§ 1 Оценка приращения дифференциального отображения

Утверждение 1. *Пусть* $f\colon \mathbb{R}^n \to \mathbb{R}^m$, $m \geqslant 2$. *Тогда формула Лагранжа*

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

не работает.

E.g. Пусть

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

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

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

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

□ «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

$$\begin{array}{lll} \varphi : \mathbb{R} \to \mathbb{R}^n & \varphi(t) := t(b-a) + a, & t \in [0,1] \\ \psi : \mathbb{R}^m \to \mathbb{R} & \psi(y) := \langle y, \ell \rangle, & \ell = f(b) - f(a) \end{array}$$

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

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

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

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

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

$$\mathrm{d}L(x,h) = f'(x) \cdot h$$

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. $\varphi'(t) = (b-a) \Rightarrow \|\varphi'(c)\| = \|b-a\|$
2. $\psi(y) = \langle y, l \rangle$, $\|\psi\| = \|\ell\|$

Так что

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

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 $f\colon G \subset \mathbb{R}^n \to \mathbb{R}$, $\forall\, x \in G\ \exists\, \partial^k_{i_1,\dots,i_k} f(x)$. Тогда

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

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*

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

Тогда $\partial^p_i f(x) = \partial^p_j f(x)$

☐ 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

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

Аналогично

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

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

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

По теореме о предельном переходе в равенствах $\ddot{\smile}$ 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 $\widetilde{f} = \partial_{i'} f$. По доказанному утверждению для двух переменных, $\partial_{j,j+1}\widetilde{f} = \partial_{j+1,j}\widetilde{f}$. 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

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

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. $\alpha! = \prod_{i=1}^n \alpha_i!$
- $a^\alpha = \prod_{i=1}^n a_i^{\alpha_i}$
- $\partial_\alpha = \partial^n_{\alpha_1,\dots,\alpha_n} = \dfrac{\partial^n f}{\partial x_{\alpha_1}\dots\partial x_{\alpha_n}}$
- $C_\alpha = C_{\alpha_1,\dots,\alpha_n}$

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

▼

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 $f\colon G \subset \mathbb{R}^n \to \mathbb{R}$, $f \in C^p(G)$. Тогда

$$\mathrm{d}^p f(x) := \sum_{1 \leqslant i_1 \leqslant \cdots \leqslant i_p \leqslant n} \frac{\partial^p f}{\partial x_{i_1} \ldots \partial x_{i_p}} \,\mathrm{d}x_{i_1} \cdots \mathrm{d}x_{i_p}$$

Или ещё можно вот так записать

$$\left(\sum_i \mathrm{d}x_i \,\partial_i\right)^p f(x)$$

Утверждение 1. *Если частные производные можно переставлять, то*

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

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

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

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

1. *В форме Пеано: $R_p(h) = o(\|h\|^p)$*
2. *В форме Лагранжа: $R_p(h) = \dfrac{1}{(p+1)!} \,\mathrm{d}^{p+1} f(a + \theta h, h)$, $\theta \in (0, 1)$*

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

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

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

база: $F(0) = f(a)$

переход: Пусть $F^{(k-1)}(t) = \sum\limits_{1 \leqslant i_1, \ldots, i_{k-1} \leqslant n} \partial_{i_{(k-1)}} f(a + th)\, h_{i_{(k-1)}}$. 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

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

Поскольку $\forall\, i \ \ |h_i| \leqslant \|h\|$

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

Следовательно, $R_p = o(\|h\|^p)$ ■

§ 5 Понятие экстремума, необходимое условие

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

$$\exists\, U(a) \colon \forall\, x \in U \ \ f(x) \leqslant f(a)$$

Когда неравенство строгое, а окрестность проколотая, то максимум — строгий Для минимума нужно $\geqslant$.

