Лекция 5.

Градиент, его геометрический смысл и физические применения.
Однородные функции и соотношение Эйлера.
Метод размерностей.

Краткое содержание предыдущих лекций:

Лекция 1. Локальная линеаризация, примеры: отыскание линейных поправок к значениям величин A^{-1}, $\exp A$, $\det E$, $\langle a, b\rangle$ при малых изменениях аргументов (A, E — матрицы, a, b — векторы); упрощение уравнения движения маятника в случае малых колебаний; подсчет перемещения за малое время: корректировка стрельбы. Абсолютная и относительная погрешности приближения. Асимптотические соотношения с символами o, O.

Лекция 2. Линейная функция (отображение). Матрица линейного отображения и ее изменение при замене базиса. Матрица оператора $\frac{d}{dx} : P_n \to P_{n-1}$ в базисе $\{1, x, \dots, x^n\}$ пространства P_n полиномов степени n от x. $\mathbb{C}$- и $\mathbb{R}$- линейность.

Норма в векторном пространстве. Примеры норм в $\mathbb{R}^n$, $C[a, b]$, $C^{(n)}[a, b]$. Метрика, порожденная нормой и норма, порожденная скалярным произведением.

Норма линейного отображения и ее геометрический смысл. Непрерывность линейного отображения и конечность его нормы. Интегрирование и дифференцирование как линейные отображения. Оценка нормы линейного отображения $A : \mathbb{R}^m \to \mathbb{R}^n$ через элементы его матрицы. Пространство $L(X, Y)$ линейных непрерывных отображений.

Лекция 3. Дифференцируемость отображения в точке. Дифференциал, его область определения и область значений. Координатная запись дифференциала (матрица Якоби) отображения $f : \mathbb{R}^m \to \mathbb{R}^n$. Взаимоотношения между непрерывностью, наличием частных производных и дифференцируемостью функции в точке. Дифференцирование линейного отображения и функций $A^{-1}, \exp A, \det A, a \cdot b, a \times b$ (A — матрица; a, b — векторы).

$\mathbb{C}$ и $\mathbb{R}$ линейность и $\mathbb{C}$ и $\mathbb{R}$ дифференцируемость ($\mathbb{C}$-$\mathbb{R}$ условия). Производная по вектору.

Лекция 4. Основные общие законы дифференцирования (дифференциального исчисления): линейность, дифференциал композиции отображений и обратного отображения; дифференцирование полилинейной функции. Вычисление частных производных сложной функции.

Итак, лекция 5.

1. Производная по вектору и градиент.

Хотя в основном здесь речь будет о градиенте, напомним сначала понятие производной по вектору или, как ее иногда называют, производной по направлению, когда длина вектора единичная.

Пусть $f(x)$ — переменная величина (например атмосферная температура) зависящая от точки пространства, где происходит ваше движение по закону $x = x(t)$, t — вещественный параметр (время). Изменение величины f на траектории вашего движения (изменение вашей местной температуры) будет происходить по закону $f(x(t))$. Найдем скорость изменения этой величины

$$\frac{d(f \circ x)}{dt} = f'(x(t))x'(t) = f'(x)v$$

где $v = x'(t)$ — вектор скорости движения в точке $x = x(t)$.

Мы видим, что если f — дифференцируемая функция, то результат зависит только от x и v и является значением дифференциала функции f в точке x на векторе v. От остальных, кроме v, элементов траектории в этот момент он не зависит и, значит, одинаков для всех траекторий, проходящих через точку x с фиксированной скоростью v. В частности, можно взять простейший закон равномерного движения $x + vt$ и принять следующее

Определение. *Производной функции f в точке x по вектору v называется величина*

$$D_v f(x) = \lim_{t \to 0} \frac{f(x + vt) - f(x)}{t}$$

Если функция f дифференцируема в точке x, то, как было показано,

$$D_v f(x) = f'(x)v$$

Упражнение. Покажите, что $D_{\lambda v} f(x) = \lambda D_v f(x)$, а если f — дифференцируема, то и $D_{v_1 + v_2} f(x) = D_{v_1} f(x) + D_{v_2} f(x)$.

Если e_i — единичный вектор i-го координатного направления, то, очевидно, $D_{e_i} f(x) = \frac{\partial f}{\partial x^i}(x)$.

Внимание! Наличие производной по любому направлению еще не означает, что функция дифференцируема.

Пример. Рассмотрите функцию $f(x,y)$ на плоскости $\mathbb{R}^2$, равную 1 на параболе $y = x^2$ без вершины (0,0), а в остальных точках плоскости $f(x,y) \equiv 0$.

Чтобы перейти теперь к градиенту, напомним кое-что из линейной алгебры и геометрии.

Если в векторном пространстве V имеется невырожденная билинейная форма $B(x,y)$, то каждой линейной функции $g^* \in V^*$ на этом пространстве отвечает единственный вектор g, такой, что $g^*(v) = B(g,v)$ для любого вектора $v \in V$.

Упражнение. Если $V = \mathbb{R}^n$, $B(x,y) = b_{ij}x^i x^j$, $g^* v = g_i v^i$, то вектор g имеет координаты $g^j = b^{ij} g_i$, где (b^{ij}) — матрица, обратная матрице (b_{ij}).

Чаще всего в качестве билинейной формы $B(\cdot,\cdot)$ выступает стандартное симметричное скалярное произведение $\langle\cdot,\cdot\rangle$ в евклидовой геометрии или кососкалярное произведение $\omega(\cdot,\cdot)$ (когда форма B кососимметрична) в симплектической геометрии.

Упражнение. Пусть $B(v_1, v_2) = \begin{vmatrix} v_1^1 & v_1^2 \\ v_2^1 & v_2^2 \end{vmatrix}$ — ориентированная площадь параллелограмма, натянутого на векторы $v_1, v_2 \in \mathbb{R}^2$. Найдите вектор $g = (g^1, g^2)$, отвечающий линейной функции $g^* = (g_1, g_2)$.

Теперь вернемся к анализу.

Определение. *Вектор, соответствующий дифференциалу функции* $f : \mathbb{R}^n \to \mathbb{R}$ *в точке* x *относительно скалярного произведения* $\langle , \rangle$ *евклидова пространства* $\mathbb{R}^n$, *называется градиентом функции* f *в этой точке и обозначается* $\operatorname{grad} f(x)$.

Итак,

$$df(x)v = \langle \operatorname{grad} f(x), v \rangle$$

для любого приложенного к x вектора $v \in T_x\mathbb{R}^n \sim \mathbb{R}^n$.

Значит,

$$f'(x)v = \frac{\partial f}{\partial x^1}(x)v^1 + \cdots + \frac{\partial f}{\partial x^n}(x)v^n = \langle \operatorname{grad} f(x), v \rangle = |\operatorname{grad} f(x)||v| \cos\varphi.$$

Отсюда видно, во-первых, что в стандартном ортонормированном базисе, т.е. в декартовых координатах,

$$\operatorname{grad} f(x) = \left(\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n} \right)(x)$$

Во-вторых, если $\operatorname{grad} \neq 0$ и e — единичный вектор, приложенный к x, то производная $D_e f(x)$ принимает наибольшее значение, равное $|\operatorname{grad} f(x)|$, когда e сонаправлен с $|\operatorname{grad} f(x)|$.

Т.е. скорость роста функции f при движении из точки x с единичной скоростью максимальна, когда направление движения совпадает с направлением градиента функции в этой точке, и равна $|\operatorname{grad} f(x)|$.

При движении в направлении, перпендикулярном вектору $\operatorname{grad} f(x)$, функция не меняется.

Если нарисовать, например, на плоскости $\mathbb{R}^2$, как на картах, линии уровня функции $f(x, y)$ (т.е. кривые $f(x, y) = \text{const}$), то градиент в данной точке будет перпендикулярен линии уровня и укажет направление наиболее крутого подъема к вершине (см. рисунок).

На этом наблюдении основаны так называемые градиентные численные методы поиска экстремумов функций многих переменных.

Вопрос. Чему равен градиент в точке, отвечающей на приведенном рисунке седлу между вершинами?

Задача. Получите уравнение $xx_0/a^2 + yy_0/b^2 + zz_0/c^2 = 1$ плоскости, касательной к эллипсоиду $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$ в точке (x_0, y_0, z_0).

Чтобы детальнее разобраться с алгебраическим и геометрическим аспектами градиента, полезно сделать еще следующие задачи:

Задача.

