Лекция 6

Элементы вариационного исчисления

В приложениях возникают задачи оптимизации на множествах различной природы, то есть необходимо выделить решения с некоторыми заданными свойствами в соответствии с данным критерием. В большинстве задач этот критерий имеет числовую природу и задача оптимизации сводится к нахождению наибольших и наименьших значений некоторого функционала.

Функционал – это обобщение понятия функции. Если на некотором множестве произвольной природы указано правило, которое ставит в соответствие каждому элементу этого множества некоторое число, то на этом множестве задан функционал.

Будем рассматривать функционалы, действующие из линейного нормированного пространства в пространство действительных чисел.

Например, если x_0 – фиксированный вектор гильбертова (евклидова) пространства H, то формула $f(x) = (x, x_0)$ задает линейный функционал на H.

Проблема экстремума функционала. Общая постановка задачи

$$\Phi(y) \to \min, \qquad (1)$$

$$T(y) \geq 0, \qquad (2)$$

$$S(y) = 0, \qquad (3)$$

где $\Phi(y)$ – заданный действительный функционал; переменная y либо свободно меняется в области определения функционала Φ (*задача без дополнительных условий*), либо на y наложены дополнительные условия. Эти условия могут быть представлены как в виде уравнений (3), так и в виде неравенств (2).

Некоторые частные случаи таких задач.

1. **Задачи на максимум и минимум с конечным числом переменных**. Например, для функции $f(x)$, заданной на отрезке $[a, b]$, в мат. анализе решаются следующие вопросы:

- необходимое условие экстремума;
- достаточные условия экстремума;
- существование экстремума (теорема Вейерштрасса);
- единственность экстремума (выпуклость функции на отрезке).

2. **Линейная и нелинейная оптимизация**. Например, задача линейного программирования – задача отыскания максимума (или минимума) линейной функции

$$z = c_1x_1 + c_2x_2 + \ldots + c_nx_n$$

по переменным $x_1, x_2, \ldots, x_n$, удовлетворяющим линейным неравенствам

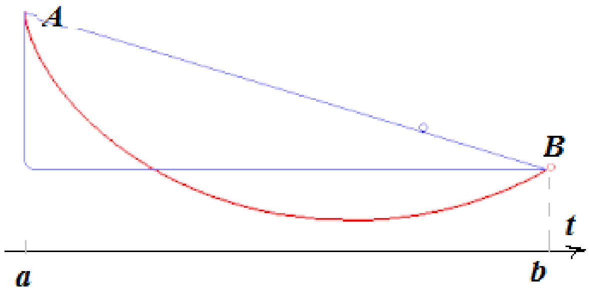
$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \leq b_1, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \leq b_m, \end{cases} \qquad x_j \geq 0, \quad j = \overline{1, n},$$

где c_j, a_{ij}, b_j, $i = \overline{1, m}$; $j = \overline{1, n}$ — постоянные числа.

3. **Комбинаторные задачи.** Сюда относятся задачи, в которых среди возможных комбинаций нужно отыскать комбинацию, минимизирующую некоторый заданный функционал. Например, алгоритм Дейкстры – найти кратчайший путь, который проходит через все данные n точек плоскости.

4. **Задачи вариационного исчисления** – задачи минимизации (максимизации) различных функционалов на заданных подмножествах функциональных пространств (*оптимизация в бесконечномерных пространствах*). В зависимости от вида функционала и пространства, на котором он задан, различают несколько специальных видов вариационных задач, самым распространенным и простым из которых является *классическая (простейшая) вариационная задача*.

Исторически первая вариационная задача была поставлена в 1696 году Иоганном Бернулли. Это **«задача о брахистохроне»**.



Даны две точки A и B, лежащие в вертикальной плоскости. Какова траектория точки, движущейся только под действием силы тяжести, которая начинает двигаться из A и достигает B за кратчайшее время?

Например, при какой форме крыши вода по ней стекает наиболее быстро?

С учетом связи между скоростью движения точки и производной функции, задающей ее траекторию, решение такой задачи естественно искать в пространстве $C^1[a,b]$ непрерывно-дифференцируемых на отрезке $[a,b]$ функций, графики которых проходят через точки A и B. Если обозначим траекторию движения точки $x = x(t)$, $t \in [a,b]$, то математическая модель задачи может быть записана в виде

$$\begin{cases} \Phi[x(t)] = \displaystyle\int_a^b \sqrt{\frac{1+\left(x'(t)\right)^2}{2gx(t)}}\,dt \to \min, \\ x \in C^1[a,b], \\ x(a) = A, \qquad x(b) = B. \end{cases}$$

Решением задачи является дуга циклоиды.

Примером вариационной задачи также является задача нахождения плоской линии, соединяющей две заданные точки и имеющей наименьшую длину.

Исследуемый функционал – длина линии.

