СОДЕРЖАНИЕ И ТЕХНОЛОГИИ МАТЕМАТИЧЕСКОГО ОБРАЗОВАНИЯ В ВУЗЕ

УДК 51(07)

ИЗЛОЖЕНИЕ ТЕМЫ "ВЫПУКЛЫЕ ФУНКЦИИ" В УНИВЕРСИТЕТСКОМ КУРСЕ МАТЕМАТИЧЕСКОГО АНАЛИЗА

В. И. Гаврилов, А. В. Субботин

Московский государственный университет им. М. В. Ломоносова, Россия, 119899, г. Москва, Ленинские горы, 1; e-mail: awsubbotin@mail.ru

Мы предлагаем оптимальный способ чтения лекции на тему "Выпуклые функции" в курсе математического анализа для вуза. Изложение основано на определении выпуклости сначала для дифференцируемых функций. Этот путь позволяет прочитать всю тему в течение обычных двух академических часов. Полностью все детали содержатся в учебнике [1].

Ключевые слова: выпуклые функции, дифференцируемость, разностное отношение, неравенство Йенсена.

В университетском курсе математического анализа и в вузовских курсах высшей математики тему "Выпуклые функции" обычно помещают в конце раздела "Дифференциальное исчисление функций одного действительного переменного", когда учащимся уже сообщены основные свойства производных функций и их приложения к изучению свойств монотонности, локальных экстремумов, тейлоровских разложений функций. В настоящей статье мы постараемся убедить читателя, что такое положение вещей не есть дань традиции, а выражает суть предмета, позволяющее изложить его оптимальным способом. Сразу договоримся о терминологии: термин "выпуклая функция" используется для выпуклой вниз функции, а термин "вогнутая функция" — для выпуклой вверх функции.

1. ВЫПУКЛЫЕ (ВОГНУТЫЕ) ДИФФЕРЕНЦИРУЕМЫЕ ФУНКЦИИ

Определение 1. Функция f, дифференцируемая на невырожденном интервале (a,b), называется выпуклой (вогнутой) на нём, если для любой точки $x_0 \in (a,b)$ неравенство

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0) \tag{1}$$

$$\Big(f(x) \leq f(x_0) + f'(x_0)(x - x_0)\Big) \tag{1'}$$

справедливо для всех $x \in (a,b)$.

Геометрически свойство выпуклости дифференцируемой функции f на (a, b) означает, что её график в пределах этого интервала располагается выше касательной, проведённой в любой точке графика; для вогнутой дифференцируемой функции картина противоположная (см. рис. 1, а), b)).

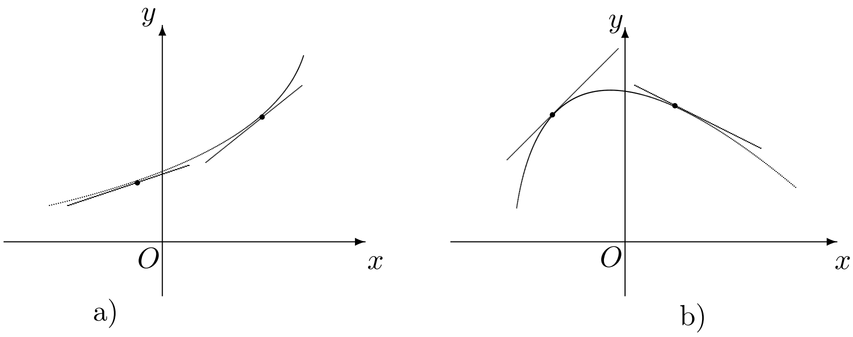

Рис. 1

Замечание. Если обозначить

$$\Delta(x;\, x_0) = \Delta(x) = f(x) - f(x_0) - f'(x_0)(x - x_0),$$

то свойство выпуклости (вогнутости) дифференцируемой функции f на (a, b) равносильно свойству, что для любой точки $x_0 \in (a, b)$ неравенство $\Delta(x) \geq 0$ справедливо для всех $x \in (a, b)$. Отметим, что $\Delta(x_0;\, x_0) = \Delta(x_0) = 0$.