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*

$$\mathrm{d}f(a) = 0 \Leftrightarrow \forall\, i \ \ \partial_i f(a) = 0$$

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

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

1. $\mathrm{d}f(a) = 0$, $\mathrm{d}^2 f(a) > 0 \Rightarrow f$ *имеет в a* min
2. $\mathrm{d}f(a) = 0$, $\mathrm{d}^2 f(a) < 0 \Rightarrow f$ *имеет в a* max
3. $\mathrm{d}f(a) = 0$, $\mathrm{d}^2 f(a) \lesseqgtr 0 \Rightarrow$ *ничего нет*
4. $\mathrm{d}f(a) = 0$, $\mathrm{d}^2 f(a) \leqslant 0 \Rightarrow f$ *не имеет в a* min
5. $\mathrm{d}f(a) = 0$, $\mathrm{d}^2 f(a) \geqslant 0 \Rightarrow f$ *не имеет в a* max

§ 6 Понятие о неявной функции

Определение 1. Пусть $F\colon G \subset \mathbb{R}^2 \to \mathbb{R}$. Рассмотрим уравнение

$$F(x,y) = 0 \tag{1}$$

Пусть $a = (x_0, y_0)$ удовлетворяет (1), а U — окрестность a: $U = U_x \times U_y$. Тогда будем говорить, что уравнение (1) определяет неявную функцию f в U, если

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

Теорема 1 (О неявной функции)**.** *Пусть $F\colon G \subset \mathbb{R}^2 \to \mathbb{R}$, $F \in C^1(x_0, y_0)$, а $a = (x_0, y_0)$:*

1. *$F(x_0, y_0) = 0$*
2. *$F'_y(x_0, y_0) \neq 0$*

Тогда $\exists\, P(x_0), Q(y_0)$: в $U = P \times Q$ уравнение (1) задаёт неявную функцию $f\colon P \to Q$. При этом

$$f \in C^1 \wedge f'(x) = -\frac{F'_x(x,y)}{F'_y(x,y)}$$

□

1. (Доказательство существования)
   Рассмотрим $\varphi(y) = F(x_0, y)$. Пусть НУО $F'_y(x_0, y_0) > 0$. Тогда

   $$\exists\, U_\varepsilon(x_0, y_0)\colon \forall\, x, y \in U\ \ F'_y(x,y) > 0$$

   Обозначим соответствующие проекции U (шар) на координатные оси за U_x, U_y Получается, что $\varphi \uparrow U_y = (y_0 - \varepsilon; y_0 + \varepsilon)$. Тогда

   $$\begin{aligned} &\exists\, V_1(x_0)\colon \forall\, x \in V_1 F(x, y+\varepsilon) > 0 \\ &\exists\, V_2(x_0)\colon \forall\, x \in V_2 F(x, y-\varepsilon) < 0 \\ &P = V_1 \cap V_2 \end{aligned}$$

   Тогда из теоремы Больцано-Коши и монотонности φ

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

   В итоге получилось определение неявной функции.

2. Непрерывность в (x_0, y_0) вроде очевидна, мы же каждому x из P сопоставили 1 y из Q. Принадлежность классу C можно установить проведя аналогичные рассуждения для $x \in P(x_0)$

3. С гладкостью что-то странное. Можно наверное сделать как в Зориче.

4. $F(x, f(x)) \equiv 0 \Rightarrow F'_x \cdot 1 + F'_2 f'(x) = 0$

■

§ 7 Полнота пространства $\mathbb{R}^n$

§ 8 Теорема о сжимающем отображении

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

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

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

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

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

§ 9 Метод Ньютона

потом

§ 10 Теорема об обратном отображении(формулировка)

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

Рассмотрим, например, линейное отображение.

$$y = F(x) = Ax \Leftrightarrow \begin{cases} y_1 = a_{11}x_1 + \cdots + a_{1n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m = a_{m1}x_1 + \cdots + a_{mn}x_n \end{cases}$$

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.

