Матанализ. Конспект 2 сем.

Мастера Конспектов

(по материалам лекций Белова Ю. С.,
а также других источников)

16 февраля 2021 г.

Некоторые записи по матанализу.

Содержание

1 Лекция 1.

В этом чеместре мы будем занимать анализом функций от многих переменных, то есть, $f : \mathbb{R}^n \to \mathbb{R}^m$, и если $m = 1$, то такая функция называется функцией многих переменных.

Определение 1. *Кривые в* $\mathbb{R}^n$ - непрерывное отображение $f : [a, b] \to \mathbb{R}^n$.

Основная проблема состоит в том, что образ может выглядеть очень и очень сложно, потому нам хотелось бы более точно понять, как всё это устроено. Потому начнём рассматривать *спрямляемые кривые*, то есть, кривые с конечной длиной. Введём следующее определение:

Определение 2. *Вариация функции* - $V_f([a, b]) = \sup_{a=x_0 \le x_2 \le \ldots \le x_n = b} \sum_{k=0} |f(x_{k+1}) - f(x_k)|$.

$(x - y)$ - евклидово расстояние.

Утверждение 1. Если $f : \mathbb{R} \to \mathbb{R}$ монотонна, то $V_f([a, b]) = |f(a) - f(b)|$.

Утверждение 2. $V_f([a, b]) = 0 \Leftrightarrow f = \text{const}$.

Утверждение 3. $V_{f+g} \le V_f + V_g$.

Утверждение 4. V_f аддитивна на промежутке: $a \le b \le c$, тогда $V_f([a, c]) = V_f([a, b]) + V_f([b, c])$.

Определение 3. Вариация ограничена, если $V_f < \infty$ на $[a, b]$.

Лемма 1.

- *$\mathbb{R} \to \mathbb{R}$, f_1 и f_2 монотонны, тогда $f_1 - f_2$ имеют ограниченную вариацию.*
- *f имеет ограниченную вариацию тогда и только тогда, когда $f = f_1 - f_2$ на отрезке $[a, b]$, причём эти две функции монотонно возрастают.*

Доказательство. Пусть у нас есть f, а также $V_f([a, b]) < \infty$. Рассмотрим $\varphi(x) = V_f([a, x])$. φ определа корректно, причём возрастает. $f = \varphi - (\varphi - f)$, скажем, что $(\varphi - f) = h$, тогда $h(x) \le h(y)$ при $x \le y$. Но это нетрудно показать, $\varphi(x) - f(x) \le \varphi(y) - f(y)$ равносильно $f(y) - f(x) \le \varphi(y - \varphi(x)) = V_f([x, y])$. □

По сути, если понимать определение вариации геометрически, то это просто длина кривой на отрезке. Перейдём теперь к способам обхода кривой.

Лемма 2. *Пусть* $g : [a, b] \to [c, d]$ *- непрерывная биекция (тогда и монотонная). Тогда* $V_f[c, d] = V_{f \circ g}([a,b])$.

Доказательство. Левая и правая части равны соответственно $\sup \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)|$ и $\sup \sum_{k=0}^{n-1} |f(g(y_{k+1})) - f(g(y_k))|$. Это, очевидно, одно и то же. □

Теперь стоит задаться вопросом: а когда же это V_f (или же, длину кривой) можно посчитать. Если f - гладкая функция (гладкая покоординатно f_k). $f := [a, b] \to \mathbb{R}^n$, $f = (f_1, \ldots, f_n)$, $f_k : [a, b] \to \mathbb{R}$. Тогда

$$V_f([a, b]) = \int_a^b \sqrt{(f_1')^2(x) + \ldots + (f_n')^2(x)}\, dx.$$

Рассмотрим

$$\sup_{a=x_0,\dots,x_n=b}\sum_{k=0}^{n-1}\sqrt{(f_1(x_{k+1})-f_1(x_k))^2+\ldots+f_n(x_{k+1})-f_n(x_k))^2}=$$
$$=\sum_{k=0}^{n-1}(x_k+1-x_k)\sqrt{f_1'^2(\xi_{1,k})+\ldots+f_n'^2(\xi_{n,k}))}$$

