§ 7.5. Дифференцируемая функция. Касательная плоскость

Для простоты будем рассматривать трехмерный случай; в n-мерном случае рассуждения аналогичны. Случай $n=1$ был специально рассмотрен в § 5.2.

Пусть на открытом множестве $G \subset R_3$ задана функция

$$u = f(x, y, z),$$

имеющая в точке $(x, y, z) \in G$ непрерывные частные производные первого порядка. Отсюда автоматически следует, что эти частные производные существуют в некоторой окрестности (x, y, z), хотя, быть может они в точках, отличных от (x, y, z), не являются непрерывными. Рассмотрим приращение f в (x, y, z), соответствующее приращению $(\Delta x, \Delta y, \Delta z)$, где $|\Delta x|$, $|\Delta y|$, $|\Delta z| < \delta$ и δ достаточно мало́, чтобы точка $(x+\Delta x, y+\Delta y, z+\Delta z)$ не выходила из указанной окрестности. Имеют место равенства (пояснения ниже):

$$\Delta u = f(x+\Delta x, y+\Delta y, z+\Delta z) - f(x, y, z) = \tag{1}$$

$$= f(x+\Delta x, y+\Delta y, z+\Delta z) - f(x, y+\Delta y, z+\Delta z) + \tag{2}$$

$$+ f(x, y+\Delta y, z+\Delta z) - f(x, y, z+\Delta z) + \tag{3}$$

$$+ f(x, y, z+\Delta z) - f(x, y, z) = \tag{4}$$

$$= f'_x(x+\theta_1\Delta x, y+\Delta y, z+\Delta z)\,\Delta x + f'_y(x, y+\theta_2\Delta y, z+\Delta z)\Delta y + \\ + f'_z(x, y, z+\theta_3\Delta z)\,\Delta z = \tag{5}$$

$$= \left(f'_x(x, y, z) + \varepsilon_1\right)\Delta x + \left(f'_y(x, y, z) + \varepsilon_2\right)\Delta y + \\ + \left(f'_z(x, y, z) + \varepsilon_3\right)\Delta z = \tag{6}$$

$$= f'_x(x, y, z)\,\Delta x + f'_y(x, y, z)\,\Delta y + f'_z(x, y, z)\,\Delta z + o(\rho) \quad (\rho \to 0), \tag{7}$$

$$0 < \theta_1, \theta_2, \theta_3 < 1, \; \rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}, \; \varepsilon_1, \varepsilon_2, \varepsilon_3 \to 0 \quad (\rho \to 0). \tag{8}$$

Отметим, что соотношение $\rho \to 0$ эквивалентно трем соотношениям: $\Delta x \to 0$, $\Delta y \to 0$, $\Delta z \to 0$.

Переход от (2) к первому члену (5) обосновывается так: функция $f(\xi, y+\Delta y, z+\Delta z)$ от ξ (при фиксированных $y+\Delta y$, $z+\Delta z$) имеет, по условию, производную (по ξ) на отрезке $[x, x+\Delta x]$ и к ней применима теорема Лагранжа о среднем. Аналогичное пояснение ко второму и третьему членам (5). Переход от (5) к (6) чисто формальный: мы положили, например,

$$f'_x(x+\theta_1\Delta x, y+\Delta y, z+\Delta z) = f'_x(x, y, z) + \varepsilon_1.$$

Но не формален здесь факт, что $\varepsilon_1 \to 0$ при $\rho \to 0$. Он следует из предположенной непрерывности f'_x в (x, y, z). Наконец, переход от (6) к (7) сводится к утверждению, что имеет место равенство

$$\varepsilon_1\Delta x + \varepsilon_2\Delta y + \varepsilon_3\Delta z = o(\rho) \quad (\rho \to 0).$$

В самом деле (см. § 6.2, (9)) при $\rho \to 0$

$$|\varepsilon_1\Delta x + \varepsilon_2\Delta y + \varepsilon_3\Delta z|/\rho \leqslant \rho\sqrt{\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2}/\rho = \sqrt{\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2} \to 0.$$

