Лекция 17. Теорема об обратной функции (обоснования)

Вся эта лекция посвящена доказательству теоремы об обратной функции (теорема 16.1) для случая отображений класса C^1. Начнем с напоминаний про нормы линейных отображений (см. лекцию 13).

Нормой линейного отображения $A\colon E \to F$, где E и F — нормированные векторные пространства, называется число

$$\|A\| = \sup_{v\neq 0} \frac{\|Av\|}{\|v\|} = \sup_{\|v\|=1} \|Av\|.$$

Очевидно, $\|A\| = 0$ тогда и только тогда, когда $A = 0$. Если E и F конечномерны, то $\|A\|$ обязательно конечна (см. предложение 13.9). Функция $A \mapsto \|A\|$ непрерывна на $\operatorname{Hom}(E, F)$ (это следует, например, из очевидного неравенства $\|A+B\| \leqslant \|A\| + \|B\|$, показывающего, что указанная функция является нормой на векторном пространстве $\operatorname{Hom}(E, F)$).

Предложение 17.1. *Пусть $A\colon E \to F$ — изоморфизм конечномерных нормированных линейных пространств. Тогда для любых $v, w \in E$ имеют место неравенства*

$$\|A^{-1}\|^{-1} \cdot \|v - w\| \leqslant \|Av - Aw\| \leqslant \|A\| \cdot \|v - w\|.$$

Доказательство. Правое неравенство тривиально следует из определения нормы оператора, левое неравенство — это правое, примененное к отображению A^{-1} и векторам $A^{-1}v$ и $A^{-1}w$. □

Следующее предложение, интересное и само по себе, является многомерным аналогом теоремы о среднем.

Предложение 17.2. *Пусть E и F — конечномерные нормированные пространства, $U \subset E$ — открытое множество, содержащее отрезок $[a; a+h]$, и $f\colon U \to F$ — отображение класса C^1, заданное (при некотором выборе изоморфизмов E с $\mathbb{R}^m$ и F с $\mathbb{R}^n$) формулой*

$$(x_1, \ldots, x_m) \mapsto (\varphi_1(x_1, \ldots, x_m), \ldots, \varphi_n(x_1, \ldots, x_m)). \tag{17.1}$$

Тогда

$$\|f(a+h) - f(a) - f'(a)(h)\| \leqslant C \cdot \sup_{\substack{0\leqslant t\leqslant 1\\ 1\leqslant i\leqslant n\\ 1\leqslant j\leqslant m}} \left| \frac{\partial \varphi_i}{\partial x_j}(a+th) - \frac{\partial \varphi_i}{\partial x_j}(a) \right| \cdot \|h\|, \tag{17.2}$$

где константа C зависит от m, n и выбора нормы в F, но не от отображения f.

Доказательство. Выберем какой-нибудь номер i, $1 \leqslant i \leqslant n$, и положим $\varphi(t) = \varphi_i(a + th)$. Очевидно,

$$\varphi_i(a+h) = \varphi_i(a) + \int_0^1 \varphi'(t)\,dt, \tag{17.3}$$

$$\varphi'(t) = \sum_{j=1}^{m} \frac{\partial \varphi_i}{\partial x_j}(a+th)\cdot h_j. \tag{17.4}$$

Принимая во внимание очевидное неравенство

$$\left| \int_0^1 \psi(t)\,dt - \psi(0) \right| \leqslant \sup_{0 \leqslant t \leqslant 1} |\psi(t) - \psi(0)|$$

(для непрерывной функции ψ) и подставляя (17.4) в (17.3), имеем

$$\begin{aligned} \left| \varphi_i(a+h) - \varphi_i(a) - \sum_{j=1}^{m} \frac{\partial \varphi_i}{\partial x_j}(a) \right| &\leqslant \sum_{j=1}^{m} \sup_{0 \leqslant t \leqslant 1} \left| \frac{\partial \varphi_i}{\partial x_j}(a+th) - \frac{\partial \varphi_i}{\partial x_j}(a) \right| \\ &\leqslant m \cdot \sup_{\substack{0 \leqslant t \leqslant 1 \\ 1 \leqslant j \leqslant m}} \left| \frac{\partial \varphi_i}{\partial x_j}(a+th) - \frac{\partial \varphi_i}{\partial x_j}(a) \right|. \end{aligned}$$

Беря в получившемся неравенстве максимум по всем i, получаем, что неравенство (17.2) выполнено с константой $C = m$, если в качестве нормы в F выбрать максимум модулей координат в базисе, выбранном при записи формулы (17.1). Если же выбрать в F произвольную норму $\|\cdot\|_?$, то будет выполнено неравенство $\|w\|_? \leqslant K\|w\|$ для некоторой константы K (в силу замечания 13.8), так что неравенство (17.2) выполнено (с константой $C = Km$) и для этой нормы. $\square$

