Матанализ. Конспект 2 сем.

Кабашный Иван (@keba4ok)

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

$(x - y)$ - 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. $V_{f+g} \leq V_f + V_g$.

Утверждение 4. V_f аддитивна на промежутке: $a \leq b \leq c$, 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 $h(x) \leq h(y)$ при $x \leq y$. Но это нетрудно показать, $\varphi(x) - f(x) \leq \varphi(y) - f(y)$ равносильно $f(y) - f(x) \leq \varphi(y - \varphi(x)) = V_f([x, y])$. $\square$

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. $\square$

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

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

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 $\Gamma:\mathbb{R}\to\mathbb{T}=\{z:|z|=1$ 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

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

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. Днačит, 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 $f \geq \varepsilon > 0$, $||Ax|| \geq \varepsilon ||x||$, $||x|| = 1$. □

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) 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 $(f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$ f_k - гладкая $\mathbb{R}^n \to \mathbb{R}$ для всех k (можно написать доказательство).

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

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 5.

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

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_m} & \dots & \frac{\partial f_m}{\partial x_m} \end{pmatrix}$$

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:

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

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 множестве, причём частные производные существуют и непрерывны в каждой точке (условно говоря, f гладкая). Предположим, что точка x^0 - локальный максимум или минимум. Тогда* $\operatorname{grad} f|_{x^0}\equiv 0$.

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

$$(x_1^0,\dots,x_{k-1}^0,x,x_{k+1}^0,\dots,x_n^0)-(x_1^0,\dots,x_m^0)=$$
$$\left.\frac{\partial f}{\partial x_k}\right|_{x^0}(x_k-x_k^0)+o(|x_k-x_k^0|),$$

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 непрерывные частные производные, f - в точке x^0 имеет дифференциал A, A - невырожденная. Тогда в некоторой V_{x^0} существует обратная функция g - гладкая, задана в как минимум какой-то окрестности $f(x^0)$, $g(f(x))=x$, 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 $|f(x)-f(y)|\leq(||A||+\varepsilon)||x-y||$ $\forall\varepsilon>0$, $A=A_x$. $||A_x||\leq\sum_{k=1}^n\left|\left.\frac{\partial f}{\partial x_k}\right|_x\right|$.

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

$||A_{x^0}z|| \geq \varepsilon||z||$, $\forall z \in \mathbb{R}^n$, $||A_x z||$ - $A_x z = A_{x^0}z + (A_x - A_{x^0})z$. 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 $\operatorname{grad} F(w) = 0$, $= ||f(x) - y||^2 = \sum_{k=1}^n (f_k(x) - y_k)^2$,

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

И теперь, если подставить в обе стороны w, то получатся нули. Посмотрим на правую часть как на СЛУ. Дроби фиксированы - числа из матрицы Якоби (изнач. - A_w), $f_k(x)$ - какие-то неизвестные. Матрица невырождена, так как нврожденность не меняется от приведённого шевеления. Значит, решение этой системы единственно, но одно из решений мы уже знаем: $f_k(w) = y_k$. Поэтому, если мы подставим точку минимума функции F, то окажется, что эта точка переводится как раз в точку y, и тогда $F = 0$, а мы в точности нашли прообраз. $\square$

6 Лекция 6.

Заканчиваем доказательство теоремы об обратном отображении. Мы уже установили, что f - билипшицева, что f в какой-то окрестность $f(V_{x_0})$ содержит V_{g_0}, а также, что обратное отображение g по крайней мере существует в какой-то окрестности.

Осталось лишь доказать один небольшой оставшийся момент. Пусть f - гладкая в x^0, тогда мы можем расписать $f(x) = f(x^0) + A(x - x^0) + R(x)$, где $|R(x)| = o(|x - x^0|)$ (модули над векторами с данного момента - естественно, нормы). Воспользуемся тем, что любая точка y, достаточно близкая к y^0, то, по доказанному ране, у неё есть прообраз. Напишем тогда ввиду прообразов: $y = y^0 + A(g(y) - g(y^0)) + R(g(y))$. Она выполнена для любого y в некоторой окрестности y^0.

Нам хотелось бы выразить $g(y^0)$. Рассмотрим $A(g(y) - g(y^0)) = y - y^0 - R(g(y))$. Это равенство двух векторов $\mathbb{R}^n$, матрица A невырожденная, потому у неё есть обратная, применим это знание: $g(y) - g(y^0) = A^{-1}(y - y^0) - A^{-1}(R(g(y)))$. Надо оценить остаток, так как всё остальное уже хорошее. Из того, что мы уже знаем, $\forall \varepsilon > 0, V^{\varepsilon}_{x^0} |R(x)| \leq \varepsilon|x - x^0|$ (кажется, в некоторой окрестности, поэтому размер окрестности должен быть другой переменной). Тогда $|R(g(y| \leq \varepsilon|g(y) - g(y^0)|$. Мы знаем, что f билипшицева, как и обратная, поэтому продолжаем неравенство $\leq \varepsilon C|y - y^0|$. Но у нас изначально есть $|A^{-1}(R(g(y)))| \leq ||A^{-1}||\varepsilon C|y - y^0|$. И теперь, собирая всё назад, получаем, что $g(y) = g(y^0) + A^{-1}(y - y^0) + o(|y - y^0|)$, это нам и нужно было: дифференцируемость в y^0 и явный дифференциал.