Мы доказали важную теорему:

Теорема 1. *Если функция $u = f$ имеет непрерывные частные производные (первого порядка) в точке (x, y, z), то ее приращение в этой точке, соответствующее достаточно малому приращению $(\Delta x, \Delta y, \Delta z)$, можно записать по формуле*

$$\Delta u = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{\partial f}{\partial z}\Delta z + o(\rho)\ (\rho \to 0), \quad \rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}, \tag{9}$$

где частные производные взяты в точке (x, y, z).

Так как значения частных производных в правой части (9) не зависят от Δx, Δy, Δz, то из условий теоремы 1 следует, что приращение f в (x, y, z), соответствующее приращению $(\Delta x, \Delta y, \Delta z)$, может быть записано по формуле

$$\Delta u = A\Delta x + B\Delta y + C\Delta z + o(\rho) \quad (\rho \to 0), \tag{10}$$

где числа A, B, C не зависят от Δx, Δy, Δz.

Сделаем следующее определение: если приращение функции f в точке (x, y, z) для достаточно малых $(\Delta x, \Delta y, \Delta z)$ может быть записано в виде суммы (10), где A, B, C — числа, не зависящие от Δx, Δy, Δz, то говорят, что функция f *дифференцируема* в точке (x, y, z). Таким образом, *дифференцируемость функции f в (x, y, z) заключается в том, что ее приращение Δf в этой точке можно записать в виде суммы двух слагаемых: первое слагаемое есть линейная функция $A\Delta x + B\Delta y + C\Delta z$ от $(\Delta x, \Delta y, \Delta z)$ — она называется главной линейной частью приращения Δf, второе же слагаемое, вообще, сложно зависит от приращений Δx, Δy, Δz, но если стремить их к нулю, то оно будет стремиться к нулю, быстрее, чем $\rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$.*

Легко видеть, что если функция f дифференцируема в точке (x, y, z), т. е. представляется равенством (10), то она имеет в этой точке производные, равные:

$$\frac{\partial f}{\partial x} = A, \quad \frac{\partial f}{\partial y} = B, \quad \frac{\partial f}{\partial z} = C. \tag{11}$$

Например, первое равенство (11) доказывается так. Пусть приращение f в (x, y, z) записывается по формуле (10). Если считать в последней $\Delta x = h$, $\Delta y = \Delta z = 0$, то получим равенство $\Delta_{xh}u = Ah + o(h)$ $(h \to 0)$. После деления его на h и перехода к пределу, получим

$$\lim_{h\to 0}\frac{\Delta_{xh}u}{h} = \frac{\partial f}{\partial x} = A.$$

Из сказанного следует

Теорема 2. *Для того чтобы функция f была дифференцируемой в точке, необходимо, чтобы она имела в этой точке частные производные, и достаточно, чтобы она имела в этой точке непрерывные частные производные.*

Из (10) следует, что *если функция дифференцируема в точке, то она непрерывна в этой точке.*

Пример 1. Функция $f(x, y, z)$, равная нулю на координатных плоскостях $x = 0$, $y = 0$, $z = 0$ и единице в остальных точках R_3, имеет, очевидно, частные производные, равные нулю в точке $(0, 0, 0)$, но она, очевидно, разрывна в этой точке и потому не может быть в ней дифференцируемой. Таким образом, одного существования частных производных в точке недостаточно для дифференцируемости и даже непрерывности в этой точке.

