

Методы оптимизации.

Семинар 6. Выпуклые функции.

Александр Катруца

Московский физико-технический институт,
Факультет Управления и Прикладной Математики

8 октября 2017 г.

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

Определения функций

Выпуклая функция

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

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

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

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

$$C_\gamma = \{\mathbf{x} | f(\mathbf{x}) \leq \gamma\}.$$

Замкнутая функция

Функция f называется замкнутой, если её надграфик замкнутое множество.

Квазивыпуклая функция

Функция f называется квазивыпуклой, если её область определения и множество подуровней выпуклое множество.

Критерии выпуклости

Дифференциальный критерий первого порядка

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

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f^T(\mathbf{x})(\mathbf{y} - \mathbf{x})$$

Дифференциальный критерий второго порядка

Непрерывная и дважды дифференцируемая функция f выпукла тогда и только тогда, когда определена на выпуклом множестве X и $\forall \mathbf{x}, \mathbf{y} \in \text{relint}(X) \subset \mathbb{R}^n$ выполнено:

$$\nabla^2 f(\mathbf{x}) \succeq 0.$$

Связь с надграфиком

Функция выпукла тогда и только тогда, когда её надграфик выпуклое множество.

Ограничение на прямую

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

Критерии сильной выпуклости

Дифференциальный критерий первого порядка

Функция f сильно выпукла с константой m тогда и только тогда, когда определена на выпуклом множестве X и $\forall \mathbf{x}, \mathbf{y} \in X \subset \mathbb{R}^n$ выполнена:

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f^\top(\mathbf{x})(\mathbf{y} - \mathbf{x}) + \frac{m}{2} \|\mathbf{y} - \mathbf{x}\|^2$$

Дифференциальный критерий второго порядка

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

$$\nabla^2 f(\mathbf{x}) \succeq m\mathbf{I}.$$

Примеры

1. Квадратичная функция: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{P} \mathbf{x} + \mathbf{q}^T \mathbf{x} + r$,
 $\mathbf{x} \in \mathbb{R}^n$
2. Нормы в \mathbb{R}^n
3. $f(\mathbf{x}) = \log(e^{x_1} + \dots + e^{x_n})$, $\mathbf{x} \in \mathbb{R}^n$ — 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: $f(\mathbf{x}) = \max\{f_1(\mathbf{x}), f_2(\mathbf{x})\}$,
 $\text{dom } f = \text{dom } f_1 \cap \text{dom } f_2$
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 $\sup_{\mathbf{y} \in \mathcal{A}} f(\mathbf{x}, \mathbf{y})$ выпукла по \mathbf{x}
8. Максимальное собственное значение:
 $f(\mathbf{X}) = \lambda_{\max}(\mathbf{X}) = \sup\{\mathbf{y}^T \mathbf{X} \mathbf{y} \mid \|\mathbf{y}\|_2 = 1\}$

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 $\sum_{i=1}^n \alpha_i = 1$.

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

$$f\left(\int_X p(x) x dx\right) \leq \int_X f(x) p(x) dx$$

при условии, что интегралы существуют.

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

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