Методы оптимизации. Семинар 5. Выпуклые функции.

Корнилов Никита Максимович

Московский физико-технический институт

1 октября 2024г

Definition (Выпуклые функции)

Пусть U - вещественное векторное пространство, Q - непустое выпуклое множество в U. Функция $f : Q \to \mathbb{R}$ называется выпуклой, если для любых $x, y \in Q$ и любого $0 \le \alpha \le 1$ выполняется

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha)f(y). \tag{1}$$

Выпуклые функции

Definition (Выпуклые функции)

Пусть U - вещественное векторное пространство, Q - непустое выпуклое множество в U. Функция $f : Q \to \mathbb{R}$ называется выпуклой, если для любых $x, y \in Q$ и любого $0 \leq \alpha \leq 1$ выполняется

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y). \qquad (1)$$

Если неравенство (1) выполняется как строгое для всех $x \neq y$ и $0 < \alpha < 1$, то функция f называется строго выпуклой функцией на Q.

Definition (Выпуклые функции)

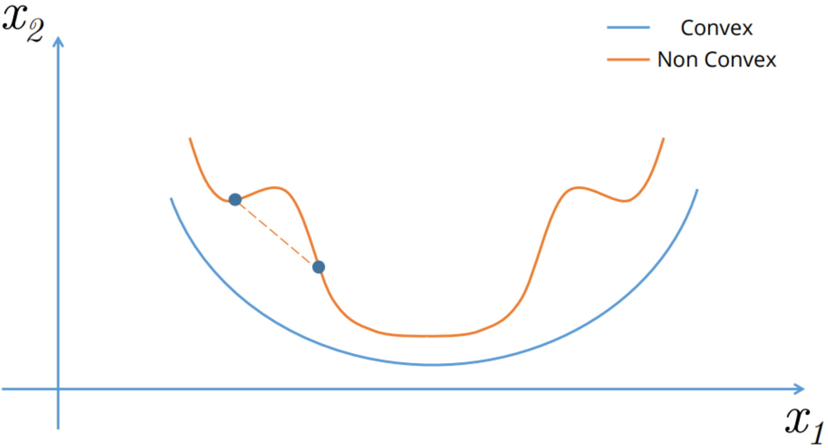
Пусть U - вещественное векторное пространство, Q - непустое выпуклое множество в U. Функция $f : Q \to \mathbb{R}$ называется выпуклой, если для любых $x, y \in Q$ и любого $0 \leq \alpha \leq 1$ выполняется

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y). \tag{1}$$

Если неравенство (1) выполняется как строгое для всех $x \neq y$ и $0 < \alpha < 1$, то функция f называется строго выпуклой функцией на Q.

$\forall x, y \in Q$ нужно вычислять функцию f в любой точке отрезка $[x, y]$. Поэтому требуем, чтобы область определения Q функции являлась выпуклым множеством.

Пример выпуклого множества

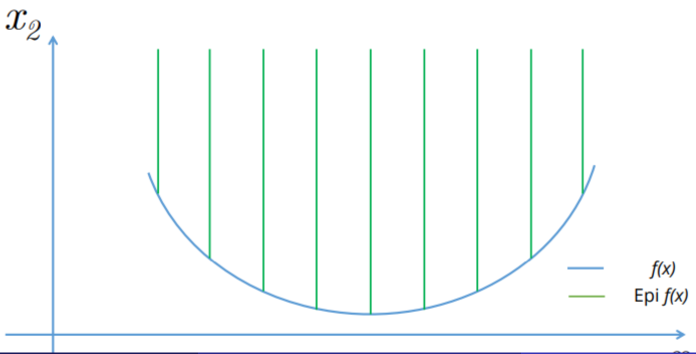


Эпиграф

Definition

Пусть U - вещественное векторное пространство, Q - непустое множество в U. Надграфиком (или эпиграфом) функции $f : Q \to \mathbb{R}$ называется множество

$$\mathrm{Epi}\, f := \{(x,t) \in Q \times \mathbb{R} : f(x) \le t\}\,.$$



Альтернативное определение выпуклости

Следующее утверждение можно считать альтернативным определением выпуклости функции.

