

Методы оптимизации

Лекция 4: Выпуклые функции

Александр Катруца

Московский физико-технический институт

23 февраля 2021 г.

На прошлой лекции

- ▶ Проекция

На прошлой лекции

- ▶ Проекция
- ▶ Критерий проекции и понятие о проксимальном операторе

На прошлой лекции

- ▶ Проекция
- ▶ Критерий проекции и понятие о проксимальном операторе
- ▶ Отделимость выпуклых множеств

На прошлой лекции

- ▶ Проекция
- ▶ Критерий проекции и понятие о проксимальном операторе
- ▶ Отделимость выпуклых множеств
- ▶ Векторное дифференцирование

Выпуклая функция (convex function)

Определение

Функция $f : \mathcal{X} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ называется выпуклой (*строго выпуклой*), если \mathcal{X} — *выпуклое множество* и для

$\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$ и $\alpha \in [0, 1]$ ($\alpha \in (0, 1)$) выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Выпуклая функция (convex function)

Определение

Функция $f : \mathcal{X} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ называется выпуклой (*строго выпуклой*), если \mathcal{X} — *выпуклое множество* и для

$\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$ и $\alpha \in [0, 1]$ ($\alpha \in (0, 1)$) выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Определение

Функция f вогнута (concave), если функция $-f$ выпукла.

Выпуклая функция (convex function)

Определение

Функция $f : \mathcal{X} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ называется выпуклой (*строго выпуклой*), если \mathcal{X} — *выпуклое множество* и для

$\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$ и $\alpha \in [0, 1]$ ($\alpha \in (0, 1)$) выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq (<) \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Определение

Функция f вогнута (concave), если функция $-f$ выпукла.

Примеры выпуклых функций

- ▶ x^p для $x \geq 0$ и $p \geq 1$
- ▶ $x \log x$, где $x > 0$
- ▶ $\max\{x_1, \dots, x_n\}$
- ▶ $\|\mathbf{x}\|$
- ▶ $\log \left(\sum_{i=1}^n e^{x_i} \right)$
- ▶ $-\log \det \mathbf{X}$ для $\mathbf{X} \in \mathbf{S}_{++}^n$

Надграфик и выпуклость

Определение

Множество $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f .

Надграфик и выпуклость

Определение

Множество $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f .

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Надграфик и выпуклость

Определение

Множество $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f .

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Доказательство

Надграфик и выпуклость

Определение

Множество $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f .

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Доказательство

1. Пусть f выпуклая функция

Надграфик и выпуклость

Определение

Множество $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f .

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Доказательство

1. Пусть f выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа (\mathbf{x}_1, t_1) и (\mathbf{x}_2, t_2) , где $t_1 \geq f(\mathbf{x}_1)$ и $t_2 \geq f(\mathbf{x}_2)$

Надграфик и выпуклость

Определение

Множество $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f .

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Доказательство

1. Пусть f выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа (\mathbf{x}_1, t_1) и (\mathbf{x}_2, t_2) , где $t_1 \geq f(\mathbf{x}_1)$ и $t_2 \geq f(\mathbf{x}_2)$
- ▶ Проверим принадлежность эпиграфу точки $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha t_1 + (1 - \alpha) t_2)$

Надграфик и выпуклость

Определение

Множество $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f .

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Доказательство

1. Пусть f выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа (\mathbf{x}_1, t_1) и (\mathbf{x}_2, t_2) , где $t_1 \geq f(\mathbf{x}_1)$ и $t_2 \geq f(\mathbf{x}_2)$
- ▶ Проверим принадлежность эпиграфу точки $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha t_1 + (1 - \alpha) t_2)$
- ▶ В силу выпуклости функции
$$\alpha t_1 + (1 - \alpha) t_2 \geq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \geq f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2).$$

Надграфик и выпуклость

Определение

Множество $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f .

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Доказательство

1. Пусть f выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа (\mathbf{x}_1, t_1) и (\mathbf{x}_2, t_2) , где $t_1 \geq f(\mathbf{x}_1)$ и $t_2 \geq f(\mathbf{x}_2)$
- ▶ Проверим принадлежность эпиграфу точки $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha t_1 + (1 - \alpha) t_2)$
- ▶ В силу выпуклости функции
$$\alpha t_1 + (1 - \alpha) t_2 \geq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \geq f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2).$$

2. Пусть надграфик выпуклое множество

Надграфик и выпуклость

Определение

Множество $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f .

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Доказательство

1. Пусть f выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа (\mathbf{x}_1, t_1) и (\mathbf{x}_2, t_2) , где $t_1 \geq f(\mathbf{x}_1)$ и $t_2 \geq f(\mathbf{x}_2)$
- ▶ Проверим принадлежность эпиграфу точки $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha t_1 + (1 - \alpha) t_2)$
- ▶ В силу выпуклости функции
$$\alpha t_1 + (1 - \alpha) t_2 \geq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \geq f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2).$$

2. Пусть надграфик выпуклое множество

- ▶ $(\mathbf{x}_1, f(\mathbf{x}_1))$ и $(\mathbf{x}_2, f(\mathbf{x}_2)) \in \text{epi } f$, то $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)) \in \text{epi } f$

Надграфик и выпуклость

Определение

Множество $\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid t \geq f(\mathbf{x})\}$ называется надграфиком (эпиграфом) функции f .

Теорема

Функция f выпукла $\Leftrightarrow \text{epi } f$ выпуклое множество.

Доказательство

1. Пусть f выпуклая функция

- ▶ Рассмотрим две точки из эпиграфа (\mathbf{x}_1, t_1) и (\mathbf{x}_2, t_2) , где $t_1 \geq f(\mathbf{x}_1)$ и $t_2 \geq f(\mathbf{x}_2)$
- ▶ Проверим принадлежность эпиграфу точки $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha t_1 + (1 - \alpha) t_2)$
- ▶ В силу выпуклости функции
$$\alpha t_1 + (1 - \alpha) t_2 \geq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \geq f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2).$$

2. Пусть надграфик выпуклое множество

- ▶ $(\mathbf{x}_1, f(\mathbf{x}_1))$ и $(\mathbf{x}_2, f(\mathbf{x}_2)) \in \text{epi } f$, то $(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)) \in \text{epi } f$
- ▶ Из определения надграфика следует выпуклость f

Линии подуровня и квазивыпуклость

Определение

Пусть $f(\mathbf{x})$ функция, определённая на множестве \mathcal{X} . Тогда множество $\mathcal{L}_\beta = \{\mathbf{x} \in \mathcal{X} \mid f(\mathbf{x}) \leq \beta\}$ называется множеством подуровня (sublevel set).

