

§1 Оценка приращения дифференциального отображения

Утверждение 1. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \geq 2$. Тогда формула Лагранжа

$$f(b) - f(a) = f'(c)(b - a)$$

не работает.

Е.г. Пусть

$$f(t) := (\cos t, \sin t), \quad b - a = 2\pi$$

Теорема 2 (об оценке приращения отображения). Пусть $f: G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, G — выпуклое, f — дифференцируема,

$$\forall x \in G \quad \|f'(x)\| \leq M$$

Тогда $\forall a, b \in G \quad \|f(b) - f(a)\| \leq M\|b - a\|$

□ «Окружим» исходную функцию:

$$F = \psi \circ f \circ \varphi$$

Вот тут как раз нужна выпуклость, иначе отрезок $[a; b]$ может и не лежать в G

где

$$\varphi: \mathbb{R} \rightarrow \mathbb{R}^n$$

$$\varphi(t) := t(b - a) + a,$$

$$t \in [0, 1]$$

$$\psi: \mathbb{R}^m \rightarrow \mathbb{R}$$

$$\psi(y) := \langle y, \ell \rangle,$$

$$\ell = f(b) - f(a)$$

Заметим, что F — обычная вещественнозначная функция. Так что для неё работает формула Лагранжа:

$$\exists c \in [0, 1]: F(1) - F(0) = F'(c)(1 - 0) = F'(c)$$

Тогда из свойств нормы (по ходу дела обозначим $\varphi(c)$ за x):

$$\|F'(c)\| = \|\psi'(f(x)) \cdot f'(x) \cdot \varphi'(c)\| \leq \|\psi'(f(x))\| \cdot \|f'(x)\| \cdot \|\varphi'(c)\|$$

Здесь тонкость в обозначениях. Производные — вроде матрицы, поэтому их нормы — что-то странное на первый взгляд. На самом деле смысл немного иной.

$$dL(x, h) = f'(x) \cdot h$$

Таким образом, дифференциал — неплохое линейное отображение. А под «нормой производной» имеется в виду норма соответствующего линейного отображения.

Теперь давайте что-нибудь скажем про эти нормы.

$$1. \quad \varphi'(t) = (b - a) \Rightarrow \|\varphi'(c)\| = \|b - a\|$$

$$2. \quad \psi(y) = \langle y, \ell \rangle, \quad \|\psi\| = \|\ell\|$$

Так что

$$\|F'(c)\| \leq M \cdot \|\ell\| \cdot \|b - a\|$$

С другой стороны:

$$F(1) - F(0) = \psi(f(b)) - \psi(f(a)) = \langle f(b), \ell \rangle - \langle f(a), \ell \rangle = \langle \ell, \ell \rangle = \|\ell\|^2$$

В итоге, совмещая оба выражения, приходим к утверждению теоремы. ■

§ 2 Частные производные высших порядков

Определение 1. Пусть $f: G \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\forall x \in G \exists \partial_{i_1, \dots, i_k}^k f(x)$. Тогда

$$\partial_{i_1, \dots, i_{k+1}}^{k+1} f(x) := \partial_{i_{k+1}} (\partial_{i_1, \dots, i_k}^k f)(x)$$

Замечание 1. $C^p(G)$ — класс функций, определённых в G с непрерывной производной до p -го порядка включительно. Функции из C^1 ещё называются гладкими.

Теорема 1 (Зависимость производных p -го порядка от перестановки переменных). Пусть $f \in C^p(G)$, $x \in G$. При этом

$$\begin{aligned} i &= \{i_1, \dots, i_p \mid i_k \in \{1, \dots, n\}\} \\ j &= \{j_1, \dots, j_p \mid j_k \in \{1, \dots, n\}\} \\ j &= \pi(i) \end{aligned}$$

$$\text{Тогда } \partial_i^p f(x) = \partial_j^p f(x)$$

□ Сначала докажем всё для $p = 2$, $n = 2$, т. е.

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

Пусть $(x, y) \in G$, $(x_0 + \Delta x, y) \in G$, $(x, y + \Delta y) \in G$, $(x + \Delta x, y + \Delta y) \in G$. Введём ещё 2 функции:

$$\begin{aligned} \varphi(t) &= f(t, y + \Delta y) - f(t, y) \\ \psi(t) &= f(x + \Delta x, t) - f(x, t) \end{aligned}$$

Тогда $\varphi(x + \Delta x) - \varphi(x) = \varphi'(c_1) \Delta x = W$, $c_1 \in [x, x + \Delta x]$. При этом

$$W = \varphi'(x) \Delta x = \left(\frac{\partial f}{\partial x}(c_1, y + \Delta y) - \frac{\partial f}{\partial x}(c_1, y) \right) \Delta x = \frac{\partial^2 f}{\partial x \partial y}(c_1, c_2) \Delta x \Delta y, \quad c_2 \in [y, y + \Delta y]$$

Аналогично

$$W = \frac{\partial^2 f}{\partial y \partial x}(c_3, c_4) \Delta x \Delta y, \quad c_4 \in [y, y + \Delta y], \quad c_3 \in [x, x + \Delta x]$$

По непрерывности второй производной f ($f \in C^2$)

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial y}(c_1, c_2) &\xrightarrow[\Delta y \rightarrow 0]{\Delta x \rightarrow 0} \frac{\partial^2 f}{\partial x \partial y}(x, y) \\ \frac{\partial^2 f}{\partial y \partial x}(c_3, c_4) &\xrightarrow[\Delta y \rightarrow 0]{\Delta x \rightarrow 0} \frac{\partial^2 f}{\partial y \partial x}(x, y) \end{aligned}$$

По теореме о предельном переходе в равенствах \smile смешанные производные равны.

Теперь поймём, что делать в случае произвольных p, n .

Представим подстановку π как произведение транспозиций соседних элементов. Будем дальше разбираться с такими транспозициями.

Пусть $\tau_k = (j, j+1)$. Сначала посчитаем производные по $x_1, \dots, x_{j-1} = i'$. А теперь обозначим $\tilde{f} = \partial_{i'} f$. По доказанному утверждению для двух переменных, $\partial_{j,j+1} \tilde{f} = \partial_{j+1,j} \tilde{f}$. А теперь продифференцируем это равенство по оставшимся переменным.

Таким образом, для произвольной транспозиции $\tau_k = (j, j+1)$ верно утверждение теоремы. А значит, и для произвольной подстановки $\pi = \prod_k \tau_k$ теорема верна. ■

Замечание 1. Тут важно, что $f \in C^p(G)$. Одной точки бы не хватило, мы ведь рассматриваем маленький параллелепипед в $U(x)$ и используем одномерную теорему Лагранжа внутри него. А для неё нужна дифференцируемость на интервале.

