УДИВИТЕЛЬНЫЕ СВОЙСТВА ВЫПУКЛЫХ ФУНКЦИЙ*

В. Н. Малозёмов
v.malozemov@spbu.ru

5 сентября 2017 г.

1°. Будем изучать свойства конечных выпуклых функций, заданных на $\mathbb{R}^n$. Напомним, что функция $f(x)$ называется выпуклой, если для любых x_0, x_1 из $\mathbb{R}^n$ и любых $t \in [0,1]$ выполняется неравенство

$$f\big(tx_1 + (1-t)x_0\big) \leqslant tf(x_1) + (1-t)f(x_0). \tag{1}$$

Характерным примером выпуклой функции является *полиэдральная функция* вида

$$f(x) = \max_{k \in 1:m} \big\{ \langle a_k, x \rangle + b_k \big\}.$$

При $n = 1$ полиэдральная функция

$$f(x) = \max_{k \in 1:m} \big\{ a_k x + b_k \big\}$$

называется *ломаной Фурье*.

При $n = 1$, $m = 2$ получаем

$$f(x) = \max\{a_1 x + b_1, a_2 x + b_2\}.$$

При $n = 1$, $m = 2$, $a_1 = 1$, $a_2 = -1$, $b_1 = b_2 = 0$ последняя функция принимает вид

$$f(x) = \max\{x, -x\} = |x|.$$

2°. Для выпуклой функции неравенство (1) допускает естественное обобщение.

ТЕОРЕМА 1. *Пусть $f(x)$ — выпуклая функция. Тогда для произвольных точек $x_0, x_1, \ldots, x_p$ из $\mathbb{R}^n$ и любых неотрицательных чисел $t_0, t_1, \ldots, t_p$, в сумме равных единице, выполняется неравенство*

$$f\Big(\sum_{k=0}^{p} t_k x_k\Big) \leqslant \sum_{k=0}^{p} t_k f(x_k). \tag{2}$$

*Семинар по конструктивному негладкому анализу и недифференцируемой оптимизации «CNSA & NDO»: `http://www.apmath.spbu.ru/cnsa/`

Доказательство. При $p = 1$ неравенство (2) соответствует определению (1). Сделаем индукционный переход от $p \geqslant 1$ к $p+1$.

Возьмём точки $x_0, \dots, x_p, x_{p+1}$ и неотрицательные числа $t_0, \dots, t_p, t_{p+1}$ в сумме равные единице. Если хотя бы одно t_k равно нулю, то требуемое неравенство будет следовать из индукционного предположения. Поэтому можно считать, что все t_k положительны. Учитывая, что их сумма равна единице, заключаем, что $t_k \in (0,1)$ при всех $k \in 0 : p+1$.

Введём обозначения $\lambda = t_0 + t_1 + \ldots + t_p$, $\lambda_k = t_k/\lambda$ при $k \in 0 : p$ и

$$\hat{x} = \lambda_0 t_0 + \lambda_1 t_1 + \ldots + \lambda_p t_p.$$

Очевидно, что числа $\lambda_0, \lambda_1, \dots, \lambda_p$ положительны и в сумме равны единице, при этом $\lambda + t_{p+1} = 1$. По определению выпуклой функции имеем

$$f\Big(\sum_{k=0}^{p+1} t_k x_k\Big) = f(\lambda\hat{x} + t_{p+1}x_{p+1}) \leqslant \lambda f(\hat{x}) + t_{p+1} f(x_{p+1}). \tag{3}$$

В силу индукционного предположения

$$f(\hat{x}) \leqslant \sum_{k=0}^{p} \lambda_k f(t_k). \tag{4}$$

Объединив (3) и (4), получим

$$f\Big(\sum_{k=0}^{p+1} t_k x_k\Big) \leqslant \lambda \sum_{k=0}^{p} \lambda_k f(t_k) + t_{p+1} f(x_{p+1}) = \sum_{k=0}^{p+1} t_k f(x_k).$$

Теорема доказана. □

Неравенство (2) называется *неравенством Йенсена.*

3°. Неравенство (1) можно переписать в эквивалентном виде

$$f\big(x_0 + t(x_1 - x_0)\big) \leqslant f(x_0) + t\big[f(x_1) - f(x_0)\big].$$

ТЕОРЕМА 2. *Если $f(x)$ — выпуклая функция, то функция $f(x_0 + th)$ выпукла как по $t \in \mathbb{R}$, так и по $h \in \mathbb{R}^n$.*

