Лекция 13. Выпуклые функции и формула Тейлора

1 Выпуклые и вогнутые C^2-гладкие функции.

Определение 1 *Функция называется выпуклой (вогнутой), если ее надграфик (подграфик) выпуклая область.*

Пример 1 *x^2 и $|x|$ - выпуклые, $-x^2$ и $-|x|$ - вогнутые на $\mathbb{R}$;*
$\sin x$ *- выпуклая функция на $[0, \pi]$, и вогнутая - на $[-\pi, 0]$.*

Определение 2 *C^2-гладкая функция на отрезке $[a, b]$ называется выпуклой (вогнутой), если $f'' \geq 0$ (соответственно, $f'' \leq 0$) на $[a, b]$.*

Теорема 1 *Для C^2-гладких функций определения 1 и 2 равносильны.*

Доказательство Пусть $f'' \geq 0$ на I. Возьмем $a, b \in I$ и точки $A = (a, c)$ и $B = b, d$, принадлежащие надграфику Γ^+ функции f над I. Докажем, что $[A, B] \subset \Gamma^+$. Пусть $A' = (a, f(a)), B' = (b, f(b))$. Поскольку надграфик вместе с каждой точкой содержит положительно направленный вертикальный луч с вершиной в этой точке, достаточно доказать, что $[A', B'] \subset \Gamma^+$. Пусть L - линейная функция, график которой проходит через точки $A', B_,$, и пусть

$$g = f - L.$$

Тогда $g(a) = g(b) = 0$, и $g'' = f'' \geq 0$. Функция g не может принимать положительных значений на (a, b). В противном случае, ее максимум C находился бы внутри $[a, b]$, и в нем выполнялись бы соотношения:

$$g'(c) = 0, \;\; g(c0 > 0,$$

поскольку $g(b) = 0$. Тогда, по теореме Лагранжа для функции g, на интервале (c, b) нашлась бы точка C_1 такая, что

$$g'(c_1) < 0.$$

Снова по теореме Лагранжа для g' на отрезке $[C, C_1]$ нашлась бы точка C_2, в котрой $g''(C_2) < 0$ - противоречие.

Пусть теперь $f : I \to \mathbb{R}$ - класса C^2, и надграфик f выпуклый. Докажем, что $f'' \geq 0$ на I. Предположим, противное:

$$\exists\, C \in I : f''(C) < 0.$$

В силу непрерывности функции f, существует отрезок $[a, b] \ni C$, на котором $f'' < 0$. По доказанному выше, *подграфик* функции над выпуклый. Надграфик функции f над $[a, b]$ может быть тоже выпуклым, только если общая граница надграфика и подграфика - график функции f на $[a, b]$ - отрезок. Но в этом случае $f'' \equiv 0$ на $[a, b]$ - противоречие. $\square$

2 Единственность экстремума.

Определение 3 *C^2-функция на интервале I называется строго выпуклой, если $f'' > 0$ на I.*

Теорема 2 *Строго выпуклая функция на интервале имеет не более одного минимума, и ни одного максимума.*

Доказательство Пусть строго выпуклая функция f на интервале имеет две критические точки. Тогда по теореме Ролля, примененной к f', эта функция имеет критическую точку. В ней $f'' = 0$ - противоречие.

В единственной критической точке функции f ее вторая производная положительна. Значит, эта точка - минимум по теореме 4 лекции 12. $\square$

3 Среднее арифметическое.

Теорема 3 *Пусть $f : I \to \mathbb{R}$ - выпуклая функция класса $C^2 : f'' > 0$. Тогда для любых $a, b \in I$,*

$$f\left(\frac{a+b}{2}\right) \leq \frac{f(a)+f(b)}{2}. \tag{1}$$

Доказательство Точка $C = \left(\frac{a+b}{2}, \ \frac{f(a)+f(b)}{2}\right)$ - середина отрезка с вершинами $A = (a, f(a))$ и $B = b, f(b)$. Эти точки лежат на графике функции f. По теореме 1 весь отрезок $[A, B]$ принадлежит надграфику f. Отсюда следует (1). $\square$

Это соотношение допускает сильное обобщение - неравенство Йенсена. Оно включено в занятие как задача.

Теорема 4 *Пусть положительные числа $q_1, \ldots, q_n$ удовлетворяют соотношению $q_1 + \ldots + q_n = 1$. Тогда для любых $x_1, \ldots, x_n \in I$ выполнено неравенство*

$$f(q_1 x_1 + \ldots + q_n x_n) \leq q_1 f(x_1) + \ldots + q_n f(x_n).$$

Мы переходим к центральному факту дифференциального исчисления - формуле Тейлора. Она позволяет проводить исчерпывающий локальный анализ для подавляющего большинства функций одного переменного. Пространство функций, имеющих n непрерывных производных (от первой до nй) обозначается C^n. Функции, имеющие бесконечное число производных, называются бесконечно дифференцируемыми; пространство таких функций обозначается C^∞.

4 Многочлен Тейлора или тейлоровский многочлен.