Если f_i непрерывна, то f_i^2 равномерно непрерывна. $f_i'^2(\xi_{i,k})\le\min_{[x_k,x_{k+1}]}f_i'^2+\varepsilon^2$ (для достаточно мелких разбиений и любого эпсилон, большего нуля). Тогда можно получить верхнюю оценку: $\le\sum_{k=0}^{n-1}(x_{k+1}-x_k)\sqrt{\sum_{l=1}^{n}\min_{[x_k]}(f_l'^2)}+\varepsilon\sqrt{n}(b-a)\le\int_a^b\sqrt{\ldots}+\varepsilon\sqrt{n}(b-a)$ (устремляем разбиение к бесконечно малому). А затем делаем аналогично снизу и получаем требуемое равенство.

2 Лекция 2.

Пусть φ - функция, которая определялась на прошлой лекции, а ψ - обратная ей. ψ - биекция, рассматриваем $f\circ\psi$. Посмотрим на $\psi([0,\beta])=[a,b]$, тогда для любых $c,d\in[0,b]$ $V_{f\circ\psi}([c,d])=d-c$.

Естественная параметризация гладкого пути практически не отличается от того, что мы уже рассматривали за одним небольшим исключением.

$$\varphi(x)=V_f([a,x])=\int_a^x|f'(s)|ds=\int_a^x\sqrt{f_1'^2+\ldots+f_n'^2}ds,$$

причём предпоследнее вырежение равно $|(f_1',\ldots,f_n')|$, а под корнем все функции от s. Рассмотрим опять ψ, и как выглядит вектор $f(\psi(x))=(f_1(\psi(x)),\ldots,f_n(\psi(x)))$, рассмотрим его производную, берём покоординатно: $f'(\psi(x))=(f_1'(\psi(x)),\ldots,f_n'(\psi(x)))$. Но $\psi'(x)=\frac{1}{\varphi(\psi(x))}$, тогда $\varphi(s)=|f'(s)|$, а также $|f'(\psi(x))|=1$.

Примечание 1. Если f - гладкая на $[a,b)$ и существует $\int_a^b|f'(s)|ds$, тогда выполнено то же самое, просто $\varphi(x)=\int_a^x|f'(s)|ds$.

Перейдём теперь к тригонометрии. Рассмотрим окружность $x^2+y^2=1$, мы планируем её обходить (то есть, через каждую точку по разу, с одинаковой скоростью, и так далее). Введём попутно также комплексное обозначение (мы не будем заниматься комплексным анализом, просто это удобно). Отождествим $\mathbb{R}^2$ с $\mathbb{C}$ понятно каким образом. Тогда какое вращение мы хотим? Мы хотим найти функцию $\Gamma:\mathbb{R}\to\mathbb{T}=\{z:|z|=1\text{ или }x^2+y^2=1, z=x+iy\}$, а хотим потребовать также следующее:

- $\Gamma\in C^1$ (гладкая),
- $\Gamma(0)=1$, $\Gamma'(0)=i$ (место старта и начальная скорость, с которой мы идём),
- $|\Gamma'(t)|=1$ для любого t (постоянная скорость 1).

Сформулируем теорему:

Теорема 1. *Функция с данными свойствами существует и единственна.*

Доказательство. $\Gamma(t) \in \mathbb{T}$ тогда и только тогда, когда $\Gamma(t)\overline{\Gamma(t)} = 1$. Продифференцируем последнее, получим

$$\Gamma'(t)\overline{\Gamma(t)} + \Gamma(t)\overline{\Gamma'(t)} = 0,$$

что также равно

$$2\operatorname{Re}(\overline{\Gamma'(t)}\Gamma(t)) = 0.$$

То есть, мы получили, что $\Gamma(t)\overline{\Gamma'(t)} = ih(t)$, $h(t) \in \mathbb{R}$. Применим теперь оставшееся неиспользованное условие: $|\Gamma(t)| = 1$, а чтобы параметризация была естественна, $|\Gamma(t|$ должно быть равно 1. То есть, $h(t) = \pm 1$. Подставим теперь нуль и получим, что функция в этой точке должна быть равна единице, а производная - i. Тогда остаётся один вариант: $h(t) \equiv 1$.