Попытаемся обобщить Формулу Тейлора для многих переменных. Для начала, разберёмся с тем, как дифференцировать композицию функций многих переменных. Пусть $f:\mathbb{R}^n\to\mathbb{R}^m$, $g:\mathbb{R}^m\to\mathbb{R}^k$, f - гладкая в x^0, g - гладкая в $f(x^0)$, тогда $g(f(x^0))$ - гладкая в x^0, дифференциал - BA.

Пример(ы) 1. $f:\mathbb{R}^n\to\mathbb{R}$, $f(x_1,\ldots,x_n)$, $g_k=g_k(x_1,\ldots,x_n)$ $(1\le k\le n)$, $g=(g_1,\ldots,g_n)$. Тогда $f(g_1(x_1,\ldots,x_n),\ldots,g_n(x_1,\ldots,x_n))=F(x_1,\ldots,x_n)$, $\tilde{x}=(x_1,\ldots,x_n)$ и

$$\begin{aligned}\left.\frac{\partial F}{\partial x_1}\right|_{\tilde{x}} &= \lim_{\varepsilon\to0}\frac{f(g_1(x_1+\varepsilon,\ldots),\ldots,g_n(x_1+\varepsilon,\ldots))-f(\ldots)}{\varepsilon}=\\ &=\lim_{\varepsilon\to0}\frac{f(g_1(x_1,\ldots,m_n)+\varepsilon\frac{\partial g}{\partial x_1}+o(\varepsilon),\ldots)-f(g(\ldots),\ldots)}{\varepsilon}=\\ &=\lim\frac{\varepsilon\left.\frac{\partial f}{\partial x_1}\right|_{g_1(x_1,\ldots,x_n)}\cdot\frac{\partial g}{\partial x_1}+\varepsilon\frac{\partial f}{\partial x_2}\frac{\partial g_2}{\partial x_1}+\ldots+\varepsilon\frac{\partial f}{\partial x_n}\frac{\partial g_n}{\partial x_1}+o(\varepsilon)}{\varepsilon}=\\ &=\sum_i\frac{\partial f}{\partial x_i}\frac{\partial g_i}{\partial x_1}=^?\sum_i\frac{\partial f}{\partial g_i}\frac{\partial g_i}{\partial x_1}\end{aligned}$$

Последнее равенство - просто вопрос обозначений. А в итоге у нас просто получается предпоследнее. $f=(\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_n})$, $BA=\left(\sum_{k=1^n}^n\frac{\partial f}{\partial x_k}\cdot\frac{\partial g_k}{\partial x_1},\ldots,\sum_{k=1}^n\frac{\partial f}{\partial x_k}\cdot\frac{\partial g_k}{\partial x_l},\ldots\right)$

Пример(ы) 2. $f(e^{x_1+x_2+x_3},x_1-x_2+x_3,x_1x_3)=F(x_1,x_2,x_3)$. Тогда

$$\frac{\partial F}{\partial x_2}=\frac{\partial f}{\partial x_1}\cdot e^{x_1+x_2+x_3}+\frac{\partial f}{\partial x_2}(-1).$$

Формула Лагранжа. Пусть $f:\mathbb{R}^n\to\mathbb{R}$, $(x_1,\ldots,x_n)$, $(y_1,\ldots,y_n)$, мы хотим научиться как-то выражать $f(y_1,\ldots,y_n)-f(x_1,\ldots,x_n)$ через производную в какой-то точке посередине.

Путь $x(t)=x+th$, $0\le t\le 1$, $h=y-x$ (покоординатно). Тогда $\varphi(t):=f(x+th)=f(x_1+t(y_1-x_1),\ldots,x_n+t(y_n-x_n))=f(x_1+th_1,\ldots,x_n+th_n)$, $f(y)-f(x)=\varphi(1)-\varphi(0)=\varphi'(\xi)$, $\xi\in[0,1]$ (по формуле одной переменной),

$$\varphi'(t)=\left.\frac{\partial f}{\partial x_1}\right|_{x(t)}(y_1-x_1)+\ldots+\left.\frac{\partial f}{\partial x_n}\right|_{x(t)}(y_n-x_n)=(\operatorname{grad}f,y-x).$$

Таким образом, формула Лагранжа: $f(y)-f(x)=(\operatorname{grad}f|_\xi,y-x)$, $\xi\in[x,y]$.

