лектор **А.Н.Подкорытов**

ОБРАТНОЕ И НЕЯВНОЕ ОТОБРАЖЕНИЯ

Терминология и обозначения

Как обычно, $\mathbb{R}^m$ — линейное пространство векторов $x = (x_1, \ldots, x_m)$, координаты $x_1, \ldots, x_m$ которых — произвольные вещественные числа; $\mathbb{O}_m$ — нулевой вектор, т.е. вектор с нулевыми координатами. Векторы, у которых одна координата равна 1, а остальные — нулю, образуют базис пространства $\mathbb{R}^m$. Его будем называть каноническим. Скалярное произведение $\langle x, y\rangle$ векторов $x = (x_1, \ldots, x_m)$ и $y = (y_1, \ldots, y_m)$ определяется равенством $\langle x, y\rangle = x_1y_1 + \ldots + x_my_m$. Оно порождает в $\mathbb{R}^m$ евклидову норму $\|x\|_m = \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + \ldots + x_m^2}$. Множество $B_r(x_0) = \{x \in \mathbb{R}^m \mid \|x - x_0\|_m < r\}$ называется открытым шаром радиуса r, $r > 0$, с центром в точке x_0, $x_0 \in \mathbb{R}^m$. Символом $\overline{B}_r(x_0)$ будем обозначать замкнутый шар: $\overline{B}_r(x_0) = \{x \in \mathbb{R}^m \mid \|x - x_0\|_m \leqslant r\}$. Точка x_0 называется внутренней точкой множества E, $E \subset \mathbb{R}^m$, если $B_r(x_0) \subset E$ для некоторого r, $r > 0$. Внутренностью множества E — обозначение $\operatorname{Int}(E)$ — называется множество всех его внутренних точек. Множество E открыто, если все его точки внутренние, т.е. $E = \operatorname{Int}(E)$.

Для натуральных m и M символом $\mathscr{L}_{mM}$ обозначим пространство всех линейных отображений из $\mathbb{R}^m$ в $\mathbb{R}^M$. Как известно из курса высшей алгебры, всякое отображение L из $\mathscr{L}_{m,M}$ единственным образом представимо в виде произведения некоторой матрицы (имеющей m столбцов и M строк) на вектор x. Точнее, существует такая единственная матрица

$$\begin{Vmatrix} l_{11} & \ldots & l_{1m} \\ \ldots & \ldots & \ldots \\ l_{M1} & \ldots & l_{Mm} \end{Vmatrix},$$

что для любого вектора x из $\mathbb{R}^m$ координаты вектора $y = L(x)$ связаны с координатами x (в канонических базисах пространств $\mathbb{R}^m$ и $\mathbb{R}^M$) равенством

$$\begin{pmatrix} y_1 \\ \ldots \\ y_M \end{pmatrix} = \begin{Vmatrix} l_{11} & \ldots & l_{1m} \\ \ldots & \ldots & \ldots \\ l_{M1} & \ldots & l_{Mm} \end{Vmatrix} \begin{pmatrix} x_1 \\ \ldots \\ x_m \end{pmatrix}.$$

Эту матрицу обозначим $[L]$. Будем также использовать запись $[L] = (l_{kj})_{\substack{1\leqslant j\leqslant m\\ 1\leqslant k\leqslant M}}$. Равенство $y = L(x)$ означает, что вектор – столбец y^T получается умножением справа матрицы $[L]$ на вектор – столбец x^T, т.е. $y^T = [L]\, x^T$.

Ясно, что $\mathscr{L}_{mM}$ — линейное пространство размерности mM. Равенством

$$\|L\|_{\mathscr{L}} = \sup_{\|h\|_m \leqslant 1} \|L(h)\|_M$$

в нём вводится норма. Напомним, что

$$\|L(h)\|_M \leqslant \|L\|_{\mathscr{L}}\, \|h\|_m \quad \text{для любого} \quad h \quad \text{из} \quad \mathbb{R}^m;$$

$$\|\widetilde{L} \circ L\|_{\mathscr{L}} \leqslant \|\widetilde{L}\|_{\mathscr{L}}\, \|L\|_{\mathscr{L}} \quad \text{для} \quad L \in \mathscr{L}_{mM},\ \widetilde{L} \in \mathscr{L}_{MN};$$

и

$$\|L\|_{\mathscr{L}} \leqslant \sqrt{\sum_{k=1}^{M}\sum_{j=1}^{m} l_{kj}^2}, \quad \text{если} \quad [L] = (l_{kj})_{\substack{1\leqslant j\leqslant m\\ 1\leqslant k\leqslant M}}.$$

Из определения $\|L\|_{\mathscr{L}}$ сразу следует, что $\|L\|_{\mathscr{L}} \geqslant \|L(x)\|_M$, если $\|x\|_m = 1$. Поэтому

$$\|L\|_{\mathscr{L}} \geqslant \|L(e_j)\|_M = \|\,[L]\,e_j^T\|_M = \sqrt{\sum_{k=1}^{M} l_{kj}^2}$$

для любого вектора e_j канонического базиса в $\mathbb{R}^m$. Это позволяет дополнить оценку сверху нормы $\|L\|_{\mathscr{L}}$ с помощью элементов матрицы $[L]$ оценкой снизу:

$$\|L\|_{\mathscr{L}} \geqslant |l_{kj}| \qquad \text{для любых} \;\; j = 1, \dots, m \quad \text{и} \quad k = 1, \dots, M.$$

Отметим ещё, что в силу первого свойства линейное отображение удовлетворяет условию Липшица:

$$\|L(\widetilde{h}) - L(h)\|_M = \|L(\widetilde{h} - h)\|_M \leqslant \|L\|_{\mathscr{L}}\, \|\widetilde{h} - h\|_m,$$

и поэтому оно равномерно непрерывно в $\mathbb{R}^m$.

Нам потребуется несложное утверждение из линейной алгебры.

Лемма (о взаимно однозначных линейных отображениях). *Пусть $L \in \mathscr{L}_{mM}$. Следующие четыре утверждения равносильны*

I. L взаимно однозначно (т.е. $L(h) \neq L(\widetilde{h})$, если $h \neq \widetilde{h}$);

II. $L(h) \neq \mathbb{O}_M$, если $h \neq \mathbb{O}_m$ (иначе говоря, $\operatorname{Ker} L = L^{-1}(\{\mathbb{O}_M\}) = \{\mathbb{O}_m\}$);

III. существует такое число $\mu > 0$, что $\|L(h)\|_M \geqslant \mu\,\|h\|_m$ для любого вектора h из $\mathbb{R}^m$;

IV. если векторы $v_1, \dots, v_n$ линейно независимы в $\mathbb{R}^m$, то их образы $L(v_1), \dots, L(v_n)$ линейно независимы в $\mathbb{R}^M$.

Доказательство. I $\Rightarrow$ II. В силу линейности отображения $L(\mathbb{O}_m) = \mathbb{O}_M$. Поэтому, взяв $\widetilde{h} = \mathbb{O}_m$ в утверждении I, мы получим $L(h) \neq L(\mathbb{O}_m) = \mathbb{O}_M$ при $h \neq \mathbb{O}_m$.

II $\Rightarrow$ III. Рассмотрим в $\mathbb{R}^m$ функцию $h \mapsto \|L(h)\|_M$. Это квадратный корень из неотрицательной квадратичной формы $\|L(h)\|_M^2$. Поэтому эта функция всюду непрерывна. По теореме Вейерштрасса она достигает наименьшего значения на единичной сфере S^{m-1}: существует такой вектор h_*, $\|h_*\|_m = 1$, что $\|L(h)\|_M \geqslant \|L(h_*)\|_M$ для любого вектора h из S^{m-1}. Так как $\|h_*\|_m = 1$, то $h_* \neq \mathbb{O}_m$. Поэтому в силу условия II число $\mu = \|L(h_*)\|_M$ положительно. Проверим, что оно удовлетворяет неравенству III. Можно считать при этом, что $h \neq \mathbb{O}_m$ (для нулевого вектора обе части этого неравенства обращаются в нуль). Тогда вектор $\overline{h} = \frac{h}{\|h\|_m}$ принадлежит S^{m-1}. Следовательно, $\|L(\overline{h})\|_M \geqslant \|L(h_*)\|_M = \mu$. В то же время

$$\|L(\overline{h})\|_M = \left\| L\left(\frac{h}{\|h\|_m}\right)\right\|_M = \left\|\frac{1}{\|h\|_m}\, L(h)\right\|_M = \frac{1}{\|h\|_m}\, \|L(h)\|_M.$$

Таким образом, $\|L(h)\|_M = \|h\|_m\, \|L(\overline{h})\|_M \geqslant \mu\, \|h\|_m$.

III $\Rightarrow$ IV. Так как отображение L линейно, то для любых коэффициентов $c_1, \dots, c_n$ справедливо равенство $c_1 L(v_1) + \ldots + c_n L(v_n) = L(c_1 v_1 + \ldots + c_n v_n)$. С помощью утверждения III мы получаем отсюда, что

$$\|c_1 L(v_1) + \ldots + c_n L(v_n)\|_M \geqslant \mu \|c_1 v_1 + \ldots + c_n v_n\|_m.$$

Поэтому, если $c_1 L(v_1) + \ldots + c_n L(v_n) = \mathbb{O}_M$, то $c_1 v_1 + \ldots + c_n v_n = \mathbb{O}_m$, а это возможно лишь при $c_1 = \ldots = c_n = 0$, поскольку векторы $v_1, \dots, v_n$ линейно независимы.

IV $\Rightarrow$ I. Надо доказать, что $L(\widetilde{h}) \neq L(h)$, если $\widetilde{h} \neq h$. Возьмём произвольный базис $v_1, \dots, v_m$ (например, канонический) в $\mathbb{R}^m$ и разложим по нему ненулевой вектор $\widetilde{h} - h$:

$$\widetilde{h} - h = c_1 v_1 + \ldots + c_m v_m.$$

Векторы $v_1, \ldots, v_m$ линейно независимы, поскольку они образуют базис. По условию IV таковы и их образы $L(v_1), \ldots, L(v_m)$. Учитывая, что среди коэффициентов $c_1, \ldots, c_m$ имеется хотя бы один ненулевой (поскольку $\widetilde{h} - h \neq \mathbb{O}_m$), получаем

$$L(\widetilde{h}) - L(h) = L(\widetilde{h} - h) = L(c_1 v_1 + \ldots + c_m v_m) = c_1 L(v_1) + \ldots + c_m L(v_m) \neq \mathbb{O}_M,$$

что и требовалось. Лемма доказана.

В утверждении IV число n не больше $\min\{m, M\}$, так как число линейно независимых векторов не превосходит размерности пространства. В то же время, векторы произвольного базиса в $\mathbb{R}^m$ (их m штук) линейно независимы. Поэтому $m \leqslant \min\{m, M\} \leqslant M$. Следовательно, условия I – IV леммы выполняются лишь при $m \leqslant M$. Отсюда вытекает, что отображение L из $\mathscr{L}_{mM}$ может быть обратимым только в случае $m = M$ (неравенство $m \leqslant M$ мы только что установили, а для получения неравенства $M \leqslant m$ надо вместо отображения L рассмотреть отображение L^{-1}, действующее из $\mathbb{R}^M$ в $\mathbb{R}^m$). Таким образом, мы приходим к факту, известному из курса линейной алгебры: для обратимости линейного отображения из $\mathbb{R}^m$ в $\mathbb{R}^M$ необходимо, чтобы размерности этих пространств совпадали.

Символом $\mathrm{Isom}\,(\mathbb{R}^m)$ обозначим подмножество пространства $\mathscr{L}_{mm}$, образованное обратимыми линейными отображениями (изоморфизмами). Ясно, что всякое такое отображение удовлетворяет условию I леммы, а следовательно, и остальным её условиям (с $M = m$). Легко видеть, что верно и обратное: всякое отображение из $\mathscr{L}_{mm}$, удовлетворяющее условиям I – IV леммы, является изоморфизмом. Действительно, условие I обеспечивает инъективность, а из условия IV вытекает сюръективность (множество $L(\mathbb{R}^m)$ совпадает с $\mathbb{R}^m$, так это линейное подпространство в $\mathbb{R}^m$, содержащее m линейно независимых векторов — образов базисных в $\mathbb{R}^m$ векторов). Следовательно, при $m = M$ лемма даёт описание изоморфизмов.

Из курса высшей алгебры известно, что линейная независимость образов $L(e_1), \ldots, L(e_m)$ векторов канонического базиса (т.е. условие IV) равносильна тому, что квадратная матрица $[L]$ неособенная, т.е. её определитель ненулевой: $\det[L] \neq 0$. Это неравенство даёт ещё один способ описания отображений из $\mathrm{Isom}\,(\mathbb{R}^m)$.

Поскольку определитель матрицы непрерывно зависит от её элементов, из теоремы о стабилизации знака следует, что любое достаточно малое изменение элементов неособенной матрицы приводит к новой неособенной матрице. Поэтому, если отображение L, $L \in \mathscr{L}_{mm}$, обратимо, то обратимы и все достаточно близкие к L линейные отображения, т.е. $\mathrm{Isom}\,(\mathbb{R}^m)$ — открытое подмножество пространства $\mathscr{L}_{mm}$. Отметим частный случай леммы:

для любого L из $\mathrm{Isom}\,(\mathbb{R}^m)$ существует такое положительное число $\mu = \mu(L)$, что

$$\|L(h)\|_m \geqslant \mu\,\|h\|_m \quad \textit{для всех} \;\; h \quad \textit{из} \;\; \mathbb{R}^m.$$

Таким образом, оценку сверху $\|L(h)\|_m \leqslant \|L\|_{\mathscr{L}}\,\|h\|_m$, справедливую для *любого* линейного отображения из $\mathbb{R}^m$ в $\mathbb{R}^m$, для *изоморфизма* можно дополнить аналогичной оценкой снизу. Из леммы следует, что это свойство не только необходимо, но и достаточно для того, чтобы отображение из $\mathscr{L}_{mm}$ было изоморфизмом.

Каждому отображению F, действующему из какого-то множества E в пространство $\mathbb{R}^M$, соответствует единственный набор функций $F_1, \ldots, F_M$ (т.е. отображений из E в $\mathbb{R}$), обладающий свойством: $F(x) = (F_1(x), \ldots, F_M(x))$ для всех x из E. Эти функции называются координатными функциями отображения F (в каноническом базисе пространства $\mathbb{R}^M$).