Посмотрим теперь ещё раз на начальные уравнение: $\Gamma'(t)\overline{\Gamma(t)} \equiv i$, то есть,

$$\Gamma'(t) = i\Gamma(t). \tag{1}$$

Таким образом, мы уже пришли к тому, что если вращение существует, то оно должно удовлетворять последнему уравнению, а также $\Gamma(0) = 1$. Это означает, что вращение, которое мы получаем, будет дифференцируемо бесконечно много раз.

Пока что, казалось бы, ни единственности, ни существования, однако из последних утверждений легко получается единственность. Пусть у нас есть $\Gamma_{1,2}$ - два простых вращения. Днначит, они оба удовлетворяют (1). Тогда завайте запишем их частное через сопряжённые и возьмём производную: $\left(\Gamma_1(t)\overline{\Gamma_2(t)}\right)' = \Gamma_1'(t)\overline{\Gamma_2(t)} + \Gamma_1(1)\overline{\Gamma_2'(t)}$, что равно $i\Gamma_1\overline{\Gamma_2} + \Gamma_1\overline{i\Gamma_2} = 0$.

Таким образом, мы получили, что $\Gamma_1\Gamma_2 = \mathrm{const}$, но поскольку $\Gamma(0) = 1$, то эта константа и равна единице. То есть, $\Gamma_1\overline{\Gamma_2} = 1$, следовательно, эти функции равны, единственность доказана.

Докажем теперь существование. Предъявим сначала произвольную параметрицацию окружности, а затем постараемся сделать в ней замену переменной, чтобы получить хорошую функцию (которая должна быть, конечно, гладкой). Давайте параметризуем верхнюю половину $\mathbb{T}$ самым естественным образом: примем $x = t$, $y = \sqrt{1-t^2}$, $-1 \le t \le 1$ (двигаемся по часовой стрелке). Теперь нам нужно отпараметризовать нижнюю половину, возьмём для этого $x = -t$, $y = -\sqrt{1-t^2}$, $-1 \le t \le 1$, двигаться мы теперь будем по нижней половине, но в другом направлении, то есть, одну из половин нужно перевернутьт и "склеить" в один целостный проход. Тогда в нижней половине "сдвинем" рассмотрение на $1 \le t \le 3$, и преобразуем: $y = -\sqrt{1-(2-t)^2}$.

Осталось проверить, что полученная функция гладкая. Вообще, это почти везде очевидно, кроме ± 1, это и проверим. $f(t) = (t, \sqrt{1-t^2})$, а вектор $f'(t) = (1, \frac{-t}{\sqrt{1-t^2}})$. Функция $\varphi(x)$ на $(-1, 1)$ выглядит как

$$\int_{-1}^{x} |f'(s)| ds = \int_{-1}^{x} \sqrt{1 + \frac{t^2}{1-t^2}} dt = \int_{-} 1^x \frac{dt}{\sqrt{1-t^2}}.$$

Функция $\varphi(x)$ - возрастающая биекция, значит, мы можем посмотреть на обратную функцию $\psi(x) = \varphi^-1(x)$. Рассмотрим теперь для $x \in (-1, 1)$,

$$(f^{-1}(\psi(x)))' = (f_1'(\psi(x))\psi'(x), f_2'(\psi(x))\psi'(x)).$$

Тогда, так как $\psi'(x) = \frac{1}{\varphi'(\varphi(x))}$, это также и равно $\sqrt{1-\psi^2(x)}$, что также равно

$$(\psi'(x), \frac{-\psi(x)}{\sqrt{1-\psi(x)}}\sqrt{1-\psi^2(x)}).$$

В последнем также можно сократить числитель и знаменатель. Итого, $f(\psi(x))$ - гладкая на $(-1,1)$, и более того, если $x \to \pm 1$, производная имеет конечный предел. Получается, дифференцируема на интервале, и производная имеет предел в крайних точках, тогда она в них также дифференцируема. Таким образом, для верхней половины мы всё показали, для нижней - аналогично, всего лишь с линейной заменой.