1. Выше в упражнении была рассмотрена кососимметрическая форма $B(v_1, v_2)$ ориентированной площади параллелограмма в $\mathbb{R}^2$. Линейной функции $g^* \in L(\mathbb{R}^2, \mathbb{R})$, определяемой своими коэффициентами (g_1, g_2), посредством этой кососимметрической формы ω сопоставляется вектор $g_\omega = (g^1, g^2)$. Сравните координатное выражение вектора g_ω и вектора $g_{\langle,\rangle}$, полученного в случае, когда $B(,) = \langle,\rangle$ — стандартное скалярное произведение.

2. Если вектор, отвечающий $df(x)$ относительно симметрической формы $\langle,\rangle$ называют градиентом $\operatorname{grad} f(x)$, то вектор, отвечающий

$df(x)$ относительно кососимметрической формы ω называют косым градиентом и обозначают $\operatorname{sgrad} f(x)$ (от английского "skew" — косой).

Запишите $\operatorname{grad} f(x)$ и $\operatorname{sgrad} f(x)$ в декартовых координатах $\mathbb{R}^2$.

Задача.

1. Как изменятся координаты вектора $\operatorname{grad} f(x)$, если в $\mathbb{R}^2$ вместо ортонормированного базиса (e_1, e_2), взять ортогональный базис $(\tilde{e}_1, \tilde{e}_2) = (\lambda_1 e_1, \lambda_2 e_2)$?

2. Наличие скалярного произведения $\langle , \rangle$ в $\mathbb{R}^2$ естественно позволяет считать его заданным и в любом пространстве $T_x\mathbb{R}^2$ векторов, приложенных к точке $x \in \mathbb{R}^2$.

   Дифференциал отображения $\Phi : \tilde{\mathbb{R}}^2 \to \mathbb{R}^2$ в точке, как мы знаем, есть линейное отображение, соответствующих касательных пространств. Введение полярных координат (r, φ) посредством соотношений $x = r\cos\varphi$, $y = r\sin\varphi$ можно рассматривать как отображение плоскости $\tilde{\mathbb{R}}^2 = \mathbb{R}^2(r, \varphi)$ с декартовыми координатами r, φ в плоскость $\mathbb{R}^2 = \mathbb{R}^2(x, y)$ с декартовыми координатами (x, y).

   Пусть $(x, y) \in \mathbb{R}^2_{(x,y)}$ точка с полярными координатами (r, φ), т.е. точка, отвечающая точке $(r, \varphi) \in \mathbb{R}^2_{(r,\varphi)}$, при указанном отображении Φ. В пространстве $T_{(x,y)}\mathbb{R}^2$, изобразите векторы e'_r, e'_φ, отвечающие векторам (1,0) и (0,1) пространства $T_{(r,\varphi)}$ (здесь $(x, y) = \Phi(r, \varphi)$ и $d\Phi(r, \varphi) : T_{(r,\varphi)}\tilde{\mathbb{R}}^2 \to T_{(x,y)}\mathbb{R}^2$).

   Найдите $\langle e'_i, e'_j \rangle$ и скажите, образуют ли векторы e'_r, e'_φ ортонормированный базис в $T_{(x,y)}\mathbb{R}^2$.

   Проверьте, что ортонормированный базис e_r, e_φ получится, если вместо векторов (1,0), (0,1) в $T_{(r,\varphi)}$ взять векторы (1,0), (0,r).

3. Как вычислять $\operatorname{grad} f$ в полярных координатах? (Имеется в виду, что заданная в $\mathbb{R}^2_{(x,y)}$ функция f рассматривается как функция переменных (r, φ), связанных с (x, y) указанной выше полярной заменой $(x, y) = \Phi(r, \varphi)$, а вектор $\operatorname{grad} f$ в каждой точке $(x, y) = \Phi(r, \varphi)$ плоскости раскладывается по ортонормированному базису e_r, e_φ полярных координатных направлений).

   Ответ: $\left(\frac{\partial f}{\partial r}, \frac{1}{r}\frac{\partial f}{\partial \varphi}\right)$.

   Указание. Один из путей — вспомнить как меняется выражение билинейной формы $B(,)$ и, в частности, скалярного произведения $\langle , \rangle$ при переходе к новым координатам.

Задача (о косом градиенте).

1. Покажите, что в $\mathbb{R}^3$ (и вообще в $\mathbb{R}^{2n+1}$) нет невырожденной кососимметрической билинейной формы.

2. В ориентированном $\mathbb{R}^2$, как мы видели, есть невырожденная кососимметрическая билинейная форма (ориентированная площадь параллелограмма). В $\mathbb{R}^{2n}$ с координатами $(x^1, \ldots, x^n, \ldots, x^{2n}) = (p^1, \ldots, p^n, q^1, \ldots, q^n)$ такая билинейная форма ω тоже есть: если $v_i = (p_i^1, \ldots, p_i^n, q_i^1, \ldots, q_i^n)$ $(i = 1, 2)$, то

$$\omega(v_1, v_2) = \begin{vmatrix} p_1^1 & q_1^1 \\ p_2^1 & q_2^1 \end{vmatrix} + \cdots + \begin{vmatrix} p_1^n & q_1^n \\ p_2^n & q_2^n \end{vmatrix}.$$

   Т.е. $\omega(v_1, v_2)$ — это сумма ориентированных площадей проекций натянутого на v_1, v_2 параллелограмма в координатных плоскостях $(p^j, q^j) \qquad (j = 1, \ldots, n)$.

   Пусть g^* — линейная функция в $\mathbb{R}^{2n}$, заданная своими коэффициентами $g^* = (p_1, \ldots, p_n, q_1, \ldots, q_n)$. Найдите координаты вектора g, сопоставляемого функции g^* посредством формы ω.

3. Дифференциалу функции $f : \mathbb{R}^{2n} \to \mathbb{R}$ в точке $x \in \mathbb{R}^{2n}$ посредством кососимметрической формы ω сопоставляется вектор, называемый косым градиентом функции f в этой точке и обозначаемый $\operatorname{sgrad} f(x)$. Найдите выражение $\operatorname{sgrad} f(x)$ в канонических декартовых координатах пространства $\mathbb{R}^{2n}$.

4. Найдите скалярное произведение $\langle \operatorname{grad} f(x), \operatorname{sgrad} f(x) \rangle$.

5. Покажите, что вектор $\operatorname{sgrad} f(x)$ направлен вдоль поверхности уровня функции f.

6. Закон движения $x = x(t)$ точки в пространстве $\mathbb{R}^{2n}$ таков, что $\dot{x}(t) = \operatorname{sgrad} f(x(t))$. Покажите, что $f(x(t)) = \text{const}$.

7. Запишите уравнение $\dot{x} = \operatorname{sgrad} f(x)$ в канонических обозначениях $(p^1, \ldots, p^n, q^1, \ldots, q^n)$ для координат и $H = H(p, q)$ для функции f. Полученная система, называемая системой уравнений Гамильтона, является одним из центральных объектов механики.

2. Примеры физических приложений градиента.

а. Потенциальные поля.

Векторное поле F (обычно это силы или напряженности), заданное в некоторой области пространства, называется *потенциальным*, если существует функция U, определенная в этой области, такая, что $F = -\operatorname{grad} U$.

Функцию U называют *потенциалом поля*, а функцию $-U$ (для которой $F = \operatorname{grad}(-U)$) обычно называют *силовой функцией поля.*

Задача.

1. Проверьте, что гравитационное поле — потенциально; укажите его потенциал (потенциал Ньютона).

2. Любое непрерывное векторное поле, действующее вдоль прямой, — потенциально.

3. Поле вида $F(x) = \lambda(|x|)x$ $(\lambda : \mathbb{R}_+ \to \mathbb{R})$ называют *центральным.* Гравитационное поле массы и поле напряженности электрического заряда — примеры центральных полей. Покажите, что любое центральное поле — потенциально.

Потенциальные силовые поля среди остальных полей выделяются тем, что работа такого поля (или в таком поле) вдоль некоторого пути зависит только от начала и конца пути, но не от самого пути и равна разности значений потенциала в конце и в начале.