Отметим отличие многомерного случая от одномерного. При $n = 1$ свойство дифференцируемости f в x записывается в виде равенства $\Delta f = A\Delta x + o(\Delta x)$, следовательно, если $A \neq 0$, то остаток стремится к нулю при $\Delta x \to 0$ быстрее главной части. При $n > 1$ это уже не так; например, при $n = 3$, каковы бы ни были числа A, B, C, одновременно не равные нулю, всегда можно стремить Δx, Δy, Δz к нулю так, чтобы при этом постоянно выполнялось равенство $A\Delta x + B\Delta y + C\Delta z = 0$, но тогда в (10) остаточный член $o(\rho)$ вообще больше главного. Впрочем, если мы заставим Δx, Δy, Δz стремиться к нулю так, чтобы выполнялась пропорциональность $\Delta x : \Delta y : \Delta z = A : B : C$, то тогда главная часть приращения будет величиной, имеющей строго порядок ρ, и остаток будет стремиться к нулю быстрее главной части.

Если функция f дифференцируема в точке (x, y, z), то главная линейная часть ее приращения в этой точке называется еще *дифференциалом f в этой точке, соответствующим приращениям $(\Delta x, \Delta y, \Delta z)$ независимых переменных.*

Он записывается так: $df = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{\partial f}{\partial z}\Delta z$. О других обозначениях мы будем еще говорить.

Рассмотрим поверхность S, описываемую функцией $z = f(x, y)$, заданной в окрестности точки (x_0, y_0).

Плоскость L_0 называется *касательной плоскостью к поверхности S в ее точке $P_0 = (x_0, y_0, z_0)$ $(z_0 = f(x_0, y_0))$*, если расстояние $r(P, L_0)$ подвижной точки $P = (x, y, z) \in S$ до L_0 стремится к нулю быстрее расстояния ρ от P до P_0:

$$r(P, L_0) = o(\rho) \quad (\rho \to 0). \tag{12}$$

Теорема 3. *Если функция f дифференцируема в точке (x_0, y_0), то описываемая ею поверхность имеет, и притом единственную, касательную плоскость в точке P_0, определяемую уравнением*

$$z - z_0 = \left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0) \tag{13}$$

($(\ )_0$ обозначает, что в скобках надо положить $x = x_0$, $y = y_0$).

Доказательство. Пусть функция f дифференцируема в точке (x_0, y_0), S — описываемая ею поверхность и L_0 — плоскость, определяемая уравнением (13). Произвольная точка $P \in S$ имеет координаты $(x, y, f(x, y))$. Из аналитической геометрии известно, что ее расстояние до L_0 выражается формулой (пояснения ниже)

$$r(P, L_0) = \frac{1}{M}\left| f(x,y) - f(x_0, y_0) - \left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) - \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0)\right| = \\ = o(r) = o(\rho), \quad r \to 0, \ \rho \to 0, \quad (14)$$

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2},$$

$$\rho = \sqrt{(x - x_0)^2 + (y - y_0)^2 + [f(x, y) - f(x_0, y_0)]^2}.$$

Здесь $M = \sqrt{\left(\frac{\partial f}{\partial x}\right)_0^2 + \left(\frac{\partial f}{\partial y}\right)_0^2 + 1}$ есть нормирующий множитель плоскости L_0. Второе равенство (14) имеет место вследствие предположенной дифференцируемости f в точке (x_0, y_0). Последнее же равенство говорит, что величина вида $o(r)$ $(r \to 0)$ обладает тем свойством, что ее отношение к ρ стремится к нулю при $\rho \to 0$. Ведь если $\rho \to 0$, то тогда и $r \to 0$ $(0 \leqslant r \leqslant \rho!)$ и потому

$$\left|\frac{o(r)}{\rho}\right| = \left|\frac{o(r)}{r} \cdot \frac{r}{\rho}\right| \leqslant \left|\frac{o(r)}{r}\right| \to 0.$$

Мы доказали, что плоскость, определяемая уравнением (13), есть касательная плоскость к S в точке (x_0, y_0).

