Лекция 16. Теорема об обратной функции (применения)

Теорема об обратной функции — важный технический результат многомерного анализа.

Теорема 16.1. *Пусть E и F — конечномерные векторные пространства, $U \subset E$ — открытое подмножество и $f : U \to F$ — отображение класса C^r. Предположим, что для некоторой точки $p \in U$ отображение $Df(p) \colon E \to F$ является изоморфизмом. Тогда существует такое содержащее точку p открытое подмножество $U_1 \subset U$, что $f(U_1) \subset F$ — открытое множество и ограничение $f|_{U_1} \colon U_1 \to f(U_1)$ является диффеоморфизмом класса C^r.*

В бесконечномерном случае это утверждение также будет верно (по крайней мере при $r = 1$), если предположить, что E и F — *полные* нормированные пространства.

Замечание 16.2. Поскольку изоморфные векторные пространства имеют одинаковую размерность, в условиях этой теоремы обязательно будет выполнено равенство $\dim E = \dim F$. Далее, пусть при некотором выборе базисов в E и F отображение f записывается формулой

$$(x_1, \ldots, x_n) \mapsto (\varphi_1(x_1, \ldots, x_n), \ldots, \varphi_n(x_1, \ldots, x_n)).$$

Тогда $Df(p)$ является изоморфизмом тогда и только тогда, когда определитель

$$\begin{vmatrix} \dfrac{\partial \varphi_1}{\partial x_1}(p) & \ldots & \dfrac{\partial \varphi_1}{\partial x_n}(p) \\ \ldots & \ldots & \ldots \\ \dfrac{\partial \varphi_n}{\partial x_1}(p) & \ldots & \dfrac{\partial \varphi_n}{\partial x_n}(p) \end{vmatrix}$$

отличен от нуля (см. предложение 14.6 по поводу записи производных в матричном виде). Такого рода определители, составленные из частных производных, называются *якобианами.*

На интуитивном уровне теорема об обратной функции выглядит совершенно естественно: по определению $Df(p)$ является приближением к отображению f в окрестности точки p, так что если отображение $Df(p)$ имеет обратное, то, наверное, и само отображение f в какой-то окрестности p должно иметь обратное. Аккуратное доказательство, однако же, непросто. На данный момент мы это доказательство отложим и посмотрим, как можно теорему об обратной функции применять.

Рассмотрим такую задачу. Пусть переменные $x_1, \ldots, x_n$ и y связаны соотношением $f(x_1, \ldots, x_m, y) = 0$. Спрашивается, можно ли из этого соотношения выразить y как функцию от $x_1, \ldots, x_n$? Чтоб сделать задачу осмысленной, надо ее уточнить. Предположим, что f принадлежит к классу C^r на каком-то открытом множестве $U \subset \mathbb{R}^{m+1}$, и пусть $p = (a_1, \dots, a_m, b) \in U$ — такая точка, что $f(a_1, \dots a_m, b) = 0$. По самому́ определению производной наше соотношение можно (неформально, разумеется) записать в виде

$$0 = f(x_1, \dots, x_m, y) \approx$$
$$\approx f(p) + \frac{\partial f}{\partial x_1}(p)(x_1 - a_1) + \dots \frac{\partial f}{\partial x_m}(p)(x_m - a_m) + \frac{\partial f}{\partial y}(p)(y - b).$$

Если знак бы здесь вместо $\approx$ стояло $=$, вопрос о том, можно ли выразить y через $x_1, \dots, x_m$, решился бы с очевидностью: если коэффициент при $(y - b)$, то есть $\dfrac{\partial f}{\partial y}(p)$, отличен от нуля, то y выражается через $x_1, \dots, x_m$, в противном случае — нет. С помощью теоремы об обратной функции мы сейчас покажем, что *локально* дело обстоит так же, даже если не заменять функцию f на ее линеаризацию; если $\dfrac{\partial f}{\partial y}(p) \neq 0$, то существуют такие открытые множества $W \ni (a_1, \dots, a_m)$ и $V \ni b$, что для всякого $(x_1, \dots, x_m) \in W$ существует и единственно такое $y \in V$, что $f(x_1, \dots, x_m, y) = 0$. Более того, получающаяся функция $(x_1, \dots, x_m) \mapsto y$ также принадлежит к классу C^r.