Напишем формулу для подсчета работы поля F на пути γ, проходимом по закону $x = x(t)$ из точки $x_0 = x(0)$ в точку $x_1 = x(1)$ (см. рисунок). Элементарная работа на малом участке от x, до $x + \Delta x$, где $\Delta x \approx \dot{x}(t)\Delta t$, равна $\Delta A = \langle F(x), \Delta x\rangle = \langle F(x(t)), \dot{x}(t)\rangle \Delta t$. Значит

$$dA = \langle F(x(t)), \dot{x}(t)\rangle dt$$

и

$$A = \int_0^1 \langle F(x(t)), \dot{x}(t)\rangle \, dt$$

Задача.

1. Покажите, что в потенциальном поле $F = -\operatorname{grad} U$ работа равна $A = U(x_1) - U(x_0)$. Указание: $dU(x(t)) = \ldots ?$

2. Если работа поля F вдоль любого пути зависит только от начала и конца пути, то поле F — потенциально.

Указание: Потенциал определен с точностью до аддитивной постоянной. Фиксировав начальную точку x_0 найдите работу $A_{x_0}(x)$ как функцию конца x пути; $U(x) = A_{x_0}(x)$.

b. Уравнение Эйлера в механике сплошной среды.

Перепад давления обусловливает перемещение сплошной среды (атмосферы или жидкости). Давление в данной точке x идеально невязкой среды задается одним числом $p(x)$. Рассмотрим маленький цилиндр среды (см. рисунок), ориентированный в направлении $\operatorname{grad} p(x)$.

Не уравновешенны здесь только силы давления, действующие на основания цилиндра (почему?)

Если ΔS — площадь основания, Δh — высота, а $\Delta V = \Delta S \cdot \Delta h$ — объем цилиндра, то действующая на цилиндр сила по величине равна $\Delta p \cdot \Delta S$ и направлена в сторону меньших давлений, т.е. в сторону, противоположную направлению $\operatorname{grad} p(x)$.

Пусть ρ — плотность среды в рассматриваемой области и a — ускорение цилиндра. Тогда по закону Ньютона

$$\rho \cdot \Delta V \cdot a = \Delta p \cdot \Delta S \cdot e$$

где $e = -\operatorname{grad} p(x)/|\operatorname{grad} p(x)|$ — единичный вектор нужного направления. Учитывая, что $\Delta p = |\operatorname{grad} p(x)|\Delta h + o(\Delta h)$ (вспомните производную по вектору и ориентацию цилиндра), при $\Delta h \to 0$ отсюда получаем уравнение Эйлера

$$\rho a = -\operatorname{grad} p.$$

Плотность ρ, ускорение a и давление p, вообще говоря, зависят и от точки и от времени: $\rho(x,t)$, $a(x,t)$, $p(x,t)$. Градиент давления $\operatorname{grad} p$, однако, вычисляется только по пространственной переменной x. Если величины ρ, a, p от времени не зависят, то течение среды называют *установившимся или стационарным.*

Если помимо давления имеются еще какие-то действующие на среду силы, например, гравитационные или электро-магнитные, то их надо добавить в правую часть уравнения Эйлера.

c. Законы сохранения.

Пусть точка массы m движется в потенциальном поле $F = -\operatorname{grad} U$ и $x = x(t)$ — закон ее движения. В соответствии с динамическим уравнением Ньютона $m\ddot{x}(t) = -\operatorname{grad} U(x(t))$. Пусть $v = \dot{x}(t)$. Проверим, что величина $\frac{1}{2}mv^2 + U$ (сумма кинетической и потенциальной энергии) остается постоянной вдоль траектории движения.

Действительно,

$$\frac{d}{dt}\left(\frac{1}{2}m\langle \dot{x}(t), \dot{x}(t)\rangle + U(x(t))\right) = m\langle \ddot{x}(t), \dot{x}(t)\rangle + \langle \operatorname{grad} U(x(t)), \dot{x}(t)\rangle =$$
$$\langle m\ddot{x}(t) + \operatorname{grad} U(x(t)), \dot{x}(t)\rangle = \langle 0, \dot{x}(t)\rangle = 0$$

Задача:

1. Закон Даниила Бернулли.

В сплошной среде (установившемся потоке идеальной жидкости) постоянной плотности $\rho \equiv 1$ будем следить за движением частицы (пылинки). Пусть $x = x(t)$ — траектория ее движения. В любой момент t скорость $\dot{x}(t)$ частицы равна скорости потока в точке $x = x(t)$, где находится в этот момент частица. Линии, касательные к которым совпадают с направлением скорости потока в соответствующих точках, называют линиями тока данного течения. Так что траектория движения частицы — линия тока. Пользуясь уравнением Эйлера, покажите, что величина $\frac{1}{2}\rho v^2 + p$ постоянна вдоль траектории (хотя может меняться при переходе на другую линию тока).

2. Работа крыла в ламинарном потоке.

Из закона Бернулли следует, что на участках траектории, где скорость больше, давление — меньше. В безвихревом (ламинарном) потоке рассмотрим две близкие линии тока, одна из которых обходит крыло сверху, а другая снизу. Поскольку линии близки, то в далекой от крыла зоне в их близких точках величины v и p почти одинаковы. Если считать, что вихрей, разрывов и смещений нет, т.е. близкие частицы p_-, p_+ после обтекания крыла снова сомкнуться как близкие, то частице, обходящей крыло сверху, придется за то же время проделать путь больший, чем частице, обходящей его снизу. Т.е. обходя крыло, верхняя частица имеет скорость большую, чем нижняя. По закону Бернулли это приведет к понижению давления сверху и возникновению подъемной силы крыла. (Рассуждение, конечно, относится к очень упрощенной модели и среды, и обтекания).

3. Интеграл Коши.

Пусть течение происходит в поле сил F. Как было замечено выше, уравнение Эйлера в этом случае имеет вид $\rho a = -\operatorname{grad} p + F$. Если поле F потенциально (например, для течения жидкость у поверхности Земли F — поле силы тяжести), то вдоль линий тока сохраняется величина, подобная рассмотренной в законе Бернулли. Укажите ее (она называется интегралом Коши).

3. Однородные функции и соотношение Эйлера.

В качестве еще одного примера использования законов дифференциального исчисления (или, попросту, правил дифференцирования) получим соотношение Эйлера для однородных функций и укажем затем одно его полезное применение.

Определение. *Функция f называется* однородной функцией степени *n, если $f(tx) = t^n f(x)$.*

При этом $t \in \mathbb{R}$, а x — точка векторного пространства (например, $\mathbb{R}^m$), в котором (или в части которого) задана функция.

Продифференцировав это равенство по t, получим $f'(tx)x = nt^{n-1}f(x)$, что при $t = 1$ дает соотношение Эйлера $f'(x)x = nf(x)$ для однородных функций.

Если $x = (x^1, \dots, x^m) \in \mathbb{R}^m$, оно имеет вид

$$\frac{\partial f}{\partial x^1}x^1 + \cdots + \frac{\partial f}{\partial x}^m(x)x^m = nf(x^1, \dots, x^m).$$

Упражнение. Какова степень однородности следующих функций $f(x,y,z)$: $xy + yz + zy$; $z(x+y)^{-1}$; $z(x^2+y^2)^{-1}$; $z^{3/2}(x+y)^{-1}$? Проверьте для них соотношение Эйлера.

Задача:

1. Покажите, что если $f'(x)x = nf(x)$, то функция f — однородная степени n.

   Указание: Продифференцируйте $\psi(t) = t^{-n}f(tx)$.

2. Попробуем обобщить понятие однородной функции. Пусть функция f такова, что $f(tx) = \varphi(t)f(x)$. Покажите, что $\varphi(t) = t^n$.

3. Пусть $t = (t^1, \dots, t^m)$, $x = (x^1, \dots, x^m)$, $tx = (t^1x^1, \dots, t^mx^m)$ и $f(tx) = \varphi(t)f(x)$. Покажите, что $\varphi(t^1, \dots, t^m) = (t^1)^{n_1} \cdot \ldots \cdot (t^m)^{n_m}$.

4. Анализ размерностей физических величин.

Основная предпосылка теории размерности состоит в том, что претендующая на физическую значимость зависимость

$$x_0 = f(x_1, \dots, x_k, x_{k+1}, \dots, x_n) \qquad (*)$$

должна быть такой, чтобы при изменении масштабов единиц измерения численные значения всех одноименных величин, участвующих в формуле $(*)$, менялись в одно и то же число раз.