Получим ещё одно крутое обобщение:

$$\varphi'(0)=\left.\frac{\partial f}{\partial x_1}\right|(y_1-x_1)+\ldots+\left.\frac{\partial f}{\partial x_n}\right|(y_n-x_n),$$

А теперь возьмём производную ещё раз.

$$\varphi''(t)=(\varphi'(t))'=\sum_{k=1}^n\frac{\partial f}{\partial x_1\partial x_k}(y_k-x_k)(y_1-x_1)+\ldots=\sum_{q\le k,l\le n}\frac{\partial f}{\partial x_l\partial x_k}\cdot(y_k-x_k)(y_l-x_l)=\sum_{1\le k,l\le n}\frac{\partial f}{\partial x_l\partial x_k}\cdot h_kh_l,$$

такая вещь называется *квадратичной формой*. Теперь для получения формулы Тейлора n-го порядка, нам нужно кучу раз дифференцировать, и будет всё это состоять из всевозможных переборов приращений по различным переменным. Закончилась лекция вспоминаниями о том, что такое квадратичная форма, но за этим лучше перейти в лекции по алгебре.

7 Лекция 7.

Утверждение 5. Пусть $f : G \subset \mathbb{R}^2 \to \mathbb{R}$, $f = f(x, y)$. $(x_0, y_0) \in G$, $\frac{\partial f}{\partial x \partial y}$ и $\frac{\partial^2 f}{\partial y \partial x}$ - существуют в окрестности точки. Тогда $\frac{\partial f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_9)$.

Доказательство. Составим на плоскости прямоугольник со стороной ε в положительные стороны от изначальной точки, расставим в шахматном порядке в углах плюсы и минусы, в изначальной точке - плюс.

Рассмотрим теперь разностную сумма: $f(x_0, y_0) + f(x_0 + \varepsilon_1, y_0 + \varepsilon_2) - f(x_0 + \varepsilon_1, y_0) - f(x_0, y_0 + \varepsilon_2) = \Phi$. Введём некоторые вспомогательные функции: $F_1(y) = f(x_0 + \varepsilon_1, y) - f(x_o, y)$, тогда $\Phi = -F_1(y_0+) + F_1(y_0 + \varepsilon_2) = \varepsilon_2 F_1'(\xi_1)$, $\xi \in [y_0, y_0 + \xi_1]$. $F_1' = \frac{\partial f}{\partial y}(x_0 + \varepsilon_1, y) - \frac{\partial f}{\partial y}(x_0, y) = \varepsilon_1 \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}(\xi_2)\right)$, а это равно $\varepsilon_1 \varepsilon_2 \cdot \frac{\partial f}{\partial x \partial y}(\xi_1, \xi_2)$, ξ_1, ξ_2 лежат в прямоугольничке.

Теперь рассмотрим то же самое, только введём $F_2(x) = f(x, y_0 + \varepsilon_2) - f(x, y_0)$, опять поприменяем теорему Лагранжа, и получим $\varepsilon_1 \varepsilon_2 \frac{\partial^2 f}{\partial x \partial y}(\xi_1, \xi_2) = \varepsilon_1 \varepsilon_2 \frac{\partial^2 f}{\partial y \partial y}(\eta_1, \eta_2)$, а теперь воспользуемся непрерывностью производной, устремим к изначальной точке и получим требуемое. $\square$

Примечание 5. Менять можно любые частные производные в функциях от многих переменных (фиксируем остальные и меняем по лемме).

Перейдём, наконец, к формуле Тейлора для многих переменных. Выглядит она сложно, но идея простая: у нас есть некоторый центр - точка $x \in \mathbb{R}^n$ и ещё она точка $y \in \mathbb{R}^n$, мы тогда пойдём по прямой от второй точки к центра, напишем естественную параметрицазия $x(t) = x + t(y - x)$, $0 \le t \le 1$, $h = (y - x)$. Введём тогда функцию $\varphi(t) = f(x + th)$, попробуем написать для φ обычную формулу Тейлора и подставим h.

$\varphi'(t) = (\operatorname{grad} f, h) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x + th) \cdot h_k$. Запишем теперь вторую производную: $\varphi''(t) = \sum_{1 \le k,l \le n} \frac{\partial^2 f}{\partial x_k \partial x_l}(x + th) h_k h_l$, и, в итоге,

$$\varphi^{(s)}(t) = \sum_{1 \le k_1, \ldots, k_s \le n} \frac{\partial^s f}{\partial x_{k_1} \ldots \partial x_{k_s}} h_{k_1} \ldots h_{k_s} = \left(\frac{\partial}{\partial x_1} h_1 + \ldots + \frac{\partial}{\partial x_n} h_n\right)^s f(x + th).$$

В общем и целом, для написания формулы Тейлора уже всё готово. Остслось вспомнить, как она устроена от одной переменной. $\varphi(\tau) = \varphi(0) + \frac{\varphi'(0)}{1!}\tau + \ldots + \frac{\varphi^m(0)}{m!}\tau^m + R_m(\tau, f)$, тогда если $\tau \equiv 1$, то

$$\varphi(1) = \sum_{k-0}^{m} \frac{\varphi^{(k)}(0)}{k!} + R_m(1, f).$$

Рассмотрим теперь остаток: $R_m(\tau, \varphi) =?$ (остаток надо дописать).

