Вариационное исчисление и оптимальное управление

В.Ю. Протасов *

МГУ, мех-мат.

1 Лекция

Конечномерные гладкие экстремальные задачи

Условия существования точек минимума. Компактность и коэрцитивность. Теорема Ферма. Теорема Лагранжа для задач с ограничениями типа равенств. Примеры.

Рассмотрим задачу
$$f \;\to\; \min\,, \qquad x \in G\,, \tag{1.1}$$
где G – некоторое подмножество пространства $\mathbb{R}^d$. Начнём со случая, когда G – компакт, а функция f непрерывна на G.

Теорема 1.1 *Непрерывная функция на компактном множестве достигает своего наименьшего и наибольшего значения.*

Эта теорема принадлежит Карлу Вейерштрассу (1815-1897), она доказывается в курсе математического анализа. При некоторых дополнительных условиях она может быть распространена также на незамкнутые или неограниченные подмножества. Мы сформулируем ее для открытых множеств. Далее через ∂A будет обозначаться граница множества A, через $\operatorname{dist}(x, A)$ – расстояние от точки x до множества A.

Определение 1.2 *Функция, заданная на открытом множестве $G \subset \mathbb{R}^d$ называется коэрцитивной, если она непрерывна и $f(x) \to +\infty$ при $\operatorname{dist}(x\,,\,\partial G) \;\to\; 0$ и при $x \to \infty$.*

Лемма 1.3 *Коэрцитивная функция достигает на G своего наименьшего значения.*

*Московский Государственный Университет, Механико-математический факультет, Воробьевы Горы, Москва, 119992, e-mail: v-protassov@yandex.ru

`Доказательство.` Пусть $z \in G$ – произвольная точка. Для данного $\varepsilon > 0$ положим $G_\varepsilon = \{ x \in G \mid \operatorname{dist}(x, \partial G) \ge \varepsilon, |x| \le 1/\varepsilon \}$. Так как f коэрцитивна, то $\inf_{x \in G \setminus G_\varepsilon} f(x) \to +\infty$ при $\varepsilon \to 0$. Поэтому при достаточно малых $\varepsilon > 0$ имеем $f(x) > f(z)$ для всех $x \in G \setminus G_\varepsilon$. Следовательно, $\inf_{x \in G} f(x) = \inf_{x \in G_\varepsilon} f(x)$, а последний инфимум достигается, поскольку G_ε – компакт.

$\square$

Точка $x \in G$ называется точкой *локального минимума* в задаче (1.1), если при некотором $\varepsilon > 0$ для любой точки $y \in G$, $|y - x| < \varepsilon$ имеем $f(y) \ge f(x)$. Множество точек локального минимума данной задачи обозначим locmin.

Напомним, что производной функции $f : G \to \mathbb{R}$ в точке $x \in \operatorname{int} G$ называется элемент $f'(x) = a \in \mathbb{R}^d$ такой, что $f(x + h) = f(x) + (a, h) + o(h)$, $h \to 0$. Как известно из курса анализа $f'(x) = \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_d}(x) \right)$. Более того, если все частные производные существуют и непрерывны в точке x, то f дифференцируема в точке x.

Теорема 1.4 *(Теорема Ферма) Если точка $x \in \operatorname{int} G$ доставляет локальный минимум в задаче (1.1), и функция f дифференцируема в точке x, то $f'(x) = 0$.*

`Доказательство.` Если $f'(x) \ne 0$, то найдется $h \in \mathbb{R}^d$ такое, что $(f'(x), h) < 0$. Тогда $f\big(x + th\big) = f(x) + t\big(f'(x), h\big) + o(t)$, и при малых положительных t получаем $f\big(x + th\big) < f(x)$, что противоречит тому, что $x \in \mathrm{locmin}$.

$\square$

С помощью теоремы 1.4 можно находить точки, "подозрительные" на локальные минимум в задаче (1.1). Однако, это не даёт достаточных условий. Кроме того, могут быть ещё локальные минимумы на границе области определения G, а на них не распространяется теорема 1.4.

Функция f называется выпуклой (строго выпуклой), если $f((1-t)x + ty) \le (1-t)f(x) + tf(y)$ (соответственно $f((1-t)x + ty) < (1-t)f(x) + tf(y)$) для любых $x, y \in G$, $x \ne y$, $t \in (0, 1)$.

Примеры выпуклых функций: $f(x) = (a, x) + b$, $x \in \mathbb{R}^d$ (аффинная функция), $f(x) = (Ax, x)$, $x \in \mathbb{R}^d$, где A – положительно определенная квадратичная форма. Длина вектора $f(x) = \sqrt{x_1^2 + \ldots + x_d^2}$, а также любая другая норма $f(x) = \|x\|$ – выпуклые. Функции одной переменной: $x^{2n}, n \in \mathbb{N}$; e^x, $|x|, -\ln x$, $x > 0$ – выпуклые. Линейная комбинация выпуклых функций с неотрицательными коэффициентами – выпуклая функция. Поточечный супремум произвольного семейства выпуклых функций – выпуклая функция. Композиция выпуклых функций $f(g(x))$ не обязательно выпукла, но если f выпукла, а g – аффинная функция, то такая композиция выпукла.

Задача (1.1) называется выпуклой, если множество G и функция f выпуклы.

Лемма 1.5 *Для выпуклой задачи каждый локальный минимум является её абсолютным минимумом.*

`Доказательство.` Пусть $x \in G$ – точка локального минимума функции G. Если он не является её абсолютным минимумом, то существует точка $y \in G$, для которой $f(y) < f(x)$. Тогда для любого $t \in (0, 1)$ имеем $f((1-t)x + ty) \le (1-t)f(x) + tf(y) = f(x) - t(f(x) - f(y)) < f(x)$. Поскольку $(1-t)x + ty \to x$ при $t \to 0$, точка x не даёт и локального минимумума функции f.

$\square$

Лемма 1.6 *Если задача выпукла, функция f дифференцируема в точке $x \in \operatorname{int} G$ и $f'(x) = 0$, то эта точка даёт абсолютный минимум.*

`Доказательство.` Если существует точка $y \in G$, для которой $f(y) < f(x)$, то для любого $t \in (0,1)$ имеем $\frac{f((1-t)x+ty)\,-\,f(x)}{t} \le f(y) - f(x) < 0$. С другой стороны, $f((1-t)x+ty) = f(x+t(y-x)) = f(x) + t\big(f'(x)\,,\,y-x\big) + o(t) = f(x) + o(t)$ при $t \to 0$. Следовательно, $\frac{f((1-t)x+ty)\,-\,f(x)}{t} \to 0$ при $t \to 0$. Получили противоречие.

□

Лемма 1.7 *Если в выпуклой задаче функция f является строго выпуклой, то она имеет не более одной точки минимума.*

`Доказательство.` Если x и y – две точки минимума, то $f(x) = f(y)$. Из строгой выпуклости получаем $f\big(\frac{x+y}{2}\big) < \frac{1}{2}\,f(x) + \frac{1}{2}\,f(y) < f(x)$, что противоречит предположению.

□

Пример 1.8 Рассмотрим следующую геометрическую задачу. *Для данных точек $a_1,\dots,a_n \in \mathbb{R}^d$ найти точку, для которой сумма расстояний $f(x) = \sum_{k=1}^n |x-a_k|$ наименьшая.*

Решение.Функция f непрерывна и коэрцитивна на $\mathbb{R}^d$, следовательно она достигает своего минимума. Кроме того, f выпукла. Функция $g(x) = |x|$ дифференцируема во всех точках кроме 0, её производная равна $\frac{x}{|x|}$ (докажите это!) Поэтому функция $f(x)$ дифференцируема во всех точках кроме $a_1,\dots,a_n$, и её производная равна сумме n единичных векторов, сонаправленных векторам $x-a_k$, $k = 1,\dots,n$. Поэтому минимум достигается либо в точке, где сумма таких векторов равна x, либо в одной из точек a_k, $k = 1,\dots,n$. Если не все точки лежат на одной прямой, то функция f строго выпукла. Следовательно, в этом случае точка минимума единственна.

Приступим к формулировке принципа Лагранжа для задач с ограничениями типа равенств. Он был впервые сформулирован Жозефом Луи Лагранжем (1736-1813) в 1759 г. Пусть

$f_j\,:\,\mathbb{R}^d \to \mathbb{R}$ – некоторые функции, $j = 0,\dots,m$. Рассмотрим задачу

$$\left\{\begin{array}{lcl} f_0(x) & \to & \min \\ f_j(x) & = & 0\,, \qquad j = 1,\dots,m\,. \end{array}\right. \tag{1.2}$$

Функция f_0 – *целевая функция*, уравнения $f_j(x) = 0$, $j = 1,\dots,m$ – *условия (уравнения) связей.* Будем называть точку x допустимой, если в ней выполнены все условия связей. Задача (1.2), таким образом, состоит в том, чтобы отыскать наименьшее значение целевой функции f_0 на множестве всех допустимых точек. Будем говорить, что точка x доставляет локальный минимум в задаче (1.2) (обозначение $x \in \mathrm{locmin}$), если существует $\varepsilon > 0$, при котором $f_0(y) \ge f_0(x)$ для всех допустимых точек y таких, что $|y-x| \le \varepsilon$. Функция $\mathcal{L}(x,\lambda) = \sum_{j=0}^m \lambda_j\, f_j(x)$ переменных $\lambda = (\lambda_0,\dots,\lambda_m)$ и $x = (x_1,\dots,x_d)$ называется *функцией Лагранжа* или *лагранжианом* данной задачи. Мы будем всегда предполагать, что множители $\lambda_0,\dots,\lambda_m$ **не р**авны **о**дновременно **н**улю (сокращённо НЕРОН).

Теорема 1.9 *(Теорема Лагранжа). Если $x \in \mathrm{locmin}$ в задаче (1.2) и все функции f_j , $j = 0,\dots,m$ дифференцируемы в точке x, то существуют такие НЕРОН множители $\lambda_0,\dots,\lambda_m$, что $\mathcal{L}_x\,(x,\lambda) = \sum_{j=0}^m \lambda_j\, f'(x) = 0$.*

Условие $\sum_{j=0}^m \lambda_j\, f'(x) \;=\; 0$ означает, что частные производные лагранжиана $\mathcal{L}\,(x,\lambda) = \sum_{j=0}^m \lambda_j\, f(x)$ по каждой из переменных $x_1,\ldots,x_d$ равны нулю. Поэтому данное условие фактически содержит не одно уравнение, а систему из d уравнений. Таким образом, для того, чтобы найти все точки $x \in \mathrm{locmin}$ нужно решить систему из $d+m$ уравнений ($\mathcal{L}_x(x,\lambda)=0$ – d уравнений, $f_j(x)=0,\ j=1,\ldots,m$ – m уравнений) с $d+m+1$ неизвестным (точка x – это d неизвестных, и множители $\lambda_0,\ldots,\lambda_m$). Неизвестных на одно больше, чем уравнений. Однако, все уравнения однородны по λ. За счёт этого можно понизить число неизвестных на 1 (разделив уравнения на ненулевой множитель λ_j). На практике обычно рассматривают два случая: $\lambda_0 = 0$ и $\lambda_0 = 1$ (или любому другому положительному числу). Доказательство теоремы Лагранжа, как правило, осуществляется с привлечением мощных результатов анализа, таких как теорема о неявной функции или теорема о неподвижной точке. Перед доказательством сформулируем несколько вспомогательных результатов.

Через $B(x,r)$ будем обозначать замкнутый евклидов шар радиуса r с центром в точке x.

Теорема 1.10 *(Теорема Брауэра о неподвижной точке). Любое непрерывное отображение $F : B(x,r) \to B(x,r)$, где $B(x,r)$ шар пространства $\mathbb{R}^d$, имеет неподвижную точку.*

Теорема Брауэра имеет несколько доказательств, основанных на различных идеях, см., например, [6].

Лемма 1.11 *(Лемма об ε-сдвиге). Если непрерывное отображение $\varphi : B(0,r) \to \mathbb{R}^d$ таково, что $|\varphi(y) - y| \le \varepsilon$ для всех $y \in B(0,r)$, то у любой точки $a \in B(0,r-\varepsilon)$ есть прообраз $\tilde{y} \in B(0,r) : \varphi(\tilde{y}) = a$.*

Таким образом, если отображение сдвигает каждую точку шара радиуса r на вектор длины не более ε, то его образ покрывает концентрический шар радиуса $r-\varepsilon$.

`Доказательство.` Отображение $F(y) = a + y - \varphi(y)$ переводит шар $B(0,r)$ в себя. В самом деле, $|F(y)| \le |a| + |y - \varphi(y)| \le r - \varepsilon + \varepsilon = r$ для любого $y \in B(0,r)$. Следовательно, F имеет неподвижную точку $\tilde{y}$. Для этой точки $\varphi(\tilde{y}) = a$.

