§ 4. *Основные факты дифференциального исчисления функций многих переменных*

4.5. Теорема об обратной функции.

Теорема об обратной функции — одна из фундаментальных теорем дифференциального исчисления функций многих переменных. Мы докажем ее пользуясь *принципом сжимающих отображений*, т. е. сводя уравнение $f(x) = y$ к уравнению $x = \varphi(x)$, где φ — сжимающее отображение. Теорема о неявной функции будет получена как следствие теоремы об обратной функции.

Метод последовательных приближений. Предположим, что нам нужно решить уравнение $x = \varphi(x)$, где $\varphi : M \to M$ — отображение метрического пространства M. Иногда это можно сделать методом последовательных приближений. Возьмем $x_1 \in M$ произвольно и далее положим $x_{n+1} = \varphi(x_n)$. Наша задача будет решена, если мы докажем, что

– x_n сходится;
– $\lim x_n$ — решение уравнения $x = \varphi(x)$.

Второе легко следует из первого, если $\varphi(x)$ непрерывно — достаточно перейти к пределу в равенстве $x_{n+1} = \varphi(x_n)$. Сходимость же x_n имеет место лишь при дополнительных предположениях.

Определение полного метрического пространства. Метрическое пространство M с метрикой $d(x, y)$ называется *полным*, если любая фунаментальная последовательность в нем сходится, т.е. из того, что

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad | \quad \forall n, m \geq N \quad d(x_n, x_m) < \varepsilon,$$

следует, что x_n имеет предел $x \in M$.

Определение сжимающего отображения. Отображение $\varphi : M \to M$, где M — метрическое пространство, называется *сжимающим*, если

$$\exists q \quad (0 < q < 1) \quad | \quad \forall x, y \in M \quad d(\varphi(x), \varphi(y)) \leq q\, d(x, y).$$

Определение неподвижной точки. Пусть $\varphi : M \to M$ — отображение метрического пространства M. Точка $x \in M$ называется *неподвижной точкой* отображения φ, если $x = \varphi(x)$.

Теорема 1 (о неподвижной точке / принцип сжимающих отображений). *Пусть B — замкнутое подмножество полного метрического пространства M, а $\varphi : B \to M$ — сжимающее отображение такое, что $\varphi(B) \subset B$. Тогда φ имеет единственную неподвижную точку в B.*

Доказательство. Докажем сначала единственность. Пусть x и y — две неподвижные точки. Тогда

$$d(\varphi(x), \varphi(y)) = d(x, y).$$

С другой стороны,

$$d(\varphi(x), \varphi(y)) \leq q\, d(x, y), \quad \text{где } q < 1.$$

Это возможно только в случае $d(x, y) = 0$.

Докажем существование, используя метод последовательных приближений. Возьмем $x_1 \in B$ произвольно и построим последовательность x_n следующим образом: $x_{n+1} = \varphi(x_n)$.

Докажем фундаментальность x_n, которая в силу полноты будет означать сходимость. Имеем

$$d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \leq q\, d(x_n, x_{n-1}) \leq \dots \leq q^{n-1} d(x_2, x_1),$$

где $0 < q < 1$. Далее

$$\begin{aligned} d(x_n, x_{n+k}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{n+k-1}, x_{n+k}) \\ &\leq (q^{n-1} + q^n + \dots + q^{n+k-2}) d(x_2, x_1) \\ &= q^{n-1}(1 + q + \dots + q^{k-1}) d(x_2, x_1) \\ &= \frac{q^{n-1}(1-q^k)}{1-q} d(x_2, x_1) \leq \frac{q^{n-1}}{1-q} d(x_2, x_1) \to 0 \quad \text{при } n \to \infty. \end{aligned}$$

Иначе говоря, x_n фундаментальна, и значит, имеет предел $x \in M$. Кроме того, вся последовательность лежит в B и предел в силу замкнутости тоже.

Осталось понять, что φ, будучи сжимающим отображением, непрерывно, и перейти к пределу в равенстве $x_{n+1} = \varphi(x_n)$. Теорема доказана.

Замечание о картах. Предыдущую теорему можно проиллюстрировать следующим образом. Предположим, что у нас есть две карты одной и той же местности, но разного масштаба. Наложим меньшую карту (большего масштаба) на большую. Утверждается, что можно проткнуть обе карты таким образом, что место прокола на обих картах будет соответствовать одной и той же точке. Действительно, если рассмотреть отображение точек большей карты в соответствующие точки меньшей карты, то получиться сжимающее отображение, удовлетворяющее условиям теоремы.

