

Лагранжева механика 5 (2016/2017)

В.А. Побережный

1 Предварительные сведения

1.1 Происхождение вариационных задач

Начнём опять с простейших задач. Материальная точка падает в поле тяжести земли $F = mg$ с высоты h . Нас интересует полное время падения и скорость точки в конечный момент. Уравнения банальные, считали неоднократно – время $T(h) = \sqrt{2h/g}$, скорость $V(h) = \sqrt{2gh}$.

Теперь, пусть точка падает не свободно, а соскальзывает без трения под действием той же силы с высоты h по клину или наклонной плоскости образующей угол α с горизонтальной плоскостью. Задача не сильно усложнилась, надо просто спроецировать силы, тоже уже решали. Ответы $T(h) = (\sqrt{2h/g})/\sin \alpha$ и $V(h) = \sqrt{2gh}$. Заметим, что скорость не изменилась, этого конечно же следовало ожидать, так как полная энергия сохраняется, сила, действующая на точку потенциальна, работа (изменение потенциальной энергии) зависит только от разности высот, а не от конкретного пути, но тогда и кинетическая энергия, а вместе с ней и модуль вектора скорости зависят только от разности высот.

Следующий шаг – задача и результат Галилея. В вертикальной плоскости лежит окружность радиуса большего h . Из нижней её точки (точка C) проводим хорду такую чтобы её конец (точка A) лежал на высоте h . На вырезанной хордой дуге круга берём любую точку B . Утверждение: время соскальзывания по ломанной ABC меньше времени соскальзывания по хорде AC . Задача вполне счётная, в рамках элементарной геометрии, но всё-таки требует некоторых выкладок и аккуратности, для сбережения времени оставим её в качестве упражнения. Следствие – время падения по любой вписанной ломанной меньше времени падения по хорде. Что, в пределе, даёт окончательный результат – соскальзывание по дуге круга происходит быстрее чем по хорде или любой вписанной ломанной. Конечная скорость всё та же $V(h) = \sqrt{2gh}$. Интересное элементарно проверяемое замечание – время соскальзывания по любой хорде одно и то же (второй конец хорды должен лежать не выше центра круга).

Следующее естественное обобщение уже приводит нас к одному из истоков вариационного исчисления - знаменитой задаче о брахистохроне, сформулированной Бернулли вскоре после задачи Галилея. Пусть даны лежащие на разной высоте в вертикальной плоскости точки A и B . Найти кривую, по которой материальная точка за минимальное время соскользнёт из одной точки в другую под действием силы тяжести. То есть по сути нам требуется найти точку минимума отображения из “всех кривых из A в B ” в действительные числа, сопоставляющего кривой время соскальзывания по

ней. Пространство “всех кривых” устроено сложнее чем привычное нам \mathbb{R}^n , оно, например, бесконечномерно, и для работы с ним и интересующим нас отображением следует ввести определённый формализм, чем мы сейчас и займёмся.

Вообще задачи подобного рода по поиску экстремумов отображений в разнообразных нетривиальных пространствах называются вариационными и часто возникают в самых различных областях от алгебраической геометрии до теоретической физики. Самые очевидные примеры относятся к дифференциальной геометрии : задача о геодезических – среди всех кривых на поверхности P соединяющих точки A и B найти кривую минимальной длины; изопериметрическая задача – среди всех замкнутых плоских кривых заданной длины l найти кривую ограничивающую максимальную площадь; проблема Плато – затянуть заданный контур в \mathbb{R}^3 плёнкой наименьшей площади.

1.2 Касательные пространства и дифференциалы

Итак, мы хотим научиться искать экстремумы у отображений из каких-то множеств в вещественные числа. Для начала вспомним, что известно по этому вопросу для обычных функций на \mathbb{R} и \mathbb{R}^n . Для таких функций нам хорошо известны необходимые и достаточные условия минимумов и максимумов в том случае, когда эти функции дифференцируемы. Вспомним как устроены эти условия и посмотрим что требуется для обобщения на случай более сложных чем \mathbb{R}^n пространств.