§ 3 Обобщение бинома

«Обычный» бином Ньютона:

$$(a + b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k}$$

Его нетрудно обобщить до полинома Ньютона

$$(a_1 + \dots + a_n)^p = \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_p=1}^n a_{i_1} \dots a_{i_p} = \sum_{\substack{\alpha_i \in \{0, \dots, p\} \\ \sum \alpha_i = p}} a_1^{\alpha_1} \dots a_n^{\alpha_n} \cdot C_{\alpha_1, \dots, \alpha_n}$$

Введём обозначения:

- $a = (a_1, \dots, a_n)$
- $\alpha = (\alpha_1, \dots, \alpha_n)$ — мультииндекс, по сути вектор индексов, для которого вводится своя куча обозначений, дабы упростить жизнь

$$1. \alpha + \beta = (\alpha_i + \beta_i)_{i=1}^n$$

$$2. |\alpha| = \sum_{i=1}^n \alpha_i$$

$$3. \alpha! = \prod_{i=1}^n \alpha_i!$$

$$\bullet a^\alpha = \prod_{i=1}^n a_i^{\alpha_i}$$

$$\bullet \partial_\alpha = \partial_{\alpha_1, \dots, \alpha_n}^n = \frac{\partial^n f}{\partial x_{\alpha_1} \dots \partial x_{\alpha_n}}$$

$$\bullet C_\alpha = C_{\alpha_1, \dots, \alpha_n}$$

Утверждение 1. $C_\alpha = \frac{p!}{\alpha!}$



Ну, это просто число перестановок с повторениями. Нужно взять по множителю из p скобок, причём перестановки одинаковых множителей входят в одно слагаемое. В итоге нужно делить общее число перестановок на число перестановок одинаковых множителей. А дальше можно сказать, что в каждое слагаемое входит каждый множитель, просто некоторые в нулевой степени.



§ 4 «Многомерный» дифференциал высоких порядков. Формула Тейлора для функций многих переменных

Определение 1. Пусть $f: G \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^p(G)$. Тогда

$$d^p f(x) := \sum_{1 \leq i_1 \leq \dots \leq i_p \leq n} \frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}} dx_{i_1} \dots dx_{i_p}$$

Или ещё можно вот так записать

$$\left(\sum_i dx_i \partial_i \right)^p f(x)$$

Утверждение 1. Если частные производные можно переставлять, то

$$d^p f(x) = \sum_{\substack{0 \leq \alpha_i \leq p \\ |\alpha| = p}} \frac{p!}{\alpha!} \partial_\alpha f(x) dx^\alpha$$

Теорема 2. Пусть $f \in C^{p+1}(G)$, $G \subset \mathbb{R}^n$, G — выпуклая, $a \in G$. Пусть также $h \in \mathbb{R}^n$: $a + h \in G$. Тогда

$$f(a + h) = \sum_{k=0}^p \frac{1}{k!} d^k f(a, h) + R_p(h)$$

Остаток $R_p(h)$ можно представить несколькими способами:

1. В форме Пеано: $R_p(h) = o(\|h\|^p)$

2. В форме Лагранжа: $R_p(h) = \frac{1}{(p+1)!} d^{p+1} f(a + \theta h, h)$, $\theta \in (0, 1)$

□ Рассмотрим $\varphi(t) = a + th$, $t \in [0, 1]$, $F(t) = f(\varphi(t))$, $F: [0, 1] \rightarrow \mathbb{R}$. По одномерной теореме Тейлора

$$F(1) = F(0) + 1 \cdot F'(0) + \frac{1}{2!} F''(0) 1^2 + \dots + \frac{1}{p!} F^{(p)}(0) \cdot 1^p + R_p$$

Докажем, что $F^{(k)}(0) = d^k f(a, h)$. Проще всего по индукции, давайте ещё такое нестандартное обозначение введём: $(k) = (1, \dots, k)$, и будем понимать под $i_{(k)}$ вектор индексов, а под $h_{i_{(k)}}$ — произведение соответствующих h .

база: $F(0) = f(a)$

переход: Пусть $F^{(k-1)}(t) = \sum_{1 \leq i_1, \dots, i_{k-1} \leq n} \partial_{i_{k-1}} f(a + th) h_{i_{k-1}}$. При дифференцировании по t всякие h_{i_j} в каждом слагаемом вынесутся за знак производной, а частная производная от f даст $\sum_{i_k} \partial_{i_k} f(a + ht) h_{i_k}$. Если скомпоновать все суммы и подставить $t = 0$, как раз получается $d^k f(a, h)$

Теперь разберёмся с остатком

$$R_p = \frac{1}{(p+1)!} F^{(p+1)}(\theta) \cdot 1^{p+1} = \frac{1}{(p+1)!} d^{(p+1)} f(a + h\theta, h)$$

Поскольку $\forall i \ |h_i| \leq \|h\|$

$$d^{(p+1)} f(a + \theta h, h) = O(\|h\|^{p+1}) = o(\|h\|^p)$$

Следовательно, $R_p = o(\|h\|^p)$ ■

§ 5 Понятие экстремума, необходимое условие

Определение 1. Пусть $f: G \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $a \in G$. Тогда говорят, что f имеет в a максимум (нестрогий), если

$$\exists U(a): \forall x \in U \ f(x) \leq f(a)$$

Когда неравенство строгое, а окрестность проколота, то максимум — строгий. Для минимума нужно \geq .

Теорема 1 (Необходимое условие экстремума). Пусть a внутренняя точка $G \subset \mathbb{R}^n$, $f \in C^1(a)$. Тогда если f имеет в a экстремум, то

$$df(a) = 0 \Leftrightarrow \forall i \ \partial_i f(a) = 0$$

□ Рассмотрим $\varphi_i(t) = f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n)$. Тогда у такой функции есть экстремум в a_i . А тогда, из одномерной теоремы Ферма $d\varphi_i(t) = 0$. А значит $\partial_i f = 0$ ■

§ 6 Про квадратичные формы

Определение 1. Функция $A: V \times V \rightarrow \mathbb{R}$, где V — векторное пространство, называется *билинейной формой*, если она линейна по обоим своим аргументам.

Определение 2. Билинейная форма A называется *симметрической*, если $\forall x, y \ A(x, y) = A(y, x)$.

Определение 3. Пусть A — билинейная форма, e_1, \dots, e_n — базис в векторном пространстве. Тогда

$$A(x, y) = \sum_{i,j=1}^n A(e_i, e_j) x^i y^j$$

и матрица A , элементы которой $a_{ij} = A(e_i, e_j)$ называется матрицей билинейной формы.

тут изложение больше по
[?]

Определение 4. Пусть A — симметрическая билинейная форма. Тогда $A(x) = A(x, x)$ — *квадратичная форма*. При этом $A(x, y)$ называется *полярной формой* по отношению к $A(x)$.

Определение 5. Матрица квадратичной формы — матрица соответствующей полярной формы.

Определение 6 («Определённость» формы). Если что-то верно, то про форму $A(x, y)$ говорят:

- $\forall x, y \neq 0 \quad A(x, y) > 0$ — положительно определена
- $\forall x, y \neq 0 \quad A(x, y) < 0$ — отрицательно определена
- $\forall x, y \neq 0 \quad A(x, y) \geq 0$ — полуопределена в положительном смысле
- $\forall x, y \neq 0 \quad A(x, y) \leq 0$ — полуопределена в отрицательном смысле

Е.г. Скалярное произведение — положительно определённая билинейная форма.