$\square$

После доказательства теоремы, можно, наконец, ввести определения:

Определение 4.

$$\cos(x) = \mathrm{Re}(\Gamma(x)),$$

$$\sin(x) = \mathrm{Im}(\Gamma(x)).$$

Далее уже можно поговорить о бесконечной дифференцируемомти и формуле Муавра, этим, вместе с доказательством, что мы нашли привычные функции, мы, кажется, и планируем заниматься далее.

3 Лекция 3.

Для начала, закончим с тригонометрией. Мы научились строить синус и косинус через вращение окружности. Немного не помню, обговаривали ли мы это на прошлой лекции, но Юрий Сергеевич кратуо цпомянул, что мы можем разложить $\Gamma(x)$ в ряд Тэйлора в $\sum_{n=0}^{\infty} \frac{(ix)^n}{n!}$ в силу свойства $\Gamma'(x) = i\Gamma(x)$ и того, что остаточный член в форме Лагранжа будет стремиться к нулю при стремлении n к бесконечности.

Тогда

$$\cos x = \mathrm{Re}\,\Gamma(x) \Rightarrow \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}(-1)^n$$

и аналогично синус по нечётным степеням.

Мнимая экспонента обладает свойствами, аналогичным обыкновенной экспоненте, поэтому покажем, что $\Gamma(x+y) = \Gamma(x)\Gamma(y)$. Рассмотрим $\Gamma(x+y)\overline{\Gamma(y)}$ - функцию от x, а y - параметр. Это - некоторый обход окружности, который также удовлетворяет всем нормировочным условиям. $\varphi(0) = 1$, $|\varphi'(x| = 1$, и, наконец, $\varphi'(0) = \Gamma'(0) = i$.

Теперь все прекрасные формулы косинуса и синуса суммы и разностей легко выводятся из доказанной формулы. Через мнимую экспоненту запишем: $e^{i(x+y)=e^{ix}\cdot e^{iy}}$, а там уже просто надо посмотреть на мнимые и действительные части.

Из полученных свойств получим, что $\Gamma(x)\Gamma(-x) = \Gamma(0 = 1)$, тогда $\Gamma(-x) = \overline{\Gamma(x)}$, откуда мы получаем чётность косинуса и нечётность синуса.

Можно упомянуть и формулу муавра. Распишем

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2},\ \sin(x) = \frac{e^{ix} - e^{-ix}}{2i},$$

это формулы Муавра. Также можно получить и периодичность, это, вообщем-то очевидно и завершает наш разговор об элементарных функциях.

Перейдём теперь к многочерному анализу. Мы бы хотели точно также уметь анализировать функции и делать всё то, что мы уже умеем делать для одномерных функций, в том числе, решать экстремальные задачи. Нас интересуют функции $f : \mathbb{R}^m \to \mathbb{R}^n$.

Начнём с того, что в евклидовом пространстве $\mathbb{R}^m$ расстояние задаётся как

$$d(x,y) = \sqrt{\sum_{k=1}^{m}(x_k - y_k)^2} = ||x-y||.$$

И если у нас имеется точка $x = (x_1, \dots, x_m)$, то её норма есть $||x|| = \sqrt{\sum_{k=1}^{m} x_k^2}$. Вообще, норму можно задать как угодно, если она удовлетворяет таким свойствам:

- норма - функция $\mathbb{R}^m \to \mathbb{R}_{+,0}$,
- $||x|| = 0 \Leftrightarrow x \equiv (0, \dots, 0)$,
- $||\alpha x|| = |\alpha| \cdot ||x||$, $\alpha \in \mathbb{R}$,
- $||x+y|| \leq ||x|| + ||y||$.

Разберёмся с понятием *гладкости.* Для начала, алгебраически. Пусть у нас есть функция нескольких переменных $f : \mathbb{R}^m \to \mathbb{R}$, $f(x_1, \dots, x_m)$.

Определение 5. f *дифференцируема* в точке $(x_1, \dots, x_m)$, если $f(y) = f(x) + L(y-x) + o(||x-y||)$, где L - линейное отображение $\mathbb{R}^m \to \mathbb{R}$, причём однородное, то есть, $L(0) = 0$.

Определение 6. Это линейное отображение L называется *дифференциалом* в точке x.