Во-первых, раз уж мы ограничиваемся дифференцируемыми функциями, стоит вспомнить их определение. Стандартный “школьный” вариант – определённая на \mathbb{R} функция f дифференцируема в точке x_0 если существует предел

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

Определение не очень удобное не только для наших целей, но даже и для обобщения на функции на \mathbb{R}^n . Перепишем его в более подходящем нам виде, но сначала выясним в каких пространствах живут участвующие в определении объекты. Итак, точки $x_0 + h$ и x_0 в которых мы смотрим значения функции это точки **аффинного (не векторного!)** пространства \mathbb{R} . То есть складывать их вообще нельзя, а вот их разность, то есть h определена хорошо и является элементом **векторного** (вещественного одномерного) пространства V_{x_0} . Пространство V_{x_0} это **касательное** пространство к \mathbb{R} в точке x_0 .

Теперь мы готовы написать более правильное определение дифференцируемости: заданная на \mathbb{R} функция f дифференцируема в точке x_0 если существует такой линейный оператор $A_{x_0} : V_{x_0} \rightarrow V_{f(x_0)}$ что

$$f(x_0 + h) = f(x_0) + A_{x_0}(h) + o(|h|)$$

Как устроены все на свете линейные отображения из одномерного векторного пространства в одномерное? Пространства векторные, значит достаточно задать действие на базисе, пространства одномерны, значит действие на базисе задаётся одним числом, и отображение это умножение на константу. Эта константа есть не что иное как значение производной в точке x_0 . То есть $A_{x_0}(h) = f'(x_0) \cdot h$.

В чём выигрыш от нашей переформулировки? В том, что новое определение элементарно обобщается не только на функции на \mathbb{R}^n , но и на функции на многообразиях, на бесконечномерных векторных пространствах, причём не только на функции, но и на отображения между этими объектами.

Например, для \mathbb{R}^n получается: функция $f : \mathbb{R}^n \rightarrow \mathbb{R}$ дифференцируема в точке \mathbf{x}_0 если существует линейное отображение $A_{\mathbf{x}_0} : V_{\mathbf{x}_0}^n \rightarrow V_{f(\mathbf{x}_0)}$ такое что

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + A_{\mathbf{x}_0}(\mathbf{h}) + o(\|\mathbf{h}\|)$$

Как тут устроено отображение $A_{\mathbf{x}_0}$? Это линейное отображение из n -мерного векторного пространства в одномерное. Все линейные отображения между конечномерными векторными пространствами определяются действием на базисе и значит задаются постоянными матрицами, в этом нашем случае матрица будет просто строчкой, а её элементы суть частные производные f по соответствующим координатам:

$$A_{\mathbf{x}_0}(\mathbf{h}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right) \cdot \mathbf{h}$$

Для всех случаев участвующее в определении линейное отображение A называется **дифференциалом** отображения f в точке x_0 , иногда также обозначается через $df_{\mathbf{x}_0}$. По сути, наше определение говорит, что отображение является дифференцируемым если у него существует дифференциал.

А как теперь устроены условия экстремума? Очень просто, в этих терминах необходимое условие – если отображение имеет экстремум в точке \mathbf{x}_0 , то его дифференциал в этой точке равен нулю. Достаточное условие тоже имеет простой вид, но дифференциала уже недостаточно, оно формулируется в терминах квадратичного отображения касательного пространства, то есть второго члена в разложении функции f в ряд Тейлора. Оставляем как упражнение.