Теорема 6. *Пусть f имеет частные производные до $(m+1)$-ой в некоторой окрестности точки x, $h = y - x$, тогда $f(y) = f(x) + \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x) h_k + \frac{1}{2!} \sum_{1 \le k_1, k_2 \le n} \frac{\partial^2 f}{\partial x_{k_1} \partial x_{k_2}}(x) h_{k_1} h_{k_2} + \ldots + \int_0^1 \frac{(\ldots)(x+th)}{m!}(1 - t)^m dt$, также можно написать в форме Лагранжа или Пеано.*

Определение 10. $f : G \subset \mathbb{R}^n \to \mathbb{R}$, имеет локальный минимум в точке x^0, если $f(y) \ge f(x^0)$, $\forall y \in V_{x^0}$. Аналогично и максимум.

Определение 11. Экстремум: $f : \mathbb{R}^n \to \mathbb{R}$ (дописать надо).

Теорема 7. *Если квадратичная форма* $\sum_{1\le k,l\le n}\frac{\partial^2 f}{\partial x_k \partial x_l}(x^0)h_k h_l$, $\mathrm{grad}(x^0)=0$ *определена положительно, то локальный минимум, если определена отрицательно, то локальный максимум.*

Рассмотрим квадратную матрицу

$$A=\{a_{k,l}\}_{1\le k,l\le n},$$

$Q(x)=\sum_{1\le k,l\le n}a_{kl}x_k x_l$, $x=(x_1,\ldots,x_n)$. $Q(x)=(x_1,\ldots,x_n)A(x_1\ldots x_n)^T$.

Теорема 8. *(Критерий Сильвестра). Пусть у нас имеется матрица* Q, $n\times n$, *тогда* Q *положительно определена тогда и только тогда, когда все угловые определители (два на главной диагонали и квадрат с ними, 4 элемента) не меньше нуля.*

8 Лекция 8.

В прошлый раз мы остановились на расссотрении поля экстремумов $f: G\to\mathbb{R}, G\subset\mathbb{R}^n$. Мы нашли два условия:

- $\mathrm{grad}\, f=0$ (необходимое);
- если (квадратичная?) форма $\sum_{q\le k,l\le n}\frac{\partial^2 f}{\partial x_k\partial x_l}$ положительно определа, то локальный минимум, если отрицательно, то локальный максимум.

Докажем второй пункт. Пусть $x^0=(x_1^0,\ldots,x_n^0)$, распишем формулу Тейлора $f(x)=f(x^0)+\sum_{k=1}^n\frac{\partial f}{\partial x_k}(x_k-x_k^0)+\frac{1}{2}\sum_{1\le k,l\le n}\frac{\partial^2 f}{\partial x_k\partial x_l}(x_k-x_k^0)(x_l-x_l^0)+o((x-x^0)^2)$. Предпоследнее слагаемое - квадратичная форма, пусть A. Тогда запишем

$$(A(x-x^0),(x-x^0))\ge\varepsilon||x-x^0||^2,$$

что следует из положительности функции. $(Ax,x)>0$, $x\ne0$, тогда $(Ax,x)>\varepsilon$ при $||x||=1$, тогда $(Ax,x)\ge\varepsilon||x||^2$.

Квадратичные формы, сопряжённые операторы. Каждая линейная операция $A:\mathbb{R}^n\to\mathbb{R}^n$ задаётся матрицей $n\times n$. Самосопряжённая матрица - матрица, которая удовлетворяет $(Ax,y)=(x,Ay)$, рассмотрим её свойства. Она диагонализируема, потому её собственные вектора это просто столбцы из нулей и единички на одном из мест. Собственные же числа - как раз числа на диагонали после диагонализации.

Пример(ы) 3. $f(x,y)=x^4+y^4-2x^2, \mathbb{R}^2\to\mathbb{R}$. $\frac{\partial f}{\partial x}=4x^3-4x=0$, $\frac{\partial f}{\partial y}=4y^3$. Под подозрением к экстремуму имеются точки $(0,1)$, $(1,0)$, $(-1,0)$. Рассмотрим $\frac{\partial^2 f}{\partial x^2}=12x^2-4$, $\frac{\partial^2 f}{\partial y^2}=12y^2$, $\frac{\partial^2 f}{\partial x\partial y}=0$. Тогда для соответствующих точек мы получаем матрицы 2×2, состоящие из нулей, кроме левого верхнего нуля - (-4), 8 и 8 соответственно. Так как глоюальный минимум должен существовать, то он минимальный по значению в этих трёх точках, а это - две последние пары (-1).