Отображение $F : E \to \mathbb{R}^M$, $E \subset \mathbb{R}^m$, дифференцируемо в точке x_0, $x_0 \in \mathrm{Int}\,(E)$, если существуют такие отображения L, $L \in \mathscr{L}_{mM}$, и $\alpha : E \to \mathbb{R}^M$, что $\lim\limits_{x_0} \alpha = \mathbb{O}_M$ и

$$F(x) = F(x_0) + L(x - x_0) + \|x - x_0\|_m\,\alpha(x) \quad \text{для всех} \;\; x \quad \text{из} \;\; E.$$

Отображение L называется *дифференциалом* отображения F в точке x_0 и обозначается $d_{x_0}F$ (или $dF(x_0)$). Соответствующая ему матрица $[d_{x_0}F]$ называется матрицей Якоби отображения F в точке x_0 и обозначается $F'(x_0)$. Она образована частными производными координатных функций отображения F, вычисленными в точке x_0:

$$F'(x_0) = \begin{Vmatrix} \frac{\partial F_1}{\partial x_1}(x_0) & \dots & \frac{\partial F_1}{\partial x_m}(x_0) \\ \dots & \dots & \dots \\ \frac{\partial F_M}{\partial x_1}(x_0) & \dots & \frac{\partial F_M}{\partial x_m}(x_0) \end{Vmatrix}.$$

Если рассматривается функция нескольких переменных, т.е. $M = 1$, то её матрица Якоби состоит из единственной строки, которая называется *градиентом* функции $F : E \to \mathbb{R}$, $E \subset \mathbb{R}^m$, в точке x_0, $x_0 \in \operatorname{Int}(E)$. Его обычно обозначают $\operatorname{grad} F(x_0)$. Эта строка образована частными производными функции F:

$$\operatorname{grad} F(x_0) = \left(\frac{\partial F}{\partial x_1}(x_0), \dots , \frac{\partial F}{\partial x_m}(x_0) \right).$$

Из определения дифференциала сразу следует равенство $d_{x_0}L = L$ в любой точке $x_0 \in \mathbb{R}^m$, если $L \in \mathscr{L}_{mM}$.

Отметим ещё, что значение $\alpha(x_0)$ не играет роли (оно умножается на нуль). Поэтому не умаляя общности, можно считать, что $\alpha(x_0) = \mathbb{O}_M$, т.е. α непрерывно в точке x_0. Всюду в дальнейшем мы будем придерживаться этого соглашения. Непрерывность α влечёт непрерывность F в точке x_0. Разумеется, непрерывность отображения лишь необходима, но не достаточна для его дифференцируемости.

§1. **Неявная функция.**

Всюду далее мы отождествляем произведение пространства $\mathbb{R}^m$ и прямой $\mathbb{R}$ с пространством $\mathbb{R}^{m+1}$, так что любому вектору x из $\mathbb{R}^m$ и числу y соответствует вектор (x, y) из $\mathbb{R}^{m+1}$ с координатами $x_1, \dots , x_m, y$.

П о с т а н о в к а з а д а ч и. Ещё в школе много раз приходилось решать уравнения и системы уравнений с параметрами. Требовалось выписать решения при каждом допустимом наборе параметров. Сейчас мы займёмся близким вопросом. С одной стороны, он более общий — мы не будем конкретизировать уравнения. Важно, что они задаются непрерывно дифференцируемыми функциями. С другой стороны, наша цель здесь значительно скромнее. Мы не ставим задачу указать все решения. Это осмысленно лишь для конкретных уравнений. Наша цель другая — выяснить, в каком случае можно гарантировать существование и "хорошую" (гладкую) зависимость решения от параметров. В этом параграфе мы рассмотрим простой случай — одно уравнение с одним неизвестным (число параметров произвольно). Здесь рассуждения достаточно наглядны, так как можно воспользоваться монотонностью.

Итак, рассмотрим одно уравнение относительно неизвестной y, $y \in \mathbb{R}$,

$$F(x_1, \dots , x_m, y) = 0$$

с вещественными параметрами $x_1, \dots , x_m$. Верно ли, что найдётся его решение, если при $x_1 = x_1^\circ, \dots , x_m = x_m^\circ$ имеется решение y_0 (т.е. $F(x_1^\circ, \dots , x_m^\circ, y_0) = 0$), а все разности $x_1 - x_1^\circ, \dots , x_m - x_m^\circ$ достаточно малы? Как скажется на решении уравнения такое изменение параметров?

Определение. Пусть функция F определена на непустом подмножестве E произведения $X \times \mathbb{R}$. Говорят, что $\varphi : X \to \mathbb{R}$ — *неявная функция*, задаваемая уравнением $F(x, y) = 0$, если $(x, \varphi(x)) \in E$ и $F(x, \varphi(x)) = 0$ для любого $x \in X$.

Пример 1. Рассмотрим уравнение $x^2 + y^2 - 1 = 0$ относительно неизвестной y (x — параметр). Очевидно, при $|x| > 1$ нет решений. В то же время при $|x| < 1$ уравнение задаёт две непрерывные функции $\varphi_\pm(x) = \pm\sqrt{1-x^2}$. Если интересоваться непрерывной неявной функцией, принимающей в точке x_0 заданное значение y_0 ($x_0^2 + y_0^2 = 1$), то такая функция существует (это $\varphi(x) = \operatorname{sign}(y_0)\sqrt{1-x^2}$) и она единственна (разрывных решений бесконечно много).

Сложнее ситуация около точек $x_0 = \pm 1$. В этом случае, разумеется, $y_0 = 0$. Однако, ни в какой окрестности точек ± 1 уравнение не имеет решения — оно существует лишь в полуокрестности (слева от точки 1 и справа от -1). При этом непрерывных решений два, но они не дифференцируемы в точке x_0. В определённом смысле это вырожденный случай. Одна из наших целей — научиться определять (не решая уравнения!), в каких случаях могут возникать подобные ситуации, а в каких случаях они исключены.

Отметим ещё, что в этом примере требуются минимальные изменения для рассмотрения более общего уравнения $x_1^2 + \ldots + x_m^2 + y^2 = 1$, задающего неявную функцию от m переменных.

Пример 2. Более сложное уравнение $F(x,y) = (x^2+y^2)^2 - x^2 + y^2 = 0$ описывает лемнискату Бернулли. Её легко нарисовать, используя полярные координаты: $x = r\cos\theta$, $y = r\sin\theta$. Тогда $r = \sqrt{\cos 2\theta}$. Получаем “восьмёрку на боку”. Точки $x_0 = \pm 1$ играют ту же роль, что и в предыдущем примере. Если $0 < |x_0| < 1$, то и здесь ситуация аналогична уже рассмотренной — каждому числу y_0, удовлетворяющему уравнению $\left(x_0^2 + y_0^2\right)^2 = x_0^2 - y_0^2$ соответствует единственная непрерывная неявная функция, принимающая в точке x_0 значение y_0 (нетрудно понять, что она не только непрерывна, но бесконечно дифференцируема). Но в этом примере появляется новая особенность — точка $x_0 = 0$ (тогда и $y_0 = 0$). Из рисунка хорошо видно, что через точку $(0,0)$ проходят графики четырёх непрерывных функций $y = \varphi(x)$, удовлетворяющих уравнению $F(x,\varphi(x)) = 0$. Две из них (меняющие знак в нуле) дифференцируемы, а другие две (сохраняющие знак) не дифференцируемы в нуле.

Лемма 1 (существование неявной функции). *Пусть функция F задана на множестве $E = X \times [y_*, y^*]$, где X — непустое множество и $-\infty < y_* < y^* < +\infty$. Допустим, что для любого x из X*

1) функция $y \mapsto F(x,y)$ непрерывна (т.е. F непрерывна по второй переменной);

2) значения $F(x,y_*)$ и $F(x,y^*)$ имеют разные знаки (т.е. $F(x,y_*)F(x,y^*) < 0$).

Тогда уравнение $F(x,y) = 0$ задаёт некоторую неявную функцию φ: для каждого x из X, существует такое число $\varphi(x)$ из интервала (y_*, y^*), что $F(x,\varphi(x)) = 0$.

Если F удовлетворяет дополнительному условию

3) при каждом x, $x \in X$, функция $y \mapsto F(x,y)$ строго монотонна,

то уравнение $F(x,y) = 0$ задаёт на множестве X единственную неявную функцию $\varphi : X \to (y_*, y^*)$ (т.е. $F(x,\varphi(x)) = 0$ при каждом x и других решений уравнение $F(x,y) = 0$ не имеет).

Для доказательства достаточно сослаться на теорему Больцано – Коши о промежуточном значении непрерывной на промежутке функции.

Отметим, что в лемме природа множества X безразлична. Естественно, что при столь общих предположениях ничего не говорится о свойствах неявной функции. Немного увеличив предположения, можно добиться её непрерывности.

Упражнение. Пусть функция F непрерывна на прямоугольнике $E = (x_*, x^*) \times [y_*, y^*]$, где $-\infty \leqslant x_* < x^* \leqslant +\infty$, $-\infty < y_* < y^* < +\infty$, и удовлетворяет условиям 2) и 3) леммы. Тогда неявная функция φ, задаваемая уравнением $F(x,y) = 0$, непрерывна на промежутке (x_*, x^*).

Чтобы гарантировать дифференцируемость неявной функции, нужны более сильные предположения. Один из возможных вариантов даётся следующей теоремой. При её первом чтении для простоты следует считать $m = 1$. Это позволяет сделать полезный рисунок. После этого остаётся проследить, что замена вещественного параметра x на вектор почти ничего не меняет в доказательстве. Предварительно установим несложное вспомогательное утверждение.

Лемма 2 (условие Липшица для неявной функции). *Пусть X — непустое множество в $\mathbb{R}^m$, $-\infty < y_* < y^* < +\infty$. Допустим, что функция F, заданная на произведении $E = X \times [y_*, y^*]$ удовлетворяет условиям 1) и 2) леммы 1, и, кроме того, существуют такие положительные числа* С *и μ, что*

$$а) \quad |F(x'', y) - F(x', y)| \leqslant C\|x'' - x'\|_m \quad \text{для всех} \;\; x'', x' \in X \quad \text{и} \;\; y \in [y_*, y^*];$$

$$б) \quad |F(x, y'') - F(x, y')| \geqslant \mu\,|y'' - y'| \quad \text{для всех} \;\; x \in X \quad \text{и} \;\; y'', y' \in [y_*, y^*].$$

Тогда уравнение $F(x, y) = 0$ задаёт единственную неявную функцию φ на множестве X и она удовлетворяет условию Липшица на нём.

Доказательство. Уравнение $F(x, y) = 0$ имеет решение в силу условий 1), 2). Неравенство б) обеспечивает единственность решения. Остаётся проверить, что неявная функция φ, задаваемая этим уравнением, удовлетворяет условию Липшица. Для этого возьмём произвольные точки x', x'' из X и рассмотрим число $F(x', \varphi(x''))$. С одной стороны, так как $F(x'', \varphi(x'')) = 0$, то с помощью неравенства а) мы получаем

$$|F(x', \varphi(x''))| = |F(x', \varphi(x'')) - F(x'', \varphi(x''))| \leqslant C\|x' - x''\|_m.$$

С другой стороны, так как $F(x', \varphi(x')) = 0$, то неравенство б) даёт нам

$$|F(x', \varphi(x''))| = |F(x', \varphi(x'')) - F(x', \varphi(x'))| \geqslant \mu|\varphi(x') - \varphi(x'')|.$$

Поэтому

$$|\varphi(x') - \varphi(x'')| \leqslant \frac{1}{\mu}|F(x', \varphi(x''))| \leqslant \frac{C}{\mu}\|x' - x''\|_m.$$

Лемма доказана.

Теорема (дифференцируемость неявной функции). *Пусть X — открытое множество в $\mathbb{R}^m$, $-\infty < y_* < y^* < +\infty$, а функция F непрерывна на произведении $E = X \times [y_*, y^*]$ и непрерывно дифференцируема внутри его.*

Тогда, если F удовлетворяет условию 2) леммы 1 и $F'_{m+1} = \frac{\partial F}{\partial y} \neq 0$ внутри E, то уравнение $F(x, y) = 0$ задаёт на множестве X единственную неявную функцию φ и она непрерывно дифференцируема на X.

Доказательство. Функция $y \mapsto F(x, y)$ строго монотонна для любого x из X, так как производная F'_{m+1} отлична от нуля. Поэтому выполнены все условия леммы 1, гарантирующей существование и единственность неявной функции φ. Её значения принадлежат интервалу (y_*, y^*).

Убедимся в непрерывности φ в произвольной точке x_0 из X. Положим $y_0 = \varphi(x_0)$. Тогда $y_* < y_0 < y^*$ и, следовательно, существует столь малое число $\Delta > 0$, что $[y_0 - \Delta, y_0 + \Delta] \subset (y_*, y^*)$. Строго монотонная функция $y \mapsto F(x_0, y)$ меняет знак в точке y_0: $F(x_0, y_0 - \Delta)F(x_0, y_0 + \Delta) < 0$. Поскольку функция F непрерывна, существует такое число $\delta > 0$, что $B = B_\delta(x_0) \subset X$ и

$$F(x, y_0 - \Delta)F(x, y_0 + \Delta) < 0 \;\; \text{для всех} \;\; x \;\; \text{из} \;\; B.$$

Таким образом, все значения неявной функции φ, принимаемые ею в шаре B, лежат между $y_0 - \Delta$ и $y_0 + \Delta$.

Положим $\mu = \frac{1}{2}\left|F'_{m+1}(x_0, y_0)\right|$, $\mu > 0$, и $C = 2\|\operatorname{grad} F(x_0, y_0)\|_{m+1}$. Так как функция F непрерывно дифференцируема, то числа δ и Δ можно выбрать столь малыми, что

$$\left|F'_{m+1}(x, y)\right| \geqslant \mu \quad \text{и} \quad \|\operatorname{grad} F(x, y)\|_{m+1} \leqslant C, \quad \text{если} \;\; x \in B \;\; \text{и} \;\; |y - y_0| < \Delta.$$

Первое неравенство даёт нам оценку снизу: согласно формуле конечных приращений при некотором $\overline{y}$ из промежутка $[y', y'']$ мы имеем

$$|F(x,y'') - F(x,y')| = \left|F'_{m+1}(\overline{y})\right| \cdot |y''-y'| \geqslant \mu|y''-y'|, \quad \text{если} \quad x \in B \quad \text{и} \quad y', y'' \in [y_0-\Delta, y_0+\Delta],$$

а второе — оценку сверху

$$|F(x'',y) - F(x',y)| \leqslant C\|(x'',y)-(x',y)\|_{m+1} = C\|x''-x'\|_m, \quad \text{если} \quad x', x'' \in B \quad \text{и} \quad |y-y_0| < \Delta.$$

Поэтому можно воспользоваться леммой 2, заменив в ней множество X шаром B, а промежуток $[y_*, y^*]$ — промежутком $[y_0-\Delta, y_0+\Delta]$. Следовательно, в шаре B неявная функция φ удовлетворяет условию Липшица.

