Простейшая задача вариационного исчисления

В курсе математического анализа были рассмотрены задачи отыскания экстремумов функций как одной, так и многих переменных.

Еще в Древней Греции было известно, что линией заданной фиксированной длины, ограничивающей на плоскости фигуру с максимальной площадью, является окружность.

Другой оптимизационной задачей, не решаемой методами конечномерного математического анализа, которую сформулировал в конце XVII века Иоганн Бернулли, является так называемая *задача о брахистохроне*, имеющая следующую формулировку.

> Пусть в вертикальной плоскости даны две не лежащие на одной вертикали точки A и B. Требуется найти гладкую траекторию, соединяющею эти точки, по которой материальная точка под действием силы тяжести скатится из A в B за минимальное время.

Приведем формализованную постановку этой задачи.

Пусть начало координат находится в точке A, ось Ox направлена горизонтально влево, а ось Oy – вертикально вниз. Пусть точка B имеет координаты $\{P,\ Q\}$.

Требуется найти гладкую линию, являющуюся графиком функции $y(x)$ (параметрически $\left\{\begin{array}{l} x = x(t), \\ y = y(t), \end{array}\right.$), начав по которой при $t = 0$ движение из A, под действием силы тяжести материальная точка попадет в B за минимально возможное время.

Как известно из курса физики, скорость движения материальной точки в однородном поле тяжести равна $v = \dfrac{dS}{dt} = \sqrt{2gy}$, в то время как дифференциал длины дуги траектории $dS = \sqrt{1 + y'^2}\, dx$.

Откуда

$$dt = \frac{dS}{v} = \frac{\sqrt{1+y'^2}\,dx}{\sqrt{2gy}}.$$

Окончательно получаем, что решением задачи является функция $y(x)$, минимизирующая выражение вида

$$J\Big(y(x)\Big) = \frac{1}{\sqrt{2g}}\int\limits_0^P \sqrt{\frac{1+y'^2(x)}{y(x)}}\,dx$$

при условиях: $y(0)=0$ и $y(P)=Q$.

Рассмотрим теперь следующую задачу.

Пусть $F(x, y, p)$ непрерывно дифференцируемая при всех $x \in [a, b]$, $y \in (-\infty, +\infty)$ и $p \in (-\infty, +\infty)$ функция.

Рассмотрим функционал

$$J(y) = \int\limits_a^b F\Big(x, y(x), y'(x)\Big)dx \tag{7.1.2}$$

на множестве $\mathcal{C}^1_{AB}[a, b] \subset \mathcal{C}^1[a, b]$ функций $y(x)$, удовлетворяющих условиям $y(a) = A$ и $y(b) = B$.

Здесь мы использовали обозначения:

$\mathcal{C}^1[a,b]$ — множество всех непрерывно дифференцируемых на $[a,b]$ вещественных функций, расстояние между которыми определяется формулой

$$\rho\,(y_1,y_2)=\max_{x\in[a,b]}|y_1(x)-y_2(x)|+\max_{x\in[a,b]}|y_1'(x)-y_2'(x)|\quad\forall y_1,y_2\in\mathcal{C}^1[a,b]\,.$$

Заметим, что множество $\mathcal{C}^1[a,b]$ является линейным нормированным пространством с нормой

$$\langle y\rangle=\max_{x\in[a,b]}|y(x)|+\max_{x\in[a,b]}|y'(x)|\ . \tag{7.1.1}$$

Проверьте самостоятельно (это будет полезно для дальнейшего), что множество $\mathcal{C}^1_{00}[a,b]$ является линейным пространством, а множество $\mathcal{C}^1_{AB}[a,b]$ при $|A|+|B|\neq 0$ – нет.

