Выпуклые функции. Семинар

Выгузов Александр

10 октября 2025

Выпуклые функции

Definition (Выпуклая функция)

Функция f называется *выпуклой* на множестве D, если для любых $x, y \in D$ и любого $\lambda \in [0, 1]$ выполняется неравенство

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y).$$

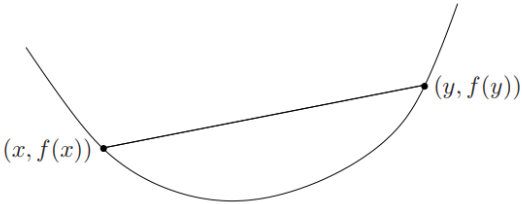

Рис.: Демонстрация неравенства выпуклости.

Критерии выпуклых функций.

Критерий 1. (Дифференцируемая функция) Пусть f — дифференцируемая функция. Она называется *выпуклой* на множестве D, если для любых $x, y \in D$ выполняется неравенство

$$f(y) \geq f(x) + \langle \nabla f(x), \, y - x \rangle,$$

Критерий 2. (Дважды дифференцируемая функция) Если функция f дважды непрерывно дифференцируема на D,то она выпукла на D тогда и только тогда, когда её матрица Гессе является положительно полуопределённой во всех точках $x \in D$

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

, то есть

$$\langle \nabla^2 f(x) v, \, v \rangle \geq 0 \quad \text{для всех } v \in \mathbb{R}^n.$$

Выпуклые функции. Простые примеры выпуклых функций.

Примеры выпуклых и вогнутых функций

- $f(x) = ax + b$ выпукла,
- $f(x) = e^{ax}$ выпукла для $a, x \in \mathbb{R}$,
- $f(x) = x^a$ для $x \in \mathbb{R}_{++}$ выпукла, когда $a \geq 1$ или $a \leq 0$, и вогнута на $0 \leq a \leq 1$,
- $f(x) = (1/2)x^\top Px + q^\top x + r$ выпукла только если $P \succeq 0$
- $f(x) = \log x$ вогнутая функций на $\mathbb{R}_{++}$
- Любые нормы $f(x) = \|x\|$ выпуклы
- $f(x) = \max\{x_1, ..., x_n\}$ выпуклая в $\mathbb{R}^n$

Выпуклые функции. Простые примеры выпуклых функций.

Некоторые обоснования

- $(e^{ax})'' = (ae^{ax})' = a^2 e^{ax} \geq 0$
- $(x^a)'' = (ax^{a-1})' = (a(a-1)x^{a-2}) \geq 0$, т.к. $a \geq 1$
- $\|\alpha x_1 + (1-\alpha)x_2\| \leq \|\alpha x_1\| + \|(1-\alpha)x_2\| = \alpha\|x_1\| + (1-\alpha)\|x_2\|$

Откуда следует неравенство?

Выпуклые функции. Функция max.

Функция max. Докажем выпуклость функции $f(x) = \max_i x_i$ при $0 \leq \theta \leq 1$,

$$f(\theta x + (1-\theta)y) = \max_i \left(\theta x_i + (1-\theta)y_i\right)$$

$$\leq \theta \max_i x_i + (1-\theta) \max_i y_i$$

$$= \theta f(x) + (1-\theta) f(y).$$

Выпуклые функции. LogSumExp.

Задача

Докажем выпуклость функции $\mathrm{LSE} : \mathbb{R}^n \to \mathbb{R}$,

$$f(x) = \mathrm{LSE}(x) = \log\left(\sum_{i=1}^{n} e^{x_i}\right), \qquad x = (x_1, \dots, x_n)^\top \in \mathbb{R}^n.$$

Выпуклые функции. LogSumExp.

Доказательство.

Log-sum-exp. Гессиан функции log-sum-exp имеет вид (см. Boyd, Stephen P., and Lieven Vandenberghe. Convex optimization. Cambridge university press, 2004. Appendix A.4)

$$\nabla^2 f(x) = \frac{1}{(1^T z)^2}\left((1^T z)\,\mathrm{diag}(z) - zz^T\right),$$

где $z = (e^{x_1}, \ldots, e^{x_n})$.

Выпуклые функции. LogSumExp.