Теорема 1. Пусть в некотором базисе f_1, \dots, f_n квадратичная форма A имеет матрицу (a_{ij}) . Пусть к тому же все «северо-западные» миноры Δ_i отличны от нуля. Тогда существует базис e_1, \dots, e_n , в котором матрица A имеет вид

$$\begin{pmatrix} \frac{1}{\Delta_1} & & & \\ & \frac{\Delta_1}{\Delta_2} & & \\ & & \ddots & \\ & & & \frac{\Delta_{n-1}}{\Delta_n} \end{pmatrix}$$

□ По сути, нам нужно построить ортогональный базис, только вместо скалярного произведения используется билинейная форма A (причём не обязательно положительно определённая). За подробностями см. [?]. ■

Теорема 2 (Правило Сильвестра). Квадратичная форма положительно определена, если все миноры из теоремы 0.6.1 положительны, и отрицательно определена, если их знаки чередуются, начиная с «-».

□ Следствие предыдущей теоремы. Определённость формы не зависит от выбора базиса. ■

§ 7 Достаточное условие экстремума

Теорема 1 (Достаточное условие экстремума). Пусть $a \in G \subset \mathbb{R}^n$, a — внутренняя точка, $f \in C^2(a)$.

1. $df(a) = 0, d^2f(a, h) > 0 \Rightarrow f$ имеет в a min
2. $df(a) = 0, d^2f(a, h) < 0 \Rightarrow f$ имеет в a max
3. $df(a) = 0, d^2f(a, h) \leq 0 \Rightarrow$ ничего нет
4. $df(a) = 0, d^2f(a, h) \leq 0 \Rightarrow f$ не имеет в a min

5. $df(a) = 0, d^2f(a, h) \geq 0 \Rightarrow f$ не имеет в a max

□ Поскольку $df(a) = 0, \Delta f(a) = \frac{1}{2}(d^2f(a) + \alpha)$, где $\alpha = o(\|h\|)$. Для упрощения жизни примем $t = \frac{h}{\|h\|}$. Тогда приращение функции можно переписать в виде

$$\Delta f = \|h\|^2 \left(\frac{1}{2} \sum b_{ij} t_i t_j + \frac{\alpha}{\|h\|^2} \right)$$

Поскольку $\frac{\alpha}{\|h\|^2} \rightarrow 0$, существует $\overset{\circ}{U}_\varepsilon(a)$ в которой знак приращения определяется лишь первым слагаемым. Нетрудно заметить, что все значения t лежат на единичной сфере, которая компакт. Причём значения t покрывают всю сферу, ведь направление h можно выбирать в окрестности a произвольно. Так что можно просто сделать второе слагаемое меньшим минимума квадратичной формы на единичной сфере.

Таким способом можно расправиться с пунктами 1–2.

Для пункта 3 отыщем $h_1: d^2(a, h_1) > 0, h_2: d^2(a, h_1) < 0$. Заметим, что если A — квадратичная форма, то $A(h) > 0 \Rightarrow \forall s A(sh) = s^2 A(h) > 0$. По сути, мы считаем значение формы вдоль прямой, проходящей через a . Если, как и выше, записать приращение в виде

$$\Delta f = s^2 \left(\frac{1}{2} d^2(a, h_{1|2}) + \frac{\alpha}{s^2} \right)$$

то видно, что можно получить в окрестности $\overset{\circ}{U}(a)$ всё, что угодно, просто $s \rightarrow 0$.

4–5 легко доказываются от противного. ■

§ 8 Полнота пространства \mathbb{R}^n

Определение 1. Последовательность (x_n) называется фундаментальной (последовательностью Коши), если

$$\forall \varepsilon > 0 \exists N: \forall m, n > N \rho(x_n - x_m) < \varepsilon$$

Определение 2. Метрическое пространство (X, ρ) — полное, если всякая фундаментальная последовательность в нём сходится

Е.г. $\mathbb{R} \setminus 0$ — не полное метрическое пространство, $x_n = 1/n$ тому пример.

Замечание 1. Если (X, ρ) — полно, то X вообще-то замкнуто. Хорошо видно на примере выше.

Замечание 2. Если (X, ρ) — полно, $Y \subset X$ — замкнуто. Тогда и Y — полно.

Утверждение 1. \mathbb{R}^n — полное метрическое пространство.

▼

◁ произвольный $\varepsilon > 0$. Тогда

$$\rho(x_n - x_m) < \varepsilon \Rightarrow \forall i \rho(x_m^i e_i - x_n^i e_i) < \varepsilon \Rightarrow \forall i |x_m^i - x_n^i| < \varepsilon$$

Таким образом $(x_n^i) \in \mathbb{R}$ — фундаментальная. А в \mathbb{R} по теореме из первого семестра фундаментальные последовательности сходятся. Тогда $\forall x_n^i \rightarrow a^i$. Значит и $x_n \rightarrow a$.



Кусок дальше не шибко нужен

Давайте введём метрику на пространстве непрерывных функций

Определение 3. Пусть $f, g, h \in C([a; b])$. Тогда

$$\rho(f, g) := \sup_{x \in [a; b]} |f(x) - g(x)| = \|f - g\|$$

Здесь супремум можно заменить на максимум по теореме Вейерштрасса.

Докажем что это правда расстояние:

- $\rho(f, g) = \rho(g, f)$ — очевидно
- $\rho(f, g) \geq 0$ — тоже очевидно
- $\rho(f, g) = \rho(f, h) + \rho(h, g)$ — не так очевидно

$$\begin{aligned} \max_{x \in [a; b]} |f(x) - g(x)| &= |f(x_0) - g(x_0)| \leq |f(x_0) - h(x_0)| + |h(x_0) - g(x_0)| \\ &\leq \max_{x \in [a; b]} |f(x) - h(x)| + \max_{x \in [a; b]} |h(x) - g(x)| = \rho(f, h) + \rho(h, g) \end{aligned}$$

Утверждение 2. Пространство $C([a; b])$ с указанной выше метрикой полно.



Поточечная сходимость очевидна из полноты \mathbb{R} . А равномерную можно получить, устремив m к ∞ , зафиксировав n . А из равномерной сходимости следует непрерывность предельной функции.



Замечание. Если взять в качестве метрики $\int_a^b |f - g|$, то полнота ломается. Пополнение будет пространством суммируемых функций.

§ 9 Теорема о сжимающем отображении

Определение 1. Пусть (X, ρ) — метрическое пространство. Тогда отображение $T: X \rightarrow X$ называется сжимающим, если

$$\exists C \in (0, 1): \forall x', x'' \rho(T(x'), T(x'')) \leq C \cdot \rho(x', x'')$$

Теорема 1 (Банах). Пусть (X, ρ) — полное метрическое пространство, а отображение $T: X \rightarrow X$ — сжимающее. Тогда $\exists! x_* \in X: Tx_* = x_*$ (неподвижная точка).