$\square$

`Доказательство теоремы 1.9.` Пусть $x \in \mathrm{locmin}$. Не ограничивая общности, считаем, что $f_0(x)=0$. Рассмотрим множество векторов $\bigl((f_0'(x),h),\ldots,(f_m'(x),h)\bigr)$, $h \in \mathbb{R}^d$. Оно является линейным подпространством пространства $\mathbb{R}^{m+1}$. Если оно не совпадает со всем $\mathbb{R}^{m+1}$, то найдется ненулевой вектор $\lambda = \bigl(\lambda_0,\ldots,\lambda_m\bigr) \in \mathbb{R}^{m+1}$, ортогональный этому подпространству. Тогда $\sum_{j=0}^m \lambda_j\,\bigl(f_j'(x)\,,\,h\bigr) = 0$ для любого $h \in \mathbb{R}^d$. Следовательно $\sum_{k=0}^m \lambda_j\, f_j'(x) = 0$, что и требовалось. Покажем, теперь, что случай, когда данное пространство совпадает с $\mathbb{R}^{m+1}$ невозможен, т.е., точка x не будет доставлять локальный минимум в задаче (1.2). Найдутся такие h_j, $j=0,\ldots,m$, что $\bigl(f_j'(x),h_j\bigr) = 1$ и $\bigl(f_k'(x),h_j\bigr) = 0$ при $k \ne j$. Для произвольного $r>0$ рассмотрим отображение $\varphi : B(0,r) \mapsto \mathbb{R}^{m+1}$, заданное формулой

$$y \;=\; (y_0,\ldots,y_m) \quad\mapsto\quad \left(f_0\Bigl(x+\sum_{j=0}^m y_j\,h_j\Bigr)\,,\ \ldots\ ,\ f_m\Bigl(x+\sum_{j=0}^m y_j\,h_j\Bigr)\right)$$

Для каждого k имеем

$$f_k\Bigl(x+\sum_{j=0}^m y_j\,h_j\Bigr) \;-\; y_k \;=\; 0 \;+\; \Bigl(f_k'(x)\,,\,y_k\,h_k\Bigr) \;+\; o\bigl(y\bigr) \;-\; y_k \;=\; o\bigl(y\bigr)\,, \quad y \to 0\,.$$

Таким образом, $\varphi(y) - y \ = o(y)\,,\ y \to 0$. Тогда при достаточно малых r имеем $|\varphi(y) - y| \le \frac{r}{2}$ для любого $y \in B(0,r)$. По лемме об ε-сдвиге, применённой к шару $B(0,r)$ и к $\varepsilon = \frac{r}{2}$ существует точка $\tilde{y} \in B(0,r)$ такая, что $\varphi(\tilde{y}) \ = \ \bigl(-\frac{r}{2}\,, 0\,, \ldots\,, 0\bigr)$. Для точки $\tilde{x} \ = \ x + \sum_{j=0}^{m} \tilde{y}_j\, h_j$ получаем $f_j(\tilde{x}) = 0\,,\ j = 1, \ldots, m$ (значит точка $\tilde{x}$ допустимая), и $f_0(\tilde{x}) \ \le \ -\frac{r}{2}$, следовательно $x \notin \mathrm{locmin}$.
□

Правило множителей Лагранжа даёт чрезвычайно удобный метод для решения экстремальных задач, доказательства теорем существования и т.д. Объясняется это двумя свойствами:

1) правило множителей позволяет не выражать одни переменные через другие (что часто бывает затруднительно), а напрямую дифференцировать уравнения связей;

2) правило множителей сохраняет симметрию задачи относительно входящих в него переменных. Это свойство полезно, например, при доказательстве неравенств.

Начнём с простого численного примера.

Пример 1.12 *Найти минимальное значение функции*
$f(x,y) = 5x^2 + 2xy + 3y^2$ *при условии* $g(x,y) = 7x^2 + 2xy + 4y^2 - 1 = 0$.

`Решение.` Используя уравнение $g(x,y) = 0$ мы можем выразить y через x (или x через y). Однако, прямая подстановка этого выражения в уравнение $f'(x, y(x)) = 0$ приводит к уравнению шестой (!) степени. На самом деле – даже к двум уравнениям, из за знака $\pm$ перед дискриминантом в выражении для $y(x)$. Правило множителей даёт такое решение: сохраняем обе переменные x, y и составляем лагранжиан: $\mathcal{L} \ = \ f(x,y) + \lambda\, g(x,y)$. Дифференцируя по x и по y, получаем систему

$$\begin{cases} \lambda_0(10x + 2y) + \lambda_1(14x + 2y) = 0 \\ \lambda_0(6y + 2x) + \lambda(8y + 2x) = 0. \end{cases}$$

Если $\lambda_0 \ = \ 0$, то $\lambda_1 \ne 0$ (НЕРОН !), откуда $x \ = \ y \ = \ 0$, что невозможно по условию. Если $\lambda_0 = 1$, то, выражая λ_1 из обеих уравнений и приравнивая, получаем $\frac{10x+2y}{14x+2y} = \frac{6y+2x}{8y+2x}$. Решая это однородное уравнение, имеем $\frac{y}{x} = -1$ или 2. Теперь подставляя в уравнение $g(x,y) = 0$, получаем несколько "подозрительных" на экстремум точек (x,y). Вычисляя значение функции f в каждой из них, убеждаемся, что минимум достигается в точках $(x,y) = (-\frac{1}{3}, \frac{1}{3})$ и $(x,y) = (\frac{1}{3}, -\frac{1}{3})$. Далее остается сослаться на существование точки минимума (эллипс $g(x,y) = 0$ – компактное множество). Получаем, что минимум равен $\frac{2}{3}$.

Пример 1.13 С помощью правила множителей докажем теорему линейной алгебры: *Симметрическая матрица всегда имеет действительный собственный вектор.*
`Решение.` Для симметрической матрицы A рассмотрим экстремальную задачу

$$\begin{cases} -(Ax, x) \quad \to \quad \min \\ (x,x) \, = \, 1\,. \end{cases}$$

Она имеет решение, поскольку единичная сфера $\bigl\{ x \in \mathbb{R}^d \ \big|\ |x| = 1 \bigr\}$ компактна. Имеем $\mathcal{L}_x \ = \ -\,2\,\lambda_0\, Ax \ + \ 2\,\lambda_1\, x \ = \ 0$. Поскольку $\lambda_0 \ne 0$ (иначе $x = 0$, что противоречит условию $|x| = 1$), получаем $Ax \ = \ \lambda_1\, x$.

Пример 1.14 . *Луч света пересекает прямую границу двух сред, входя под углом α_1, и выходя под углом α_2 (оба угла отсчитываются от нормали к границе). Тогда $\frac{\sin\alpha_1}{v_1} \ = \ \frac{\sin\alpha_2}{v_2}$, где v_1 и v_2 – скорости света в этих средах.*

Решение. Считаем известным, что луч света при своем движении от одной точки к другой выбирает самый быстрый путь. Если взять на луче точки x_1 и x_2 по разные стороны от границы, а саму границу обозначить через $l = \{x \in \mathbb{R}^2\,,\ (n,x) = c\}$, где n – вектор нормали к прямой, получим задачу

$$\begin{cases} f(x) \ = \ |x - x_1|/v_1 \ + \ |x - x_2|/v_2 \ \to \ \min \\ g(x) \ = \ (n,x) \ - \ c \ = \ 0\,. \end{cases} \qquad (1.3)$$

Минимум достигается, в некоторой точке x т.к., функция f коэрцитивна. Производная функции $|x{-}x_1|$ в точке x является вектором u_1 единичной длины, сонаправленным с вектором $x{-}x_1$ (пример 1.8). Аналогично определяем вектор u_2 – производную длины $x - x_2$. Итак, $f'(x) = \frac{1}{v_1}u_1 + \frac{1}{v_2}\,u_2$. Далее, $g'(x) = n$. Правило множителей даёт $\lambda_0\left(\frac{1}{v_1}\,u_1 + \frac{1}{v_2}\,u_2\right) + \lambda_1\,n = 0$. Таким образом, вектор $\frac{1}{v_1}\,u_1 + \frac{1}{v_2}\,u_2$ коллинеарен вектору n, т.е., перпендикулярен прямой l. Это значит, что проекция данного вектора на l равна нулю, что равносильно условию $\frac{\sin\alpha_1}{v_1} = \frac{\sin\alpha_2}{v_2}$.

2 Лекция

Дифференцирование в нормированных пространствах.
Элементы выпуклого анализа.

Вариация по Лагранжу. Производная по Гато и по Фреше. Теорема отделимости в нормированных пространствах. Лемма Фаркаша. Теорема Лагранжа для задач с ограничениями типа равенств и неравенств.

Определение 2.1 *Пусть X, Y – нормированные пространства, $G \subset X$, $x \in \operatorname{int} G$, $h \in X$. Вариацией по Лагранжу отображения $F:\ G \to Y$ в точке x по направлению h называется предел $\delta_F(x,h) = \lim_{t\to 0}\frac{F(x+th)-F(x)}{t}$.*

Отображение называется *дифференцируемым по Лагранжу* в точке x, если вариация по Лагранжу в данной точке существует по любому направлению $h \in X$. Следующая теорема является обобщением теоремы Ферма на нормированные пространства.

Теорема 2.2 *(Теорема Ферма для нормированных пространств). Если $x \in \operatorname{int} G \subset X$ – точка локального минимума функции $f: G \to \mathbb{R}$, и функция f дифференцируема по Лагранжу в этой точке, то $\delta_f(x,h) = 0$ для любого $h \in X$.*

`Доказательство.` Для любого h рассмотрим функцию $F_h(t) = f(x{+}th)$. Так как $0 \in \operatorname{locmin} F_h$, то $F'_h(0) = 0$. Остаётся заметить, что $F'_h(0) = \delta_F(x,h)$. □

Теперь мы можем обобщить два главных свойства выпуклых функций на произвольные нормированные пространства. Лемма 1.2 не меняется вовсе, ни формулировка ни доказательство: *для выпуклой задачи каждый локальный минимум является её абсолютным минимумом.* Лемма 1.3 теперь переформулируется так:

Предложение 2.3 *Если задача выпукла, функция f дифференцируема по Лагранжу в точке $x \in \operatorname{int} G$ и $\delta_f(x,h) = 0$ для всех $h \in X$, то эта точка даёт абсолютный минимум.*

Доказательство. Предположим, что существует точка $y \in G$, для которой $f(y) < f(x)$. Проведем прямую через точки x и y: $\left\{ x + th \;\middle|\; t \in \mathbb{R} \right\}$, где $h = y - x$. Ограничение функции f на эту прямую является выпуклой функцией одной переменной, её производная в точке x равна $\delta_f(x,h) = 0$, следовательно (лемма 1.3), абсолютный минимум этой функции на прямой достигается в точке x, а значит $f(y) \geq f(x)$.

□

Определение 2.4 *Отображение $F: G \to Y$, где $G \subset X$, называется дифференцируемым по Гато в точке $x \in \operatorname{int} G$, если оно дифференцируемо по Лагранжу в этой точке и существует линейный непрерывный оператор $A : X \to Y$ такой, что $\delta_F(x,h) = Ah$. Оператор A называется производной по Гато отображения F в точке x.*

Следствие 2.5 *В условиях теоремы 2.2 функция f дифференцируема по Гато в точке x, и производная по Гато равна нулю.*

Определение 2.6 *Отображение $F: G \to Y$, где $G \subset X$, называется дифференцируемым по Фреше в точке $x \in \operatorname{int} G$, если существует линейный непрерывный оператор $A : X \to Y$ такой, что $F(x+h) = F(x) + Ah + o(h)$, $h \to 0\,, h \in X$. Оператор A называется производной по Фреше отображения F в точке x и обозначается $A = F'(x)$.*

Таким образом, $F(x+h) = F(x) + F'(x)[h] + o(h)$, $h \to 0$. Из этого следует, что функция, дифференцируемая по Фреше в точке x, непрерывна в x. Для дифференцируемости по Гато это может не выполняться (пример 2.8). Заметим, что производная по Фреше, если существует, однозначно определена. В противном случае, если найдутся два оператора A_1, A_2, для которых выполнено соотношение $F(x+h) = F(x) + A_i[h] + o(h)$, $h \to 0\,,\ i = 1,2$, то, вычитая, получим, $(A_1 - A_2)[h] = o(h)\,, h \to 0$. Последнее означает, что $A_1 = A_2$. Иначе было бы $(A_1 - A_2)[\tilde{h}] \neq 0$ для некоторого $\tilde{h} \in X$, и тогда $t\,(A_1 - A_2)[\tilde{h}] = (A_1 - A_2)[t\tilde{h}] = o(t)\,, t \to 0$, что невозможно.

Как связана производная по Фреше с вариацией по Лагранжу ? Имеем

$$\delta_F(x,h) \quad = \quad \lim_{t \to 0} \frac{F(x+th) - F(x)}{t} = \lim_{t \to 0} \frac{F'(x)[th] + o(th)}{t} \quad = \quad F'(x)[h].$$

Следовательно, отображение, дифференцируемое по Фреше, дифференцируемо и по Лагранжу. Более того, так как ого вариация по Лагранжу равна $\delta_F(x,h) = Ah$, где $A = F'(x)$, то приходим к выводу, что отображение, дифференцируемое по Фреше, дифференцируемо и по Гато (с той же производной). Итак,

Фреше $\Rightarrow$ **Гато** $\Rightarrow$ **Лагранж**

Обратные импликации не выполняются, как показывают следующие примеры:

Пример 2.7 Функция $F : \mathbb{R}^2 \to \mathbb{R}$, заданная формулой $F(x_1,x_2) = \left(x_1^2 x_2\right)^{1/3}$, дифференцируема по Лагранжу в точке $x = (0,0)$: $\delta_F(x,h) = \left(h_1^2 h_2\right)^{1/3}$. Если $e_1 = (1,0), e_2 = (0,1)$, то $\delta_F(x,e_1) = \delta_F(x,e_2) = 0$. Однако, $\delta_F(x,e_1+e_2) = 1$. Если отображение дифференцируемо по Гато, то $\delta_F(x,h)$ линейно зависит от h. Значит, должно выполняться $\delta_F(x,e_1+e_2) = \delta_F(x,e_1) + \delta_F(x,e_2)$, что неверно. Поэтому F не дифференцируемо по Гато.