Пример(ы) 4. Рассмотрим $f(x,y)=xy\log(x^2+y^2)$. Зададим $f(0,0)=0$, там с непрерывностью всё хорошо, а с гладкостью, конечно, нет. $\frac{\partial f}{\partial y}=y\log(x^2+y^2)+\frac{2x^2y}{x^2+y^2}$, по иксу симметрично. Приравняем оба выражения к нелю и бедем решать систему. Пусть $x,y\ne0$.

Поделим и получим систему поприятнее, вычтем из одного другое и получим, что $x = \pm y$. Тогда на одну переменную получем $\log(2x^2) + 1 = 0$, тогда $x = \pm\frac{1}{\sqrt{2e}}$, отсюда 4 точки, а также с нулями ещё точки $(0, \pm 1)$, и симметрично.

Как устроена $\frac{\partial^2 f}{\partial x^2} = \frac{2xy}{(x^2+y^2)} + \frac{4xy(x^2+y^2)-4x^3y}{(x^2+y^2)^2}$. Вторая производная по y симметрична. $\frac{\partial^2 f}{\partial x \partial y} = \log(x^2+y^2) + \frac{2y^2}{(x^2+y^2)} + \frac{2x^2}{(x^2+y^2)} - \frac{4x^2y^2}{(x^2+y^2)^2}$. Попытаемся теперь разобраться, что происходит с "подозрительными" точками. С точкой $(0, 0)$ ничего не понятно, посмотрим на остальные. $(1, 0)$ имеет квадратичную форму

$$\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$$

Про точку $(-1, 0)$ и так всё понятно в силу нечётночти функции, она не подходит. Для точки $(0, 2)$ получаем аналогичную матрицу. Посмотрим теперь на $(\frac{1}{\sqrt{2e}}, \frac{1}{\sqrt{2e}})$, получаем матрицу

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Если у одной из координат поменяем знак, то получим

$$\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

Осталась только точка $(0, 0)$, в которой ничего не понятно. Однако посмотрим на изменение знаков при фиксированных значениях и поймём, что она не обладает какими-то интересными свойствами.

Есть у нас в анализе одной переменной касательная к графику функции. Аналогичные вещи могут быть и в срарших размерностях, посмотрим на $\mathbb{R}^3$. Напишем касательную $f(x_0,y_0) + \frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0)$. По аналогичным аппромаксимацонным моображениям, получим, что эта вещь, как и на плоскости, прекрасно аппроксимирует график функции.

Вектор нормали: $(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, -1)$.

Пример(ы) 5. Параметризуем сферу широтой и долготой. Сфера в $\mathbb{R}^3$ - точки, сумма квадратов координат которых равна единице. Ну там по углам параметризуем, в общем, география 0 класс.

9 Лекция 9.

(три минуты начала лекции проспаны)

Теорема 9. *Пусть $F : G \to \mathbb{R}$, $G \subset \mathbb{R}^2$ - открытое. Также выполнено:*

- $F(x_0, y_0) = 0$;
- $F \in C^1(G)$;
- $F'_y(x_0, y_0) \neq 0$.

Тогда $x_0 \in I_x$, $y_0 \in I_y$, $I_x \times I_y \subset \mathbb{R}^2$, $\exists!\ y = f(x)$ такая. что $F(x, y) = 0$ тогда и только тогда, когда $y = f(x)$.

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

Доказательство. Рассмотрим $F(x, f(x)) = 0$, $F'_x(x, f(x)) + F'_y(x, f(x))' f(x) = 0$, это - неформальная мысль о том, почему производная именно такая, какая она и должна быть, а теперь перейдём к доказательству.

Для определённости будем считать, что $F'_y(x_0, y_0) > 0$. Поскольку мы предположили непрерывность проиводных, то $F'_y > 0$ в некоторой окрестности этой точки. Зафискируем одну из координат и посмотрим на функцию $G(y) = F(x_0, y)$, мы знаем, что $G(y_0) = 0$, $G' > 0$, потому G возрастает, а $G(y_1) < 0 < G(y_2)$ при $y_1 < y_0 < y_2$.

Сама функция F является непрерывной функцией двух переменных, потому если мы посмотрим $G_x(y) = F(x, y)$, то G - возрастающая. Давайте обозначим, что G задана (?) на $[y_0 - \beta, y_0 + \beta]$, тогда фиксируя x, $G_x(y_0 - \beta)$, $G_x(y_0 + \beta) > 0$ для $x \in [x_0 - \alpha, x_0 + \alpha]$. В силу того, что функция возрастает, $\forall x \in [x_0 - \alpha, x_0 + \alpha]$, $\exists! y$, $G_x(y) = F(x, y) = 0$.