Theorem

Пусть U — вещественное векторное пространство, Q - непустое выпуклое множество в U. Функция $f : Q \to \mathbb{R}$ является выпуклой тогда и только тогда, когда её надграфик $\mathrm{Epi}\, f$ является выпуклым множеством в пространстве $U \times \mathbb{R}$

Альтернативное определение выпуклости

Следующее утверждение можно считать альтернативным определением выпуклости функции.

Theorem

Пусть U — вещественное векторное пространство, Q - непустое выпуклое множество в U. Функция $f : Q \to \mathbb{R}$ является выпуклой тогда и только тогда, когда её надграфик $\mathrm{Epi}\, f$ является выпуклым множеством в пространстве $U \times \mathbb{R}$

Больше теоретическое нежели практическое определение.

Вогнутые функции

Definition

Пусть U — вещественное векторное пространство, Q - непустое выпуклое множество в U. Функция $f : Q \to \mathbb{R}$ называется вогнутой, если для любых $x, y \in Q$ и любого $0 \leq \alpha \leq 1$ выполняется

$$f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha) f(y). \tag{2}$$

Вогнутые функции

Definition

Пусть U — вещественное векторное пространство, Q - непустое выпуклое множество в U. Функция $f : Q \to \mathbb{R}$ называется вогнутой, если для любых $x, y \in Q$ и любого $0 \leq \alpha \leq 1$ выполняется

$$f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha)f(y). \tag{2}$$

Если это неравенство (2) выполняется как строгое для всех $x \neq y$ и $0 < \alpha < 1$, то функция f называется строго вогнутой функцией на Q.

Вогнутые функции

Definition

Пусть U — вещественное векторное пространство, Q - непустое выпуклое множество в U. Функция $f : Q \to \mathbb{R}$ называется вогнутой, если для любых $x, y \in Q$ и любого $0 \leq \alpha \leq 1$ выполняется

$$f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha)f(y). \tag{2}$$

Если это неравенство (2) выполняется как строгое для всех $x \neq y$ и $0 < \alpha < 1$, то функция f называется строго вогнутой функцией на Q.

Функция f является (строго) выпуклой тогда и только тогда, когда функция $-f$ является (строго) вогнутой.

Докажем по определению

Example (Афинная функция)

Пусть в пространстве U задано (произвольное) скалярное произведение и $f : Q \to \mathbb{R}$ - аффинная функция

$$f(x) = \langle a, x \rangle + b,$$

где $a \in U$ и $b \in \mathbb{R}$. Проверьте f на выпуклость/вогнутость.

Example (Норма)

Пусть в пространстве U задана (произвольная) норма $||\cdot||$. Функция $f : Q \to \mathbb{R}$ задана формулой

$$f(x) = ||x||.$$

Проверьте f на выпуклость/вогнутость.

Расширение выпуклой функции

Удобно считать, что функция задана не только на своей истинной области определения Q, но так же и за ее пределами, считая, что там функция принимает значение $+\infty$.

Definition

Пусть U - векторное пространство, и $f : U \to \mathbb{R}^*$ - функция, принимающая значения во множестве расширенных вещественных чисел $\mathbb{R}^* = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. Будем называть эффективной областью определения функции f множество всех точек, в которых функция принимает конечные значения:

$$dom\, f = \{x \in U : |f(x)| < +\infty\}$$

Расширение выпуклой функции

Definition (Выпуклые расширеннозначные функции)

Пусть U — вещественное векторное пространство и функция $f : U \to \mathbb{R} \cup \{+\infty\}$ задана на всем пространстве и принимает расширенные вещественные значения. Функция f называется выпуклой, если для любых $x, y \in U$ и любого $\alpha \in [0, 1]$ выполняется

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

Согласно определению выпуклая расширеннозначная функция может принимать только одно расширенное значение - $+\infty$

Расширение выпуклой функции

Definition (Выпуклые расширеннозначные функции)

Пусть U — вещественное векторное пространство и функция $f : U \to \mathbb{R} \cup \{+\infty\}$ задана на всем пространстве и принимает расширенные вещественные значения. Функция f называется выпуклой, если для любых $x, y \in U$ и любого $\alpha \in [0, 1]$ выполняется