Пример 2.8 Функция $F:\mathbb{R}^2\to\mathbb{R}$, заданная формулой

$$F(x) \quad = \quad \begin{cases} 1\,, & x_2=x_1^2\,,\ x_1\neq 0 \\ 0\,, & \text{иначе} \end{cases}$$

дифференцируема по Лагранжу в точке $x=(0,0)$: $\delta_F(x,h)=0$ для любого $h\in\mathbb{R}^2$. Значит, она дифференцируема и по Гато, её производная по Гато равна 0. Но по Фреше она не дифференцируема, так как она разрывна в точке x.

Класс функций, дифференцируемых по Фреше в точке x обозначим $\mathcal{D}(x)$, а дифференцируемых в каждой точке области G – через $\mathcal{D}(G)$. Далее, если не оговорено обратное, под дифференцируемыми функциями мы будем понимать именно дифференцируемые по Фреше.

Наша следующая задача – сформулировать и доказать принцип Лагранжа для задач с ограничениями типа равенств и неравенств в нормированных пространствах. Это мы сделаем на следующей лекции, а пока проделаем предварительную работу и установим несколько вспомогательных результатов выпуклого анализа.

Определение 2.9 *Подмножества A и B нормированного пространства X называются отделимыми, если существует функционал $x^*\in X$ такой, что $\sup_{a\in A}(x^*,a)\ \le\ \inf_{b\in B}(x^*,b)$. Они называются строго отделимыми, если $\sup_{a\in A}(x^*,a)\ <\ \inf_{b\in B}(x^*,b)$.*

Теорема 2.10 *(Первая теорема отделимости). Если A и B – непустые выпуклые подмножества нормированного пространства X, причём $\operatorname{int}B\neq\emptyset$, $(\operatorname{int}B)\cap A=\emptyset$, то A и B отделимы.*

Эта теорема следует из теоремы Хана-Банаха, ее доказательства мы не приводим.

Теорема 2.11 *(Вторая теорема отделимости). Если A и B – непустые выпуклые замкнутые подмножества нормированного пространства X, причём A – компакт и $A\cap B=\emptyset$, то A и B строго отделимы.*

`Доказательство теоремы 2.11.` Обозначим через P открытый единичный шар в пространстве X. Покажем, что для достаточно малых r множества $A+rP$ и B не пересекаются. В противном случае найдутся последовательности точек $\{a_k\}\subset A$, $\{b_k\}\subset B$, для которых $\|a_k-b_k\|\to 0$ при $k\to\infty$. Поскольку A компактно, переходя к подпоследовательности, считаем, что последовательность a_k сходится к некоторой точке $a\in A$. Но тогда и $b_k\to a$, $k\to\infty$, что, в силу замкнутости B влечёт $a\in B$. Таким образом, $A\cap B\neq\emptyset$, что противоречит условию. Итак, открытое выпуклое множество $A+rP$ не пресекает B, поэтому, в силу теоремы 2.10 найдётся функционал $x^*\in X^*$, разделяющий эти множества:

$$\sup_{a\in A,x\in P}(x^*,a+rx)\ \le\ \inf_{b\in B}(x^*,b).$$

Заметим, что $\sup_{a\in A,x\in P}(x^*,a+rx)=\sup_{a\in A}(x^*,a)+\sup_{x\in P}(x^*,rx)=\sup_{a\in A}(x^*,a)+r\|x^*\|$. Итак, $\sup_{a\in A}(x^*,a)\ \le\ -r\|x^*\|+\inf_{b\in B}(x^*,b)$. Таким образом, x^* строго разделяет A и B.

Следствие 2.12 *Любая точка нормированного пространства X, не принадлежащая выпуклому замкнутому множеству $A\subset X$, строго отделима от A.*

Упражнение 2.13 Приведите пример выпуклого подмножества A гильбертова пространства H, которое имеет пустую внутренность, но не лежит ни в каком замкнутом подпространстве пространства H.

Упражнение 2.14 * Приведите пример двух непересекающихся выпуклых подмножеств гильбертова пространства, которые нельзя отделить.

Лемма 2.15 (Лемма Фаркаша). *Пусть $a_1, \dots, a_k \in X^*$ – линейные функционалы на нормированном пространстве X. Тогда либо существует элемент $h \in X$, для которого $(a_j, h) < 0$, $j = 1, \dots, k$, либо найдутся неотрицательные НЕРОН множители $\lambda_1, \dots, \lambda_k$ такие, что $\sum_{j=1}^{k} \lambda_j a_j = 0$.*

`Доказательство.` Предположим сначала, что $X = \mathbb{R}^d$ – конечномерное пространство. Тогда, как известно, $(X^*)^* = X$. Если $A = \mathrm{Conv}\{a_1, \dots, a_k\}$, выпуклая оболочка точек $a_1, \dots a_k$, содержит ноль, то $0 = \sum_{j=1}^{k} \lambda_j a_j$ для некоторых $\lambda_j \ge 0, \sum_j \lambda_j = 1$. Если же $0 \notin A$, то согласно следствию 2.12 найдётся $h \in (X^*)^* = X$, для которого $\sup_{a \in A}(a, h) < (0, h) = 0$, а значит и $(a_j, h) < 0\,,\ j = 1, \dots, k$.

В общем случае, для произвольного X, мы рассмотрим конечномерное линейное пространство $\tilde{X}$, состоящее из векторов $\tilde{x} = \big((a_1, x), \dots, (a_k, x)\big) \in \mathbb{R}^k$ для всевозможных $x \in X$. Функционалу a_i соответствует функционал $\tilde{a}_i$ на пространстве $\tilde{X}$, ставящий в соответствие вектору его i-тую координату. Воспользовавшись доказанным конечномерным случаем, имеем: либо $\sum_j \lambda_j \tilde{a}_j = 0$ для некоторых $\lambda_j \ge 0$, а значит и $\sum_j \lambda_j a_j = 0$, либо $(\tilde{a}_j, \tilde{h}) < 0\,,\ j = 1, \dots, k$ при некотором $\tilde{h} \in \tilde{X}$, а значит $(a_j, h) < 0\,,\ j = 1, \dots, k$ для соответствующего $h \in X$.

□

3 Лекция

Принцип Лагранжа для задач с ограничениями типа равенств и неравенств. Пример: доказательство теоремы Минковского о многогранниках

Мы приступаем к формулировке принципа Лагранжа для задач с ограничениями типа равенств и неравенств. Пусть X – нормированное пространство, $f_i : X \to \mathbb{R}$ – некоторые функции, $i = 0, \dots, n + m$, $n \ge 0, m \ge 0$. Рассмотрим задачу

$$\begin{cases} f_0(x) & \to & \min \\ f_i(x) & \le & 0\,, \qquad i = 1, \dots, n\,; \\ f_j(x) & = & 0\,, \qquad j = n+1, \dots, n+m\,. \end{cases} \tag{3.4}$$

Теорема 3.1 *Если $\hat{x} \in \mathrm{locmin}$ в задаче (3.4) и все функции $f_k\,,\ k = 0, \dots, n+m$ дифференцируемы по Гато в точке $\hat{x}$, то существуют такие НЕРОН множители $\lambda_0, \dots, \lambda_{n+m}$, что*

1) $\mathcal{L}_x(\hat{x}, \lambda) = \sum_{k=0}^{n+m} \lambda_k f'(\hat{x}) = 0$ (условие стационарности. Все производные – по Гато.);

2) $\lambda_i \ge 0\,,\quad i = 0, 1, \dots, n$ (условие неотрицательности);

3) $\lambda_i f_i(\hat{x}) = 0\,,\quad i = 1, \dots, n$ (условие дополняющей нежёсткости).

Таким образом, для задач с ограничениями типа неравенств верен то же принцип, что и для задач с равенствами: для точки локального минимума найдётся набор множителей, при котором производная лагранжиана в этой точке равна нулю. Возникают два новых условия: для целевой функции и для всех ограничений типа неравенств множители неотрицательны (условие 2), и каждое ограничение типа неравенств либо обращается в равенство, либо соответствующий множитель обращается в ноль (условие 3). Назовем ограничение $f_i(x) \le 0$ *активным*, если в данной точке $\hat{x}$ оно обращается в равенство, в противном случае (если неравенство строгое) назовём его *пассивным*. Условие дополняющей нежёсткости означает, что для всех пассивных неравенств множители Лагранжа обращаются в ноль. Рассмотрим для простоты случай $X = \mathbb{R}^d$. Если в точке $\hat{x} \in \mathrm{locmin}$ из n неравенств ровно k активных (например, первые k), то для поиска точки $\hat{x}$ мы имеем систему из $d+m+k$ уравнений ($\mathcal{L}_x(\hat{x},\lambda) = 0$ – это d уравнений, $f_j(x) = 0$, $j = 1,\ldots,k$ и $j = n+1,\ldots n+m$ – это $m+k$ уравнений) с $d+m+k+1$ неизвестным (точка x – это d неизвестных, а также все множители λ_j, кроме $\lambda_1 = \cdots = \lambda_k = 0$, это $m+k+1$ неизвестных). Таким образом, неизвестных на одно больше, чем уравнений. Однако, все уравнения однородны по λ. Засчёт этого можно понизить число неизвестных на 1 (разделив уравнения на ненулевой множитель λ_j). Итак, чтобы найти точки локального минимума в задаче (3.4) нужно рассмотреть 2^n случаев (в зависимости от того, какие из n неравенств $f_i(x) \le 0$ активные), в каждом случае возникает система, в которой число уравнений равно числу неизвестных. Решение задачи (3.4), как правило, состоит в разумно организованном переборе этих случаев.

Теорему 3.1 мы докажем в частном случае, когда $m = 0$, т.е., когда все ограничения – типа неравенств.

`Доказательство теоремы 3.1.` Пусть $\hat{x} \in \mathrm{locmin}$. После возможной перестановки индексов, можем считать, что неравенства $f_i(x) \le 0$ активные при $i \le k$ и пассивные при $i > k$. Если активных неравенств нет, то положим $k = 0$, а если нет пассивных, то $k = n$. Полагаем $\lambda_j = 0$, $j = k+1,\ldots,n$ (т.е., для всех пассивных неравенств), и рассмотрим функционалы $a_i = f_i'(\hat{x}) \in X^*$, $i = 0,\ldots,k$. По Лемме Фаркаша, либо существуют неотрицательные множители $\lambda_0,\ldots,\lambda_k$, для которых $\sum_{i=0}^k \lambda_i = 1$ (значит, выполнен НЕРОН), и $\sum_{i=0}^k \lambda_i f_i'(\hat{x}) = 0$, т.е., выполнены условия 1) - 3) теоремы, либо существует $h \in X$, для которого $\big(f_i'(\hat{x}), h\big) < 0$, $i = 0,\ldots,k$. В первом случае теорема доказана. Во втором случае, из равенства $\big(f_i'(\hat{x}), h\big) = \delta_{f_i}(\hat{x}, h) = \lim_{t\to 0} \frac{f_i(\hat{x}+th) - f_i(\hat{x})}{t}$, получаем, что $f_i(\hat{x}+th) < f_i(\hat{x})$ для всех достаточно малых $t > 0$, при каждом $i = 0,\ldots,k$. С другой стороны, это неравенство верно и при каждом $i = k+1,\ldots,n$, поскольку для пассивных неравенств $f_i(\hat{x}) < 0$. Таким образом, при всех малых $t > 0$ в точке $\hat{x}+th$ выполнены все ограничения $f_i \le 0$, и при этом $f_0(\hat{x}+th) < f_0(\hat{x})$, а значит $\hat{x} \notin \mathrm{locmin}$.

$\square$

Замечание 3.2 Теорема 3.1 впервые была установлена для случая $m = 0$ немецко-американским математиком Фрицем Джоном (1910-1994), представителем знаменитой Геттингенской математической школы.

Среди огромного числа приложений теоремы 3.1 мы рассмотри одно: доказательство теоремы Минковского о выпуклых многогранниках с данными площадями и направлениями граней.

Пример 3.3 Пусть многогранник в $\mathbb{R}^d$ имеет k $d-1$-мерных граней, обозначим через S_i площадь ($d-1$-мерный объем) i-той грани, а через $\mathbf{n}_i$ вектор единичной внешней нормали к этой грани.

Теорема 3.4 *(Г.Минковский, 1897). Для любого выпуклого многогранника* $\sum_{i=1}^{k} S_i \mathbf{n}_i = 0$. *Обратно, для любых положительных чисел* $a_1, \dots, a_k$ *таких, что* $\sum_{i=1}^{k} a_i \mathbf{n}_i = 0$ *существует единственный выпуклый многогранник с внешними нормалями* $\mathbf{n}_i$, *для которого* $S_i = a_i, i = 1, \dots, n$.

В трехмерном пространстве равенство $\sum_{i=1}^{k} S_i \mathbf{n}_i = 0$ имеет физическое объяснение, причем для любого многогранника, не обязательно выпуклого. Если заполнить многогранник идеальным газом под давлением p, то сила, действующая на i-тую грань, равна pS_i. Так как равнодействующая всех сил равна нулю (иначе эта сила будет двигать многогранник, образуя "вечный двигатель"), то $\sum_{i=1}^{k} pS_i \mathbf{n}_i = 0$.

Теорема 3.4 содержит три утверждения: необходимость условия $\sum_{i=1}^{k} S_i \mathbf{n}_i = 0$ для существования многогранника, его достаточность и единственность многогранника. Мы докажем все, кроме единственности.