Докажем теперь дифференцируемость φ в $x_0 = \left(x_1^\circ, \ldots, x_m^\circ\right)$. Для этого воспользуемся дифференцируемостью функции F в точке (x_0, y_0):

$$F(x,y) = F(x_0,y_0) + \sum_{j=1}^{m} l_j \cdot (x_j - x_j^\circ) + l_{m+1} \cdot (y-y_0) + \|(x,y)-(x_0,y_0)\|_{m+1}\, \alpha(x,y),$$

где $(l_1, \ldots, l_m, l_{m+1}) = \operatorname{grad} F(x_0,y_0)$, а функция α непрерывна в точке (x_0,y_0) и $\alpha(x_0,y_0)=0$. Взяв $y=\varphi(x)$, где $x \in B$, мы получим (так как $F(x,\varphi(x)) = F(x_0,\varphi(x_0)) = 0$):

$$0 = \sum_{j=1}^{m} l_j \cdot (x_j - x_j^\circ) + l_{m+1} \cdot (\varphi(x)-\varphi(x_0)) + \|(x,\varphi(x))-(x_0,\varphi(x_0))\|_{m+1}\, \alpha(x,\varphi(x)).$$

Поскольку $l_{m+1} = F'_{m+1}(x_0,y_0) \neq 0$, отсюда следует, что

$$\varphi(x) = \varphi(x_0) + \sum_{j=1}^{m} A_j \cdot (x_j - x_j^\circ) + \|x-x_0\|_m\, \beta(x),$$

где

$$A_j = -\frac{l_j}{l_{m+1}} = -\frac{F'_j(x_0,y_0)}{F'_{m+1}(x_0,y_0)}$$

и при $x \neq x_0$

$$\beta(x) = -\frac{\alpha(x,\varphi(x))}{l_{m+1}\|x-x_0\|_m}\|(x,\varphi(x))-(x_0,\varphi(x_0))\|_{m+1}.$$

Проверим, что $\lim_{x_0}\beta = 0$. Так как функция φ удовлетворяет условию Липшица в шаре B, то

$$\|(x,\varphi(x))-(x_0,\varphi(x_0))\|_{m+1} = \sqrt{\|x-x_0\|_m^2 + (\varphi(x)-\varphi(x_0))^2} = O\big(\|x-x_0\|_m\big).$$

Таким образом,

$$|\beta(x)| \leqslant \text{const}\,|\alpha(x,\varphi(x))| \underset{x\to x_0}{\longrightarrow} \text{const}\,|\alpha(x_0,\varphi(x_0))| = \text{const}\,|\alpha(x_0,y_0)| = 0.$$

Итак, неявная функция φ дифференцируема в произвольной точке x_0 из X и

$$\varphi'_j(x_0) = A_j = -\frac{F'_j(x_0,y_0)}{F'_{m+1}(x_0,y_0)} = -\frac{F'_j(x_0,\varphi(x_0))}{F'_{m+1}(x_0,\varphi(x_0))}.$$

В числителе и знаменателе стоят суперпозиции непрерывных функций, причём по условию знаменатель нигде не обращается в нуль. Следовательно, дробь непрерывна, т.е. функция φ непрерывно дифференцируема в произвольной точке x_0 из X. Теорема доказана.

Существенный недостаток теоремы — слишком сильные предположения. В реальной ситуации чаще всего решается (относительно y) уравнение $F(x_1, \dots, x_m, y) = 0$ с гладкой функцией F, а информации о промежутке $[y_*, y^*]$ нет (см. примеры 1 и 2). Её можно извлечь из других предположений о функции F — равенства $F(x_0, y_0) = 0$ и неравенства $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$. Однако за это ослабление предположений приходится платить — гарантируется существование неявной функции лишь вблизи точки x_0.

Теорема (дифференцируемость неявной функции — локальный вариант). *Пусть функция $F: E \to \mathbb{R}$ непрерывно дифференцируема на открытом множестве E, $E \subset \mathbb{R}^{m+1}$, а точка (x_0, y_0) из E ($x_0 \in \mathbb{R}^m$, $y_0 \in \mathbb{R}$) такова, что*

$$F(x_0, y_0) = 0 \quad \textit{и} \quad F'_{m+1}(x_0, y_0) = \frac{\partial F}{\partial y}(x_0, y_0) \neq 0.$$

Тогда вблизи (x_0, y_0) уравнение $F(x, y) = 0$ задаёт единственную функцию, равную y_0 в точке x_0, и эта функция гладкая: существуют такой шар B, $B \subset \mathbb{R}^m$, с центром в x_0 и такое число $\Delta > 0$, что

1) $B \times (y_0 - \Delta, y_0 + \Delta) \subset E$;

2) для любого x из B существует единственное число $\varphi(x) \in (y_0 - \Delta, y_0 + \Delta)$, удовлетворяющее уравнению $F(x, \varphi(x)) = 0$; в частности $\varphi(x_0) = y_0$;

3) функция φ непрерывно дифференцируема на B;

4) для всех $j = 1, \dots, m$ и $x \in B$

$$\varphi'_j(x) = \frac{\partial \varphi}{\partial x_j}(x) = -\frac{\frac{\partial F}{\partial x_j}(x, \varphi(x))}{\frac{\partial F}{\partial y}(x, \varphi(x))} = -\frac{F'_j(x, \varphi(x))}{F'_{m+1}(x, \varphi(x))}.$$

Говоря иными словами, вблизи (x_0, y_0) множество $\{(x, y) \in E \mid F(x, y) = 0\}$ устроено просто — это график гладкой функции m переменных.

Следует особо отметить роль условия $F'_{m+1}(x_0, y_0) \neq 0$. В примере 1 оно нарушено при $x_0 = \pm 1$, а в примере 2 — ещё и при $x_0 = 0$. Это привело к отсутствию неявной функции в окрестностях точек ± 1 (в обоих примерах) и её неединственности вблизи нуля (в примере 2).

Доказательство. Поскольку (x_0, y_0) — внутренняя точка множества E, для достаточно малых положительных чисел δ и Δ мы получим $B_\delta(x_0) \times [y_0 - \Delta, y_0 + \Delta] \subset E$. Так как $F'_{m+1}(x_0, y_0) \neq 0$, то числа δ и Δ можно выбрать столь малыми, что $F'_{m+1}(x, y) \neq 0$, если $x \in B_\delta(x_0)$ и $|y - y_0| \leqslant \Delta$. Тогда для любого x из шара $B_\delta(x_0)$ функция $y \mapsto F(x, y)$ строго монотонна на $[y_0 - \Delta, y_0 + \Delta]$, а так как $F(x_0, y_0) = 0$, то $F(x_0, y_0 - \Delta) F(x_0, y_0 + \Delta) < 0$. По теореме о стабилизации знака это неравенство сохраняется при замене x_0 достаточно близкой точкой x. Уменьшив в случае необходимости радиус δ, мы можем считать, что

$$F(x, y_0 - \Delta) F(x, y_0 + \Delta) < 0, \text{ если } \ x \in B_\delta(x_0).$$

В результате мы оказываемся в условиях предыдущей теоремы с $[y_*, y^*] = [y_0 - \Delta, y_0 + \Delta]$ и $X = B_\delta(x_0)$. Она гарантирует существование единственной неявной функции $\varphi : B_\delta(x_0) \to (y_0 - \Delta, y_0 + \Delta)$ и её непрерывную дифференцируемость в шаре $B_\delta(x_0)$. Теорема доказана.

Следствие. Используя выражения частных производных неявной функции φ через производные функции F, уравнение касательной к графику φ в точке $(x_0, y_0) = (x_0, \varphi(x_0))$

$$y = \varphi(x_0) + \langle \operatorname{grad} \varphi(x_0), x - x_0 \rangle = y_0 + \sum_{j=1}^{m} \varphi'_j(x_0)(x_j - x_j^\circ)$$

можно записать в виде

$$F'_{m+1}(x_0, y_0)(y - y_0) = -\sum_{j=1}^{m} F'_j(x_0, y_0)(x_j - x_j^\circ) \qquad \text{или} \qquad \langle \operatorname{grad} F(x_0, y_0), (x - x_0, y - y_0)\rangle = 0.$$

Пример 3. Рассмотрим алгебраическое уравнение $y^m + a_1 y^{m-1} + \ldots + a_{m-1}y + a_m = 0$ с вещественными коэффициентами $a_1, \ldots, a_m$. Пусть $y = b$ — его *вещественный* корень. Как он изменится, если немного изменить коэффициенты уравнения?

Уже при $m = 2$ на примере уравнения $y^2 + a = 0$ видно, что при $a < 0$ малое изменение этого коэффициента приводит к плавному изменению корней — они остаются вещественными и "хорошо зависят" от a (функция $a \mapsto \sqrt{-a}$ бесконечно дифференцируема в любой точке a, $a < 0$). Но при $a = 0$ ситуация совсем другая: всё зависит от такого, в какую сторону будет изменён нулевой коэффициент. Если он стал отрицательным, то вместо единственного вещественного корня $y = 0$ мы получим пару вещественных корней, которые "плохо зависят" от a (функция $a \mapsto \sqrt{-a}$ не дифференцируема в нуле). Если же нулевой коэффициент заменить положительным (сколь угодно малым!), то вещественные корни исчезают вовсе.

С уравнениями более высоких степеней ситуация вполне аналогичная. Нарисовав эскиз графика многочлена, нетрудно догадаться, что *простой* (т.е. не кратный) корень уравнения устойчив к малому изменению коэффициентов многочлена — он равен значению некоторой непрерывно дифференцируемой функции от переменных $a_1, \ldots, a_m$. Если же корень *кратный*, то малое изменение коэффициентов многочлена может привести как к исчезновению этого корня, так и к замене его несколькими новыми корнями. Например, уравнение $y^3 = 0$ имеет единственный корень $y = 0$ (кратности три), а близкое уравнение $y^3 + ay = 0$ с отрицательным (и сколь угодно малым) коэффициентом a имеет три различных корня (два из них "плохо зависят" от a).

Воспользуемся локальным вариантом теоремы о неявной функции, чтобы придать строгость этим догадкам. Для этого определим на $\mathbb{R}^{m+1}$ бесконечно дифференцируемую функцию F:

$$F(x, y) = y^m + x_1\, y^{m-1} + \ldots + x_{m-1}\, y + x_m \quad \text{для любых} \;\; x \in \mathbb{R}^m \quad \text{и} \;\; y \in \mathbb{R}.$$

Пусть при некотором выборе коэффициентов $x_0 = a = (a_1, \ldots, a_m)$ уравнение $F(a, y) = 0$, т.е.

$$y^m + a_1\, y^{m-1} + \ldots + a_{m-1}\, y + a_m = 0$$

имеет вещественный корень $y = b$: $F(a, b) = 0$. Согласно теореме, если $F'_{m+1}(a, b) \neq 0$, то для набора коэффициентов $x = (x_1, \ldots, x_m)$ достаточно близкого к $a = (a_1, \ldots, a_m)$ уравнение $F(x, y) = 0$ имеет единственный корень $y = \varphi(x)$, лежащий около точки b, причём функция φ непрерывно дифференцируема вблизи точки a (корень уравнения плавно меняется при изменении коэффициентов). Остаётся заменить, что неравенству $F'_{m+1}(a, b) \neq 0$ удовлетворяют все простые корни уравнения $F(a, y) = 0$. Если же корень кратный, то эта производная равна нулю, т.е. условие теоремы не выполнено.

§2. **Обратное отображение**

Постановка задачи. Пусть E — открытое множество в $\mathbb{R}^m$ и $F : E \to \mathbb{R}^m$ — непрерывно дифференцируемое отображение. В каком случае можно гарантировать существование обратного отображения F^{-1} и его непрерывную дифференцируемость?

При $m = 1$ (E интервал в $\mathbb{R}$) ответ на этот вопрос очевиден — нужно, чтобы производная функции F не обращалась в нуль на этом интервале. Тогда она сохраняет знак на нём, поэтому функция F строго монотонна и, следовательно, существует обратная функция $G = F^{-1}$, заданная на промежутке $F(E)$. Поскольку F строго монотонна на интервале E, промежуток $F(E)$, очевидно, открыт. При этом функция G всюду дифференцируема и $G'(y) = \frac{1}{F'(G(y))}$ для любого y из $F(E)$. В частности, функция G' непрерывна. Таким образом, уравнение $F(x) = y$ имеет “хорошее” решение: для всех чисел y из $F(E)$ найдётся такое число $x = G(y)$, что $F(G(y)) = y$, это решение единственно и гладким образом зависит от параметра y, стоящего в правой части уравнения.

В кратном случае (т.е. при $m > 1$) вместо одного уравнения рассматривается система m уравнений относительно переменных $x_1, \dots, x_m$

$$\begin{cases} F_1(x_1, \dots, x_m) = y_1 \\ \dots\dots\dots\dots\dots \\ F_m(x_1, \dots, x_m) = y_m. \end{cases}$$

Сразу возникают естественные вопросы. Для каких значений параметров $y_1, \dots, y_m$ она имеет решение? Локальный вариант этого вопроса: если система имеет решение $x_1^\circ, \dots, x_m^\circ$ при некотором наборе параметров $y_1^\circ, \dots, y_m^\circ$, стоящих в правых частях уравнений, то можно ли утверждать, что их небольшое изменение не сделает систему неразрешимой? Иначе говоря, верно ли, что для всякого вектора $y = (y_1, \dots, y_m)$ достаточно близкого к вектору $y_0 = (y_1^\circ, \dots, y_m^\circ)$ (т.е. такого, что норма $\|y - y_0\|_m$ мала) рассматриваемая система уравнений по-прежнему будет иметь какое-то решение $x = (x_1, \dots, x_m)$? Таким образом, рассмотрев отображение F с координатными функциями $F_1, \dots, F_m$, заданное вблизи $x_0 = (x_1^\circ, \dots, x_m^\circ)$, мы приходим к задаче обращения этого отображения вблизи точки $y_0 = F(x_0)$. Разумеется, решение уравнения $y = F(x)$ зависит от y: $x = G(y)$, где G — некоторое отображение, определённое вблизи точки y_0, $G(y_0) = x_0$. Будет ли это решение $x = (x_1, \dots, x_m)$ “хорошо зависеть” от параметров $y = (y_1, \dots, y_m)$? Точнее, будет ли отображение G непрерывно дифференцируемым в окрестности точки y_0? Разумеется, в первую очередь необходимо обеспечить существование обратного отображения, а затем выяснять, какие дополнительные предположения гарантируют его непрерывную дифференцируемость.

Отметим существенное отличие кратной задачи ($m > 1$) от одномерной. На прямой из наличия *локального* решения уравнения вблизи каждой точки вытекает существование *глобального* решения (непрерывная функция взаимно однозначная вблизи каждой точки промежутка взаимно однозначна на всём промежутке, поскольку локальная строгая монотонность функции влечёт её монотонность на всём промежутке). В кратном случае ситуация значительно сложнее.

Пример 1. Пусть $E = \mathbb{R} \times (0, +\infty)$ — верхняя полуплоскость. Рассмотрим бесконечно дифференцируемое отображение F этого открытого множества в $\mathbb{R}^2$, действующее по правилу $F(\theta, r) = (r\cos\theta, r\sin\theta)$ для любой точки (θ, r) из E. Геометрический смысл этого преобразования очевиден — восстановление декартовых координат точки по её полярным координатам. Ясно, что $F(E) = \mathbb{R}^2 \setminus \{\mathbb{O}_2\}$ и отображение F не взаимно однозначно (из-за периодичности по переменной θ). Поэтому оно не обратимо на всём множестве $F(E)$. В то же время для любых чисел θ_0 и $r_0 > 0$ существует такая окрестность U точки (θ_0, r_0), что сужение F на эту окрестность взаимно однозначно. Например, можно взять открытый прямоугольник $U = (\theta_0 - \pi, \theta_0 + \pi) \times (0, 2r_0)$.