Рассмотрим $y = f(x)$ и попробуем установить непрерывность f в окрестности точки x_0. Мы знаем, что $y_0 = f(x_0)$, и нам нужно понять, почему при довольно малом сдвиге по x, f также изменится минимально. $y_0 = f(x_0)$. Рассмотрим множество решений уравнения $F(x, y) = 0$ (внутри рассматриваемо прямоугольника). Мы уже доказали, что на каждом вертикальном отрезке (или горизонтальном, смотря как повернуть) есть ровно одна точка, которая удовлетворяет условию, тогда получим противоречие (не услышал в чём).

f непрерывна в любой точке x из $(x_0 - \alpha, x_0 + \alpha)$. Рассмотрим теперь $h \in \mathbb{R}$, а также $F(x, f(x)) = 0$, x - фиксированный. $F(x + h, f(x + h)) = 0$, $h \to 0$. Тогда $F(x + h, f(x + h)) - F(x, f(x)) = F(x_h, f(x) + (f(x + h) - f(x)) - F(x, f(x))) = h(F'_x(x, f(x))) + h \cdot \frac{f(x+h) - f(x)}{h} F'_y(x, f(x)) + o(h) = 0$ - по формуле тейлора при $h \to 0$ (o можно расписать получше). Тогда $\exists \lim_{h \to 0} \frac{f(x+h) - f(x)}{x} = -\frac{F'_x}{F'_y}(x, f(x))$.

Примечание 6. Если $F \in C^k(G)$, то $f \in C^k$, так как дробь первой производной можно как угодно кучу раз дифференцировать.

□

Начнём повышать размерности.

Теорема 10. *Пусть $F : G \in \mathbb{R}$, G - открыто в $\mathbb{R}^{m+1}$, а также выполнено:*

- $F \in C'(G)$;
- $F(x_1, \ldots, x_m, y_0) = 0$;
- $F'_y(x_1^0, \ldots, x_m^0, y_0) \neq 0$.

Тогда $I_x^m \times I_y$, $F(x_1, \ldots, x_m, y) = 0$ тогда и только тогда, когда $y = f(x) = f(x_1, \ldots, x_m)$. Опять-таки, записать можно ткаую же штуку для производной.

Доказательство. Да и доказательство не сильно поменяется, в общем-то. □

Посмотрим теперь, что это нам даёт. Это нам даёт, что поверхность может быть задана неявным образом. Пусть $z = f(x, y)$, $F(x, y, z) = 0$, тогда если мы нашли какую-то точку $F(x_0, y_0, z_0) = 0$, $F'_z(x_0, y_0, z_0) \neq 0$. Тогда $F(x, y, z) = 0$ тогда и только тогда, когда $z = f(x, y)$.

Вектор нормали должен был выглядеть как $(f'_x, f'_y, -1)$, но тут много чего непонятного, запишем как $\left(-\frac{F'_x}{F'_z}, -\frac{F'_y}{F'_z}, -1\right)$, при $F'_z \neq 0$. Тогда домножим и получим $(F'_x, F'_y, F'_z) = \operatorname{grad} F$.

Попробуем начать формулировать теорему о неявной функции. Пусть $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, и F_i(всех переменных) при $1 \leq n$. Тогда если $x^0 = (x_1^0, \ldots, x_m^0)$, $y^0 = (y_1^0, \ldots, y_n^0)$ так что $F_k(x^0, y^0) = 0$ для любого $1 \leq k \leq n$. $F_k(x, y) = 0$ тогда и только тогда, когда $y = f(x)$, $y^0 = f(x^0)$.

Зафиксируем $(x_1, \ldots, x_m)$, Рассмотрим для F_k, $F : \mathbb{R}^n \to \mathbb{R}^n$, тогда дифференциал F - матрица A обратима.

Теорема 11. *$F : G \to \mathbb{R}^n$, где G - открытое в $\mathbb{R}^{m+n}$ и выполнено следующее:*

- *$F \in C^1(G)$;*
- *$F(x^0, y^0) = 0$, $x^0 \in \mathbb{R}^m$, $y^0 \in \mathbb{R}^n$;*
- *$F'_y(x^0, y^0)$ - обратимая матрица.*

Тогда окрестность (x^0, y^0), $\exists!\, f : \mathbb{R}^m \to \mathbb{R}^n$, $F(x, y) = 0$ тогда и только тогда, когда $y = f(x)$.

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

10 Лекция 10.

Продолжаем изучение теоремы о неявной функции.