`Доказательство необходимости и достаточности в Теореме 3.4.` *Необходимость.* Пусть G – данный многогранник. Для произвольной точки $x \in G$ обозначим через $h_i(x)$ расстояние от нее до i-той грани. Так как объем пирамиды с вершиной x и основанием, совпадающим с i-той гранью, равен $\frac{1}{d}\, h_i(x) S_i$, то $\sum_{i=1}^{k} h_i(x) S_i \;=\; d\,\mathrm{Vol}\,(G)$. Продифференцировав обе части данного равенства по x, получим $\sum_{i=1}^{k} \mathbf{n}_i S_i = 0$.

Достаточность. Для любого набора неотрицательных чисел $h = (h_1, \dots, h_k)$ рассмотрим многогранник $G(h) = \bigl\{x \in \mathbb{R}^d \;\bigm|\; (\mathbf{n}_i, x) \le h_i\,,\, i = 1, \dots, k\bigr\}$ и обозначим через $V(h)$ его объем. Заметим, что $0 \in G(h)$ для всех h. Рассмотрим задачу

$$\begin{cases} -V(h) \;\to \min\,, \\ \sum_{i=1}^{k} a_i h_i = 1\,, \\ -h_i \le 0\,,\; i = 1, \dots, k. \end{cases} \tag{3.5}$$

Иначе говоря, среди всех многогранников, ограниченных гиперплоскостями, перпендикулярными векторам $\mathbf{n}_i$ и проходящими на расстояниях h_i от них, так, что $\sum_{i=1}^{k} a_i h_i = 1$, мы ищем многогранник максимального объема.

Решение задачи (3.5) существует, поскольку множество допустимых векторов h компактно. Обозначим решение через $\hat{h}$. По теореме 3.1 имеем $\mathcal{L}_h(\lambda, \hat{h}) = 0$, где

$$\mathcal{L} \;=\; -\lambda_0 V(h) \;-\; \sum_{i=1}^{k} \lambda_i h_i \;+\; \lambda_{k+1}\Bigl(\sum_{i=1}^{k} a_i h_i - 1\Bigr)$$

– лагранжиан. При этом $\lambda_i \ge 0$ и $\lambda_i \hat{h}_i = 0$, $i = 1, \dots, k$. Дифференцируя по h_i, получаем

$$-\lambda_0 V_{h_i} \;-\; \lambda_i \;+\; \lambda_{k+1} a_i \;=\; 0, \quad i = 1, \dots, k\,.$$

Начнем с того, что производная $V_{h_i}(h)$ численно равна площади ($(d-1)$-мерному объему) пересечения многогранника $G(h)$ с плоскостью $L_i = \{x \in \mathbb{R}^d \mid (\mathbf{n}_i, x) = h_i\}$. Поэтому, если при некотором j плоскость L_j не пересекает $G(\hat{h})$ по $(d-1)$-мерной грани, т.е., является "лишней", то $V_{h_j} = 0$, откуда $\lambda_j = \lambda_{k+1} a_j$. То же верно и в случае $\lambda_0 = 0$. Если $\lambda_j = 0$, то $\lambda_{k+1} = 0$, а значит $\lambda_i = -\lambda_0 V_{h_i}$ для всех i. Если $\lambda_0 = 0$, то получаем противоречие с НЕРОН, а если $\lambda_0 = 1$, то $\lambda_i = -V_{h_i}$ для всех i. Если i-тая плоскость пересекает многогранник по $d-1$-мерной грани (а такая плоскость, естественно, найдется), то $V_{h_i} = S_i > 0$, а значит $\lambda_i < 0$, что противоречит условию неотрицательности. Итак, $\lambda_0 = 1$, и все k плоскостей образуют грани многогранника $G(\hat{h})$, т.е., $V_{h_i} = S_i > 0$ для всех i. Тогда

$$-S_i - \lambda_i + \lambda_{k+1} a_i = 0, \quad i = 1, \ldots, k.$$

Умножив данное равенство на вектор $\mathbf{n}_i$, сложив по всем i, и приняв во внимание, что $\sum_{i=1}^k S_i \mathbf{n}_i = \sum_{i=1}^k a_i \mathbf{n}_i = 0$, получим $\sum_{i=1}^k \lambda_i \mathbf{n}_i = 0$. Обозначим через I множество индексов i, для которых $\lambda_i > 0$. Тогда $\sum_{i \in I} \lambda_i \mathbf{n}_i = 0$, а в силу условия дополняющей нежёсткости, $\hat{h}_i = 0$ для всех $i \in I$. С другой стороны, многогранник $G(\hat{h})$ имеет хотя бы одну внутреннюю точку z, для которой $(\mathbf{n}_i, z) < \hat{h}_i$, $i = 1, \ldots, k$. Умножив на λ_i и сложив по всем $i \in I$, получим $0 < \sum_{i \in I} \lambda_i \hat{h}_i = 0$. Данное противоречие означает, что I пусто, т.е., $\lambda_i = 0$, а значит $S_i = \lambda_{k+1} a_i$ для всех $i = 1, \ldots, k$. Таким образом, у многогранника $G(\hat{h})$ площади граней пропорциональны числам a_i. Сделав гомотетию с нужным коэффициентом, получим многогранник, у которого площади граней равны этим числам.

$\square$

Упражнение 3.5 * Докажите, что функция $-\ln V(h)$ строго выпукла по h.

Упражнение 3.6 Докажите, с помощью результата предыдущего упражнения, единственность в теореме Минковского.

4 Лекция

Негладкие выпуклые задачи.

Субдифференциал и его свойства. Субдифференциальное исчисление. Примеры.

Ещё одним преимуществом выпуклых экстремальных задач является то обстоятельство, что методы их решения распространяются на все выпуклые функции, даже не дифференцируемые. Для этого нужно обобщить понятие производной выпуклой функции.

Определение 4.1 *Пусть функция f определена в некоторой окрестности точки $x \in X$, где X – нормированное пространство. Функционал $x^* \in X^*$ называется субградиентом функции f в точке x, если $(x^*, h) \le f(x+h) - f(x)$ для любого $h \in X$, для которого f определена в точке $x + h$. Множество всех субградиентов функции f в точке x называется её субдифференциалом и обозначается $\partial f(x)$.*

Подчеркнём, что $\partial f(x)$ – не функционал, а множество функционалов. Из определения субдифферениала следует, что это множество всегда выпукло и замкнуто. В самом деле, для произвольного h множество элементов $x^* \in X^*$, удовлетворяющих неравенству $(x^*, h) \le$

$f(x+h) - f(x)$ есть замкнутое полупространство. Поскольку пересечение любого числа выпуклых замкнутых множеств также выпукло и замкнуто, субдифференциал $\partial f(x)$, как пересечение данных полупространств по всем $h \in X$, является выпуклым и замкнутым. Далее мы увидим (предложение 4.7), что для выпуклой функции множество $\partial f(x)$ ограничено.

Везде далее $f : G \to \mathbb{R}$ – выпуклая функция, определенная на выпуклом множестве $G \subset X$. Начнём с того, что в случае дифференцируемой функции её производная совпадает с субдифференциалом.

Предложение 4.2 *Если f дифференцируема по Гато в точке $x \in \mathrm{int}\,(G)$, то $\partial f(x) = \bigl\{ f'(x) \bigr\}$.*

`Доказательство.` Рассмотрим выпуклую функцию $g(h) = f(x+h) - (f'(x), h)$. Так как $g'(0) = f'(x) - f'(x) = 0$, то, в силу предложения 2.3, имеем $0 \in \mathrm{absmin}\, f$. Следовательно, $g(h) \geq g(0) = f(x)$, т.е., $f(x+h) - f(x) \geq (f'(x), h)$ для любого h. Таким образом, $f'(x) \in \partial f(x)$.

С другой стороны, для любого $x^* \in \partial f(x)$ и любого $h \in X$ имеем

$$\bigl(x^*\,,\, t\,h\bigr) \;\leq\; f(x+th) \;-\; f(x) \;=\; t\,f'(x)[h] \,+\, o(t)\,, \qquad \text{при } t \,\to\, 0\,.$$

Следовательно, $\bigl(x^*\,,\, h\bigr) \;\leq\; f'(x)[h]$ для любого $h \in X$. Применяя это неравенство для элементов h и $-h$, получаем, что $\bigl(x^*\,,\, h\bigr) \;=\; f'(x)[h]$ для всех h, а значит $x^* \;=\; f'(x)$. Таким образом, $\partial f(x) \;=\; \bigl\{ f'(x) \bigr\}$.

□

Итак, для дифференцируемых функций субдифференциал совпадает с производной. А для недифференцируемых ? Оказывается, что субдифференциал существует у любой выпуклой функции. Для доказательства нам понадобится ввести еще одно понятие. *Надграфиком* (epigraph) функции f называется множество

$$\textbf{epi}\, f \quad = \quad \Bigl\{ \,(x,u) \in X \times \mathbb{R} \quad \Bigl| \quad x \in G\,,\; u \geq f(x)\,\Bigr\}.$$

Теорема 4.3 *Субдифференциал выпуклой функции непуст в любой внутренней точке области определения, в которой функция непрерывна.*

`Доказательство.` Пусть $f : G \to \mathbb{R}$ – выпуклая функция, $x \in \mathrm{int}(G)$. Поскольку f непрерывна в x, она ограничена в некоторой окрестности U этой точки. Итак, $M \;=\; \sup_{x \in U}\; f(x) \;<\; \infty$. Тогда открытое множество $\Bigl\{ \,(u,v) \;\in\; U \times \mathbb{R} \,\Bigl| \quad u \in U\,,\; v \;>\; M \Bigr\}$ лежит в $\textbf{epi}\, f$. Итак, надграфик $\textbf{epi}\, f$ имеет непустую внутренность, причем точка $z \;=\; \bigl(x, f(x)\bigr)$ не принадлежит этой внутренности. В самом деле, точка $\bigl(x, f(x) - \varepsilon\bigr)$ стремится к z при $\varepsilon \to +0$ и не лежит в $\textbf{epi}\, f$. Согласно теореме 2.10 выпуклые множества $\textbf{epi}\, f$ и $\{z\}$ отделимы. Это значит, что существует непрерывный ненулевой функционал y^* на пространстве $X \times R$, для которого $\sup_{y \in \textbf{epi}\, f}\,(y^*, y) \;\leq\; (y^*, z)$. Из курса функционального анализа известно, что любой непрерывный линейный функционал y^* на прямом произведении $Y \;=\; X \times \mathbb{R}$ имеет вид $y^*[(u,v)] \;=\; (x^*, u) + c\,v$, где $x^* \in X^*$ и $c \in \mathbb{R}$. Таким образом,

$$\sup_{u \in G\,,\; v \,\geq\, f(u)} \Bigl[\, \bigl(x^*, u\bigr) \;+\; c\,v \,\Bigr] \quad \leq \quad \bigl(x^*, x\bigr) \;+\; c\,f(x)$$

Следовательно, для любых $u \in G\,,\; v \;\geq\; f(u)$ имеем

$$-\,c\Bigl(\,v \;-\; f(x)\Bigr) \quad \geq \quad \bigl(\,x^*\,,\, u - x\,\bigr)\,, \tag{4.6}$$

а значит

$$-c\Big(f(u) - f(x)\Big) \quad \geq \quad \big(x^*\,,\, u-x\big)\,.$$

Обозначая $u - x = h$, получаем

$$-c\Big(f(x+h) - f(x)\Big) \quad \geq \quad \big(x^*\,,\, h\big)\,. \tag{4.7}$$

Так как x – внутренняя точка множества G, элемент h принимает любые значения из некоторого шара $B(0,\varepsilon)$. Заметим, что $c \leq 0$, в противном случае левая часть неравенства (4.6) обращается в $-\infty$ при $v \to +\infty$. Если $c = 0$, то $\big(x^*\,,\, h\big) \leq 0$ для всех $h \in X, \|h\| \leq \varepsilon$. Применяя это неравенство для элементов h и $-h$, получаем $\big(x^*\,,\, h\big) = 0$ для всех $h \in X, \|h\| \leq \varepsilon$, следовательно $x^* = 0$, что невозможно, поскольку $y^* \neq 0$. Таким образом, $c < 0$, и, поделив обе части неравенства (4.7) на число $-c$, получаем

$$f(x+h) - f(x) \quad \geq \quad \Big(-\frac{1}{c}\,x^*\,,\, h\Big)\,.$$

Следовательно, $-\frac{1}{c}\,x^* \in \partial f(x)$.

□

Мы используем без доказательства следующий факт: при $\dim X < \infty$ выпуклая функция непрерывна в любой точке $x \in \mathrm{int}\,(G)$.

Упражнение 4.4 Докажите это.

Следствие 4.5 *При* $\dim X < \infty$ *выпуклая функция имеет непустой субдифференциал в любой точке* $x \in \mathrm{int}\,(G)$.

Доказательство следующих утверждений мы не приводим.