Ещё часто ссылаются на следующий факт, появляющийся в процессе доказательства:

$$\forall x_0 \in X \exists \lim_{n \rightarrow \infty} T^n x_0 = x_*$$

□ $\triangleleft x_n = T^n x_0$, где $x_0 \in X$ — произвольное. Докажем, что

1. $x_n \rightarrow x_*$
2. $Tx_* = x_*$
3. других таких x_* нет.

Поехали

1. (x_n) сходится в себе, ведь $C \in (0, 1)$.

$$\rho(x_m, x_{m+p}) \leq \rho(x_m, x_{m+1}) + \dots + \rho(x_{m+p-1}, x_{m+p}) \leq \rho(x_0, x_1) C^m (1 + \dots + C^{p-1}) < \rho(x_0, x_1) \frac{C^m}{1-C} \xrightarrow{m \rightarrow \infty} 0$$

раз пространство полное, $\exists \lim_{n \rightarrow \infty} x_n$

2. отображение T непрерывно $\Rightarrow Tx_n \rightarrow Tx_*$. Но по теореме о подпоследовательности и единственности предела $Tx_* = x_*$.
3. Пусть x_{**} — другая неподвижная точка. Но тогда

$$\rho(x_{**}, x_*) = \rho(Tx_{**}, Tx_*) \leq C \rho(x_{**}, x_*) \Rightarrow (C-1) \rho(x_{**}, x_*) \geq 0 \xrightarrow{C < 1} \rho(x_{**}, x_*) = 0$$

■

§ 10 Метод Ньютона

Пусть $f \in C([a; b])$, $f(a) \cdot f(b) < 0$, $f(x_*) = 0$, $f'(x_*) \neq 0$. Сам метод выглядит как-то так:
Проводится касательная к графику в текущей точке, ищется её пересечение с осью x , отсюда восстанавливается перпендикуляр, пересечение которого с графиком — новая точка.

$$x - Tx = \frac{f(x)}{f'(x)} \Leftrightarrow Tx = x - \frac{f(x)}{f'(x)}$$

Докажем, что это вообще работает.

Теорема 1. Пусть $f \in C^2([a; b])$, $x_* \in [a; b]$:

- a) $f(x_*) = 0$
- b) $f'(x_*) \neq 0$

Тогда $\exists U(x_*)$: $\forall x_0 \in U$, такая что $T^n x_0 \rightarrow x_*$ и $x_{n+1} - x_* = O((x_n - x_*)^2)$

□ Сначала оценим T' .

$$T'(x) = 1 - \frac{f'(x)^2 - f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2} \xrightarrow[n \rightarrow \infty]{(f \in \mathcal{C}^2)} \frac{f(x_*)f''(x_*)}{f'(x_*)^2} = 0$$

Так что $\exists U(x_*) : c \in \bar{U} \Rightarrow |T'(c)| \leq 1/2$.

Теперь покажем, что $T(\bar{U}) \subset \bar{U}$. Из вышесказанного

$$|T_X - T_{X_*}| \leq \frac{1}{2}|X - X_*| \quad (1)$$

Поскольку $Tx_* = x_* - 0 = x_*$, то (1) равносильно

$$|Tx - x_*| \leq \frac{1}{2}|x - x_*|$$

А это как раз то, что надо. Значит T как раз сжимающее отображение, и по теореме 0.9.1 такой $x_*: f(x_*) = 0$ единственный.

Вторая часть тривиально получается из разложения f в ряд Тейлора в окрестности x_* . ■

§ 11 Теорема об обратном отображении(формулировка)

Пусть $F : G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ — гладкое. Порассуждаем, когда может существовать F^{-1} .

Рассмотрим, например, линейное отображение.

[illegible]

Понятно, что в таком случае задача поиска обратного отображения сводится к поиску обратной матрицы. Тогда из линала ясно, что для того, чтобы у нас всё вышло, нужно

$$m = n \wedge \det A \neq 0$$

Теперь попытаемся обобщить на остальные функции.

Пусть $a \in G$, $b = F(a)$

$$(?) \exists U(a), V(b) : F : U \leftrightarrow V$$

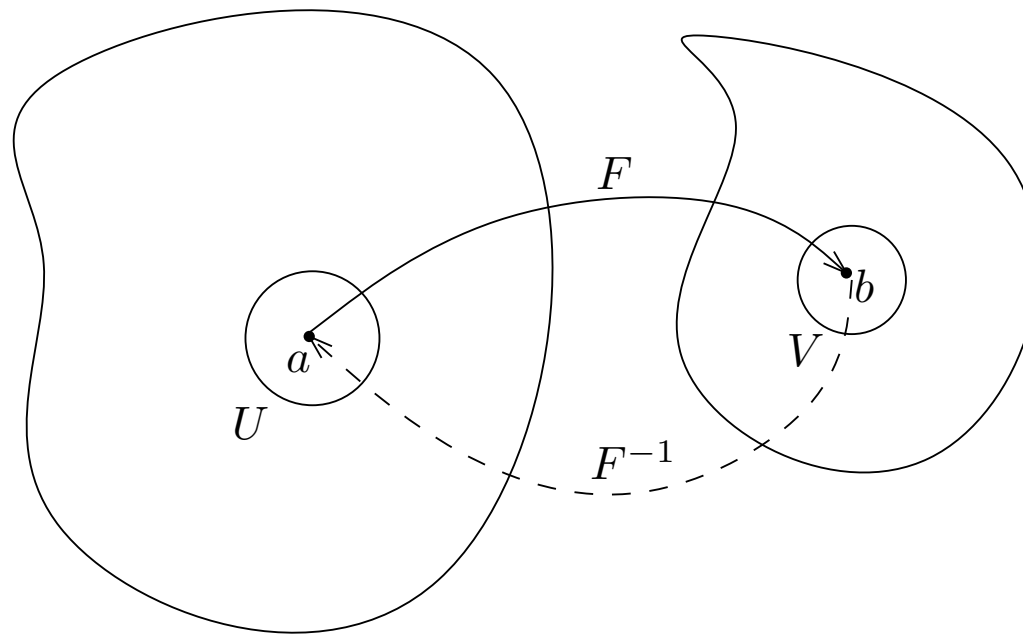
$$\Delta F = F(x) - F(a) = y - b = \Delta y \quad (1)$$

$$\Delta F = F'(a)dx + o(dx) \quad (2)$$

$$dF(a) = d\gamma(b) \quad (3)$$

Условие разрешимости (3) — $\det(F'(a)) \neq 0$. Утверждается, что (3) \Rightarrow (1).

Соответственно, формулировка



Теорема 1. Пусть $F: G \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, $a \in G$, $b = F(a)$. Пусть ещё $F \in C^1$, $\det(F'(a)) \neq 0$
Тогда

$$\begin{aligned} \exists U(a), V(b) : F: U &\leftrightarrow V \\ \exists F^{-1}V \rightarrow U, F^{-1} &\in C^1 \end{aligned}$$

§ 12 Доказательство теоремы об обратимости

□ (Теорема об обратимости отображения) Введём обозначения:

$$\begin{aligned} F'(a) &= \Gamma \\ \Phi(x) &= x - \Gamma^{-1}(F(x) - y) \end{aligned}$$

Нетрудно заметить, что x — неподвижная точка Φ (что $\Leftrightarrow F(x) = y$). Очень хотелось бы подогнать всё под теорему Банаха (0.9.1). Тогда отображение в окрестности a будет взаимно-однозначным.

