СТРОГО ВЫПУКЛЫЕ ФУНКЦИИ

В.Н.Малозёмов
(Доклад на семинаре 24.11.98)

1^o. Пусть $\Omega \subset \mathbb{R}^n$ — выпуклое множество. Функция f, определённая на Ω, называется *строго выпуклой*, если для любых x_0, x_1 из Ω, $x_0 \neq x_1$, и всех $t \in (0,1)$ выполняется неравенство

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

Это неравенство можно переписать в равносильном виде

$$f(x_0 + t(x_1 - x_0)) < f(x_0) + t[f(x_1) - f(x_0)]. \tag{2}$$

Строгая вогнутость функции f означает строгую выпуклость функции $(-f)$.

ТЕОРЕМА 1. *Пусть $U \subset \mathbb{R}^n$ — открытое выпуклое множество и f — непрерывно дифференцируемая на U функция. Для того чтобы f была строго выпуклой на U, необходимо и достаточно, чтобы*

$$f(x_1) - f(x_0) > \langle f'(x_0), x_1 - x_0 \rangle \tag{3}$$

для всех x_0, x_1 из U, $x_0 \neq x_1$.

Доказательство. Необходимость. Зафиксируем две различные точки x_0, x_1 из U. Поскольку строго выпуклая функция является просто выпуклой, то

$$f(x_1) - f(x_0) \geq \langle f'(x_0), x_1 - x_0 \rangle.$$

Покажем, что на самом деле здесь выполняется строгое неравенство. Допустим противное:

$$f(x_1) - f(x_0) = \langle f'(x_0), x_1 - x_0 \rangle.$$

Зафиксируем $t \in (0,1)$ и положим $x(t) = tx_1 + (1-t)x_0$. Точка $x(t)$ принадлежит U. Согласно (2) запишем

$$f(x(t)) - f(x_0) < t[f(x_1) - f(x_0)] = t\langle f'(x_0), x_1 - x_0 \rangle.$$

С другой стороны,

$$f(x(t)) - f(x_0) \geq t\langle f'(x_0), x_1 - x_0 \rangle.$$

Противоречие завешает доказательство.

Д о с т а т о ч н о с т ь. Пусть x_0, x_1 — две различные точки из U и $t \in (0,1)$. В силу (3) имеем

$$f(x_0) - f(x(t)) > \langle f'(x(t)), x_0 - x(t)\rangle,$$

$$f(x_1) - f(x(t)) > \langle f'(x(t)), x_1 - x(t)\rangle.$$

Умножив первое неравенство на t, второе — на $(1-t)$ и сложив их, придем к (1). Теорема доказана. $\blacksquare$

ТЕОРЕМА 2. *В условиях предыдущей теоремы функция f будет строго выпуклой на U тогда и только тогда, когда*

$$\langle f'(x_1) - f'(x_0), x_1 - x_0\rangle > 0 \tag{4}$$

при всех x_0, x_1 из U, $x_0 \neq x_1$.

Д о к а з а т е л ь с т в о. Достаточно проверить эквивалентность условий (3) и (4). Если выполнено условие (3), то

$$f(x_0) - f(x_1) > \langle f'(x_1), x_0 - x_1\rangle = -\langle f'(x_1), x_1 - x_0\rangle.$$

Сложив это неравенство с (3), получим (4).

Наоборот, пусть выполнено условие (4). Положим $h = x_1 - x_0$. По теореме о среднем

$$f(x_1) - f(x_0) = \langle f'(x_0 + \theta h), h\rangle =$$

$$= \frac{1}{\theta}\langle f'(x_0 + \theta h) - f'(x_0), \theta h\rangle + \langle f'(x_0), h\rangle.$$

Здесь $\theta \in (0,1)$. В силу (4)

$$f(x_1) - f(x_0) > \langle f'(x_0), x_1 - x_0\rangle,$$

т. е. выполняется условие (3). Теорема доказана. $\blacksquare$

ТЕОРЕМА 3. *Пусть функция f дважды непрерывно дифференцируема на U. Если при всех $x \in U$ и $h \in \mathbb{R}^n$, $h \neq \mathbf{0}$, выполняется неравенство*

$$\langle f''(x)h, h\rangle > 0, \tag{5}$$

то f — строго выпуклая на U функция.