Пусть среди величин $x_1, \dots, x_k, \dots, x_n$ подобно длине, массе и времени, величины $x_1, \dots, x_k$ независимы. При изменении масштабов основных единиц (длины, массы, времени, ...) или взаимно независимых производных единиц $x_1, \dots, x_k$ численное выражение физической величины x_0 изменится на множитель $\alpha_0 = \varphi(\alpha_1, \dots, \alpha_k)$, определяемый только коэффициентами $\alpha_1, \dots, \alpha_k$ изменения масштабов независимых величин $x_1, \dots, x_k$:

$$f(\alpha_1 x_1, \dots, \alpha_k x_k, \dots, \alpha_n x_n) = \varphi(\alpha_1, \dots, \alpha_k) f(x_1, \dots, x_k, \dots, x_n).$$

В силу результата предыдущей задачи

$$\varphi(\alpha_1, \dots, \alpha_k) = \alpha_1^{d_1} \cdot \dots \cdot \alpha_k^{d_k}.$$

Эту функцию или набор показателей $(d_1, \dots, \dots d_k)$ называют размерностью данной физической величины по отношению к независимым физическим величинам $x_1, \dots, x_k$.

Пример. Если L — длина, M — масса, T — время, то ускорение a имеет размерность $[a] = L^1 M^0 T^{-2}$, т.е. (1, 0, -2).

Задача. Покажите, что должным выбором масштабов независимых величин $x_1, \dots, x_k$ (а именно, полагая $\alpha_1 = 1/x_1, \dots, \alpha_k = 1/x_k$) соотношение $(*)$ приводится к следующему соотношению между безразмерными величинами

$$\Pi = f(\underbrace{1, \dots, 1}_{k}, \Pi_1, \dots, \Pi_{n-k}). \qquad (**)$$

Это утверждение называется П-теоремой теории размерностей. Оно полезно во многих отношениях.

Пример. Пользуясь анализом размерностей попробуем найти период колебаний маятника (см. рисунок).

Интересующая нас величина T априори может зависеть от угла φ_0 начального отклонения, от ускорения g силы тяжести (которое различно на Земле и на Луне), от массы m тела и длины l подвеса: $T = f(l, m, , g, \varphi_0)$. При этом φ_0 — безразмерная величина, а размерности l, m и g по отношению к основным единицам L, M, T — длине, массе и времени — таковы:$[l] = L$, $[m] = M$, $[g] = LT^{-2}$. Значит величины l, m, g независимы. Повторим теперь еще раз общее рассуждение:

$$f(\alpha_1 l, \alpha_2 m, \alpha_3 g, \varphi_0) = \varphi(\alpha_1, \alpha_2, \alpha_3) f(l, m, g, \varphi_0) = \alpha_1^{d_1} \alpha_2^{d_2} \alpha_3^{d_3} f(l, m, g, \varphi_0)$$

Полагая здесь $\alpha_1 = 1/l$, $\alpha_2 = 1/m$, $\alpha_3 = 1/g$ (что можно делать, учитывая независимость l, m, g), находим

$$\frac{T}{\alpha_1^{d_1} \alpha_2^{d_2} \alpha_3^{d_3}} = f(1, 1, 1, \varphi_0)$$

что соответствует безразмерному соотношению (**) П-теоремы.

Итак,

$$T = c(\varphi_0) l^{d_1} m^{d_2} g^{d_3}$$

Чтобы найти показатели d_1, d_2, d_3 сопоставим размерности левой и правой частей равенства. Поскольку $[T] = T = [g]^{-1/2}[l]^{-1/2}$, а l, m, g — независимы, заключаем, что

$$T = c(\varphi_0)\sqrt{\frac{l}{g}}$$

Коэффициент $c(\varphi_0)$ можно найти в лабораторном эксперименте. Из общего анализа размерностей его получить уже нельзя. Более полное, но не такое простое исследование этого вопроса приводит к формуле

$$T = 4K(k)\sqrt{\frac{l}{g}},$$

где

$$K(k) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin\theta}}, \text{ а } k^2 = \sin^2(\varphi_0/2).$$

Если $|\varphi_0| \ll 1$, то $K(k) \approx \pi/2$ и

$$T \approx 2\pi\sqrt{\frac{l}{g}}.$$

.

Лекция 6.

Теорема о конечном приращении и некоторые ее применения.

1. Обсуждение известного и наводящие соображения.

Длина пройденного пути не больше величины максимальной скорости движения, умноженной на время движения.

Это физически очевидное утверждение в математическом оформлении составляет содержание очень употребительной и технически полезной теоремы о конечных приращениях.

Приращение функции и дифференциал, перемещение и мгновенная скорость по самому смыслу этих понятий известным образом связаны в малом (для бесконечно малых приращений аргумента или интервалов времени). Теорема о конечном приращении говорит о том, как эти объекты связаны в большом (при конечных, а не бесконечно малых изменениях аргумента).

Один пример такой теоремы вам уже известен.

Теорема Лагранжа (о среднем). *Если функция $f \in C([a,b],\mathbb{R})$ дифференцируема на промежутке $]a,b[$, то найдется такая точка $\xi \in]a,b[$, что*

$$f(b) - f(a) = f'(\xi)(b-a) \tag{1}$$

Эта теорема немедленно обобщается на вещественнозначные функции $f : D \to \mathbb{R}$, заданные в областях любого пространства (например $\mathbb{R}^m$), а не только в $\mathbb{R}$.

Обобщенная теорема Лагранжа (о среднем). *Если точки $x, x+h$ лежат в области D вместе с соединяющим их отрезком $[x, x+h]$, а функция $f : D \to \mathbb{R}$ непрерывна на нем и дифференцируема в его внутренних точках, то найдется точка $\xi = x + \theta h$, $0 < \theta < 1$ этого отрезка, такая, что*

$$f(x+h) - f(x) = f'(\xi)h \tag{2}$$

Доказательство. Достаточно к функции $\varphi(t) = f(x+th)$ на отрезке $[0,1]$ применить классическую теорему Лагранжа.

Упражнение. Для функции $f : \mathbb{R}^m \to \mathbb{R}$ дайте развернутую координатную запись формулы (2).

Простейшие примеры отображений $f:\mathbb{R}\to\mathbb{R}^2$ показывают, однако, что для векторнозначных функций теорема о среднем, вообще говоря, неверна.

Пример. Равномерное движение по единичной окружности задается законом $z=e^{it}$ (в комплексной записи) или парой функций $x=\cos t$, $y=\sin t$ в вещественной форме.

Вектор скорости $v=ie^{it}$ (или $v=(-\sin t,\cos t)$) касателен к окружности и по величине постоянно равен 1. В моменты $t=0$ и $t=2\pi$ точка занимает одно и то же положение на окружности, сделав один оборот за время 2π. Значит какую бы точку $\xi\in[0,2\pi]$ ни взять, для рассматриваемого отображения $f:\mathbb{R}\to\mathbb{R}^2$ имеем

$$0=f(2\pi)-f(0)\neq f'(\xi)2\pi$$

Вместе с тем, очевидно, что все же верна оценка

$$|f(b)-f(a)|\leq\mathcal{L}_{[a,b]}\leq\max_{\xi\in[a,b]}|f'(\xi)||b-a|, \tag{3}$$

где a и b — произвольные моменты времени t, а $\mathcal{L}_{[a,b]}$ — длина дуги, пройденной за промежуток времени между моментами a и b.

Задача.

1. Выпуклой оболочкой $\operatorname{convex}E$ множества $E\subset\mathbb{R}^m$ называется наименьшее выпуклое замкнутое множество, содержащее E. Покажите, что $\operatorname{convex}E$ существует и изобразите $\operatorname{convex}E$ в случае, когда E — пара точек; три точки; m точек; окружность.

2. Пусть $f:\mathbb{R}\to\mathbb{R}^2$ отображение прямой в окружность, описанное в приведенном выше примере, и пусть $E=\{f'(t)|a\leq t\leq b\}$ — множество векторов скорости движения в промежутке $[a,b]$. Покажите, что существует вектор $v\in\operatorname{convex}E$ такой, что

$$f(b)-f(a)=v(b-a) \tag{4}$$

Рассмотрим теперь отображение $f:\mathbb{R}^m\to\mathbb{R}^n$ или вообще отображение $f:D\to Y$, определенное в области D какого-то векторного нормированного пространства X в другое такое пространство Y. Будем считать, что f — дифференцируемо в D и что D вместе с точками a, b содержит весь отрезок $I=[a,b]$. Прикинем как сравнить длину $|I|=|b-a|$ отрезка I и длину $\mathcal{L}$ образа $f(I)$ этого отрезка. Если $x\in[a,b]$ и $\lambda(x)$ — коэффициент локального (около x) изменения длин малых промежутков $[x,x+\Delta x]\subset[a,b]$ при отображении f (точнее $\lambda(x)=\limsup\limits_{\Delta x\to0}\frac{|\Delta f(x)|}{|\Delta x|}$), то, ясно, что $\mathcal{L} \le \left(\sup\limits_{x\in[a,b]} \lambda(x)\right) |I|$ и даже $\mathcal{L} \le \int_I \lambda(x)\,dx$, если этот интеграл определен. Поскольку $|f(b) - f(a)| \le \mathcal{L}$, то мы заодно получаем оценку приращения функции.