1. Сначала оценим $\|\Phi'\|$.

$$\Phi'(x) = E - \Gamma^{-1}(F'(x)) = \Gamma^{-1}(F'(a) - F'(x))$$

Можно норму оценить

$$\|\Phi'(x)\| = \|\Gamma^{-1}\| \cdot \|F'(a) - F'(x)\|$$

Последний множитель явно $\xrightarrow{x \rightarrow a} 0$ (так как $F \in C^1$) Тогда и $\|\Phi'(x)\| \rightarrow 0$. А значит найдётся $U_\varepsilon(a)$: $\|\Phi'(x)\| \leq \frac{1}{2}$.

Тогда по теореме 0.1.2

$$x, x' \in U_\varepsilon(a) \Rightarrow \|\Phi(x) - \Phi(a)\| \leq \frac{1}{2} \|x - x'\|$$

Собственно, почти победа. Осталось лишь выбрать внутри U_ε компакт $\overline{U_{\varepsilon_1}}$ (иначе множество не очень полное).

2. Теперь покажем, что

$$\exists \overline{U}: \Phi(U) \subset U$$

Попутно примем $\|y - b\| < \delta$, это потом поможет доказать непрерывность.

$$\begin{aligned} \|\Phi(x) - a\| &= \|x - a - \Gamma^{-1}(F(x) - y)\| \leq \|\Gamma^{-1}\| \cdot \|\Gamma(x - a) - F(x) + y + b - b\| \\ &\leq \|\Gamma^{-1}\| \cdot (\underbrace{\| - (F(x) - F'(a)(x - a) - F(a)) \|}_{\alpha} + \|y - b\|) \end{aligned}$$

Выберем произвольный $\varepsilon: 0 < \varepsilon < \varepsilon_1$.

Однако мы ещё можем подкрутить ε_1 .

$$\exists U_{\varepsilon_1}: \frac{\|\alpha\|}{\|x - a\|} < \frac{1}{2\|\Gamma^{-1}\|}$$

Это следует из формулы Тейлора (0.4.2), а применять её можно, так как шар — выпуклое множество. Ещё выберем $\delta = \frac{\varepsilon}{2\|\Gamma^{-1}\|}$. Там правда ε , а не ε_1 .

Тогда цепочка неравенств выше преобразуется к такому виду

$$\dots < \|\Gamma^{-1}\| \cdot \frac{\|x - a\|}{2\|\Gamma^{-1}\|} + \frac{\varepsilon}{2\|\Gamma^{-1}\|} \cdot \|\Gamma^{-1}\|$$

А теперь положим $\|x - a\| \leq \varepsilon$ (неравенство нужно нестрогое для полноты). Тогда

$$x \in \overline{U_\varepsilon}(a) \Rightarrow \Phi(x) \in U_\varepsilon(a) \subset \overline{U_\varepsilon}(a)$$

А теперь по теореме Банаха

$$\exists! x_0 \in \overline{U_\varepsilon}(a): \Phi(x_0) = x_0 \Leftrightarrow F(x_0) = y_0$$

Видимо, осталось пересечь окрестность a с прообразом $V(b): U = F^{-1}(V) \cap U_\varepsilon(a)$

3. Заодно получилась и непрерывность:

$$\forall U \exists V_\delta(b): F^{-1}(V_\delta) \subset U_\varepsilon$$



§ 13 Теорема о дифференцируемости обратного отображения

Теорема 1 (о дифференцируемости F^{-1}). Пусть $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^n$, $F: U \leftrightarrow V$. Пусть также F — дифференцируемо в $a \in U$, $F(a) = b$, $\det F'(a) \neq 0$. Тогда F^{-1} дифференцируемо в b .

□ То, что есть обратное отображение, доказали выше. Пусть $y = F(x)$. Обозначим: $h = x - a$, $k = y - b$. Отображение биективно, значит $h \neq 0 \Leftrightarrow k \neq 0$. Из дифференцируемости F

$$k = y - b = F(x) - F(a) = Ah + \alpha, \alpha = o(h) \quad (h \rightarrow 0)$$

$A = F'(a) \neq 0$, следовательно $\exists A^{-1}$

$$A^{-1}k = A^{-1}Ah + A^{-1}\alpha \Rightarrow \Delta F^{-1} = h = A^{-1}k - A^{-1}\alpha$$

Докажем, что $-A^{-1}\alpha =: \beta = o(k) \quad (k \rightarrow 0)$

$$\beta \leq \frac{-\alpha}{\|k\|} = \frac{-\alpha}{\|h\|} \cdot \frac{\|h\|}{\|k\|}$$

Покажем, что последний член — ограничен

$$\frac{\|h\|}{\|k\|} = \frac{\|h\|}{\|Ah + \alpha\|} \leq \frac{\|h\|}{\|Ah\| - \|\alpha\|} = \frac{1}{\left| \frac{\|Ah\|}{\|h\|} - \frac{\|\alpha\|}{\|h\|} \right|}$$

А последнее выражение ограничено при $\|h\| < \delta$

■

Следствие. $(F^{-1})'(b) = (F'(a))^{-1}$

§ 14 Теорема о гладкости обратного отображения

Теорема 1. Пусть $F: U \leftrightarrow V$, биективна, $\in C^p$. Пусть к тому же $\det F'(x) \neq 0$. Тогда $F^{-1} \in C^p$

□ Введём обозначения (оно всё существует по предыдущим теоремам хоть где-то)

$$F'(x) = \left(\frac{\partial F_i}{\partial x_j} \right)_{i,j=1}^n = (a_{ij}) = A$$

$$(F^{-1})'(y) = \left(\frac{\partial F_i^{-1}}{\partial y_j} \right)_{i,j=1}^n = (b_{ij}) = B$$

Вполне ясно, что $B = A^{-1}$. Из алгебры $b_{ij} = \frac{A_{ji}}{\det A}$ (здесь \mathcal{A} — алгебраическое дополнение).

Заметим, что из последнего выражения следует, что b_{ij} — рациональная функция от $\{a_{lk}\}$. Следовательно, $\widetilde{b_{ij}} = b_{ij}(a_{11}, \dots, a_{kl}, \dots, a_{nn}) \in C^\infty$. С другой стороны

$$b_{ij}(y) = \frac{\partial F^{-1}}{\partial y_j}(y) = \frac{\partial F^{-1}}{\partial y_j}(F(x)) \Leftrightarrow b_{ij}(y) = \widehat{b_{ij}}(x)$$

Так что $\widehat{b_{ij}} = b_{ij} \circ F$.