Д о к а з а т е л ь с т в о. Зафиксируем две различные точки x_0, x_1 из U и покажем, что для них справедливо неравенство (4). Обозначим $h = x_1 - x_0$, $h \neq \mathbf{0}$, и введем функцию одной переменной $\varphi(t) = \langle f'(x_0 + th), h\rangle$. Имеем

$$\varphi'(t) = \left[\sum_{k=1}^{n} f'_k(x_0 + th)h_k\right]'_t =$$

$$= \sum_{k=1}^{n} h_k \sum_{j=1}^{n} f''_{kj}(x_0 + th)h_j = \langle f''(x_0 + th)h, h\rangle.$$

Учитывая эту формулу и (5), получаем

$$\langle f'(x_1) - f'(x_0), x_1 - x_0\rangle = \varphi(1) - \varphi(0) =$$

$$= \varphi'(\theta) = \langle f''(x_0 + \theta h)h, h\rangle > 0.$$

Теорема доказана. ■

Неравенство (5) является *достаточным* условием строгой выпуклости функции f на U, но в отличие от (3) и (4) оно не является необходимым условием.

ПРИМЕР. Рассмотрим функцию одной переменной $g(t) = t^4$. Для неё при $t_0 \neq t_1$ имеем

$$(g'(t_1) - g'(t_0))(t_1 - t_0) = 4(t_1^3 - t_0^3)(t_1 - t_0) =$$

$$= 4(t_1 - t_0)^2(t_1^2 + t_1 t_0 + t_0^2) = (t_1 - t_0)^2((2t_1 + t_0)^2 + 3t_0^2) > 0.$$

Согласно теореме 2, g — строго выпуклая на $\mathbb{R}$ функция. Вместе с тем, $g''(0) = 0$, так что условие (5) не выполняется.

Для функции одной переменной критерий строгой выпуклости в терминах второй производной выглядит следующим образом.

ТЕОРЕМА 4. *Пусть функция $f(t)$ дважды непрерывно дифференцируема на интервале (a, b). Для строгой выпуклости $f(t)$ на (a, b) необходимо и достаточно, чтобы $f''(t) \geq 0$ при всех $t \in (a, b)$ и чтобы ни при каких $t_0 < t_1$ из (a, b) не выполнялось тождество $f''(t) \equiv 0$ на (t_0, t_1).*

Доказательство. Необходимость. Согласно теореме 2, производная $f'(t)$ строго возрастает на (a, b). В частности, $f''(t) \geq 0$ при всех $t \in (a, b)$. Допустив, что $f''(t) \equiv 0$ при $t \in (t_0, t_1)$, получим, что $f'(t) \equiv \mathrm{const}$ на (t_0, t_1). Это противоречит строгой монотонности $f'(t)$.

Достаточность. Условие $f''(t) \geq 0$ гарантирует монотонное неубывание $f'(t)$ на (a, b). Покажем, что $f'(t_0) < f'(t_1)$ при $t_0 < t_1$. В противном случае $f'(t_0) = f'(t_1)$. Более того, в силу монотонного неубывания $f'(t) = f'(t_0)$ при всех $t \in (t_0, t_1)$. Отсюда следует, что $f''(t) \equiv 0$ на (t_0, t_1). Это противоречит условию теоремы.

Установлено, что производная $f'(t)$ строго возрастает на (a, b). Остается сослаться на теорему 2. ■

СЛЕДСТВИЕ. *Для строгой выпуклости дважды непрерывно дифференцируемой на (a, b) функции $f(t)$ достаточно, чтобы вторая производная $f''(t)$ была положительной на (a, b) всюду, за исключением конечного числа точек, в которых $f''(t)$ обращается в ноль.*

2^o. Очевидно, что квадратичная функция $q(t) = at^2 + bt + c$ строго выпукла на $\mathbb{R}$ при $a > 0$. линейна при $a = 0$ и строго вогнута на $\mathbb{R}$ при $a < 0$. Аналогичный по характеру результат справедлив для дробно-рациональной функции $h(t)$ вида

$$h(t) = \frac{at^2 + bt + c}{t - p},$$

рассматриваемой на множестве $U = \{t \in \mathbb{R} \,|\, t > p\}$.