Итак, возвращаясь к изначально нас интересовавшему вопросу, если у нас имеется отображение из некоторого пространства B например в вещественные числа, что надо потребовать от устройства B чтобы смочь адаптировать к поиску экстремумов на нём методы конечномерного евклидова пространства? Давайте для краткости, чтобы не упоминать лишний раз касательные пространства сразу положим что B это векторное пространство, очевидно это не ограничивает общности и элементарно модифицируется к общему случаю. Далее, для прямого обобщения необходимо определить на B дифференцируемые функции, ведь условия экстремума мы знаем именно для них. То есть, нам надо определить что такое дифференциал отображения или функции на B . В определении дифференциала входит о маленькое от нормы, значит, как минимум нам нужна какая-то норма на пространстве B . То есть пространство B должно быть нормированным. Ну и поскольку в определении фигурирует не только норма, но и предельный переход, стоит также попросить чтобы B относительно этой нормы было полно. Этого оказывается вполне достаточно для построения дифференциального исчисления на B .

1.3 Банаховы пространства

Определение 1 *Векторное пространство, являющееся нормированным и относительно этой нормы полным, называется Банаховым пространством.*

ством.

Очевидный пример банахова пространства это обычное \mathbb{R}^n с евклидовой нормой. Также банаховыми являются пространства l^p и $L^p[a, b]$. Для наших задач наиболее важными будут пространства $C^k[a, b]$.

Определение 2 $C^k[a, b]$ это пространство функций $\varphi(x)$ непрерывных на отрезке $[a, b]$ вместе с производными до k -ой включительно, с нормой

$$\|\varphi\|_k = \max_{x \in [a, b]} |\varphi(x)| + \max_{x \in [a, b]} |\varphi'(x)| + \dots + \max_{x \in [a, b]} |\varphi^{(k)}(x)|$$

Предложение 1 Пространство $C^0[a, b]$ банахово.

Пространство векторное, норма предъявлена, надо только проверить полноту. Проверяем: пусть $\varphi_n \in C^0[a, b]$ - фундаментальная последовательность: $\forall \varepsilon > 0, \exists N_\varepsilon$ такое что

$$\max_{x \in [a, b]} |\varphi_n(x) - \varphi_m(x)| < \varepsilon, \quad n, m > N_\varepsilon$$

Зафиксируем точку $x_0 \in [a, b]$. Тогда последовательность чисел $\varphi_n(x_0)$ лежащая в \mathbb{R} фундаментальна. Пространство \mathbb{R} банахово, следовательно существует предел $\lim_{n \rightarrow \infty} \varphi_n(x_0)$. Обозначим его $\varphi(x_0)$. Остаётся показать, что при изменении x_0 величина $\varphi(x_0)$ меняется непрерывно. Ну очевидно что делать следует - выбираем n_ε так чтобы при $n > n_\varepsilon$ выполнялось $|\varphi_n(x) - \varphi(x)| < \varepsilon/3$, фиксируем какое-нибудь $n_0 > n_\varepsilon$ и выбираем δ такое чтобы $|\varphi_{n_0}(x) - \varphi_{n_0}(x_0)| < \varepsilon/3$ при $|x - x_0| < \delta$. Получаем:

$$\begin{aligned} |\varphi(x) - \varphi(x_0)| &= |(\varphi(x) - \varphi_{n_0}(x)) + (\varphi_{n_0}(x) - \varphi_{n_0}(x_0)) + (\varphi_{n_0}(x_0) - \varphi(x_0))| \leq \\ &\leq |\varphi(x) - \varphi_{n_0}(x)| + |\varphi_{n_0}(x) - \varphi_{n_0}(x_0)| + |\varphi_{n_0}(x_0) - \varphi(x_0)| \leq \\ &\leq \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon \end{aligned}$$

Предложение 2 Пространство $C^1[a, b]$ банахово.

То есть теперь от предельной функции требуется не только непрерывность, но ещё и непрерывная дифференцируемость на отрезке.

Нам потребуется элементарный вспомогательный результат - если последовательность $\varphi_n(x)$ сходится в $C^0[a, b]$, то последовательность чисел $\int_a^b \varphi_n(x) dx$ сходится в \mathbb{R} и предел равен $\int_a^b \varphi(x) dx$.