Определение 7.1.1

Будем говорить, что функционал (7.1.2) достигает на функции $y^*(x) \in \mathcal{C}^1_{AB}[a,b]$ *слабого локального минимума* (*максимума*), если найдется число $\varepsilon > 0$ такое, что

$$\forall y(x) \in \mathcal{C}^1_{AB}[a,b] \qquad \text{с} \qquad \rho\,(\, y(x),\, y^*(x)\,) < \varepsilon$$

выполняется неравенство

$$J(y) \geq J(y^*) \quad (\, J(y) \leq J(y^*)\,)\ .$$

Если неравенства строгие при $y \neq y^*$, то говорят о *строгом* экстремуме. В случае, когда неравенства удовлетворяются *для всех* функций $y(x) \in \mathcal{C}^1_{AB}[a,b]$, экстремум *абсолютный.*

Задачу отыскания слабого локального экстремума при использовании нормы (7.1.1) называют *простейшей вариационной задачей* или же *задачей с закрепленными концами.*

Основным инструментом при исследовании на экстремум функционала вида (7.1.2) служит его *вариация* – функционал, являющийся аналогом производной по направлению от функции многих переменных.

Для его определения нам потребуется еще одно подмножество в пространстве $\mathcal{C}^1[a,b]$, а именно $\mathcal{C}^1_{00}[a,b]$ – множество непрерывно дифференцируемых на $[a,b]$ функций $h(x)$ таких, что $h(a)=h(b)=0$.

Заметим, что при любом вещественном параметре α функция $y(x,\alpha) = y(x) + \alpha h(x) \in \mathcal{C}^1_{AB}[a,b]$, если $h(x) \in \mathcal{C}^1_{00}[a,b]$. Это свойство дает основание называть $\alpha h(x)$ *допустимым приращением* (или, как иногда говорят, *допустимой вариацией*) $y(x)$ – аргумента исследуемого функционала (7.1.2).

Наконец, рассматривая при малых по модулю α множество значений функционала

$$J(y+\alpha h) = \int\limits_a^b F\left(x, y(x) + \alpha h(x), y'(x) + \alpha h'(x)\right) dx\,, \tag{7.1.3}$$

можно делать заключения об экстремальных свойствах исходного функционала (7.1.2) в малой окрестности функции $y(x)$ (в пространстве $\mathcal{C}^1[a,b]$).

Более конкретно, величину и направление изменения $J(y+\alpha h)$ (как функции параметра α при фиксированных $y(x)$ и $h(x)$) можно оценивать числом $\left.\dfrac{dJ(y+\alpha h)}{d\alpha}\right|_{\alpha=0}$.

В сделанных предположениях (для малых по модулю значений α) функционал $J(y+\alpha h)$ является непрерывно дифференцируемой функцией α.

С другой стороны, (7.1.3) можно рассматривать как собственный интеграл, зависящий от параметра α, для которого справедлива теорема Лейбница, утверждающая, что

$$\left.\frac{dJ(y+\alpha h)}{d\alpha}\right|_{\alpha=0} =$$

$$= \int\limits_a^b \left(\frac{\partial F\left(x, y(x), y'(x)\right)}{\partial y} h(x) + \frac{\partial F\left(x, y(x), y'(x)\right)}{\partial y'} h'(x) \right) dx \; . \quad (7.1.4)$$

Определение 7.1.2	Выражение $$\left.\frac{dJ(y+\alpha h)}{d\alpha}\right\|_{\alpha=0}$$ называется *первой вариацией* функционала $J(y)$ на функции $y(x)$ при $\forall h(x) \in \mathcal{C}^1_{00}[a,b]$. Первую вариацию принято обозначать $\delta J(y,h)$.

Обратите внимание на структурное сходство формулы (7.1.4) с формулой

$$\frac{\partial F}{\partial l} = \sum_{k=1}^{n} \frac{\partial F}{\partial \xi_k}\,\omega_k\ ,$$

определяющей в E^n величину производной функции $F\,(\xi_1,\,\xi_2,\,\dots\,\xi_n)$ по направлению $\|l\| = \|\,\omega_1,\,\omega_2,\,\dots,\,\omega_n\|^{\mathrm{T}}$.

Необходимое условие существования слабого экстремума, используя понятие первой вариации, формулирует

Теорема 7.1.1 **Если $y^*(x)$ есть решение простейшей вариационной задачи, то $\delta J(y^*,h) = 0\ \ \forall h(x) \in \mathcal{C}^1_{00}[a,b]$.**

Доказательство.