Положим $A = ap^2 + bp + c$.

ТЕОРЕМА 5. *Функция $h(t)$ строго выпукла на U при $A > 0$, линейна при $A = 0$ и строго вогнута на U при $A < 0$.*

Доказательство. По формуле Лейбница имеем

$$h''(t) = \frac{2a}{t-p} - \frac{2(2at+b)}{(t-p)^2} + \frac{2(at^2+bt+c)}{(t-p)^3} =$$

$$= 2(t-p)^{-3}[a(t-p)^2 - (2at+b)(t-p) + at^2 + bt + c] =$$

$$= 2\frac{ap^2+bp+c}{(t-p)^3} = \frac{2A}{(t-p)^3}.$$

Отсюда и из теоремы 3 следует, что функция $h(t)$ строго выпукла на U при $A > 0$ и строго вогнута на U при $A < 0$.

Пусть $A = 0$. Тогда $c = -ap^2 - bp$ и

$$h(t) = \frac{at^2 + bt - ap^2 - bp}{t-p} = a(t+p) + b.$$

Теорема доказана. ■

Приведем таблицу функций одной переменной, строго выпуклых на интервале (a, b).

ТАБЛИЦА СТРОГО ВЫПУКЛЫХ ФУНКЦИЙ

Вид функции	Интервал (a, b)
t^{2n} при $n = 1, 2, \ldots$	$(-\infty, +\infty)$
$\sqrt{c^2 + t^2}$ при $c \neq 0$	$(-\infty, +\infty)$
$-\sqrt{1 - t^2}$	$(-1, 1)$
$-\ln t$	$(0, +\infty)$
$t \ln t$	$(0, +\infty)$
$-\sin t$	$(0, \pi)$
$-\ln \sin t$	$(0, \pi)$
$\mathrm{tg}\, t$	$(0, \frac{\pi}{2})$
$-\ln \mathrm{tg}\, t$	$(0, \frac{\pi}{4})$
$-\ln(1 - t^2)$	$(-1, 1)$

3^o. Введем обозначение $x = (t_1, \ldots, t_n)$ и пусть

$$V = \{x \,|\, t_1 > 0, \ldots, t_n > 0\}.$$

ТЕОРЕМА 6. *Функция*

$$f(x) = -\sum_{k=1}^{n} \alpha_k \ln t_k, \tag{6}$$

где α_k — положительные константы, строго выпукла на V.

Доказательство. Зафиксируем на множестве V две различные точки $x_0 = (t_1^{(0)}, \ldots, t_n^{(0)})$ и $x_1 = (t_1^{(1)}, \ldots, t_n^{(1)})$. Имеем

$$\langle f'(x_1) - f'(x_0), x_1 - x_0 \rangle = -\sum_{k=1}^{n} \alpha_k \left(\frac{1}{t_k^{(1)}} - \frac{1}{t_k^{(0)}} \right) (t_k^{(1)} - t_k^{(0)}) =$$

$$= \sum_{k=1}^{n} \alpha_k (t_k^{(0)} t_k^{(1)})^{-1} (t_k^{(1)} - t_k^{(0)})^2 > 0.$$

Остается сослаться на теорему 2. ■

Теорему 6 можно переформулировать иначе: *функция*

$$F(x) = t_1^{\alpha_1} t_2^{\alpha_2} \ldots t_n^{\alpha_n}$$

является логарифмически строго вогнутой на V.

В связи с теоремой 6 сделаем общее замечание. Пусть $f_k(t)$, $k \in 1:n$, — строго выпуклые при $t > 0$ функции. Тогда функция n переменных

$$F(x) = \sum_{k=1}^{n} \alpha_k \, f_k(t_k),$$

где α_k — положительные константы, будет строго выпуклой на множестве V. Это следует из определений.