Теорема 2 (об обратной функции). *Пусть $f : \Omega \to \mathbb{R}^n$, где Ω — область в $\mathbb{R}^n$, все компоненты f принадлежат $C^1(\Omega)$, а точка $x_0 \in \Omega$ такова, что* $\det\left(\frac{\partial f}{\partial x}(x_0)\right) \neq 0$. *Тогда*

1. *существует такая окрестность U точки x_0, в которой f обратима, подробнее, множество $V = f(U)$ — окрестность $y_0 = f(x_0)$ и существует такое отображение $g : V \to U$, что $f(g(y)) = y$ для всех $y \in V$ и $g(f(x)) = x$ для всех $x \in U$;*
2. $\frac{\partial g}{\partial y}(y) = \left(\frac{\partial f}{\partial x}(x)\right)^{-1}$, *где $y = f(x)$, $x \in U$;*
3. *если $f^j \in C^k(\Omega)$, $j = 1, \dots, n$, то $g^i \in C^k(V)$, $i = 1, \dots, n$.*

Доказательство. Идея доказательства. Очевидно, что для линейного отображения вопрос о его обратимости решается просто: если его матрица невырожденная, то отображение обратимо. Нужно просто взять обратную матрицу. Если функция имеет непрерывные частные производные, то локально она представляет собой линейное отображение плюс малое возмущение. Нужно показать, что малое возмущение локально не портит обратимость.

Шаг 1 (переформулировка задачи). Надо решить уравнение $f(x) = y$ для y близких к $y_0 = f(x_0)$. Имеем

$$f(x) = f(x_0) + df(x_0)(x - x_0) + \alpha(x),$$

где $\alpha(x) = o(x - x_0)$, и уравнение принимает вид

$$y_0 + L(x - x_0) + \alpha(x) = y,$$

где $L := df(x_0)$ для краткости. По условию L обратимо, и уравнение можно переписать так:

$$x = L^{-1}(y - y_0) + x_0 - L^{-1}\alpha(x).$$

Обозначая $z := L^{-1}(y - y_0) + x_0$ и $\theta(x) := -L^{-1}\alpha(x)$, еще раз перепишем уравнение:

$$x = \theta(x) + z.$$

Таким образом, решение уравнения $f(x) = y$ сводится к поиску неподвижной точки отображения $\varphi(x) = \theta(x) + z$. Отметим, что z здесь — фиксированный вектор, причем когда y близко к y_0, z близко к x_0, а $\theta(x)$ мало при x близких к x_0.

Выясним сразу более подробно свойства $\theta(x)$. Имеем

$$\alpha(x) = f(x) - f(x_0) - df(x_0)(x - x_0).$$

Отсюда следует, что $\alpha^j(x)$ имеют непрерывные частные производные

$$\frac{\partial \alpha^j}{\partial x^i}(x) = \frac{\partial f^j}{\partial x^i}(x) - \frac{\partial f^j}{\partial x^i}(x_0).$$

Причем $\frac{\partial \alpha^j}{\partial x^i}(x_0) = 0$, $i, j = 1, \dots, n$. Значит, все компоненты $\theta^i(x)$ тоже имеют непрерывные частные производные, и $\frac{\partial \theta^j}{\partial x^i}(x_0) = 0$, $i, j = 1, \dots, n$.

Шаг 2 (оценка приращения $\theta(x)$). Покажем, что точка x_0 имеет окрестность $B(x_0, \delta)$, в которой верна оценка

$$\|\theta(x_1) - \theta(x_2)\|_2 \le \frac{1}{2}\|x_1 - x_2\|_2, \quad x_1, x_2 \in B(x_0, \delta).$$

Для каждой компоненты по теореме о среднем

$$\theta^i(x_1) - \theta^i(x_2) = \sum_{l=1}^{n} \frac{\partial \theta^i(\xi_i)}{\partial x^l}(x_1^l - x_2^l)$$

для некоторого ξ_i из отрезка $[x_1, x_2]$. В силу непрерывности $\frac{\partial \theta^i}{\partial x^l}$ в точке x_0 и того, что $\frac{\partial \theta^j}{\partial x^i}(x_0) = 0$, $i, j = 1, \dots, n$, для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что $\left|\frac{\partial \theta^i}{\partial x^l}(\xi_i)\right| < \varepsilon$, как только $x_1, x_2 \in B(x_0, \delta)$. Значит, для таких x_1 и x_2 верны оценки

$$|\theta^i(x_1) - \theta^i(x_2)| \le \left(\sum_{l=1}^{n} \left|\frac{\partial \theta^i}{\partial x^l}(\xi_i)\right|^2\right)^{1/2} \left(\sum_{l=1}^{n} |x_1^l - x_2^l|^2\right)^{1/2} \le \varepsilon\sqrt{n}\,\|x_1 - x_2\|_2,$$

суммируя которые получаем

$$\|\theta(x_1) - \theta(x_2)\|_2 = \left(\sum_{l=1}^{n} |\theta^i(x_1) - \theta^i(x_2)|^2\right)^{1/2} \le \varepsilon\sqrt{n}\sqrt{n}\,\|x_1 - x_2\|_2.$$

Выбирая $\varepsilon = \frac{1}{2n}$, получаем требуемую оценку в шаре $B(x_0, \delta)$.

