Выпуклые функции. Семинар 2. 10 сентября 2019 г.

Семинарист: Данилова М.

Выпуклые функции

Определение 1. Функция $f(x)$, определенная на **выпуклом** множестве G называется **выпуклой на** G, если

$$\forall\alpha \in [0,1] \quad \forall x,y \in \mathrm{G} \quad f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha)f(y).$$

(если знак $<$ для всех $x \ne y,\ \alpha \in (0,1)$, то **строго выпуклой**)

Замечание 1. Необходимость выпуклости области определения возникает в связи с тем, что в определение подразумевается, что мы смотрим на значение функции $f(x)$ в любой точке между x и y.

Определение 2. Функция $f(x)$, определенная на **выпуклом** множестве G называется **сильно выпуклой на** G, если

$$\forall\alpha \in [0,1] \quad \forall x,y \in \mathrm{G} \quad f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha)f(y) - \tfrac{\mu}{2}\alpha(1-\alpha)\|x-y\|^2,$$

где μ – константа сильной выпуклости.

График (строго) выпуклой функции лежит (строго) выше касательной гиперплоскости, а для сильно выпуклой функции график лежит выше некоторого параболоида.

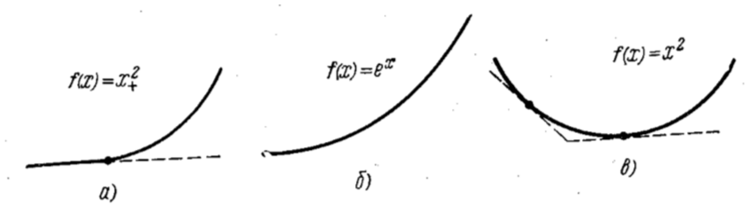


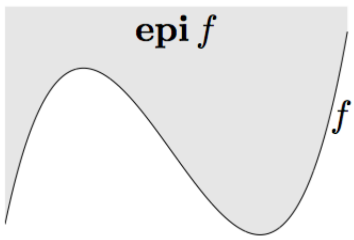
Рис. 1: а) выпуклая функция б) строго выпуклая функция в) сильно выпуклая функция

Примеры:

1. $f(x) = x^2$
2. $f(x) = e^x$
3. $f(x) = -\log x$
4. $f(x) = \|x\|$ – произвольная норма
   - Пусть $x, y \in \mathbb{R}^n,\ \alpha \in [0,1]$, тогда
   - $f(\alpha x + (1-\alpha)y) = \|\alpha x + (1-\alpha)y\| \le \alpha\|x\| + (1-\alpha)\|y\| = \alpha f(x) + (1-\alpha)f(y)$

Определение 3. Эпиграфом (надграфиком) функции называется множество

$$\mathrm{epi}f = \{(x, \mu) \mid x \in \mathbb{R}^n,\ \mu \in \mathbb{R},\ \mu \geq f(x)\} \subseteq \mathbb{R}^{n+1}$$



Определение 4. Функция называется **замкнутой**, если ее надграфик — замкнутое множество.

Определение 5. Функция $f(x)$, заданная на выпуклом множестве, называется **вогнутой (строго вогнутой)**, если $g(x) = -f(x)$ является выпуклой (строго выпуклой) функцией.

Определение 6. Множеством подуровня (множеством Лебега) функции называется

$$\mathrm{C}_\gamma = \{x \in \mathrm{X} \subseteq \mathbb{R}^n \mid f(x) \leq \gamma,\ \gamma = const\}$$

Свойства

1. Выпуклые функции непрерывны во всех внутренних точках области определения.
2. Сильно выпуклая функция, очевидно, строго выпукла. Обратное неверно.
3. Функция выпуклая, тогда и только тогда, когда её надграфик выпуклое множество.
4. Если $f(x)$ выпуклая функция, то множество C_γ является выпуклым.
5. Для выпуклой функции $f(x)$ выполнено **неравенство Йенсена**

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

   для любых $\lambda_i \geq 0$ и $\sum\limits_{i=1} \lambda_i = 1$.