2. КРИТЕРИЙ ВЫПУКЛОСТИ (ВОГНУТОСТИ) ДИФФЕРЕНЦИРУЕМОЙ ФУНКЦИИ НА ИНТЕРВАЛЕ

Теорема 1. *Для того, чтобы функция f, дифференцируемая на интервале (a, b), была выпуклой (вогнутой) на (a, b), необходимо и достаточно, чтобы её производная функция f' возрастала (убывала) на (a, b). При этом производная функция f' непрерывна на (a, b).*

$\lhd$ *Необходимость.* Пусть, для определённости, функция f выпукла на (a, b). Рассмотрим произвольные точки $x_1, x_2 \in (a, b)$, $x_1 < x_2$. Применяя определение 1 к точке $x_0 = x_1$ и считая $x = x_2$, получим неравенство $f(x_2) \geq$ $\geq f(x_1) + f'(x_1)(x_2 - x_1)$, а применяя его к точке $x_0 = x_2$ и считая $x = x_1$, получаем неравенство $f(x_1) \geq f(x_2) + f'(x_2)(x_1 - x_2)$, на основании которых, с учётом условия $x_2 - x_1 > 0$, имеем

$$f'(x_2) \geq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq f'(x_1);$$

т. е., имеем утверждение, что производная функция f' возрастает на (a, b).

При этом, производная f' как монотонная функция может иметь в интервале (a, b) точки разрыва только первого рода, а с другой стороны, согласно теореме о точках разрыва производной функции, все точки разрыва производной функции f' обязаны быть точками разрыва второго рода. Поэтому, производная функция f' выпуклой функции f обязана быть непрерывной функцией на интервале (a, b).

Достаточность. Пусть, для определённости, производная функция f' возрастает на интервале (a, b). Как и выше, заключаем, что f' непрерывна на (a, b). Фиксируем произвольную точку $x_0 \in (a, b)$. Функция $\Delta(x) = f(x)- -f(x_0) - f'(x_0)(x - x_0)$ определена и дифференцируема в интервале (a, b) и её производная $\Delta'(x) = f'(x) - f'(x_0)$. Так как $\Delta'(x) \geq 0$ для всех $x \in (x_0, b)$, $x > x_0$, и $\Delta'(x) \leq 0$ для всех $x \in (a, x_0)$, $x < x_0$, то функция $\Delta(x)$ возрастает на промежутке $[x_0, b)$ и убывает на промежутке $(a, x_0]$, так что $\Delta(x) \geq \Delta(x_0) = 0$ для всех $x \in (a, b)$. Согласно определению 1, функция f выпукла на (a, b). $\square$

3. КРИТЕРИЙ ВЫПУКЛОСТИ (ВОГНУТОСТИ) ДВАЖДЫ ДИФФЕРЕНЦИРУЕМОЙ ФУНКЦИИ НА ИНТЕРВАЛЕ

Теорема 2. *Для того, чтобы функция f, дважды дифференцируемая в интервале (a, b), была выпуклой (вогнутой) на (a, b), необходимо и достаточно, чтобы $f''(x) \geq 0$ ($f''(x) \leq 0$) во всех точках $x \in (a, b)$.*

$\lhd$ Согласно критерию монотонности функции на промежутке, условие $f''(x) \geq 0$ ($f''(x) \leq 0$) для всех $x \in (a, b)$ является необходимым и достаточным условием возрастания (убывания) производной функции f' на (a, b). Последнее свойство, согласно теореме предыдущего пункта, является необходимым и достаточным условием выпуклости (вогнутости) функции f на интервале (a, b). $\square$

4. НЕРАВЕНСТВО ЙЕНСЕНА

Теорема 3. *Пусть дифференцируемая функция f выпукла (вогнута) на интервале (a, b). Тогда, для любых точек $x_1, x_2, \ldots, x_n$ из интервала (a, b) и любых неотрицательных чисел $\alpha_1, \alpha_2, \ldots, \alpha_n$, у которых $\sum\limits_{k=1}^{n} \alpha_k = 1$, выполняется неравенство*

$$\sum_{k=1}^{n} \alpha_k f(x_k) \geq f\left(\sum_{k=1}^{n} \alpha_k x_k\right) \tag{2}$$

$$\left(\sum_{k=1}^{n} \alpha_k f(x_k) \leq f\left(\sum_{k=1}^{n} \alpha_k x_k\right)\right). \tag{2'}$$

$\lhd$ Обозначим $x_0 = \sum\limits_{k=1}^{n} \alpha_k x_k$ и покажем, что $x_0 \in (a, b)$. Так как $\alpha_k \geq 0$, $k = \overline{1, n}$ и $\sum\limits_{k=1}^{n} \alpha_k = 1$, то хотя бы одно $\alpha_k > 0$. По условию, $a < x_k < b$, $k = \overline{1, n}$, и значит, $a \cdot \alpha_k \leq \alpha_k x_k \leq b \cdot \alpha_k$, $k = \overline{1, n}$, и хотя бы для одного k