Доказательство. Обозначим $\varphi(t) = f(x_0 + th)$. При $\lambda \in [0,1]$ в силу выпуклости функции f имеем

$$\begin{aligned} \varphi\big(\lambda t_1 + (1-\lambda)t_0\big) &= f\big(\lambda(x_0 + t_1 h) + (1-\lambda)(x_0 + t_0 h)\big) \leqslant \\ &\leqslant \lambda f(x_0 + t_1 h) + (1-\lambda) f(x_0 + t_0 h) = \lambda\varphi(t_1) + (1-\lambda)\varphi(t_0). \end{aligned}$$

Выпуклость $f(x_0 + th)$ по t установлена. Аналогично проверяется выпуклость $f(x_0 + th)$ по h. □

4°. Следующее утверждение вскрывает неожиданную особенность определения выпуклой функции (см. [1], пункт 142).

ТЕОРЕМА 3. *Пусть* $f(x)$ *— выпуклая функция и* $x_0 \neq x_1$. *Если при некотором* $t \in (0,1)$ *неравенство* (1) *выполняется как равенство, то*

$$f\bigl(tx_1 + (1-t)x_0\bigr) = tf(x_1) + (1-t)f(x_0) \quad \forall t \in [0,1]. \tag{5}$$

Доказательство. Обозначим

$$r(t) = f\bigl(x_0 + t(x_1 - x_0)\bigr) - f(x_0) - t\bigl[f(x_1) - f(x_0)\bigr].$$

В силу выпуклости функции f имеем $r(t) \leqslant 0$ для всех $t \in [0,1]$. При этом $r(0) = r(1) = 0$ и по условию теоремы существует точка $t_0 \in (0,1)$, в которой также $r(t_0) = 0$. Нужно показать, что $r(t) = 0$ при всех $t \in [0,1]$.

Допустим противное. Тогда найдётся точка $t_1 \in (0,1)$, в которой $r(t_1) < < 0$. Точка t_1 отлична от 0, t_0 и 1. Для определённости будем считать, что $t_0 < t_1$. В этом случае $t_0 \in (0, t_1)$. Воспользуемся представлением

$$t_0 = \frac{t_0}{t_1} \cdot t_1 + \Bigl(1 - \frac{t_0}{t_1}\Bigr) \cdot 0.$$

Коэффициент t_0/t_1 принадлежит интервалу $(0,1)$. Согласно теореме 2 функция $f\bigl(x_0 + t(x_1 - x_0)\bigr)$ выпукла по t. Выпуклой будет и функция $r(t)$, поэтому

$$r(t_0) \leqslant \frac{t_0}{t_1} r(t_1) + \Bigl(1 - \frac{t_0}{t_1}\Bigr) r(0) = \frac{t_0}{t_1} r(t_1) < 0.$$

Получили противоречие с условием $r(t_0) = 0$.

Случай $t_1 < t_0$, когда $t_0 \in (t_1, 1)$, рассматривается аналогично. Нужно воспользоваться представлением

$$t_0 = \frac{1 - t_0}{1 - t_1} \cdot t_1 + \frac{t_0 - t_1}{1 - t_1} \cdot 1,$$

в котором коэффициенты положительны и в сумме равны единице, и неравенством

$$r(t_0) \leqslant \frac{1 - t_0}{1 - t_1} r(t_1) + \frac{t_0 - t_1}{1 - t_1} r(1) = \frac{1 - t_0}{1 - t_1} r(t_1) < 0.$$

Противоречие с условием $r(t_0) = 0$ завершает доказательство теоремы. □

СЛЕДСТВИЕ. *Если при некотором* $t \in (0,1)$ *неравенство* (1) *выполняется как строгое, то*

$$f\bigl(tx_1 + (1-t)x_0\bigr) < tf(x_1) + (1-t)f(x_0) \quad \forall t \in (0,1). \tag{6}$$

Действительно, в противном случае найдётся $t \in (0,1)$, при котором неравенство (1) выполняется как равенство. По теореме 3 неравенство (1) будет выполняться как равенство при всех $t \in [0,1]$. Но это противоречит условию следствия.

5°. Конечная на $\mathbb{R}^n$ функция $f(x)$ называется *строго выпуклой*, если при всех $x_0 \neq x_1$ и любых $t \in (0,1)$ выполняется строгое неравенство (6).

Очевидно, что строго выпуклая функция является выпуклой.