   Неравенство Йенсена обобщается и на случай выпуклой комбинации бесконечного (счетного или несчетного) числа точек.
6. В бесконечномерном случае: $p(x) \geq 0,\ \int_{\mathrm{S}} p(x)dx = 1,\ \mathrm{S} \subseteq \mathrm{dom}f$

   $$f\left(\int_{\mathrm{S}} p(x)x dx\right) \leq \int_{\mathrm{S}} f(x)p(x)dx.$$

7. **Неравенство Гёльдера**: пусть $p > 1,\ \frac{1}{p} + \frac{1}{q} = 1$, тогда

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

8. Линейная функция $f(x) = a^\top x + b,\ a \in \mathbb{R}^n,\ b \in \mathbb{R}$ - выпуклая и вогнутая одновременно.

Пример 1. Неравенство Йенсена

- $f(x) = -\log x$ - выпукла на $\mathbb{R}_{++}$
- Используем неравенство Йенсена:

$$-\log \left(\sum_{i=1}^{m} \lambda_i x_i \right) \leq -\sum_{i=1}^{m} \lambda_i \log x_i = -\log \left(\prod_{i=1}^{m} x_i^{\lambda_i} \right) \quad \Rightarrow \quad \sum_{i=1}^{m} \lambda_i x_i \geq \prod_{i=1}^{m} x_i^{\lambda_i}.$$

- Положим $\lambda_i = \frac{1}{m}$ и получим классическое неравенство: $\frac{1}{m} \sum\limits_{i=1}^{m} x_i \geq \sqrt[m]{\prod\limits_{i=1}^{m} x_i}$.

Проверка выпуклости функции

1. По определению
2. Проверка выпуклости эпиграфа (свойство 3)
3. Операции, сохраняющие выпуклость
4. Дифференциальные критерии выпуклости

Операции, сохраняющие выпуклость

1. Положительная взвешенная сумма

Пусть $f_1, .., f_m$ - выпуклые функции на выпуклом множестве G; $\lambda_1, .., \lambda_m$ – неотрицательные числа.

Тогда функция

$$f(x) = \sum_{i=1}^{m} \lambda_i f_i(x) \quad \text{– выпукла на G.}$$

Доказательство. Докажем по определению, пусть $x, y \in \mathrm{G},\ \alpha \in [0, 1]$.

$$f(\alpha x + (1-\alpha)y) = \sum_{i=1}^{m} \lambda_i f_i(\alpha x + (1-\alpha)y) \leq \sum_{i=1}^{m} \lambda_i \left[\alpha f_i(x) + (1-\alpha) f_i(y)\right] =$$

$$\alpha \sum_{i=1}^{m} \lambda_i f_i(x) + (1-\alpha) \sum_{i=1}^{m} \lambda_i f_i(y) = \alpha f(x) + (1-\alpha) f(y)$$

□

2. Максимум выпуклых функций

Пусть $f_1, .., f_m$ - выпуклые функции на выпуклом множестве G.

Тогда функция

$$f(x) = \max_{i=1,m} f_i(x) \quad \text{– выпукла на G.}$$

Доказательство. Докажем по определению, пусть $x, y \in \mathrm{G},\ \alpha \in [0, 1]$.

$$f(\alpha x + (1-\alpha)y) = \max\{\ f_1(\alpha x + (1-\alpha)y),\ f_2(\alpha x + (1-\alpha)y)\ \} \leq$$

$$\max\{\ \alpha f_1(x) + (1-\alpha) f_1(y),\ \alpha f_2(x) + (1-\alpha) f_2(y)\ \} \leq$$

$$\alpha \max\{\ f_1(x), f_2(x)\ \} + (1-\alpha)\max\{\ f_1(y), f_2(y)\ \} = \alpha f(x) + (1-\alpha) f(y)$$

□

3. Поточечный супремум

Если функция двух аргументов $g(x, y)$ выпукла по $x \in \mathbb{R}^n$ для любого $y \in \mathrm{Y} \subseteq \mathbb{R}^m$, то функция

$$f(x) = \sup_{y \in \mathrm{Y}} g(x, y) \quad \text{– выпукла по } x.$$

Пример 2. Функция расстояния до наиболее удаленной точки множества G:

$$f(x) = \sup_{y \in \mathrm{G}} \|x - y\| \quad \text{– выпукла,}$$

даже когда G - не выпукло.