Теперь, пусть последовательность $\varphi_n(x)$ фундаментальна в $C^1[a, b]$. Тогда в силу

$$\begin{aligned} \|\varphi_n - \varphi_m\|_0 &\leq \|\varphi_n - \varphi_m\|_1 \\ \|\varphi'_n - \varphi'_m\|_0 &\leq \|\varphi_n - \varphi_m\|_1 \end{aligned}$$

последовательности $\varphi_n(x)$ и $\varphi'_n(x)$ фундаментальны в $C^0[a, b]$. Значит, существуют предельные функции $\varphi(x) = \lim_{n \rightarrow \infty} \varphi_n(x)$ и $\psi(x) = \lim_{n \rightarrow \infty} \varphi'_n(x)$. Остаётся показать, что $\varphi(x)$ дифференцируема, и её производная равна $\psi(x)$. Смотрим, с одной стороны (вспомогательное утверждение)

$$\lim_{n \rightarrow \infty} \int_a^{x_0} \varphi'_n(x) dx = \int_a^{x_0} \psi(x) dx,$$

а с другой, (по определению интеграла и производной)

$$\lim_{n \rightarrow \infty} \int_a^{x_0} \varphi'_n(x) dx = \lim_{n \rightarrow \infty} (\varphi_n(x_0) - \varphi_n(a)) = \varphi(x_0) - \varphi(a)$$

Итак,

$$\varphi(x) - \varphi(a) = \int_a^x \psi(t) dt$$

и значит, из непрерывности $\psi(t)$ следует

$$\varphi'(x) = \frac{d}{dx} \int_a^x \psi(t) dt = \psi(x)$$

Всё. Доказательство общего случая (индукцией по k) остаётся в качестве упражнения.

Теорема 1 При всех $k \in \mathbb{N}$ пространство $C^k[a, b]$ банахово.

2 Дифференциальное исчисление в банаховом пространстве

Продолжим построение техники работы с экстремальными задачами в банаховом пространстве. Нам предстоит определить для них понятие дифференцируемости отображения и соответственно дифференциала.

2.1 Линейные отображения

Важно отметить, что отличие от конечномерного случая для банаховых пространств B_1 и B_2 из линейности отображения $A : B_1 \rightarrow B_2$ между ними еще не следует его непрерывность.

Определим норму линейного отображения A следующим образом:

$$\|A\| = \sup_{h \neq 0} \frac{\|Ah\|_2}{\|h\|_1} = \sup_{\|h\|_1=1} \|Ah\|_2$$

Теорема 2 Линейное отображение $A : B_1 \rightarrow B_2$ непрерывно если и только если $\|A\| < \infty$

В одну сторону. Пусть $\|Ah\|_2 \leq \|A\| \|h\|_1$. Тогда при $\delta = \varepsilon / \|A\|$ получаем $\|Ah\|_2 \leq \varepsilon$ при $\|h\|_1 \leq \delta$.

В обратную. Пусть A непрерывно. Тогда существует такое $\delta \geq 0$ что при $\|h\|_1 \leq \delta$ выполняется $\|Ah\|_2 \leq 1$. Но тогда

$$\|A\| = \sup_{\|h\|_1 \leq 1} \|Ah\|_2 = \sup_{\|h\|_1 = \delta} \left\| A \frac{h}{\delta} \right\|_2 = \frac{1}{\delta} \sup_{\|h\|_1 = \delta} \|Ah\|_2 \leq \frac{1}{\delta} < \infty$$

Следствие 1 Если линейное отображение A непрерывно в точке x_0 банахова пространства B_1 то оно непрерывно и во всех прочих его точках

Потому что норма A считается в нуле и вообще ни про какие точки ничего не знает

2.2 Дифференциал, вариации

Всё готово для определения дифференциала и дифференцируемости.