Пусть для определенности функционал $J(y)$ имеет на функции $y^*(x) \in \mathcal{C}^1_{AB}[a,b]$ слабый локальный минимум. Тогда, согласно определению 7.1.1, в $\mathcal{C}^1_{AB}[a,b]$ существует некоторая окрестность $U_\varepsilon(y^*)$ такая, что $\forall y(x) \in U_\varepsilon(y^*)$ выполнено неравенство

$$J(y(x)) \geq J(y^*(x)).$$

При этом $y(x)$ может быть представлена в виде $y^*(x)+\alpha h(x)$, где $h(x) \in \mathcal{C}^1_{00}[a,b]$. И для *непрерывно дифференцируемой по параметру* α функции $J(y^*(x)+\alpha h(x))$ верно неравенство

$$J\left(y^*(x)+\alpha h(x)\right) \geq J(y^*(x)) \quad \forall h(x) \in \mathcal{C}^1_{00}[a,b].$$

Значит функция $J(y^*(x)+\alpha h(x))$ имеет минимум в $\alpha = 0$ и, следовательно, ее производная

$$\left.\frac{dJ(y^*+\alpha h)}{d\alpha}\right|_{\alpha=0} = \quad \delta J(y^*,h) \quad = \quad 0$$

(в силу (7.1.4) и определения 7.1.2) для любой фиксированной функции $h(x) \in \mathcal{C}^1_{00}[a,b]$.

Теорема доказана.

При использовании теоремы 7.1.1 необходимо убеждаться в равенстве нулю первой вариации $\delta J(y^*, h)$ одновременно *для всех* функций $h(x) \in \mathcal{C}^1_{00}[a, b]$, что может оказаться непростой задачей.

Более удобное для практического использования необходимое условие экстремума в случае простейшей вариационной задачи можно получить (следуя Лагранжу), проверив, что верна (часто называемая *основной леммой вариационного исчисления*)

Лемма 7.1.1

Если $f(x)$ **непрерывна на** $[a, b]$ **и**

$$\int\limits_a^b f(x)h(x)\,dx = 0 \qquad \forall h(x) \in \mathcal{C}^1_{00}[a, b],$$

то $f(x) \equiv 0$ **на** $[a, b]$.

Доказательство.

Предположим противное. Пусть $f(x) \not\equiv 0,\ x \in [a,b]$. Тогда $\exists\, x_0 \in [a,b]$ такое, что $f(x_0) \neq 0$. Например, пусть $f(x_0) > 0$.

В силу непрерывности $f(x)$ найдутся $x_0 \in (a,b)$ и $\Delta > 0$ такие, что

$$f(x) \geq \frac{1}{2} f(x_0) \quad \forall x \in [x_0 - \Delta, x_0 + \Delta] \ \subseteq (a,b).$$

В $\mathcal{C}^1_{00}[a,b]$ выберем функцию $h(x) =$

$$= \begin{cases} (x - (x_0 - \Delta))^2 (x - (x_0 + \Delta))^2, & x \in [x_0 - \Delta, x_0 + \Delta], \\ 0, & x \notin [x_0 - \Delta, x_0 + \Delta]. \end{cases}$$

Согласно интегральной теореме о среднем имеем оценку

$$\int\limits_a^b f(x) h(x)\, dx = \int\limits_{x_0-\Delta}^{x_0+\Delta} f(x) h(x)\, dx =$$

$$= f(\xi) \int\limits_{x_0-\Delta}^{x_0+\Delta} h(x)\, dx \quad \geq \quad \frac{1}{2} f(x_0) \int\limits_{x_0-\Delta}^{x_0+\Delta} h(x)\, dx > 0\ ,$$

где $\xi \in (x_0 - \Delta, x_0 + \Delta)$. Но это противоречит условию леммы.

Лемма доказана.