Пример 3. Функция (максимальное собственное число симметричной матрицы):

$$f\left(\mathrm{X}\right) = \lambda_{\max}\left(\mathrm{X}\right), \ \mathrm{dom} f = \mathbb{S}^m \quad - \text{выпукла.}$$

Доказательство. $f\left(\mathrm{X}\right)$ можно представить в виде

$$f\left(\mathrm{X}\right) = \sup\left\{y^\top \mathrm{X} y \mid \|y\|_2 = 1\right\},$$

то есть как поточечный супремум семейства линейных функций от X □

4. Аффинная подстановка аргумента

- Пусть $\varphi(x)$ - выпуклая функция на выпуклом множестве $\mathrm{G} \subseteq \mathbb{R}^m$.
- Пусть $\mathrm{A} \in \mathbb{R}^{m\times n}, \ b \in \mathbb{R}^m, \ \mathrm{X} = \{x \in \mathbb{R}^n \mid \mathrm{A}x + b \in \mathrm{G}\}, \ \mathrm{X} \neq \varnothing$.
- Тогда $f(x) = \varphi(\mathrm{A}x + b)$ - выпуклая функция на X.

Доказательство. Докажем по определению, пусть $x, y \in \mathrm{X}, \ \alpha \in [0, 1]$.

$$f(\alpha x + (1-\alpha)y) = \varphi\left(\mathrm{A}(\alpha x + (1-\alpha)y) + b\right)$$

$$\varphi\left(\alpha(\mathrm{A}x + b) + (1-\alpha)(\mathrm{A}y + b)\right) \leq \alpha\varphi(\mathrm{A}x + b) + (1-\alpha)\varphi(\mathrm{A}y + b) = \alpha f(x) + (1-\alpha) f(y)$$

□

5. Суперпозиция выпуклых функций

- Пусть $h(x)$ - выпуклая функция, $P(y)$ - выпуклая и неубывающая.
  Тогда $f(x) = P(h(x))$ - выпуклая функция.

  Доказательство. Докажем по определению, пусть $x, y \in \mathrm{X}, \ \alpha \in [0, 1]$.

  $$f(\alpha x + (1-\alpha)y) = P(h(\alpha x + (1-\alpha)y)) \leq P(\alpha h(x) + (1-\alpha)h(y)) \leq$$
  $$\alpha P(h(x)) + (1-\alpha)P(h(y)) \leq \alpha f(x) + (1-\alpha) f(y)$$

  □

- Пусть $h(x)$ - вогнутая функция, $P(y)$ - выпуклая и невозрастающая.
  Тогда $f(x) = P(h(x))$ - выпуклая функция.

Пример 4. Является ли выпуклой функция

$$f(x) = \exp\left(\sum_{i=1}^m |a_i^\top x - b_i|\right)?$$

- $g_i(x) = |a_i^\top x - b_i|$ - выпукла, так как модуль $|\cdot|$ - выпуклая функция, а функция $a_i^\top x - b_i$ – линейная, то есть аффинная подстановка аргумента
- $h(x) = \sum\limits_{i=1}^m g_i(x)$ - выпукла как сумма выпуклых функций
- $f(x) = P(h(x))$ - выпукла как суперпозиция выпуклой, неубывающей и выпуклой функций

Дифференциальные критерии выпуклости

Для краткости формулировок выпуклые функции рассматриваются ниже как сильно выпуклые с константой $\mu = 0$.