Теорема 4.6 *(Критерий дифференцируемости выпуклой функции). В случае* $\dim X < \infty$ *выпуклая функция* f *дифференцируема в точке* $x \in \mathrm{int}\,(G)$ *тогда и только тогда когда её субдифференциал* $\partial f(x)$ *состоит из одной точки.*

Предложение 4.7 *Для любой выпуклой функции* f *в любой точке непрерывности субдифференциал* $\partial f(x)$ *– ограниченное выпуклое замкнутое множество.*

Следствие 4.8 *В случае* $\dim X < \infty$ *субдифференциал выпуклой функции* f *в любой точке* $x \in \mathrm{int}\,(G)$ *– выпуклый компакт.*

Пример 4.9 Функция $f : \mathbb{R} \to \mathbb{R}\,,\ f(x) = x$ при $x > 0\,,\ f(x) = 0$ при $x \leq 0$ выпукла. Она дифференцируема везде, кроме нуля, поэтому $\partial f(x) = \{f'(x)\}$ при $x \neq 0$. Таким образом, $\partial f(x) = \{1\}$ при $x > 0$ и $\partial f(x) = \{0\}$ при $x < 0$. Найдём субдифференциал в точке $x = 0$. Точка a принадлежит $\partial f(0)$ тогда и только тогда когда $ah \leq f(h) - f(0)$ для всех $h \in \mathbb{R}$. При $h > 0$ имеем $ah \leq h$, откуда $a \leq 1$. При $h \leq 0$ имеем $ah \leq 0$, откуда $a \geq 0$. Таким образом $\partial f(0) = [0,1]$.

Пример 4.10 *(Субдифференциал нормы в нуле).* В нормированном пространстве X норма $f(x) = \|x\|$ является выпуклой функцией. Найдём $\partial f(0)$. Имеем $x^* \in \partial f(0) \Leftrightarrow (x^*, h) \leq \|h\|\ \forall\, h \in X$. В частности, $(x^*, h) \leq 1$ для всех h таких, что $\|h\| = 1$. Итак, $\sup_{\|h\|=1}(x^*, h) \leq 1$, т.е., $\|x^*\| \leq 1$. Таким образом, $\partial f(0) = \big\{\, x^* \in X^* \ \big|\ \|x^*\| \leq 1\,\big\}$ – единичный шар в пространстве X^*. Например, для евклидовой нормы в $\mathbb{R}^d$, т.е., $|x| = \big[\sum_{i=1}^{d}(x^i)^2\big]^{1/2}$ её субдифференциал в нуле – единичный шар с центром в нуле.

Следующая теорема объясняет роль субдифференциалов в выпуклых экстремальных задачах.

Теорема 4.11 *Если f – выпуклая функция и $x \in \mathrm{int}\,(G)$, то $x \in \mathrm{absmin}\, f \Leftrightarrow 0 \in \partial f(x)$.*

`Доказательство.` Соотношение $0 \in \partial f(x)$ означает, что $0 = (0, h) \le f(x+h) - f(x)$ для любого $h \in X$, т.е., $x \in \mathrm{absmin}\, f$.

□

Таким образом, для решения выпуклых экстремальных задач нужно уметь вычислять субдифференциалы выпуклых функций. Основу субдифференциального исчисления составляют теоремы Моро-Ракафеллара (о субдифференциале суммы) и Дубовицкого-Милютина (о субдифференциале максимума). Мы примем обе теоремы без доказательства.

Напомним, что суммой по Минковскому двух множеств $A, B \subset X$ называется множество $A + B = \{ a + b \mid a \in A,\, b \in B \}$. Лёгким упражнением для читателя будет показать, что если два множества выпуклы, то и их сумма выпукла.

Теорема 4.12 *(Моро, Рокафеллар). Если $f_1, f_2 : X \to \mathbb{R}$ – выпуклые функции, имеющие непустые субдифференциалы в точке x, причём хотя бы одна из них в этой точке непрерывна, то $\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x)$.*

Пример 4.13 Если функции f_1, f_2 выпуклы, f_1 дифференцируема в точке x, и $\partial f_2(x) \ne \emptyset$, то $\partial(f_1 + f_2)(x) = f_1'(x) + \partial f_2(x)$. Например, если $f_1(x) = (a, x)$, $a \in X^*$, то $\partial(f_1 + f_2)(x) = a + \partial f_2(x)$. Если $f_1 \equiv \mathrm{const}$, то $\partial(f_1 + f_2)(x) = \partial f_2(x)$

Теорема 4.14 *(Дубовицкий, Милютин) Если $f_i : X \to \mathbb{R}$ – выпуклые функции, непрерывные в точке x, $f(u) = \max_{i=1,\dots,m} f_i(u)$, то*

$$\partial f(x) \quad = \quad \mathrm{co} \left(\bigcup_{f_i(x) = f(x)} \partial f_i(x) \right).$$

Пример 4.15 . Рассмотрим задачу

$$f(x, y) \ = \ \sqrt{x^2 + y^2} \ + \ |x| \ + \ e^{x+y} \ - \ 2y \qquad \to \qquad \min .$$

Функция f выпукла. Кроме того, она непрерывна и коэрцитивна, поэтому её минимум достигается в некоторой точке $(\hat x, \hat y)$. В этой точке имеем $0 \in \partial f(\hat x, \hat y)$. Функция f дифференцируема везде кроме точек вида $(0, y)$ (в них не дифференцируем модуль), а в точке $(0, 0)$ не дифференцируем и корень. Если начать разбор вариантов с "простейшего" случая $\hat x \ne 0$, когда f дифференцируема, то получим систему уравнений $f_x(\hat x, \hat y) = f_y(\hat x, \hat y) = 0$, которую, однако не удастся решить. Как же найти минимум ? Рассмотрим случай $(\hat x, \hat y) = 0$. В этой точке субдифференциал будет наиболее массивным, поэтому есть надежда что он будет содержать ноль. Имеем $\partial(\sqrt{x^2 + y^2})\,(0, 0)$ – круг радиуса 1 с центром в нуле (пример 4.10), $\partial\big(e^{x+y} - 2y\big)\,(0, 0) = \big(e^{x+y} - 2y\big)'\,(0, 0) = \big(1, -1\big)$. Для вычисления субдифферециала модуля применим теорему Дубовицкого-Милютина. Поскольку $|x| = \max\{x, -x\}$, получаем $\partial|x|\,(0, 0) = \mathrm{co}\,\{(-1, 0)\,,\, (1, 0)\}$ – отрезок с концами $(-1, 0)$ и $(1, 0)$. Теперь применяем теорему Моро-Рокафеллара. Множество $\partial f\,(0, 0)$ – есть сумма круга $B(0, 1)$, отрезка $[-1, 1]$ оси OX и вектора $\big(1, -1\big)$. Изобразив это множество (оно является квадратом со стороной 2 с центром в точке $\big(1, -1\big)$ с построенными на его вертикальных сторонах во внешнюю сторону полукругами), получаем, что оно содержит ноль. Таким образом, $(0, 0) \in \mathrm{absmin}$, и минимум равен нулю. Разбор остальных случаев оказался не нужен.

Пример 4.16 . Обратимся еще раз к примеру 1.8 о минимизиции функции $f(x) = \sum_{i=1}^{n} |x - a_i| \to \min$ для заданных точек $a_1, \dots, a_n \in \mathbb{R}^d$, не лежащих на одной прямой. Мы установили, что эта задача всегда имеет единственное решение $\hat{x}$ и охарактеризовали его в случае, когда x не совпадает ни с одной из точек a_i: сумма единичных векторов u_i, сонаправленных векторам $\hat{x} - a_i$, равна нулю. Теперь мы можем завершаем решение, рассмотрев оставшийся случай.

Если $\hat{x} = a_j$ при некотором j, то, обозначив $f_i(x) = |x - a_i|$, имеем: $\partial f_j(\hat{x}) = B(0,1)$ – шар радиуса 1 с центром в нуле (пример 4.10), и $\partial f_i(\hat{x}) = \{u_i\}$ для всех $i \neq j$, поскольку все функции f_i диффересцируемы в точке $\hat{x}$. Тогда по теореме Моро-Рокафеллара (теорема 4.12) имеем $\partial f(a_j) = B(0,1) + \sum_{i\neq j} u_i = B\big(\sum_{i\neq j} u_i\,,\,1\big)$. Тогда (теорема 4.11) $a_j \in \mathrm{absmin}$ в том и только том случае, когда шар с центром $\sum_{i\neq j} u_i$ и радиусом 1 содержит точку 0. Это означает, что $\big|\sum_{i\neq j} u_i\big| \le 1$. Итак,

Для произвольных n точек $a_1, \dots, a_n \in \mathbb{R}^d$, не лежащих на одной прямой, либо найдётся единственная точка $\hat{x}$, не совпадающая ни с одной из них, для которой $\sum_{i=1}^{n} u_i = 0$, либо среди данных точек найдётся единственная, для которой $\big|\sum_{i\neq j} u_i\big| \le 1$. В каждом случае данная точка является решением задачи $\sum_{i=1}^{n} |x - a_i| \to \min$.

Например, при $d = 2, n = 3$, получаем, что точка, сумма расстояний от которой до вершин треугольника минимальна 1) совпадает с вершиной, если угол при этой вершине не меньше 120^o; 2) если таких вершин нет, то совпадает с точкой Ферма-Торичелли-Штейнера, из которой все стороны видны под углом 120^o,

5 Лекция

Теорема Каруша-Куна-Таккера.

В условиях следующей теоремы X – произвольное линейное пространство, не обязательно нормированное, $A \subset X$ – произвольное выпуклое множество.

Теорема 5.1 *(Каруш, Кун, Таккер). Пусть $f_0(x), \dots, f_n(x) : A \to \mathbb{R}$ – выпуклые функции, заданные на выпуклом множестве $A \subset X$. Рассмотрим следующую задачу*

$$\begin{cases} f_0(x) \;\to\; \min \\ \textit{при условиях} \quad x \in A\,,\; f_i(x) \;\le 0, \quad i = 1, \dots, n. \end{cases} \tag{5.8}$$

Если $\hat{x}$ – решение задачи, то существуют НЕРОН множители $\lambda = (\lambda_0, \dots, \lambda_n)$, удовлетворяющие условиям:

а) (стационарность) $\min\limits_{x\in A} \mathcal{L}(x,\lambda) = \mathcal{L}(\hat{x},\lambda)$, где $\mathcal{L}(x,\lambda) = \sum\limits_{i=0}^{n} \lambda_i f_i(x)$ – лагранжиан;

б) (неотрицательность): $\lambda_i \ge 0,\; i = 0, ..., n$;

в) (дополняющая нежёсткость): $\lambda_i f_i(\hat{x}) = 0, i = 1, \dots, n$.

Обратно, если для некоторой точки $\hat{x} \in A$ нашёлся набор множителей λ, удовлетворяющий условиям (а)-(в), причём $\lambda_0 > 0$, то $\hat{x}$ – решение задачи (5.8).

Если выполнено условие Слейтера: существует точка $\tilde{x} \in A$ в которой $f_i(\tilde{x}) < 0,\; i = 1, \dots, n$, то $\lambda_0 > 0$.

`Доказательство`. Без ограничения общности будем считать, что $f_0(\hat{x})=0$, в противном случае мы можем прибавить или отнять константу. Определим множества $B \subset R^{n+1}$ и $C \subset R^{n+1}$:

$$\begin{array}{rcl} B & = & \Big\{ v = (v_0,\dots,v_n) : \ v_i < 0, \ i=0,\dots,n \Big\} \\ C & = & \Big\{ u = (u_0,\dots,u_n) : \ \exists\, x \in A \ : \ f_i(x) \le u_i \, , \ i=0,\dots,n \Big\} \end{array}$$

Множество B выпукло и открыто. Проверим, что и множество C выпукло. Возьмём произвольное $t \in [0,1]$. Пусть $u, z \in C$. Это означает, что существуют такие $x, y \in A$, что

$$f_i(x) \ \le \ u_i, \ f_i(y) \ \le z_i \, , \qquad i=0,\dots,n\, .$$

Тогда

$$f_i\big((1-t)x+ty\big) \ \le \ (1-t)f_i(x) \ + tf_i(y) \ \le \ (1-t)u_i \ + \ tz_i \, , \qquad i=0,\dots,n\, .$$

Следовательно, $(1-t)u + tz \in C$. Мы доказали, что C выпукло. Далее, множества B и C не пересекаются. Действительно, если $u=(u_0,\dots,u_n) \in B \cap C$, то существует такой x, что $f_i(x) \le u_i < 0$. Тогда x – допустимая точка в задаче (5.8), $f_0(x)<0$, следовательно, $\hat{x}$ – не решение. Итак, B и C выпуклы, и B открыто, значит эти множества отделимы (теорема 2.10). Вспоминая общий вид линейного функционала в $\mathbb{R}^{n+1}$, получаем, что существует вектор $\lambda \in \mathbb{R}^{d+1} \setminus \{0\}$, для которого $\sup_{v\in B} \big(\lambda, v\big) \ \le \ \inf_{u \in C} \big(\lambda, u\big)$. Заметим, сначала, что $\lambda_i \ge 0$ для всех $i=0,\dots,n$. В самом деле, если $\lambda_j < 0$ для некоторого j, то можем взять $v_j \to -\infty$ и $v_i \to 0$ для остальных i, тогда получим $\sup_{v\in B} \big(\lambda, v\big) = +\infty$. Это невозможно, поскольку для любой точки $\bar{u} \in C$ имеем $\inf_{u\in C} \big(\lambda, u\big) \ \le \ \big(\lambda, \bar{u}\big) \ < \ +\infty$. Итак, $\lambda \ge 0$. Так как при $v \to 0$ имеем $\big(\lambda, v\big) \to 0$, то $\sup_{v \in C} \big(\lambda, v\big) \ \ge \ 0$, откуда $\inf_{u\in C} \big(\lambda, u\big) \ \ge \ 0$. Однако, если взять $\hat{u} = \big(f_0(\hat{x}), \dots, f_n(\hat{x})\big)$, то $\hat{u}_i \le 0$, и следовательно $(\lambda, \hat{u}) \le 0$. Таким образом,

$$\inf_{u\in C} \big(\lambda, u\big) \quad = \quad \big(\, \lambda \, , \, \hat{u} \, \big) \quad = \quad 0\, .$$

С другой стороны, для любого $x \in A$ имеем $u = \big(f_0(x), \dots, f_n(x)\big) \in C$, поэтому $\mathcal{L}\big(\, x \, , \, \lambda \, \big) = \big(\, \lambda \, , \, u\big) \ge 0$. Таким образом, $\min_{x\in A} \mathcal{L}\big(\, x\, , \, \lambda\, \big) \ = \ \mathcal{L}\big(\, \hat{x}\, , \, \lambda\, \big) = 0$. Итак, набор $\lambda = (\lambda_0, \dots, \lambda_n)$ удовлетворяет условиям (а) и (б) (стационарности и неотрицательности). Проверим условие (в) дополняющей нежёсткости. Заметим сначала, что $\lambda_i \ge 0$ и $\hat{u}_i \le 0$, а значит $\lambda_i\, \hat{u}_i \ \le 0$ для любого $i=0,\dots,n$. Поскольку $\sum_{i=0}^n \lambda_i\, \hat{u}_i \ = \ \big(\, \lambda\, , \, \hat{u}, \big) \ = \ 0$, получаем $\lambda_i\, \hat{u}_i \ = 0$ для каждого i.