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

Согласно определению выпуклая расширеннозначная функция может принимать только одно расширенное значение - $+\infty$

1. $\forall \alpha \in \mathbb{R}^n : \quad -\infty < \alpha < +\infty,$
2. $\forall \alpha \in \mathbb{R}^n : \quad \alpha + \infty = +\infty, \quad \alpha - \infty = -\infty,$
3. $\forall \alpha > 0 \in \mathbb{R}^n : \quad \alpha \cdot (+\infty) = +\infty, \quad \alpha \cdot (-\infty) = -\infty.$

Пример расширеннозначной выпуклой функции

Example

Функция индикатор множества Q:

$$I_Q(x) = \begin{cases} 1, & x \in Q, \\ +\infty, & x \notin Q. \end{cases}$$

Для выпуклых множеств Q функция I_Q выпуклая.

Theorem (Критерий выпуклости 1-го порядка)

Пусть $\mathrm{dom} f$ *является открытым множеством и функция* f *дифференцируема всюду на* $\mathrm{dom} f$. *Функция* f *является выпуклой тогда и только тогда, когда* $\mathrm{dom} f$ *является выпуклым множеством и*

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle \tag{3}$$

для всех $x, y \in \mathrm{dom} f$.

Example

Проверьте критерий на функции $f(x) = \|x\|_2, x \in \mathbb{R}^n$.

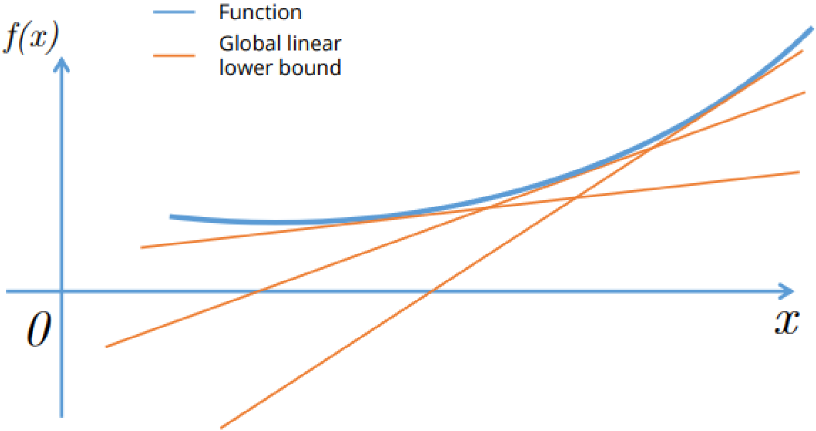


График функции $f(x)$ лежит выше касательной в любой точке $dom\, f$

Theorem (Дифференциальное условие оптимальности для выпуклой функции)

Пусть $f : U \to \mathbb{R} \cup \{+\infty\}$ *- выпуклая функция* $\mathrm{dom}\, f$ *является открытым множеством, и пусть* $x^* \in \mathrm{dom}\, f$*. Тогда* x^* *является глобальным минимумом функции* f*, если и только если* $\nabla f(x^*) = 0$*. Другими словами любая стационарная точка автоматически является глобальным минимумом функции* f*.*

Критерий выпуклости второго порядка

Theorem (Критерий выпуклости 2-го порядка)

Пусть $\mathrm{dom}\, f$ *является открытым множеством и функция* f *дважды дифференцируема на* $\mathrm{dom}\, f$*. Функция* f *является выпуклой тогда и только тогда, когда* $\mathrm{dom}\, f$ *является выпуклым множеством и*

$$\textit{билинейная форма } d^2 f(x) \textit{ неотрицательно определена}$$

для всех $x \in \mathrm{dom}\, f$*.*

В случае $f : \mathbb{R}^n \to \mathbb{R}$ условие имеет вид

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

Докажем по критериям

Example

- $f(x) = \exp(ax)$ выпукла на $\mathbb{R}$ для любого $a \in \mathbb{R}$,
- $f(x) = -\ln x$ выпукла на $\mathbb{R}_{++}$,
- $f(x) = x \ln x$ выпукла на $\mathbb{R}_{+}$,
- $f(x) = x^p$ для $p \geq 1$ или $p \leq 0$ выпукла на $\mathbb{R}_{++}$ и вогнута для $0 \leq p \leq 1$.