Доказательство.

Чтобы проверить, что $\nabla^2 f(x) \succeq 0$, необходимо показать, что для любого v

$$v^T \nabla^2 f(x) v \geq 0,$$

то есть

$$v^T \nabla^2 f(x) v = \frac{1}{(1^T z)^2} \left(\left(\sum_{i=1}^n z_i \right) \left(\sum_{i=1}^n v_i^2 z_i \right) - \left(\sum_{i=1}^n v_i z_i \right)^2 \right) \geq 0.$$

Это следует из неравенства Коши — Буняковского

$$(a^T a)(b^T b) \geq (a^T b)^2$$

применённого к векторам с компонентами $a_i = v_i \sqrt{z_i}$, $b_i = \sqrt{z_i}$. $\square$

Надграфик выпуклых функций.

Определение. Пусть $f : X \to \mathbb{R}$ — функция, определённая на множестве $X \subseteq \mathbb{R}^n$. *Надграфиком* функции f называется множество

$$\operatorname{epi} f = \{(x, t) \in X \times \mathbb{R} \mid t \geq f(x)\}.$$

Иными словами, надграфик состоит из всех точек, расположенных на графике функции и выше него.

Надграфик выпуклых функций.

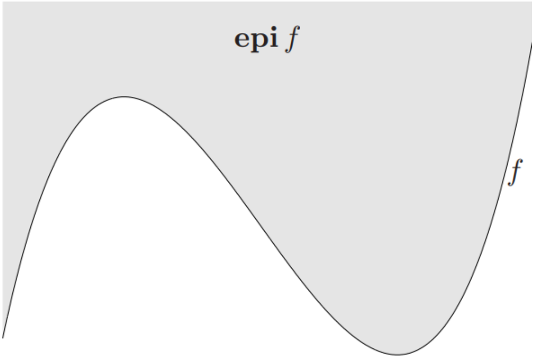

Рис.: Надграфик функции f. Ист. Boyd, Stephen P., and Lieven Vandenberghe. Convex optimization. Cambridge university press, 2004.

Надграфик выпуклых функций.

Функция выпукла тогда и только тогда, когда их надграфик - выпуклое множество.

Доказательство.

1) Допустим f выпукла, тогда

$\alpha t_1 + (1-\alpha)t_2 \geq \alpha f(x_1) + (1-\alpha)f(x_2) \overset{\text{вып.}}{\geq} f(\alpha x_1 + (1-\alpha)x_2)$,

2) Допустим epi f выпукло, тогда по определению надграфика имеем при $t_1 = f(x_1)$, $t_2 = f(x_2)$ и $(\lambda x_1 + (1-\lambda)x_2, \lambda f(x_1) + (1-\lambda)f(x_2))$ имеем следующее неравенство $\lambda f(\alpha x_1 + (1-\alpha)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$, а это и есть выпуклость. $\square$

Операции сохраняющие выпуклость.

- Взвешенная сумма: $\omega_0 x_0 + ... + \omega_n x_n$
- Композиция выпуклой функции f и афинной: $g(x) = f(Ax + b)$
- Поточечный максимум выпуклых функций $f(x) = \max\{f_1(x), ..., f_n(x)\}$

Пример. Сумма r наибольших компонент.

Задача (Сумма r наибольших компонент.)

Для $x \in \mathbb{R}^n$ обозначим через $x_{[i]}$ i-ую наибольшую компоненту вектора x, то есть

$$x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}$$

— это компоненты x, упорядоченные по невозрастанию. Тогда функция

$$f(x) = \sum_{i=1}^{r} x_{[i]},$$

то есть сумма r наибольших элементов вектора x, является выпуклой функцией.

Пример. Сумма r наибольших компонент.

Доказательство.

Это можно увидеть, записав её в виде

$$f(x) = \sum_{i=1}^{r} x_{[i]} = \max\{x_{i_1} + \cdots + x_{i_r} \mid 1 \leq i_1 < i_2 < \cdots < i_r \leq n\},$$

то есть как максимум всех возможных сумм r различных компонент вектора x. Так как это есть поточечный максимум

$$\frac{n!}{r!(n-r)!}$$

линейных функций, то $f(x)$ является выпуклой. $\square$

Операции сохраняющие выпуклость.