Теперь установим достаточность. Если вектор $\lambda \ne 0$ таков, что в некоторой точке $\hat{x}$ выполнены условия (а) – (в), то

$$(\lambda\, , \, \hat{u}) \ = \ \lambda_0\, f_0(\hat{x}) \ = \ \inf_{x\in A} \ \mathcal{L}(x\, , \, \lambda) \ \le \ \inf_{u\in C} (\lambda\, , \, u) \ \le \ (\lambda\, , \, \hat{u})\, .$$

Первое равенство выполнено в силу свойства (в), второе – в силу свойства (а), третье неравенство – поскольку для любого $u \in C$ найдётся $x \in A$, для которого $u_i \ge f_i(x)\, , \ i=1,\dots,n$, что в силу неотрицательности λ, влечёт $(\lambda, u) \ \ge \ \mathcal{L}(x, \lambda)$, четвёртое – по определению inf. Далее, прибавляя или вычитая константу из функции f_0, можно считать, что $f_0(\hat{x}) \ = \ 0$. Следовательно,

$$0 \ = \ \lambda_0\, f_0(\hat{x}) \ = \ (\lambda, \hat{u}) \ = \ \min_{u\in C} (\lambda, u)\, . \tag{5.9}$$

Если $\lambda_0 > 0$, то $\hat{x}$ – решение задачи. Иначе найдётся $\bar{x}$, для которого $f_0(\bar{x}) \ < \ 0$ и $f_i(\bar{x}) \ \le \ 0\, , \ i = 1,\dots,n$. Тогда для вектора $\bar{u} = \big(f_0(\bar{x}), \dots, f_n(\bar{x})\big)$ имеем $(\lambda, \bar{u}) < 0$, что противоречит (5.9).

Наконец, если выполнено условие Слейтера: $f_i(\tilde{x}) < 0,\ i = 1, \dots, n$ в некоторой точке $\tilde{x} \in A$, то в случае $\lambda_0 = 0$ для вектора $\tilde{u} = \big(f_0(\tilde{x}), \dots, f_n(\tilde{x})\big)$ получим $(\lambda, \tilde{u}) < 0$, что противоречит (5.9). В самом деле, в силу условия НЕРОН $\lambda_j > 0$ для некоторого $j \geq 1$, а значит $(\lambda, \tilde{u}) = \sum_{i=1}^{n} \lambda_i \tilde{u}_i \leq \lambda_j \tilde{u}_j < 0$.

□

Напомним, что неравенство $f_i(x) \leq 0$ называется активным, если $f_i(\hat{x}) = 0$ и *пассивным* если $f_i(\hat{x}) < 0$. Условие (в) дополняющей нежёсткости означает, что $\lambda_i = 0$ для всех пассивных неравенств. Обозначим $I(\hat{x}) = \{0\} \cap \{\, i \mid f_i(\hat{x}) = 0\}$, т.е., это множество состоит из нуля и индексов всех активных неравенств. Все множители λ_i неотрицательны (условие (б)), и по условию НЕРОН имеем $\sum_{i=0}^{n} \lambda_i > 0$. Разделив все множители на их сумму, можем считать, что $\sum_{i=0}^{n} \lambda_i = 1$. Условие (а) равносильно условию $0 \in \partial \mathcal{L}_x(\hat{x}, \lambda)$ (теорема 4.11), которое, в силу теоремы Моро-Рокафеллара (теорема 4.12), означает, что $0 \in \sum_{i \in I(\hat{x})} \lambda_i \, \partial f_i(\hat{x})$. Это значит, что найдутся такие точки $z_i \in \partial f_i(\hat{x})\,,\ i \in I(\hat{x})$, что $\sum_{i \in I(\hat{x})} \lambda_i \, z_i = 0$, т.е., точка $\{0\}$ является выпуклой комбинацией точек $\{z_i\}$. А это, в свою очередь, означает, что $0 \in \mathrm{co} \big\{\, \partial f_i(\hat{x})\,,\ i \in I(\hat{x}) \,\big\}$. Таким образом, доказана

Теорема 5.2 *Если X – нормированное пространство, $\hat{x} \in \operatorname{int} A$ и все функции $f_0, \dots, f_m$ непрерывны в точке $\hat{x}$, то в теореме Каруша-Куна-Таккера (теорема 5.1) условия (а) – (в) равносильны следующему условию: выпуклая оболочка субдифференциалов целевой функции f_0 и активных неравенств f_i в точке $\hat{x}$ содержит ноль.*

6 Лекция

Простейшая задача вариационного исчисления. Уравнения Эйлера-Лагранжа. Интегралы импульса и энергии.

Простейшая задача вариационного исчисления. Слабый минимум. Необходимые условия первого порядка. Уравнения Эйлера-Лагранжа. Импульс и энергия. Пример Гильберта. Задача о минимальной площади поверхности.

Простейшей задачей вариационного исчисления называется следующая задача:

$$\begin{cases} \mathcal{J}(x) \;=\; \int\limits_{t_0}^{t_1} L\big(t, x, \dot{x}\big)\, dt \quad \to \quad \min\,, \\ \\ x \in C^1\big([t_0, t_1]\,, \mathbb{R}^n\big)\,, \\ x(t_0) = x_0\,, \quad x(t_1) = x_1\,, \end{cases} \tag{6.10}$$

где $x(t) = \big(x_1(t), \ldots, x_n(t)\big)$ – непрерывно-дифференцируемая вектор-функция из отрезка $[t_0, t_1]$ в $\mathbb{R}^n$, $x_0, x_1 \in \mathbb{R}^n$ – заданные точки (граничные условия), $L \in C\left([t_0,t_1]\times\mathbb{R}^n\times\mathbb{R}^n, \mathbb{R}\right)$ – заданная функция, называемая *интегрантом.* Таким образом, среди всех непрерывно-дифференцируемых функций, принимающих данные значения на концах отрезка, найти такую, которая доставляет минимум интегральному функционалу $\mathcal{J}(x)$. Функции $x \in C^1\left([t_0,t_1], \mathbb{R}^n\right)$, удовлетворяющие данным граничным условиям, будем называть *допустимыми.*

Определение 6.1 *Допустимая функция $\hat{x} \in C^1\left([t_0,t_1], \mathbb{R}^n\right)$ доставляет слабый локальный минимум в задаче (6.10), если существует $\varepsilon > 0$ такой, что $\mathcal{J}(x) \ \geq\ \mathcal{J}(\hat{x})$ для любой допустимой функции x, удовлетворяющей условию $\|x - \hat{x}\|_{C^1[t_0,t_1]} < \varepsilon$.*

Термин *слабый* относится к метрике пространства C^1, которая определяет окрестность для данного локального минимума. Таким образом, точка $\hat{x}$ доставляет слабый локальный минимум, если для всех допустимых x, расположенных достаточно близко от $\bar{x}$ в метрике пространства C^1, имеем $\mathcal{J}(x) \ \geq\ \mathcal{J}(\hat{x})$. Следующая теорема даёт необходимые условия слабого локального минимума.

Теорема 6.2 *Предположим, что в задаче (6.10) функции L, L_x и $L_{\dot{x}}$ непрерывны. Если функция $\hat{x}$ доставляет слабый локальный минимум, то выполнено уравнение Эйлера-Лагранжа:*

$$\frac{d}{dt}\, L_{\dot{x}}\,(t, \hat{x}, \dot{\hat{x}}) \quad = \quad L_{x}\,(t, \hat{x}, \dot{\hat{x}})\,. \tag{6.11}$$

Любое решение $\hat{x}(t)$ уравнения (6.11) называется *экстремалью.* Уравнение (6.11) является фактически системой из n уравнений $\frac{d}{dt} L_{\dot{x}_i} \ = \ L_{x_i}\,,\ i = 1, \ldots, n$. Таким образом, для поиска экстремали имеется система из n дифференциальных уравнений второго порядка и $2n$ граничных условий $x(t_0) = x_0, x(t_1) = x_1$ (т.е., $x_i(t_0) = x_{0i}, x_i(t_1) = x_{1i}\ \ i = 1, \ldots, n$). Согласно теореме 6.2 любая функция, доставляющая слабый локальный минимум, является экстремалью. Но, вообще говоря, не наоборот. Мы докажем теорему 6.2 в случае $n = 1$. Многомерный случай полностью аналогичен, мы оставляем его читателю. Доказательство состоит из двух лемм. Пространство функций из $C^1[t_0, t_1]$, удовлетворяющих условию $h(t_0) = h(t_1) = 0$ будем обозначать $C^1_0[t_0, t_1]$ или просто C^1_0.

Лемма 6.3 *Если функции L, L_x и $L_{\dot{x}}$ непрерывны, то для любой функции $h \in C^1_0[t_0, t_1]$ имеем*

$$\delta_{\mathcal{J}}\,(x, h) \quad = \quad \int_{t_0}^{t_1} \Big(\, L_x\, h \,+\, L_{\dot{x}}\, \dot{h} \Big)\, dt\,.$$

`Доказательство.` Так как частные производные L_x и $L_{\dot{x}}$ непрерывны, то функция $\varphi(\lambda) = L(t, x + \lambda h, \dot{x} + \lambda \dot{h})$ дифференцируема по λ. Пользуясь правилом дифференцирования функции многих переменных и дифференцированием интеграла по параметру, имеем

$$\delta_{\mathcal{J}}\,(x, h) \,=\, \lim_{\lambda \to 0} \frac{\mathcal{J}(x + \lambda h) - \mathcal{J}(x)}{\lambda} \;=\; \int_{t_0}^{t_1} \lim_{\lambda \to 0} \left(\frac{L(t, x + \lambda h, \dot{x} + \lambda \dot{h}) \,-\, L(t, x, \dot{x})}{\lambda} \right) dt \;=$$

$$\int_{t_0}^{t_1} \frac{d}{d\lambda} L(t, x + \lambda h, \dot{x} + \lambda \dot{h})\Big|_{\lambda = 0} dt \quad = \quad \int_{t_0}^{t_1} \Big(\, L_x\, h \,+\, L_{\dot{x}}\, \dot{h} \Big)\, dt\,.$$

□

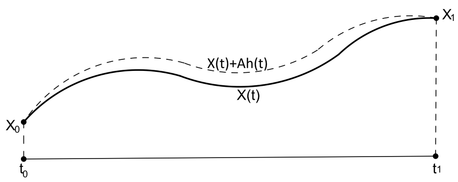

Рис. 1: Вариация функции $x(t)$

Лемма 6.4 *(Дюбуа-Реймона). Пусть $a(t), b(t) \in C\,[t_0, t_1]$ и для любой функции $h \in C_0^1[t_0, t_1]$ имеем*

$$\int_{t_0}^{t_1} \Big(a(t)\, h(t) \;+\; b(t)\, \dot h(t) \Big)\, dt \;=\; 0\,.$$

Тогда функция $b(t)$ непрерывно дифференцируема и $b'(t) \,=\, a(t)$.

`Доказательство.` Пусть $A(t)$ – любая первообразная функции $a(t)$, т.е., $A(t) \;=\; \int_{t_0}^{t} a(\tau)\, d\tau + K$. Интегрируя по частям, имеем

$$0 \;=\; \int_{t_0}^{t_1} \Big(a(t)\, h(t) \;+\; b(t)\, \dot h(t) \Big)\, dt \;=\; A(t)h(t)\Big|_{t_0}^{t_1} + \int_{t_0}^{t_1} \Big(-A(t) \;+\; b(t) \Big)\, \dot h(t)\; dt\,.$$

Теперь учитываем, что $h(t_0) \;=\; h(t_1) \;=\; 0$:

$$\int_{t_0}^{t_1} \Big(-A(t) \;+\; b(t) \Big)\, \dot h(t)\; dt \quad = \quad 0\,.$$

Это равенство выполнено для любой функции $h \in C_0^1[t_0, t_1]$. Выберем теперь нужную функцию h. Для этого подберем константу K таким образом, чтобы интеграл функции $-A(t) + b(t)$ по отрезку $[t_0, t_1]$ был равен нулю. Тогда функция $h(t) \;=\; \int_{t_0}^{t} \big(-A(\tau) \;+\; b(\tau) \big)\, d\tau$ принадлежит $C_0^1[t_0, t_1]$, и при этом

$$\int_{t_0}^{t_1} \Big(-A(t) \;+\; b(t) \Big)\, \dot h(t)\, dt \quad = \quad \int_{t_0}^{t_1} \Big(-A(t) \;+\; b(t) \Big)^2 dt \;=\; 0\,.$$

Следовательно, $b(t) \;\equiv\; A(t)$. Поэтому $b \in C^1[t_0, t_1]$ и $b'(t) \;=\; a(t)$.