Линии подуровня и квазивыпуклость

Определение

Пусть $f(\mathbf{x})$ функция, определённая на множестве \mathcal{X} . Тогда множество $\mathcal{L}_\beta = \{\mathbf{x} \in \mathcal{X} \mid f(\mathbf{x}) \leq \beta\}$ называется множеством подуровня (sublevel set).

Определение

Функция называется квазивыпуклой, если все её множества подуровня выпуклы.

Линии подуровня и квазивыпуклость

Определение

Пусть $f(\mathbf{x})$ функция, определённая на множестве \mathcal{X} . Тогда множество $\mathcal{L}_\beta = \{\mathbf{x} \in \mathcal{X} \mid f(\mathbf{x}) \leq \beta\}$ называется множеством подуровня (sublevel set).

Определение

Функция называется квазивыпуклой, если все её множества подуровня выпуклы.

Q: приведите пример квазивыпуклой, но не выпуклой функции

Сильно выпуклая функция (strongly convex function)

Определение

Функция $f : \mathcal{X} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ называется сильно выпуклой с константой $m > 0$, если \mathcal{X} — выпуклое множество и для $\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$ и $\alpha \in [0, 1]$ выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) - \frac{m}{2} \alpha (1 - \alpha) \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$$

Сильно выпуклая функция (strongly convex function)

Определение

Функция $f : \mathcal{X} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ называется сильно выпуклой с константой $m > 0$, если \mathcal{X} — выпуклое множество и для $\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$ и $\alpha \in [0, 1]$ выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) - \frac{m}{2} \alpha (1 - \alpha) \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$$

- Выпуклость \supset строгая выпуклость \supset сильная выпуклость

Сильно выпуклая функция (strongly convex function)

Определение

Функция $f : \mathcal{X} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ называется сильно выпуклой с константой $m > 0$, если \mathcal{X} — выпуклое множество и для $\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$ и $\alpha \in [0, 1]$ выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) - \frac{m}{2} \alpha (1 - \alpha) \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$$

- ▶ Выпуклость \supset строгая выпуклость \supset сильная выпуклость
- ▶ Для сильно выпуклых функций многие утверждения о методах оказываются более сильными, чем просто для выпуклых функций

Сильно выпуклая функция (strongly convex function)

Определение

Функция $f : \mathcal{X} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ называется сильно выпуклой с константой $m > 0$, если \mathcal{X} — выпуклое множество и для $\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$ и $\alpha \in [0, 1]$ выполнено:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) - \frac{m}{2} \alpha (1 - \alpha) \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$$

- ▶ Выпуклость \supset строгая выпуклость \supset сильная выпуклость
- ▶ Для сильно выпуклых функций многие утверждения о методах оказываются более сильными, чем просто для выпуклых функций
- ▶ Подробности в следующем семестре

Дифференциальные критерии

Будем считать выпуклую функцию сильно выпуклой с $m = 0$.

Дифференциальные критерии

Будем считать выпуклую функцию сильно выпуклой с $m = 0$.

Теорема

Пусть функция $f(\mathbf{x})$ дифференцируема и определена на выпуклом множестве $\mathcal{X} \subseteq \mathbb{R}^n$. Тогда $f(\mathbf{x})$ сильно выпукла с константой $m \geq 0$ в том и только том случае, если

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in \mathcal{X}.$$

Дифференциальные критерии

Будем считать выпуклую функцию сильно выпуклой с $m = 0$.

Теорема

Пусть функция $f(\mathbf{x})$ дифференцируема и определена на выпуклом множестве $\mathcal{X} \subseteq \mathbb{R}^n$. Тогда $f(\mathbf{x})$ сильно выпукла с константой $m \geq 0$ в том и только том случае, если

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in \mathcal{X}.$$

Доказательство: пусть f выпукла

- По определению:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Дифференциальные критерии

Будем считать выпуклую функцию сильно выпуклой с $m = 0$.

Теорема

Пусть функция $f(\mathbf{x})$ дифференцируема и определена на выпуклом множестве $\mathcal{X} \subseteq \mathbb{R}^n$. Тогда $f(\mathbf{x})$ сильно выпукла с константой $m \geq 0$ в том и только том случае, если

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in \mathcal{X}.$$

Доказательство: пусть f выпукла

- По определению:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

- Перепишем в виде

$$f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) \leq f(\mathbf{x}_2) + \alpha(f(\mathbf{x}_1) - f(\mathbf{x}_2)) \text{ или}$$

$$\frac{f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) - f(\mathbf{x}_2)}{\alpha} \leq f(\mathbf{x}_1) - f(\mathbf{x}_2)$$

Дифференциальные критерии

Будем считать выпуклую функцию сильно выпуклой с $m = 0$.

Теорема

Пусть функция $f(\mathbf{x})$ дифференцируема и определена на выпуклом множестве $\mathcal{X} \subseteq \mathbb{R}^n$. Тогда $f(\mathbf{x})$ сильно выпукла с константой $m \geq 0$ в том и только том случае, если

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \langle f'(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}^*\|_2^2, \quad \forall \mathbf{x}, \mathbf{x}^* \in \mathcal{X}.$$

Доказательство: пусть f выпукла

- По определению:

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

- Перепишем в виде

$$f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) \leq f(\mathbf{x}_2) + \alpha(f(\mathbf{x}_1) - f(\mathbf{x}_2)) \text{ или}$$

$$\frac{f(\mathbf{x}_2 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)) - f(\mathbf{x}_2)}{\alpha} \leq f(\mathbf{x}_1) - f(\mathbf{x}_2)$$

- При $\alpha \rightarrow 0$ получим

$$\langle f'(\mathbf{x}_2), \mathbf{x}_1 - \mathbf{x}_2 \rangle \leq f(\mathbf{x}_1) - f(\mathbf{x}_2)$$

Доказательство: пусть выполнено неравенство

- Рассмотрим $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$

Доказательство: пусть выполнено неравенство

- ▶ Рассмотрим $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$
- ▶ Запишем два неравенства для \mathbf{z}, \mathbf{x}_1 и \mathbf{z}, \mathbf{x}_2

$$f(\mathbf{x}_1) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

Доказательство: пусть выполнено неравенство

- ▶ Рассмотрим $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$
- ▶ Запишем два неравенства для \mathbf{z}, \mathbf{x}_1 и \mathbf{z}, \mathbf{x}_2

$$f(\mathbf{x}_1) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

- ▶ Сложим эти равенства

$$f(\mathbf{z}) = f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Доказательство: пусть выполнено неравенство

- ▶ Рассмотрим $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$
- ▶ Запишем два неравенства для \mathbf{z}, \mathbf{x}_1 и \mathbf{z}, \mathbf{x}_2

$$f(\mathbf{x}_1) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

- ▶ Сложим эти равенства

$$f(\mathbf{z}) = f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Сильно выпуклый случай

Для случая сильной выпуклости необходимо применить аналогичные выкладки к функции $f(\mathbf{x}) - \frac{m}{2} \|\mathbf{x}\|_2^2$.

