square \frac{1}{2\lambda}\| \cdot \|_2^2 \right)(\mathbf{x})$$

- $M_{\lambda f}(\mathbf{x})$ – выпукла
- $M^*_{\lambda f}(\mathbf{y}) = f^*(\mathbf{y}) + \frac{\lambda}{2}\|\mathbf{y}\|_2^2$ – сильно выпукла с параметром λ
- $M_{\lambda f} = M^{**}_{\lambda f} = (f^* + \frac{\lambda}{2}\| \cdot \|_2^2)^*$
- Сопряжённая функция к сильно выпуклой функции является гладкой $\Rightarrow M_{\lambda f}$ – гладкая функция и

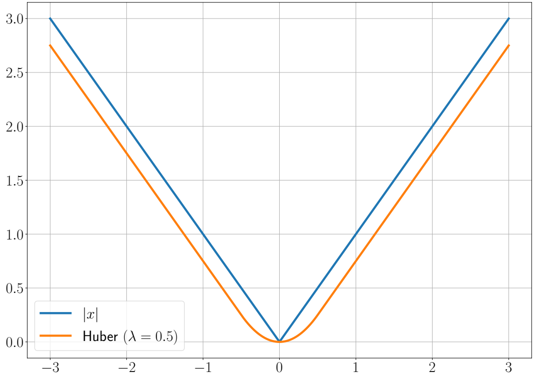
$$M'_{\lambda f}(\mathbf{x}) = \frac{1}{\lambda}(\mathbf{x} - \mathbf{u}^*), \quad \mathbf{u}^* = \arg\min_{\mathbf{u}} \left(f(\mathbf{u}) + \frac{1}{2\lambda}\|\mathbf{x} - \mathbf{u}\|_2^2 \right)$$

Важное свойство

Множество точек минимума f и $M_{\lambda f}$ совпадает.
Доказательство далее в курсе...

Пример

- $f(x) = |x|$
- $M_{\lambda f}(x) = \begin{cases} \frac{x^2}{2\lambda} & |x| \leq \lambda \\ |x| - \lambda/2 & |x| \geq \lambda \end{cases}$ — получите это выражение!



Главное

- Условия оптимальности для безусловных задач

Главное

- Условия оптимальности для безусловных задач
- Сопряжённое функция

Главное

- Условия оптимальности для безусловных задач
- Сопряжённое функция
- Свойства сопряжённых функций

Главное

- Условия оптимальности для безусловных задач
- Сопряжённое функция
- Свойства сопряжённых функций
- L-гладкость и μ-сильная выпуклость в контексте вычисления сопряжённых функций