Докажем по критериям

Example

Пусть $A \in \mathbb{S}^n, b \in \mathbb{R}^n$. Рассмотрим квадратичную функцию

$$f(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle.$$

Она является выпуклой в том и только в том случае, когда $A \succeq 0$.

Example

В частности, в $\mathbb{R}^2$ рассмотрим матрицу $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$.

Докажем по критериям

Example

Функция

$$f(X) = -\ln det(X)$$

является выпуклой на $\mathbb{S}^n_{++}$.

Example

Функция

$$f(x) = \ln\left(e^{x_1} + e^{x_2} + ... + e^{x_n}\right)$$

является выпуклой на $\mathbb{R}^n$.

Неравенство Йенсена

Theorem

Пусть $f(x)$ — выпуклая функция на выпуклом множестве $X \subseteq \mathbb{R}^n$. Пусть также $x_1, ..., x_k$ — точки, принадлежащие множеству X и коэффициенты $\alpha_1, ..., \alpha_k$ таковы, что $\alpha_i \geq 0$ и $\sum\limits_{i=1}^{k} \alpha_i = 1$. Тогда справедливо следующее неравенство:

$$f\left(\sum_{i=1}^{k} \alpha_i x_i\right) \leq \sum_{i=1}^{k} \alpha_i f(x_i), \tag{4}$$

причем равенство достигается тогда и только тогда, когда функция f является аффинной или когда все точки x_i совпадают.

Следствия

1. Для вектора чисел $x \in \mathbb{R}^n_{++}$ верно
$$\frac{x_1 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 \cdot \ldots \cdot x_n}.$$

Следствия

1. Для вектора чисел $x \in \mathbb{R}^n_{++}$ верно

$$\frac{x_1 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 \cdot \cdots \cdot x_n}.$$

2. Неравенство Гельдера в частности Коши-Буняковского:
Для векторов $x, y \in \mathbb{R}^d$ и чисел $p \geq 1, \frac{1}{q} + \frac{1}{p} = 1$ выполняется неравенство

$$\langle x, y \rangle \leq |\langle x, y \rangle| \leq \|x\|_p \|y\|_q. \tag{5}$$

Следствия

1. Для вектора чисел $x \in \mathbb{R}^n_{++}$ верно
$$\frac{x_1 + \dots + x_n}{n} \geq \sqrt[n]{x_1 \cdot \dots \cdot x_n}.$$

2. Неравенство Гельдера в частности Коши-Буняковского:
Для векторов $x, y \in \mathbb{R}^d$ и чисел $p \geq 1, \frac{1}{q} + \frac{1}{p} = 1$ выполняется неравенство
$$\langle x, y\rangle \leq |\langle x, y\rangle| \leq \|x\|_p \|y\|_q. \tag{5}$$

3. Для выпуклой функции f и случайной величины X верно,
$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

Операции, сохраняющие выпуклость

- Неотрицательная взвешенная сумма

Proposition

Пусть функции $f_1, ...f_m \; : \mathbb{R}^n \to \mathbb{R}$ *выпуклы,* $c_1, ..., c_n \in \mathbb{R}_+$ *Тогда функция*

$$f(x) = \sum_{i=1}^{k} c_i f_i(x)$$

является выпуклой.

Операции, сохраняющие выпуклость

- Неотрицательная взвешенная сумма

Proposition

Пусть функции $f_1, \dots f_m \ : \mathbb{R}^n \to \mathbb{R}$ *выпуклы,* $c_1, \dots, c_n \in \mathbb{R}_+$ *Тогда функция*

$$f(x) = \sum_{i=1}^{k} c_i f_i(x)$$

является выпуклой.