На топологии мы доказывали, что в конечномерном пространстве различные норма липшицево-эквивалентны, потому мы просто во всех рассуждениях будем использовать именно евклидовы нормы, потому что они удобные. А теперь перейдём к базовым свойствам.

Примечание 2. L - единственно.

Примечание 3. Если у нас есть две функции: f и g, то дифференциал $\alpha f + \beta g$, α, $\beta \in \mathbb{R}$ есть $\alpha L_1 + \beta L_2$, где L_1 и L_2 - дифференциалы f и g.

Рассмотрим теперь отображение общего вида: $f : \mathbb{R}^m \to \mathbb{R}^n$. Тогда

Определение 7. (Гладкость). $f(y) = f(x) + L(y-x) + o(||x-y||)$, где L - линейное отображение $\mathbb{R}^m \to \mathbb{R}^n$, $L(x+y) = L(x) + L(y)$. o-малое в данном случае можно понять как

$$\frac{f(y) - f(x) - L(y-x)}{||y-x||} \to 0,$$

то есть, элемент $\mathbb{R}^n$ стремится у нулю, но для удобства можно взять евклидову норму этого выражения.

Какой вид имеет общее линейное отображение из $\mathbb{R}^m \to \mathbb{R}^n$? Естественно, это - матрица, это мы знаем из алгебры и умеем расписывать переход в тривиальном базисе.

Перейдём к свойствам линейных отображений. Мы умеем их складывать, умножать, а также, совершать композиции в случае согласованности размерностей, которая соответствует перемножению матриц.

Пусть теперь, опять же, у нас есть отображение $L : \mathbb{R}^m \to \mathbb{R}^n$, то $L(\mathbb{R}^m) \subset \mathbb{R}^n$ - подпространство, которое имеет размерность от 0 до n, эту размерность мы понимаем как *ранг* линейного отображения. Если же мы берём композицию линейных отображений, то ранг не может вырасти (куда растягивать-то). Также, легоко видеть, что если $m < n$, то $\dim(L(\mathbb{R}^m)) \leq m < n$.

Зададимся теперь вопросом, какая существует естественная метрика на линейных отображениях $\mathbb{R}^m \to \mathbb{R}^n$. По сути, эти линейные отображения представляют собой евклидово пространство размерности $m \cdot n$. Задать на нём мы можем евклидову метрику: под корнем будут квадраты всех матричных элементов. Эта норма вычисляется проще, но зато гораздо менее естественна, чем следующая (например, относительно вопроса о композиции). $||L|| = \sup_{||x||<1} ||Lx||$, $x \in \mathbb{R}^m$, $LX \in \mathbb{R}^n$. Эта вещь конечна, так как она не превосходит $\sum_{k=1}^m ||Le_k||$, а также выполняются все свойства нормы.

Геометрический смысл у данной нормы очень простой: мы смотрим, насколько сильно она растягивает расстояние в зависимости от направления.

Завершаем лекцию несколькими переопределениями нормы:

- $\sup_{||x||<1} ||Lx||$,
- $\sup_{||x||\leq 1} ||Lx||$,
- $\sup_{||x||\neq 0} \frac{||Lx||}{||x||}$,
- $\sup_{||x||<=} ||Lx||$.

4 Лекция 4.

Продолжаем с операторами, пусть $A : \mathbb{R}^n \to \mathbb{R}^n$ - линейный, $||A|| = \sup_{||x||\leq 1} ||Ax||$ - норма, где $||x||$ - Евклидово. $A \cong \mathbb{R}^{nm}$, так как можно выносить константу, не меньше нуля (притом равна тогда и только тогда, когда сам оператор - нуль), а также, норма суммы не превосходит сумму норм.

Определение 8. $||A||$ - *операторная норма*, притом супремум всегда достигается.

Операторная норма есть самое большое по модулю собственное число. Предположим, что у A есть n различных λ_i собственных чисел, у которых есть соответственные x^i собственные векторы. Запишем тогда $x = \sum_{k=1}^n a_k x^k$, $Ax = \sum_{k=1}^n \lambda_k a_k x^k$, тогда $||Ax|| \leq \max_k |\lambda_k| \cdot ||x||$, но это мы объяснить не смогли.