Дальше немного магии. Введём ещё одну функцию

$$\overline{b_{ij}}(x) = b_{ij}(a_{11}(x), \dots, a_{kl}(x), \dots, a_{nn}(x))$$

Заметим, что каждая $a_{ij}(x) \in C^{p-1} \Rightarrow \overline{b_{ij}} \in C^{p-1}$. Хорошо, тогда

$$b_{ij}(y) = (\overline{b_{ij}} \circ F^{(-1)})(y)$$

Раньше доказали, что $F^{-1} \in C^0$. Теперь разматываем цепочку дальше:

$$F^{-1} \in C^i \Rightarrow \overline{b_{ij}} \circ F^{-1} \in C^i \Rightarrow b_{ij} \in C^i$$

Значит, частные производные F^{-1} принадлежат C^i . Тогда сама $F^{-1} \in C^{i+1}$. Таким бобром мы доберёмся до C^p . Дальше не выйдет, так как не хватит гладкости $\overline{b_{ij}}$. ■

§ 15 Гладкая зависимость корней многочлена от его коэффициентов

Теорема 1. Пусть $P(x) \in \mathbb{R}[x]$ имеет n корней (x_j^0) , $x_j^0 \in \mathbb{R}$, таких что $\forall i, j \ x_i^0 \neq x_j^0$. Тогда

$$x_i = x_i(a_0, \dots, a_{n-1}) \in C^\infty$$

□ Пусть $P(x) = (x - x_1) \cdots (x - x_n)$. Вспомним теорему Виета (из алгебры)

$$a_0 = (-1)^n x_1 \cdots x_n$$

$$a_1 = (-1)^{n-1} \sum_i \prod_{j \neq i} x_j$$

.....

$$a_{n-1} = (-1) \sum_i x_i$$

Рассмотрим P как отображение $(x_1, \dots, x_n) \mapsto (a_0, \dots, a_{n-1})$.

$$P'(x) = \begin{pmatrix} \frac{\partial P_1}{\partial x_1} & \cdots & \frac{\partial P_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial P_n}{\partial x_1} & \cdots & \frac{\partial P_n}{\partial x_n} \end{pmatrix} = \begin{pmatrix} (-1)^n \prod_{i \neq 1} x_i & (-1)^n \prod_{i \neq 2} x_i & \cdots & (-1)^n \prod_{i \neq n} x_i \\ \dots & \dots & \dots & \dots \\ -1 & -1 & \cdots & -1 \end{pmatrix}$$

Посчитаем $\det(F')$. Этот определитель можно рассмотреть как многочлен $\in R[x_1, \dots, x_n]$. Его степень не превосходит $0 + 1 + \dots + (n-1) = \frac{n(n-1)}{2}$. Заметим, что если хоть какая-то пара столбиков равны, то определитель равен нулю. Так что $\det(F')$ делится на всевозможные многочлены вида $x_i - x_j$. А их как раз $\frac{n(n-1)}{2}$ и они неприводимые. Следовательно,

$$\det(F')(x_1, \dots, x_n) = C \prod_{i < j} (x_i - x_j)$$

А значит при условии неравенства корней он ненулевой.

Дальше можно воспользоваться теоремой о гладкости обратного отображения. ■

§ 16 Теорема о неявном отображении

20:07 2016-10-15

Определение 1. Пусть $F: \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^m$. Рассмотрим уравнение

$$F(x, y) = 0 \quad (1)$$

Пусть $x^0 \in \mathbb{R}^n$, $y^0 \in \mathbb{R}^k$ такие, что $F(x^0, y^0) = 0$.

Тогда если $\exists P(x^0) \subset \mathbb{R}^n$, $Q(y^0) \subset \mathbb{R}^m$, такие что

$$\forall x \in P \exists! y \in Q: F(x, y) = 0$$

то говорят, что уравнение (1) задаёт неявную функцию $f: P \rightarrow Q$.

Сначала всякие комментарии.

$$(1) \Leftrightarrow \begin{cases} F_1(x_1, \dots, x_n, y_1, \dots, y_m) = 0 \\ \vdots \\ F_k(x_1, \dots, x_n, y_1, \dots, y_m) = 0 \end{cases}$$

Как обычно, главная идея состоит в том, чтобы всё линеаризовать

$$\begin{cases} dF_1 = 0 \\ \dots\dots\dots \\ dF_k = 0 \end{cases} \Leftrightarrow \begin{cases} \sum_{j=1}^m \frac{\partial F_1}{\partial y_j} dy_j = \sum_{j=1}^m \frac{\partial F_1}{\partial x_j} dx_j \\ \dots\dots\dots \\ \sum_{j=1}^m \frac{\partial F_k}{\partial y_j} dy_j = \sum_{j=1}^m \frac{\partial F_k}{\partial x_j} dx_j. \end{cases} \quad (2)$$

При этом dy_j мы хотим выразить через dx_j . Какие-то шансы обратить всё это дело есть лишь при условиях:

1. $k = m$
2. $\det \left(\frac{\partial(F_1, \dots, F_k)}{\partial(y_1, \dots, y_m)} \right) \neq 0$

Сейчас будем доказывать, что (2) \Rightarrow (1).

Теорема 1 (Теорема о неявном отображении). Пусть $F: G \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, $F \in C^p$, $p \geq 1$.

$$F(x, y) = 0, \quad (x_0, y_0) \in G$$

$$1. F(x_0, y_0) = 0$$

$$2. \det F'_y(x_0, y_0) \neq 0$$

Тогда $\exists P(x_0), Q(x_0)$, такие, что (1) задаёт неявное отображение $f: P \rightarrow Q$. При этом $f \in C^p$ и

$$f'(x) = -(F'_y(x, y))^{-1} \cdot F'_x(x, y)$$

□ Доказательство — «обёртка» над теоремой об обратном отображении. К слову, в [?, с. 673] сразу доказывается утверждение о неявном отображении.

Итак, обозначения:

1. $\Phi: G \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m$. Работает как-то так:

$$(x, y) \mapsto (u, v), \quad \begin{cases} u = x, & u \in \mathbb{R}^n \\ v = F(x, y), & v \in \mathbb{R}^m \end{cases}$$

2. $i: \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^m$ такого сорта $x \mapsto (x, 0)$

3. $\pi: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ такого сорта $(x, y) \mapsto y$

Теперь найдём определитель $\Phi'(x, y)$. Посчитав как-то частные производные, получим

$$\Phi'(x, y) = \left(\begin{array}{c|c} E_n & 0 \\ \hline F'_x & F'_y \end{array} \right) \Rightarrow \det(\Phi'(x_0, y_0)) = \det E_n \cdot \det F'_y(x_0, y_0) \neq 0$$

Чудно, значит по теореме об обратном отображении (0.11.1) $\exists \Phi^{-1}(x_0, y_0)$ и ещё окрестности $U(x_0, y_0), V(x_0, 0)$. Теперь определим окрестности из условия теоремы:

$$P(x_0) = i^{-1}(V) \cap Q(y_0) = \pi(U)$$

По сути — проекции.

В таких обозначениях $f = \pi \circ \Phi^{-1} \circ i$. Вполне очевидно, что $f \in C^p$. Ну $i, \pi \in C^\infty$, $\Phi^{-1} \in C^p$.