Ещё одно преимущество одномерной ситуации обеспечивается теоремой о сохранении промежутка. Она позволяет легко описать образ промежутка для любой непрерывной функции. В частности, если функция F строго монотонна, а промежуток E открыт, то $F(E)$ — тоже открытый промежуток. В кратном случае нет понятия монотонности, и поэтому мы лишены такого простого и эффективного средства для описания образа даже для “простых” множеств из $\mathbb{R}^m$ (шары, прямоугольные параллелепипеды и т.п.). Например, вызывает немалое затруднение вопрос “Какие точки открытого множества E, $E \subset \mathbb{R}^m$, переводятся гладким отображением F во внутренние точки множества $F(E)$?” При этом предположения о строгой монотонности координатных функций по каждой переменной ничего не дают. Чтобы убедиться в этом, достаточно рассмотреть линейное отображение, которому соответствует матрица с положительными элементами (как следует из леммы о взаимно однозначных линейных отображениях, важны не знаки элементов матрицы, а её ранг).

Отмеченные затруднения вынуждают в кратном случае ограничиваться получением локальных результатов. При этом основная идея дальнейших рассуждений довольна проста. Это идея *линеаризации* — поскольку всякое "хорошее" отображение локально "почти линейно", можно надеяться, что ситуация значительно прояснится, если разобраться с упрощённой задачей, получающейся заменой (вблизи фиксированной точки x_0) сложно устроенного отображения $F(x)$ более простым $\widetilde{F}(x) = F(x_0) + \big(d_{x_0}F\big)(x - x_0)$. Вместо того, чтобы решать (относительно x) уравнение $F(x) = y$ с правой частью близкой к $y_0 = F(x_0)$ рассматривается "похожее уравнение" $\widetilde{F}(x) \approx y$:

$$F(x_0) + \big(d_{x_0}F\big)(x - x_0) \approx y.$$

Чтобы выписать его решение, надо предположить обратимость линейного отображения $L = d_{x_0}F$, $L \in \mathscr{L}_{mm}$. Тогда

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

Таким образом, в этой упрощённой задаче естественно появляется условие $L \in \operatorname{Isom}(\mathbb{R}^m)$, которое будет играть решающую роль во всех дальнейших рассуждениях. В одномерной ситуации его смысл совершенно очевиден — производная не обращается в нуль. Без этого условия невозможно обеспечить гладкость обратной функции.

К сожалению, замена расплывчатого соотношения " $\approx$ " знаком " $=$ " потребует значительных усилий. Для получения основного результата этого параграфа (теорема об обратном отображении) мы установим предварительные результаты, в которых будем последовательно уменьшать априорные предположения об обратном отображении F^{-1}, стремясь выводить их из предположений об исходном отображении F. Наш план таков.

1) Научиться дифференцировать обратное отображение, если оно существует и дифференцируемо.

2) Найти условия на отображение, гарантирующие его локальную обратимость.

3) Найти условия, обеспечивающие открытость множества $F(E)$.

4) Выяснить, в каком случае обратное отображение непрерывно дифференцируемо.

Лемма (дифференцирование обратного отображения). *Пусть $E \subset \mathbb{R}^m$, $F : E \to \mathbb{R}^m$ — взаимно однозначное отображение, $x_0 \in \operatorname{Int}(E)$, $y_0 = F(x_0) \in \operatorname{Int}\big(F(E)\big)$. Тогда, если отображение F дифференцируемо в точке x_0, а отображение F^{-1} дифференцируемо в точке y_0, то $d_{x_0}F \in \operatorname{Isom}(\mathbb{R}^m)$ и*

$$d_{y_0}F^{-1} = \left(d_{x_0}F\right)^{-1}.$$

В частности, для дифференцируемости отображения, обратного к дифференцируемому, необходимо, чтобы матрица Якоби $F'(x_0)$ была неособенной. В терминах частных производных это необходимое условие сводится к неравенству

$$\begin{vmatrix} \frac{\partial F_1}{\partial x_1}(x_0) & \dots & \frac{\partial F_1}{\partial x_m}(x_0) \\ \dots & \dots & \dots \\ \frac{\partial F_m}{\partial x_1}(x_0) & \dots & \frac{\partial F_m}{\partial x_m}(x_0) \end{vmatrix} \neq 0.$$

Д о к а з а т е л ь с т в о. По условию отображение F биективно отображает множество E на $F(E)$. Пусть $G = F^{-1}$. Тогда $x = G\big(F(x)\big)$ для всех x из E, т.е. на E суперпозиция $G \circ F$

совпадает с тождественным отображением Id. Аналогично, $\mathrm{Id} = F \circ G$ на $F(E)$. Поскольку $d\,\mathrm{Id} = \mathrm{Id}$, мы получаем (с помощью теоремы о дифференцировании суперпозиции) равенства

$$\mathrm{Id} = d_{x_0}\big(G \circ F\big) = d_{y_0}G \circ d_{x_0}F,$$
$$\mathrm{Id} = d_{y_0}\big(F \circ G\big) = d_{x_0}F \circ d_{y_0}G.$$

Поэтому отображения $d_{y_0}G$ и $d_{x_0}F$ обратимы и $d_{y_0}G = \big(d_{x_0}F\big)^{-1}$. Лемма доказана.

Из леммы следует *необходимое* условие дифференцируемости отображения, обратного к дифференцируемому — надо, чтобы дифференциал $d_{x_0}F$ был обратим. Это можно проиллюстрировать таким простым примером.

Пример 2. Пусть $m = 1$, $E = \mathbb{R}$ и $F(x) = x^3$. Это, очевидно, биекция и $F^{-1}(y) = \sqrt[3]{y}$ для любого вещественного y. Кроме того, оба отображения F и F^{-1} непрерывны на $\mathbb{R}$. Тем не менее, хотя функция F непрерывно дифференцируема всюду на $\mathbb{R}$, обратная функция не дифференцируема в нуле.

Определение. Пусть x_0 — внутренняя точка множества E, $E \subset \mathbb{R}^m$, и отображение $F : E \to \mathbb{R}^m$ дифференцируемо в x_0. Точку x_0 будем называть *критической точкой* отображения F, если матрица Якоби $F'(x_0)$ необратима, т.е. если $\det F'(x_0) = \det\big(\frac{\partial F_k}{\partial x_j}(x_0)\big)_{1\leqslant k,j\leqslant m} = 0$.

Таким образом, для дифференцируемости обратного отображения в точке $y_0 = F(x_0)$ необходимо, чтобы x_0 была не критической. В примере 2 таковы все точки $x \neq 0$, а 0 — критическая точка. К сожалению некритичность точки — лишь необходимое, но не достаточное условие дифференцируемости обратного отображения. Оно не обеспечивает обратимости (даже вблизи точки x_0) F. В этом можно убедиться уже в одномерном случае.

Пример 3. Пусть $m = 1$, $E = \mathbb{R}$ и $F(x) = x + x^2\,\sin\frac{\pi}{x}$ при $x \neq 0$, $F(0) = 0$. Ясно, что

$$\frac{F(x) - F(0)}{x} = 1 + x\sin\frac{\pi}{x} \to 1 \qquad \text{при} \quad x \to 0.$$

Поэтому $F'(0) = 1$ и, следовательно, d_0F — тождественное отображение прямой $\mathbb{R}$. В частности, $d_0F \in \mathrm{Isom}\,(\mathbb{R})$, т.е. нуль не критическая точка. Тем не менее, функция F не обратима ни в какой, сколь угодно малой окрестности нуля. Действительно,

$$F'(x) = 1 + 2x\sin\frac{\pi}{x} - \pi\cos\frac{\pi}{x} \qquad \text{при} \quad x \neq 0.$$

Для любого номера n мы имеем

$$F'\Big(\frac{1}{n}\Big) = 1 - (-1)^n\pi.$$

Взяв номер n чётным, получим, что сколь угодно близко к нулю есть точки, в которых производная F' отрицательна, и поэтому вблизи этих точек F убывает. Аналогично, взяв нечётный номер n, получим возрастание F на интервалах, сколь угодно близких к нулю. Таким образом, непрерывная функция F меняет характер монотонности в любой окрестности нуля и, следовательно, она не инъективна в каждой такой окрестности.

Нетрудно понять причину, из–за которой это произошло. Дело в том, что производная терпит разрыв в точке 0 (хотя $F'(0) = 1$, тем не менее $F'(x) < 0$ для некоторых, но сколь угодно малых x). Если бы она была непрерывной, то сохраняла бы знак в некоторой окрестности нуля и, следовательно, там функция F была бы строго монотонной. Поэтому в дальнейшем мы ограничимся рассмотрением лишь “достаточно хороших” отображений.

Теорема (о гладких отображениях). *Пусть в открытом множестве $E \subset \mathbb{R}^m$ задано всюду дифференцируемое отображение F, действующее в $\mathbb{R}^M$, и $F_1, \dots, F_M$ — его координатные функции. Тогда следующие два утверждения равносильны*

1) *для всех $k = 1, \dots, M$ и всех $j = 1, \dots, m$ функция $\dfrac{\partial F_k}{\partial x_j}$ непрерывна на E;*

2) *для каждой точки $x_0 \in E$ и любого числа $\varepsilon > o$ существует такое число $\delta > 0$, что $\|d_x F - d_{x_0} F\|_{\mathscr{L}} < \varepsilon$, как только $x \in E$ и $\|x - x_0\|_m < \delta$.*

Утверждение 2) означает просто, что операция дифференцирования, т.е. отображение $x \mapsto d_x F$ из E в $\mathscr{L}_{mM}$, непрерывна всюду на E.

Определение. Отображение F, удовлетворяющее (равносильным) условиям 1) и 2), называется *гладким* или *непрерывно дифференцируемым* на множестве E.

Доказательство. 1)$\Rightarrow$2). Пусть $x_0 \in E$ и $\varepsilon > 0$. Так как частная производная $\frac{\partial F_k}{\partial x_j}$ непрерывна в точке x_0, то найдётся такое число $\delta_{jk} > 0$, что для всех $x \in E$, $\|x - x_0\|_m < \delta_{jk}$, выполняется неравенство
$$\left|\frac{\partial F_k}{\partial x_j}(x) - \frac{\partial F_k}{\partial x_j}(x_0)\right| < \frac{\varepsilon}{\sqrt{mM}}.$$
Возьмём $\delta = \min_{jk} \delta_{jk}$. Ясно, что это положительное число, причём последние неравенства справедливы одновременно для всех $k = 1, \dots, M$ и $j = 1, \dots, m$, если $x \in E$ и $\|x - x_0\|_m < \delta$. Убедимся, что δ — требуемое число, т.е. $\|d_x F - d_{x_0} F\|_{\mathscr{L}} < \varepsilon$ для $x \in E \cap B_\delta(x_0)$. Отображению $d_x F - d_{x_0} F$ соответствует матрица
$$F'(x) - F'(x_0) = \left(\frac{\partial F_k}{\partial x_j}(x) - \frac{\partial F_k}{\partial x_j}(x_0)\right)_{\substack{1 \leqslant j \leqslant m \\ 1 \leqslant k \leqslant M}}.$$
Поэтому (см. третье неравенство после определения нормы линейного отображения во введении)
$$\|d_x F - d_{x_0} F\|_{\mathscr{L}} \leqslant \sqrt{\sum_{k=1}^{M}\sum_{j=1}^{m}\left(\frac{\partial F_k}{\partial x_j}(x) - \frac{\partial F_k}{\partial x_j}(x_0)\right)^2} < \sqrt{\sum_{k=1}^{M}\sum_{j=1}^{m}\frac{\varepsilon^2}{mM}} = \varepsilon,$$
если $x \in E$ и $\|x - x_0\|_m < \delta$. Так как здесь x_0 — любая точка из E, а ε — произвольное положительное число, то утверждение 2) доказано.

2)$\Rightarrow$1). Пусть как и на предыдущем шаге x_0 — произвольная точка из E, ε — любое положительное число. Согласно утверждению 2) существует такое $\delta > 0$, что $\|d_x F - d_{x_0} F\|_{\mathscr{L}} < \varepsilon$ для всех x из пересечения $E \cap B_\delta(x_0)$. Иначе говоря, для таких x справедливо неравенство $\sup\limits_{\|h\|_m \leqslant 1} \|(d_x F - d_{x_0} F)(h)\|_M < \varepsilon$. Взяв в качестве вектора h произвольный вектор канонического базиса e_j ($j = 1, \dots, m$), мы видим, что
$$\|(d_x F - d_{x_0} F)(e_j)\|_M = \|(F'(x) - F'(x_0)) \cdot e_j^T\|_M = \sqrt{\sum_{k=1}^{M}\left(\frac{\partial F_k}{\partial x_j}(x) - \frac{\partial F_k}{\partial x_j}(x_0)\right)^2} < \varepsilon$$
и тем более для всех $k = 1, \dots, M$
$$\left|\frac{\partial F_k}{\partial x_j}(x) - \frac{\partial F_k}{\partial x_j}(x_0)\right| < \varepsilon, \quad \text{если} \ \ x \in E \cap B_\delta(x_0).$$

Поскольку ε — произвольное положительное число, это означает, что каждая частная производная $\frac{\partial F_k}{\partial x_j}$ непрерывна в любой наперёд заданной точке x_0 из E. Теорема доказана.

Непрерывная дифференцируемость (или короче гладкость) отображения — условие, которое обычно выполняется в прикладных задачах. Желательно, чтобы таким оказалось и обратное отображение (если оно существует). Отображения, которые не только сами непрерывно дифференцируемы, но и имеют столь же хорошее обратное, представляют особый интерес и для теории, и для приложений. Поэтому естественно выделить класс таких отображений.

Определение. *Диффеоморфизмом* двух открытых множеств E и $\widetilde{E}$ из $\mathbb{R}^m$ называется такая биекция $F: E \to \widetilde{E}$, что оба отображения F и F^{-1} непрерывно дифференцируемы во всех точках множеств E и $\widetilde{E}$ соответственно.

Например, функция x^3 — не диффеоморфизм прямой, но её сужение на правую полуось, очевидно, является диффеоморфизмом. Привести примеры диффеоморфизмов всей прямой на какой-то интервал совсем несложно: e^x, $\operatorname{arctg} x$, $2x + \sin x$ и др. С помощью этих примеров строятся диффеоморфизмы в кратном случае, например, $(x, y) \mapsto (e^x, \operatorname{arctg} y)$ — диффеоморфизм между плоскостью $\mathbb{R}^2$ и полуполосой $(0, +\infty) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, а отображение $(x, y) \mapsto (\ln x, 2y + \sin y)$ диффеоморфно переводит правую полуплоскость $(x > 0)$ на $\mathbb{R}^2$. Привести примеры в кратном случае, в которых не используется "разделение переменных", значительно сложнее (так как непросто установить биективность отображения).