неравенства строгие, так что $\sum\limits_{k=1}^{n} a \cdot \alpha_k < \sum\limits_{k=1}^{n} \alpha_k x_k < \sum\limits_{k=1}^{n} b \cdot \alpha_k$. Поскольку $\sum\limits_{k=1}^{n} a\,\alpha_k = a\sum\limits_{k=1}^{n} \alpha_k = a$ и $\sum\limits_{k=1}^{n} b\,\alpha_k = b$, то заключаем, что $a < x_0 < b$; т. е. $x_0 \in (a, b)$. Отметим, что $\sum\limits_{k=1}^{n} \alpha_k(x_k - x_0) = 0$.

Пусть сначала f выпукла на (a, b). Так как $\alpha_k \geq 0$, $k = \overline{1, n}$, то из неравенства (1), записанного для точки $x_0 = \sum\limits_{k=1}^{n} \alpha_k x_k$, следует, что неравенства

$$\alpha_k\, f(x_k) \geq \alpha_k\, f(x_0) + f'(x_0)(x_k - x_0)\alpha_k \tag{3}$$

выполняются для каждого k, $1 \leq k \leq n$. Сложив почленно неравенства (3), получим соотношение

$$\sum_{k=1}^{n} \alpha_k\, f(x_k) \geq f(x_0) \sum_{k=1}^{n} \alpha_k + f'(x_0) \sum_{k=1}^{n} \alpha_k(x_k - x_0) = f(x_0),$$

которое совпадает с (2). Аналогично (с изменением знаков всех неравенств в (3) и ниже на противоположные) устанавливается справедливость неравенства (2′) для вогнутой дифференцируемой функции f на (a, b). □

Замечание. Коснёмся вопроса о переходе неравенств (2) и (2′) в равенства. Это произойдёт, во-первых, в том случае, когда точки x_k, $k = \overline{1, n}$, лежат на каком-либо отрезке линейности дифференцируемой функции f; т. е., на таком отрезке интервала (a, b), на котором монотонная производная функция f' постоянна. Если же производная функция f' строго монотонна на (a, b) и все $\alpha_k > 0$, $k = \overline{1, n}$, то при выполнении равенства $\sum\limits_{k=1}^{n} \alpha_k f(x_k) = f(x_0)$ каждое из неравенств (3), $1 \leq k \leq n$, обязано перейти в равенство, так что $x_0 = x_1 = x_2 = \ldots = x_n$. Таким образом, знак равенства в (2) и (2′) во втором случае достигается только при $x_1 = x_2 = \ldots = x_n$.

5. ГЕОМЕТРИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ НЕРАВЕНСТВА ЙЕНСЕНА

Рассмотрим в неравенстве Йенсена случай $n = 2$. Тогда, для дифференцируемой и выпуклой (вогнутой) на интервале (a, b) функции f, для любых точек x_1, $x_2 \in (a, b)$, $x_1 < x_2$, и любого числа λ, $0 \leq \lambda \leq 1$, справедливо неравенство

$$(1 - \lambda)f(x_1) + \lambda f(x_2) \geq f\Big((1 - \lambda)x_1 + \lambda x_2\Big) \tag{4}$$

$$\bigg((1 - \lambda)f(x_1) + \lambda f(x_2) \leq f\Big((1 - \lambda)x_1 + \lambda x_2\Big)\bigg). \tag{4′}$$

В правой части неравенства (4) стоит значение функции f в произвольной точке $x = (1 - \lambda)x_1 + \lambda x_2$, расположенной на отрезке $[x_1, x_2]$, содержащемся в интервале (a, b). Левая часть в (4) выражает собой ординату точки координатной плоскости, абсцисса которой равна $x = (1 - \lambda)x_1 + \lambda x_2$, $0 \leq \lambda \leq 1$, и которая лежит на прямолинейном отрезке (хорде), соединяющем точки $M_1\Big(x_1, f(x_1)\Big)$ и $M_2\Big(x_2, f(x_2)\Big)$ графика функции f.

Итак, если дифференцируемая функция f выпукла на интервале (a, b), то для любых его точек $x_1, x_2 \in (a, b)$, $x_1 < x_2$, график функции f на отрезке $[x_1, x_2]$ расположен *ниже* хорды, стягивающей концевые точки графика на этом отрезке (см. рис. 2, a)).