К тому же

$$\forall x \in P \quad x \xrightarrow{i} (x, 0) \xrightarrow{\Phi^{-1}} (x, y) \xrightarrow{\pi} y \in Q$$

При этом такой y — единственный. В итоге получилось задать неявно отображение f .

Из вышесказанного, оно сколько нужно раз дифференцируемо. Так что

$$\frac{\partial}{\partial x} F(x, f(x)) = F'_x \cdot E + F'_y \cdot f'(x) = 0$$

По условию F'_y — обратима. Следовательно,

$$f'(x) = -(F'_y(x, y))^{-1} \cdot F'_x(x, y)$$

■

§ 17 Функциональная зависимость системы функций

Определение 1. Пусть $f_1, \dots, f_m, g: G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ — гладкие функции, $x_0 \in G$. Тогда эти функции называются функционально зависимыми в $V(x_0)$, если

$$\exists \varphi: U(f(x_0)) \rightarrow \mathbb{R}, \varphi \in C^1 : g(x) = \varphi(f(x)) \text{ в } V(x_0)$$

Определение 2. Пусть $f_1, \dots, f_m, g: G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ — гладкие функции. Тогда эти функции называются функционально независими, если определение выше не выполняется ни для какой $V \subset G$

Замечание. Тут лажа какая-то с определениями, не отрицание же нифига.

Теорема 1. (о функциональной зависимости) Пусть $f_1, \dots, f_m, g: G \subset \mathbb{R}^n \rightarrow \mathbb{R}$ — гладкие функции. К тому же $a \in G$,

$f = (f_i)_i, y = f(x), \operatorname{rk} \begin{pmatrix} f'_1 \\ \vdots \\ f'_m \end{pmatrix} = m$ в точке $x \in U(a)$. Тогда, если $\operatorname{rk} \begin{pmatrix} f'_1 \\ \vdots \\ f'_m \\ g' \end{pmatrix} = m$ в точке $x \in U(a)$, то $\exists V(a)$ в которой g

функционально зависит от f_1, \dots, f_m .

□ Пусть сразу $n \geq m$, иначе условие теоремы не выполняется совсем никогда (ну там m векторов всегда ЛЗ).

Введём обозначения:

$$x = (\underbrace{x_1, \dots, x_m}_{\bar{x}}, \underbrace{x_{m+1}, \dots, x_n}_{\bar{\bar{x}}}), \quad \bar{y} = (y_1, \dots, y_m, \bar{\bar{x}})$$

Из алгебры в $f'(x), x \in U(a)$ существует ненулевой минор порядка m . Можно НУО считать, что он соответствует \bar{x} .

Тогда это равносильно тому, что $\det \left(\frac{\partial f}{\partial \bar{x}}(a) \right) \neq 0$.

Рассмотрим такую неявную функцию

$$F: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m, F(\bar{y}, \bar{x}) = y - f(\bar{x}, \bar{\bar{x}}) = 0$$

Оно всё по условию гладкое. Тогда по теореме о неявном отображении существует пара окрестностей P, Q и

$$\exists \varphi: P \subset \mathbb{R}^n \rightarrow Q \subset \mathbb{R}^m, \bar{x} = \varphi(\bar{y})$$

В этих окрестностях $F \equiv 0 \Leftrightarrow y \equiv f(\varphi(y, \bar{x}), \bar{x})$. Заметим, что здесь y, \bar{x} — независимые переменные. Так что если $j > m$, то

$$\frac{\partial}{\partial x_j} f_i(\varphi(\bar{y}), \bar{x}) = \sum_{k=1}^m \partial_k f_i \cdot \frac{\partial \varphi_k}{\partial x_j} + \partial_j f_i \equiv 0$$

Из условия на ранг известно, что

$$g'(x) = \sum_{i=1}^m \lambda_i f'_i(x), \quad x \in U(a)$$

Нам для того чтобы показать, что g функционально зависит от f , необходимо приравнять в окрестности точки a g к функции от y . Пусть снова $j > m$, тогда

$$g(x) = g(\bar{x}, \bar{x}) = g(\varphi(\bar{y}), \bar{x})$$

$$\frac{\partial g}{\partial x_j} = \sum_{k=1}^m \partial_k g \cdot \frac{\partial \varphi_k}{\partial x_j} + \partial_j g$$

А вот теперь нужно воспользоваться условием на ранги. Тут очень важно, что это условие работает в окрестности a — ведь какие-то тождества в точке нам ничего интересного не дадут.

$$\frac{\partial g}{\partial x_j} = \sum_{i=1}^m \lambda_i \left(\partial_k f \cdot \frac{\partial \varphi_k}{\partial x_j} + \partial_j f_i \right) = 0$$

Из того, что g, φ — гладкие получаем, что и функция, нужная в определении [0.17.1](#) тоже гладкая. Осталось только пересечь много окрестностей (из неявного отображения, условия на ранг etc). ■

§ 18 Геометрический смысл ранга матрицы Якоби

Определение 1 (Коразмерность). Пусть V — подпространство U . Тогда $\text{codim } V = \dim U - \dim V$.

Теорема 1. Пусть $F: G \subset \mathbb{R}^n \rightarrow \mathbb{R}^m, a \in G, b = F(a), \exists V(a): \forall x \in V \text{ rk } F'(x) = r$. Тогда

1. $\exists U(a): F(U)$ имеет вид графика $\bar{y} = \varphi(\bar{x})$
2. $\exists U(a) \subset F^{-1}(b)$ имеет вид графика $\bar{x} = \psi(\bar{y})$

□ Аккуратное следствие [0.17.1](#) и [0.16.1](#). ■

Замечание. Ранг, собственно, показывает сколько есть степеней свободы у значений функции, причём в довольно механическом смысле. Мало ли, вдруг мы отобразили пространство в какую-то кривую в другом.

Есть кстати шансы, что в этой теореме попутно определили размерность образа при отображении, но неточно. Собственно, график $\bar{y} = \varphi(\bar{x})$ можно считать заданным на $y \in \mathbb{R}^m$, которое уже «прямое», а дальше размерностью объявить $\dim\{\bar{y}\}$.

§ 19 Три способа локального задания поверхности

1. Параметрическое

$$f: D \subset \mathbb{R}^k \rightarrow \mathbb{R}^n : \operatorname{rk} f' = k \forall x \in D (\geq k)$$

Тогда $M = f(D)$ — поверхность размерности k .

Условие на ранг означает, что нигде нету изломов, параметр же по сути — скорость.

2. Задание графиком

$$D \subset \mathbb{R}^k, f: D \rightarrow \mathbb{R}^{n-k} \text{ — гладкое}$$

Тогда $M = \{(t, f(t)) \mid t \in D\} = \Gamma_f$.

Определение 1 (Поверхность(нестрого)). Множество $S \subset \mathbb{R}^m$ можно называть k -мерной гладкой поверхностью, если в окрестности любой своей точки оно задаётся графиком гладкого отображения $f: D \subset \mathbb{R}^k \rightarrow \mathbb{R}^{n-k}$.