Отметим, что суперпозиция диффеоморфизмов — диффеоморфизм, а их сумма (если они определены на одном множестве) — гладкое, но не обязательно обратимое отображения.

Ясно, что определитель $\det F'(x)$ матрицы Якоби непрерывно зависит от x, если F — гладкое отображение. Из леммы о дифференцировании обратного отображения сразу следует, что для диффеоморфизма этот определитель всюду отличен от нуля, т.е. диффеоморфизм не имеет критических точек. Если он задан в некоторой области (т.е. на открытом линейно связном множестве), то $\det F'(x)$ сохраняет в ней знак.

Очевиден недостаток леммы — сделаны очень сильные предположения об обратном отображении (оно существует и дифференцируемо, а точка $y_0 = F(x_0)$ лежит внутри множества $F(E)$). Эти три условия трудно проверить, зная лишь координатные функции исходного отображения F. В первую очередь постараемся избавиться от предположения о его обратимости. Хотя формальное доказательство потребует от нас некоторых усилий, его идея совсем проста: дифференцируемое отображение "почти линейно", поэтому его приращение в малом шаре почти линейно. Точнее, если r — малое положительное число, то для всех x'' и x' из шара $B_r(x_0)$, $B_r(x_0) \subset E$, мы имеем $F(x'') \approx F(x_0) + L(x'' - x_0)$ и $F(x') \approx F(x_0) + L(x' - x_0)$, где $L = d_{x_0}F$. Поэтому $F(x'') - F(x') \approx L(x'' - x')$. Если $L \in \operatorname{Isom}(\mathbb{R}^m)$, то $L(x'' - x') \neq \mathbb{O}_m$ при $x'' \neq x'$ и можно надеяться, что $F(x'') - F(x') \neq \mathbb{O}_m$ для различных точек x'' и x' достаточно близких к не критической точке x_0.

Наша ближайшая цель — не только формализовать эти нестрогие рассуждения, но и дополнить результат оценкой уклонения $\|F(x'') - F(x')\|_m$ снизу. Для этого воспользуемся многомерным аналогом классической теоремы о среднем.

Теорема (неравенство Лагранжа). *Пусть отображение F, действующее из шара B пространства $\mathbb{R}^m$ в пространство $\mathbb{R}^M$, дифференцируемо во всех точках этого шара. Положим $C = \sup\limits_{x \in B} \|d_x F\|_{\mathscr{L}}$. Тогда*

$$\|F(x'') - F(x')\|_M \leqslant C\|x'' - x'\|_m \quad \textit{для всех} \ \ x'', x' \ \textit{из} \ \ B.$$

Таким образом, если $C < +\infty$, то отображение F удовлетворяет условию Липшица. Если

$F_1, \dots, F_M$ — координатные функции отображения, то

$$\|d_x F\|_{\mathscr{L}} \leqslant \sqrt{\sum_{k=1}^{M}\sum_{j=1}^{m}\left(\frac{\partial F_k}{\partial x_j}(x)\right)^2}.$$

Поэтому условие $C < +\infty$ заведомо выполнено, если все частные производные ограничены в шаре B. Верно и обратное. Действительно, $\left|\frac{\partial F_k}{\partial x_j}(x)\right| \leqslant \|d_x F\|_{\mathscr{L}}$ для любых j, k и x (как было отмечено во введении, $|l_{kj}| \leqslant \|L\|_{\mathscr{L}}$, если линейному отображению L соответствует матрица с элементами l_{kj}). Поэтому всюду в шаре B выполняются неравенства $\left|\frac{\partial F_k}{\partial x_j}(x)\right| \leqslant C$.

В отличие от одномерного в кратном случае утверждается лишь неравенство — приращение отображения оценивается сверху через приращение аргумента. Нам этого вполне достаточно, поскольку с помощью этого результата оказывается возможным получить и оценку снизу (правда лишь локально — вблизи не критических точек). Такая оценка гарантирует локальную обратимость гладкого отображения и непрерывность обратного отображения.

Теорема (оценка снизу приращения гладкого отображения). *Пусть E — открытое множество в $\mathbb{R}^m$ и $F : E \to \mathbb{R}^m$ — непрерывно дифференцируемое отображение. Тогда для любой не критической точки x_0 из E, существуют такие положительные числа r и μ, что $B_r(x_0) \subset E$ и*

$$\|F(x'') - F(x')\|_m \geqslant \mu\|x'' - x'\|_m \quad \text{для всех} \;\; x'', x' \;\; \text{из} \;\; B_r(x_0).$$

Доказательство. По условию линейное отображение $L = d_{x_0}F$ — изоморфизм. Следовательно, существует такое положительное число μ, что

$$\|L(h)\|_m \geqslant 2\mu\,\|h\|_m \quad \text{для всех} \;\; h \;\; \text{из} \;\; \mathbb{R}^m.$$

Поскольку отображение F гладкое, можно подобрать такой радиус r, что $B_r(x_0) \subset E$ и

$$\|d_x F - L\|_{\mathscr{L}} < \mu \quad \text{для всех} \;\; x \;\; \text{из} \;\; B_r(x_0).$$

Проверим, что числа r и μ требуемые. Для этого определим в шаре $B_r(x_0)$ отображение Φ, действующее по формуле

$$\Phi(x) = F(x) - L(x).$$

Тогда

$$\Phi(x'') - \Phi(x') = F(x'') - F(x') - L(x'' - x').$$

Для оценки приращения отображения Φ заметим, что

$$d_x\Phi = d_x F - d_x L = d_x F - L,$$

и следовательно, $\|d_x\Phi\|_{\mathscr{L}} < \mu$ для всех x из шара $B_r(x_0)$. Поэтому неравенство Лагранжа даёт нам оценку $\|\Phi(x'') - \Phi(x')\|_m \leqslant \mu\|x'' - x'\|_m$. Итак,

$$\begin{aligned}\|F(x'') - F(x')\|_m &= \|L(x'') - L(x') + \Phi(x'') - \Phi(x')\|_m \geqslant \\ &\geqslant \|L(x'') - L(x')\|_m - \|\Phi(x'') - \Phi(x')\|_m \geqslant \\ &\geqslant 2\mu\|x'' - x'\|_m - \mu\|x'' - x'\|_m = \mu\|x'' - x'\|_m.\end{aligned}$$

Теорема доказана.

Следствие. *В условиях теоремы отображение F взаимно однозначно на шаре $B_r(x_0)$, а обратное отображение $G = F^{-1}$ удовлетворяет условию Липшица с коэффициентом $\frac{1}{\mu}$.*

Доказательство. Инъективность сужения F на шар $B_r(x_0)$ вытекает из неравенства $\|F(x'') - F(x')\|_m \geqslant \mu\|x'' - x'\|_m$. Из него же следует, что $\|y'' - y'\|_m \geqslant \mu\|G(y'') - G(y')\|_m$ для любых y'' и y' из $F(B_r(x_0))$, т.е. условие Липшица для отображения G.

Для функции $F(x) = x^3$ из примера 2 оценки, установленной в теореме, около критической точки $x_0 = 0$, очевидно, нет: разность $|F(x) - F(0)| = |x|^3$ невозможно оценить снизу величиной $\mathrm{const}\,|x|$ при малых x. Как показывает следующее упражнение, это верно и в общем случае.

Упражнение. Если в условиях теоремы x_0 — критическая точка, то утверждение теоремы заведомо неверно. Поэтому обратное отображение, если оно существует, не удовлетворяет условию Липшица.

Класс непрерывных отображений F, у которых существует и непрерывно обратное отображение F^{-1} часто возникает в различных задачах. Для таких отображений нет необходимости требовать совпадения размерностей пространств, в которых заданы F и F^{-1}.

Определение. *Гомеоморфизмом* двух множеств E, $E \subset \mathbb{R}^m$, и $\widetilde{E}$, $\widetilde{E} \subset \mathbb{R}^M$, называется такая биекция $F : E \to \widetilde{E}$, что оба отображения F и F^{-1} непрерывны во всех точках множеств E и $\widetilde{E}$ соответственно.

Функция $F(x) = x^3$ из примера 2 — гомеоморфизм, но не диффеоморфизм прямой.

Очевидно, суперпозиция гомеоморфизмов также гомеоморфизм.

Теорема (о диффеоморфизме). *Пусть F — гомеоморфизм между открытыми множествами E и $\widetilde{E}$ из $\mathbb{R}^m$, удовлетворяющий условию: F — гладкое отображение, не имеющее критических точек. Тогда F^{-1} — гладкое на $\widetilde{E}$ отображение и, следовательно, F — диффеоморфизм.*

Доказательство. Докажем сначала дифференцируемость отображения $G = F^{-1}$ в произвольной точке y_0 из $\widetilde{E}$. Так как $\widetilde{E} = F(E)$, то $y_0 = F(x_0)$ для некоторой точки x_0 из E. Пусть $L = d_{x_0}F$. По условию $L \in \mathrm{Isom}\,(\mathbb{R}^m)$.

Докажем, что $d_{y_0}G = L^{-1}$ (тот факт, что никакое другое линейное отображение не может быть дифференциалом G следует из леммы о дифференцировании обратного отображения; но в этой лемме *предполагается* дифференцируемость G, а нам нужно её *доказать*).

В силу оценки снизу найдутся такие положительные числа r и μ, что

$$B_r(x_0) \subset E \quad \text{и} \quad \|F(x) - F(x_0)\|_m \geqslant \mu\|x - x_0\|_m \quad \text{для всех} \quad x \quad \text{из шара} \quad B_r(x_0).$$

Поскольку отображение G непрерывно, найдётся такой радиус ρ, что

$$B_\rho(y_0) \subset \widetilde{E} \quad \text{и} \quad G(y) \in B_r(x_0) \quad \text{при} \quad y \in B_\rho(y_0).$$

Тогда оценка снизу для F даёт оценку сверху для G:

$$\|G(y) - G(y_0)\|_m \leqslant \frac{1}{\mu}\|y - y_0\|_m, \quad \text{если} \quad \|y - y_0\|_m < \rho.$$

Так как L — дифференциал F в точке x_0, то существует такое отображение $\alpha : E \to \mathbb{R}^m$, что $\lim_{x_0} \|\alpha\|_m = 0$ и

$$F(x) - F(x_0) = L(x - x_0) + \|x - x_0\|_m\,\alpha(x) \quad \text{для всех} \quad x \in E$$

(при этом мы можем и будем считать, что $\alpha(x_0) = \mathbb{O}_m$, т.е. отображение α непрерывно в точке x_0). Возьмём в этом равенстве $x = G(y)$, где y — произвольная точка шара $B_\rho(y_0)$. Мы получим

$$y - y_0 = L\left(G(y) - G(y_0)\right) + \|G(y) - G(y_0)\|_m \, \alpha\big(G(y)\big).$$

Следовательно,

$$L^{-1}(y - y_0) = G(y) - G(y_0) + \|G(y) - G(y_0)\|_m \, L^{-1}\big(\alpha\left(G(y)\right)\big),$$

т.е.

$$G(y) = G(y_0) + L^{-1}(y - y_0) - \|G(y) - G(y_0)\|_m \, L^{-1}\left(\alpha\left(G(y)\right)\right).$$

Чтобы убедиться в том, что L^{-1} — действительно дифференциал G в точке y_0, нам осталось проверить соотношение

$$\|G(y) - G(y_0)\|_m \, L^{-1}\left(\alpha\left(G(y)\right)\right) = o\left(\|y - y_0\|_m\right) \quad \text{при} \quad y \to y_0.$$

Но оно очевидно, так как $\|G(y) - G(y_0)\|_m \leqslant \frac{1}{\mu}\,\|y - y_0\|_m$ и при $y \to y_0$

$$L^{-1}\left(\alpha\left(G(y)\right)\right) \to L^{-1}\left(\alpha\big(G(y_0)\big)\right) = L^{-1}\left(\alpha(x_0)\right) = L^{-1}\left(\mathbb{O}_m\right) = \mathbb{O}_m,$$

поскольку все три отображения L^{-1}, α и G непрерывны.

Итак, дифференцируемость отображения G и равенство $d_{y_0}G = (dF_{x_0})^{-1}$ установлены. Для завершения доказательства теоремы надо проверить гладкость обратного отображения. Для этого докажем непрерывность каждой частной производной $\frac{\partial G_k}{\partial y_j}$ на множестве $\widetilde{E}$. Пусть y — произвольная точка из $\widetilde{E}$ и $x = G(y)$. Тогда $x \in E$, $y = F(x)$ и $G'(y) = (F'(x))^{-1}$. Матрица Якоби $G'(y)$ образована частными производными $\frac{\partial G_k}{\partial y_j}(y)$. Так как она совпадает с матрицей $(F'(x))^{-1}$, то пользуясь выражением элементов обратной матрицы через определитель и алгебраические дополнения обращаемой матрицы, мы получаем (для всех j, k)

$$\frac{\partial G_k}{\partial y_j}(y) = \frac{1}{\det F'(x)}\, A_{jk}\left(F'(x)\right) = \frac{1}{\det F'(G(y))}\, A_{jk}\left(F'(G(y))\right).$$

Символ A_{kj} обозначает алгебраическое дополнение элемента $\frac{\partial F_k}{\partial x_j}(x)$ матрицы Якоби $F'(x)$ (напомним, что при построении обратной матрицы матрица, образованная алгебраическими дополнениями, транспонируется). Все частные производные координатных функций отображения F непрерывны на E, а поэтому там непрерывны определители $\det F'$ и $A_{jk}\left(F'\right)$. Следовательно, на $\widetilde{E}$ непрерывны суперпозиции $\det F'(G)$ и $A_{jk}\left(F'(G)\right)$. Первая из них не обращается в нуль, так как всюду $\det F' \neq 0$ (отображение F не имеет критических точек). Поэтому производная $\frac{\partial G_k}{\partial y_j}$, равная отношению этих суперпозиций, непрерывна на $\widetilde{E}$. Теорема доказана.

Вернёмся к основной теме этого параграфа — получению результата, позволяющего судить о существовании и гладкости обратного отображения, исходя лишь из информации об отображении F. Нам осталось избавиться от последнего мешающего предположения — множество $F(E)$ открыто. Какие свойства отображения F гарантируют это? Поскольку мы можем надеяться на получение лишь локального результата, этот вопрос следует уточнить: какие свойства отображения F гарантируют, что его значение в произвольной точке x_0 окажется внутри множества $F\big(B_r(x_0)\big)$? Например, не может ли случиться так, что шар $B_r(x_0)$ переводится отображением F в какой-то другой шар, но при этом точка $F(x_0)$ оказалась на его границе?

