Методы оптимизации. Семинар 4. Выпуклые функции.

Александр Катруца

Московский физико-технический институт

4 октября 2021 г.

Напоминание

- Производная по скаляру
- Производная по вектору
- Производная по матрице
- Производная сложной функции
- Автоматическое дифференцирование

Определения функций

Выпуклая функция

Функция $f : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}$ 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:

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

Вогнутая функция

Функция f вогнутая (строго вогнутая), если $-f$ выпуклая (строго выпуклая).

Сильно выпуклая функция

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

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

для минимально возможного $m > 0$.

Определения множеств

Надграфик (эпиграф)

Надграфиком функции f называется множество
$\text{epi} f = \{(\mathbf{x}, y) : \mathbf{x} \in \mathbb{R}^n,\ y \in \mathbb{R},\ y \geq f(\mathbf{x})\} \subset \mathbb{R}^{n+1}$

Множество подуровней (множество Лебега)

Множество подуровня функции f называется следующее множество
$$C_\gamma = \{\mathbf{x} | f(\mathbf{x}) \leq \gamma\}.$$

Квазивыпуклая функция

Функция f называется квазивыпуклой, если её область определения и множество подуровней для любых γ выпуклые множества.

Критерии выпуклости

Дифференциальный критерий первого порядка

Функция f выпукла $\Leftrightarrow$ она определена на выпуклом множестве $\mathcal{X}$ и $\forall \mathbf{x}, \mathbf{y} \in \mathcal{X} \subset \mathbb{R}^n$ выполнено:

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

Дифференциальный критерий второго порядка

Непрерывная и дважды дифференцируемая функция f выпукла $\Leftrightarrow$ она определена на выпуклом множестве $\mathcal{X}$ и $\forall \mathbf{x} \in \mathrm{relint}\,(\mathcal{X}) \subset \mathbb{R}^n$ выполнено:

$$f''(\mathbf{x}) \succeq 0.$$

Связь с надграфиком

Функция выпукла $\Leftrightarrow$ её надграфик выпуклое множество.

Ограничение на прямую

Функция $f : \mathcal{X} \to \mathbb{R}$ выпукла тогда и только тогда, когда $\mathcal{X}$ выпуклое множество и выпукла функция $g(t) = f(\mathbf{x} + t\mathbf{v})$ на множестве $\{t \mid \mathbf{x} + t\mathbf{v} \in \mathcal{X}\}$ для всех $\mathbf{x} \in \mathrm{dom}(f)$ и $\mathbf{v} \in \mathbb{R}^n(\mathbf{S}^n)$.

Критерии сильной выпуклости

Дифференциальный критерий первого порядка

Функция f сильно выпукла с константой m $\Leftrightarrow$ она определена на выпуклом множестве $\mathcal{X}$ и $\forall \mathbf{x}, \mathbf{y} \in \mathcal{X} \subset \mathbb{R}^n$ выполнено:

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

Дифференциальный критерий второго порядка

Непрерывная и дважды дифференцируемая функция f сильно выпукла с константой m $\Leftrightarrow$ она определена на выпуклом множестве $\mathcal{X}$ и $\forall \mathbf{x} \in \mathrm{relint}\,(\mathcal{X}) \subset \mathbb{R}^n$ выполнено:

$$f''(\mathbf{x}) \succeq m\mathbf{I}.$$

Примеры

1. Квадратичная функция: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$

Примеры

1. Квадратичная функция: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$
2. Нормы в $\mathbb{R}^n$

Примеры

1. Квадратичная функция: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$
2. Нормы в $\mathbb{R}^n$
3. $f(\mathbf{x}) = \log\left(e^{x_1} + \ldots + e^{x_n}\right)$, $\mathbf{x} \in \mathbb{R}^n$ — гладкое приближение максимума

Примеры

1. Квадратичная функция: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$
2. Нормы в $\mathbb{R}^n$
3. $f(\mathbf{x}) = \log\left(e^{x_1} + \ldots + e^{x_n}\right)$, $\mathbf{x} \in \mathbb{R}^n$ — гладкое приближение максимума
4. Логарифм детерминанта: $f(\mathbf{X}) = -\log\det \mathbf{X}$, $\mathbf{X} \in \mathbf{S}^n_{++}$

Примеры

1. Квадратичная функция: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$
2. Нормы в $\mathbb{R}^n$
3. $f(\mathbf{x}) = \log\left(e^{x_1} + \ldots + e^{x_n}\right)$, $\mathbf{x} \in \mathbb{R}^n$ — гладкое приближение максимума
4. Логарифм детерминанта: $f(\mathbf{X}) = -\log\det\mathbf{X}$, $\mathbf{X} \in \mathbf{S}^n_{++}$
5. Множество выпуклых функций — выпуклый конус