Доказательство: пусть выполнено неравенство

- ▶ Рассмотрим $\mathbf{z} = \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$
- ▶ Запишем два неравенства для \mathbf{z}, \mathbf{x}_1 и \mathbf{z}, \mathbf{x}_2

$$f(\mathbf{x}_1) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_1 \rangle \mid \cdot \alpha$$

$$f(\mathbf{x}_2) \geq f(\mathbf{z}) + \langle f'(\mathbf{z}), \mathbf{z} - \mathbf{x}_2 \rangle \mid \cdot (1 - \alpha)$$

- ▶ Сложим эти равенства

$$f(\mathbf{z}) = f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2)$$

Сильно выпуклый случай

Для случая сильной выпуклости необходимо применить аналогичные выкладки к функции $f(\mathbf{x}) - \frac{m}{2} \|\mathbf{x}\|_2^2$.

Упражнение

Покажите, что f сильно выпукла $\Leftrightarrow f(\mathbf{x}) - \frac{m}{2} \|\mathbf{x}\|_2^2$ выпукла.

Теорема

Дважды непрерывно дифференцируемая функция f сильно выпукла с константой $m \geq 0$ iff $f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

Теорема

Дважды непрерывно дифференцируемая функция f сильно выпукла с константой $m \geq 0$ iff $f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

- Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle$$

Теорема

Дважды непрерывно дифференцируемая функция f сильно выпукла с константой $m \geq 0$ iff $f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

- Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle$$

- Если $f''(\mathbf{x}) \succeq m\mathbf{I}$, то $\frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$,
и по критерию первого порядка f выпукла

Теорема

Дважды непрерывно дифференцируемая функция f сильно выпукла с константой $m \geq 0$ iff $f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

- Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle$$

- Если $f''(\mathbf{x}) \succeq m\mathbf{I}$, то $\frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$,
и по критерию первого порядка f выпукла
- Если найдётся точка \mathbf{z} такая, что $f''(\mathbf{z}) \not\succeq m\mathbf{I}$, тогда найдётся направление \mathbf{d} такое, что $\mathbf{d}^\top f''(\mathbf{z}) \mathbf{d} < m \|\mathbf{d}\|_2^2$

Теорема

Дважды непрерывно дифференцируемая функция f сильно выпукла с константой $m \geq 0$ iff $f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

- Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle$$

- Если $f''(\mathbf{x}) \succeq m\mathbf{I}$, то $\frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$,
и по критерию первого порядка f выпукла
- Если найдётся точка \mathbf{z} такая, что $f''(\mathbf{z}) \not\succeq m\mathbf{I}$, тогда найдётся направление \mathbf{d} такое, что $\mathbf{d}^\top f''(\mathbf{z}) \mathbf{d} < m \|\mathbf{d}\|_2^2$
- В таком случае в силу критерия первого порядка f невыпукла — противоречие

Теорема

Дважды непрерывно дифференцируемая функция f сильно выпукла с константой $m \geq 0$ iff $f''(\mathbf{x}) \succeq m\mathbf{I}$

Доказательство

- Разложение в ряд Тейлора до второго порядка

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle$$

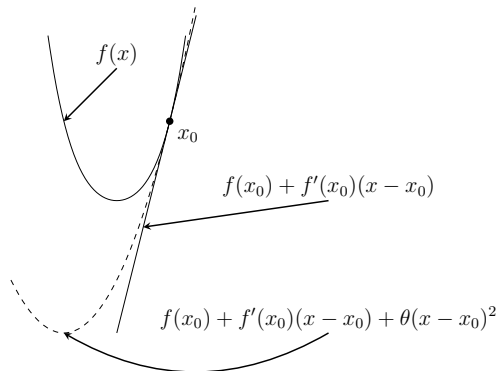
- Если $f''(\mathbf{x}) \succeq m\mathbf{I}$, то $\frac{1}{2} \langle \mathbf{y} - \mathbf{x}, f''(\mathbf{x}_\alpha)(\mathbf{y} - \mathbf{x}) \rangle \geq \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$, и по критерию первого порядка f выпукла
- Если найдётся точка \mathbf{z} такая, что $f''(\mathbf{z}) \not\succeq m\mathbf{I}$, тогда найдётся направление \mathbf{d} такое, что $\mathbf{d}^\top f''(\mathbf{z})\mathbf{d} < m\|\mathbf{d}\|_2^2$
- В таком случае в силу критерия первого порядка f невыпукла — противоречие

Напоминание

Для проверки определённости матрицы нужно использовать определение или критерий Сильвестра

Иллюстрация дифференциальных критериев

Пусть $f(x) = x^2 + \frac{1}{2}x^4$



- ▶ Линейная глобальная оценка снизу для **выпуклой функции**
- ▶ Квадратичная глобальная оценка снизу для **сильно выпуклой функции**

Ещё один критерий выпуклости первого порядка

Теорема

Пусть $f(\mathbf{x})$ дифференцируемая функция на множестве \mathcal{X} . Она является сильно выпуклой с константой $m \geq 0$ iff

$$\langle f'(\mathbf{x}) - f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \geq m \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

Ещё один критерий выпуклости первого порядка

Теорема

Пусть $f(\mathbf{x})$ дифференцируемая функция на множестве \mathcal{X} . Она является сильно выпуклой с константой $m \geq 0$ iff

$$\langle f'(\mathbf{x}) - f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \geq m \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

Доказательство

Ещё один критерий выпуклости первого порядка

Теорема

Пусть $f(\mathbf{x})$ дифференцируемая функция на множестве \mathcal{X} . Она является сильно выпуклой с константой $m \geq 0$ iff

$$\langle f'(\mathbf{x}) - f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \geq m \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

Доказательство

1. Пусть выполнено неравенство

Ещё один критерий выпуклости первого порядка

Теорема

Пусть $f(\mathbf{x})$ дифференцируемая функция на множестве \mathcal{X} . Она является сильно выпуклой с константой $m \geq 0$ iff

$$\langle f'(\mathbf{x}) - f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \geq m \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

Доказательство

1. Пусть выполнено неравенство

► По формуле Ньютона-Лейбница

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \int_0^1 \langle f'(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)), \mathbf{x} - \mathbf{x}_0 \rangle dt$$