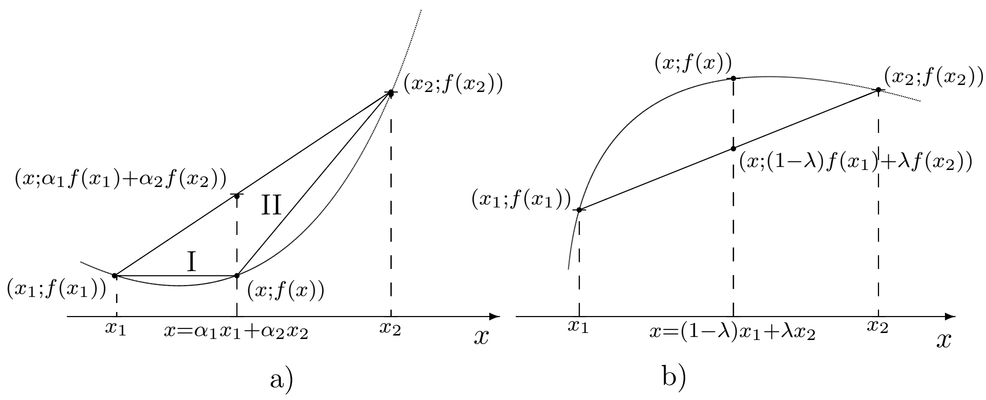

Рис. 2

Аналогично, заключаем, что если дифференцируемая функция f вогнута на интервале (a, b), то для любых его точек $x_1, x_2 \in (a, b)$, $x_1 < x_2$, график функции f на отрезке $[x_1, x_2]$ расположен *выше* хорды, стягивающей концевые точки графика на этом отрезке (см. рис. 2, b)).

6. ВЫПУКЛЫЕ (ВОГНУТЫЕ) ФУНКЦИИ

Неравенство Йенсена, точнее, его частный случай для двух точек, позволяет определить понятие выпуклости (вогнутости) для произвольной функции, заданной в интервале.

Определение 2. Функция f, определённая на невырожденном интервале (a, b), называется выпуклой (вогнутой) на (a, b), если для любых точек $x_1, x_2 \in (a, b)$ и любых чисел $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, для которых $\alpha_1 + \alpha_2 = 1$, справедливо неравенство

$$\alpha_1 f(x_1) + \alpha_2 f(x_2) \geq f(\alpha_1 x_1 + \alpha_2 x_2) \tag{5}$$

$$\Big(\alpha_1 f(x_1) + \alpha_2 f(x_2) \leq f(\alpha_1 x_1 + \alpha_2 x_2)\Big). \tag{5'}$$

Выше отмечалось, что геометрически условие выпуклости (вогнутости) функции f в интервале (a, b) означает, что точки любой дуги графика функции лежат под (над) хордой, стягивающей эту дугу.

Неравенству (5) можно придать другой вид в предположении, что $a < x_1 < x < x_2 < b$ и $x = \alpha_1 x_1 + \alpha_2 x_2$. Из соотношений $x = \alpha_1 x_1 + \alpha_2 x_2$ и $\alpha_1 + \alpha_2 = 1$ имеем $\alpha_1 = \dfrac{x_2 - x}{x_2 - x_1}$, $\alpha_2 = \dfrac{x - x_1}{x_2 - x_1}$, и значит, неравенство (5) принимает вид

$$f(x) \le \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2),$$

или, после умножения обеих частей неравенства на множитель $x_2 - x_1 > 0$,

$$(x_2 - x)f(x_1) + (x_1 - x_2)f(x) + (x - x_1)f(x_2) \ge 0. \tag{6}$$

Поскольку $x_2 - x_1 = (x_2 - x) + (x - x_1)$, то после элементарных преобразований неравенство (6) переходит в неравенство

$$\frac{f(x) - f(x_1)}{x - x_1} \le \frac{f(x_2) - f(x)}{x_2 - x} = \frac{f(x) - f(x_2)}{x - x_2}\,, \tag{7}$$

справедливое для любого x, $x_1 < x < x_2$.

Неравенство (7) является другой формой записи определения выпуклости функции f на интервале (a, b). Геометрически (7) означает (см. рис. 2, а)), что угловой коэффициент хорды I, соединяющей точки $\big(x_1, f(x_1)\big)$ и $\big(x, f(x)\big)$ графика функции f, не больше углового коэффициента хорды II, соединяющей точки $\big(x, f(x)\big)$ и $\big(x_2, f(x_2)\big)$ графика.

В случае функции f, вогнутой на интервале (a, b), знак неравенства в (7) меняется на противоположный.