Решением данной задачи является отрезок прямой, соединяющей эти точки. Уравнение такой прямой находится однозначно по заданным точкам.

Будем рассматривать пространства функций:

- $C[a,b]$ – пространство непрерывных функций $y(x)$ отрезке $[a,b]$ с нормой

$$\|y(x)\|_C = \max_{x\in[a,b]} |y(x)|;$$

- $C^1[a,b]$ – пространство непрерывно-дифференцируемых функций $y(x)$ отрезке $[a,b]$ с нормой

$$\|y(x)\|_{C^1} = \max\left\{\max_{x\in[a,b]}|y(x)|, \max_{x\in[a,b]}|y'(x)|\right\}$$

или

$$\|y(x)\|_{C^{(1)}} = \max_{x\in[a,b]}|y(x)| + \max_{x\in[a,b]}|y'(x)|.$$

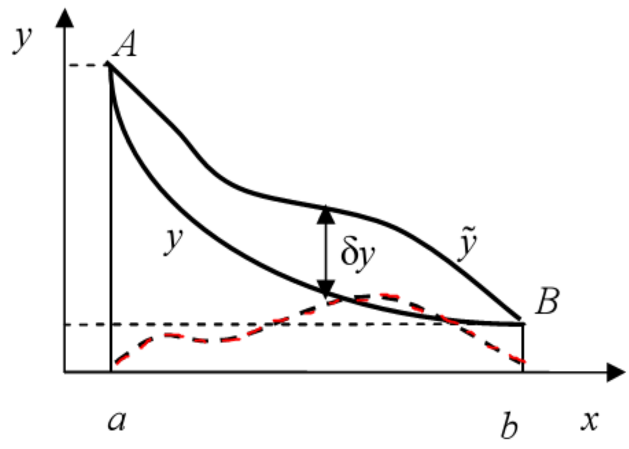
Общая задача вариационного исчисления, вариация функционала

Дан функционал $\Phi[y(x)]$, определенный на множестве функций $y \in Y$, удовлетворяющих заданным граничным условиям.

Требуется среди этих функций найти такую функцию y^*, на которой функционал достигает своего минимального (или максимального) значения, т.е. найдется число $r > 0$ такое, что для любого $y(x) \neq y^*(x)$: $\|y - y^*\| < r$ выполнено неравенство

$$\Phi[y(x)] > (<)\Phi[y^*(x)].$$

Кривая $y^*(x)$ в этом случае называется ***экстремалью***.



Вариацией функции $\delta y(x)$ называется разность между двумя функциями $y(x)$ и $\tilde{y}(x)$, принадлежащих выбранному классу Y:

$$\delta y(x) = \tilde{y}(x) - y(x).$$

Для близких функций $y(x)$ и $\tilde{y}(x)$ в смысле выбранной нормы в Y, вариация $\delta y(x)$ (сама является функцией) близка к нулевой функции.

Приращением функционала $\Phi[y(x)]$, соответствующим приращению $\delta y(x)$ аргумента, называется величина

$$\Delta\Phi = \Delta\Phi[y(x)] = \Phi[y(x) + \delta y(x)] - \Phi[y(x)].$$

Пусть приращение функционала $\Delta\Phi$ можно представить в виде

$$\Delta\Phi = L[\delta y(x)] + o(\|\delta y(x)\|),$$

где $L[\delta y(x)]$ – линейный функционал от $\delta y(x)$ и $\|\delta y(x)\| \to 0$. Тогда главная линейная по отношению к $\delta y(x)$ часть приращения функционала $\Delta\Phi$ называется **(сильной вариацией) вариацией функционала** $\Phi[y(x)]$ и обозначается $\delta\Phi = L[\delta y(x)]$.

Часто используют определение, которое ввел французский математик Лагранж. Согласно этому определению, **вариация функционала (слабая вариация)** может быть получена как значение производной функционала $\Phi[y(x) + t\delta y(x)]$ по параметру t при $t = 0$:

$$\delta\Phi = \left.\frac{\partial\Phi[y(x) + t\delta y(x)]}{\partial t}\right|_{t=0}.$$

Если существует сильная вариация, то существует и слабая вариация. Обратное, вообще говоря, не верно.

Теорема (*необходимое условие экстремума функционала*)

Пусть $y^* \in Y$ – экстремум функционала $\Phi[y]$ и существует вариация $\delta\Phi$. Тогда $\delta\Phi[y^*] = 0$.

Общий метод нахождения минимума функционала: найти его вариацию, после чего решить получившееся функциональное (чаще дифференциальное) уравнение $\delta\Phi[y^*(x)] = 0$.

Если же вариацию найти не удалось или же полученное функциональное уравнение слишком сложное для решения, то используют самый универсальный и во многих случаях единственно возможный метод, предложенный Ритцем (1908 г.). Чаще всего этот метод реализуется следующим образом.