Однако разговор сейчас шёл о различных собственных числах, бывают же *кратные* собственные числа. Что происходит?

Важный момент, почему важна операторная норма. Пусть $A : \mathbb{R}^n \to \mathbb{R}^m$, $B : \mathbb{R}^m \to \mathbb{R}^k$, тогда $||BA|| \leq ||B|| \cdot ||A||$, так как левая часть по определению равна $\sup_{||x||\leq 1(\text{в } \mathbb{R}^n)} \leq \sup_{||y||\leq ||A||} ||By|| \leq ||B|| \cdot ||A||$. Заметим также две следующие вещи для линейного $A : \mathbb{R}^n \to \mathbb{R}^m$ равносильны:

- $\ker A = \{0\}$
- $||Ax|| \geq \varepsilon ||x||, \exists \varepsilon > 0$.

Доказательство. $\{x : ||x|| = 1\}$ - единичная сфера в $\mathbb{R}^n$. Пусть $f(x) : x \to ||Ax||$, f - непрерывная (?), $f \neq 0$ на единичной сфере, тогда $f \geq \varepsilon > 0$, $||Ax|| \geq \varepsilon ||x||$, $||x|| = 1$. $\square$

Вообще, нам все эти операторы нужны для рассуждений о гладкости, сформулируем теорему:

Теорема 2. *$f : G \to \mathbb{R}^m$, $G \subset \mathbb{R}^n$ - открытое, f - гладкая в окрестности x^0 (верхние индексы), $y^0 = f(x^0)$, $g : V_{f(x^0)} \to \mathbb{R}^k$, гладкая в $f(x^0)$, для f и g существуют линейные операторы A $(x_0 0)$ и B $(f(x_0))$. Тогда $g(f(x))$ - гладкое (?) отображение в x_0 с линейным оператором (?) $BA : \mathbb{R}^n \to \mathbb{R}^k$.*

Доказательство. Мы знаем, что существует представление $f(x) = f(x^0)+A(x-x^0)+o(||x-x^0||)$. Применим g, получим

$$g(f(x)) = g(y^0 + A(x - x^0) + o(||x - x^0||)). \tag{2}$$

Также мы знаем, что g гладкая, то есть, также представима в виде $g(y) = g(y^0) + B(y - y^0) + o(||y - y^0||)$, тогда приняв аргумент правой части (1) за y, получим продолжение тождества:

$$\begin{aligned} g(y^0) + B(A(x - x^0) + o(||x - x^0||)) + o(A(x - x^0) + o(||x - x^0||)) = \\ g(y^0) + BA(x - x^0) + o(||x - x^0||). \end{aligned} \tag{3}$$

□

Нам много чего хочется от анализа многих переменных, но тут всё, конечно, гораздо сложнее. Перейдём к *частным производным.*

Примечание 4. $f : \mathbb{R}^n \to \mathbb{R}^m$ - гладкая в x^0 тогда и только тогда, когда при записи $(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$ f_k - гладкая $\mathbb{R}^n \to \mathbb{R}$ для всех k (можно написать доказательство).

Определение 9. *Частная производная.* Пусть имеется $f : \mathbb{R} \to \mathbb{R}$, $f(x_1, \ldots, x_n)$, $x^0 = (x_1^0, \ldots, x_n^0)$. Тогда частная производная по x_k, $f(x_1^0, \ldots, x_{k-1}^0, x, x_{k+1}^0, \ldots, x_m^0) = g(x), g'(x_k^0)$. $\left.\frac{\partial f}{\partial x_k}\right|_{x^0} := g'(x_k^0) = \lim_{\varepsilon \to 0} \frac{f(\ldots, x_k^0 + \varepsilon, \ldots) - f(\ldots)}{\varepsilon}$.

Рассмотрим теперь *производную по направлению.* Пусть направление задаётся $e \in \mathbb{R}^n$, $||e|| = 1$, f - дифференцируема по направлению e, если $g(t) = f(x^0 + te)$, $t \in \mathbb{R}$ и существует $g'(0)$, то производная по направлению e - $g'(0) = \lim_{t \to 0} \frac{f(x^0+te)-f(x^0)}{t}$.