Ещё один критерий выпуклости первого порядка

Теорема

Пусть $f(\mathbf{x})$ дифференцируемая функция на множестве \mathcal{X} . Она является сильно выпуклой с константой $m \geq 0$ iff

$$\langle f'(\mathbf{x}) - f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \geq m \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

Доказательство

1. Пусть выполнено неравенство

- По формуле Ньютона-Лейбница

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \int_0^1 \langle f'(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)), \mathbf{x} - \mathbf{x}_0 \rangle dt$$

- Тогда $f(\mathbf{x}) - f(\mathbf{x}_0) - \langle f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle =$

$$\int_0^1 \langle f'(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)), \mathbf{x} - \mathbf{x}_0 \rangle dt - \langle f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle =$$

$$\int_0^1 \frac{1}{t} \langle f'(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) - f'(\mathbf{x}_0), t(\mathbf{x} - \mathbf{x}_0) \rangle dt \geq$$

$$\int_0^1 \frac{1}{t} m t^2 \|\mathbf{x} - \mathbf{x}_0\|_2^2 dt = \frac{m}{2} \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

Ещё один критерий выпуклости первого порядка

Теорема

Пусть $f(\mathbf{x})$ дифференцируемая функция на множестве \mathcal{X} . Она является сильно выпуклой с константой $m \geq 0$ iff

$$\langle f'(\mathbf{x}) - f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \geq m \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

Доказательство

1. Пусть выполнено неравенство

- По формуле Ньютона-Лейбница

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \int_0^1 \langle f'(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)), \mathbf{x} - \mathbf{x}_0 \rangle dt$$

- Тогда $f(\mathbf{x}) - f(\mathbf{x}_0) - \langle f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle =$

$$\int_0^1 \langle f'(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)), \mathbf{x} - \mathbf{x}_0 \rangle dt - \langle f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle =$$

$$\int_0^1 \frac{1}{t} \langle f'(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) - f'(\mathbf{x}_0), t(\mathbf{x} - \mathbf{x}_0) \rangle dt \geq$$

$$\int_0^1 \frac{1}{t} m t^2 \|\mathbf{x} - \mathbf{x}_0\|_2^2 dt = \frac{m}{2} \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

- По критерию первого порядка функция сильно выпукла

2. Пусть функция сильно выпукла

2. Пусть функция сильно выпукла

- ▶ Запишем критерий первого порядка для пары \mathbf{x}, \mathbf{x}_0 и для пары \mathbf{x}_0, \mathbf{x}

$$f(\mathbf{x}) - f(\mathbf{x}_0) \geq \langle f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

$$f(\mathbf{x}_0) - f(\mathbf{x}) \geq \langle f'(\mathbf{x}), \mathbf{x}_0 - \mathbf{x} \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

2. Пусть функция сильно выпукла

- ▶ Запишем критерий первого порядка для пары \mathbf{x}, \mathbf{x}_0 и для пары \mathbf{x}_0, \mathbf{x}

$$f(\mathbf{x}) - f(\mathbf{x}_0) \geq \langle f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

$$f(\mathbf{x}_0) - f(\mathbf{x}) \geq \langle f'(\mathbf{x}), \mathbf{x}_0 - \mathbf{x} \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

- ▶ Сложим эти неравенства

$$0 \geq \langle f'(\mathbf{x}_0) - f'(\mathbf{x}), \mathbf{x} - \mathbf{x}_0 \rangle + m \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

2. Пусть функция сильно выпукла

- ▶ Запишем критерий первого порядка для пары \mathbf{x}, \mathbf{x}_0 и для пары \mathbf{x}_0, \mathbf{x}

$$f(\mathbf{x}) - f(\mathbf{x}_0) \geq \langle f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

$$f(\mathbf{x}_0) - f(\mathbf{x}) \geq \langle f'(\mathbf{x}), \mathbf{x}_0 - \mathbf{x} \rangle + \frac{m}{2} \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

- ▶ Сложим эти неравенства

$$0 \geq \langle f'(\mathbf{x}_0) - f'(\mathbf{x}), \mathbf{x} - \mathbf{x}_0 \rangle + m \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

- ▶ Перенесём в левую часть

$$\langle f'(\mathbf{x}) - f'(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle \geq m \|\mathbf{x} - \mathbf{x}_0\|_2^2$$

Выпуклые композиции функций

- ▶ Если $f(\mathbf{x})$ — выпукла, то $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$ также выпукла

Выпуклые композиции функций

- ▶ Если $f(\mathbf{x})$ — выпукла, то $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$ также выпукла
- ▶ Если f_i — выпуклы, то $\max_{i=1,\dots,m} f_i$ также выпукла

Выпуклые композиции функций

- ▶ Если $f(\mathbf{x})$ — выпукла, то $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$ также выпукла
- ▶ Если f_i — выпуклы, то $\max_{i=1,\dots,m} f_i$ также выпукла
- ▶ Сумма выпуклых функций с неотрицательными коэффициентами — выпуклая функция

Выпуклые композиции функций

- ▶ Если $f(\mathbf{x})$ — выпукла, то $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$ также выпукла
- ▶ Если f_i — выпуклы, то $\max_{i=1,\dots,m} f_i$ также выпукла
- ▶ Сумма выпуклых функций с неотрицательными коэффициентами — выпуклая функция
- ▶ Скалярная композиция $h(f(\mathbf{x}))$

Сохранение выпуклости при взятии инфимума

Теорема

Пусть $f(\mathbf{x}, \mathbf{y})$ выпукла относительно обоих аргументов, множество \mathcal{C} выпукло. Тогда функция $g(\mathbf{x}) = \inf_{\mathbf{y} \in \mathcal{C}} f(\mathbf{x}, \mathbf{y})$ также выпукла.

Сохранение выпуклости при взятии инфимума

Теорема

Пусть $f(\mathbf{x}, \mathbf{y})$ выпукла относительно обоих аргументов, множество \mathcal{C} выпукло. Тогда функция $g(\mathbf{x}) = \inf_{\mathbf{y} \in \mathcal{C}} f(\mathbf{x}, \mathbf{y})$ также выпукла.

Доказательство

Сохранение выпуклости при взятии инфимума

Теорема

Пусть $f(\mathbf{x}, \mathbf{y})$ выпукла относительно обоих аргументов, множество \mathcal{C} выпукло. Тогда функция $g(\mathbf{x}) = \inf_{\mathbf{y} \in \mathcal{C}} f(\mathbf{x}, \mathbf{y})$ также выпукла.