Названный многочлен решает следующую задачу: для данной функции, данной точки и данного n найти такой многочлен степени n, все производные которого до порядка n в этой точке равны соответствующим производным данной функции в данной точке. Тейлоровский многочлен степени n, приближающий функцию f в точке a, как описано выше, обозначается через $T_{a,n}(x)$. Этот многочлен удовлетворяет требованию:

$$f^{(k)}(a) = T_{a,n}(k)(a), \ k = 0, 1, ..., n. \tag{2}$$

Напомним, что нулевая производная функции - это она сама.

Поставленную задачу решает формула Тейлора, которую мы сейчас выведем. Без ограничения общности, будем считать, что $a - 0$. Ключевое замечание: моном x^k имеет только одну производную, отличную от нуля в нуле, и она равна $k!$:

$$(x^k)^{(m)}(0) = \delta_{km} k!, \tag{3}$$

здесь δ_{km} - символ Кронекера: $\delta_{km} = 0$ при $k \neq m, \delta_{kk} = 1$. Проверьте!

Будем искать многочлен $T = T_{0,n}$ в виде:

$$T(x) = a_0 + a_1 x + ... + a_n x^n. \tag{4}$$

По формуле (3),

$$T^{(k)}(0) = a_k k!.$$

Следовательно, по формуле (2),

$$a_k = \frac{f^(k)(0)}{k!}. \tag{5}$$

Следовательно, задача (2) имет единственное решение:

$$T_{0,n}(x) = f(0) + \frac{f'(0)}{1} x + ... + \frac{f^{(n)}}{n!} x^n. \tag{6}$$

Возникает естественный вопрос: с какой точностью приближает функцию ее тейлоровский многочлен? Другими словами, насколько быстро убывает остаточный член - разность между функцией и ее тейлоровским многочленом?

5 Формула Тейлора с остаточным членом в форме Пеано.

Опуская детали, можно сказать, что тейлоровский многочлен приближает соответствующую функцию с точностью до $o(x^n)$ в нуле. Более точно это формулируется так.

Теорема 5 *Пусть* $f : \mathbb{R} \to \mathbb{R}$ *- функция класса* $C^n, T = T_{0,n}$ *- ее тейлоровский многчлен степени* n *в нуле. Тогда*

$$f(x) = T(x) + o(x^n). \tag{7}$$

Доказательство Доказательство проводится индукцией по n.

База индукции: $n = 1$. При $n = 1$ формула (7) превращается в определние производной, и тем самым, верна:

$$f(x) = f(0) + f'(0)x + o(x).$$

Шаг индукции: переход от $n-1$ к n. Обозначим остаточный член через $R: \; R(x) = f(x) - T(x)$. По формуле (2),

$$R(0) = R'(0) = ... = R^{(n)}(0) = 0. \tag{8}$$

Докажем, что отсюда следует утверждение теоремы: $R(x) = o(x^n)$. Положим: $Q = R'$. Эта функция удовлетворяет условию:

$$Q(0) = Q'(0) = ... = Q^{(n-1)}(0) = 0.$$

По предположению индукции,

$$Q(x) = o(x^{(n-1)}),$$

то есть

$$\lim_{x \to 0} \frac{Q(x)}{x^{n-1}} = 0.$$

Теперь воспользуемся правилом Лопиталя:

$$\lim_{x \to 0} \frac{R(x)}{x^n} = \lim_{x \to 0} \frac{Q(x)}{nx^{n-1}} = 0.$$

Последнее равенство следует из предположения индукции. Теорема доказана. □

6 Локальное исследование функций одного переменного.

Формула Тейлора полностью решает вопрос о локальном исследовании функций при условии, что ни в одной точке все производные функции не обращаются одновременно в ноль.

Определение 4 C^∞ *- функция называется* плоской *в точке* a*, если все ее производные в этой точке равны* 0.

Пример 2 *Функция* $2^{-\frac{1}{|x|}}$ *- плоская в нуле.*

Теорема 6 *Пусть* $f - C^{\infty}$ *- функция с критической точкой* 0*, не плоская в нуле. Пусть* $f^{(k)}$ *- первая ненулевая производная функции* f *в нуле. Тогда, если* k *четно, то* 0 *- локальный экстремум функции* f*; при этом, если* $f^{(k)} > 0 (< 0)$*, то* 0 *- локальный мимнимум (соответственно, локальный максимум) функции* f*. Если* k *нечетно, то* 0 *точка перегиба.*

Доказательство По теореме 5,

$$f(x) - f(0) = ax^k + o(x^k),\ a = \frac{f^{(k)}(0)}{k!}$$

или, что равносильно,

$$f(x) - f(0) = x^k(a + o(1)) \text{ при } x \to 0.$$

Функция $a + r(x)$ при $r = o(1)$ сохраняет тот же знак, что и a при малых x.

Поэтому, если k четно, то $x^k > 0$ на $\mathbb{R} \setminus \{0\}$, и разность $f(x) - f(0)$ - того же знака, что и $a = f^{(k)}(0)/k!$ при малых x. Это доказывает теорему для четных k.

Если k нечетно, то функция x^k меняет знак в нуле. Значит, график функции f в малой окрестности нуля лежит по разные стороны от своей касательной в нуле. Значит 0 - точка перегиба для $f(x)$ в нуле. $\square$