Вспомним теперь геометрический смысл нормы линейного отображения и воспользуемся им применительно к отображению $f'(x)$, касательному к f в точке x. Поскольку величина $\|f'(x)\|$ мажорирует коэффициент линейного растяжения по любому направлению в точке x, то $\lambda(x) \le \|f'(x)\|$ и значит, надо полагать,

$$|f(b) - f(a)| \le \mathcal{L} \le \int_{[a,b]} \|f'(x)\|\,dx \le \sup_{x\in[a,b]} \|f'(x)\||b-a|. \tag{5}$$

Задача.

1. Пусть $J = \int_0^1 \varphi(t)\,dt$ — интеграл от векторнозначной функции $\varphi : [0,1] \to Y$ и пусть $E = \{\varphi(t) \in Y | t \in [0,1]\}$. Покажите, что $J \in \operatorname{convex} E$.

2. Пусть $f \in C^{(1)}([a,b], Y)$. Используя формулу Ньютона-Лейбница и оценку интеграла, получите указанное неравенство между крайними членами соотношения (5).

3. Покажите, что если $f \in C^{(1)}([a,b], Y)$, то найдется линейное отображение $v \in \operatorname{convex}\{f'(x) | x \in [a,b]\}$ такое, что выполняется равенство (4).

2. Формулировка и доказательство теоремы о конечном приращении.

Нам ничего не остается как подвести итог, т.е. правильно сформулировать и формально доказать следующее утверждение.

Теорема о конечном приращении.

Пусть $f : D \to Y$ — дифференцируемое отображение, определенное в области D нормированного пространства Y (например, $\mathbb{R}^m$) со значениями в нормированном пространстве Y (например, $\mathbb{R}^n$).

Если отрезок $[x, x+h]$ содержится в D, то

$$|f(x+h) - f(x)| \le \sup_{0\le\theta\le 1} \|f'(x+\theta h)\||h| \tag{6}$$

Норма каждого из участвующих в этой формуле объектов h, $f(x+h) - f(x)$, $f'(x+\theta h)$ разумеется, берется в том пространстве X, Y, $L(X,Y)$ соответственно, которому он принадлежит.

Доказательство. Пусть, для краткости, $M = \sup\limits_{0\le\theta\le1} \|f'(x+\theta h)\|$. Мы хотим показать, что

$$|f(x+h)-f(x)| \le M|h| \tag{7}$$

Простая, использующая только неравенство треугольника и свойство отрезка, выкладка

$$\begin{aligned}
|f(x_3)-f(x_1)| &\le |f(x_3)-f(x_2)|+|f(x_2)-f(x_1)| \le \\
&\le M|x_3-x_2|+M|x_2-x_1| = M(|x_3-x_2|+|x_2-x_1|) = \\
&= M|x_3-x_1|
\end{aligned}$$

показывает, что если на частях $[x_1,x_2]$, $[x_2,x_3]$ отрезка $[x_1,x_3]$ справедлива оценка вида (7), то она справедлива и на отрезке $[x_1,x_3]$.

Значит, если оценка (7) неверна для отрезка $[x,x+h]$, то последовательным делением его пополам можно получить последовательность стягивающихся к некоторой точке $x_0\in[x,x+h]$ отрезков $[a_k,b_k]\subset[x,x+h]$, на каждом из которых (7) нарушено. Поскольку $x_0\in[a_k,b_k]$, то, рассмотрев отрезки $[a_k,x_0]$, $[x_0,b_k]$, по тем же соображениям можно считать, что нашлась последовательность отрезков вида $[x_0,x_0+h_k]\subset[x,x+h]$, где $h_k\to0$ при $k\to\infty$, на которых

$$|f(x_0+h_k)-f(x_0)| > M|h_k| \tag{8}$$

Если (7) доказать с заменой M на $M+\varepsilon$, где ε — любое положительное число, то при $\varepsilon\to0$ все равно получится (7), поэтому (8) тоже можно заменить на

$$|f(x_0+h_k)-f(x_0)| > (M+\varepsilon)|h_k| \tag{8$'$}$$

и теперь показать, что это несовместимо с дифференцируемостью f в точке x_0.

Действительно,

$$|f(x_0+h_k)-f(x_0)| = |f'(x_0)h_k+o(h_k)| \le \|f'(x_0)\||h_k|+o(|h_k|) \le (M+\varepsilon)|h_k|$$

при $h_k\to0$.

Задача.

Покажите, что в выпуклой оболочке множества $\{f'(x+\theta h)|0\le\theta\le 1\}\subset L(X,Y)$ найдется линейное отображение v, для которого справедливо равенство $f(x+h)-f(x)=vh$ вида (4).

Задача.

1. Пусть f — функция, удовлетворяющая условиям теоремы о конечном приращении, а $L \in L(X, Y)$. Покажите, что
$$|f(x+h) - f(x) - Lh| \leq \sup_{0 \leq \theta \leq 1} \|f'(x+\theta h) - L\||h| \tag{9}$$

2. Покажите, что гладкая на выпуклом компакте K функция, удовлетворяет на нем условию Липшица:
$$\exists M \in \mathbb{R}\ \forall x_1, x_2 \in K\ (|f(x_1) - f(x_2)| \leq M|x_1 - x_2|).$$

3. Примеры использования теоремы о конечном приращении.

a. Условия постоянства функции в области.

Утверждение. *Функция постоянна в области тогда и только тогда, когда ее дифференциал нулевой в любой точке этой области.*

Нетривиально только соотношение $f' \equiv 0 \Longrightarrow f \equiv \mathrm{const}$. Докажем его. Напомним, что область $D \subset X$ в пространстве X — это открытое и связное множество. Ввиду открытости область вместе с любой своей точкой x содержит и некоторую шаровую окрестность точки x. Применяя неравенство (6), устанавливаем постоянство f в шаровой окрестности x. Осталось воспользоваться связностью D, чтобы убедиться, что $f(x_0) = f(x_1)$ для любы точек $x_0, x_1 \in D$. Ввиду связности D из x_0 в x_1 можно пройти по некоторому пути $\gamma \subset D$, который, без ограничения общности, можно даже считать конечнозвенной ломаной (Проверьте!). На каждом звене этой ломаной, в силу теоремы о конечном приращении (6), функция постоянна. Значит, действительно, $f(x_0) = f(x_1)$.

Замечание. На паре областей эти константы могут быть различными: $f|_{D_1} \equiv c_1$, $f|_{D_2} \equiv c_2$.

Задача.

1. Частная производная $\frac{\partial f}{\partial y}$ гладкой функции $f : D \to \mathbb{R}$, заданной в области $D \subset \mathbb{R}^2$ плоскости переменных (x, y), тождественно равна нулю в D. Верно ли, что f не зависит от y в этой области?

2. Опишите области на плоскости, для которых это верно.

b. Достаточное условие дифференцируемости функции в терминах ее частных производных.

Известно, что наличие частных производных у функции в точке не влечет не только дифференцируемости, но даже непрерывности функции в этой точке. На этом фоне важно и полезно следующее

Утверждение. *Если частные производные функции существуют в окрестности некоторой точки и непрерывны в самой этой точке, то функция дифференцируема в этой точке.*

Доказательство. Доказательство проведем для функции пары переменных. Как будет видно, ценой более громоздких выражений его можно повторить и в общем случае.