□

`Доказательство теоремы 6.2.` Если $\bar x$ – точка локального минимума функционала $\mathcal{J}(x)$ в пространстве $C_0^1[t_0, t_1]$ то, согласно теореме 2.1 (лекция 2), $\delta_{\mathcal{J}}(\bar x, h) \;=\; 0$ для любой функции $h \in C_0^1[t_0, t_1]$. Применяя теперь лемму 6.3, а затем лемму 6.4 для $a = L_x$, $b = L_{\dot x}$, завершаем доказательство.

□

Применив теперь предложение 2.1, получаем

Следствие 6.5 *Если интегрант в простейшей задаче является выпуклым функционалом от x, т.е.,*

$$L\Big(t\,,\, (1-\lambda)x + \lambda y\,,\, (1-\lambda)\dot x + \lambda \dot y \Big) \quad \le \quad (1-\lambda)\, L(t, x, \dot x) \;+\; \lambda\, L(t, y, \dot y)$$

для любых допустимых $x, y \;\in\; C^1[t_0, t_1]$ в любой точке $t \;\in\; [t_0, t_1]$, то уравнение Эйлера-Лагранжа является достаточным условием абсолютного минимума.

В некоторых частных случаях уравнение Эйлера-Лагранжа может быть сведено у уравнению первого порядка. Введем ещё два определения: интеграла импульса $p(t) \ = \ L_{\dot{x}}$ и интеграла энергии $H(t) \ = \ \dot{x}\, L_{\dot{x}} \ - \ L$.

Предложение 6.6 *Если интегрант $L(t)$ не зависит явно от $\dot{x}$, т.е., $L(t,x,\dot{x}) \ = \ L(t,x)$, то уравнение Эйлера-Лагранжа равносильно уравнению $\hat{L}_x(t) \ \equiv \ 0$.*

Если интегрант $L(t)$ не зависит явно от x, то уравнение Эйлера-Лагранжа равносильно уравнению $\hat{p}(t) \ = \ \hat{L}_{\dot{x}}(t) \ \equiv \ \mathrm{const}$.

Если интегрант $L(t)$ не зависит явно от t, то из уравнения Эйлера-Лагранжа следует, что $\hat{H}(t) \ = \ \hat{L}_{\dot{x}}\dot{\hat{x}} - L \ \equiv \ \mathrm{const}$. Если известно, что экстремаль $\hat{x}$ не является тождественной константой ни на каком интервале, то верно и обратное: из уравнения $\hat{H} \ \equiv \ \mathrm{const}$ следует уравнение Эйлера-Лагранжа.

`Доказательство.` Первые два пункта очевидны, докажем третий. Имеем

$$H' \ = \ \frac{d}{dt}\Big(L_{\dot{x}}\Big)\,\dot{x} \ + \ L_{\dot{x}}\,\ddot{x} \ - \ L_{x}\,\dot{x} \ - \ L_{\dot{x}}\,\ddot{x} \ = \ \Big(\frac{d}{dt}\,L_{\dot{x}} \ - \ L_{x}\Big)\,\dot{x}.$$

Поэтому из уравнения Эйлера-Лагранжа следует, что $\hat{H}' \equiv 0$, т.е., $\hat{H} \ \equiv \ \mathrm{const}$. Если функция $\dot{\hat{x}}(t)$ не обращается в ноль ни на каком интервале, то функция $\frac{d}{dt}\,\hat{L}_{\dot{x}} \ - \ \hat{L}_{x}$ равна нулю на всюду плотном подмножестве отрезка $[t_0, t_1]$, а значит (в силу непрерывности), и на всём отрезке.

□

Пример 6.7 *(Пример Гильберта).* Рассмотрим задачу

$$\begin{cases} J(x) \ = \ \int\limits_0^1 t^{2/3}\,\dot{x}^2\,dt \quad \to \ \min\,; \\ x(0) = 0\,, \ x(1) = 1\,. \end{cases} \tag{6.12}$$

Так как интегрант не зависит явно от x, то уравнение Эйлера-Лагранжа дает $\frac{d}{dt}\,L_{\dot{x}} \ = \ 0$, откуда $L_{\dot{x}} = \mathrm{const}$, следовательно $2t^{2/3}\dot{x} = c$. Единственное решение этого дифференциального уравнения, удовлетворяющее начальным условиям, $\hat{x}(t) = t^{1/3}$. Докажем, что $\hat{x} \in \mathrm{absmin}$. Для любой допустимой вариации $h \in C_0^1[0,1]$ имеем

$$J(\hat{x}+h) - J(\hat{x}) \ = \ \int_0^1 [t^{2/3}(\dot{\hat{x}}+\dot{h})^2 - t^{2/3}\dot{\hat{x}}^2]\,dt \ = \ \int_0^1 [2t^{2/3}\dot{\hat{x}}\dot{h} \ + \ t^{2/3}\dot{h}^2]\,dt \ \geq$$

$$\int_0^1 2t^{2/3}\dot{\hat{x}}\dot{h}\,dt \ = \ \int_0^1 2t^{2/3}\frac{1}{3}t^{-2/3}\,\dot{h}\,dt \ = \ \frac{2}{3}\big(h(1)-h(0)\big) \ = \ 0\,.$$

Таким образом, $\hat{x}$ доставляет абсолютный минимум. Тем не менее, эта функция не является экстремалью, поскольку она не принадлежит $C^1[0,1]$. Поэтому, в данной задаче вообще нет допустимых экстремалей. Этот пример показывает, что в некоторых случаях пространство C^1 слишком узко для решения простейшей задачи, и имеет смысл искать экстремали в более широких пространствах, например в пространствах Соболева.

Пример 6.8 *(Задача о минимальной площади поверхности вращения).* Общая задача Лагранжа-Плато состоит в нахождении поверхности минимальной площади, содержащей заданное компактное множество в $\mathbb{R}^3$. Мы рассмотрим случай, когда это множество – два круга радиусом 1, причем отрезок между их центрами равен $2a$ и перпендикулярен плоскостям кругов. Мы ограничимся только поверхностями вращения (что выглядит естественно, но не так просто обосновывается). Если поверхность образована вращением графика функции $x(t)$ такой, что $x(-a) = x(a) = 1$, вокруг оси абсцисс, то задача формализуется в виде:

$$\begin{cases} J(x) \;=\; \int\limits_{-a}^{a} 2\pi x\sqrt{1+\dot{x}^2}\,dt \;\to\; \min\,; \\ x(-a) = x(a) = 1\,. \end{cases} \tag{6.13}$$

Так как интегрант не зависит явно от t, можем воспользоваться интегралом энергии: $H =$

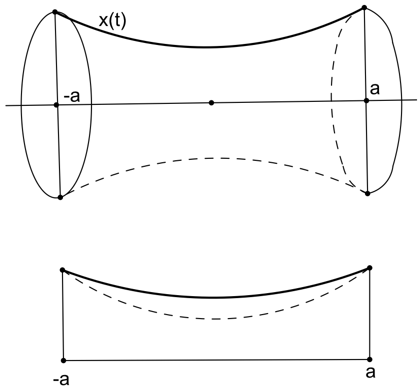

Рис. 2:

$\dot{x}L_{\dot{x}} - L \;=$ const. Вычислив производные и проделав очевидные преобразования, получаем $\frac{x}{\sqrt{1+\dot{x}^2}} = c$, откуда $\frac{dx}{\sqrt{\left(\frac{x}{c}\right)^2-1}} = dt$. С помощью замены $x = c\,\mathrm{ch}\,\tau$, находим решение данного дифференциального уравнения, удовлетворяющее начальным условиям: $\hat{x}(t) = c\,\mathrm{ch}\left(\frac{t}{c}\right)$. Остается найти параметр c из краевого условия $\hat{x}(a) = 1$ (условие $\hat{x}(-a) = 1$ будет тогда выполнено в силу чётности функции). Обозначив $s = \frac{a}{c}$, получаем уравнение $\mathrm{ch}\,s = \frac{s}{a}$. Пусть число a_0 таково, что прямая $y = \frac{s}{a_0}$ касается графика функции $y = \mathrm{ch}\,s$ (рис. 3). Имеем $a_0 = 0.662\ldots$. При каждом $a < a_0$ прямая пересекает график в двух точках, поэтому существует два значения параметра c, т.е., задача имеет две допустимые экстремали. При $a = a_0$ экстремаль единственна, а при $a > a_0$ экстремалей нет. Последнее объясняется тем, что при $a > a_0$ площадь любой поверхности вращения становится больше суммы площадей двух кругов радиусом 1, поэтому минимальная поверхность "распадается" в объединение двух кругов. В случае $a < a_0$ из двух экстремалей абсолютный минимум дает одна, соответствующая меньшему из двух значений c, т.е., меньшему корню уравнения $c\,\mathrm{ch}\left(\frac{a}{c}\right) = 1$; вторая не доставляет даже локального минимума. Это мы доказывать не будем.

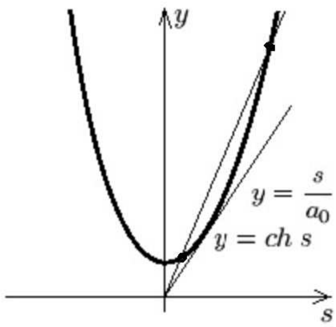

Рис. 3:

7 Лекция

Задача Больца. Условия трансверсальности.
Изопериметрическая задача.

Условия первого порядка в задаче Больца. Условия трансверсальности.
Изопериметрическая задача. Задача Дидоны.

Задачей Больца называется задача минимизации следующего функционала на пространстве $C^1[t_0, t_1]$:

$$\mathcal{J}(x) \quad = \quad \int\limits_{t_0}^{t_1} L\left(t, x, \dot{x}\right) dt \ + \ l(x_0, x_1) \quad \to \quad \min\,, \tag{7.14}$$

где $x(t) \in C^1\big([t_0, t_1], \mathbb{R}^n\big)$, $x_0 = x(t_0), x_1 = x(t_1)$; *интегрант* $L \in C\left([t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}\right)$ и *терминант* $l \in C\big(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}\big)$ – заданные функции. В отличие от простейшей задачи, задача Больца является безусловной, т.е., никаких дополнительных условий (например, граничных) на функцию x не накладывается. Вместо этого граничные значения появляются в терминанте.

Определение 7.1 *Функция $\hat{x} \in C^1\left([t_0, t_1], \mathbb{R}^n\right)$ доставляет слабый локальный минимум в задаче (7.14), если существует $\varepsilon > 0$ такой, что $\mathcal{J}(x) \ \geq \ \mathcal{J}(\hat{x})$ для любой функции x, удовлетворяющей условию $\|x - \hat{x}\|_{C^1[t_0, t_1]} < \varepsilon$.*

Теорема 7.2 *Предположим, что в задаче (7.14) функции $L, L_x, L_{\dot{x}}, l_{x_0}$ и l_{x_1} непрерывны. Если функция $\hat{x}$ доставляет слабый локальный минимум, то выполнено уравнение Эйлера-Лагранжа (6.11) и условия трансверсальности:*

$$L_{\dot{x}}(t_0, \hat{x}(t_0), \dot{\hat{x}}(t_0)) \ = \ l_{x_0}(\hat{x}_0, \hat{x}_1)\,; \qquad L_{\dot{x}}(t_1, \hat{x}(t_1), \dot{\hat{x}}(t_1)) \ = \ -\,l_{x_1}(\hat{x}_0, \hat{x}_1), \tag{7.15}$$

или в короткой записи $\hat{L}_{\dot{x}}(t_k) \ = \ (-1)^k \hat{l}_{x_k}\,,\ k = 0, 1.$

`Доказательство` проведём для случая $n = 1$. Имеем

$$\delta_{\mathcal{J}}\left(x, h\right) = \int_{t_0}^{t_1} \Big(L_x h + L_{\dot x} \dot h \Big)\, dt + l_{x_0} h(t_0) + l_{x_1} h(t_1).$$

В точке локального минимума $\hat x$ должно быть $\delta_{\mathcal{J}}\left(\hat x, h\right) = 0$ для любой функции $h \in C^1[t_0, t_1]$. Ограничимся сначала функциями h из пространства $C_0^1[t_0, t_1]$. Для всех таких функций имеем $\int_{t_0}^{t_1} \big(\hat L_x h + \hat L_{\dot x} \dot h \big)\, dt = 0$. Тогда из леммы Дюбуа-Реймона заключаем, что на экстремали $\hat x$ функция $\hat L_{\dot x}$ непрерывно дифференцируема и выполнено уравнение Эйлера-Лагранжа $\frac{d}{dx} \hat L_{\dot x} = \hat L_x$. Берём теперь произвольную функцию $h \in C^1[t_0, t_1]$. Интегрируя по частям и учитывая уравнение Эйлера-Лагранжа, получаем $\int_{t_0}^{t_1} \Big(\hat L_x h + \hat L_{\dot x} \dot h \Big)\, dt = \hat L_{\dot x}(t)\, h(t) \Big|_{t_0}^{t_1}$. Таким образом

$$\hat L_{\dot x}(t)\, h(t) \Big|_{t_0}^{t_1} + \hat l_{x_0} h(t_0) + \hat l_{x_1} h(t_1) \quad = \quad 0\,.$$

Итак,

$$\Big(\hat l_{x_0} - \hat L_{\dot x}(t_0)\Big) h(t_0) \quad + \quad \Big(\hat l_{x_1} + \hat L_{\dot x}(t_1)\Big) h(t_1) \quad = \quad 0$$

для любой функции $h \in C^1[t_0, t_1]$. Значит данная сумма равна нулю при всех значениях $h(t_0)$ и $h(t_1)$, следовательно, $\hat l_{x_0} - \hat L_{\dot x}(t_0) = 0$ и $\hat l_{x_1} + \hat L_{\dot x}(t_1) = 0$.