Теорема 12. *Пусть $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, $F : G \to \mathbb{R}^n$, $G \subset \mathbb{R}^{m_1}$. Тогда если*

- *$F \in C^1(G)$;*
- *$F(x^0, y^0) = 0$;*
- *$F'(x^0, y^0)$ - обратимо ($n \times n$ - матрица).*

Тогда в некоторой окрестности (x^0, y^0), $F(x, y) = 0$ тогда и только тогда, когда $y = f(x)$, $f : \mathbb{R}^m \to \mathbb{R}^n$.

Доказательство. $f'(x) = -(F'_y(x, f(x)))^{-1} F'_x(x, f(x))$. В правой части как раз матрицы $n \times n$ и $n \times m$.

Мы сейчас находимся в положении, что если $n = 1$, то теорема доказана (на прошлой лекции). Докажем теперь всё остальное индукцией по n. Попробуем уменьшить на 1 число переменных. Мы знаем, что $F = (F_1, \ldots, F_n)$, $F_k : \mathbb{R}^{m+1} \to \mathbb{R}$. $F_k = F_k(x^1, \ldots, x^m, y^1, \ldots, y^n)$, и эта вещь равна нулю из кучи соотношений. Нам хочется выделить одну переменную и посмотреть на это как на неявное соотношение между игреками и иксами. Попытаемся зафиксировать всю совокупность переменных, кроме одной, и, как раз, получится то, что надо.

Обозначим цель: выразить игреки через иксы, то есть найти $y^i = f_i(x_1, \ldots, x_m)$. Начинаем с малого, и выразим одну через все остальные. Выбирать мы будет не абы какую переменную. F'_y можно записать как матрицу Якоби. Поскольку мы знаем, что эта матрица невырожденная, то у неё есть хотя бы одна ненулевая строка, у которой есть ненулевой элемент. Не умаляя общности, это - последняя строка и последний в ней элемент. $\frac{\partial F_n}{\partial y_n}(x^0, y^0) \neq 0$ (даже в некоторой окрестности). y_n мы выражать и будем.

Применим теперь теорему о неявной функции в случае с одной переменной, и тогда $F_k(x_1,\dots,x_m,y_1,\dots,y_{n-1}) = 0$, $1 \le k \le n$, у нас есть столько соотношений, и нам бы хотелось их записать в виде того, что y_n является функцией от остальных переменных. Чтобы это сделать, нам достаточно только одного. Количество соотношений должно быть равно количеству переменных, поэтому нас из всех этих соотношений нужно только последнее - n-ое. У нас есть от него начальный вектор длины $M+n-1$. Подставим

$$F_n(x_1^0,\dots,x_m^0,y_1^0,\dots,y_{n-1}^0,y_n^0) = 0,$$

это нам дано.

Тогда согласно теореме, у нас существует f^* такая, что $y_n = f^*(x_1,\dots,x_m,y_1,\dots,y_{n-1})$. У этой формулы есть, конечно, одна беда, было бы совсем круто, если бы мы выразили это всё только через иксы, а тут ещё куча игреков. Но тогда спустимся индукцией вниз, и дойдём до того, что как раз останутся только иксы. Однако, тут нужно следить за многими важными моментами, например, за невырожденностью матриц. Заведём следующую функцию: $\varphi_1 = F_1(x_1,\dots,x_m,y_1,\dots,y_{n-1},f^*(x_1,\dots,x_m,y_1,\dots,y_{n-1}))$. Точно также мы можем сделать для $1 \le k \le n-1$. Тогда $\varphi_n \equiv 0$.

$\varphi_k(x_1,\dots,x_m,y_1,\dots,y_{n-1}) = 0$, $1 \le k \le n-1$. Нам нужно проверить, что матрица отображения $\varphi = (\varphi_1,\dots,\varphi_{n-1})$ невырожденна в точке. Тогда нам ещё нужно научиться считать производные φ_k по игреку. Попытаемся продифференцировать $\frac{\partial \varphi_k}{\partial y_l}$, где и k и l - от 1 до $n-1$. Эта производная равна $\frac{\partial F_k}{\partial y_l} + \frac{\partial F_k}{\partial y_n} \cdot \frac{\partial f^*}{\partial y_l}$, в f^* мы подставляем всё, кроме последней координаты, а в F подставляем всё, да на месте последней координаты - f^* от всего остального.

