Методы оптимизации. Семинар 6. Выпуклые функции.

Корнилов Никита Максимович

Московский физико-технический институт

9 октября 2025г

Расширеннозначные функции

Definition

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

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

Операции с бесконечностями:

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

Если функция задана только на области определения $Q \subset U$, то удобно доопределить её за пределами Q на всем U, считая, что там функция принимает значение $+\infty$.

Выпуклые функции

Definition (Выпуклые функции)

Пусть U - линейное пространство. Функция $f : U \to \overline{\mathbb{R}}$ называется *выпуклой*, если для любых $x, y \in U$ и любого $\alpha \in [0, 1]$ выполняется

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

Выпуклые функции

Definition (Выпуклые функции)

Пусть U - линейное пространство. Функция $f : U \to \overline{\mathbb{R}}$ называется *выпуклой*, если для любых $x, y \in U$ и любого $\alpha \in [0, 1]$ выполняется

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

Если неравенство (1) выполняется как строгое для всех $x \ne y$ и $\alpha \in (0, 1)$, то функция f называется *строго выпуклой* функцией.

Выпуклые функции

Definition (Выпуклые функции)

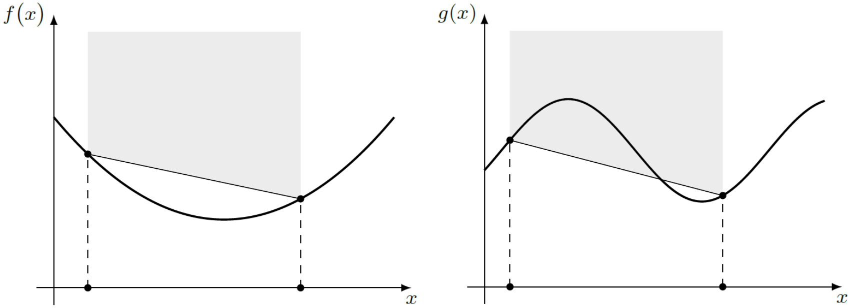
Пусть U - линейное пространство. Функция $f : U \to \overline{\mathbb{R}}$ называется *выпуклой*, если для любых $x, y \in U$ и любого $\alpha \in [0, 1]$ выполняется

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

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

- Выпуклая функция может принимать только одно расширенное значение - $+\infty$ (или быть $f \equiv -\infty$). Поэтому введём определение *собственной функции*: $f(x) > -\infty, \forall x \in U$.
- $\forall x, y \in \operatorname{dom} f$ значение выпуклой функции f в любой точке отрезка $[x, y]$ должно быть конечным. Поэтому у выпуклых функций $\operatorname{dom} f$ — выпуклое множество.

Примеры функций



Вогнутые функции

Definition

Пусть U — линейное пространство. Функция $f : U \to \overline{\mathbb{R}}$ называется *вогнутой*, если для любых $x, y \in U$ и любого $\alpha \in [0, 1]$ выполняется

$$f(\alpha x + (1 - \alpha)y) \geq \alpha f(x) + (1 - \alpha)f(y). \quad (2)$$

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

- Функция f является (строго) выпуклой тогда и только тогда, когда функция $-f$ является (строго) вогнутой.
- У вогнутой функции dom f все также выпуклое множество.
- Вогнутая функция может принимать только значение $-\infty$ (или быть $f \equiv +\infty$.)

Докажем по определению

Example (Афинная функция)

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

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

Проверьте f на выпуклость/вогнутость.

Пример расширеннозначной выпуклой функции

Example

Функция индикатор множества $Q \subset U$:

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

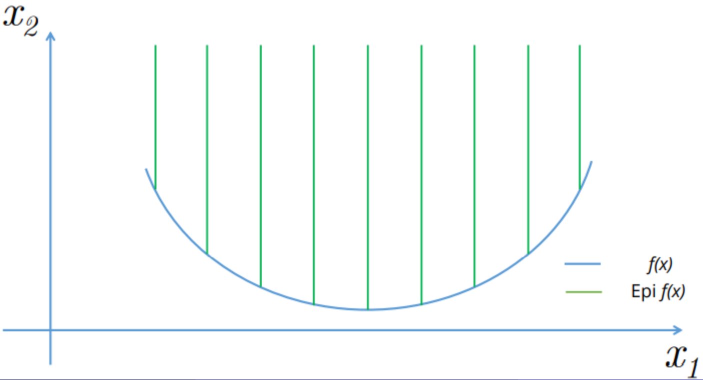
Для выпуклых множеств Q функция I_Q выпуклая.