Шаг 3 (решение уравнения $x = \varphi(x)$). Докажем, что если $z \in B(x_0, \delta/2)$ (заметьте, что z берется в шаре половинного радиуса!), то $\varphi(x)$ удовлетворяет условиям теоремы о неподвижной точке, т. е.

(1) $\varphi(x) = z + \theta(x)$ переводит шар $B(x_0, \delta)$ в себя;

(2) $\varphi(x)$ — сжимающее отображение на шаре $B(x_0, \delta)$.

Действительно, возьмем $x \in B(x_0, \delta)$ и покажем, что $\varphi(x) \in B(x_0, \delta)$:

$$\|z + \theta(x) - x_0\|_2 \le \|z - x_0\|_2 + \|\theta(x) - \underbrace{\theta(x_0)}_{=0}\| \le \frac{\delta}{2} + \frac{1}{2}\|x - x_0\| \le \delta.$$

Выполнение второго условия очевидно:

$$\|\varphi(x_1) - \varphi(x_2)\|_2 = \|\theta(x_1) - \theta(x_2)\|_2 \le \frac{1}{2}\|x_1 - x_2\|_2.$$

По теореме о неподвижной точке уравнение $x = \varphi(x)$ имеет единственное решение $x \in B(x_0, \delta)$, которое будет и решением уравнения $f(x) = y$, коль скоро z, определяемый по y, лежит в $B(x_0, \delta/2)$.

Шаг 4 (непрерывность обратного отображения и выбор окрестностей U и V). Для решения, полученного на предыдущем шаге, имеет место оценка

$$\|x - x_0\|_2 = \|z + \theta(x) - x_0\|_2 \le \|z - x_0\|_2 + \|\theta(x) - \theta(x_0)\|_2 \le \|z - x_0\|_2 + \frac{1}{2}\|x - x_0\|_2.$$

Перенося второе слагаемое в левую часть, получаем

$$\|x - x_0\|_2 \le 2\|z - x_0\|_2.$$

Вспоминая, как z связано с y: $z = x_0 + L^{-1}(y - y_0)$, оценим

$$\|z - x_0\|_2 \le \|L^{-1}(y - y_0)\|_2 \le c\|y - y_0\|_2$$

с некоторой константой c, зависящей от L. Если выбрать $V = B(y_0, \delta/2c)$, то для $y \in V$ соответствующий z лежит в $B(x_0, \delta/2)$ и имеет место оценка

$$\|x - x_0\|_2 \le 2c\|y - y_0\|_2,$$

которая означает, что обратное отображение $x = g(y)$, заданное на V, непрерывно в точке y_0. Итак, окрестность V выбрана. Положим $U = g(V)$. Будучи прообразом открытого множества V при непрерывном отображении f, U также открыта (осознать!).

Шаг 5 (дифференцируемость обратного отображения). Дифференцируемость g в точке y_0 и формула для матрицы Якоби $\frac{\partial g(y_0)}{\partial y}$ немедленно получаются из теоремы о дифференцировании обратного отображения. Чтобы доказать дифференцируемость g в других точках y, возможно придется сузить окрестности U и V. Именно, уменьшим δ так, чтобы $\det\left(\frac{\partial f(x)}{\partial x}\right) \ne 0$ для всех $x \in U$.

Это возможно, так как определитель матрицы Якоби непрерывно зависит от x, а $U = g(V)$, где g непрерывно в y_0. После сужения окрестностей мы можем повторить рассуждения предыдущих шагов для произвольной точки $x \in U$, доказав непрерывность и следовательно дифференцируемость g в $y = f(x)$.

Шаг 6 (гладкость класса C^k). Так как

$$\frac{\partial g(y)}{\partial y} = \left(\frac{\partial f}{\partial x}(g(y))\right)^{-1},$$

частные производные g выражаются через частные производные f при помощи алгебраических операций:

$$\frac{\partial g^i(y)}{\partial y^j} = G^i_j\Big(\frac{\partial f^1(g(y))}{\partial x^1}, \frac{\partial f^1(g(y))}{\partial x^2}, \dots, \frac{\partial f^n(g(y))}{\partial x^n}\Big),$$

где $G^i_j : \mathbb{R}^{n\times n} \to \mathbb{R} \in C^\infty$, когда ее аргумены составляют невырожденную матрицу. Отсюда по теореме о композиции и классах C^k немедленно получаем

$$\begin{cases} g \in C(V) \\ f \in C^1(\Omega) \\ G^i_j \in C^\infty \end{cases} \quad \longrightarrow \quad \frac{\partial g^i}{\partial y^j} \in C(V) \quad \text{т. е. } g \in C^1(V).$$

Аналогично

$$\begin{cases} g \in C^1(V) \\ f \in C^2(\Omega) \\ G^i_j \in C^\infty \end{cases} \quad \longrightarrow \quad \frac{\partial g^i}{\partial y^j} \in C^1(V) \quad \text{т. е. } g \in C^2(V).$$

Продолжая по индукции получаем $g \in C^k(V)$. Теорема доказана.