Выбираем в области определения Y функционала Φ некоторый базис, т.е. набор линейно независимых функций $\{\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)\}$, обладающих свойством полноты: любая функция $y(x)$ из области допустимых решений может быть представлена в виде $y(x) = \lim\limits_{n\to\infty} \sum_{k=1}^{n} a_k\, \varphi_k(x)$.

Будем искать приближение к функции, доставляющей минимум функционала $\Phi[y(x)]$, в виде $y^*(x) \approx y_n(x) = \sum_{k=1}^{n} a_k\, \varphi_k(x)$. После подстановки $y_n(x)$ в функционал $\Phi[y(x)]$ получим функцию n переменных:

$$f(a_1, a_2, \dots, a_n) = \Phi[y_n] = \Phi\left(\sum_{k=1}^{n} a_k\, \varphi_k(x)\right).$$

Неизвестные значения коэффициентов разложения $a_1, a_2, \dots, a_n$ искомого решения по функциям базиса находия из условия $\min\limits_{a_1, a_2, \dots, a_n} f(a_1, a_2, \dots, a_n)$.

Таким образом, задача вариационного исчисления сводится к нахождению минимума функции n переменных. Алгоритмы решения этой задачи разработаны. ????? Какие знаете?

Конкретизируем вид функционала и множество допустимых функций.

Предположим, задана функция $F\colon \mathbf{R}^3 \to \mathbf{R}$, такая, что для каждой функции $y(x) \in C^1[a,b]$ существует интеграл $\int_a^b F\big(x, y(x), y'(x)\big)dx$. С помощью этой функции определим на пространстве $\mathrm{C}^1[a,b]$ функционал $J\colon \mathrm{C}^1[a,b] \to \mathbf{R}$ формулой

$$J[y(x)] = \int_a^b F\big(x, y(x), y'(x)\big)dx. \tag{4}$$

Этот функционал будем называть ***интегральным функционалом***.

Определение. Классической (простейшей) задачей вариационного исчисления называется следующая задача условной оптимизации:

$$\begin{cases} J[y(x)] = \displaystyle\int_a^b F\big(x, y(x), y'(x)\big)dx \to \min\ (\max), \\ \qquad y \in C^1[a,b], \\ y(a) = A, \qquad y(b) = B. \end{cases} \tag{5}$$

При этом условия

$$y(a) = A, \qquad y(b) = B \qquad (6)$$

называются ***краевыми условиями*** в задаче, а функции $y(x)$ из пространства $C^1[a,b]$, удовлетворяющие краевым условиям (6), называются ***допустимыми***.

Таким образом, классическая вариационная задача – это задача минимизации интегрального функционала на множестве допустимых функций.

Определение. Говорят, что допустимая функция $y^*(x)$ доставляет в классической вариационной задаче локальный минимум (максимум), если существует $\Delta > 0$ такое, что

$$J[y^*(x)] \leq (\geq) J[y(x)] \qquad (7)$$

для всех $y(x) \in C^1[a,b]$, таких, что

$$\|y(x) - y^*(x)\|_1 \leq \Delta.$$

Точки, являющиеся точками локального минимума или максимума, называются точками ***локального экстремума***.

Если неравенство (7) выполняется для всех допустимых функций $y \in C^1[a,b]$, то y^* называется ***глобальным*** минимумом (максимумом).

Замечание. В вариационном исчислении также рассматривают задачи с подвижными концами, т.е. A и (или) B – произвольные.

Решение задачи проводится в рамках необходимых условий, сформулированных в следующей теореме.

Теорема. Если функция $y = y(x)$ удовлетворяет условиям (6) и доставляет функционалу (7) экстремум, то она является решением уравнения Эйлера:

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0. \qquad (8)$$

В подробной записи уравнение (8) имеет вид

$$\frac{\partial F}{\partial y} - \frac{\partial^2 F}{\partial y' \, \partial y} y' - \frac{\partial^2 F}{\partial y'^2} y'' = 0. \qquad (9)$$

Кривые $\breve{AB}$ $y = y^*(x), x \in [a,b]$, являющиеся графиками функций – решений уравнения Эйлера, называются ***экстремалями***.

Таким образом, решение простейшей задачи вариационного исчисления свелось к решению дифференциального уравнения (8) Эйлера при краевых условиях (6).

После нахождения допустимых экстремалей остается проверить, действительно ли найденные экстремали доставляют экстремум в рассматриваемой задаче.

Достаточное условие отсутствия экстремума. Если функция $F''_{y'y'}(x, y^*(x), y^{*\prime}(x))$ меняет знак на отрезке $[a,b]$, то функция $y^*(x)$ не доставляет локального экстремума для интегрального функционала (5).

