Методы Оптимизации. Даниил Меркулов. Выпуклые функции. Сильно выпуклые функции.

Convexity

Выпуклая функция

Функция $f(x)$, **определенная на выпуклом множестве** $S \in \mathbb{R}^n$, называется выпуклой на S, если:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

для любых $x_1, x_2 \in S$ и $0 \leq \lambda \leq 1$.
Если данное неравенство выполняется как строгое при $x_1 \neq x_2$ и $0 < \lambda < 1$, то функция называется строго выпуклой на S

КАРТИНКА С ВЫПУКЛЫМИ ФУНКЦИЯМИ

Примеры

Ниже приведены примеры выпуклых функций:

- $f(x) = x^p, p > 1, \quad S = \mathbb{R}_+$
- $f(x) = \|x\|^p, \quad p > 1, S = \mathbb{R}$
- $f(x) = e^{cx}, \quad c \in \mathbb{R}, S = \mathbb{R}$
- $f(x) = -\ln x, \quad S = \mathbb{R}_{++}$
- $f(x) = x \ln x, \quad S = \mathbb{R}_{++}$
- $f(x) = x_{(1)} + \ldots + x_{(k)}, \quad S = \mathbb{R}^n$ - сумма наибольших k координат
- $f(X) = \lambda_{max}(X), \quad X = X^T$
- $f(X) = -\log \det X, \quad S = S^n_{++}$

Epigraph

Надграфик (эпиграф)

Для функции $f(x)$, определенной на множестве $S \subseteq \mathbb{R}^n$, множество:

$$\text{epi } f = \{[x, \mu] \in S \times \mathbb{R} : f(x) \leq \mu\}$$

называется эпиграфом или надграфиком функции $f(x)$

КАРТИНКА С НАДГРАФИКОМ ПОПОНЯТНЕЕ

Множество подуровней

Для функции $f(x)$, определенной на множестве $S \subseteq \mathbb{R}^n$, множество:

$$\mathcal{L}_\beta = \{x \in S : f(x) \le \beta\}$$

называется множеством Лебега или множеством подуровня функции $f(x)$

КАРТИНКА

Замкнутая функция

Функциф $f(x)$, определенная на множестве $S \subseteq \mathbb{R}^n$ называется замкнутой, если ее надграфик $\mathbf{epi}\ f$ является замкнутым множеством.

КАРТИНКА С ВЫПУКЛОЙ РАЗРЫВНОЙ ФУНКЦИЕЙ

Criteria

Дифференциальный критерий выпуклости 1-ого порядка

Функция $f(x)$, определенная на выпуклом множестве $S \subseteq \mathbb{R}^n$ выпукла тогда и только тогда, когда $\forall x, y \in S$:

$$f(y) \ge f(x) + \nabla f^T(x)(y - x)$$

Пусть $y = x + \Delta x$, тогда критерий перепишется в более интуитивной формулировке:

$$f(x + \Delta x) \ge f(x) + \nabla f^T(x)\Delta x$$

КАРТИНКА С ПОДПИРАНИЕМ ПРЯМОЙ

Дифференциальный критерий выпуклости 2-ого порядка

Дважды непрерывно дифференцируемая функция $f(x)$, определенная на выпуклом множестве $S \subseteq \mathbb{R}^n$ выпукла тогда и только тогда, когда $\forall x \in \mathbf{relint}(S)$:

$$\nabla^2 f(x) \succeq 0$$

Или, что то же самое $\forall y \in \mathbb{R}^n$:

$$\langle y, \nabla^2 f(x) y\rangle \ge 0$$

Связь с надграфиком

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

Связь с линиями уровня

Если $f(x)$ - выпуклая функция, определенная на выпуклом множестве $S \subseteq \mathbb{R}^n$, то для любого β множество подуровня $\mathcal{L}_\beta$ выпукло.
Функция $f(x)$, определенная на множестве $S \subseteq \mathbb{R}^n$, является замкнутой тогда и только тогда, когда все множества подуровня $\mathcal{L}_\beta$ замкнуты.

Ограничение на прямую

Функция $f : S \to \mathbb{R}$ выпукла тогда и только тогда, когда S - выпуклое множество и выпукла функция $g(t) = f(x + tv)$ на множестве $\{t \mid x + tv \in S\}$ для всех $x \in S, v \in \mathbb{R}^n$
PROFIT!!! Можно проверять выпуклость функции многих переменных, проверив выпуклость функции от одной переменной.