Доказательство

- ▶ Пусть $\mathbf{x}_1, \mathbf{x}_2 \in \text{dom } g$. Тогда для любого $\varepsilon > 0$ найдётся точки $\mathbf{y}_1, \mathbf{y}_2$ такие что $f(\mathbf{x}_1, \mathbf{y}_1) \leq g(\mathbf{x}_1) + \varepsilon$ и $f(\mathbf{x}_2, \mathbf{y}_2) \leq g(\mathbf{x}_2) + \varepsilon$

Сохранение выпуклости при взятии инфимума

Теорема

Пусть $f(\mathbf{x}, \mathbf{y})$ выпукла относительно обоих аргументов, множество \mathcal{C} выпукло. Тогда функция $g(\mathbf{x}) = \inf_{\mathbf{y} \in \mathcal{C}} f(\mathbf{x}, \mathbf{y})$ также выпукла.

Доказательство

- ▶ Пусть $\mathbf{x}_1, \mathbf{x}_2 \in \text{dom } g$. Тогда для любого $\varepsilon > 0$ найдётся точки $\mathbf{y}_1, \mathbf{y}_2$ такие что $f(\mathbf{x}_1, \mathbf{y}_1) \leq g(\mathbf{x}_1) + \varepsilon$ и $f(\mathbf{x}_2, \mathbf{y}_2) \leq g(\mathbf{x}_2) + \varepsilon$
- ▶ Рассмотрим точку $\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$:
$$\begin{aligned} g(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) &= \inf_{\mathbf{y} \in \mathcal{C}} f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \mathbf{y}) \leq \\ &f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha \mathbf{y}_1 + (1 - \alpha) \mathbf{y}_2) \leq \\ &\alpha f(\mathbf{x}_1, \mathbf{y}_1) + (1 - \alpha) f(\mathbf{x}_2, \mathbf{y}_2) \leq \alpha g(\mathbf{x}_1) + (1 - \alpha) g(\mathbf{x}_2) + \varepsilon \end{aligned}$$

Сохранение выпуклости при взятии инфимума

Теорема

Пусть $f(\mathbf{x}, \mathbf{y})$ выпукла относительно обоих аргументов, множество \mathcal{C} выпукло. Тогда функция $g(\mathbf{x}) = \inf_{\mathbf{y} \in \mathcal{C}} f(\mathbf{x}, \mathbf{y})$ также выпукла.

Доказательство

- ▶ Пусть $\mathbf{x}_1, \mathbf{x}_2 \in \text{dom } g$. Тогда для любого $\varepsilon > 0$ найдётся точки $\mathbf{y}_1, \mathbf{y}_2$ такие что $f(\mathbf{x}_1, \mathbf{y}_1) \leq g(\mathbf{x}_1) + \varepsilon$ и $f(\mathbf{x}_2, \mathbf{y}_2) \leq g(\mathbf{x}_2) + \varepsilon$
- ▶ Рассмотрим точку $\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$:
$$\begin{aligned} g(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) &= \inf_{\mathbf{y} \in \mathcal{C}} f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \mathbf{y}) \leq \\ &f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \alpha \mathbf{y}_1 + (1 - \alpha) \mathbf{y}_2) \leq \\ &\alpha f(\mathbf{x}_1, \mathbf{y}_1) + (1 - \alpha) f(\mathbf{x}_2, \mathbf{y}_2) \leq \alpha g(\mathbf{x}_1) + (1 - \alpha) g(\mathbf{x}_2) + \varepsilon \end{aligned}$$
- ▶ Так как это неравенство выполнено для любого $\varepsilon > 0$, то $g(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha g(\mathbf{x}_1) + (1 - \alpha) g(\mathbf{x}_2)$

Ограничение на прямую

Теорема

Функция f выпукла iff функция $g(t) = f(\mathbf{x} + t\mathbf{y})$ выпукла на множестве $\{t \mid \mathbf{x} + t\mathbf{y} \in \text{dom } f\}$, где $\mathbf{x} \in \text{dom } f$ и \mathbf{y} произвольное направление

Ограничение на прямую

Теорема

Функция f выпукла iff функция $g(t) = f(\mathbf{x} + t\mathbf{y})$ выпукла на множестве $\{t \mid \mathbf{x} + t\mathbf{y} \in \text{dom } f\}$, где $\mathbf{x} \in \text{dom } f$ и \mathbf{y} произвольное направление

Доказательство

Ограничение на прямую

Теорема

Функция f выпукла iff функция $g(t) = f(\mathbf{x} + t\mathbf{y})$ выпукла на множестве $\{t \mid \mathbf{x} + t\mathbf{y} \in \text{dom } f\}$, где $\mathbf{x} \in \text{dom } f$ и \mathbf{y} произвольное направление

Доказательство

1. Пусть f выпукла

Ограничение на прямую

Теорема

Функция f выпукла iff функция $g(t) = f(\mathbf{x} + t\mathbf{y})$ выпукла на множестве $\{t \mid \mathbf{x} + t\mathbf{y} \in \text{dom } f\}$, где $\mathbf{x} \in \text{dom } f$ и \mathbf{y} произвольное направление

Доказательство

1. Пусть f выпукла

- Возьмём точки t_1 , t_2 и $t_\alpha = \alpha t_1 + (1 - \alpha)t_2$, $\alpha \in [0, 1]$

Ограничение на прямую

Теорема

Функция f выпукла iff функция $g(t) = f(\mathbf{x} + t\mathbf{y})$ выпукла на множестве $\{t \mid \mathbf{x} + t\mathbf{y} \in \text{dom } f\}$, где $\mathbf{x} \in \text{dom } f$ и \mathbf{y} произвольное направление

Доказательство

1. Пусть f выпукла

- ▶ Возьмём точки t_1, t_2 и $t_\alpha = \alpha t_1 + (1 - \alpha)t_2, \alpha \in [0, 1]$
- ▶ Тогда $g(\alpha t_1 + (1 - \alpha)t_2) = f(\mathbf{x} + (\alpha t_1 + (1 - \alpha)t_2)\mathbf{y}) = f(\alpha(\mathbf{x} + t_1\mathbf{y}) + (1 - \alpha)(\mathbf{x} + t_2\mathbf{y})) \leq \alpha f(\mathbf{x} + t_1\mathbf{y}) + (1 - \alpha)f(\mathbf{x} + t_2\mathbf{y}) = \alpha g(t_1) + (1 - \alpha)g(t_2)$

Ограничение на прямую

Теорема

Функция f выпукла iff функция $g(t) = f(\mathbf{x} + t\mathbf{y})$ выпукла на множестве $\{t \mid \mathbf{x} + t\mathbf{y} \in \text{dom } f\}$, где $\mathbf{x} \in \text{dom } f$ и \mathbf{y} произвольное направление