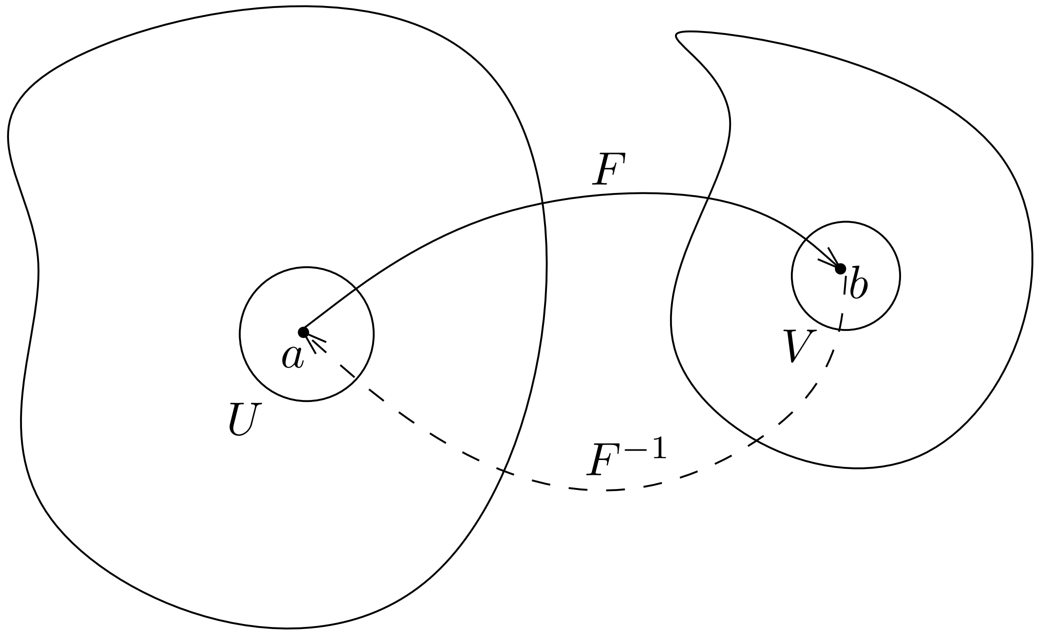
Пусть $a \in G,\ b = F(a)$

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

$$\Delta F = F(x) - F(a) = y - b = \Delta y \tag{1}$$

$$\Delta F = F'(a)dx + o(dx) \tag{2}$$

$$\mathrm{d}F(a) = \mathrm{d}y(b) \tag{3}$$



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

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

$$\exists\, U(a), V(b):\;\; F\colon U \leftrightarrow V$$
$$\exists\, F^{-1}V \to U,\; F^{-1} \in C^1$$

§ 11 Доказательство теоремы об обратимости

□ (*Теорема об обратимости отображения*) Введём обозначения:

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

Нетрудно заметить, что x — неподвижная точка Φ (что $\Leftrightarrow F(x) = y$). Очень хотелось бы подогнать всё под теорему Банаха (0.8.1). 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 $\xrightarrow[x\to a]{} 0$ (так как $F \in C^1$) Тогда и $\|\Phi'(x)\| \to 0$. А значит найдётся $U_\varepsilon(a)\colon \|\Phi'(x)\| \leqslant \frac{1}{2}$.

Тогда по теореме 0.1.2

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

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

2. Теперь покажем, что

$$\exists\, \overline{U}\colon \Phi(U) \subset U$$

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

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

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

Однако мы ещё можем подкрутить ε_1.

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

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

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

А теперь положим $\|x - a\| \leqslant \varepsilon$ (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

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

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

3. Заодно получилась и непрерывность:

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

$\blacksquare$

§ 12 Теорема о дифференцируемости обратного отображения

Теорема 1 (о дифференцируемости F^{-1})**.** *Пусть* $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^n$, $F\colon U \leftrightarrow V$. *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

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

$A = F'(a) \neq 0$, 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 $-A^{-1}\alpha =: \beta = o(k)\ (k \to 0)$