5 Лекция 4.

Введём несколько дополнительных терминологий. Пусть у нас есть отображение $f : \mathbb{R}^n \to \mathbb{R}^m$, $f = (f_1, \ldots, f_m)$, $\frac{\partial f_k}{\partial x_l}$, $1 \le k \le m$, $1 \le l \le n$, тогда *матрица Якоби* выглядит как

Теорема 3. *Пусть у нас есть отображение $f : \mathbb{R}^n \to \mathbb{R}^m$, $V_{x_0} \to \mathbb{R}^m$, причём существуют все частные производные в V_{x^0} и они непрерывные в x^0. Тогда f дифференцируема в точке x^0.*

Доказательство. Для начала, мы можем полагать, что $m = 1$, так как можно доказывать, по сути, покомпонентно. Пусть $x^0 = (x_1^0, x_2^0, x_3^0)$ (докажем для 3, потом обсудим общий случай), ну а $x = (x_1, x_2, x_3)$. Нас интересует $f(x_1, x_2, x_3) - f(x_1^0, x_2^0, x_3^0)$. Действуем стандартным образом, будем двигать координаты по одной (так как все сразу двигать не можем). Меняя по одной координате, представим разности из частных производных. Разность равна

$$f(x_1, x_2, x_3) - f(x_1^0, x_2, x_3) + f(x_1^0, x_2, x_3) - f(x_1^0, x_2^0, x_3^0)$$

Разбивается в две подряд идещие разности, достаточно удобные, но последняя всё равно "не айс":

$$f(x_1,x_2,x_3)-f(x_1^0,x_2,x_3)+f(x_1^0,x_2,x_3)-f(x_1^0,x_2^0,x_3)+f(x_1^0,x_2^0,x_3)-f(x_1^0,x_2^0,x_3^0)$$

Теперь уже три удобные разности, так и запишем равенство далее:

$$\begin{aligned}
&=\frac{\partial f}{\partial x_1}\bigg|_{(\xi_1,x_2,x_3)_{\xi\in[x_1^0,x_1]}}(x_1-x_1^0)+\frac{\partial f}{\partial x_2}\bigg|_{(x_1,\xi_2,x_3)}(x_2-x_2^0)+\frac{\partial f}{\partial x_3}\bigg|_{(x_1,x_2,\xi_3)}(x_3-x_3^0)=\\
&=\frac{\partial f}{\partial x_1}\bigg|_{(x_1^0,x_2^0,x_3^0)}(x_1-x_1^0)+\frac{\partial f}{\partial x_2}\bigg|_{(x_1^0,x_2^0,x_3^0)}(x_2-x_2^0)+\frac{\partial f}{\partial x_3}\bigg|_{(x_1^0,x_2^0,x_3^0)}(x_3-x_3^0)+\\
&+\left(\frac{\partial f}{\partial x_1}\bigg|-\frac{\partial f}{\partial x_1}\right)(x_1-x_1^0)+\left(\frac{\partial f}{\partial x_2}\bigg|-\frac{\partial f}{\partial x_2}\right)(x_2-x_2^0)+\left(\frac{\partial f}{\partial x_3}\bigg|-\frac{\partial f}{\partial x_3}\right)(x_3-x_3^0)
\end{aligned}$$

Последние три слагаемых - остаток, $R(x)$, тогда $\forall\varepsilon>0$, $\exists\delta$ $||x-x^0||<\delta$, $|R(x)|<\varepsilon||x-x^0||$.

□

Теорема 4. *Пусть $f:\mathbb{R}^n\to\mathbb{R}$, f - гладкая на G - открытое(???). Я нихуя не могу прочитать, что тут написано.*

Доказательство. $\operatorname{grad} f=(\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_n})$, f - локальный максимум f, x^0, $\frac{\partial f}{\partial x_k}\Big|_{x_0}\neq 0$.

$$\begin{aligned}(x_1^0,\ldots,x_{k-1}^0,x,x_{k+1}^0,\ldots,x_n^0)-(x_1^0,\ldots,x_m^0)=\\
\frac{\partial f}{\partial x_k}\bigg|_{x^0}(x_k-x_k^0)+o(|x_k-x_k^0|),\end{aligned}$$

причём первое слагаемое не нуль. □

Нам бы ещё хотелось иметь теорему об обратном отображении.