Strong convexity

Сильно выпуклая функция

Функция $f(x)$, **определенная на выпуклом множестве** $S \in \mathbb{R}^n$, называется μ-сильно выпуклой (сильно выпуклой) на S, если:

$$f(\lambda x_1 + (1 - \lambda)x_2) \le \lambda f(x_1) + (1 - \lambda)f(x_2) - \mu\lambda(1 - \lambda)\|x_1 - x_2\|$$

для любых $x_1, x_2 \in S$ и $0 \le \lambda \le 1$ и какого-то $\mu > 0$.

КАРТИНКА С ПОДПИРАНИЕМ ПАРАБОЛЛОЙ

Дифференциальный критерий сильной выпуклости 1-ого порядка

Функция $f(x)$, определенная на выпуклом множестве $S \subseteq \mathbb{R}^n$ μ-сильно выпукла тогда и только тогда, когда $\forall x, y \in S$:

$$f(y) \ge f(x) + \nabla f^T(x)(y - x) + \frac{\mu}{2}\|y - x\|^2$$

Пусть $y = x + \Delta x$, тогда критерий перепишется в более интуитивной формулировке:

$$f(x + \Delta x) \ge f(x) + \nabla f^T(x)\Delta x + \frac{\mu}{2}\|\Delta x\|^2$$

Дифференциальный критерий сильной выпуклости 2-ого порядка

Дважды непрерывно дифференцируемая функция $f(x)$, определенная на выпуклом множестве $S \subseteq \mathbb{R}^n$ μ-сильно выпукла тогда и только тогда, когда $\forall x \in \mathbf{relint}(S)$:

$$\nabla^2 f(x) \succeq \mu I$$

То есть:

$$\langle y, \nabla^2 f(x) y\rangle \ge \mu\|y\|^2$$

Важные факты о выпуклости

- Функция $f(x)$ называется (строго) вогнутой, если функция $-f(x)$ - (строго) выпукла.
- Неравенство Йенсена для выпуклых функций:

$$f\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i f(x_i)$$

для $\alpha_i \geq 0; \quad \sum\limits_{i=1}^n \alpha_i = 1$ (вероятностный симплекс)

Его аналог можно определить для бесконечномерного случая:

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

При условии существования интегралов и $p(x) \geq 0, \quad \int\limits_S p(x)dx = 1$

- Если выпуклы функция $f(x)$ и множество S, то любая точка $x^* \in S$, являющаяся точкой локального минимума, будет оптимальной для задачи минимизации функции $f(x)$ на множестве S. Если функция при этом строго выпукла, то решение единственно.

Operations that preserve convexity

- Сумма выпуклых функций с неотрицательными весами: $\alpha f(x) + \beta g(x), (\alpha > 0, \beta > 0)$
- $f(Ax + b)$ выпукла, если $f(x)$ - выпукла
- Если $f_1(x), \ldots, f_m(x)$ - выпуклы, то $f(x) = \max\{f_1(x), \ldots, f_m(x)\} - \text{выпукла}$
- Если $f(x, y)$ - выпукла по x для всех $y \in Y$: $g(x) = \sup\limits_{y \in Y} f(x, y)$ - выпукла
- Если $f(x)$ - выпукла на S, то $g(x, t) = tf(x/t)$ - выпукла при $x/t \in S, t > 0$

Examples

Пример 1

Покажите, что функция $f(x) = \|x\|$ выпукла на $\mathbb{R}^n$.

Решение:

- Докажем выпуклость $f(x)$ по определению:
  $f(\lambda x_1 + (1-\lambda)x_2) = \|\lambda x_1 + (1-\lambda)x_2\| \leq \lambda\|x_1\| + (1-\lambda)\|x_2\|$. Здесь мы использовали неравенство треугольника и $\lambda \in [0, 1]$. Видно, что определение выпуклой функции выполняется.

Пример 2

Покажите, что $f(x) = c^T x + b$ выпукла и вогнута.

Решение:

- Докажем выпуклость $f(x)$ по определению:

  С одной стороны, $f(\lambda x_1 + (1-\lambda)x_2) = c^T(\lambda x_1 + (1-\lambda)x_2) + b$.

  С другой, $\lambda f(x_1) = \lambda c^T x_1 + b; \quad (1-\lambda)f(x_2) = (1-\lambda)(c^T x_2 + b)$
- Складывая последние два равенства, убеждаемся, что определение выпуклости выполняется. Аналогично доказывается выпуклость функции $-f(x)$, что означает вогнутость функции $f(x)$

Пример 3

Покажите, что функция $f(x) = x^T A x$, где $A \succeq 0$ - выпукла на $\mathbb{R}^n$.

Решение:

- $\nabla f = (A + A^T)X$
- $f''(x) = A + A^T$ - необходимо проверить гессиан на положительную определенность, зная, что $A \succeq 0$. Так как матрица A - симметрична (только для матриц, обладающих этим свойством, вводится понятие положительной определенности), и $A^T = A$ то утверждение очевидно.