Определение 3 *Функции определённые на банаховом пространстве B будем называть функционалами (чтобы отличать от функций которыми зачастую являются сами элементы банахова пространства)*

Определение 4 *Функционал $\alpha : B \rightarrow \mathbb{R}$ называется бесконечно малым в точке $h_0 \in B$ если $\lim_{h \rightarrow h_0} \alpha(h) = 0$.*

Определение 5 *Непрерывное линейное отображение банахова пространства B в вещественные числа называется линейным функционалом на B .*

Определение 6 *Функционал $f : B \rightarrow \mathbb{R}$ называется дифференцируемым в точке $h_0 \in B$ если существует линейный функционал A_{h_0} такой что*

$$f(h_0 + h) = f(h_0) + A_{h_0}(h) + \|h\|\alpha(h)$$

Определение 7 *Линейный функционал A_{h_0} иначе обозначаемый по аналогии с евклидовым случаем δf_{h_0} называется дифференциалом или вариацией отображения f в точке h_0 .*

Как и в евклидовом конечномерном случае дифференциал определён корректно и однозначно. Иначе из

$$A_{h_0}(h) + \|h\|\alpha(h) = f(h_0 + h) - f(h_0) = \tilde{A}_{h_0}(h) + \|h\|\tilde{\alpha}(h)$$

получаем

$$A_{h_0}(h) - \tilde{A}_{h_0}(h) = \|h\|(\tilde{\alpha}(h) - \alpha(h))$$

и

$$A_{h_0}(h) - \tilde{A}_{h_0}(h) = \lim_{n \rightarrow \infty} \frac{A_{h_0}(\frac{h}{n}) - \tilde{A}_{h_0}(\frac{h}{n})}{1/n} = \lim_{n \rightarrow \infty} \left\| \frac{h}{n} \right\| \left(\tilde{\alpha}\left(\frac{h}{n}\right) - \alpha\left(\frac{h}{n}\right) \right) = 0$$

при всех h тождественно.

Приведём элементарные примеры вариаций. Как и в евклидовом пространстве если функционал f линеен, то $\delta f_{h_0} = f$ для всех h_0 . Если же $f(h) = \lambda_0$ для всех h тождественно, то $\delta f_{h_0} = 0$

2.3 Экстремумы

Определение 8 *Точка $h_0 \in B$ является точкой локального максимума функционала f если существует такое $\delta > 0$ что $f(h) \leq f(h_0)$ для всех h таких что $\|h - h_0\| \leq \delta$.*

Теорема 3 *Если всюду дифференцируемый функционал $f : B \rightarrow \mathbb{R}$ достигает своего максимума или минимума в точке $h_0 \in B$ то тогда $\delta_{h_0} f = 0$*

Доказательство тривиально следует из евклидоваго случая. Пусть h_0 точка максимума для f . Предположим, существует \tilde{h} такая что $\delta f_{h_0}(\tilde{h}) \neq 0$. Тогда вещественная функция

$$\varphi(t) = f(h_0 + \tilde{h}t) - f(h_0)$$

будет дифференцируемой при $|t| < 1$ и будет иметь максимум при $t = 0$. Значит, её производная в этой точке обязана быть равной нулю. Но

$$\varphi'(0) = \delta f_{h_0}(\tilde{h}) \neq 0$$

3 Задачи, которые надо уметь решать после этой лекции:

Задача: Показать что в задаче Галилея время соскальзывания по хорде не зависит от выбора хорды и что оно больше времени падения по двузвенной ломанной.

Задача: Выписать явный вид (формулу) функционала сопоставляющего непрерывно дифференцируемой плоской кривой с концами в точках A и B время соскальзывания материальной точки по этой кривой под действием силы тяжести. Можно пользоваться сохранением энергии.

Задача: Показать, что если функции $a(x), b(x), c(x)$ лежат в $C^1[a, b]$ то

- функционал

$$f(h) = \int_a^b (a(x)h(x) + b(x)h'(x)) dx$$

непрерывен на $C^1[a, b]$

- определённый на $C^1[a, b]$ функционал

$$g(h) = \frac{\int_a^b (a(x)h^2(x) + b(x)h(x)h'(x) + c(x)h'^2(x)) dx}{\|h\|}$$

бесконечно мал в нуле