Достаточные условия экстремума (по Лежандру) связаны с исследованием слабой вариации функционала (5). Если на экстремали $F''_{y'y'} > 0$, то на экстремали достигается минимум, если $F''_{y'y'} < 0$ – то максимум.

Пример. Найти экстремаль функционала:

$$J[y] = \int_{-1}^{0} (12xy - y'^2)dx$$

с дополнительными условиями $y(-1) = 1,\ y(0) = 0.$

Здесь $F(x, y, y') = 12xy - y'^2$, $F'_y = 12x$, $F''_{yy'} = 0$, $F'_{y'} = -2y'$, $F''_{y'y'} = -2$.
Подставляя в (9), получим уравнение Эйлера

$\frac{\partial F}{\partial y} - \frac{\partial^2 F}{\partial y' \partial y} y' - \frac{\partial^2 F}{\partial y'^2} y'' = 0 \ \Rightarrow\ 12x - 0 \cdot y' - (-2) \cdot y'' = 0 \ \Rightarrow \boldsymbol{y'' = -6x}.$

Интегрируя его, получаем

$$y(x) = -x^3 + C_1 x + C_2.$$

Используя граничные условия, получим $C_1 = C_2 = 0$, $\boldsymbol{y^*(x) = -x^3}$ **– экстремаль?**

Проверим $F''_{y'y'} = -2 < 0$, следовательно, имеет место максимум при $\boldsymbol{y^*(x) = -x^3}$

Вычислим вариацию функционала

$$\Delta J[y^*(x)] = J[y^*(x) + \delta y(x)] - J[y^*(x)] = J[-\boldsymbol{x^3} + \delta y(x)] - J[-\boldsymbol{x^3}] =$$

$$= \int_{-1}^{0} (12x(-\boldsymbol{x^3} + \delta y) - (-\boldsymbol{3x^2} + (\delta y)')^2)dx - \int_{-1}^{0} (12x(-\boldsymbol{x^3}) - (-\boldsymbol{3x^2})^2)dx =$$

$$= \int_{-1}^{0} \left(12x\delta y + \boldsymbol{6x^2}(\delta y)' - \left((\delta y)'\right)^2\right) dx$$

Вычислим по частям интеграл

$$\int_{-1}^{0} \boldsymbol{6x^2}(\delta y)'dx = \left| \begin{matrix} u = \boldsymbol{6x^2} & \boldsymbol{du = 12xdx} \\ dv = (\delta y)'dx & v = \delta y \end{matrix} \right| = \boldsymbol{6x^2}\delta y|_{\boldsymbol{-1}}^{\boldsymbol{0}} - \int_{-1}^{0} \boldsymbol{12x}(\delta y)dx.$$

Заметим, что $\delta y(-1) = \delta y(0) = 0$. Почему?

Далее вычисляем вариацию функционала

$$\Delta J[y^*(x)] = \int_{-1}^{0} 12x\delta y dx + \int_{-1}^{0} \boldsymbol{6x^2}(\delta y)'dx - \int_{-1}^{0} ((\delta y)')^2 dx =$$

$$= \int_{-1}^{0} 12x\delta y dx + 0 - \int_{-1}^{0} 12x\delta y dx - \int_{-1}^{0} \left((\delta y)'\right)^2 dx = - \int_{-1}^{0} \left((\delta y)'\right)^2 dx < 0.$$

Таким образом, доказано, что для всех $y(x)$

$$\Delta J[y^*(x)] = J[y(x)] - J[y^*(x)] < 0,$$

Это означает, что по определению имеет место максимум при $\boldsymbol{y^*(x) = -x^3}$**.**

Вычислим слабую вариацию функционала $J[y] = \int_{-1}^{0} (12xy - y'^2) dx$

$$\delta J = \frac{\partial J[y^*(x) + t\delta y(x)]}{\partial t}\bigg|_{t=0} = \frac{d}{dt} \int_{-1}^{0} \left(12x(-\boldsymbol{x^3} + t\delta y) - (-\boldsymbol{3x^2} + t(\delta y)')^2\right) dx \bigg|_{t=0} =$$

$$= \int_{-1}^{0} (12x\delta y - 2(-\boldsymbol{3x^2} + t(\delta y)')(\delta y)') dx \bigg|_{t=0} =$$

$$= \int_{-1}^{0} \left(12x\delta y + 6\boldsymbol{x^2}(\delta y)' - 2t(\delta y)'^2\right) dx \bigg|_{t=0} = \int_{-1}^{0} (12x\delta y + 6\boldsymbol{x^2}(\delta y)') dx =$$

$$= \int_{-1}^{0} 12x\delta y dx + \boldsymbol{6x^2}\delta y|_{\boldsymbol{-1}}^{\boldsymbol{0}} - \int_{-1}^{0} 12x\delta y dx = 0.$$

Таким образом, непосредственно проверено, что в точке максимума слабая вариация функционала равна нулю.