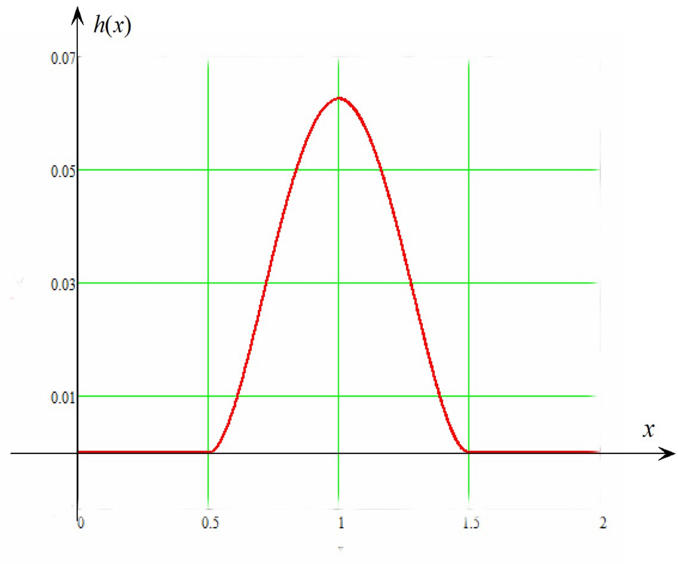

Рис. 1. К доказательству леммы 7.1.1.

Лемма 7.1.1 позволяет получить необходимое условие простейшей вариационной задачи в следующей упрощенной форме, которую дает

Теорема 7.1.2 **Пусть** $F(x, y, p)$ **дважды непрерывно дифференцируемая при всех** $x \in [a, b]$, $y \in (-\infty, +\infty)$ **и** $p \in (-\infty, +\infty)$ **функция. Если дважды непрерывно дифференцируемая функция** $y^*(x)$, $x \in [a, b]$ **есть решение простейшей вариационной задачи, то она удовлетворяет** *уравнению Эйлера*

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} = 0 \ . \tag{7.1.5}$$

Доказательство.

Поскольку $y^*(x)$ есть решение простейшей вариационной задачи, то $\delta J(y^*, h) = 0$ для любой $h(x) \in \mathcal{C}^1_{00}[a, b]$. С учетом формулы (7.1.4) это дает

$$0 = \int\limits_a^b \left(\frac{\partial F\left(x, y(x), y'(x)\right)}{\partial y} h(x) + \frac{\partial F\left(x, y(x), y'(x)\right)}{\partial y'} h'(x) \right) \Bigg|_{y=y^*} dx.$$

В условиях теоремы второе слагаемое в подынтегральной функции можно проинтегрировать по частям

$$0 = \frac{\partial F\left(x, y(x), y'(x)\right)}{\partial y'} h(x) \Bigg|_a^b + \int\limits_a^b \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right)_{y=y^*} h(x)\, dx =$$

$$= \int\limits_a^b \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right)_{y=y^*} h(x)\, dx \ ,$$

поскольку $h(a) = h(b) = 0$.

Из непрерывности функции

$$f(x) = \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right) \Bigg|_{y=y^*}$$

и основной леммы вариационного исчисления (лемма 7.1.1) следует, что $y^*(x)$ удовлетворяет уравнению Эйлера (7.1.5).

Теорема доказана.

Определение 7.1.3	Всякое решение уравнения Эйлера (7.1.5) называется *экстремалью* функционала $J(x, y, y')$. В случае, когда эта экстремаль принадлежит множеству $\mathcal{C}^1_{AB}[a, b]$, она называется *допустимой экстремалью.*

Сделанное при выводе уравнения Эйлера предположение о непрерывности второй производной допустимой экстремали оказывается излишним, если использовать, приводимою здесь без доказательства, следующую лемму.

Лемма 7.1.2 (Дюбуа-Реймона) **Если** $f(x)$ **непрерывна на** $[a,b]$ **и**

$$\int\limits_a^b f(x)h'(x)\,dx = 0 \qquad \forall h(x) \in \mathcal{C}^1_{00}[a,b],$$

то $f(x) \equiv \text{const}$ **на** $[a,b]$.