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

Покажем, что последний член — ограничен

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

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

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

§ 13 Теорема о гладкости обратного отображения

Теорема 1. *Пусть* $F\colon U \leftrightarrow V$, *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 $b_{ij} = \dfrac{\mathcal{A}_{ji}}{\det A}$ (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

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

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

Дальше немного магии. Введём ещё одну функцию

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

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

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

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 $\overline{b_{ij}}$. ■

§ 14 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)

$$\begin{aligned} a_0 &= (-1)^n x_1 \cdots x_n \\ a_1 &= (-1)^{n-1} \sum_i \prod_{j \neq i} x_j \\ &\cdots\cdots\cdots \\ a_{n-1} &= (-1) \sum_i x_i \end{aligned}$$

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

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

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,

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

Этого, конечно, не было в курсе алгебры, но там не используется ничего страшнее теоремы о делении с остатком. Вообще доказать бы надо, но лень.

А значит при условии неравенства корней он ненулевой.

Дальше можно воспользоваться теоремой о гладкости обратного отображения. ■

§ 15 Теорема о неявном отображении

20:07 2016-10-15

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

$$F(x, y) = 0 \tag{1}$$

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

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

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

Сначала всякие комментарии.

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

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

$$\begin{cases} \mathrm{d}F_1 = 0 \\ \dots\dots\dots \\ \mathrm{d}F_k = 0 \end{cases} \Leftrightarrow \begin{cases} \sum_{j=1}^m \frac{\partial F_1}{\partial y_j}\,\mathrm{d}y_j = \sum_{j=1}^m \frac{\partial F_1}{\partial x_j}\,\mathrm{d}x_j \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \sum_{j=1}^m \frac{\partial F_k}{\partial y_j}\,\mathrm{d}y_j = \sum_{j=1}^m \frac{\partial F_k}{\partial x_j}\,\mathrm{d}x_j. \end{cases} \tag{2}$$

При этом $\mathrm{d}y_j$ мы хотим выразить через $\mathrm{d}x_j$. Какие-то шансы обратить всё это дело есть лишь при условиях:

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

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

Теорема 1 (Теорема о неявном отображении). *Пусть $F\colon G \subset \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, $F \in C^p$, $p \geqslant 1$.*

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

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 $f\colon P \to Q$. При этом $f \in C^p$ и

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

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

Итак, обозначения:

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

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

2. $i\colon \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m$ такого сорта $x \mapsto (x, 0)$
3. $\pi\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ такого сорта $(x, y) \mapsto y$

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

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

Чудно, значит по теореме об обратном отображении (0.10.1) $\exists\, \Phi^{-1}(x_0, y_0)$ и ещё окрестности $U(x_0, y_0), V(x_0, 0)$. Теперь определим окрестности из условия теоремы:

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

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

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

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,

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

■

§ 16 Функциональная зависимость системы функций

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

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

Определение 2. Пусть $f_1, \ldots, f_m, g\colon G \subset \mathbb{R}^n \to \mathbb{R}$ — гладкие функции. Тогда эти функции называются функционально *не*зависимыми, 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 $f_1, \ldots, f_m, g\colon G \subset \mathbb{R}^n \to \mathbb{R}$ — 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 $f_1, \ldots, f_m$.*

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

Введём обозначения:

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

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

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

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

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

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

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

Из условия на ранг известно, что

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

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

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

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 0.16.1 тоже гладкая. Осталось только пересечь много окрестностей (из неявного отображения, условия на ранг etc). ■

§ 17 Геометрический смысл ранга матрицы Якоби

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

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

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

□ Аккуратное следствие 0.16.1 и 0.15.1. ■

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 $\bar{\bar{y}}=\varphi(\bar{y})$ можно считать заданным на $y\in\mathbb{R}^m$, которое уже «прямое», а дальше размерностью объявить $\dim\{\bar{y}\}$.