Эпиграф

Definition

Пусть U - линейное пространство. Надграфиком (или эпиграфом) функции $f : U \to \overline{\mathbb{R}}$ называется множество

$$\operatorname{Epi} f := \left\{(x, t) \in U \times \overline{\mathbb{R}} : f(x) \leq t\right\}.$$

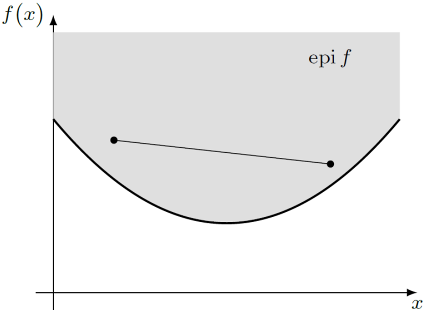


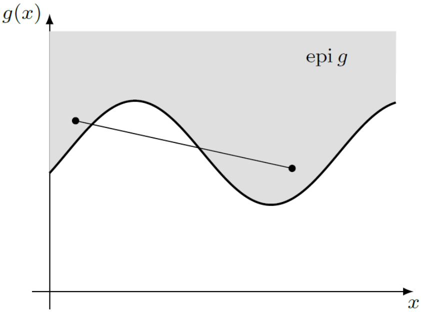
Альтернативное определение выпуклости

Следующее утверждение можно считать альтернативным определением выпуклости функции.

Theorem

Пусть U — линейное пространство. Функция $f : U \to \overline{\mathbb{R}}$ является выпуклой тогда и только тогда, когда её надграфик $\operatorname{Epi} f$ является выпуклым множеством в пространстве $U \times \overline{\mathbb{R}}$.





Критерий выпуклости первого порядка

Theorem (Критерий выпуклости 1-го порядка)

Пусть dom f является открытым множеством и собственная функция f дифференцируема всюду на dom f. Функция f является выпуклой тогда и только тогда, когда dom f является выпуклым множеством и

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

Example

Проверьте критерий на функции $f(x) = \sum_{i=1}^{d} \frac{1}{x_i}, x \in \mathbb{R}^n$ с dom $f = \mathbb{R}^n_{++}$.

Example

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

Критерий выпуклости первого порядка

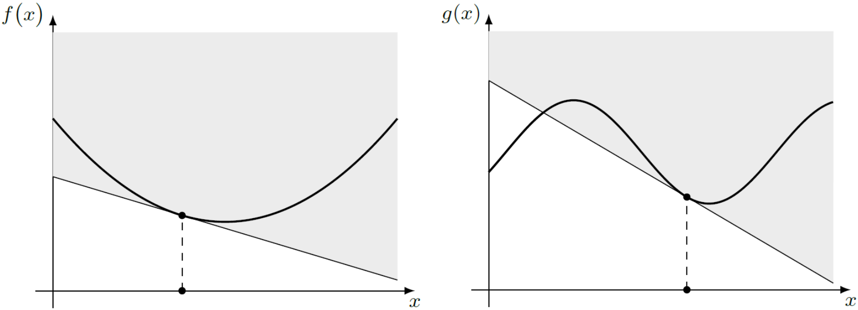


График функции $f(x)$ лежит выше касательной $f(y) + \langle \nabla f(y), x - y \rangle$ в любой точке dom f.

