Лекция 8. Производная

Как известно, производной функции $f\colon (a;b) \to \mathbb{R}$ в точке $x \in (a;b)$ называется предел

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$$

(если он существует). Производная функции f в точке x обозначается $f'(x)$, а функция, имеющая производную в точке x, называется *дифференцируемой* в этой точке.

Легко видеть, что число c является производной функции f в точке x тогда и только тогда, когда разность

$$R(x) = f(x+h) - f(x) - ch$$

обладает тем свойством, что $\lim_{h\to 0}(R(h)/|h|) = 0$. Из этой переформулировки следует, что если функция $f\colon (a;b) \to \mathbb{R}$ дифференцируема в точке x, то она и непрерывна в точке x.

В математике распространено следующее (несколько вольное) сокращенное обозначение: всякая функция от h, обладающая тем свойством, что ее отношение к $|h|$ стремится к нулю при $h \to 0$, обозначается $o(h)$. (Аналогично, например, про последовательность $\{a_n\}$, для которой $\lim_{n\to\infty} a_n/n^2 = 0$, говорят, что $a_n = o(n^2)$.)Пользуясь этим обозначением, можно записать определение производной в виде

$$f(x+h) = f(x) + ch + o(h) \qquad \text{при } h \to 0.$$

Вот пример того, как работает переформулировка с $o(h)$.

Предложение 8.1 (Теорема Ферма). *Пусть $f\colon (a;b) \to \mathbb{R}$ — дифференцируемая функция, и пусть $x \in (a;b)$ — точка локального экстремума функции f. Тогда $f'(x) = 0$.*

Доказательство. Покажем, что если $f'(x) = c \neq 0$, то в точке x локального экстремума быть не может. В самом деле, пусть $R(x) = f(x+h) - f(x) - ch$; тогда существует такое $\varepsilon > 0$, что при $|h| < \varepsilon$ имеем $|R(h)| \leqslant |c| \cdot |h|/2$. Следовательно, при $|h| < \varepsilon$ знак разности $f(x+h) - f(h)$ такой же, как у числа ch, так что, например, при $c > 0$ имеем $f(x+h) > f(x)$, если $0 < h < \varepsilon$ и $f(x+h) < f(x)$, если $-\varepsilon < h < 0$. Значит, x не является экстремумом. $\square$

С помощью этого же рассуждения (вкупе с соображениями, примененными нами при доказательстве компактности и связности отрезка) можно показать, что функция, производная которой на отрезке положительна, явдляется возрастающей; ниже мы наметим другое доказательство этого факта.

Перечислим известные из школы свойства производных:

Предложение 8.2. *1) Если функция $f\colon (a;b) \to \mathbb{R}$ постоянна, то $f'(c) = 0$ для всех $c \in (a;b)$.*

2) Пусть $f\colon (a;b) \to \mathbb{R}$ и $g\colon (a;b) \to \mathbb{R}$ — функции, дифференцируемые в точке c, и пусть $\lambda, \mu \in \mathbb{R}$. Тогда функция $\lambda f + \mu g$ дифференцируема в точке c и ее производная равна $\lambda f'(c) + \mu g'(c)$.

3) Пусть $f\colon (a;b) \to \mathbb{R}$ и $g\colon (a;b) \to \mathbb{R}$ — функции, дифференцируемые в точке c. Тогда функция fg дифференцируема в точке c и ее производная равна $f'(c)g(c) + f(c)g'(c)$.

4) Пусть $f\colon (a;b) \to \mathbb{R}$ и $g\colon (a;b) \to \mathbb{R}$ — функции, дифференцируемые в точке c, и пусть $g(c) \neq 0$. Тогда функция f/g дифференцируема в точке c и ее производная равна $(f'(c)g(c) - f(c)g'(c))/(g(c)^2)$.

Предложение 8.3. *Пусть $g\colon (a;b) \to (c;d) \subset \mathbb{R}$ и $f\colon (c;d) \to \mathbb{R}$ — две функции, причем g дифференцируема в точке $p \in (a;b)$, а f дифференцируема в точке $g(p) \in (c;d)$. Тогда* композиция *$f \circ g\colon (a;b) \to \mathbb{R}$, действующая по правилу $(f \circ f)(x) = f(g(x))$, дифференцируема в точке p, и ее производная в точке p равна $f'(g(p)) \cdot g'(p)$.*

Вы легко докажете эти факты самостоятельно (для предложения 8.3 удобно воспользоваться переформулировкой с $o(|h|)$).

В связи с понятием дифференцируемой функции уместно ввести следующие полезные обозначения:

Обозначение 8.4. Пусть $(a;b)$ — открытый интервал на вещественной оси и $r > 0$ — целое число. Будем говорить, что функция *принадлежит классу C^r*, если она r раз дифференцируема на $(a;b)$ и ее r-я производная непрерывна (непрерывные функции будем называть функциями класса C^0). Если функция принадлежит классу C^r для всех $r \geqslant 0$, то будем говорить, что функция *принадлежит классу C^∞* (или бесконечно дифференцируема).