Проверим, что единственный претендент на дифференциал — линейная функция $\frac{\partial f}{\partial x^1}(x^1,x^2)h^1+\frac{\partial f}{\partial x^2}(x^1,x^2)h^2$, действительно им является:

$$
\begin{aligned}
& |f(x^1+h^1,x^2+h^2)-f(x^1,x^2)-\frac{\partial f}{\partial x^1}(x^1,x^2)h^1-\frac{\partial f}{\partial x^2}(x^1,x^2)h^2| \le \\
\le\ & |f(x^1+h^1,x^2+h^2)-f(x^1,x^2+h^2)-\frac{\partial f}{\partial x^1}(x^1,x^2)h^1| + \\
+\ & |f(x^1,x^2+h^2)-f(x^1,x^2)-\frac{\partial f}{\partial x^2}(x^1,x^2)h^2| \le \\
\le\ & \sup_{0\le\theta_1\le 1}\|\frac{\partial f}{\partial x^1}(x^1+\theta_1h^1,x^2+h^2)-\frac{\partial f}{\partial x^1}(x^1,x^2)\||h^1| + \\
+\ & \sup_{0\le\theta_2\le 1}\|\frac{\partial f}{\partial x^2}(x^1,x^2+\theta_2h^2)-\frac{\partial f}{\partial x^2}(x^1,x^2)\||h^2|.
\end{aligned}
$$

Учитывая непрерывность частных производных в точке (x^1,x^2) и то, что если $h=(h^1,h^2)$, то $|h^1|\le|h|$, $|h^2|\le|h|$, теперь заключаем, что начальный член этой цепочки неравенств есть $o(|h|)$ при $h\to 0$.

Замечание. Можно считать, что мы дважды воспользовались теоремой о конечном приращении в форме (9). Само неравенство (9) сводится к (6), если перейти к функции $\Phi(t)=f(x+th)-Lth$ на отрезке [0,1].

Доказанное утверждение имеет полезное прямое

Следствие. *Если частные производные функции непрерывны в некоторой области, то сама функция непрерывно дифференцируема в этой области (т.е. функция $f'(x)$ определена и непрерывна в любой точке этой области).*

Упражнение. Проверьте непрерывность f'.

Таким образом, принадлежность функции классу $C^{(1)}$ (непрерывно дифференцируемых функций) равносильна принадлежности ее частных производных классу C (непрерывных функций).

Лекция 7.

Производные высшего порядка. Формула Тейлора. Экстремумы функций.

1. Высшие производные (порядка выше первого)

a. Смешанные производные

Пусть $f(x,y) = xy^3 - \sin(x^2+y)$. Вычислим частные производные: $\frac{\partial f}{\partial x} = y^3 - 2x\cos(x^2+y)$, $\frac{\partial f}{\partial y} = 3xy^2 - \cos(x^2+y)$. Они тоже являются функциями двух переменных.

Найдем их частные производные

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = -2\cos(x^2+y) + 4x^2\sin(x^2+y),$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y\partial x} = 3y^2 + 2x\sin(x^2+y),$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x\partial y} = 3y^2 + 2x\sin(x^2+y),$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = 6xy + \sin(x^2+y).$$

Видим, что в смешанных производных результат получился один и тот же. Возникает вопрос, случайно ли совпадение, или всегда $\frac{\partial^2 f}{\partial x\partial y} = \frac{\partial^2 f}{\partial y\partial x}$?

b. Дифференциалы высшего порядка

Вспомним, что если $f : X \to Y$ — дифференцируемое отображение, то его производное отображение $f' : X \to Y'$ сопоставляло каждой точке $x \in X$ точку уже не в пространстве Y, а в пространстве $Y' = L(X, Y)$ линейных отображений, причем $f'(x)$ — дифференциал функции f в точке x. (Считая X и Y векторными пространствами, мы позволили себе отождествление $T_xX \sim X$, $T_yY \sim Y$ и $L(T_xX, T_yY) \sim L(X, Y)$.)

Относительно отображения $f' : X \to Y'$ тоже можно поставить вопрос о его дифференцируемости. Если оно дифференцируемо, то его производное отображение $(f')' : X \to Y''$ называют *вторым производным отображением, второй производной* или *вторым дифференциалом* исходного отображения f и обозначают f'' или $d^2 f$.

По индукции получается определение производной или дифференциала любого порядка $n \in \mathbb{N}$, обозначаемого $f^{(n)}$ или $d^n f$.

Таким образом, переход от пространства аргументов к пространству значений рассматриваемых функций следующий:

$$\begin{tikzcd}
X & Y & L(X,Y) & L(X,L(X,Y)) & \cdots \\
x \arrow[r, mapsto] & f(x) & & & \\
x \arrow[rr, mapsto] & & f'(x) \arrow[dl, "h"'] & & \\
& f'(x)h & & & \\
x \arrow[rrr, mapsto] & & & f''(x) \arrow[dl, "h_1"'] & \\
& (f''(x)h_1)h_2 & f''(x)h_1 \arrow[l, "h_2"'] & & \\
& & \cdots\cdots\cdots & &
\end{tikzcd}$$

Значит, подобно тому, как $f'(x)h$ — функция, линейная по $h \in T_xX \sim X$, со значениями в $T_yY \sim Y$, функция $(f''(x)h_1)h_2 =: f''(x)(h_1, h_2)$, как видно, линейна и по h_1, и по h_2, т.е. билинейна, и, вообще, $f^{(n)}(x)(h_1, \ldots, h_n)$ — n-линейная форма на пространстве $T_xX \sim X$ со значениями в пространстве $T_{y=f(x)}Y \sim Y$.

c. Вычислительные формулы.

Учитывая, что

$$f'(x)h = D_h f(x) = \lim_{t \to 0} \frac{1}{t}(f(x + th) - f(x)) \in Y,$$

покажем, что

$$f^{(n)}(x)(h_1, \ldots, h_n) = D_{h_1} \ldots D_{h_n} f(x)$$

Сделаем проверку при $n = 2$:

$$
\begin{aligned}
f''(x)(h_1,h_2) \ &:= \ f''(x)(h_1,h_2) = (D_{h_1} f'(x))h_2 = \lim_{t\to 0}\frac{1}{t}(f'(x+th) - f'(x))h_2 = \\
&= \ \lim_{t\to 0}\frac{1}{t}(f'(x+th_1)h_2 - f'(x)h_2) = \\
&= \ \lim_{t\to 0}\frac{1}{t}(D_{h_2} f(x+th_1) - D_{h_2} f(x)) = D_{h_1} D_{h_2} f(x).
\end{aligned}
$$

Замечание. В спаривании $(x, x^*) = x^*(x)$ линейной функции x^* и ее аргумента x каждую из величин x и x^* можно рассматривать как линейную функцию по отношению к другой. Если на действие h_2 смотреть с этой точки зрения, то, учитывая перестановочность линейной функции и дифференцирования ($(L\circ\varphi)' = L\circ\varphi'$), можно было бы, минуя предельные переходы, сразу написать ответ.

d. Симметричность высших дифференциалов.

Равенство $D_{h_1} D_{h_2} f(x) = D_{h_2} D_{h_1} f(x)$

Подобно взаимоотношению $D_h f(x)$ и $f'(x)$, величина $D_{h_1} D_{h_2} f(x)$ может существовать и быть вычислена даже если $f''(x)$ в точке x не существует.

Возвращаясь к примеру равенства смешанных производных, с которого мы начали лекцию, покажем, что при условии непрерывности величин $D_{h_1} D_{h_2} f(x)$, $D_{h_2} D_{h_1} f(x)$ в некоторой точке x, они совпадают в этой точке.

Напомним обозначение

$$\Delta f(x)h := f(x+h) - f(x) \quad \text{— первая разность.}$$

Тогда

$$
\begin{aligned}
&\Delta^2 f(x)(h_1,h_2) := \Delta(\Delta f(x)(h_1))(h_2) = \Delta(f(x+h_1) - f(x))h_2 = \\
&= f(x+h_1+h_2) - f(x+h_2) - f(x+h_1) + f(x) \quad \text{— вторая разность.}
\end{aligned}
$$

Вторая разность, очевидно, симметрична относительно h_1, h_2.
Утверждение. *Если функция $D_{h_2} D_{h_1} f(x)$ непрерывна в точке $x\in X$, то справедливо соотношение*

$$D_{h_2} D_{h_1} f(x) = \lim_{t\to 0}\frac{1}{t^2}\Delta^2 f(x)(th_1, th_2).$$

Поскольку вторая разность симметрична, то отсюда получается

Следствие. *Если функции* $D_{h_1}D_{h_2}f(x)$, $D_{h_2}D_{h_1}f(x)$ *непрерывны в точке* $x \in X$, *то они совпадают в этой точке.*

Доказательство. Докажем утверждение. Напомним, что $D_{\lambda h}f(x) = \lambda D_h f(x)$. Оценим $|\Delta^2 f(x)(th_1, th_2) - D_{t_2h_2}D_{t_1h_1}f(x)|$. Если показать, что при $t_1 = t_2 = t \to 0$ это величина $o(t^2)$, то утверждение будет доказано.