Теорема 4. *Для любой дифференцируемой функции f на интервале (a, b) определения 1 и 2 эквивалентны.*

$\lhd$ Проверку справедливости утверждения этой теоремы проведём для выпуклых функций.

Установленное в пункте 4 неравенство Йенсена показывает, что в случае дифференцируемых функций из определения 1 следует определение 2.

Рассмотрим теперь произвольную дифференцируемую функцию f, выпуклую на интервале (a, b) в смысле определения 2, и покажем, что её производная функция f' возрастает на (a, b). Для этого возьмём произвольные точки $x_1, x_2 \in (a, b)$, $x_1 < x_2$, так что неравенство (7) справедливо для всех x, $x_1 < x < x_2$. Так как функция f дифференцируема в точке x_1, т. е., $f \in D(x_1)$, то f непрерывна в x_1, т. е., $f \in C(x_1)$, и у неё существуют левая и правая производные $D^- f(x_1)$ и $D^+ f(x_1)$ в точке x_1 и $D^- f(x_1) = D^+ f(x_1) = f'(x_1)$. Аналогично, $f \in C(x_2)$ и $D^- f(x_2) = D^+ f(x_2) = f'(x_2)$. Поэтому, переходя в неравенстве (7) к пределу при $x \to x_1 + 0$, получим

$$\begin{aligned} f'(x_1) &= D^+ f(x_1) = \lim_{x \to x_1 + 0} \frac{f(x) - f(x_1)}{x - x_1} \le \lim_{x \to x_1 + 0} \frac{f(x) - f(x_2)}{x - x_2} = \\ &= \frac{f(x_1) - f(x_2)}{x_1 - x_2} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}\,, \end{aligned} \tag{8}$$

а переходя в (7) к пределу при $x \to x_2 - 0$, получим

$$\begin{aligned}\frac{f(x_2)-f(x_1)}{x_2-x_1} &= \lim_{x\to x_2-0}\frac{f(x)-f(x_1)}{x-x_1} \le \lim_{x\to x_2-0}\frac{f(x)-f(x_2)}{x-x_2} = \\ &= D^- f(x_2) = f'(x_2). \qquad (9)\end{aligned}$$

Объединяя (8) и (9), имеем

$$f'(x_1) \le \frac{f(x_2)-f(x_1)}{x_2-x_1} \le f'(x_2), \quad x_1 < x_2,$$

что означает возрастание производной функции f' на интервале (a, b).

Согласно критерию выпуклости дифференцируемой функции на интервале (см. пункт 2), функция f выпукла на (a, b) в смысле определения 1. □

7. УПРАЖНЕНИЯ

В качестве задач для самостоятельного решения студентам предлагается: 1) доказать неравенство Йенсена (2) для произвольной выпуклой функции, удовлетворяющей определению 2; 2) опираясь на неравенство (7), доказать, что произвольная выпуклая функция, удовлетворяющая определению 2, имеет в каждой точке интервала конечные односторонние производные (и следовательно, непрерывна в этом интервале), и односторонние производные суть возрастающие функции на интервале.

В качестве темы для курсовой работы можно предложить доказать, что последнее свойство 2) является также достаточным условием выпуклости функции на интервале.

8. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

По предложенной схеме эта тема изложена в учебном пособии [1], рекомендованном в качестве базового учебника по математическому анализу в Национальном Университете Черногории и используемом в других университетах Югославии. Многолетний опыт преподавания курсов математического анализа и высшей математики в Московском государственном университете имени М. В. Ломоносова, в Национальном Университете Черногории и экспериментального курса математического анализа в Московском государственном техническом университете имени Н. Э. Баумана в 1987–89 гг. показывает, что прочтение всей темы укладывается в объём обычной аудиторной лекции из двух академических часов.

ЛИТЕРАТУРА

1. Gavrilov V. I., Pavičevič Ž. Matematička Analiza I. — Podgorica: PMF Unirex, 1994. 631 p. (на сербском языке)

PRESENTATION OF THE THEME “CONVEX FUNCTIONS” IN THE UNIVERSITY COURSE OF MATHEMATICAL ANALYSIS

V. I. Gavrilov, A. V. Subbotin

We propose an optimal way of lecturing on the theme “Convex functions” in the course of mathematical analysis at university. It presupposes first the definition of convexity for differentiable functions. This allows the tutor to present the whole theme during two academic hours lecture.

Keywords: convex functions, differentiability, difference quotient, Jensen’s inequality.