Действительно, пусть

$$G(x) = \int\limits_a^x \frac{\partial F(u, y(u), y'(u))}{\partial y}\, du\,.$$

Тогда, в силу теоремы 7.1.1 и формулы 7.1.3, имеем

$$\delta J(y^*, h) = 0 = \int\limits_a^b \left(G'(x)h + \frac{\partial F}{\partial y'} h' \right) dx =$$

(интегрируя первое слагаемое по частям)

$$= G(x)h\Bigg|_a^b - \int\limits_a^b G(x)h'dx + \int\limits_a^b \frac{\partial F}{\partial y'} h'dx = \int\limits_a^b \left(\frac{\partial F}{\partial y'} - G(x) \right) h'dx\,,$$

если учесть, что $h(x) \in \mathcal{C}^1_{00}[a, b]$.

Применив теперь утверждение леммы Дюбуа-Реймона, получаем *интегральную форму* уравнения Эйлера

$$\frac{\partial F}{\partial y'} - \int\limits_a^x \frac{\partial F}{\partial y}\,du \equiv const\,.$$

Откуда, окончательно следует, что

$$\frac{d}{dx}\frac{\partial F}{\partial y'} - \frac{\partial F}{\partial y} = 0\,.$$

Полученное условие экстремальности является необходимым, но не достаточным. Поэтому, в тех случаях, когда допустимая экстремаль однозначно определяется уравнением Эйлера и граничными условиями, целесообразно попытаться выполнить исследование на экстремальность непосредственно по его определению.

Использование этого метода иллюстрирует

Задача 7.1.1 Решить простейшую вариационную задачу

$$J(y) = \int\limits_0^1 \left(x^3 + \frac{1}{2} y^2 + 2y'^2 \right) dx \ , \quad y(0) = 0, \ y(1) = 2 \ .$$

`Решение.` Заметим, что

$$J(y)=\frac{1}{4}+\int\limits_0^1\left(\frac{1}{2}y^2+2y'^2\right)\,dx$$

и исследование на экстремальность достаточно выполнить для функционала $J(y)=\int\limits_0^1\left(\frac{1}{2}y^2+2y'^2\right)\,dx$.

Составим и решим уравнение Эйлера. Имеем

$$\frac{\partial F}{\partial y} = y \ , \quad \frac{\partial F}{\partial y'} = 4y' \ , \quad \frac{d}{dx}\frac{\partial F}{\partial y'} = 4y'' \ .$$

Тогда, $\dfrac{d}{dx}\dfrac{\partial F}{\partial y'} - \dfrac{\partial F}{\partial y} = 0 \quad \Longrightarrow \quad 4y'' - y = 0 \ .$

Общее решение этого уравнения

$$y(x) = C_1 \exp\left(\frac{x}{2}\right) + C_2 \exp\left(-\frac{x}{2}\right)$$

есть множество всех экстремалей, в том числе и допустимых. Граничные условия есть система уравнений

$$\left\{ \begin{array}{ccccc} C_1 & + & C_2 & = & 0 \ , \\ C_1 e^{\frac{1}{2}} & + & C_2 e^{-\frac{1}{2}} & = & 2 \ . \end{array} \right.$$

Откуда находим, что

$$C_1 = -C_2 = \frac{1}{\operatorname{sh}\dfrac{1}{2}}\,,$$

и единственная *допустимая* экстремаль дается формулой

$$y^*(x) = \frac{2\operatorname{sh}\dfrac{x}{2}}{\operatorname{sh}\dfrac{1}{2}}\,.$$

Исследуем найденную допустимую экстремаль на оптимальность. Пусть $h(x)$ – произвольная пробная функция из класса $\mathcal{C}^1_{00}[0,1]$.

Оценим знак приращения функционала

$$J(y^* + h) - J(y^*) =$$

$$= \int\limits_0^1 \left(\frac{(y^* + h)^2}{2} + 2((y^*)' + h')^2 - \frac{(y^*)^2}{2} - 2((y^*)')^2 \right) dx =$$

$$= \int\limits_0^1 y^* h \, dx + 4 \int\limits_0^1 (y^*)' h' \, dx + \int\limits_0^1 \left(\frac{h^2}{2} + 2h'^2 \right) dx =$$

(проинтегрировав второй интеграл по частям и перегруппировав слагаемые)

$$= \int\limits_0^1 (y^* - 4(y^*)'') \, h \, dx + 4(y^*)' h \Big|_0^1 + \int\limits_0^1 \left(\frac{h^2}{2} + 2h'^2 \right) dx \geq 0,$$

поскольку первый интеграл равен нулю в силу равенства $y^* - 4(y^*)'' = 0$, а проинтегрированная часть есть ноль по свойству пробной функции $h(0) = h(1) = 0$.