Примеры

1. Квадратичная функция: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$
2. Нормы в $\mathbb{R}^n$
3. $f(\mathbf{x}) = \log\left(e^{x_1} + \ldots + e^{x_n}\right)$, $\mathbf{x} \in \mathbb{R}^n$ — гладкое приближение максимума
4. Логарифм детерминанта: $f(\mathbf{X}) = -\log\det\mathbf{X}$, $\mathbf{X} \in \mathbf{S}^n_{++}$
5. Множество выпуклых функций — выпуклый конус
6. Поэлементный максимум выпуклых функций: $f(\mathbf{x}) = \max\{f_1(\mathbf{x}), f_2(\mathbf{x})\}$, dom f = dom $f_1 \cap$ dom f_2

Примеры

1. Квадратичная функция: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$
2. Нормы в $\mathbb{R}^n$
3. $f(\mathbf{x}) = \log\left(e^{x_1} + \ldots + e^{x_n}\right)$, $\mathbf{x} \in \mathbb{R}^n$ — гладкое приближение максимума
4. Логарифм детерминанта: $f(\mathbf{X}) = -\log\det\mathbf{X}$, $\mathbf{X} \in \mathbf{S}^n_{++}$
5. Множество выпуклых функций — выпуклый конус
6. Поэлементный максимум выпуклых функций: $f(\mathbf{x}) = \max\{f_1(\mathbf{x}), f_2(\mathbf{x})\}$, dom f = dom $f_1 \cap$ dom f_2
7. Расширение на бесконечное множество функций: если для $\mathbf{y} \in \mathcal{A}$ функция $f(\mathbf{x}, \mathbf{y})$ выпуклая функция по $\mathbf{x}$, тогда $\sup\limits_{\mathbf{y} \in \mathcal{A}} f(\mathbf{x}, \mathbf{y})$ выпукла по $\mathbf{x}$

Примеры

1. Квадратичная функция: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{P} \in \mathbf{S}^n$
2. Нормы в $\mathbb{R}^n$
3. $f(\mathbf{x}) = \log\left(e^{x_1} + \ldots + e^{x_n}\right)$, $\mathbf{x} \in \mathbb{R}^n$ — гладкое приближение максимума
4. Логарифм детерминанта: $f(\mathbf{X}) = -\log\det\mathbf{X}$, $\mathbf{X} \in \mathbf{S}^n_{++}$
5. Множество выпуклых функций — выпуклый конус
6. Поэлементный максимум выпуклых функций: $f(\mathbf{x}) = \max\{f_1(\mathbf{x}), f_2(\mathbf{x})\}$, dom f = dom $f_1 \cap$ dom f_2
7. Расширение на бесконечное множество функций: если для $\mathbf{y} \in \mathcal{A}$ функция $f(\mathbf{x}, \mathbf{y})$ выпуклая функция по $\mathbf{x}$, тогда $\sup\limits_{\mathbf{y} \in \mathcal{A}} f(\mathbf{x}, \mathbf{y})$ выпукла по $\mathbf{x}$
8. Максимальное собственное значение: $f(\mathbf{X}) = \lambda_{\max}(\mathbf{X})$

Неравенство Йенсена

Неравенство Йенсена

Для выпуклой функции f выполнено следующее неравенство:

$$f\left(\sum_{i=1}^{n} \alpha_i \mathbf{x}_i\right) \leq \sum_{i=1}^{n} \alpha_i f(\mathbf{x}_i),$$

где $\alpha_i \geq 0$ и $\sum\limits_{i=1}^{n} \alpha_i = 1$.

или в бесконечномерном случае: $p(x) \geq 0$ и $\int\limits_X p(x) = 1$

$$f\left(\int\limits_X p(x) x dx\right) \leq \int\limits_X f(x) p(x) dx$$

при условии, что интегралы существуют.

Примеры

1. Неравенство Гёльдера
2. Неравенство о среднем арифметическом и среднем геометрическом
3. $f(\mathrm{E}[\mathbf{x}]) \leq \mathrm{E}[f(\mathbf{x})]$
4. Выпуклость множества $\left\{\mathbf{x} \mid \prod\limits_{i=1}^{n} x_i \geq 1\right\}$

Резюме

- Выпуклая функция
- Надграфик и множество подуровня функции
- Критерии выпуклости функции
- Неравенство Йенсена