Следствие 17.3. *Пусть U — выпуклое открытое подмножество в конечномерном нормированном пространстве E, и пусть $f\colon U \to E$ — отображение класса C^1, заданное (при некотором выборе координат) формулой (17.1). Если $\left| \frac{\partial f_i}{\partial x_j}(p) - \frac{\partial f_i}{\partial x_j}(q) \right| \leqslant \mu$ для всех $p, q \in U$ и всех i, j, то для $p, q \in U$ выполнено неравенство*

$$\|f(q) - f(p) - f'(p)(q-p)\| \leqslant C\mu \cdot \|q - p\|,$$

где константа C зависит от размерностей пространств и от выбора нормы в F, но не от p, q или отображения f. $\square$

Теперь можно приступить непосредственно к доказательству теоремы об обратной функции. Оно будет разбито на ряд лемм, каждая из которых логически слабее, чем утверждение самой теоремы. Вот первый шаг.

Лемма 17.4. *Пусть E и F — конечномерные пространства, $U \subset E$ — открытое подмножество, $f\colon U \to F$ — отображение класса C^1. Если $f'(a)$ является изоморфизмом для некоторой точки $a \in U$, то существует такая окрестность $V \ni a$, что $f(p) \neq f(q)$, как только $p, q \in V$ и $p \neq q$.*

Доказательство. Идея доказательства проста: по определению производной разность $f(q) - f(p)$ равна $f'(p)(q-p) + R$, где «остаточный член» R пренебрежимо мал по сравнению с первым слагаемым, а первое слагаемое при $p \neq q$ отлично от нуля, если $f'(p)$ является невырожденным оператором. Чтобы оформить эту идею технически, мы воспользуемся следствием 17.3.

Именно, выберем базисы в E и F и запишем отображение f в виде

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

($n = \dim E = \dim F$). Выберем также какие-нибудь нормы в E и F.

Поскольку отображение $p \mapsto f'(p)$ непрерывно по условию, а множество невырожденных $n \times n$-матриц открыто в пространстве всех $n \times n$-матриц, и отображение, сопоставляющее невырожденной матрице обратную к ней, также непрерывно, существуют такие окрестность $V_1 \ni a$ и число M, что для всех $p \in V_1$ отображение $f'(p)$ является изоморфизмом, причем $\|(f'(p))^{-1}\| \leqslant M$. Далее, поскольку все частные производные функций φ_i непрерывны в U, у точки a существует такая окрестность $V_2 \subset V_1$, что для любых $p, q \in V_2$ значения любой частной производной любой функции φ_i в точках p и q отличаются по модулю не более, чем на $1/2MC$, где C — константа из следствия 17.3. Очевидно, можно считать окрестность V_2 выпуклой (например, открытым шаром относительно некоторой нормы). Теперь применим следствие 17.3: если $p, q \in V_2$, то

$$f(q) - f(p) = f'(p)(q-p) + R, \tag{17.5}$$

где $\|R\| \leqslant C \cdot (1/2MC)\|q-p\| = (1/2M)\|q-p\|$. С другой стороны, предложение 17.1 показывает, что

$$\|f'(p)(q-p)\| \geqslant \|f'(p)^{-1}\|^{-1}\|q-p\| \geqslant (1/M)\|q-p\|.$$

Следовательно, норма второго слагаемого в правой части (17.5) строго меньше нормы первого слагаемого, так что сумма в правой части (17.5) отлична от нуля. Значит, $f(p) \neq f(q)$, и в качестве V можно взять V_2. □

Итак, мы уже знаем, что отображение класса C^1 с невырожденной производной будет локально взаимно однозначно. Следующий (и самый трудный) шаг — установить, что такое отображение будет, как говорят, *открыто* — переводить открытые множества в открытые.