Theorem (Дифференциальное условие оптимальности для выпуклой функции)

Пусть f — собственная выпуклая функция, $\mathrm{dom}\, f$ является открытым множеством, на котором f дифференцируемая, и $x^* \in \mathrm{dom}\, f$. 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 dom f является открытым множеством и собственная функция f дважды дифференцируема на dom f. Функция f является выпуклой тогда и только тогда, когда dom f является выпуклым множеством и

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

для всех $x \in$ dom f.

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

$$\nabla^2 f(x) \succeq 0, \quad x \in \mathrm{dom}\ f.$$

Докажем по критериям

Example

- $f(x) = \exp(ax)$ выпукла для любого $a \in \mathbb{R}$,
- $f(x) = -\ln x$ выпукла с $\operatorname{dom} f = \mathbb{R}_{++}$,
- $f(x) = x \ln x$ выпукла с $\operatorname{dom} f = \mathbb{R}_{+}$,
- $f(x) = -\sqrt{x}$ выпукла с $\operatorname{dom} f = \mathbb{R}_{++}$
- $f(x) = x^p$ для $p \geq 1$ или $p \leq 0$ выпукла с $\operatorname{dom} f = \mathbb{R}_{++}$ 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

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

является выпуклой на $\mathbb{R}^n$.

Неравенство Йенсена

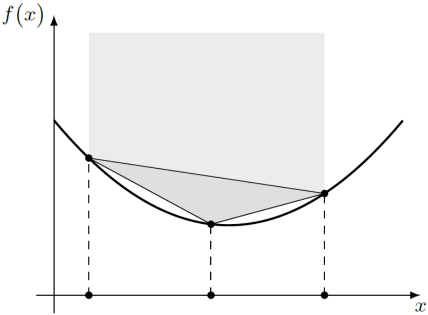
Theorem

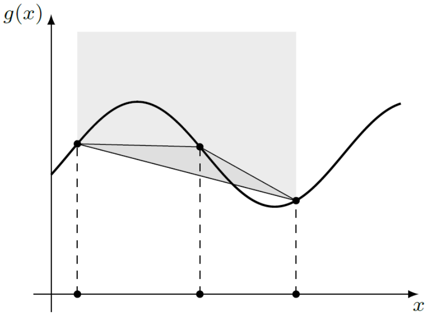
Пусть $f(x) : U \to \overline{\mathbb{R}}$ — выпуклая функция. Пусть также $x_1, \ldots, x_k$ — точки из U и коэффициенты $\alpha_1, \ldots, \alpha_k$ таковы, что $\alpha_i \geq 0$ и $\sum_{i=1}^{k} \alpha_i = 1$. Тогда справедливо следующее неравенство:

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

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

Иллюстрация





Следствия

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

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

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

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 $f_1, \ldots f_m$ *выпуклы,* $c_1, \ldots, c_n \in \mathbb{R}_+$ *Тогда функция*

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

является выпуклой.

Операции, сохраняющие выпуклость

- Неотрицательная взвешенная сумма

Proposition

Пусть функции $f_1, \ldots f_m$ *выпуклы,* $c_1, \ldots, c_n \in \mathbb{R}_+$ *Тогда функция*

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

является выпуклой.

- Аффинная подстановка аргумента

Proposition

Пусть $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *- 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 и f_2, приходим к новому выпуклому множеству, которое является эпиграфом функции

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

Пересекая произвольное число выпуклых множеств, мы опять получаем выпуклое множество.

Theorem

Если функция двух аргументов $g(x, y)$ выпукла по x 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

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

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

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

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 $h : \mathbb{R}^m \to \overline{\mathbb{R}}$ 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 $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}$ — выпуклые функции для $i = \overline{1, m}$, а $h : \mathbb{R}^m \to \overline{\mathbb{R}}$ - 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 $f(x)$ выпукла на $\mathbb{R}^n$. Тогда функция

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

является выпуклой на $\mathbb{R}^n$.