Таким образом, приходим к заключению о том, что допустимая экстремаль $y^*(x)$ доставляет исследуемому функционалу абсолютный минимум.

Решение получено.

Допустимая экстремаль, находимая из уравнения Эйлера, вообще говоря, не обязательно является решением простейшей вариационной задачи. Этот факт демонстрирует

Задача 7.1.2 Решить простейшую вариационную задачу

$$J(y) = \int\limits_0^{\pi} \left(y'^2 - \frac{25}{4} y^2 \right) dx \ , \quad y(0) = 1, \ y(\pi) = 2 \ .$$

Решение. Уравнение Эйлера в данном случае имеет вид

$$y'' + \frac{25}{4} y = 0 \; .$$

Его общее решение есть

$$y(x) = C_1 \cos\frac{5x}{2} + C_2 \sin\frac{5x}{2} \; ,$$

а допустимая экстремаль

$$y^*(x) = \cos\frac{5x}{2} + 2\sin\frac{5x}{2} \; .$$

Возьмем $h(x) \in \mathcal{C}^1_{00}[0,\pi]$ вида $h(x) = \dfrac{1}{k}\sin nx$, где k и n – произвольные натуральные числа. Тогда, выполнив преобразования аналогичные сделанным при решении задачи 7.1.1, получим

$$\Delta J = J(y^* + h) - J(y^*) =$$

$$= \int\limits_0^\pi \left(h'^2 - \frac{25}{4} h^2 \right) dx = \left(n^2 - \frac{25}{4} \right) \frac{\pi}{2k^2}.$$

Откуда следует, что $\forall k \in \mathbb{N}$

$\Delta J < 0$ при $n = 1;\ 2$

и $\Delta J > 0$ при $n \geq 3\,.$

Решение получено. Значит $y^*(x)$ не является решением данной вариационной задачи.

В заключение обсуждения постановки простейшей задачи вариационного исчисления обратим внимание на то, что в определении 7.1.1 в названии типа экстремума присутствует прилагательное «слабый». Возникает вопрос: какой смысл в использовании этого термина?

Ответ заключается в том, что линейное пространство, образованное функциями непрерывно дифференцируемыми на $[a, b]$ не является конечномерным. Действительно, при исследовании функционалов на экстремум необходимо иметь количественную оценку степени близости произвольной пары элементов в этом пространстве. Как известно, одним из способов построения такой оценки является введение нормы элемента (например по формуле (7.1.1), т.е. превращение рассматриваемого линейного пространства в нормированное. Однако, различные типы норм в линейных пространствах, не имеющих базиса, не эквивалентны друг другу и могут приводить при решении экстремальных задач для одного и того же функционала к различным результатам.

Поясним сказанное следующим примером. Рассмотрим как альтернативу линейному пространству непрерывно дифференцируемых на $[a,b]$ функций с нормой (7.1.1) нормированное линейное пространство функций, имеющих непрерывную производную, кроме быть может, конечного числа точек на $[a,b]$, в которых производная имеет разрыв первого рода. Норму во втором пространстве определим по формуле

$$\langle y \rangle = \max_{x \in [a,b]} |y(x)| \ .$$

По сложившейся исторически традицией экстремум с такой с нормой принято называть «сильным».

В качестве упражнения найдите допустимую экстремаль для задачи с условиями

$$J(y) = \int\limits_0^1 {y'}^3 \, dx \ , \qquad y(0) = 0, \ \ y(1) = 1 \ .$$

Используя определение 7.1.1 с разными нормами, покажите, что в этой задаче на допустимой экстремали $y(x) = x$ имеется слабый экстремум, но нет сильного.

В более общем виде связь между условиями существования сильного и слабого экстремума на допустимой экстремали можно сформулировать так:

> необходимое условие слабого экстремума является необходимым условием сильного, а достаточное условие сильного экстремума есть достаточное условие слабого.