Другой касательной плоскости к S в точке (x_0, y_0) не существует. В самом деле, пусть касательная плоскость к S в P_0 имеет уравнение

$$A(x - x_0) + B(y - y_0) - C(z - z_0) = 0, \quad A^2 + B^2 + C^2 = 1. \quad (15)$$

Тогда

$$A(x - x_0) + B(y - y_0) - C(f - f_0) = o(\rho) = o(r), \quad \rho \to 0, \ r \to 0. \quad (16)$$

Последнее равенство в (16) объясняется следующим образом. В силу дифференцируемости f в точке (x_0, y_0)

$$|f - f_0| = \left|\left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0) + \varepsilon r\right| \leqslant \\ \leqslant r\sqrt{\left(\frac{\partial f}{\partial x}\right)_0^2 + \left(\frac{\partial f}{\partial y}\right)_0^2} + r|\varepsilon| < cr,$$

где c — константа, не зависящая от r. Надо учесть, что ε ограничено, потому что $\varepsilon \to 0$ при $r \to 0$. Но тогда

$$\rho = \sqrt{r^2 + |f - f_0|^2} < c_1 r,$$

где c_1 не зависит от r и $\rho \to 0$ при $r \to 0$, следовательно,

$$\left|\frac{o(\rho)}{r}\right| = \left|\frac{o(\rho)}{\rho}\,\frac{\rho}{r}\right| \leqslant c_1 \left|\frac{o(\rho)}{\rho}\right| \to 0, \quad r \to 0, \quad \rho \to 0.$$

Если положить в (16) $y = y_0$, разделить на $x - x_0$ и перейти к пределу при $x \to x_0$, то получаем

$$A - C\left(\frac{\partial f}{\partial x}\right)_0 = 0.$$

Аналогично, полагая в (16) $x = x_0$, деля на $y - y_0$ и переходя к пределу при $y \to y_0$, получим

$$B - C\left(\frac{\partial f}{\partial y}\right)_0 = 0.$$

Но тогда $C \neq 0$, потому что иначе было бы $A = B = C = 0$, и, следовательно, наша плоскость имеет вид (13).

Пример 2. Функция ($\alpha > 0$)

$$f(x, y) = \begin{cases} (x^2 + y^2)^{(1+\alpha)/2} & \text{в рациональных точках,} \\ 0 & \text{в остальных точках,} \end{cases}$$

очевидно, разрывна в любой точке, отличной от нулевой, в нулевой же точке она дифференцируема:

$$f(x, y) - f(0, 0) = f(x, y) = 0x + 0y + \rho^{1+\alpha},$$

где $\rho^{1+\alpha} = o(\rho)$ $(\rho \to 0)$. Таким образом, f есть пример функции, дифференцируемой в точке, но не имеющей непрерывных частных производных в этой точке.

Примеры 1 и 2 показывают, что свойство функции быть дифференцируемой в точке слабее свойства иметь непрерывные частные производные в точке, но сильнее свойства иметь частные производные в точке.

§ 7.6. Производная сложной функции; производная по направлению; градиент

Ограничимся рассмотрением функции трех переменных, определенной на открытом множестве $G \subset R_3$. Распространение излагаемых здесь фактов на n-мерный случай производится аналогично.

Теорема 1. *Пусть функция*

$$u = f(x, y, z) \tag{1}$$

дифференцируема в точке $(x, y, z) \in G$, *а функции*

$$x = \varphi(t), \ y = \psi(t), \ z = \chi(t), \tag{2}$$

зависящие от скалярного параметра t, *имеют производную в* t. *Тогда производная по* t *от сложной функции (производная от* f *вдоль кривой* (2)) $u = F(t) = f(\varphi(t), \psi(t), \chi(t))$ *вычисляется по формуле*

$$F'(t) = f'_x(\varphi(t), \psi(t), \chi(t))\varphi'(t) + f'_y(\varphi(t), \psi(t), \chi(t))\psi'(t) + \\ + f'_z(\varphi(t), \psi(t), \chi(t))\chi'(t),$$

или, короче:

$$\frac{du}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}. \tag{3}$$