Доказательство

1. Пусть f выпукла

- ▶ Возьмём точки t_1, t_2 и $t_\alpha = \alpha t_1 + (1 - \alpha)t_2, \alpha \in [0, 1]$
- ▶ Тогда $g(\alpha t_1 + (1 - \alpha)t_2) = f(\mathbf{x} + (\alpha t_1 + (1 - \alpha)t_2)\mathbf{y}) = f(\alpha(\mathbf{x} + t_1\mathbf{y}) + (1 - \alpha)(\mathbf{x} + t_2\mathbf{y})) \leq \alpha f(\mathbf{x} + t_1\mathbf{y}) + (1 - \alpha)f(\mathbf{x} + t_2\mathbf{y}) = \alpha g(t_1) + (1 - \alpha)g(t_2)$

2. Пусть g выпукла

Ограничение на прямую

Теорема

Функция f выпукла iff функция $g(t) = f(\mathbf{x} + t\mathbf{y})$ выпукла на множестве $\{t \mid \mathbf{x} + t\mathbf{y} \in \text{dom } f\}$, где $\mathbf{x} \in \text{dom } f$ и \mathbf{y} произвольное направление

Доказательство

1. Пусть f выпукла

- ▶ Возьмём точки t_1, t_2 и $t_\alpha = \alpha t_1 + (1 - \alpha)t_2, \alpha \in [0, 1]$
- ▶ Тогда $g(\alpha t_1 + (1 - \alpha)t_2) = f(\mathbf{x} + (\alpha t_1 + (1 - \alpha)t_2)\mathbf{y}) = f(\alpha(\mathbf{x} + t_1\mathbf{y}) + (1 - \alpha)(\mathbf{x} + t_2\mathbf{y})) \leq \alpha f(\mathbf{x} + t_1\mathbf{y}) + (1 - \alpha)f(\mathbf{x} + t_2\mathbf{y}) = \alpha g(t_1) + (1 - \alpha)g(t_2)$

2. Пусть g выпукла

- ▶ Возьмём произвольную пару точек $\mathbf{x}_1, \mathbf{x}_2$, тогда $\mathbf{y} = \mathbf{x}_2 - \mathbf{x}_1$

Ограничение на прямую

Теорема

Функция f выпукла iff функция $g(t) = f(\mathbf{x} + t\mathbf{y})$ выпукла на множестве $\{t \mid \mathbf{x} + t\mathbf{y} \in \text{dom } f\}$, где $\mathbf{x} \in \text{dom } f$ и \mathbf{y} произвольное направление

Доказательство

1. Пусть f выпукла

- ▶ Возьмём точки t_1, t_2 и $t_\alpha = \alpha t_1 + (1 - \alpha)t_2, \alpha \in [0, 1]$
- ▶ Тогда $g(\alpha t_1 + (1 - \alpha)t_2) = f(\mathbf{x} + (\alpha t_1 + (1 - \alpha)t_2)\mathbf{y}) = f(\alpha(\mathbf{x} + t_1\mathbf{y}) + (1 - \alpha)(\mathbf{x} + t_2\mathbf{y})) \leq \alpha f(\mathbf{x} + t_1\mathbf{y}) + (1 - \alpha)f(\mathbf{x} + t_2\mathbf{y}) = \alpha g(t_1) + (1 - \alpha)g(t_2)$

2. Пусть g выпукла

- ▶ Возьмём произвольную пару точек $\mathbf{x}_1, \mathbf{x}_2$, тогда $\mathbf{y} = \mathbf{x}_2 - \mathbf{x}_1$
- ▶ $f(\mathbf{x}_1) = g(\mathbf{x}_1 + 0 \cdot \mathbf{y}) = g(0)$ и $f(\mathbf{x}_2) = g(\mathbf{x}_1 + 1 \cdot \mathbf{y}) = g(1)$

Ограничение на прямую

Теорема

Функция f выпукла iff функция $g(t) = f(\mathbf{x} + t\mathbf{y})$ выпукла на множестве $\{t \mid \mathbf{x} + t\mathbf{y} \in \text{dom } f\}$, где $\mathbf{x} \in \text{dom } f$ и \mathbf{y} произвольное направление

Доказательство

1. Пусть f выпукла

- ▶ Возьмём точки t_1, t_2 и $t_\alpha = \alpha t_1 + (1 - \alpha)t_2, \alpha \in [0, 1]$
- ▶ Тогда $g(\alpha t_1 + (1 - \alpha)t_2) = f(\mathbf{x} + (\alpha t_1 + (1 - \alpha)t_2)\mathbf{y}) = f(\alpha(\mathbf{x} + t_1\mathbf{y}) + (1 - \alpha)(\mathbf{x} + t_2\mathbf{y})) \leq \alpha f(\mathbf{x} + t_1\mathbf{y}) + (1 - \alpha)f(\mathbf{x} + t_2\mathbf{y}) = \alpha g(t_1) + (1 - \alpha)g(t_2)$

2. Пусть g выпукла

- ▶ Возьмём произвольную пару точек $\mathbf{x}_1, \mathbf{x}_2$, тогда $\mathbf{y} = \mathbf{x}_2 - \mathbf{x}_1$
- ▶ $f(\mathbf{x}_1) = g(\mathbf{x}_1 + 0 \cdot \mathbf{y}) = g(0)$ и $f(\mathbf{x}_2) = g(\mathbf{x}_1 + 1 \cdot \mathbf{y}) = g(1)$
- ▶ $f(\alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2) = f(\mathbf{x}_1 + (1 - \alpha)(\mathbf{x}_2 - \mathbf{x}_1)) = f(\mathbf{x}_1 + (0 \cdot \alpha + 1 \cdot (1 - \alpha))\mathbf{y}) = g(0 \cdot \alpha + 1 \cdot (1 - \alpha)) \leq \alpha g(0) + (1 - \alpha)g(1) = \alpha f(\mathbf{x}_1) + (1 - \alpha)f(\mathbf{x}_2)$

Локальный минимум является глобальным минимумом

Теорема

Если f выпуклая функция и x^* *локальный* минимум, то x^* — *глобальный* минимум.

Доказательство от противного

Локальный минимум является глобальным минимумом

Теорема

Если f выпуклая функция и x^* *локальный* минимум, то x^* — *глобальный* минимум.

Доказательство от противного

- ▶ Пусть $y^* \neq x^*$ — глобальный минимум: $f(y^*) < f(x^*)$

Локальный минимум является глобальным минимумом

Теорема

Если f выпуклая функция и \mathbf{x}^* *локальный* минимум, то \mathbf{x}^* — *глобальный* минимум.