3. Неявное

Пусть $F: \mathbb{R}^n \rightarrow \mathbb{R}^{n-k}$. Тогда

$$M = \{x \in \mathbb{R}^n \mid F(x) = 0\}$$

По сути уравнения связи.

Теорема 1. Если в некой окрестности $a \in \mathbb{R}^n$ k -мерная поверхность может быть задана один из 3 способов, то она может быть задана и всеми остальными.

□

1 → 2 см [0.18.1](#)

2 → 3 $F(t, y) = f(t) - y, F' = (f'_t \mid -E) \Rightarrow \operatorname{rk} F' = n - k$

3 → 2 см [0.18.1](#)

2 → 1 $(x, y) \mapsto (x(t), f(x(t)))$, где $t = x$. С рангами очевидно проблем нет.

■

§ 20 Условный экстремум(нестрого)

Определение 1 (Безусловный экстремум). Пусть $f: G \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $a \in G$ — внутренняя точка. Тогда в точке a \max / \min если

$$\exists U(a): \forall x \in U \quad f(x) \leqslant / \geqslant f(a)$$

Определение 2 (Экстремум на подмножестве). Пусть $f: G \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $M \subset \mathbb{R}^n$ — k -мерная поверхность, $a \in G \cap M$ — внутренняя точка. Тогда в точке a \max / \min относительно M , если

$$\exists U(a): \forall x \in U \cap M \quad f(x) \leqslant / \geqslant f(a)$$

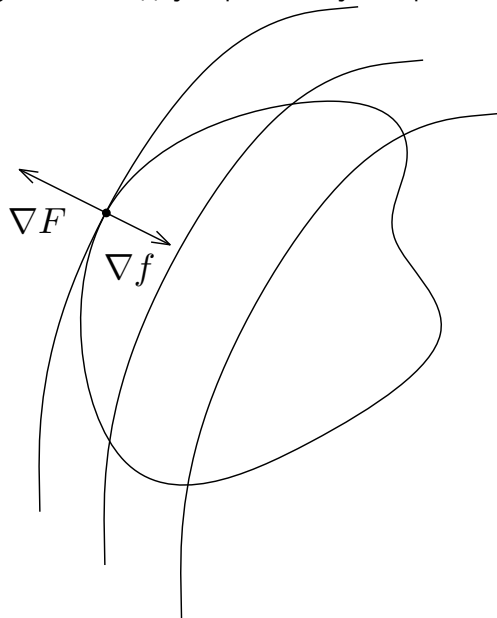
Чаще всего M задают неявно — «накладывают условия» на значения f .

Определение 3 (Условный экстремум). Пусть $f: G \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $F_1, \dots, F_m: G \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $a \in G \cap M$ — внутренняя точка, $F_1(a) = \dots = F_m(a) = 0$. Тогда в точке a *условный* \max / \min если

$$\exists U(a): \forall x \in U, F_1(x) = \dots = F_m(x) = 0 \quad f(x) \leqslant / \geqslant f(a)$$

Теорема 1. Пусть $f, F_1, \dots, F_m \in C^1(G)$, $a \in G$. Тогда если f имеет в a экстремум при условии $F(a) = 0$, то $\nabla f(a), \nabla F_1(a), \dots, \nabla F_m(a)$ линейно зависимы.

Е.г. Можно двумерный случай рассмотреть.



$$\nabla f = \lambda \nabla F$$

Следствие 1 (Правило Лагранжа). f имеет в a экстремум при условии $F_1(a) = \dots = F_m(a) = 0$, то

1. либо $\nabla F_1(a), \dots, \nabla F_m(a)$ ЛЗ
2. либо $\exists \lambda_1, \dots, \lambda_m \in \mathbb{R}: \nabla f(a) = \sum_i \lambda_i \nabla F_i(a)$

§ 21 Доказательство теоремы об условном экстремуме

□

1. Пусть $m = n - 1$. Будем доказывать от противного. Пусть в a условный \max , но $\nabla f(a), \nabla F_1(a), \dots, \nabla F_m(a) — \text{ЛНЗ}$. Рассмотрим $\Phi(x) = (f(x), F_1(x), \dots, F_m(x))$. Тогда такое $\Phi: G \rightarrow \mathbb{R}^n$. Из линейной независимости градиентов

$$\Phi'(a) = \begin{pmatrix} \nabla f(a) \\ \nabla F_1(a) \\ \vdots \\ \nabla F_m(a) \end{pmatrix}, \det \Phi'(a) \neq 0$$

Пусть $b = \Phi(a) = (f(a), 0, \dots, 0)$. Тогда по теореме об обратном отображении (0.14.1)

$$\exists U(a), V(b) : \Phi: U \rightarrow V — \text{диффеоморфизм}$$

Пусть $V \supset B_\varepsilon(b)$, $y = (f(a) + \frac{\varepsilon}{2}, 0, \dots, 0) \in V$, тогда $\exists! x \in U: \Phi(x) = y$. Получается, что $f(x) > f(a)$, $\forall i F_i(x) = 0$, что немного противоречит тому, что в a условный \max .

2. Теперь рассмотрим случай $m < n - 1$ (всё остальное неинтересно, точно будет ЛЗ). Будем доказывать от противного. Пусть в a условный \max , но $\nabla f(a), \nabla F_1(a), \dots, \nabla F_m(a) — \text{ЛНЗ}$. Тогда $\text{rk } \Phi'(a) = m + 1 < n$. Добавим ещё функций F_{m+1}, \dots, F_{n-1} таких, что $F_i(x) = x_{i+1} - ai + 1$.

Введём ещё стандартное обозначение

$$x = (\underbrace{x_1, \dots, x_{m+1}}_{\bar{x}}, \underbrace{x_{m+2}, \dots, x_n}_{\bar{\bar{x}}})$$

И не совсем стандартное

$$A = \frac{\partial(f_1, F_1, \dots, F_m)}{\partial x_1, \dots, x_{m+1}}(a)$$

Обычно такой «дробью» обозначают якобиан, но, пожалуй, сохраню обозначения с лекции.

Итак

$$\Phi'(a) = \left(\begin{array}{c|c} A & * \\ \hline 0 & E_{n-m-1} \end{array} \right) \Rightarrow \text{rk } \Phi'(a) = n$$

А теперь можно подвести всё к первому пункту. Рассмотрим $\tilde{M} = \{x \mid F_1(x) = \dots = F_{n-1}(x) = 0\}$ (а $M = \{x \mid F_1(x) = \dots = F_m(x) = 0\}$). Поскольку $\tilde{M} \subset M$, f будет иметь в a максимум и относительно \tilde{M} .

Аналогично 1 пункту получаем бред какой-то.

■

Замечание 1. Такая теорема может найти лишь точки, «подозрительные» на экстремум. Надо ещё отдельно думать. Например, вдруг там на компакте всё определено.

Замечание 2. Можно рассматривать функцию Лагранжа:

$$\mathcal{L}(x) = f(x) - \sum_{i=1}^m \lambda_i F_i(x)$$

Тогда если в a условный экстремум, то в a стационарная точка ($\mathcal{L}'(a) = 0$) функции Лагранжа.