ТЕОРЕМА 4. *Для того чтобы выпуклая функция $f(x)$ была строго выпуклой, необходимо и достаточно, чтобы ни при каких $x_0 \neq x_1$ не выполнялось условие* (5), *то есть чтобы функция $\varphi(t) = f\big(tx_1 + (1-t)x_0\big)$ не была аффинной на отрезке* $[0,1]$.

Доказательство. Необходимость очевидна.
Достаточность. Возьмём произвольные $x_0 \neq x_1$. Если для выпуклой функции $f(x)$ не выполняется условие (5), то найдётся $t \in (0,1)$, при котором

$$f\big(tx_1 + (1-t)x_0\big) < tf(x_1) + (1-t)f(x_0).$$

Но тогда по следствию из теоремы 3 выполняется условие (6). Это и гарантирует строгую выпуклость функции $f(x)$. $\square$

6°. Неравенство (6), определяющее строго выпуклую функцию, допускает естественное обобщение.

ТЕОРЕМА 5. *Пусть $f(x)$ — строго выпуклая на $\mathbb{R}^n$ функция. Тогда для любых положительных чисел $t_0, t_1, \ldots, t_p$, в сумме равных единице, и произвольных точек $x_0, x_1, \ldots, x_p$ из $\mathbb{R}^n$, среди которых хотя бы две не равны между собой, выполняется неравенство*

$$f\Big(\sum_{k=0}^{p} t_k x_k\Big) < \sum_{k=0}^{p} t_k f(x_k). \tag{7}$$

Доказательство. При $p = 1$ неравенство (7) соответствует определению строго выпуклой функции.

Пусть $p \geqslant 2$ и для определённости $x_0 \neq x_1$. Введём числа $\lambda = t_0 + t_1$, $\lambda_0 = t_0/\lambda$, $\lambda_1 = t_1/\lambda$ и точку

$$\hat{x} = \lambda_0 x_0 + \lambda_1 x_1.$$

Очевидно, что λ_0, λ_1 — положительные числа, в сумме равные единице, и

$$\lambda + t_2 + \ldots + t_p = 1.$$

Строго выпуклая функция $f(x)$ является выпуклой. По неравенству Йенсена (2)

$$f\Big(\sum_{k=0}^{p} t_k x_k\Big) = f\Big(\lambda\hat{x} + \sum_{k=2}^{p} t_k x_k\Big) \leqslant \lambda f(\hat{x}) + \sum_{k=2}^{p} t_k f(x_k). \tag{8}$$

В силу строгой выпуклости

$$f(\hat{x}) < \lambda_0 f(x_0) + \lambda_1 f(x_1). \tag{9}$$

Объединив (8) и (9), получим

$$f\Big(\sum_{k=0}^{p} t_k x_k\Big) < \lambda\big[\lambda_0 f(x_0) + \lambda_1 f(x_1)\big] + \sum_{k=2}^{p} t_k f(x_k) = \sum_{k=0}^{p} t_k f(x_k).$$

Теорема доказана. □

Теорема 5 допускает эквивалентную формулировку.

ТЕОРЕМА 5′. *Пусть $f(x)$ — строго выпуклая на $\mathbb{R}^n$ функция. Тогда для любых положительных чисел $t_0, t_1, \ldots, t_p$, в сумме равных единице, и произвольных точек $x_0, x_1, \ldots, x_p$ из $\mathbb{R}^n$ справедливо неравенство*

$$f\Big(\sum_{k=0}^{p} t_k x_k\Big) \leqslant \sum_{k=0}^{p} t_k f(x_k). \tag{10}$$

Это неравенство выполняется как равенство тогда и только тогда, когда $x_0 = x_1 = \ldots = x_p$.

Действительно, если все x_k равны между собой, то неравенство (10) выполняется как равенство. Наоборот, если неравенство (10) выполняется как равенство, то все x_k равны между собой — иначе получим противоречие с теоремой 5.

Теорема 5′ активно используется при решении экстремальных задач.

7°. Выпуклые функции обладают усиленным свойством непрерывности. Поясним это. Обозначим

$$B_1(x_0;\beta) = \big\{x \in \mathbb{R}^n \mid \|x - x_0\|_1 \leqslant \beta\big\},$$

где $\|\cdot\|_1$ — ℓ_1-норма в $\mathbb{R}^n$.