Доказательство от противного

- ▶ Пусть $\mathbf{y}^* \neq \mathbf{x}^*$ — глобальный минимум: $f(\mathbf{y}^*) < f(\mathbf{x}^*)$
- ▶ По определению локального минимума: $f(\mathbf{x}^*) \leq f(\mathbf{x})$, где $\|\mathbf{x}^* - \mathbf{x}\|_2 \leq \delta$

Локальный минимум является глобальным минимумом

Теорема

Если f выпуклая функция и \mathbf{x}^* **локальный** минимум, то \mathbf{x}^* — **глобальный** минимум.

Доказательство от противного

- ▶ Пусть $\mathbf{y}^* \neq \mathbf{x}^*$ — глобальный минимум: $f(\mathbf{y}^*) < f(\mathbf{x}^*)$
- ▶ По определению локального минимума: $f(\mathbf{x}^*) \leq f(\mathbf{x})$, где $\|\mathbf{x}^* - \mathbf{x}\|_2 \leq \delta$
- ▶ Выберем достаточно малое $\alpha \in (0, 1)$ и рассмотрим точку $\mathbf{z} = (1 - \alpha)\mathbf{x}^* + \alpha\mathbf{y}^*$ такую что $\|\mathbf{z} - \mathbf{x}^*\|_2 \leq \delta$

Локальный минимум является глобальным минимумом

Теорема

Если f выпуклая функция и \mathbf{x}^* **локальный** минимум, то \mathbf{x}^* — **глобальный** минимум.

Доказательство от противного

- ▶ Пусть $\mathbf{y}^* \neq \mathbf{x}^*$ — глобальный минимум: $f(\mathbf{y}^*) < f(\mathbf{x}^*)$
- ▶ По определению локального минимума: $f(\mathbf{x}^*) \leq f(\mathbf{x})$, где $\|\mathbf{x}^* - \mathbf{x}\|_2 \leq \delta$
- ▶ Выберем достаточно малое $\alpha \in (0, 1)$ и рассмотрим точку $\mathbf{z} = (1 - \alpha)\mathbf{x}^* + \alpha\mathbf{y}^*$ такую что $\|\mathbf{z} - \mathbf{x}^*\|_2 \leq \delta$
- ▶ $f(\mathbf{x}^*) \leq f(\mathbf{z}) \leq \alpha f(\mathbf{y}^*) + (1 - \alpha)f(\mathbf{x}^*) < f(\mathbf{x}^*)$

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

- ▶ \mathcal{C}^n выпукло

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

- ▶ \mathcal{C}^n выпукло
- ▶ $\mathbf{S}_+^n \subset \mathcal{C}^n$

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

- ▶ \mathcal{C}^n выпукло
- ▶ $\mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки $\mathbf{X} \notin \mathcal{C}^n$ является co-NP полной!

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

- ▶ \mathcal{C}^n выпукло
- ▶ $\mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки $\mathbf{X} \notin \mathcal{C}^n$ является co-NP полной!
- ▶ Задача конической оптимизации с конусом \mathcal{C}^n является NP-трудной

Пример сложной выпуклой задачи

Определение

Множество $\mathcal{C}^n = \{\mathbf{A} \in \mathbf{S}^n \mid \mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0, \mathbf{x} \geq 0\}$ называется *copositive cone*.

- ▶ \mathcal{C}^n выпукло
- ▶ $\mathbf{S}_+^n \subset \mathcal{C}^n$
- ▶ Задача проверки $\mathbf{X} \notin \mathcal{C}^n$ является co-NP полной!
- ▶ Задача конической оптимизации с конусом \mathcal{C}^n является NP-трудной

Пример

Задача определения максимального независимого множества вершин графа сводится к задаче оптимизации на множестве \mathcal{C}^n . Подробности [тут](#)

Пример простой невыпуклой задачи

Дана матрица $\mathbf{Q} \in \mathbf{S}_{++}^n$.

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

Пример простой невыпуклой задачи

Дана матрица $\mathbf{Q} \in \mathbf{S}_{++}^n$.

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

- ▶ Допустимое множество невыпукло

Пример простой невыпуклой задачи

Дана матрица $\mathbf{Q} \in \mathbf{S}_{++}^n$.

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

- ▶ Допустимое множество невыпукло
- ▶ Целевая функция выпукла

Пример простой невыпуклой задачи

Дана матрица $\mathbf{Q} \in \mathbf{S}_{++}^n$.

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

- ▶ Допустимое множество невыпукло
- ▶ Целевая функция выпукла

Q: какая интерпретация у \mathbf{x}^* и $f(\mathbf{x}^*)$?

Степенной метод (Power method)

- ▶ Степенной метод для поиска собственного вектора для максимального по модулю собственного значения

$$\mathbf{x}_{k+1} = \frac{\mathbf{A}\mathbf{x}_k}{\|\mathbf{A}\mathbf{x}_k\|_2}, \quad \hat{\lambda}_{k+1} = \mathbf{x}_{k+1}^\top \mathbf{A}\mathbf{x}_{k+1}$$

Степенной метод (Power method)

- ▶ Степенной метод для поиска собственного вектора для максимального по модулю собственного значения

$$\mathbf{x}_{k+1} = \frac{\mathbf{A}\mathbf{x}_k}{\|\mathbf{A}\mathbf{x}_k\|_2}, \quad \hat{\lambda}_{k+1} = \mathbf{x}_{k+1}^\top \mathbf{A} \mathbf{x}_{k+1}$$

- ▶ \mathbf{A} – диагонализуема, поэтому есть базис из собственных векторов $[\mathbf{v}_1, \dots, \mathbf{v}_n]$ причём предположим $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$

Степенной метод (Power method)

- ▶ Степенной метод для поиска собственного вектора для максимального по модулю собственного значения

$$\mathbf{x}_{k+1} = \frac{\mathbf{A}\mathbf{x}_k}{\|\mathbf{A}\mathbf{x}_k\|_2}, \quad \hat{\lambda}_{k+1} = \mathbf{x}_{k+1}^\top \mathbf{A}\mathbf{x}_{k+1}$$

- ▶ \mathbf{A} – диагонализуема, поэтому есть базис из собственных векторов $[\mathbf{v}_1, \dots, \mathbf{v}_n]$ причём предположим $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$
- ▶ $\|\mathbf{x}_k - \mathbf{v}_1\|_2 \leq \left(\frac{|\lambda_2|}{|\lambda_1|}\right)^k \|\mathbf{x}_0 - \mathbf{v}_1\|_2$

Степенной метод (Power method)

- ▶ Степенной метод для поиска собственного вектора для максимального по модулю собственного значения

$$\mathbf{x}_{k+1} = \frac{\mathbf{A}\mathbf{x}_k}{\|\mathbf{A}\mathbf{x}_k\|_2}, \quad \hat{\lambda}_{k+1} = \mathbf{x}_{k+1}^\top \mathbf{A}\mathbf{x}_{k+1}$$