Если функция $f(x, y)$ выпукла по x для всех $y \in Q$, то функция

$$g(x) = \sup_y f(x, y)$$

так же выпукла.

Примеры:

- Самая дальняя точка множества $Q \subseteq \mathbb{R}^n$ от заданной точки x: $f(x) = \sup_{y \in Q} \|x - y\|$
- Опорная функция к множеству Q: $S_c(x) = \sup\{x^\top y | y \in Q\}$

Операции сохраняющие выпуклость.

Задача (Максимальное собственное значение симметричной матрицы.)

Функция $f(X) = \lambda_{\max}(X)$ *с областью определения* $\mathrm{dom}\, f = S^m$ *является выпуклой.*

Операции сохраняющие выпуклость.

Доказательство.

Представим собственное число в виде $Xy = \lambda y =>$
$y^\top Xy = \lambda y^\top y = \lambda \|y\|_2^2 => (\frac{y}{\|y\|_2})^\top X(\frac{y}{\|y\|_2}) = \lambda$
Обозначим $v = \frac{y}{\|y\|}$, тогда из требования на максимум исходная задача запишется в виде

$$f(X) = \sup\{v^T Xv \mid \|v\|_2 = 1\},$$

то есть как поточечная супремумная оболочка семейства линейных функций от X (а именно, $y^T Xy$). $\square$

Выпуклость композиции функций.

Дана композиция функций $h : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}$:

$$f(x) = h(g(x)),$$

$\mathrm{dom}\ f = \{x \in \mathrm{dom}\ g \mid g(x) \in \mathrm{dom}\ h\}$, тогда

- f выпукла, если h выпукла и не убывает, g выпукла,
- f выпукла, если h выпукла и не возрастает, g вогнута,
- f вогнута, если h вогнута и не убывает, g вогнута,
- f вогнута, если h вогнута и не возрастает, g выпукла.

Выпуклость композиции функций.

Некоторые примеры выпуклых функций $h(x)$, у которых композиция $h(g(x))$ не выпукла:

- $f(x) = \exp(-x^2)$, где $g(x) = \exp x$ и $h(x) = -x^2$
- $f(x) = -(x^2)$, где $g(x) = x^2$ и $h(x) = -x$
- $f(x) = (-x + x^2)^2$, где $g(x) = -x + x^2$ и $h(x) = x^2$

Выпуклость композиции функций.

Доказательство.

Пусть $f = h \circ g$, где g и h выпуклы, h неубывает. Возьмём $x, y \in \mathrm{dom} f$, $\theta \in [0, 1]$. Тогда $x, y \in \mathrm{dom} g$, $g(x), g(y) \in \mathrm{dom} h$. Из выпуклости g имеем

$$g(\theta x + (1-\theta)y) \le \theta g(x) + (1-\theta)g(y).$$

Так как h неубывает, то

$$h(g(\theta x + (1-\theta)y)) \le h(\theta g(x) + (1-\theta)g(y)).$$

Из выпуклости h получаем

$$h(\theta g(x) + (1-\theta)g(y)) \le \theta h(g(x)) + (1-\theta)h(g(y)).$$

Объединяя предыдущие два неравенства получаем

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta)f(y).$$

Выпуклость композиции функций.

- Если g выпукла, то $\exp g(x)$ выпукла.
- Если g вогнута и положительна, то $\log g(x)$ вогнута.
- Если g вогнута и положительна, то $1/g(x)$ выпукла.
- Если g выпукла и неотрицательна и $p \geq 1$, то $g(x)^p$ выпукла.
- Если g выпукла, то $-\log(-g(x))$ выпукла на множестве $\{x \mid g(x) < 0\}$.

Выпуклость композиции векторных функций. Примеры.

Рассмотрим теперь случай функции $h : \mathbb{R}^k \to \mathbb{R}, g_i : \mathbb{R}^n \to \mathbb{R}$:

$$f(x) = h(g(x)) = h(g_1(x), ..., g_n(x))$$

- f выпукла, если h выпукла, не убывает по каждому аргументу и g_i выпуклы.
- f выпукла, если h выпукла, не возрастает по каждому аргументу и g_i вогнуты.
- f вогнута, если h вогнута, не убывает по каждому аргументу и g_i вогнуты.