В одномерном случае всё просто: у функции F, непрерывной на интервале I вещественной оси, лишь её наибольшее и наименьшее значения (если они существуют) лежат на границе множества значений $F(I)$. Поэтому для того, чтобы значение $F(x_0)$ в произвольной точке x_0, $x_0 \in I$, оказалось внутри множества $F\big((x_0 - r, x_0 + r)\big)$, нужно, чтобы F не имела локальных экстремумов на I, а это равносильно строгой монотонности F. Если функция F дифференцируема, то для этого достаточно, чтобы производная F' нигде не обращалась в нуль. В кратном случае нет понятия монотонности, но аналог неравенства $F' \neq 0$ есть — это обратимость dF. Можно надеяться, что при отсутствии критических точек значение $F(x_0)$ лежит внутри $F\big(B_r(x_0)\big)$ для всех $r > 0$. Прежде чем убедиться в этом, введём и коротко обсудим одно полезное понятие, непосредственно связанное с обсуждаемыми вопросами.

Определение. Отображение $F : E \to \mathbb{R}^m$, заданное на открытом множестве E, $E \subset \mathbb{R}^m$, называется ***открытым***, если образ $F(E_0)$ любого открытого подмножества $E_0 \subset E$ открыт.

Не следует путать это определение с определением непрерывности отображения, заданного на открытом множестве (там говорится от открытости *прообраза*).

Легко видеть, что открытость отображения равносильна интересующему нас свойству — всякое значение $F(x_0)$ лежит *внутри* множества $F\big(B_r(x_0)\big)$ (для любого радиуса r, лишь бы $B_r(x_0) \subset E$).

Очевидно, гомеоморфизм между двумя открытыми подмножествами пространства $\mathbb{R}^m$ — это открытое отображение. В частности, таково любое отображение из $\mathrm{Isom}\,(\mathbb{R}^m)$, а при $m = 1$ — любая непрерывная и строго монотонная на интервале функция. Других одномерных открытых отображений нет, так что x^2, $\sin x$, $\frac{1}{1+x^2}$ и т.п. не открыты на прямой. В кратном случае содержательные примеры можно получить с помощью следующего утверждения.

Теорема (об открытом отображении). *Пусть E — открытое множество в $\mathbb{R}^m$ и $F : E \to \mathbb{R}^m$ — гладкое отображение, не имеющее критических точек. Тогда F — открытое отображение.*

Доказательство. Достаточно проверить, что для любого шара $B = B_r(a) \subset E$ точка $b = F(a)$ лежит внутри $F(B)$. При этом мы можем считать, что замкнутый шар $\overline{B} = \overline{B}_r(a)$ содержится в E. Уменьшив радиус в случае необходимости, мы можем, опираясь на теорему об оценке снизу приращения гладкого отображения, считать также радиус r столь малым, что для некоторого положительного числа μ справедливо неравенство

$$\|F(x) - F(a)\|_m \geqslant \mu\, \|x - a\|_m \qquad (x \in \overline{B}).$$

Проверим теперь, что шар $B_\rho(b)$ радиуса $\rho = \frac{1}{2}\mu r$ содержится в $F(B)$:

для любой точки y_0, $\|y_0 - b\|_m < \rho$, существует такая точка $x_0 \in B$, что $y_0 = F(x_0)$.

Для этого зафиксируем вектор $y_0 = (y_1^\circ, \dots, y_m^\circ)$ до конца рассуждений и определим на E вспомогательную функцию:

$$\psi(x) = \|F(x) - y_0\|_m^2 = \sum_{k=1}^{m} (F_k(x) - y_k^\circ)^2.$$

Это непрерывно дифференцируемая на E функция. Найдём $\min_{\overline{B}} \psi$. По теореме Вейерштрасса этот минимум достигается в некоторой точке x_0 шара $\overline{B}$. Убедимся в том, что она лежит внутри шара $\overline{B}$. Для этого докажем, что на его границе значения функции ψ слишком большие (больше, чем в центре шара).

По выбору точки y_0 мы имеем

$$\psi(a) = \|F(a) - y_0\|_m^2 = \|b - y_0\|_m^2 < \rho^2.$$

Если же $x \in \partial\overline{B}$, т.е. $\|x-a\|_m = r$, то

$$\sqrt{\psi(x)} = \|F(x)-y_0\|_m = \|F(x)-F(a)+b-y_0\|_m \geqslant \|F(x)-F(a)\|_m - \|b-y_0\|_m > \\ > \mu\|x-a\|_m - \rho = \mu r - \rho = \rho.$$

Итак, $\psi(a) < \rho^2 < \psi(x)$ для любой граничной точки x шара $\overline{B}$. Поэтому минимум функции ψ достигается внутри него, т.е. $x_0 \in B$. Но тогда все частные производные функции ψ в точке x_0 равны нулю:

$$0 = \frac{\partial\psi}{\partial x_j}(x_0) = 2\sum_{k=1}^{m}\big(F_k(x_0) - y_k^\circ\big)\frac{\partial F_k}{\partial x_j}(x_0) \qquad (j = 1,\ldots,m).$$

Таким образом, числа $F_k(x_0) - y_k^\circ$ $(k = 1,\ldots,m)$ удовлетворяют однородной системе линейных уравнений с неособенной матрицей $F'(x_0) = \Big(\frac{\partial F_k}{\partial x_j}(x_0)\Big)_{1\leqslant j,k\leqslant m}$ (напомним, что $\det F' \neq 0$ на E). Следовательно, все эти числа равны нулю, т.е. $y_0 = F(x_0) \in F(B)$. Поскольку y_0 — произвольная точка из шара $B_\rho(b)$, мы получаем, что весь он содержится в $F(B)$. Теорема доказана.

Теперь у нас всё готово для получения основного результата.

Теорема (об обратном отображении). *Пусть E — открытое подмножество пространства $\mathbb{R}^m$ и $F : E \to \mathbb{R}^m$ — гладкое отображение, не имеющее критических точек. Тогда F — локальный диффеоморфизм, т.е. для любой точки x_0 из E найдётся такой шар $B = B_r(x_0)$, $B \subset E$, что множество $F(B)$ открыто в $\mathbb{R}^m$, а сужение $F|_B$ — диффеоморфизм между B и $F(B)$.*

В частности, обратное отображение $G := \left(\,F\,|_B\right)^{-1}$ непрерывно дифференцируемо на $F(B)$.

Доказательство. По следствию из теоремы об оценке снизу приращения гладкого отображения существует такой шар $B = B_r(x_0)$, $B \subset E$, что отображение F инъективно на B и обратное отображение $G = \left(\,F\,|_B\right)^{-1}$ непрерывно на $F(B)$, т.е. $F|_B$ — гомеоморфизм. По теореме об открытом отображении множество $F(B)$ открыто. Поэтому осталось сослаться на теорему о диффеоморфизме (заменив в ней множество E шаром B). Теорема доказана.

§3. Относительный экстремум функции нескольких переменных

Постановка задачи. На практике часто возникает такая задача: на открытом множестве E, $E \subset \mathbb{R}^m$, задана гладкая функция $f : E \to \mathbb{R}$. Требуется найти её экстремальные значения, но не на всём множестве E, а лишь на его подмножестве, образованном теми точками, которые удовлетворяют дополнительным условиям (уравнениям связи):

$$\varphi_1(x) = 0, \ldots, \varphi_n(x) = 0.$$

Например, если надо найти $\max_K f$ на компакте K, $K \subset E$, то приходится рассматривать две возможности: точка x^*, в которой f достигает наибольшего значения (она существует по теореме Вейерштрасса), находится *внутри* или *на границе* компакта K. В первом случае всё сравнительно просто — тогда

$$\operatorname{grad} f(x^*) = \mathbb{O}_m, \qquad \text{т.е.}\quad \frac{\partial f}{\partial x_1}(x^*) = \ldots = \frac{\partial f}{\partial x_m}(x^*) = 0.$$

Остаётся решить эту систему, в которой число уравнений совпадает с числом неизвестных.

Во втором случае, когда $x^* \in \partial K$, равенство $\operatorname{grad} f(x^*) = \mathbb{O}_m$, вообще говоря, не выполняется. Поскольку при $m > 1$ множество ∂K содержит бесконечно много точек, желательно найти признак, выделяющий среди них точки “подозрительные на экстремум” (в одномерном случае в этом нет необходимости, так как граница отрезка состоит лишь из двух точек; поэтому достаточно сравнить значения функции f на концах отрезка с её значениями во всех внутренних стационарных точках).

Чтобы дать некоторое представление о том, что могло бы послужить этим признаком в кратном случае, рассмотрим такую задачу. На карте холмистой местности отмечена дорога. Как найти её самую высокую и самую низкую точки? Иначе говоря, на открытом множестве E, $E \subset \mathbb{R}^2$, задана функция f (значение $f(x)$ равно высоте земной поверхности в точке x, $x \in E$). Кроме того, в E лежит гладкая кривая ℓ (дорога). Требуется найти

$$h^* = \max_{x\in\ell} f(x) \qquad \text{и} \qquad h_* = \min_{x\in\ell} f(x).$$

Нарисовав на карте линию постоянной высоты h, $h > h^*$ или $h < h_*$, мы получим, что она не имеет общих точек с кривой ℓ. Линия, соответствующая промежуточной высоте $h \in (h_*, h^*)$, пересекается с ℓ. Если параметр h приближать к h^*, то точки пересечения будут сближаться, и в предельном случае $h = h^*$ линия $\{x \in E \,|\, f(x) = h^*\}$ коснётся кривой ℓ в искомой точке x^*. Таким образом, касательная к линии $\{x \in E \,|\, f(x) = h^*\}$, т.е. прямая $\langle \operatorname{grad} f(x^*), x - x^*\rangle = 0$ является одновременно и касательной к кривой ℓ в точке x^*. Если эта кривая задаётся уравнением $\varphi(x) = 0$, то $\langle \operatorname{grad} \varphi(x^*), x - x^*\rangle = 0$ — уравнение касательной к ℓ в точке x^*. Совпадение касательных равносильно пропорциональности градиентов $\operatorname{grad} f(x^*)$ и $\operatorname{grad} \varphi(x^*)$: существует такое число λ, что $\operatorname{grad} f(x^*) = \lambda \operatorname{grad} \varphi(x^*)$, т.е.

$$\frac{\partial f}{\partial x_1}(x^*) = \lambda \frac{\partial \varphi}{\partial x_1}(x^*) \quad \text{и} \quad \frac{\partial f}{\partial x_2}(x^*) = \lambda \frac{\partial \varphi}{\partial x_2}(x^*).$$

Это и есть искомый критерий в рассматриваемой двумерной задаче: ценой введения новой неизвестной λ мы нашли два дополнительных уравнения, которые вместе с уравнением $\varphi(x_1, x_2) = 0$, связывающим переменные x_1 и x_2, образуют систему из трёх уравнений с тремя неизвестными.

Аналогичное явление, возникающее при рассмотрении более общей задачи, описывается следующей теоремой.

Теорема (необходимое условие относительного экстремума). *Пусть функция $f : E \to \mathbb{R}$ непрерывно дифференцируема на открытом множестве $E, E \subset \mathbb{R}^m$. Пусть ещё $\varphi = (\varphi_1, \dots, \varphi_n) : E \to \mathbb{R}^n$, $n < m$, — непрерывно дифференцируемое отображение и $E_\varphi = \{x \in E \,|\, \varphi(x) = \mathbb{O}_n\}$. Допустим, что $x_0 \in E_\varphi$ — точка локального экстремума f на E_φ и ранг матрицы Якоби $\varphi'(x_0)$ максимально возможный: $\operatorname{rank} \varphi'(x_0) = n$.*

Тогда градиент $\operatorname{grad} f(x_0)$ линейно зависит от градиентов $\operatorname{grad} \varphi_1(x_0), \dots, \operatorname{grad} \varphi_n(x_0)$.

Условие $\operatorname{rank} \varphi'(x_0) = n$ означает, что строки матрицы $\varphi'(x_0)$, т.е. градиенты $\operatorname{grad} \varphi_1(x_0), \dots, \operatorname{grad} \varphi_n(x_0)$, линейно независимы. Теорема утверждает, что для некоторых чисел $\lambda_1, \dots, \lambda_n$ выполняется равенство

$$\operatorname{grad} f(x_0) = \sum_{k=1}^{n} \lambda_k \operatorname{grad} \varphi_k(x_0).$$

Это система m уравнений, в которой $m + n$ неизвестных: к координатам $x_1^\circ, \dots, x_m^\circ$ искомого вектора x_0 добавились новые неизвестные коэффициенты $\lambda_1, \dots, \lambda_n$. Их называют множителями Лагранжа. Кроме этих m уравнений, у нас есть ещё n *уравнений связи*

$$\begin{cases} \varphi_1(x_1^\circ, \dots, x_m^\circ) = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \varphi_n(x_1^\circ, \dots, x_m^\circ) = 0. \end{cases}$$

Итак, для поиска точки относительного локального экстремума функции m переменных, который обусловлен n уравнениями связи, надо решить систему из $m+n$ уравнений с таким же числом неизвестных.

Доказательство. Не умаляя общности, будем считать $f(x_0)=0$ (этого можно добиться, заменив функцию f функцией $f-f(x_0)$). Допустим, что утверждение теоремы не выполнено: $\operatorname{grad} f(x_0)$ линейно не зависит от $\operatorname{grad}\varphi_1(x_0),\dots,\operatorname{grad}\varphi_n(x_0)$. Убедимся в том, что в этом случае точка x_0 не может быть точкой локального экстремума функции f на множестве E_φ.