- Аффинная подстановка аргумента

Proposition

Пусть $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *- выпуклая функция,* $A \in \mathbb{R}^{n \times m}$*, и* $b \in \mathbb{R}^n$*. Тогда*

$$g(x) = f(Ax + b),$$

с областью определения $dom\, g = \{x \,|\, Ax + b \in dom f\}$ *.*

Докажем по сохранению

Example

Пусть $a, b \in \mathbb{R}^k$ и $c \in \mathbb{R}^k_+$. Функция

$$f(x) = \sum_{i=1}^{k} c_i \exp\left(\langle a_i, x\rangle + b_i\right)$$

является выпуклой.

Поточечный максимум

Пересекая выпуклые эпиграфы двух выпуклых функций $f_1(x)$ и $f_1(x)$, определенных соответсвенно на множествах X_1 и X_2, приходим к выпуклому множеству, которое является эпиграфом функции

$$f(x) = max\left\{f_1(x), f_2(x)\right\}.$$

Данная функция определена на множестве $X = X_1 \cap X_2$.

Поточечный максимум

Пересекая выпуклые эпиграфы двух выпуклых функций $f_1(x)$ и $f_1(x)$, определенных соответсвенно на множествах X_1 и X_2, приходим к выпуклому множеству, которое является эпиграфом функции

$$f(x) = max\left\{f_1(x), f_2(x)\right\}.$$

Данная функция определена на множестве $X = X_1 \cap X_2$. Пересекая произвольное число выпуклых множеств, мы опять получаем выпуклое множество.

Theorem

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

$$f(x) = \sup_{y \in Y} g(x, y)$$

так же выпукла по x.

Примеры на максимум

Example

Кусочно-линейная функция

$$f(x) = max\left\{a_1^\top x + b_1, ..., a_m^\top x + b_m\right\},$$

где $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, $1 \le i \le m$, выпукла на $\mathbb{R}^n$.

Example (Сумма r максимальных координат)

Обозначим $x_{[i]}$ i-ю максимальную координату вектора $x \in \mathbb{R}^n$, т.е.

$$x_{[1]} \ge x_{[2]} \ge ... \ge x_{[n]}$$

Тогда функция

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

то есть сумма r максимальных координат, есть выпуклая функция.

Примеры на максимум

Example (Расстояние до наиболее удаленной точки множества)

Пусть $C \subseteq \mathbb{R}^n$, $||\cdot||$ - произвольная норма. Тогда расстояние от точки x до наиболее удаленной точки множества C

$$f(x) = \sup_{y \in C} ||x - y||,$$

— выпуклая функция.

Example (Наибольшее собственное число)

Пусть X — симметрическая матрица. Тогда

$$f(X) = \lambda_{\max(X)}$$

является выпуклой.

Монотонная суперпозиция

Definition

Функция $h : \mathbb{R}^m \to \mathbb{R}$ неубывающая, если $\forall x, y \in \mathbb{R}^m$: покоординатно $x \leq y$ верно то, что

$$h(x) \leq h(y).$$

Аналогично обобщаются другие варианты монотонности.

Монотонная суперпозиция

Definition

Функция $h : \mathbb{R}^m \to \mathbb{R}$ неубывающая, если $\forall x, y \in \mathbb{R}^m$: покоординатно $x \leq y$ верно то, что

$$h(x) \leq h(y).$$

Аналогично обобщаются другие варианты монотонности.

Proposition

Пусть $f_i : \mathbb{R}^n \to \mathbb{R}$ — выпуклые функции для $i = \overline{1, m}$, а $h : \mathbb{R}^m \to \mathbb{R}$ - выпуклая неубывающая функция. Тогда композиция этих функции $g(x) = h((f_1(x), \ldots, f_m(x)))$ является выпуклой функцией.

Примеры на монотонность

Example

Пусть $g(x)$ выпукла на $\mathbb{R}^n$. Тогда функция

$$f(x) = e^{g(x)}$$

является выпуклой на $\mathbb{R}^n$.

Example

Пусть $||\cdot||$ — произвольная норма. Тогда функция

$$f(x) = ||x||^p$$

является выпуклой на $\mathbb{R}^n$ при $p \geq 1$.