Пример 4

Показать, что функция $f(A) = \lambda_{max}(A)$ - выпукла, если $A \in S^n$

Решение:

- По определению собственного числа:

$$Ax = \lambda x \quad \to \quad x^T A x = \lambda x^T x \quad \to \quad \lambda = y^T A y, \|y\| = 1$$

Значит, $\lambda_{max}(X) = \sup y^T A y$ - выпуклая функция, как супремум линейных (по A) функций.

Пример 5

Доказать, что функция $-\log\det X$ выпукла по $X \in S^n_{++}$

Решение:

- Рассмотрим скалярную функцию одной переменной:

$$\begin{aligned} g(t) &= f(X + tV) \\ &= \log\det(X + tV) \\ &= \log\det X(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

Она является выпуклой по переменной t, т.к. является композицией выпуклых функций. А значит, исходная функция выпукла.

Пример 6

При каких a, b, c функция $f(x, y, z) = x^2 + 2axy + by^2 + cz^2$ выпукла, строго выпукла, сильно выпукла?

- Для начала составим матрицу Гессе:

$$f''(x) = \begin{pmatrix} 2 & 2a & 0 \\ 2a & 2b & 0 \\ 0 & 0 & 2c \end{pmatrix}$$

- Проверка выпуклости заключается в подсчете миноров всех порядков и задании их неотрицательности.
- Проверка сильной выпуклости происхлжит согласно дифференциальному критерию 2-го порядка.

Пример 7

Покажите с помощью критериев первого и второго порядков, что функция $f(x) = \sum\limits_{i=1}^n x_i^4$ выпукла

easy, men
распиши градиент и гессиан, как в прошлом семинаре

Пример 8

При каких $x \in \mathbb{R}^n$ функция $f(x) = \dfrac{-1}{2(1+x^Tx)}$ выпукла, строго выпукла, сильно выпукла?

Решение:

- Подсчитаем $\nabla f(x) = \dfrac{1}{(1+x^Tx)}x; \quad f''(x) = \dfrac{1}{(1+x^Tx)^2}\left(I - \dfrac{4xx^T}{(1+x^Tx)}x\right)$
- Применяя соответствующие критерии, получаем выпуклость при $x^Tx \leq \frac{1}{3}$, строгую выпуклость при $x^Tx < \frac{1}{3}$

Домашнее задание 5

1. Выпуклы ли следующие функции:
   $f(x) = e^x - 1,\ x \in \mathbb{R}; \quad f(x_1, x_2) = x_1x_2,\ x \in \mathbb{R}^2_{++}; \quad f(x_1, x_2) = 1/(x_1x_2),\ x \in \mathbb{R}^2_{++}$?
2. Докажите, что множество $S = \left\{x \in \mathbb{R}^n \mid \prod\limits_{i=1}^n x_i \geq 1\right\}$ выпукло.
3. Докажите, что функция $f(X) = \mathbf{tr}(X^{-1}), X \in S^n_{++}$ выпукла, а $g(X) = (\det X)^{1/n}, X \in S^n_{++}$ вогнута.
4. Расстоянием Кульбака - Лейблера между между $p, q \in \mathbb{R}^n_{++}$ называется:

$$D(p,q) = \sum_{i=1}^n (p_i \log(p_i/q_i) - p_i + q_i)$$

Докажите, что $D(p,q) \geq 0 \forall p, q \in \mathbb{R}^n_{++}$ и $D(p,q) = 0 \leftrightarrow p = q$
Подсказка:

$$D(p,q) = f(p) - f(q) - \nabla f(q)^T(p-q), \quad f(p) = \sum_{i=1}^n p_i \log p_i$$

5. Пусть x - действительнозначная переменная, принимающая конечный набор значений $a_1 < a_2 < \ldots < a_n$ с вероятностями $\mathbb{P}(x = a_i) = p_i$. Оцените выпуклость и вогнутость следующих функций от p на множестве $\left\{p \mid \sum\limits_{i=1}^n p_i = 1, p_i \geq 0\right\}$
   - $\mathbb{E}x$
   - $\mathbb{P}\{x \geq \alpha\}$
   - $\mathbb{P}\{\alpha \leq x \leq \beta\}$

- $\sum_{i=1}^{n} p_i \log p_i$
- $\mathbb{V}x = \mathbb{E}(x - \mathbb{E}x)^2$
- $\mathbf{quartile}(x) = \inf\{\beta \mid \mathbb{P}\{x \leq \beta\} \geq 0.25\}$