Отметим еще, что r-я производная функции f в точке x обозначается $f^{(r)}(x)$.

Основной технический факт о производных — «теорема о среднем» — является по существу утверждением о кривых на плоскости.

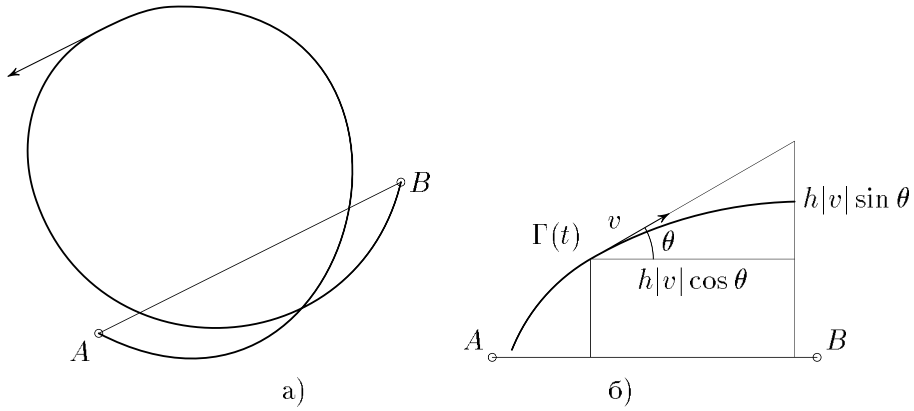

Рис. 8.3. Теорема о среднем и ее доказательство.

Непрерывное отображение интервала $(a;b)$ в $\mathbb{R}^2$ будем называть параметризованной плоской кривой. Пусть это отображение задается в виде $\Gamma\colon t \mapsto (\varphi(t), \psi(t))$; если функции φ и ψ дифференцируемы и при этом производные $\varphi'(t)$ и $\psi'(t)$ никогда не обращаются в нуль одновременно (иными словами, «вектор скорости» $(\varphi'(t), \psi'(t))$ нигде не обращается в нуль), то Γ (или $\Gamma(a;b)$) называется гладкой иммерсированной кривой. Условие необращения в нуль вектора скорости гарантирует, что у кривой нет «изломов».

Теорема 8.5 (теорема о среднем). *Пусть $\Gamma\colon t \mapsto (\varphi(t), \psi(t))$ — такое непрерывное отображение отрезка $[a;b]$ в $\mathbb{R}^2$, что функции φ и ψ дифференцируемы на $(a;b)$, причем для всех t из интервала $(a;b)$ имеем $\Gamma'(t) = (\varphi'(t), \psi'(t)) \neq (0,0)$. Положим $\Gamma(a) = A$, $\Gamma(b) = B$, и пусть $A \neq B$ Тогда для некоторого $t \in (a;b)$ вектор $\Gamma'(t)$ параллелен прямой AB.*

(См. рис. 8.3а.)

Доказательство. Ввиду компактности отрезка существует такое $t \in [a;b]$, что расстояние от $\Gamma(t)$ до прямой AB максимально (поскольку при t, равном a ил b, это рассточние равно нулю, максимум обязательно реализуется для какого-то $t \in (a;b)$). Покажем, что это t — искомое.

В самом деле, в противном случае касательный вектор в точке t (обозначим его $v = (\varphi'(t), \psi'(t))$) образует с отрезком AB ненулевой угол θ. Имеем $\Gamma(t+h) = \Gamma(t) + hv + o(h)$. Следовательно (если обозначить через d расстояние до прямой AB) имеем

$$d(\Gamma(t+h)) = d(\Gamma(t)) + h|v|\sin\theta + O(|h|)$$

(см. рис. 8.3б), так что для достаточно малого h подходящего знака имеем $d(\Gamma(t+h)) > d(\Gamma(t))$ — противоречие. □

Следствие 8.6 (теорема Лагранжа). *Пусть непрерывная функция* $f\colon [a;b] \to \mathbb{R}$ *дифференцируема на* $(a;b)$. *Тогда существует такое число* $\xi \in (a;b)$, *что* $f(b) = f(a) + f'(\xi)(b-a)$.

(Положите $\varphi(t) = t$, $\psi(t) = f(t)$.)

Следствие 8.7 (теорема Ролля). *Пусть непрерывная функция* $f\colon [a;b] \to \mathbb{R}$ *дифференцируема на* $(a;b)$. *Если* $f(a) = f(b)$, *то существует такое число* $\xi \in (a;b)$, *что* $f'(\xi) = 0$. □

Из теоремы Лагранжа немедленно следуют достаточные условия возрастания/убывания («если производная положительна, то функция возрастает...») и достаточные условия экстремума.

Основной факт о производных функций одного переменного — формула Тейлора. Начнем с мотивировок.