- ▶ \mathbf{A} – диагонализуема, поэтому есть базис из собственных векторов $[\mathbf{v}_1, \dots, \mathbf{v}_n]$ причём предположим $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$
- ▶ $\|\mathbf{x}_k - \mathbf{v}_1\|_2 \leq \left(\frac{|\lambda_2|}{|\lambda_1|}\right)^k \|\mathbf{x}_0 - \mathbf{v}_1\|_2$
- ▶ Для поиска максимального и минимального собственных значений (не по модулю) существуют более продвинутые методы (см. [метод Ланцоша](#))

Степенной метод (Power method)

- ▶ Степенной метод для поиска собственного вектора для максимального по модулю собственного значения

$$\mathbf{x}_{k+1} = \frac{\mathbf{A}\mathbf{x}_k}{\|\mathbf{A}\mathbf{x}_k\|_2}, \quad \hat{\lambda}_{k+1} = \mathbf{x}_{k+1}^\top \mathbf{A}\mathbf{x}_{k+1}$$

- ▶ \mathbf{A} – диагонализуема, поэтому есть базис из собственных векторов $[\mathbf{v}_1, \dots, \mathbf{v}_n]$ причём предположим $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$
- ▶ $\|\mathbf{x}_k - \mathbf{v}_1\|_2 \leq \left(\frac{|\lambda_2|}{|\lambda_1|}\right)^k \|\mathbf{x}_0 - \mathbf{v}_1\|_2$
- ▶ Для поиска максимального и минимального собственных значений (не по модулю) существуют более продвинутые методы (см. [метод Ланцоша](#))
- ▶ Чтобы понять, как ими пользоваться см. [документацию в SciPy](#)

Неравенство Йенсена

Теорема

Если функция f выпукла, то $f\left(\sum_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^k \alpha_i f(\mathbf{x}_i)$, где

$$\sum_{i=1}^k \alpha_i = 1, \alpha_i \geq 0.$$

Неравенство Йенсена

Теорема

Если функция f выпукла, то $f\left(\sum_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^k \alpha_i f(\mathbf{x}_i)$, где

$$\sum_{i=1}^k \alpha_i = 1, \alpha_i \geq 0.$$

Доказательство по индукции

Неравенство Йенсена

Теорема

Если функция f выпукла, то $f\left(\sum_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^k \alpha_i f(\mathbf{x}_i)$, где

$$\sum_{i=1}^k \alpha_i = 1, \alpha_i \geq 0.$$

Доказательство по индукции

- База $k = 2$ выполнена в силу определения

Неравенство Йенсена

Теорема

Если функция f выпукла, то $f\left(\sum_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^k \alpha_i f(\mathbf{x}_i)$, где

$$\sum_{i=1}^k \alpha_i = 1, \alpha_i \geq 0.$$

Доказательство по индукции

- ▶ База $k = 2$ выполнена в силу определения
- ▶ Пусть неравенство выполнено для $k = m - 1$:

$$f\left(\sum_{i=1}^{m-1} \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^{m-1} \alpha_i f(\mathbf{x}_i) \text{ и } \sum_{i=1}^{m-1} \alpha_i = 1, \alpha_i \geq 0$$

Неравенство Йенсена

Теорема

Если функция f выпукла, то $f\left(\sum_{i=1}^k \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^k \alpha_i f(\mathbf{x}_i)$, где

$$\sum_{i=1}^k \alpha_i = 1, \alpha_i \geq 0.$$

Доказательство по индукции

► База $k = 2$ выполнена в силу определения

► Пусть неравенство выполнено для $k = m - 1$:

$$f\left(\sum_{i=1}^{m-1} \alpha \mathbf{x}_i\right) \leq \sum_{i=1}^{m-1} \alpha_i f(\mathbf{x}_i) \text{ и } \sum_{i=1}^{m-1} \alpha_i = 1, \alpha_i \geq 0$$

► Рассмотрим $k = m$: $f\left(\sum_{i=1}^m \hat{\alpha}_i \mathbf{x}_i\right) = f\left(\sum_{i=1}^{m-1} \hat{\alpha} \mathbf{x}_i + \hat{\alpha}_m \mathbf{x}_m\right) =$

$$f\left((1 - \hat{\alpha}_m) \sum_{i=1}^{m-1} \frac{\hat{\alpha}_i}{1 - \hat{\alpha}_m} \mathbf{x}_i + \hat{\alpha}_m \mathbf{x}_m\right) \leq$$

$$(1 - \hat{\alpha}_m) f\left(\sum_{i=1}^{m-1} \frac{\hat{\alpha}_i}{1 - \hat{\alpha}_m} \mathbf{x}_i\right) + \hat{\alpha}_m f(\mathbf{x}_m) \leq \sum_{i=1}^k \alpha_i f(\mathbf{x}_i)$$

Следствия и обобщения

- ▶ Запись неравенства Йенсена для функции $-\log x$

$$\frac{1}{m} \sum_{i=1}^m x_i \geq \sqrt[m]{x_1 \cdot \dots \cdot x_m}$$

Следствия и обобщения

- ▶ Запись неравенства Йенсена для функции $-\log x$

$$\frac{1}{m} \sum_{i=1}^m x_i \geq \sqrt[m]{x_1 \cdot \dots \cdot x_m}$$

- ▶ Неравенство Гёльдера

$$\sum_{i=1}^n |x_i y_i| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \left(\sum_{i=1}^n |y_i|^q \right)^{1/q}$$

Следствия и обобщения

- ▶ Запись неравенства Йенсена для функции $-\log x$

$$\frac{1}{m} \sum_{i=1}^m x_i \geq \sqrt[m]{x_1 \cdot \dots \cdot x_m}$$

- ▶ Неравенство Гёльдера

$$\sum_{i=1}^n |x_i y_i| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \left(\sum_{i=1}^n |y_i|^q \right)^{1/q}$$

- ▶ Обобщение на непрерывный случай даёт неравенство для выпуклой функции от математического ожидания

$$f(\mathbb{E}(\mathbf{x})) \leq \mathbb{E}(f(\mathbf{x}))$$

- ▶ Выпуклые функции и их свойства

Главное

- ▶ Выпуклые функции и их свойства
- ▶ Критерии выпуклости

- ▶ Выпуклые функции и их свойства
- ▶ Критерии выпуклости
- ▶ Операции, сохраняющие выпуклость

- ▶ Выпуклые функции и их свойства
- ▶ Критерии выпуклости
- ▶ Операции, сохраняющие выпуклость
- ▶ Неравенство Йенсена