Сначала разберём основной случай с максимальным числом уравнений связи: $n=m-1$. Введём новое отображение $F : E \to \mathbb{R}^m$, действующее на вектор x из E по правилу $F(x) = \big(f(x),\varphi(x)\big) = \big(f(x),\varphi_1(x),\dots,\varphi_{m-1}(x)\big)$. Тогда $F(x_0)=\big(f(x_0),\varphi(x_0)\big)=\big(0,\mathbb{O}_{m-1}\big)=\mathbb{O}_m$ и

$$F'(x_0)=\begin{Vmatrix} \operatorname{grad} f(x_0) \\ \operatorname{grad}\varphi_1(x_0) \\ \dots\dots\dots\dots\dots \\ \operatorname{grad}\varphi_{m-1}(x_0)\end{Vmatrix} = \begin{Vmatrix} \frac{\partial f}{\partial x_1}(x_0) & \dots & \frac{\partial f}{\partial x_m}(x_0) \\ \frac{\partial \varphi_1}{\partial x_1}(x_0) & \dots & \frac{\partial \varphi_1}{\partial x_m}(x_0) \\ \dots & \dots & \dots \\ \frac{\partial \varphi_{m-1}}{\partial x_1}(x_0) & \dots & \frac{\partial \varphi_{m-1}}{\partial x_m}(x_0)\end{Vmatrix}.$$

Это квадратная $m\times m$ матрица. По сделанному предположению её первая строка линейно не зависит от остальных строк, которые в свою очередь линейно независимы по условию теоремы. Следовательно, $\det F'(x_0)\neq 0$. Так как функция $x\mapsto \det F'(x)$ непрерывна на E, то $\det F'(x)\neq 0$, т.е. $d_xF\in \operatorname{Isom}(\mathbb{R}^m)$ вблизи x_0. По теореме об открытом отображении для достаточно малого шара B с центром в точке x_0 его образ $F(B)$ — открытое множество в $\mathbb{R}^m$. Поскольку $\mathbb{O}_m=F(x_0)\in F(B)$, мы получаем, что $\mathbb{O}_m$ — внутренняя точка множества $F(B)$. Следовательно, при достаточно малом положительном ε векторы $\big(\varepsilon,\mathbb{O}_{m-1}\big)$ и $\big(-\varepsilon,\mathbb{O}_{m-1}\big)$ являются значениями отображения F на шаре B: существуют такие точки $x_+,x_-\in B$, что $F(x_+)=\big(\varepsilon,\mathbb{O}_{m-1}\big)$ и $F(x_-)=\big(-\varepsilon,\mathbb{O}_{m-1}\big)$. По определению отображения F это означает, что $\big(f(x_\pm),\varphi(x_\pm)\big)=\big(\pm\varepsilon,\mathbb{O}_{m-1}\big)$, т.е. $f(x_\pm)=\pm\varepsilon$ и $\varphi(x_\pm)=\mathbb{O}_{m-1}$. Второе равенство означает, что $x_\pm\in E_\varphi$. Таким образом, функция f принимает как положительные, так и отрицательные значения на множестве E_φ, причём это происходит сколь угодно близко к x_0 (точки $x_\pm$ принадлежат шару B, радиус которого можно взять сколь угодно малым). Поэтому сужение функции f на множество E_φ не имеет локального экстремума в точке x_0. Итак, допущение линейной независимости градиента $\operatorname{grad} f(x_0)$ от $\operatorname{grad}\varphi_1(x_0),\dots,\operatorname{grad}\varphi_{m-1}(x_0)$ приводит к противоречию.

Пусть теперь $n\leqslant m-2$. Сведём этот случай к уже рассмотренной нами ситуации с максимальным числом уравнений связи. В силу условия (градиенты функций $\varphi_1,\dots,\varphi_n$ в точке x_0 линейно независимы) и сделанного предположения ($\operatorname{grad} f(x_0)$ линейно независим от них) прямоугольная матрица

$$\begin{Vmatrix} \operatorname{grad} f(f_0) \\ \operatorname{grad}\varphi_1(x_0) \\ \dots \\ \operatorname{grad}\varphi_n(x_0)\end{Vmatrix} = \begin{Vmatrix} \frac{\partial f}{\partial x_1}(x_0) & \dots & \frac{\partial f}{\partial x_m}(x_0) \\ \frac{\partial \varphi_1}{\partial x_1}(x_0) & \dots & \frac{\partial \varphi_1}{\partial x_m}(x_0) \\ \dots & \dots & \dots \\ \frac{\partial \varphi_n}{\partial x_1}(x_0) & \dots & \frac{\partial \varphi_n}{\partial x_m}(x_0)\end{Vmatrix}$$

имеет максимально возможный ранг $n+1$. Не умаляя общности, будем считать, что линейно независимы её первые столбцы (в противном случае следует изменить нумерацию переменных), и поэтому квадратная матрица

$$A=\begin{Vmatrix} \frac{\partial f}{\partial x_1}(x_0) & \dots & \frac{\partial f}{\partial x_{n+1}}(x_0) \\ \frac{\partial \varphi_1}{\partial x_1}(x_0) & \dots & \frac{\partial \varphi_1}{\partial x_{n+1}}(x_0) \\ \dots & \dots & \dots \\ \frac{\partial \varphi_n}{\partial x_1}(x_0) & \dots & \frac{\partial \varphi_n}{\partial x_{n+1}}(x_0)\end{Vmatrix}$$

неособенная: $\det A\neq 0$.

Чтобы свести ситуацию к уже рассмотренной, увеличим число уравнений связи до $m-1$, добавив ещё $m-n-1$ ограничение: $\varphi_j(x)=0$, где $\varphi_j(x)=x_{j+1}-x_{j+1}^{\circ}$ при $n+1\leqslant j<m$. Таким образом, мы заменяем множество E_φ меньшим множеством $\widetilde{E}_\varphi$, рассматривая в E_φ только те точки, у которых последние $m-n-1$ координат зафиксированы и совпадают с соответствующими координатами точки x_0 ("лишние координаты заморожены"). Ясно, что $x_0\in\widetilde{E}_\varphi$ и в этой точке функция f имеет локальный экстремум на $\widetilde{E}_\varphi$ (поскольку она имеет его на большем множестве E_φ).

Теперь уравнения связи задаются равенством $\widetilde{\varphi}(x)=\mathbb{O}_{m-1}$, где $\widetilde{\varphi}: E\to\mathbb{R}^{m-1}$ — отображение с координатными функциями $\varphi_1,\ldots,\varphi_{m-1}$. Проверим, что градиенты $\operatorname{grad} f(x_0), \operatorname{grad}\varphi_1(x_0), \ldots, \operatorname{grad}\varphi_{m-1}(x_0)$ линейно независимы. Это равносильно тому, что квадратная матрица $\widetilde{A}$, образованная такими строками, неособенная. Ясно, что

$$\widetilde{A}=\begin{Vmatrix}\operatorname{grad} f(x_0)\\ \operatorname{grad}\varphi_1(x_0)\\ \ldots\ldots\ldots\ldots\\ \operatorname{grad}\varphi_{m-1}(x_0)\end{Vmatrix}=\begin{Vmatrix}\frac{\partial f}{\partial x_1}(x_0) & \ldots & \frac{\partial f}{\partial x_m}(x_0)\\ \frac{\partial \varphi_1}{\partial x_1}(x_0) & \ldots & \frac{\partial \varphi_1}{\partial x_m}(x_0)\\ \ldots & \ldots & \ldots\\ \frac{\partial \varphi_{m-1}}{\partial x_1}(x_0) & \ldots & \frac{\partial \varphi_{m-1}}{\partial x_m}(x_0)\end{Vmatrix}=\begin{Vmatrix}A & C\\ \mathbb{O} & \mathbb{I}\end{Vmatrix}.$$

Здесь $\mathbb{O}$ — нулевая матрица с $m-n-1$ строкой и $n+1$ столбцом, прямоугольная матрица C образована частными производными по $x_{n+2},\ldots,x_m$ функций $f,\varphi_1,\ldots,\varphi_n$, а $\mathbb{I}$ — единичная квадратная матрица размера $(m-n-1)\times(m-n-1)$. Поскольку $\det\widetilde{A}=\det A\neq 0$, вектор $\operatorname{grad} f(x_0)$ линейно не зависит от векторов $\operatorname{grad}\varphi_1(x_0),\ldots,\operatorname{grad}\varphi_{m-1}(x_0)$. Но тогда, как уже доказано, функция f не имеет в точке x_0 локального экстремума на множестве $\widetilde{E}_\varphi$ (и тем более на E_φ). Теорема доказана.

Пример. Дадим аналитическое доказательство полученного другим способом в курсе высшей алгебры важного утверждения: всякая квадратная вещественная симметричная матрица имеет вещественное собственное число.

Для этого матрице $A=(a_{jk})_{1\leqslant j,k\leqslant m}$, $a_{jk}=a_{kj}$, сопоставим квадратичную форму

$$Q(x)=x\,A\,x^T=\sum_{1\leqslant j,k\leqslant m}a_{j,k}x_jx_k \qquad (x\in\mathbb{R}^m).$$

По теореме Вейерштрасса она достигает наибольшего значения на единичной сфере $S^{m-1}=\{x\in\mathbb{R}^m \mid \|x\|_m=1\}$, т.е. существует решение задачи

$$\textit{найти}\quad \max Q(x)\quad \textit{при условии}\quad x_1^2+\ldots+x_m^2=1.$$

Здесь только одно уравнение связи $\varphi(x)=x_1^2+\ldots+x_m^2-1=0$ и предположение теоремы выполнено: $\operatorname{rank}\varphi'(x)=1$ всюду на S^{m-1}, так как $\operatorname{grad}\varphi(x)=(2x_1,\ldots,2x_m)=2x\neq\mathbb{O}_m$. По доказанной теореме в той точке $x_0=(x_1^{\circ},\ldots,x_m^{\circ})$, где квадратичная форма Q принимает наибольшее на S^{m-1} значение, для некоторого вещественного числа λ выполняется равенство $\operatorname{grad} Q(x_0)=\lambda\operatorname{grad}\varphi(x_0)$, т.е. $2\lambda\,x_j^{\circ}=\frac{\partial Q}{\partial x_j}(x_0)$ при $j=1,\ldots,m$. Так как

$$Q(x)=a_{11}x_1^2+\ldots+a_{mm}x_m^2+2\bigl(a_{12}x_1x_2+\ldots+a_{m-1,m}x_{m-1}x_m\bigr),$$

то мы получаем, что

$$2\lambda x_j^{\circ}=\frac{\partial Q}{\partial x_j}(x_0)=2\sum_{k=1}^m a_{jk}x_k^{\circ}\qquad (j=1,\ldots,m),\qquad \text{т.е.}\quad 2\lambda\,x_0^T=2A\,x_0^T.$$

Поэтому λ — собственное число матрицы A.

Домножив равенство $A\,x_0^T = \lambda\,x_0^T$ слева на x_0 мы видим, что $Q(x_0) = x_0\,A\,x_0^T = \lambda x_0\,x_0^T = \lambda\|x_0\|_m^2 = \lambda$, т.е. на единичной сфере наибольшее значение квадратичной формы достигается на собственном векторе и равно собственному числу.

§4. **Неявное отображение**

Здесь мы займёмся более сложной, чем в первом параграфе, задачей — исследованием системы нескольких уравнений. Рассмотрим систему из n уравнений относительно такого же числа неизвестных $y_1, \dots, y_n$

$$\begin{cases} F_1(x_1, \dots, x_m, y_1, \dots, y_n) = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ F_n(x_1, \dots, x_m, y_1, \dots, y_n) = 0 \end{cases}$$

с вещественными параметрами $x_1, \dots, x_m$.

В случае $n = 1$ мы возвращаемся к задаче о неявной функции. Обращение отображения (см. §2) — также частный случай решения системы уравнений, в которой $m = n$, а функции $F_1, \dots, F_m$ имеют специальный вид:

$$F_j(x_1, \dots, x_m, y_1, \dots, y_n) = \widetilde{F}_j(y_1, \dots, y_m) - x_j \qquad \text{для } \ j = 1, \dots, m.$$

В этом случае уравнения системы сводятся к уравнениям $\widetilde{F}_j(y_1, \dots, y_m) = x_j$, и мы приходим к задаче обращения отображения $\widetilde{F}$: зная вектор $x = (x_1, \dots, x_m)$, надо найти такой вектор $y = (y_1, \dots, y_m)$, что $\widetilde{F}(y) = x$ (по сравнению с §2 x и y поменялись местами).

Задача об обращении отображения имеет лишь локальное решение. Поэтому далее, при решении более общей задачи о неявном отображении, речь пойдёт о получении только локальных результатов: зная, что при каком-то наборе параметров система уравнений имеет решение, можно ли заключить, что она имеет решение и при малом изменении этих параметров? Можно ли это решение дифференцировать по параметрам?

Хотя изучение неявной функции укладывается в новую, более общую схему, усилия, затраченные в предыдущем параграфе, не напрасны — использование монотонности очень наглядно и даёт не только локальный результат (см. теорему о гладкости неявной функции).

При обсуждении этих вопросов целесообразно отказаться от школьных обозначений и использовать более компактную векторную запись. Тогда наша задача выглядит так: имеется отображение F множества E, $E \subset \mathbb{R}^{m+n}$, в пространство $\mathbb{R}^n$. Известно, что $F(x_0, y_0) = \mathbb{O}_n$ при некоторых x_0 из $\mathbb{R}^m$ и y_0 из $\mathbb{R}^n$. Верно ли, что для любого вектора x из $\mathbb{R}^m$ достаточно близкого к x_0 найдётся такой вектор y из $\mathbb{R}^n$ (зависящий от x), что $(x, y) \in E$ и $F(x, y) = \mathbb{O}_n$? Если это действительно так, то вблизи точки x_0 определено отображение G, сопоставляющее параметру x вектор $y = G(x)$ — решение уравнения $F(x, y) = \mathbb{O}_n$, т.е. $F(x, G(x)) = \mathbb{O}_n$. В этом случае говорят, что вблизи точки x_0 *неявное отображение* G задаётся уравнением $F(x, y) = \mathbb{O}_n$. Наше цель — получить достаточное условие существования неявного отображения и исследовать его свойства (непрерывность и дифференцируемость).

Как и в частном случае этой задачи (обращение дифференцируемого отображения) основная идея решения прежняя — это *линеаризация.* Надо воспользоваться тем, что локально дифференцируемое отображение “почти линейно”. Посмотрим сначала, не следя за строгостью рассуждений, как трансформируется наша задача, если заменить значение $F(x, y)$ более простым

$$\widetilde{F}(x, y) = F(x_0, y_0) + \big(d_{(x_0, y_0)} F\big)(x - x_0, y - y_0).$$

Матрица Якоби $F'(x_0, y_0)$ образована частными производными координатных функций F_k, $k = 1, \dots, n$, по переменным $x_1, \dots, x_m$ и $y_1, \dots, y_n$:

$$F'(x_0, y_0) = \begin{Vmatrix} \frac{\partial F_1}{\partial x_1}(x_0, y_0) & \dots & \frac{\partial F_1}{\partial x_m}(x_0, y_0) & \frac{\partial F_1}{\partial y_1}(x_0, y_0) & \dots & \frac{\partial F_1}{\partial y_n}(x_0, y_0) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial F_n}{\partial x_1}(x_0, y_0) & \dots & \frac{\partial F_n}{\partial x_m}(x_0, y_0) & \frac{\partial F_n}{\partial y_1}(x_0, y_0) & \dots & \frac{\partial F_n}{\partial y_n}(x_0, y_0) \end{Vmatrix}.$$

Частные производные по переменным $x_1, \dots, x_m$ образуют левую подматрицу — обозначим её $F'_x(x_0, y_0)$, в которой m столбцов и n строк, а производные по переменным $y_1, \dots, y_n$ образуют правую квадратную подматрицу размера $n \times n$ — её обозначим $F'_y(x_0, y_0)$:

$$F'_x(x_0,y_0) = \begin{Vmatrix} \frac{\partial F_1}{\partial x_1}(x_0,y_0) & \dots & \frac{\partial F_1}{\partial x_m}(x_0,y_0) \\ \dots & \dots & \dots \\ \frac{\partial F_n}{\partial x_1}(x_0,y_0) & \dots & \frac{\partial F_n}{\partial x_m}(x_0,y_0) \end{Vmatrix}, \qquad F'_y(x_0,y_0) = \begin{Vmatrix} \frac{\partial F_1}{\partial y_1}(x_0,y_0) & \dots & \frac{\partial F_1}{\partial y_n}(x_0,y_0) \\ \dots & \dots & \dots \\ \frac{\partial F_n}{\partial y_1}(x_0,y_0) & \dots & \frac{\partial F_n}{\partial y_n}(x_0,y_0) \end{Vmatrix}.$$

Им соответствуют линейные отображения $L_X \in \mathscr{L}_{mn}$ и $L_Y \in \mathscr{L}_{nn}$ (“дифференциалы исходного отображения F по x и y”; точнее, это сужения $d_{(x_0,y_0)}F$ на подпространста векторов из $\mathbb{R}^{m+n}$, у которых равны нулю последние n и соответственно первые m координат). Тогда для любых векторов $H \in \mathbb{R}^m$ и $h \in \mathbb{R}^n$ выполняется равенство

$$\big(d_{(x_0,y_0)}F\big)(H,h) = L_X(H) + L_Y(h).$$

Следовательно,

$$\widetilde{F}(x,y) = F(x_0,y_0) + L_X(x-x_0) + L_Y(y-y_0).$$

Поэтому уравнение $F(x,y) = \mathbb{O}_n$, которому удовлетворяет точка (x_0,y_0), заменяется “похожим уравнением”

$$L_X(x-x_0) + L_Y(y-y_0) \approx \mathbb{O}_n.$$

Чтобы решить его относительно y, нужна обратимость отображения L_Y, т.е. неравенство $\det F'_y(x_0,y_0) \neq 0$. В этом случае

$$y \approx y_0 - L_Y^{-1}\big(L_x(x-x_0)\big).$$

Точку (x_0,y_0) назовём *невырожденной*, если $\det F'_y(x_0,y_0) \neq 0$. Невырожденность точки — это многомерный аналог неравенства $F'_{m+1} \neq 0$, возникшего при исследовании неявной функции.