Лемма 17.5. *Пусть E и F — конечномерные пространства, $U \subset E$ — открытое подмножество, $f\colon U \to F$ — отображение класса C^1. Если $f'(a)$ является изоморфизмом для всякой точки $a \in U$, то множество $f(U) \subset F$ открыто.*

Доказательство. По-прежнему положим $n = \dim E = \dim F$. Пусть $a \in U$. Нам надо доказать, что существует окрестность $W \ni f(a)$, обладающая следующим свойством: для всякого $y \in W$ найдется такой $x \in U$, что $f(x) = y$. Мы построим требуемый x с помощью «метода Ньютона» — одного из вариантов метода последовательных приближений. Идея построения такова. Для всех x, близких к a, выполнено приближенное равенство $f(x) \approx f(a) + f'(a)(x-a)$. Если бы это приближенное равенство было точным (иными словами, если бы отображение f было аффинным), то решение уравнения $f(x) = y$ было бы очевидно: поскольку линейное отображение $f'(a)$ обратимо, можно (и нужно) было бы взять $x = x_1 = a + (f'(a))^{-1}(y - f(a))$. Поскольку, однако же, отображение f аффинным быть не обязано, точка x_1 является всего лишь приближенным решением уравнения $f(x) = y$; давайте теперь запишем приближенное равенство $f(x) \approx f(x_1) + f'(x_1)(x - x_1)$ и, заменяя $\approx$ на $=$, найдем решение уравнения $f(x_1) + f'(x_1)(x - x_1) = y$; это решение есть $x_2 = x_1 + (f'(x_1))^{-1}(y - f(x_1))$. Продолжая аналогичным образом, получим последовательность $\{x_n\}$ точек-последовательных приближений, заданную с помощью следующего рекурсивного определения:

$$x_0 = a; \qquad x_n = x_{n-1} + (f'(x_{n-1}))^{-1}(y - f(x_{n-1})). \tag{17.6}$$

Хочется надеяться, что эта последовательность будет сходиться к точке $x \in U$, и что $f(x) = y$.

Если при данном a выбирать точку y как попало, то эта идея не срабатывает: и сходимость последовательности не гарантируется, и вообще, какая-то из точек x_n может выйти за пределы области определения отображения f. Поэтому при выборе y придется проявить известную аккуратность.

Для начала заметим, что, ввиду непрерывности производной от f, существует такая выпуклая окрестность $V \ni a$ и константа $A > 0$, что для всех $p \in V$ имеем $\|(f'(p))^{-1}\| \leqslant A$. Положим

$$\mu = \sup_{\substack{p,q \in V\\ 1 \leqslant i,j \leqslant n}} \left| \frac{\partial \varphi_i}{\partial x_j}(p) - \frac{\partial \varphi_i}{\partial x_j}(q) \right|,$$

где $\varphi_1, \dots, \varphi_n$ — компоненты отображения f относительно некоторых базисов в E и F. Уменьшив при необходимости окрестность V, можно считать, что число μ конечно.

Пусть, далее, C — константа из следствия 17.3. Поскольку, ввиду определения (17.6), имеем для любого $n \in \mathbb{N}$

$$x_n - x_{n-1} = f'(x_{n-1})(y - f(x_{n-1})),$$

получаем следующие соотношения:

$$\begin{gathered}(x_{n-1} \in V) \Rightarrow (\|x_n - x_{n-1}\| \leqslant A\|y - f(x_{n-1})\|) \\ (x_{n-1}, x_n \in V) \Rightarrow \|y - f(x_n)\| \leqslant C\mu\|x_n - x_{n-1}\|. \qquad (17.7)\end{gathered}$$

Стало быть,

$$\|x_{n+1} - x_n\| \leqslant A\|y - f(x_{n-1})\| \leqslant AC\mu\|x_n - x_{n-1}\|. \qquad (17.8)$$

Уменьшим теперь окрестность $V \ni a$ таким образом, чтобы было выполнено неравенство $\mu < \dfrac{1}{3AC}$ (заметим, что старые значения A и C будут годиться и для этой уменьшенной окрестности). Тогда неравенство (17.8) влечет неравенство

$$\|x_{n+1} - x_n\| \leqslant \frac{1}{3}\|x_n - x_{n-1}\| \qquad (17.9)$$

при условии, что $x_{n-1}, x_n \in V$.

Уменьшая при необходимости окрестность V, можно считать, что она имеет вид

$$\{\, x \in E \mid \|x - a\| < \varepsilon \,\}.$$

Положим $W = \{\, y \in F \mid \|(f'(a)^{-1})(y - f(a))\| < \varepsilon/2 \,\}$. Покажем, что при всех y, лежащих в открытом множестве W, последовательность $\{x_n\}$, заданная соотношениями (17.6), определена при всех $n \in \mathbb{N}$, и при этом для всех $n \in \mathbb{N}$ имеем $x_n \in V$.