Рассмотрим определитель Вандермонда

$$W(x) = \prod_{1 \le j < k \le n} (t_k - t_j)$$

на множестве $U = \{x \,|\, t_1 < t_2 < \ldots < t_n\}$.

ТЕОРЕМА 7. *Определитель Вандермонда $W(x)$ является логарифмически вогнутой на U функцией.*

Доказательство. Нужно проверить, что функция

$$F(x) = -\ln W(x) = -\sum_{1 \le j < k \le n} \ln(t_k - t_j)$$

выпукла на U. Возьмём на U две точки $x_0 = (t_1^{(0)}, \dots, t_n^{(0)})$, $x_1 = (t_1^{(1)}, \dots, t_n^{(1)})$ и обозначим $h_k = t_k^{(1)} - t_k^{(0)}$. Имеем

$$\langle F'(x_1) - F'(x_0), x_1 - x_0 \rangle =$$

$$= -\sum_{k=1}^{n} h_k \sum_{j=1,\, j\neq k}^{n} \left(\frac{1}{t_k^{(1)} - t_j^{(1)}} - \frac{1}{t_k^{(0)} - t_j^{(0)}} \right) =$$

$$= \sum_{k=1}^{n} \sum_{j=1,\, j\neq k}^{n} \frac{h_k(h_k - h_j)}{(t_k^{(1)} - t_j^{(1)})(t_k^{(0)} - t_j^{(0)})} =$$

$$= (\sum_{k=1}^{n-1} \sum_{j=k+1}^{n} + \sum_{j=1}^{n-1} \sum_{k=j+1}^{n}) \frac{h_k(h_k - h_j)}{(t_k^{(1)} - t_j^{(1)})(t_k^{(0)} - t_j^{(0)})} =$$

$$= \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \frac{h_k(h_k - h_j) + h_j(h_j - h_k)}{(t_k^{(1)} - t_j^{(1)})(t_k^{(0)} - t_j^{(0)})} =$$

$$= \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \frac{(h_j - h_k)^2}{(t_j^{(1)} - t_k^{(1)})(t_j^{(0)} - t_k^{(0)})} \geq 0.$$

Теорема доказана. ■

К сожалению, определитель Вандермонда не будет логарифмически *строго* вогнутым на U. Действительно, рассмотрим прямую $x = x(\lambda) := = x_0 + \lambda e$, где x_0 — некоторая точка из U и e — вектор, все компоненты которого равны единице. Очевидно, что $x(\lambda) \in U$ при всех $\lambda \in \mathbb{R}$. При этом

$$W(x(\lambda)) = \prod_{1\leq j<k\leq n} \left[(t_k^{(0)} + \lambda) - (t_j^{(0)} + \lambda) \right] \equiv W(x_0).$$

Для логарифмически строго вогнутой функции такое невозможно.

Чтобы исправить положение, введем функцию

$$T(x) = W(0, t_1, \dots, t_n) = \prod_{k=1}^{n} t_k \prod_{1\leq j<k\leq n} (t_k - t_j).$$

Таким образом, в определителе Вандермонда один узел зафиксирован.

ТЕОРЕМА 8. *Функция $T(x)$ является логарифмически строго вогнутой на множестве $U_0 = \{x \,|\, 0 < t_1 < t_2 < \ldots < t_n\}$.*

Доказательство. Имеем

$$F(x) := -\ln T(x) = -\sum_{k=1}^{n} \ln t_k - \sum_{1\leq j<k\leq n} \ln(t_k - t_j).$$

На множестве U_0 функция $F_1(x) = -\sum_{k=1}^{n} \ln t_k$ согласно теореме 6 строго выпукла, а функция

$$F_2(x) = -\sum_{1 \le j < k \le n} \ln(t_k - t_j)$$

в силу теоремы 7 просто выпукла. Очевидно, что их сумма $F(x) = F_1(x)+$ $+F_2(x)$ будет строго выпуклой на U_0. Теорема доказана. ■