Теорема (о неявном отображении). *Пусть E — открытое подмножество пространства $\mathbb{R}^{m+n}$ и $F : E \to \mathbb{R}^n$ — гладкое на E отображение.*

Тогда уравнение $F = \mathbb{O}_n$ вблизи каждой невырожденной точки, удовлетворяющей этому уравнению, задаёт единственное неявное отображение и оно гладкое: если векторы $x_0 \in \mathbb{R}^m$ и $y_0 \in \mathbb{R}^n$ таковы, что $(x_0,y_0) \in E$, $F(x_0,y_0) = \mathbb{O}_n$ и $\det\Big(\frac{\partial F_k}{\partial y_j}(x_0,y_0)\Big)_{1\leqslant j,k\leqslant n} \neq 0$, то существуют такие положительные числа δ и Δ, что

1) $(x,y) \in E$, если $\|x-x_0\|_m < \delta$, $\|y-y_0\|_n < \Delta$;

2) для любого x, $\|x-x_0\|_m<\delta$, существует единственный вектор $G(x)\in\mathbb{R}^n$, $\|G(x)-y_0\|_n<\Delta$, удовлетворяющий уравнению $F(x,G(x)) = \mathbb{O}_n$; в частности $G(x_0) = y_0$;

3) отображение G непрерывно дифференцируемо в шаре $\|x-x_0\|_m < \delta$;

4) справедливо равенство

$$G'(x_0) = -\big(F'_y(x_0,y_0)\big)^{-1} \circ F'_x(x_0,y_0).$$

В теореме ради простоты формулировки предполагается линейная независимость *последних* n столбцов матрицы Якоби. Если какие-то другие её n столбцов линейно независимы, то соответствующие им n переменных можно выразить через остальные m переменных.

Отметим ещё, что числа δ и Δ зависят не только от E и F, но и от точки (x_0,y_0) из E.

Доказательство. Сведём доказываемое утверждение к уже рассмотренному в §2 частному случаю — теореме об обратном отображении. Это удаётся сделать за счёт увеличения размерности задачи.

Определим новое отображение $\Phi:\ E\to\mathbb{R}^{m+n}$, действующее по правилу

$$\Phi(x,y)=(x,F(x,y))\quad\text{для любой точки}\quad (x,y)\quad\text{из}\quad E.$$

Первые m координатных функций отображения Φ — линейные функции (это координатные функции тождественного отображения). Они непрерывно дифференцируемы. Последние n координатных функций отображения Φ непрерывно дифференцируемы по условию (это координатные функции отображения F). Поэтому отображение Φ непрерывно дифференцируемо на E. Его матрица Якоби в произвольной точке (x,y) множества E равна

$$\Phi'(x,y)=\begin{Vmatrix} \frac{\partial\Phi_1}{\partial x_1}(x,y) & \dots & \frac{\partial\Phi_1}{\partial x_m}(x,y) & \frac{\partial\Phi_1}{\partial y_1}(x,y) & \dots & \frac{\partial\Phi_1}{\partial y_n}(x,y)\\ \dots & \dots & \dots & \dots & \dots & \dots\\ \frac{\partial\Phi_{m+n}}{\partial x_1}(x,y) & \dots & \frac{\partial\Phi_{m+n}}{\partial x_m}(x,y) & \frac{\partial\Phi_{m+n}}{\partial y_1}(x,y) & \dots & \frac{\partial\Phi_{m+n}}{\partial y_n}(x,y)\end{Vmatrix}.$$

По определению $\Phi_k(x,y)=x_k$ для $k\leqslant m$ и $\Phi_k(x,y)=F_{k-m}(x,y)$ для $k>m$. В частности,

$$\frac{\partial\Phi_k}{\partial x_j}=\begin{cases}1, & \text{если}\quad k=j,\\ 0, & \text{если}\quad k\neq j,\end{cases}\qquad \frac{\partial\Phi_k}{\partial y_i}=0\quad\text{для}\quad k,j=1,\dots,m\quad\text{и}\quad i=1,\dots,n.$$

Следовательно, матрица $\Phi'(x,y)$ имеет клеточную структуру:

$$\Phi'(x,y)=\begin{Vmatrix}\mathbb{I}_m & \mathbb{O}_{mn}\\ A(x,y) & B(x,y)\end{Vmatrix},$$

где $\mathbb{I}_m$ — единичная матрица размера $m\times m$, $\mathbb{O}_{mn}$ — нулевая матрица (m строк и n столбцов), а матрицы $A(x,y)$ и $B(x,y)$ — левая и правая части матрицы $F'(x,y)$:

$$A(x,y)=\begin{Vmatrix}\frac{\partial F_1}{\partial x_1}(x,y) & \dots & \frac{\partial F_1}{\partial x_m}(x,y)\\ \dots & \dots & \dots\\ \frac{\partial F_n}{\partial x_1}(x,y) & \dots & \frac{\partial F_n}{\partial x_m}(x,y)\end{Vmatrix},\qquad B(x,y)=\begin{Vmatrix}\frac{\partial F_1}{\partial y_1}(x,y) & \dots & \frac{\partial F_1}{\partial y_n}(x,y)\\ \dots & \dots & \dots\\ \frac{\partial F_n}{\partial y_1}(x,y) & \dots & \frac{\partial F_n}{\partial y_n}(x,y)\end{Vmatrix}.$$

Поэтому

$$\det\Phi'(x,y)=\det B(x,y)=\det\left(\frac{\partial F_k}{\partial y_i}(x,y)\right)_{1\leqslant i,k\leqslant n}\neq 0\quad\text{вблизи точки}\quad (x_0,y_0).$$

По теореме об обратном отображении Φ — локальный диффеоморфизм: для произвольной точки (x_0,y_0) существуют столь малый шар B, $B\subset E$, с центром в (x_0,y_0), что сужение $\Phi|_B$ диффеоморфно отображает B на открытое множество $\Phi(B)$. Взяв точку (x_0,y_0), удовлетворяющую равенству $F(x_0,y_0)=\mathbb{O}_n$, получим, что $\Phi(x_0,y_0)=\big(x_0,F(x_0,y_0)\big)=(x_0,\mathbb{O}_n)$ — внутренняя точка множества $\Phi(B)$.

Пусть $\Psi=\big(\Phi|_B\big)^{-1}$, $\Psi:\ \Phi(B)\to B$. Ясно, что $\Psi(x_0,\mathbb{O}_n)=(x_0,y_0)$. У этого отображения первые m координатных функций те же, что у Φ: $\Psi_j(x,y)=x_j$ для $j=1,\dots,m$ и $(x,y)\in\Phi(B)$. Поэтому Ψ можно записать в виде $\Psi(x,y)=(x,\theta(x,y))$, где θ — гладкое отображение из $\Phi(B)$ в $\mathbb{R}^n$. При $x=x_0$ и $y=y_0$ мы получаем $(x_0,y_0)=\Psi(x_0,\mathbb{O}_n)=(x_0,\theta(x_0,\mathbb{O}_n))$, т.е. $y_0=\theta(x_0,\mathbb{O}_n)$. По определению отображений Φ, Ψ и θ для всякой точки (x,y) из $\Phi(B)$ выполняется равенство

$$(x,y)=\Phi\big(\Psi(x,y)\big)=\Phi\big(x,\theta(x,y)\big)=\big(x,F(x,\theta(x,y))\big).$$

В частности, если $(x,\mathbb{O}_n)\in\Phi(B)$, то

$$(x,\mathbb{O}_n)=\big(x,F(x,\theta(x,\mathbb{O}_n)\big),\quad\text{т.е.}\quad \mathbb{O}_n=F\big(x,\theta(x,\mathbb{O}_n)\big).$$

Поэтому вектор $G(x) := \theta(x, \mathbb{O}_n)$ удовлетворяет уравнению $F(x, G(x)) = \mathbb{O}_n$, если $(x, \mathbb{O}_n) \in \Phi(B)$. Поскольку $(x_0, \mathbb{O}_n)$ — внутренняя точка множества $\Phi(B)$, существует столь малое положительное число δ, что $(x, y) \in \Phi(B)$, если $\|x - x_0\|_m < \delta$ и $\|y\|_n < \delta$. В частности, $(x, \mathbb{O}_n) \in \Phi(B)$ для таких x, т.е. G определено во всех точках шара $\|x - x_0\|_m < \delta$. В нём отображение G гладкое, так как его координатные функции получены сужением координатных функций гладкого отображения θ.

Отметим, что $G(x_0) = \theta(x_0, \mathbb{O}_n) = y_0$ и $(x, G(x)) = (x, \theta(x, \mathbb{O}_n)) = \Psi(x, \mathbb{O}_n) \in B$, если $(x, \mathbb{O}_n) \in \Phi(B)$ (в частности, это верно для всех x из шара $\|x - x_0\|_m < \delta$).

Пусть $\Delta > 0$ таково, что $\|G(x) - y_0\|_n < \Delta$ при $\|x - x_0\|_m < \delta$. Уменьшив числа δ и Δ, если это необходимо, мы можем считать, что $(x, y) \in B$, как только $\|x - x_0\|_m < \delta$ и $\|y - y_0\|_n < \Delta$. Если для таких x и y окажется, что $F(x, y) = \mathbb{O}_n$, то

$$\Phi(x, y) = \big(x, F(x, y)\big) = (x, \mathbb{O}_n) = \big(x, F(x, G(x))\big) = \Phi(x, G(x)).$$

Тогда $y = G(x)$, так как Φ взаимно однозначно на шаре B и ему принадлежат точки (x, y), $(x, G(x))$. Таким образом, для каждого x, $\|x - x_0\|_m < \delta$, уравнение $F(x, y) = \mathbb{O}_n$ имеет единственное решение y, удовлетворяющее неравенству $\|y - y_0\|_n < \Delta$ (и это решение — $y = G(x)$).

Осталось вычислить $d_{x_0}G$. Пусть L_X и L_Y — линейные отображения, порождаемые матрицами $A(x_0, y_0)$ и $B(x_0, y_0)$ (ради простоты обозначений мы не отмечаем зависимость L_X и L_Y от x_0 и y_0). Воспользуемся дифференцируемостью отображения F в точке (x_0, y_0):

$$\begin{aligned} F(x, y) &= F(x_0, y_0) + \big(d_{(x_0, y_0)}F\big)(x - x_0, y - y_0) + \|(x - x_0, y - y_0)\|_{m+n}\alpha(x, y) = \\ &= L_X(x - x_0) + L_Y(y - y_0) + \|(x - x_0, y - y_0)\|_{m+n}\alpha(x, y), \end{aligned}$$

где $\alpha : E \to \mathbb{R}^n$ и $\alpha(x_0, y_0) = \lim\limits_{(x,y)\to(x_0,y_0)} \alpha(x, y) = \mathbb{O}_n$. Взяв $y = G(x)$, $\|x - x_0\|_m < \delta$, получим:

$$\mathbb{O}_n = F(x, G(x)) = L_X(x - x_0) + L_Y(G(x) - G(x_0)) + \big\|\big(x - x_0, G(x) - G(x_0)\big)\big\|_{m+n}\alpha(x, G(x)).$$

По условию $L_y \in \mathrm{Isom}(\mathbb{R}^n)$. Поэтому

$$G(x) = G(x_0) - L_Y^{-1}\big(L_X(x - x_0)\big) - \big\|\big(x - x_0, G(x) - G(x_0)\big)\big\|_{m+n} L_Y^{-1}\left(\alpha(x, G(x))\right).$$

Поскольку отображение G дифференцируемо в точке x_0, вблизи x_0 выполняется неравенство $\|G(x) - G(x_0)\|_n \leqslant C\|x - x_0\|_m$ с некоторым коэффициентом C. Поэтому

$$\|(x - x_0, G(x) - G(x_0))\|_{m+n} = \sqrt{\|x - x_0\|_m^2 + \|G(x) - G(x_0)\|_n^2} \leqslant \|x - x_0\|_m\sqrt{1 + C^2}.$$

Кроме того,

$$\lim_{x\to x_0} L_Y^{-1}\left(\alpha(x, G(x))\right) = L_Y^{-1}\left(\alpha(x_0, G(x_0))\right) = L_Y^{-1}\left(\mathbb{O}_n\right) = \mathbb{O}_n$$

и, следовательно, при $x \to x_0$

$$G(x) = G(x_0) - L_Y^{-1}\left(L_X(x - x_0)\right) + o\big(\|x - x_0\|_m\big),$$

т.е. $d_{x_0}G = -L_Y^{-1} \circ L_X$. Элементы матриц $[L_X] = A(x_0, y_0)$ и $[L_Y] = B(x_0, y_0)$ непрерывно зависят от (x_0, y_0), причём вторая матрица обратима не только в точке (x_0, y_0), но и во всех точках достаточно близких к (x_0, y_0). Кроме того $y_0 = G(x_0)$ непрерывно зависит от x_0. Поэтому элементы матриц $A(x_0, G(x_0))$ и $B(x_0, G(x_0))$, а следовательно, и матрицы $B^{-1}(x_0, G(x_0)) \cdot A(x_0, G(x_0))$, непрерывно зависят от x_0. Таким образом, отображение $d_{x_0}G = -L_Y^{-1} \circ L_X$ непрерывно зависит от x_0, т.е. отображение G гладкое.

Теорема доказана.