Сильная выпуклость

Definition

Пусть дана непрерывно дифференцируемая на $\mathbb{R}^d$ функция $f : \mathbb{R}^n \to \mathbb{R}$. Будем говорить, что она является μ-сильно выпуклой ($\mu > 0$), если для любых $x, y \in \mathbb{R}^n$ выполнено

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\mu}{2}\|x - y\|_2^2.$$

Definition

Будем говорить, что она является сильно выпуклой, если для любых $x, y \in \mathbb{R}^n$ и для любого $\lambda \in [0; 1]$ выполнено

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - \lambda(1-\lambda)\frac{\mu}{2}\|x - y\|_2^2.$$

Критерий сильной выпуклости

Theorem

Пусть функция $f : \mathbb{R}^n \to \mathbb{R}$ *дважды непрерывно дифференцируема на* $\mathbb{R}^n$*. Тогда функция* f *является* μ*-сильно выпуклой тогда и только тогда, когда для любого* $x \in \mathbb{R}^n$ *выполнено*

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

Гладкость

Definition

Пусть дана непрерывно дифференцируемая на $\mathbb{R}^n$ функция $f : \mathbb{R}^n \to \mathbb{R}$. Будем говорить, что данная функция имеет L-Липшицев градиент (говорить, что она является L-гладкой), если для любых $x, y \in \mathbb{R}^n$ выполнено

$$\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x - y\|_2.$$

Theorem

Пусть дана L - гладкая функция $f : \mathbb{R}^n \to \mathbb{R}$. Тогда для любых $x, y \in \mathbb{R}^n$ выполнено

$$|f(y) - f(x) - \langle \nabla f(x), y - x \rangle| \leq \frac{L}{2}\|x - y\|_2^2.$$

Выпуклые и гладкие функции

Дифференцируемая функция является μ - сильно выпуклой и L гладкой если

$$\frac{\mu}{2}\|x-y\|^2 \leq f(y) - f(x) - \langle \nabla f(x), y - x \rangle \leq \frac{L}{2}\|x-y\|_2^2, \quad x, y \in \mathbb{R}^n,$$

или эквивалентное утверждение

$$\mu I \preceq \nabla^2 f(x) \preceq LI, \quad x \in \mathbb{R}^n,$$

или же через спектр

$$\mu \leq \lambda_i(\nabla^2 f(x)) \leq L, \quad i = \overline{1, n}, \quad x \in \mathbb{R}^n.$$

Иллюстрация

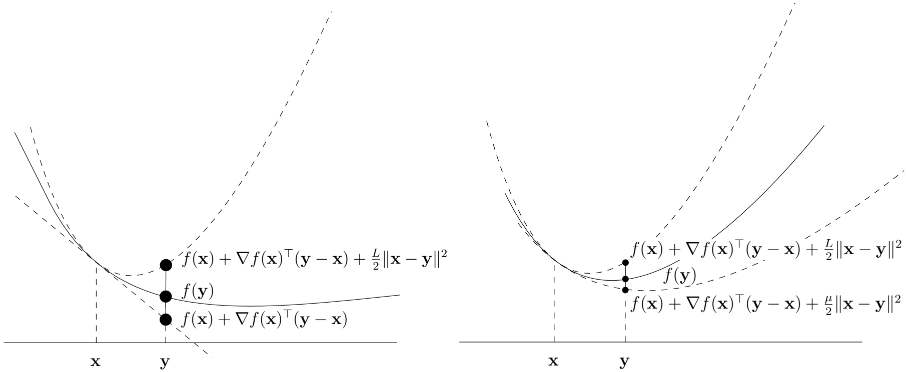


Рис.: Иллюстрация понятий L-гладкости и (μ-сильной) выпуклости

Выпуклость линии уровней

Definition

Пусть $f(x)$ — функция, определенная на множестве X. Тогда множество

$$\mathfrak{L}_\beta = \{x \in X : f(x) \leq \beta\}$$

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

Proposition

Пусть U — вещественное векторное пространство, Q - непустое множество в U и $f : Q \to \mathbb{R}$ выпуклая функция. Тогда $\forall \beta \in \mathbb{R}$ множество подуровня $\mathfrak{L}_\beta$ выпукло.

Верное ли обратное?

Линии уровня