На самом деле мы будем доказывать более общее утверждение. Именно, в уравнении $f(x_1, \dots, x_m, y) = 0$ мы будем понимать y не обязательно как число, но как точку пространства $\mathbb{R}^n$ для произвольного n. Такое обобщение, однако, влечет и другие изменения в формулировке. В самом деле, если $y = (y_1, \dots, y_n) \in \mathbb{R}^n$, то нам нужно по данному $x = (x_1, \dots, x_m)$ определить n чисел $y_1, \dots, y_n$ из уравнения $f(x_1, \dots, x_m, y_1, \dots, y_n) = 0$. Однако же одного уравнения для этого недостаточно: вообще говоря, чтобы однозначно определить n неизвестных, требуется n уравнений. Поэтому в общем случае нельзя считать, что f — это просто функция (отображение в $\mathbb{R}$): чтобы получилось n уравнений на игреки, надо чтобы $f = (\varphi_1, \dots, \varphi_n)$ было отображением из (открытого подмножества) $\mathbb{R}^{m+n}$ в $\mathbb{R}^n$. Наконец, условие $\dfrac{\partial f}{\partial y}(p) \neq 0$ заменится на условие необращения в нуль некоторого якобиана. Вот окончательная формулировка:

Теорема 16.3 (о неявной функции). *Пусть $U \subset \mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$ — открытое множество и $f \colon U \to \mathbb{R}^n$ — отображение класса C^r, записанное в координатах как*

$$(x_1, \dots, x_m; y_1, \dots, y_n) \mapsto$$
$$(\varphi_1(x_1, \dots, x_m; y_1, \dots, y_n), \dots, \varphi_1(x_1, \dots, x_m; y_1, \dots, y_n)).$$

Пусть $a = (a_1, \ldots, a_m) \in \mathbb{R}^m$, $b = (b_1, \ldots, b_n) \in \mathbb{R}^n$ и $(a; b) \in U$, причем $f(a; b) = (0, \ldots, 0) \in \mathbb{R}^n$. Если определитель

$$\begin{vmatrix} \dfrac{\partial\varphi_1}{\partial y_1}(p) & \ldots & \dfrac{\partial\varphi_1}{\partial y_n}(p) \\ \ldots & \ldots & \ldots \\ \dfrac{\partial\varphi_n}{\partial y_1}(p) & \ldots & \dfrac{\partial\varphi_n}{\partial y_n}(p) \end{vmatrix} \tag{16.1}$$

отличен от нуля, то существуют открытое множество $W \subset \mathbb{R}^m$, содержащее точку $(a_1, \ldots, a_m)$, и открытое множество $V \subset \mathbb{R}^n$, содержащее точку $(b_1, \ldots, b_n)$, такие, что для всякого $x = (x_1, \ldots, x_m) \in W$ существует и единственно $y = (y_1, \ldots, y_n) \in V$, удовлетворяющее условию $f(x; y) = (0, \ldots, 0) \in \mathbb{R}^n$. Наконец, отображение $g\colon W \to V$ ставящее в соответствие точке $x \in W$ то единственное $y \in V$, для которого $f(x; y) = (0, \ldots, 0)$, принадлежит к классу C^r, как и отображение f.

Доказательство. Рассмотрим отображение $\tilde{f}\colon \mathbb{R}^{m+n} \to \mathbb{R}^{m+n}$, заданное следующей формулой, в которой мы в некоторых местах для краткости пишем x вместо $(x_1, \ldots, x_m)$ и y вместо $(y_1, \ldots, y_n)$:

$$\tilde{f}\colon (x; y) \mapsto (x_1, \ldots, x_m, \varphi_1(x; y), \ldots, \varphi_n(x; y)) = (x; f(x, y)).$$

Очевидно, отображение $\tilde{f}$ принадлежит классу C^r на множестве U; далее, матрица его производной $D\tilde{f}(a; b)$ имеет вид

$$\begin{pmatrix} 1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ldots & \vdots \\ 0 & \ldots & 1 & 0 & \ldots & 0 \\ \dfrac{\partial\varphi_1}{\partial x_1}(a; b) & \ldots & \dfrac{\partial\varphi_1}{\partial x_m}(a; b) & \dfrac{\partial\varphi_1}{\partial y_1}(a; b) & \ldots & \dfrac{\partial\varphi_1}{\partial y_n}(a; b) \\ \dfrac{\partial\varphi_2}{\partial x_1}(a; b) & \ldots & \dfrac{\partial\varphi_2}{\partial x_m}(a; b) & \dfrac{\partial\varphi_2}{\partial y_1}(a; b) & \ldots & \dfrac{\partial\varphi_2}{\partial y_n}(a; b) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \dfrac{\partial\varphi_n}{\partial x_1}(a; b) & \ldots & \dfrac{\partial\varphi_n}{\partial x_m}(a; b) & \dfrac{\partial\varphi_n}{\partial y_1}(a; b) & \ldots & \dfrac{\partial\varphi_n}{\partial y_n}(a; b) \end{pmatrix}. \tag{16.2}$$