Пусть $f\colon \mathbb{R} \to \mathbb{R}$ — «достаточно хорошая» функция (в частности, мы будем предполагать, что у нее есть столько производных, сколько нам нужно). Попробуем «разложить f в степенной ряд» — найти для нее представление вида

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots. \tag{8.1}$$

Сумма в правой части бесконечна, и понимать эту запись следует так:

$$f(x) = \lim_{n\to\infty}(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n).$$

Правда, такое разложение вполне может не существовать (даже для бесконечно дифференцируемых функций!), но мы это проигнорируем и попробуем найти его коэффициенты. Проще всего найти a_0: полагая $x = 0$ в (8.1), получаем $a_0 = f(0)$. Чтобы найти a_1, продифференцируем обе части (8.1):

$$f'(x) = a_1 + 2a_2x + 3a_2x^2 + \cdots \tag{8.2}$$

(Такое дифференцирование ряда не всегда законно, но мы и на это не посмотрим.) Полагая теперь $x = 0$ в (8.2), находим, что $a_1 = f'(0)$. Аналогичным образом находим и прочие коэффициенты: продифференцировав ряд (8.1) n раз, получаем

$$f^{(n)}(x) = n!a_n + c_{n+1}x^{n+1} + c_{n+2}x^{n+2} + \cdots,$$

и полагая в этом равенстве $x = 0$, получаем $a_n = f^{(n)}(0)/n!$.

Итак, мы нашли формулу для коэффициентов разложения функции в степенной ряд. Оказывается, что хотя самого́ разложения может и не существовать, найденные нами коэффициенты всегда дают если не ряд, то по крайней мере асимптотику.

Теорема 8.8. *Пусть функция $f\colon (A;B) \to \mathbb{R}$ является $n+1$ раз дифференцируемой. Тогда при $a, x \in (A;B)$ имеем*

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \\ + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}, \tag{8.3}$$

где $\xi \in \mathbb{R}$ — некоторое число, лежащее между a и x.

Следствие 8.9. *Пусть $f\colon (A;B) \to \mathbb{R}$ — функция класса C^{n+1}, и пусть $a \in (A;B)$. Тогда*

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \dots \\ + \frac{f^{(n)}(a)}{n!}(x-a)^n + o(|x-a|^n) \quad \text{при } x \to a. \tag{8.4}$$

Запись с $o(|x-a|^n)$ означает, как водится, что, если обозначить через $R(x)$ разность $f(x) - \Big(f(a) + f'(a)(x-a) + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n\Big)$, то будет иметь место соотношение

$$\lim_{x \to a} R(x)/|x-a|^n = 0.$$

Доказательство теоремы 8.8. Поскольку $(n+1)$-я производная функции не изменится, если вычесть из нее многочлен степени n, можно заменить $f(x)$ на $R(x)$ и считать тем самым, что $f(a) = f'(a) = \ldots = f^{(n)}(a) = 0$; доказать при этом надо, что $f(x) = f^{(n+1)}(\xi)/(x-a)^{n+1}$ для некоторого $\xi \in (a;x)$. Теперь, считая, что $x > a$ (при $x < a$ рассуждение аналогично), применим теорему 8.5 (о среднем) для случая $b = x$, $\varphi(x) = (x-a)^{n+1}$, $\psi(x) = f(x)$. Из теоремы вытекает, что для некоторого $\xi_1 \in (a;x)$ имеем $\psi'(\xi)/\varphi'(\xi) = \frac{\psi(b)-\psi(a)}{\varphi(b)-\varphi(a)}$, т. е.

$$\frac{f'(\xi_1)}{(n+1)(\xi_1-a)^n} = \frac{f(x)}{(x-a)^{n+1}}.$$

Применим еще раз ту же теорему, положив $b=\xi_2$, $\varphi(x)=(x-a)^n/n!$, $\psi(x)=f'(x)$; получим, что для некоторого $\xi_2\in(a;\xi_1)$ выполнено равенство

$$\frac{f''(\xi_2)}{n(n+1)(\xi_2-a)^{n-1}}=\frac{f'(\xi_1)}{(n+1)(\xi_1-a)^n}=\frac{f(x)}{(x-a)^{n+1}}.$$

Продолжим этот процесс; когда дойдем до $(n+1)$-й производной, окажется, что для некоторого $\xi_{n+1}\in(a;\xi_n)\subset(a;x)$ имеем $f^{(n+1)}(\xi_{n+1})/(n+1)!=f(x)/(x-a)^{n+1}$. Остается положить $\xi=\xi_n$. $\square$

Формула (8.3) называется формулой Тейлора с остаточным членом в форме Лагранжа. При $n=0$ она превращается в теорему Лагранжа.

Формула (8.4) называется формулой Тейлора с остаточным членом в форме Пеано. Для ее справедливости достаточно, чтобы f принадлежала к классу C^r.