Рассмотрим, что у нас получается с матрицей $\frac{\partial \varphi_k}{\partial y_l}$, $(n-1)\times(n-1)$ (точнее, рассматриваем $\frac{\partial F_k}{\partial y_l}$ и размера $n \times n$). Рассматриваем теперь верхнюю левую подматрицу $(n-1)\times(n-1)$, а из большой получим при помощи элементарных преобразований получим красивую большую матрицу, которая состоит из элементов, идентичных $\frac{\partial F_k}{\partial y_l} + \frac{\partial F_k}{\partial y_n} \cdot \frac{\partial f^*}{\partial y_l}$. Внутри того куска $(n-1)\times(n-1)$ получим красоту (φ'_y), однако в последнем столбце у нас получается что-то не хорошее ($\frac{\partial F_n}{\partial y_1} + \frac{\partial F_n}{\partial y_n} \cdot \frac{\partial f^*}{\partial y_1}$, однако, так как это производные, то тождественные нули). Аналогично не очень хорошо получается в последней строке, а в правом нижнем углу стоит $\frac{\partial F_n}{\partial y_n}$.

Тогда определитель матрицы $F'_y = \det \varphi'_y \cdot \frac{\partial F_n}{\partial y_n}$. Но ни первый, ни второй множитель не равняются нулю, поэтому и искомое не нуль. $\varphi_i(\text{everything}) = 0$, тогда существуют f_k, $1 \le k \le n-1$, тогда $y_n = f^*(x_1,\dots,x_m,y_1,\dots,y_{n-1}) = f^*(x_1,\dots,x_m,f_1(\text{x's}),\dots,f_{n-1}(\text{x's}))))$.

Осталось лишь доказать формулы $f'(x)$, $f : \mathbb{R}^m \to \mathbb{R}^n$. У нас есть соотношение

$$F_k(x_1,\dots,x_m,f_1,\dots,f_n) = 0,$$

$1 \le k \le n$. Давайте продифференцируем это дело. Мы имеем право дифференуировать по m переменным. $\frac{\partial F_k}{\partial x_1} = \frac{\partial F_k}{\partial x_1} + \sum_{l=1}^{n} \frac{\partial F_k}{\partial y_l} \cdot \frac{\partial f_l}{\partial x_1} = 0$. Индексы при F и x могут быть любыми, тогда мы можем 1 заменить на s, и у нас выполнено $m \times n$ соотношений ($1 \le k \le n$, $1 \le s \le m$). $F'_x + F'_y f' = 0$, тогда посмотрим на рзмеры матриц, и пристальным вглядыванием получим из обратимости $f' = -(F'_y)^{-1} \cdot F'_x$. (немного перестал понимать, что происходит - надо вникнуть и записать). □

Перейдём теперь к теме *условные экстремумы*. Пусть $f = f(x_2,\dots,x_n)$, $G \subset \mathbb{R}^n$, $F \to \mathbb{R}$. Если у нас есть набор $f_i(x_1,\dots,x_n) = 0$ при $1 \le i \le m$, то хотелось бы, чтобы $m \le n$. Тогда мы можем сократить количество переменных. Ну, Юрий Сергеевич начал рисовать картиночки, я так не умею. Пусть у нас есть поверхность S и $f(x,y,z)$, и мы хотим найти экстремум по S - $f|_S$. Пусть у нас есть точка $(x_0, y_0, z_0) \in S$, тогда мы хотим, чтобы в точке было касание. Если в общем и целом, то нас инетересует касание, мы рассматриваем параметризацию от одной переменной t по кривой на поверхности, и дальше с ней чего-то там делаем.

11 Лекция 11.

Продолжаем с условными экстремумами. Немного не формальные рассуждения: у нас есть функция $f(x, y, z)$, поверхность S, заданная уравнением $f_1(x, y, z)$, и вот

Пусть есть кривая $(x(t), y(t), z(t)) \in S$ такая, что при $t = 0$ получается x_0, y_0, z_0. Введём $\psi(t) = f(x(t), y(t), z(t))$, тогда $\psi'(0) = 0$, $\psi'(t) = \frac{\partial f}{\partial x}x'(t) + \frac{\partial f}{\partial y}y'(t) + \frac{\partial f}{\partial z}z'(t)$, что опять равно нулю в нуле. Эта вещь равносильна утверждению о том, что $\operatorname{grad} f$ перпендикулярен касательной плоскости $T_s(x_0, y_0, z_0)t = 0$. Мы знаем, что grad $\perp$ поверхности уровня $(f(x, y, z) = f(x_0, y_0, z_0))$, потому и получаем.

Давайте теперь рассмотрим общий случай размерностей. У нас есть функция n переменных и поверхность, заданная m соотношениями: $f = f(x_1, \ldots, x_n)$, $S- \to f_k(x_1, \ldots, x_n) = 0$ - система из m уравнений $(1 \leq k \leq m)$. (?)

Пусть мы ищем $f|_s$ - имеет экстремум в x^0, тогда при $x(t) \in S$, $x(0) = x^0$, рассматриваем опять $\psi(t) = f(x(t))$, о которой мы уже знаем, что значение производной в нуле равно нулю, и равно $(\operatorname{grad}|_{t=0}, x'(0))$. Тогда для любой $x(t) \in S$, (постоянные сбои в связи, придётся пересматривать)