1. **Критерий первого порядка №1**

   Пусть $f(x)$ дифференцируемая функция на выпуклом множестве $\mathrm{X} \subseteq \mathbb{R}^n$. Тогда $f(x)$ сильно выпукла на X с константой $\mu \geq 0$ тогда и только тогда, когда

   $$f(x) - f(y) \geq \nabla f(y)^T (x-y) + \frac{\mu}{2}\|x-y\|^2 \qquad \forall x, y \in \mathrm{X}$$

   (если знак $>$, $\theta = 0$ и $\forall x, y \in \mathrm{X},\ x \neq y$, тогда $f(x)$ строго выпукла)

2. **Критерий первого порядка №2**

   Пусть $f(x)$ непрерывно дифференцируемая функция на выпуклом множестве $\mathrm{X} \subseteq \mathbb{R}^n$. Тогда $f(x)$ сильно выпукла на X с константой $\mu \geq 0$ на X тогда и только тогда, когда

   $$\left(\nabla f(x) - \nabla f(y)\right)^\top (x-y) \geq \mu\|x-y\|^2 \qquad \forall x, y \in \mathrm{X}$$

   (если знак $>$, $\theta = 0$ и $\forall x, y \in \mathrm{X},\ x \neq y$, тогда $f(x)$ строго выпукла)

3. **Критерий второго порядка**

   Пусть $\mathrm{X} \subseteq \mathbb{R}^n$ - выпуклое множество, причем $\mathrm{int}\mathrm{X} \neq \varnothing$. $f(x)$ - дважды непрерывно дифференцируемая функция на X. Тогда $f(x)$ сильно выпукла на X с константой $\mu \geq 0$ тогда и только тогда, когда

   $$h^\top \nabla^2 f(x) h \geq \mu\|h\|^2 \qquad \forall x \in \mathrm{X},\ h \in \mathbb{R}^n$$

   $$\left(\nabla^2 f(x) \succeq \mu \mathrm{I} \quad \forall x \in \mathrm{X}\right)$$

4. **Достаточное условие строгой выпуклости второго порядка**

   Пусть $f(x)$ дважды дифференцируемая на выпуклом множестве $\mathrm{X} \subseteq \mathbb{R}^n$.

   Если $\forall x \in \mathrm{X}: \quad \nabla^2 f(x) > 0$, то $f(x)$ строго выпукла на X.

Критерий Сильвестра

$\mathrm{A} = \mathrm{A}^\top \in \mathbb{S}^{n \times n}$ - симметричная матрица.

A положительно полуопределенная ($\mathrm{A} \succeq 0$), если $h^\top \mathrm{A} h \geq 0 \quad \forall h \in \mathbb{R}^n$;

A положительно определенная ($\mathrm{A} \succ 0$), если $h^\top \mathrm{A} h > 0 \quad \forall h \in \mathbb{R}^n \quad h \neq 0$.

1. $\mathrm{A} \succeq 0$, если все её **главные миноры** ($2^n - 1$ штук) положительно полуопределены;

2. $\mathrm{A} \succ 0$, если и только если все её **главные угловые миноры** строго положительны.

Напомним, что главным минором называется любая подматрица, симметричная относительно главной диагонали. Например, образуемая пересечение строк $i_1, \dots, i_k$ и столбцов $i_1, \dots, i_k$, где индексы не обязательно идут подряд. Главным угловым минором пордяка k является матрица, образованная пересечением первых k строки и первых k столбцов.

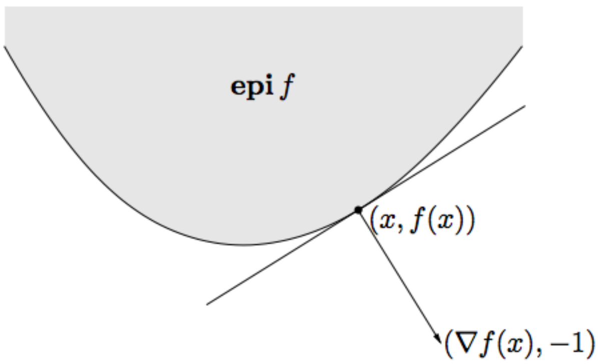
Геометрический смысл

1. Эпиграф функции $f(x)$:
$$\text{epi}f = \{(x,t) \mid f(x) \le t\}$$

2. Критерий выпуклости 1го порядка
$$t \;\geq\; f(x) \;\geq\; f(x_0) + \nabla f(x_0)^\top (x - x_0)$$