В самом деле, будем рассуждать по индукции. То, что $x_0 = a$ лежит в V, очевидно; из нашего выбора окрестности W следует что $x_1 = a + (f'(a)^{-1})(y - a)$ лежит в V и, более того, $\|x_1 - a\| < \varepsilon/2$. Пусть уже доказано, что $x_1, \ldots, x_n \in V$. Из неравенства (17.9) следует, что

$$\begin{aligned}\|x_{n+1} - a\| \leqslant \|x_1 - x_0\| + \|x_2 - x_1\| + \cdots + \|x_{n+1} - x_n\| \leqslant \\ \leqslant \|x_1 - x_0\|\Bigl(1 + \frac{1}{3} + \frac{1}{3^2} + \cdots\Bigr) < \frac{\varepsilon}{2} \cdot \frac{3}{2} < \frac{3\varepsilon}{4},\end{aligned}$$

откуда $x_{n+1} \in V$, и наше утверждение доказано.

Поскольку $\|x_{n+1} - x_n\| \leqslant \dfrac{1}{3} \cdot \|x_n - x_{n-1}\|$, последовательность $\{x_n\}$ сходится, так как сходится по норме ряд $x_0 + \sum_{i \geqslant 1}(x_i - x_{i-1})$. Пусть $\lim x_i = x \in E$; поскольку $\|x_n - a\| \leqslant \dfrac{3\varepsilon}{4}$ для всех n, имеем $x \in V$. Поскольку $\lim_{n\to\infty} \|x_n - x_{n-1}\| = 0$, неравенство (17.7) показывает, что $\lim_{n\to\infty} f(x_n) = y$; так как дифференцируемое отображение f непрерывно, $\lim f(x_n) = f(x)$ и $f(x) = y$. Итак, для всякого $y \in W$ существует такой $x \in V$, что $f(x) = y$, что и требовалось доказать. □

Завершим доказательство теоремы. Пусть $f\colon U \to F$ — отображение класса C^1, и пусть для некоторой точки $a \in U$ отображение

$f'(a)\colon E \to F$ является изоморфизмом (обозначения U, E и F имеют прежний смысл). Уменьшая при необходимости открытое множество U, можно считать, что $f'(p)$ — изоморфизм для всех $p \in U$ (ввиду непрерывности f') и что f взаимно однозначно (лемма 17.4). Лемма 17.5 гарантирует, что образ открытого множества при отображении f открыт. Следовательно, $f(U)$ открыто в F и ограничение $f|_U\colon U \to f(U)$ — гомеоморфизм. Пусть $g\colon f(U) \to U$ — обратное (обязательно непрерывное) отображение. Покажем, что g принадлежит классу C^1. Это следует из такой леммы:

Лемма 17.6. *Пусть $f\colon U \to V$ — отображение класса C^1, являющееся гомеоморфизмом и обладающее тем свойством, что $f'(a)$ — изоморфизм для всех $a \in U$ (U и V — открытые подмножества в $\mathbb{R}^n$). Тогда обратное отображение $g\colon V \to U$ также принадлежит классу C^1.*

Доказательство. Пусть $a \in U$ и $b = f(a)$. Покажем сначала, что g дифференцируемо в точке b и что его производная равна $(f'(a))^{-1}$ (ничему другому она равняться и не могла бы!). Итак, нам нужно доказать, что

$$\lim_{y\to b} \frac{\|g(y) - g(b) - ((f'(a))^{-1})(y-b)\|}{\|y-b\|} = 0.$$

Поскольку f и g — взаимно обратные гомеоморфизмы, это равносильно равенству

$$\lim_{x\to a} \frac{\|x - a - ((f'(a))^{-1})(f(x) - f(a))\|}{\|f(x) - f(a)\|} = 0. \tag{17.10}$$

Поскольку f дифференцируемо в точке a, выполнено равенство $f(x) - f(a) = f'(a)(x-a) + \psi(x)$, где $\|\psi(x)\| = o(\|x-a\|)$. Подставляя это в (17.10), получаем, что надо доказать соотношение

$$\lim_{x\to a} \frac{\|f'(a)^{-1}(\psi(x))\|}{f'(a)(x-a) + \psi(x)} = 0. \tag{17.11}$$

Числитель здесь, очевидно, есть $o(\|x-a\|)$, а знаменатель заключен между $c_1\|x-a\|$ и $c_2\|x-a\|$, где c_1 и c_2 — положительные константы (см. предложение 17.1), так что равенство (17.11), доказано, а с ним и дифференцируемость g.

Поскольку, ввиду вышесказанного, $g'(b) = (f'(g(b)))^{-1}$, непрерывная дифференцируемость g следует из непрерывности отображения g и операции взятия обратной матрицы. □

Итак, теорема об обратной функции для отображений класса C^1 доказана. Случай отображений класса C^r, где $r > 1$, вывести из доказанного нами уже несложно, но на данной лекции мы это сделать не успеем.