ТЕОРЕМА 6. *Пусть $f(x)$ — выпуклая на $\mathbb{R}^n$ функция. Тогда для произвольной точки $x_0 \in \mathbb{R}^n$ и любого $\beta > 0$ справедливо неравенство*

$$|f(x) - f(x_0)| \leqslant L\|x - x_0\|_1 \quad \forall x \in B_1(x_0;\beta). \tag{11}$$

Константа L имеет вид

$$L = \max_{k \in 1:n} \left\{ \frac{f(x_0 \pm \beta e_k) - f(x_0)}{\beta} \right\},$$

где $e_1, \ldots, e_n$ — единичные орты.

Отметим, что при $f(x) = \|x\|_1$ и $x_0 = \mathbb{O}$ неравенство (11) выполняется как равенство (с $L = 1$).

При доказательстве неравенства (11) используется только выпуклость функции f (см. доклад [2]).

Из теоремы 6 следует обычная непрерывность выпуклой функции.

8°. Обратимся к вопросу о дифференцируемости выпуклой функции.

ТЕОРЕМА 7. *Пусть у выпуклой функции $f(x)$ в точке x_0 существуют и конечны все частные производные. Тогда $f(x)$ дифференцируема в точке x_0.*

Доказательство. (см. [3], стр. 68). Обозначим

$$\varphi(h) = f(x_0 + h) - f(x_0) - \big\langle f'(x_0), h\big\rangle.$$

Нужно показать, что $\varphi(h)/\|h\| \to 0$ при $h \to \mathbb{O}$. (Здесь $\|\cdot\|$ — евклидова норма).

Очевидно, что $\varphi(0) = 0$. По теореме 2 функция $\varphi(h)$ выпукла, поэтому

$$0 = \varphi(0) = \varphi\big(\tfrac{1}{2}h + \tfrac{1}{2}(-h)\big) \leqslant \tfrac{1}{2}\varphi(h) + \tfrac{1}{2}\varphi(-h).$$

Отсюда следует, что

$$-\varphi(h) \leqslant \varphi(-h). \tag{12}$$

Запишем разложение вектора h по единичным ортам

$$h = \sum_{k=1}^{n} h_k e_k = \tfrac{1}{n}\sum_{k=1}^{n} n h_k e_k.$$

Согласно неравенству Йенсена

$$\varphi(h) \leqslant \tfrac{1}{n}\sum_{k=1}^{n} \varphi(n h_k e_k). \tag{13}$$

Введём функцию

$$\varepsilon_k(h) = \begin{cases} 0, & \text{если } h_k = 0,\\ \frac{\varphi(n h_k e_k)}{n h_k}, & \text{если } h_k \neq 0,\end{cases}$$

и вектор-функцию $\varepsilon(h) = \big(\varepsilon_1(h), \ldots, \varepsilon_n(h)\big)$. Приняв во внимание, что $\varphi(0) = 0$, перепишем неравенство (13) так:

$$\varphi(h) \leqslant \sum_{k=1}^{n} h_k \varepsilon_k(h) = \big\langle h, \varepsilon(h)\big\rangle.$$

Отметим, что

$$\frac{\varphi(nh_k e_k)}{nh_k} = \frac{f(x_0 + nh_k e_k) - f(x_0)}{nh_k} - \frac{\partial f(x_0)}{\partial x_k}.$$

Теперь ясно, что $\varepsilon_k(h) \to 0$ при $h \to \mathbb{O}$ для всех $k \in 1 : n$. Как следствие, $\big\|\varepsilon(h)\big\| \to 0$ при $h \to \mathbb{O}$.

По неравенству Коши-Буняковского

$$\varphi(h) \leqslant \|h\| \cdot \big\|\varepsilon(h)\big\|.$$

Вместе с тем, в силу (12)

$$-\varphi(h) \leqslant \|h\| \cdot \big\|\varepsilon(-h)\big\|.$$

Значит,

$$\left|\frac{\varphi(h)}{\|h\|}\right| \leqslant \max\Big\{\big\|\varepsilon(h)\big\|, \big\|\varepsilon(-h)\big\|\Big\}.$$

Это неравенство и условие $\big\|\varepsilon(h)\big\| \to 0$ при $h \to \mathbb{O}$ гарантируют дифференцируемость функции $f(x)$ в точке x_0.

Теорема доказана. □

9°. В заключение приведём дополнительные сведения о выпуклых функциях одной переменной.

ТЕОРЕМА 8. *Пусть $f(x)$ — выпуклая на $\mathbb{R}$ функция и $[x_0, x_1]$, $[y_0, y_1]$ — два отрезка, взаимное положение которых определяется условиями*

$$y_0 \geqslant x_0, \quad y_1 \geqslant x_1.$$

Тогда

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leqslant \frac{f(y_1) - f(y_0)}{y_1 - y_0}. \tag{14}$$

Доказательство теоремы опирается на следующую лемму.