Определитель этой матрицы равен, очевидно, значению ее $n \times n$-минора, лежащего в нижнем правом углу, а этот последний отличен от нуля по условию. Стало быть, $D\tilde{f}(a; b)$ является изоморфизмом, так что по теореме об обратной функции найдется открытое множество $U \ni (a; b)$, что $\tilde{f}$ индуцирует диффеоморфизм U_1 на $\tilde{f}(U_1)$. Уменьшая, если нужно, множество U, можно считать, что $U = O \times V$, где $O \subset \mathbb{R}^m$ — окрестность точки a и $V \subset \mathbb{R}^n$ — окрестность точки b; очевидно, $f(U) \supset W \times \{(0, \ldots, 0)\}$, где $W \subset \mathbb{R}^m$ — окрестность точки a. Покажем, что открытые множества W и V — именно те, существование которых утверждается теоремой о неявной функции.

Заметим, что $f(x;y) = (0,\ldots,0)$ тогда и только тогда, когда $\tilde f(x;y) = (x;0,\ldots,0)$. Поскольку отображение $\tilde f$ взаимно однозначно на U, отсюда следует, что для всякого $x\in W$ существует не более одного $y\in V$, удовлетворяющего условию $f(x;y) = (0,\ldots,0)$. Поскольку $f(U)$ содержит $W\times\{(0,\ldots,0)\}$, для всякого $x\in W$ существует такой $y\in V$, что $f(x;y) = (0,\ldots,0)$, причем в силу вышесказанного этот y единствен. Стало быть, отображение $g\colon W\to V$, о котором идет речь в формулировке теоремы, построено.

Осталось доказать, что g принадлежит к классу C^r. Пусть $h\colon \tilde f(U)\to U$ — диффеоморфизм класса C^r, обратный к $\tilde f$; обозначим через $i\colon W\to\mathbb{R}^{m+n}$ отображение, заданное формулой $x\mapsto(x;(0,\ldots,0))$ и через $p\colon U\to V$ — проекцию на второй сомножитель. Ясно, что $g = p\circ h\circ i$, и это отображение имеет класс C^r как композиция трех отображений класса C^r (два крайних имеют даже класс C^∞). Все доказано. $\square$

Коль скоро мы установили, что «неявная функция» $g\colon W\to\mathbb{R}^n$ гладка, стоит посчитать ее производную. Это делается так.

Следствие 16.4. *В обозначениях теоремы 16.3 положим*

$$P = \begin{pmatrix} \dfrac{\partial\varphi_1}{\partial y_1}(x;y) & \ldots & \dfrac{\partial\varphi_1}{\partial y_n}(x;y) \\ \ldots & \ldots & \ldots \\ \dfrac{\partial\varphi_n}{\partial y_1}(x;y) & \ldots & \dfrac{\partial\varphi_n}{\partial y_n}(x;y)\end{pmatrix}, \quad Q = \begin{pmatrix} \dfrac{\partial\varphi_1}{\partial x_1}(x;y) & \ldots & \dfrac{\partial\varphi_1}{\partial x_m}(x;y) \\ \ldots & \ldots & \ldots \\ \dfrac{\partial\varphi_n}{\partial x_1}(x;y) & \ldots & \dfrac{\partial\varphi_n}{\partial x_m}(x;y)\end{pmatrix}$$

(P — квадратная матрица порядка n, Q — матрица из m столбцов и n строк).

Тогда производная отображения g в точке $a\in W\subset\mathbb{R}^m$ записывается (в стандартных базисах в $\mathbb{R}^m$ и $\mathbb{R}^n$) матрицей $-P^{-1}Q$.

Доказательство. Мы сохраняем обозначения из доказательства теоремы о неявной функции. Матрицу отображения $D\tilde f(x;y)$, можно записать «в блочном виде» как $\left(\begin{smallmatrix} I_m & 0 \\ Q & P\end{smallmatrix}\right)$ (здесь I_m — единичная матрица размера $m\times m$). Поскольку отображение h обратно к $\tilde f$, отображение $Dh(x;f(x,y))$ является обратным к $D\tilde f(x;y)$; стало быть, матрица этого отображения обратна к (16.2) и тем самым равна $\left(\begin{smallmatrix} I_m & 0 \\ -P^{-1}Q & P^{-1}\end{smallmatrix}\right)$. Поскольку матрицы отображений Di и Dp имеют вид $\left(\begin{smallmatrix} I_m \\ 0\end{smallmatrix}\right)$ и $\left(\begin{smallmatrix} 0 & I_n\end{smallmatrix}\right)$ соответственно матрица производной их композиции равна

$$\begin{pmatrix} 0 & I_n\end{pmatrix}\begin{pmatrix} I_m & 0 \\ -P^{-1}Q & P^{-1}\end{pmatrix}\begin{pmatrix} I_m \\ 0\end{pmatrix} = -P^{-1}Q.$$

$\square$