Используем (дважды) теорему о конечном приращении. Сначала применительно к функции

$$\Phi(t) = f(x + th_1 + t_2h_2) - f(x + th_1) - D_{t_2h_2}D_{th_1}f(x)$$

на отрезке с концами 0, t_1:

$$\begin{gathered}|f(x + t_1h_1 + t_2h_2) - f(x + t_1h_1) - f(x + t_2h_2) + f(x) - t_1t_2D_{h_2}D_{h_1}f(x)| \leq \\ \sup_{0<\Theta_1<1} \|D_{h_1}f(x + \Theta_1h_1 + t_2h_2) - D_{h_1}(f(x + \Theta_1h_1)) - t_2D_{h_2}D_{h_1}f(x)\||t_1|.\end{gathered}$$

А теперь, продолжая выкладку, еще раз, но применительно к функции

$$\Psi(t) = D_{h_1}f(x + \theta_1h_1 + th_2) - tD_{h_2}D_{h_1}f(x)$$

на отрезке с концами 0, t_2:

$$\begin{gathered}\sup_{0<\theta_2<1} \|D_{h_1}f(x + \theta_1h_1 + t_2h_2) - D_{h_1}(f(x + \theta_1h_1)) - t_2D_{h_2}D_{h_1}f(x)\||t_1| \leq \\ \leq \sup_{0<\theta_1<1} \sup_{0<\theta_2<1} \|D_{h_2}D_{h_1}f(x + \theta_1h_1 + \theta_2h_2) - D_{h_2}D_{h_1}f(x)\||t_1||t_2|\end{gathered}.$$

Ввиду предположенной непрерывности $D_{h_2}D_{h_1}f(x)$ в точке x, при $t_1 \to 0$, $t_2 \to 0$ последнее выражение есть величина $o(t_1t_2)$. Полагая $t_1 = t_2 = t$, получим то, что собирались доказать.

Задача:

1. Если $f''(x)$ существует, то $f''(x)(h_1, h_2) = f''(x)(h_2, h_1)$, т.е. $f''(x)$ — симметричная билинейная форма.

   Указание. Если существует $f''(x)$, то

   $$f''(x)(h_1, h_2) = \lim_{t\to 0} \frac{\Delta^2 f(x)(th_1, th_2)}{t^2}.$$

2. Пусть $X = \mathbb{R}^m$ и все производные $D_{e_i}D_{e_j}f(x)$ непрерывны ($\frac{\partial^2 f}{\partial x^i \partial x^j}$ непрерывны). Тогда существует второй дифференциал $f''(x)$, и он непрерывен.

   Указание. Непрерывность частных производных — достаточное условие дифференцируемости.

Замечание. $f'' : X \to L(X, L(X,Y))$ и значит

$$(f''\textit{непрерывна}) \Longleftrightarrow (x_2 \to x_1 \ \Rightarrow \|f''(x_2) - f''(x_1)\|_{L(X,L(X,Y))} \to 0)$$

Свойства высших производных легко устанавливаются по аналогии со свойствами $f''(x)$. В частности, если существует $f^{(k)}(x)$, то

$$D_{h_k} \cdot \ldots \cdot D_{h_1} f(x) = D_{h_{\sigma(k)}} \cdot \ldots \cdot D_{h_{\sigma(1)}} f(x)$$

для любой перестановки σ.

Определение. *Если у функции f существует непрерывная k-я производная, то говорят, что f k-раз непрерывно дифференцируема (обозначение $f \in C^{(k)}$).*

$$\text{В частности, } f \in C^{(k)}(\mathbb{R}^m, Y) \Longleftrightarrow \frac{\partial^k f}{\partial x^{i_k} \cdots \partial x^{i_1}} = \partial_{i_1 \ldots i_k} f \in C(\mathbb{R}^m, Y)$$

Указание: см. пункт 2) приведенной выше задачи.

Отметим еще, что принято следующее обозначение:

$$f^{(k)}(x)h^k := f^{(k)}(x)(h, \ldots, h).$$

2. Формула Тейлора.

Введенные определения и обозначения позволяют осмысленно написать (а ниже и доказать) *формулу Тейлора.*

$$f(x+h) = f(x) + \frac{1}{1!} f'(x) \cdot h + \ldots + \frac{1}{(n-1)!} f^{(n-1)}(x) h^{n-1} + r_n(x,h), \ (*)$$

$$\text{где } r_n(x,h) = \frac{1}{(n-1)!} \int_0^1 f^{(n)}(x+th) h^n (1-t)^{n-1} dt,$$

применительно к функциям, действующим в любом векторном пространстве (а не только в $\mathbb{R}$).

a. Частные случаи.

Рассмотрим частный случай $f : X \to \mathbb{R}$. Пусть $[x, x+h] \subset X$ и $\varphi(t) = f(x+th)$. Если $f \in C^{(n)}(X, \mathbb{R})$, то $\varphi \in C^{(n)}([0,1], \mathbb{R})$.

К функции φ можно применить классическую формулу Тейлора:

$$\varphi(1) = \varphi(0) + \frac{1}{1!}\varphi'(0) + \ldots + \frac{1}{(n-1)!}\varphi^{(n-1)}(0) + \frac{1}{(n-1)!}\int_0^1 \varphi^{(n)}(t)(1-t)^{n-1}dt\,.$$

Подставим выражение φ через f. Получим:

$$\begin{aligned}&f(x+h) =\\ &= f(x) + \tfrac{1}{1!}f'(x)h + \ldots + \tfrac{1}{(n-1)!}f^{(n-1)}(x)h^{n-1} +\\ &+ \tfrac{1}{(n-1)!}\textstyle\int_0^1 f^{(n)}(x+th)h^n(1-t)^{n-1}dt\,.\end{aligned}$$

Таким образом для случая $f : X \to \mathbb{R}$ формула Тейлора доказана.

Эквивалентная запись:

$$\begin{aligned}&f(x+h) =\\ &= (D_h^0 + \tfrac{1}{1!}D_h^1 + \ldots + \tfrac{1}{(n-1)!}{D_h}^{n-1}f(x) + \tfrac{1}{(n-1)!}\textstyle\int_0^1 (D_h)^n f(x+th)(1-t)^{n-1}dt\,.\end{aligned}$$

Эта запись, как и сама формула Тейлора, конечно, имеет смысл не только в рассмотренном частном случае.

Пример: $f \in C^{(n)}(\mathbb{R}^m, \mathbb{R})$.

Оператор дифференцирования в этом случае имеет вид:

$$D_h f(x) = \left(h^1\frac{\partial f}{\partial x^1} + \cdots + h^m\frac{\partial}{\partial x^m}\right)f(x).$$

Если он применяется k раз, то

$$D_h^k f(x) = \left(h^1\frac{\partial}{\partial x^1} + \ldots + h^m\frac{\partial}{\partial x^m}\right)^k f(x)$$

В этих обозначениях формула Тейлора имеет вид:

$$\begin{aligned}&f(x^1+h^1, \ldots, x^m+h^m) =\\ &= f(x^1, \ldots, x^m) + \tfrac{1}{1!}\left(h^1\tfrac{\partial}{\partial x^1} + \ldots + h^m\tfrac{\partial}{\partial x^m}\right)f(x^1, \ldots, x^m) + \ldots\end{aligned}$$

Вспомним компактную запись оператора дифференцирования:

$$D_h f(x) = \frac{\partial f}{\partial x^1}(x)h^1 + \ldots + \frac{\partial f}{\partial x^m}(x)h^m \equiv \frac{\partial f}{\partial x^i}(x)h^i \equiv \partial_i f(x)h^i$$

(Суммирование производится по i от 1 до m.)

Тогда $D_h(D_h f)(x) = \partial_j \partial_i f(x) h^i h^j$ и вообще $D_h^k(f(x)) =$
$= \partial_{i_1 \dots i_k} f(x) h^{i_1} \dots h^{i_k}$ и формула Тейлора приобретает вид

$$\begin{aligned}
&f(x^1 + h^1, \dots, x^m + h^m) = f(x) + \tfrac{1}{1!}\partial_i f(x) h^i + \dots + \\
&+ \tfrac{1}{(m-1)!}\partial_{i_1 \dots i_{m-1}} f(x) h^{i_1} \dots h^{i_{m-1}} + \\
&+ \tfrac{1}{(n-1)!}\int_0^1 \partial_{i_1 \dots i_n} f(x + th) h^{i_n} \dots h^{i_n} (1-t)^{n-1}\, dt.
\end{aligned}$$

Упражнение. Напишите в развернутом виде выражение $\frac{1}{2!}\partial_{ij} f(x) h^i h^j$ для случая функций от одной, двух и трех переменных.

b. Общая формула Тейлора ($f : X \to Y$).