§ 18 Три способа локального задания поверхности

1. Параметрическое

$$f\colon D \subset \mathbb{R}^k \to \mathbb{R}^n : \quad \operatorname{rk} f' = k \forall x \in D (\geqslant k)$$

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

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

2. Задание графиком

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

   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 $f\colon D \subset \mathbb{R}^k \to \mathbb{R}^{n-k}$.

3. Неявное
   Пусть $F\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$. Тогда

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

   По сути уравнения связи.

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

□

$1 \to 2$ см 0.17.1

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

$3 \to 2$ см 0.17.1

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

■

§ 19 Условный экстремум(нестрого)

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

$$\exists\, U(a)\colon \forall\, x\in U\ \ f(x)\leqslant/\geqslant f(a)$$

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

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

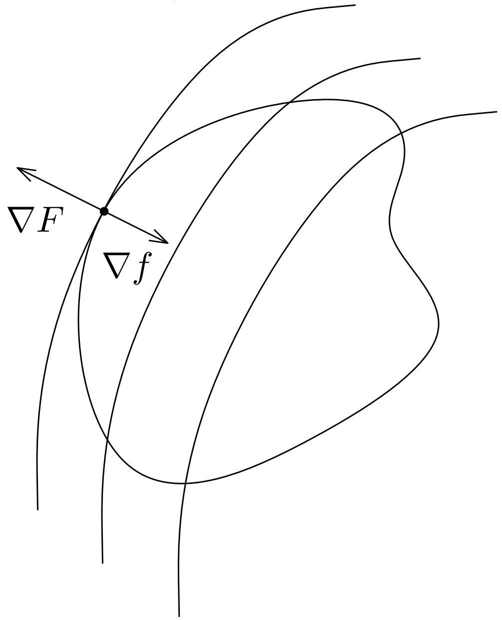
Чаще всего M задают неявно — «накладывают условия» на значения f.

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

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

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.*

E.g. 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 $\exists\,\lambda_1,\dots,\lambda_m\in\mathbb{R}\colon \nabla f(a)=\sum_i\lambda_i\nabla F_i(a)$*

§ 20 Доказательство теоремы об условном экстремуме

□

1. Пусть $m = n-1$. Будем доказывать от противного. Пусть в a условный max, но $\nabla f(a), \nabla F_1(a), \dots, \nabla F_m(a)$ — ЛНЗ. Рассмотрим $\Phi(x) = (f(x), F_1(x), \dots, F_m(x))$. Тогда такое $\Phi\colon G \to \mathbb{R}^n$. 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 теоеме об обратном отображении (0.13.1)
$$\exists\, U(a), V(b) \,:\ \Phi\colon U \to V \text{ — диффеоморфизмъ}$$
Пусть $V \supset B_\varepsilon(b)$, $y = (f(a) + \frac{\varepsilon}{2}, 0, \dots, 0) \in V$, тогда $\exists!\, x \in U\colon \Phi(x) = y$. 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 $\nabla f(a), \nabla F_1(a), \dots, \nabla F_m(a)$ — ЛНЗ. Тогда $\operatorname{rk} \Phi'(a) = m + 1 < n$. Добавим ещё функций $F_{m+1}, \dots, F_{n-1}$ таких, что $F_i(x) = x_{i+1} - ai + 1$.

   Введём ещё стандартное обозначение
$$x = (\underbrace{x, \dots, x_{m+1}}_{\bar{x}}, \underbrace{x_{m+2}, \dots, x_n}_{\bar{\bar{x}}})$$
   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
$$\Phi'(a) = \left(\begin{array}{c|c} A & * \\ \hline 0 & E_{n-m-1} \end{array}\right) \Rightarrow \operatorname{rk} \Phi'(a) = n$$
   А теперь можно подвести всё к первому пункту. Рассмотрим $\widetilde{M} = \{x \mid F_1(x) = \dots = F_{n-1}(x) = 0\}$ (а $M = \{x \mid F_1(x) = \dots = F_m(x) = 0\}$). Поскольку $\widetilde{M} \subset M$, f будет иметь в a максимум и относительно $\widetilde{M}$.

   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 Лагража.