Теорема 5. *(Об обратном отображении). Пусть $f:G\to\mathbb{R}^n$, G - открыто в $\mathbb{R}^n$, у f есть гладкие частные производные (???), f - в точке x^0 дифференцируема A, A - (сука??????????). Тогда V_{x^0} $\exists g$ - гладкая, (?????) $f(x^0)$, $g(f(x))=x$, g - дифференцируема в $f(x^0)\Rightarrow A^{-1}$.*

Доказательство. $(f^{-1}(x))'=\frac{1}{f'(f^{-1}(x))}$ - вспомнили, а теперь - к доказательству.

Утв. 1. f - гладкая в окрестности точки x^0 с непрерывными частными производными, тогда f липшицева, то есть, $|f(x)-f(y)|\leq C||x-y||$. Если мы зафиксируем точку x, то $|f(x)-f(y)|\leq(||A||+\varepsilon)||x-y||$ $\forall\varepsilon>0$, $A=A_x$. $||A_x||\leq\sum_{k=1}^n\left|\frac{\partial f}{\partial x_k}\Big|_x\right|$.

Утв. 2. Если к тому же $\operatorname{Ker}(A)=\{0\}$, то f - билипшицево (в окрестности x^0), $C_2||x-y||\leq|f(x)-f(y)|\leq C_1||x-y||$. Докажем и его. $f(y)=f(x)+A_x(y-x)+o(||x-y||)$, тогда $||A_{x^0}z||\geq\varepsilon||z||$, $\forall z\in\mathbb{R}^n$, $||A_xz||$ - $A_xz=A_{x^0}z+(A_x-A_{x^0})z$. Первый элемент не меньше $\varepsilon||z||$, а $||A_x-A_{x^0}||$ стремится к 0 в окрестности этой точки, тогда

$$|f(y)-f(x)|=|A_x(y-x)+o(||x-y||),$$

но каждый из них можно ограничить снизу $\frac{\varepsilon}{2}||x-y||$.

Тогда $f : \mathbb{R}^n \to \mathbb{R}^m$, $\operatorname{Ker} A = \{0\}$, тогда $n \leq m$.

Итого, у нас есть отображение $f : f(x) = f(x^0) + A(x - x^0) + o(||x - x^0||)$. Рассмотрим шарик $B_r(x^0) = \{x : ||x - x^0|| < r\}$, $f(B_r(x^0))$, f - биективна. Проверим, что он содержится в каком-то $B_{r'}(f(x^0))$.

Утв. 3. В условиях теоремы для любого r существует r', $f(\overline{B_r(x^0)}) \supset \overline{B_{r'}(f(x^0))}$. Для любого $y \in B_{r'}(f(x^0))$ $f(x) = y$, хотим найти x. $F(x) = ||f(x) - y||^2$, гладкая в окрестности x^0. Минимум этой функции где-то достигается (непрерывная на компакте). $F(x^0) = ||f(x^0) - y||^2 \leq r'^2$, тогда минимум не может достигаться на границе, так как иначе $||x - x^0|| = r$. Тогда с одной стороны $||f(x) - y||^2 = ||f(x) - f(x^0) + f(x^0) - y||^2$. f билипшицева, поэтому разность первых двух можно оценить чнизу $\varepsilon||x - x^0||$, а разность последних двух можно ограничить сверху r'^2, то есть, вся эта вещь как минимум r'^2.

Пусть w - минимум $F(x)$ на $B_r(x^0)$, тогда $\operatorname{grad} F(w) = 0$, $= ||f(x) - y||^2 = \sum_{k=1}^n (f_k(x) - y_k)^2$,

$$\left.\frac{\partial F}{\partial x_l}\right|_w = \sum_{k=1}^n \left.\frac{\partial f_k}{\partial x_l}\right|_w 2(f_k(x) - y_k),$$

Ну под конец не успел, слишком долго расшифровывать. □