Упражнение: Пусть X — векторное пространство, X^* — сопряженное пространство (т.е. пространство линейных функций на X). Пусть $x_1 \in X$, $x_2 \in X$. Тогда $(x_1 = x_2) \iff (\forall x^* \in X^*\ (x^*(x_1) = x^*(x_2)))$.

Докажите это для случая, когда X — конечномерное пространство.

Вернемся к формуле Тейлора $(*)$. Возьмем произвольную линейную функцию и докажем, что ее значения на левой и правой частях формулы Тейлора совпадают. (Тогда, как следует из упражнения, правая и левая части формулы Тейлора равны.)

Пусть L — линейная функция.

К числовой функции $L \circ f$ применима уже доказанная формула Тейлора:

$$\begin{aligned}
&(L \circ f)(x + h) = (L \circ f)(x) + \tfrac{1}{1!}(L \circ f)'(x) \cdot h + \dots + \\
&+ \tfrac{1}{(n-1)!}(L \circ f)^{(n-1)}(x) h^{n-1} + \tfrac{1}{(n-1)!}\int_0^1 (L \circ f)^{(n)}(x + th) h^n (1-t)^{n-1} dt
\end{aligned}$$

Т.к. L коммутирует и с оператором дифференцирования ($(L \circ f)' = L \circ f'$) и с интегралом, то, вынося L, из этой формулы получаем:

$$\begin{aligned}
&L \circ (f(x + h)) = \\
&= L\Big[f(x) + \tfrac{1}{1!}f'(x) \cdot h + \dots + \tfrac{1}{(n-1)!}f^{(n-1)}(x) h^{n-1} + \\
&+ \tfrac{1}{(n-1)!}\int_0^1 f^{(n)}(x + th) h^n (1-t)^{n-1} dt\Big]
\end{aligned}$$

Итак, для функций класса $C^{(k)}$ формула Тейлора полностью доказана.

c. Различные формы остаточного члена в формуле Тейлора.

1. *Интегральная форма*: ($f \in C^{(n)}(X)$)

$$r_n(x, h) = \frac{1}{(n-1)!}\int_0^1 f^{(n)}(x + th) h^n (1-t)^{n-1} dt$$

2. Если $f : X \to \mathbb{R}$, то, применяя теорему о среднем, получим *форму Лагранжа*:
$$r_n(x,h) = \frac{1}{n!} f^{(n)}(x + \Theta h)h^n$$

3. *Форма Пеано.* При $h \to 0$ имеем $f^{(n)}(x+th)h^n \approx (f^{(n)}(x) + o(1))h^n$. Тогда $r_n(x,h) = \frac{1}{n!}f^{(n)}(x)h^n + o(|h|^n)$.

   Задача:

   1. Если $f(0) = 0, \dots, f^{(n)}(0) = 0$, то $f(x) = o(|x|^n)$ (обратите внимание, нули в правых частях — разные).

   2. Доказать формулу Тейлора с остаточным членом в форме Пеано, при единственном условии, что существует $f^{(n)}(x)$.

Задача. (Лемма Адамара). Докажите следующие утверждения:

1. Пусть $f \in C^1(\mathbb{R})$, а $f(0) = 0$. Тогда $f(x) = g(x)x$, где $g \in C(\mathbb{R})$ и $g(0) = f'(0)$.

2. Пусть $f \in C^1(\mathbb{R}^m, \mathbb{R})$ и $f(0) = 0$. Тогда $f(x^1, \dots, x^m) = \sum_{i=1}^m g_i(x)x^i$, где g_i — непрерывны и $g_i(0) = \frac{\partial f}{\partial x^i}(0)$.

3. Если $f \in C^n(\mathbb{R})$, то тогда $f(x) = f(0) + \ldots + \frac{1}{(n-1)!}f^{(n-1)}(0)x^{n-1} + g(x)x^n$, где g — непрерывна и $g(0) = \frac{1}{n!}f^{(0)}(0)$.

4. Как выглядит 3), если $f \in C^n(\mathbb{R}^m, \mathbb{R})$?

3. Экстремумы функций.

Напомним определение.x_0 *— локальный минимум функции, точнее, точка локального минимума вещественнозначной функции , если в любой точке x некоторой окрестности точки x_0 $\quad f(x) \geq f(x_0)$.*

x_0 — локальный максимум функции, если $f(x) \leq f(x_0)$.

Такие точки называют локальными экстремумами функции.

Если неравенства строгие, то x_0 — точка строго локального экстремума.

Лемма Ферма. *Если x — внутренняя точка $U(x)$ и функция $f : U(x) \to \mathbb{R}$ имеет экстремум в x , то для любого вектора h, для которого существует $D_h f(x)$, выполнено $D_h f(x) = 0$. В частности, если существует $f'(x)$, то $f'(x) = 0$.*

Доказательство. Положим $\varphi(t) = f(x + th)$ для произвольного фиксированного h. Если x — внутренняя точка $U(x)$, то при $-\delta < t < \delta$ имеем: $x + th \in U(x)$. По классической лемме Ферма $\varphi'(0) = 0$. Но $\varphi'(0) \equiv D_h f(x) = 0$.

Лемма Ферма дает необходимое условие экстремума, но оно, вообще говоря, не достаточное. Формула Тейлора дает достаточное условие наличия или отсутствия экстремума:

Утверждение: *Пусть x — внутренняя точка $U(x)$, а функция $f : U(x) \to \mathbb{R}$ такова что $f'(x) = 0$ и существует $f''(x)$. Тогда:*

Если $\inf\limits_{|h|=1} f''(x)h^2 > c > 0$, то x — точка строго локального минимума.

Если $\sup\limits_{|h|=1} f''(x)h^2 < c < 0$, то x — точка строго локального максимума.

Если существуют h_- и h_+ такие, что $f''(x)h_- < 0$, а $f''(x)h_+ > 0$, то экстремума в x нет.

Доказательство. По формуле Тейлора при $h \to 0$

$$f(x+h) - f(x) = \frac{1}{2}f''(x)h^2 + o(|h|^2) = (\frac{1}{2}f''(x)(\frac{h}{|h|})^2 + o(1))|h|^2$$

Это доказывает первые два пункта. Третий получается аналогично:

$$f(x+th_\pm) - f(x) = \frac{1}{2}f''(x)(th_\pm)^2 + o(|th|^2) = (\frac{1}{2}f''(x)h_\pm + o(1))t^2,$$

где t — числовой параметр и $t \to 0$.

Замечание. Для функций, определенных в областях пространства $\mathbb{R}^m$, проверять указанные свойства $f''(x)h^2$ на сфере $|h| = 1$ можно, используя критерий Сильвестра положительной (отрицательной) определенности формы $f''(x)h = \partial_{ij}f(x)h^ih^j$. Действительно, если форма положительно (отрицательно) определена, то как непрерывная на компакте функция (а сфера в $\mathbb{R}^m$ — компакт) она имеет положительный (отрицательный) минимум (максимум) на сфере $|h| = 1$.

Задачи:

1. **Задача Гюйгенса.**

Между упругими шарами с массами m и M надо поставить n таких промежуточных масс, чтобы при соударении передать от большого шара маленькому наибольшую энергию.

2. **Принцип наименьших квадратов.** Функция f вычисляется экспериментально — вычислены значения функции в некоторых точках $x_1, \ldots, x_n$: $f(x_i) = y_i$. Найти наилучшую линейную аппроксимацию функции $y = f(x) \approx ax + b$, точнее, линейную эмпирическую формулу, наилучшим образом отвечающую результатам эксперимента.

   Необходимо найти минимум $\Delta(a, b) = \sum_{k=1}^{n} \delta_k^2$, где $\delta_k = (ax_k+b)-y_k$

3. Проверьте, что функция $f(x, y) = (y - x^2)(y - 3x^2)$ не имеет экстремума в точке (0,0), хотя имеет его на любой прямой, проходящей через эту точку.