Example

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

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

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

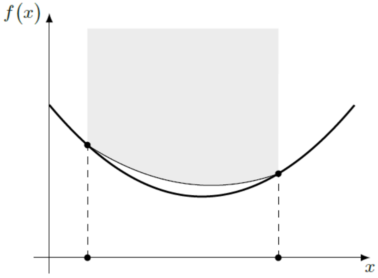
Алгоритм проверки на выпуклость

1. Для простых функций: попробовать воспользоваться определением или выпуклость эпиграфа
2. Посчитать второй дифференциал и проверить, что он положительно определен $d^2 f(x)[h, h] \geq 0$. Если функция $f : \mathbb{R}^n \to \mathbb{R}$, то можно поверить гессиан на $\nabla^2 f(x) \succeq 0$.
3. Если считать второй дифференциал сложно или невозможно, можно проверить операции, сохраняющие выпуклость: аффинная постановка, максимум, положительная сумма, неубывающие суперпозиция. Особенно полезно целевая функция состоит из максимумов или множества простых выпуклых функций.

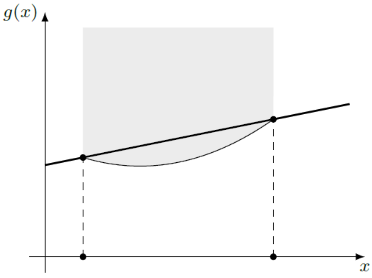
Definition

Функция $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ называется μ-сильно выпуклой, если для любых $x, y \in \mathbb{R}^n$ и $\alpha \in [0, 1]$ выполнено

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



(a) μ-сильно выпуклая парабола.



(b) Не сильно выпуклая прямая.

Критерии сильной выпуклости

Theorem

Пусть $dom\ f$ *является открытым множеством и собственная функция* f *дифференцируема всюду на* $dom\ f$. *Функция* f *является* μ*-сильно выпуклой тогда и только тогда, когда* $dom\ f$ *является выпуклым множеством и*

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

Theorem

Пусть $dom\ f$ *является открытым множеством и собственная функция* f *дважды дифференцируема на* $dom\ f$. *Функция* f *является* μ*-сильно выпуклой тогда и только тогда, когда* $dom\ f$ *является выпуклым множеством и*

$$\nabla^2 f(x) \succeq \mu I, \quad \forall x \in dom\ f.$$

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

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

Непрерывно дифференцируемая функция является μ-сильно выпуклой и L-гладкой:

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

или эквивалентное утверждение для дважды непрерывно дифференцируемой функции

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

или же через спектр

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

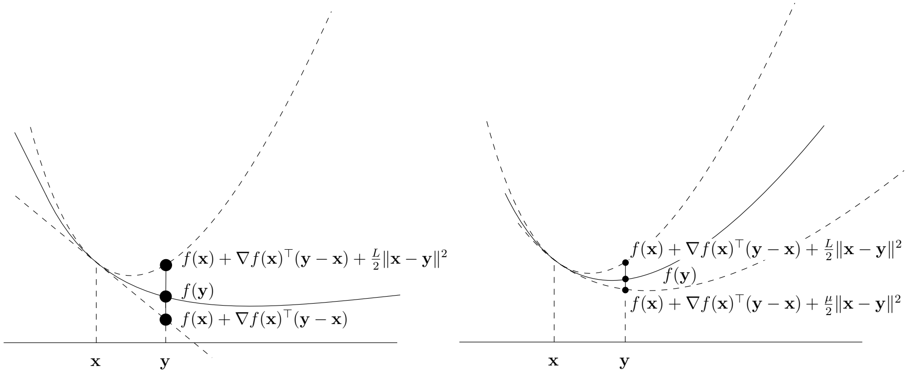


Рис.: Иллюстрация понятий L-гладкости и (μ-сильной) выпуклости

Выпуклость линии уровней

Definition

Для функции $f : U \to \overline{\mathbb{R}}$ множество $\mathfrak{L}_\beta$, определенное скаляром $\beta \in \mathbb{R}$:

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

называется множеством подуровня β функции $f(x)$.

Proposition

Пусть U — линейное пространство и $f : U \to \overline{\mathbb{R}}$ выпуклая функция. Тогда $\forall \beta \in \mathbb{R}$ множество подуровня $\mathfrak{L}_\beta$ выпукло.

Верное ли обратное?

Линии уровня