3. Преобразуем неравенство и получим полупространство
$$\begin{bmatrix} \triangledown f(x_0) \\ -1 \end{bmatrix}^\top \left(\begin{bmatrix} x \\ t \end{bmatrix} - \begin{bmatrix} x_0 \\ f(x_0) \end{bmatrix} \right) \le 0$$

4. Опорная гиперплоскость к epi$f$ в точке $(x_0, f(x_0))$
$$\begin{bmatrix} \triangledown f(x_0) \\ -1 \end{bmatrix}^\top \left(\begin{bmatrix} x \\ t \end{bmatrix} - \begin{bmatrix} x_0 \\ f(x_0) \end{bmatrix} \right) = 0$$



Пример 5. Рассмотрим простые примеры

1. $f(x) = x^2 + e^x$
   - $f'(x) = 2x + e^x$
   - $f''(x) = 2 + e^x \ge 0 \quad \forall x \in \mathbb{R}$
   - $f(x)$ - выпуклая
2. $f(x) = x^3$

- $f'(x) = 3x^2$
- $f''(x) = 6x$
- $\forall x < 0 \ \rightarrow \ f''(x) < 0.$
- $f(x)$ - не выпуклая

3. $f(x) = \frac{1}{2}\|\mathrm{A}x - b\|_2^2$
   - $\frac{1}{2}\|\mathrm{A}x - b\|_2^2 = \frac{1}{2}(\mathrm{A}x - b)^\top(\mathrm{A}x - b)$
   - $\nabla f(x) = \mathrm{A}^\top(\mathrm{A}x - b)$
   - $\nabla^2 f(x) = \mathrm{A}^\top \mathrm{A}$
   - $h^\top \nabla^2 f(x) h = h^\top \mathrm{A}^\top \mathrm{A} h = \|\mathrm{A}h\|_2^2 \geq 0 \quad \forall h \in \mathbb{R}^n$
   - $f(x)$ - выпуклая
4. $f(x) = x^4 \quad \mathrm{G} = [-1, 1]$
   - $f'(x) = 4x^3$
   - $f''(x) = 12x^2 \geq 0 \quad \forall x \in \mathrm{G}$
   - $f(x)$ - выпукла на G
   - достаточное условие строгой выпуклости не выполняется: $f''(x) \not> 0 \quad \forall x \in \mathrm{G}$
   - критерий первого порядка: $(f'(x) - f'(y))(x - y) > 0$?
   - $4(x^3 - y^3)(x - y) > 0$
   - $(x^2 + xy + y^2) > 0 \quad \forall x, y \in \mathrm{G},\ x \neq y$
   - $f(x)$ - строго выпукла на G

Упражнение 1. Докажите выпуклость следующих функций:

- $f(x) = \max\limits_i x_i$
- $f(x) = \max\{f_1(x), f_2(x)\}$, где $f_1(x), f_2(x)$ - выпуклые функции
- $f(x) = \left(\sum\limits_{i=1}^{n} e^{x_i}\right)$
- $f(x) = \frac{\|\mathrm{A}x - b\|^2}{1 - x^\top x}$, $\mathrm{X} = \{x \mid \|x\|^2 \leq 1\}$
- $f(x) = \frac{x_1^2}{x_2}$, $\mathrm{X} = \{x \mid x_2 > 0\}$

Упражнение 2. Докажите выпуклость следующих функций:

- пусть функция $f(x)$ является выпуклой и принимает только неотрицательные значения. Пусть функция $g(x)$ вогнута и принимает только положительные значения. Докажите, что функция $\mathrm{F}(x) = \frac{f^2(x)}{g(x)}$ является выпуклой;
- пусть функция $f(x)$ выпукла и $f(x) \leq 0$, функция $g(x)$ вогнута и принимает лишь положительные значения на своей области определения. Докажите, что функция $\mathrm{F}(x) = g(x) f\left(\frac{x}{g(x)}\right)$ является выпуклой.