Выпуклость композиции векторных функций. Примеры.

- Функция $h(z) = \log\left(\sum_{i=1}^{n} e^{z_i}\right)$ является выпуклой и неубывающей по каждому аргументу, поэтому $\log\left(\sum_{i=1}^{n} e^{g_i}\right)$ выпукла не зависимо от того g_i выпукла или нет.
- Геометрическое среднее $h(z) = \left(\prod_{i=1}^{n} z_i\right)^{1/n}$ на $\mathbb{R}_{++}^{n}$ является вогнутой и неубывающей по каждому аргументу функцией. Отсюда следует, что если функции g_k неотрицательны и вогнуты, то их геометрическое среднее $\left(\prod_{k=1}^{m} g_k\right)^{1/m}$ также является вогнутой функцией.

Выпуклость композиции векторных функций. Пример.

Задача

Предположим, что $p \geq 1$, *а* $g_1, \ldots, g_k$ *являются выпуклыми и неотрицательными функциями. Тогда функция*

$$f(x) = \left(\sum_{i=1}^{k} g_i(x)^p\right)^{1/p}$$

является выпуклой.

Доказательство.

Чтобы это показать, рассмотрим функцию $h : \mathbb{R}^k \to \mathbb{R}$, определённую как

$$h(z) = \left(\sum_{i=1}^{k} \max\{z_i, 0\}^p \right)^{1/p},$$

с областью определения $\operatorname{dom} h = \mathbb{R}^k$. Эта функция выпуклая и неубывающая, поэтому заключаем, что $h(g(x))$ является выпуклой функцией от x. Для $z \geq 0$ имеем

$$h(z) = \left(\sum_{i=1}^{k} z_i^p \right)^{1/p}.$$

$\square$

L-липшицев градиент

Определение 1. Пусть f — дифференцируемая функция. Говорят, что её градиент *липшицев* (или что f имеет L-липшицев градиент), если существует константа $L \geq 0$ такая, что для любых $x, y \in D$ выполняется

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|,$$

где $\| \cdot \|$ — евклидова норма в $\mathbb{R}^n$.

L-липшицев градиент.

Определение 2. Если функция f дважды непрерывно дифференцируема, то условие липшицевости градиента с константой L эквивалентно требованию

$$\|H_f(x)\|_2 \leq L \quad \text{для всех } x \in D,$$

где $H_f(x)$ — матрица Гессе функции f в точке x, а $\|\cdot\|_2$ — спектральная норма (наибольшее по модулю собственное значение).

Сильная выпуклость.

Определение 1. Пусть f — дифференцируемая функция. Она называется *μ-сильно выпуклой* на D (где $\mu > 0$), если для любых $x, y \in D$ выполняется

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\mu}{2} \|y - x\|^2,$$

где $\langle \cdot, \cdot \rangle$ — скалярное произведение, а $\| \cdot \|$ — евклидова норма.

Сильная выпуклость.

Определение 2. Если функция f дважды непрерывно дифференцируема, то она является μ-сильно выпуклой на D тогда и только тогда, когда

$$H_f(x) \succeq \mu I \quad \text{для всех } x \in D,$$

то есть

$$\langle H_f(x)v,\, v\rangle \geq \mu\|v\|^2 \quad \text{для всех } v \in \mathbb{R}^n,$$

где $H_f(x)$ — матрица Гессе, I — единичная матрица, а символ $\succeq$ обозначает положительную определённость в матричном смысле.

Условие градиентного доминирования/PL-условие.

Определение. Функция f удовлетворяет *условию Поляка–Лоясевича* (Polyak–Łojasiewicz, PL-условию) с константой $\mu > 0$, если для всех $x \in D$ выполняется неравенство

$$\frac{1}{2}\|\nabla f(x)\|^2 \geq \mu\big(f(x) - f^*\big),$$

где $f^* = \inf_{y \in D} f(y)$ — минимальное значение функции, а $\|\cdot\|$ — евклидова норма.

Ссылки I

Boyd, Stephen P., and Lieven Vandenberghe. Convex optimization. Cambridge university press, 2004. (Appendix A.4)