ЛЕММА. *Если $f(x)$ — выпуклая на $\mathbb{R}$ функция и $x_0 < x < x_1$, то*

$$\frac{f(x) - f(x_0)}{x - x_0} \leqslant \frac{f(x_1) - f(x_0)}{x_1 - x_0}, \tag{15}$$

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leqslant \frac{f(x_1) - f(x)}{x_1 - x}, \tag{16}$$

$$\frac{f(x) - f(x_0)}{x - x_0} \leqslant \frac{f(x_1) - f(x)}{x_1 - x}. \tag{17}$$

Д о к а з а т е л ь с т в о. (см. рис. 1).

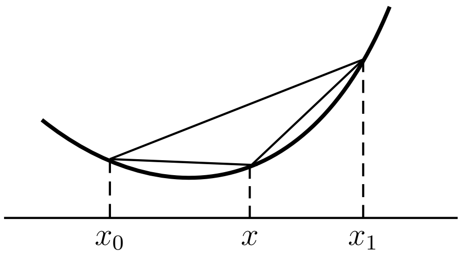

Рис. 1

Имеем

$$x = x_0 + \frac{x - x_0}{x_1 - x_0}(x_1 - x_0).$$

Коэффициент $\frac{x-x_0}{x_1-x_0}$ принадлежит интервалу $(0, 1)$. В силу выпуклости функции f получаем

$$f(x) \leqslant f(x_0) + \frac{x - x_0}{x_1 - x_0}\bigl[f(x_1) - f(x_0)\bigr]. \tag{18}$$

Отсюда непосредственно следует неравенство (15).

Перепишем (18) в виде

$$f(x) \leqslant f(x_1) - \frac{x_1 - x}{x_1 - x_0}\bigl[f(x_1) - f(x_0)\bigr].$$

Это неравенство равносильно (16).

Объединив (15) и (16), придём к (17). □

Д о к а з а т е л ь с т в о т е о р е м ы. Рассмотрим четыре случая.

1) $y_0 = x_0$. Неравенство (14) тривиально при $y_1 = x_1$ и следует из (15) при $x_0 < x_1 < y_1$.

2) $x_0 < y_0 < x_1$. При $y_1 = x_1$ неравенство (14) следует из (16). Если же $y_1 > x_1$, то (см. рис. 2)

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leqslant \frac{f(x_1) - f(y_0)}{x_1 - y_0} \leqslant \frac{f(y_1) - f(y_0)}{y_1 - y_0}.$$

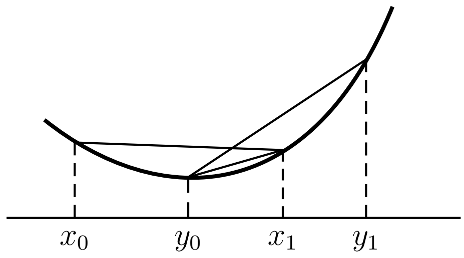

Рис. 2

3) $y_0 = x_1$. Неравенство (14) следует из (17) для тройки $x_0 < x_1 < y_1$.

4) $y_0 > x_1$. Имеем согласно (17) (см. рис. 3)

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leqslant \frac{f(y_0) - f(x_1)}{y_0 - x_1} \leqslant \frac{f(y_1) - f(y_0)}{y_1 - y_0}.$$

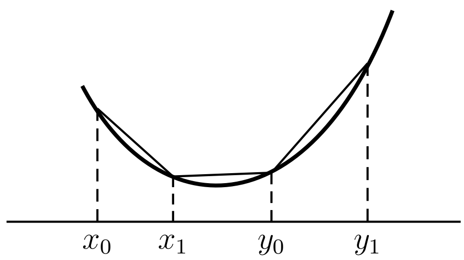

Рис. 3

Теорема доказана. □

ЛИТЕРАТУРА

1. Фихтенгольц Г. М. *Курс дифференциального и интегрального исчисления.* Том 1. Изд. 11-е, стер. СПб.: «Лань», 2017. 608 с.

2. Малозёмов В. Н., Плоткин А. В. *Липшицева непрерывность выпуклой функции* // Семинар «CNSA & NDO». Избранные доклады. 24 ноября 2016 г. (`http://apmath.spbu.ru/cnsa/reps16.shtml#1124`)

3. *Введение в нелинейное программирование.* Под ред. К.-Х. Эльстера. М.: «Наука», 1985. 264 с.