$\square$

Аналогично следствию 6.5 получаем

Следствие 7.3 *Если интегрант в задаче Больца является выпуклым функционалом от x, а терминант – выпуклой функцией на $\mathbb{R}^2$, то уравнение Эйлера-Лагранжа и условия трансверсальности достаточны для абсолютного минимума.*

Изопериметрической задачей называется задача минимизации интегрального функционала:

$$\mathcal{J}_0(x) \quad = \quad \int\limits_{t_0}^{t_1} f_0\left(t, x, \dot x\right) dt \quad \to \quad \min\,, \tag{7.16}$$

при ограничениях $\mathcal{J}_k(x) = \int\limits_{t_0}^{t_1} f_k\left(t, x, \dot x\right) dt = a_k\,,\ k = 1, \dots, m$ и при условиях $x(t_0) = x_0\,,\ \ x(t_1) = x_1\,,\ x(t) \in C^1\big([t_0, t_1], \mathbb{R}^n\big)$. При этом $f_k \in C([t_0, t_1]\,, \mathbb{R}^n\,, \mathbb{R}^n)\,,\ k = 0, \dots, m$ – заданные функции, a_k – заданные числа. Функция $\hat x \in C^1\left([t_0, t_1]\,, \mathbb{R}^n\right)$ доставляет слабый локальный минимум, если существует $\varepsilon > 0$ такой, что $\mathcal{J}_0(x) \ \geq\ \mathcal{J}_0(\hat x)$ для любой функции x, удовлетворяющей всем условиям задачи (7.16) и условию $\big\|x - \hat x\big\|_{C^1[t_0, t_1]} < \varepsilon$.

Теорема 7.4 *Предположим, что в изопериметрической задаче функции $f_k, (f_k)_x, (f_k)_{\dot x}$ непрерывны при всех $k = 0, \dots, m$. Если функция $\hat x$ доставляет слабый локальный минимум, то существуют НЕРОН множители $\lambda_0, \dots, \lambda_m$ такие, что интегрант $L(t, x, \dot x) = \sum_{k=0}^m \lambda_k f_k(t, x, \dot x)$ удовлетворяет при $x = \hat x$ уравнению Эйлера-Лагранжа (6.11): $\frac{d}{dt} \hat L_{\dot x} = \hat L_x$.*

`Доказательство`, как всегда, проведем для $n = 1$. Имеем,

$$\delta_{J_k}(x, h) \quad = \quad \int_{t_0}^{t_1} \Big((f_k)_x h + (f_k)_{\dot x} \dot h \Big)\, dt\,, \qquad k = 0, \dots, m\,.$$

Далее, согласно теореме Лагранжа (теорема 2.3), применённой к пространству $X = C_0^1[t_0, t_1]$, получаем, что существуют НЕРОН множители $\lambda_0, \dots, \lambda_m$, для которых $\sum_{k=0}^m \lambda_k \mathcal{J}_k{}'(\hat{x}) = 0$. Следовательно, интегрант $L(t, x, \dot{x}) = \sum_{k=0}^m \lambda_k f_k(t, x, \dot{x})$ в точке $x = \hat{x}$ удовлетворяет уравнению $\int_{t_0}^{t_1} \left(\hat{L}_x h + \hat{L}_{\dot{x}} \dot{h} \right) dt = 0$ для всех $h \in C_0^1[t_0, t_1]$. Применив теперь лемму 6.4, получаем, что функция $\hat{L}_{\dot{x}}$ непрерывно дифференцируема по t, и функция $\hat{L}$ удовлетворяет уравнению Эйлера-Лагранжа.

□

Пример 7.5 (Задача Дидоны). *Среди всех линий данной длины с концами на данной прямой линии найти ту, которая ограничивает фигуру наибольшей площади.*

Данную задачу можно считать первой задачей теории экстремма, поскольку она была поставлена в IX в. до н.э., задолго до Архимеда или Евклида. История, которую поведал древнеримский поэт Публий Вергилий Марон в поэме "Энеида" происходила в IX веке до н.э. на ближнем Востоке. Финикийская царевна Дидона, дочь Финикийского царя, спасаясь от преследований своего брата, отправилась на корабле в сторону запада, вдоль южного берега Средиземного моря. Остановилась в живописном месте на берегу Тунисского залива. Переговоры с предводителем местного племени Ярбом о выделении участка земли не дали результата. Ярб заявил, что даст царевне столько земли, сколько можно ограничить одной бычьей шкурой. Дидона распорядилась разрезать шкуру на тонкие полоски, из которых сплели длинную верёвку. Этой веревкой о огородили фигуру наибольшей площади от прямолинейного берега. Так была основана крепость, а рядом – город Карфаген.

Равносильная задача: основное изопериметрическое неравенство. *Среди всех замкнутых кривых данной длины наибольшую площадь ограничивает круг.*

Эта задача равносильна задаче Дидоны. Достаточно сделать осевую симметрию относительно линии берега. Строго это было доказано лишь в XX веке, в книге В.Бляшке "Круг и шар" [7]. Мы запишем задачу Дидоны в следующем виде: *Среди всех функций $x(t)$, для которых $x(t_0) = x_0, x(t_1) = x_1$, а график имеет заданную длину l найти ту, график которой ограничивает наибольшую площадь.* Если доказать, что ответ – дуга окружности, то все будет сделано. И в задаче Дидоны, и в основном изопериметрическом неравенстве получаем, что локально экстремальная кривая является окружностью (для любых двух точек кривой, расположенных достаточно близко друг от друга, дуга кривой, соединяющая их, является дугой окружности.) Отсюда следует основное изопериметрическое неравенство и решение задачи Дидоны – полукруг. Итак:

$$\begin{aligned} &\int_{t_0}^{t_1} x(t)\, dt \quad \to \quad \max, \\ &\int_{t_0}^{t_1} \sqrt{1 + \dot{x}^2(t)}\, dt \;=\; l, \\ &x(t_0) \;=\; x_0,\; x(t_1) = x_1 . \end{aligned}$$

Согласно теореме 7.4 интегрант $L = -\lambda_0 x + \lambda_1 \sqrt{1 + \dot{x}^2}$ удовлетворяет уравнению Эйлера-Лагранжа. Так как интегрант не зависит явно от t, то можно воспользоваться интегралом энергии (предложение 6.6): $\hat{H}(t) = \dot{x} \hat{L}_{\dot{x}} - \hat{L} \equiv c$. В нашем случае получаем (для удобства далее мы не ставим крышки над x):

$$\hat{H}(t) \;=\; \frac{\lambda_1 \dot{x}^2}{\sqrt{1 + \dot{x}^2}} \;-\; \lambda_1 \sqrt{1 + \dot{x}^2} \;+\; \lambda_0 x \quad \equiv \quad c,$$

откуда

$$\lambda_0 x - \frac{\lambda_1}{\sqrt{1 + \dot{x}^2}} = c\,.$$

Если $\lambda_0 = 0$, то $\hat{x}(t) \equiv \mathrm{const}$, значит $x(t)$ – линейная функция, и её график – прямая. Это – частный случай, когда расстояние между точками (t_0, x_0) и (t_1, x_1) равно l. В этом случае задача имеет только одну подходящую функцию – линейную (для остальных длина графика будет больше l). Если $\lambda_0 \neq 0$, то, полагая $\lambda_0 = 1$ и проведя элементарные преобразования, получаем

$$\frac{\lambda_1^2}{(x-c)^2} = 1 + \dot{x}^2\,.$$

Следовательно,

$$\frac{(x-c)\,dx}{\sqrt{\lambda_1^2 - (x-c)^2}} = d\,t\,.$$

Интегрируя, получаем $\sqrt{\lambda_1^2 - (x-c)^2} = -t + c_0$, где c_0 – некоторая константа. Таким образом, $(t-c_0)^2 + (x-c)^2 = \lambda_1^2$. Это – уравнение окружности с центром (c_0, c) и радиусом $|\lambda_1|$.

Заметим, что мы не доказали, что дуга окружности является решением задачи. Мы лишь доказали, что она является единственной экстремалью.

8 Лекция

Задача с подвижными концами. Вариация интегрального функционала.

Задача с подвижными концами. Формула вариации интегрального функционала с подвижными концами.

Переходим теперь к задаче, обобщающей простейшую задачу, задачу Больца и изопериметрическую задачу.

Задачей с подвижными концами называется задача минимизации функционала $\mathcal{J}_0(t_0, t_1, x)$ на пространстве $\mathbb{R} \times \mathbb{R} \times C^1[t_0, t_1]$ при условиях $\mathcal{J}_k(t_0, t_1, x) = 0\,, k = 1, \ldots, m$, где

$$\mathcal{J}_k(t_0, t_1, x) = \int_{t_0}^{t_1} f_k\left(t, x, \dot{x}\right) dt + \psi_k(t_0, t_1, x_0, x_1) \tag{8.17}$$

$f_k \in C(\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n)$, $\psi_k \in C(\mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n)$. При этом $t_0, t_1 \in \mathbb{R}$ – любые числа, предполагается только, что $t_1 > t_0$, $x(t) \in C^1\big([t_0, t_1], \mathbb{R}^n\big)$, $x_0 = x(t_0), x_1 = x(t_1)$. Так как в этой задаче ищется экстремум не только по всем функциям $x \in x(t) \in C^1\big([t_0, t_1], \mathbb{R}^n\big)$, но и по концам отрезка t_0, t_1, то мы ищем не оптимальную функцию $\hat{x}$, а оптимальную тройку $(\hat{x}, \hat{t}_0, \hat{t}_1)$. Тройка доставляет слабый локальный минимум в данной задаче, если существует $\varepsilon > 0$ такое, что $\mathcal{J}_0(t_0,t_1,x) \geq \mathcal{J}_0(\hat{t}_0,\hat{t}_1,\hat{x})$ как только $\mathcal{J}_k(t_0,t_1,x) = 0$, $k = 1,\ldots,m$, $\|x-\hat{x}\|_{C^1([t_0,t_1],\mathbb{R}^n)} < \varepsilon$ и $|t_0-\hat{t}_0| < \varepsilon$, $|t_1-\hat{t}_1| < \varepsilon$. Для любого набора множителей $\lambda = \lambda_0,\ldots,\lambda_m$ запишем лагранжиан $\mathcal{L}(\lambda,x,t_0,t_1) = \int_{t_0}^{t_1} L(t,x,\dot{x})\,dt + l(t_0,t_1,x_0,x_1)$, где $L(t,x,\dot{x}) = \sum_{k=0}^m \lambda_k\, f_k(t,x,\dot{x})$ – интегрант, а $l(t_0,t_1,x_0,x_1) = \sum_{k=0}^m \lambda_k\,\psi_k(t_0,t_1,x_0,x_1)$ – терминант.

Теорема 8.1 *Предположим, что в задаче с подвижными концами функции $f_k, (f_k)_x, (f_k)_{\dot{x}}$ и частные производные функций ψ_k по всем аргументам непрерывны, $k = 0,\ldots,m$. Если тройка $(\hat{x},\hat{t}_0,\hat{t}_1)$ доставляет слабый локальный минимум, то существуют НЕРОН множители $\lambda_0,\ldots,\lambda_m$ такие, что при $x = \hat{x}, t_0 = \hat{t}_0, \hat{t}_1$ выполнены следующие условия:*

1) уравнение Эйлера-Лагранжа:

$$\frac{d}{dt}\hat{L}_{\dot{x}} \;=\; \hat{L}_x \quad\Leftrightarrow\quad \sum_{k=0}^m \lambda_k\Big(\frac{d}{dt}(f_k)_{\dot{x}} - (f_k)_x\Big) \;=\; 0\,;$$

2) условие трансверсальности:

$$\hat{L}_{\dot{x}}(t_i) \;=\; (-1)^i\,\hat{l}_{x_i}\,,\ i = 0,1 \quad\Leftrightarrow\quad \sum_{k=0}^m \lambda_k\big[\,(f_k)_{\dot{x}}\big](\hat{t}_i) \;=\; (-1)^i\sum_{k=0}^m \lambda_k\,(\psi_k)_{x_i} \qquad i = 0,1;$$

3) условие стационарности на подвижные концы (выписываются только для подвижных концов):

$$\mathcal{L}_{t_i} \;=\; 0 \quad\Leftrightarrow\quad L\big(\hat{t}_i,\hat{x}(t_i),\dot{\hat{x}}(t_i)\big) \;=\; (-1)^i\sum_{k=0}^m \lambda_k\Big((\psi_k)_{t_i} + (\psi_k)_{x_i}\,\dot{\hat{x}}_i\Big), \quad i \;=\; 0,1\,.$$

`Доказательство` для случая $n = 1$. Так как функции f_k и ψ_k имеют непрерывные частные производные, то функционал $\mathcal{J}_k(t_0,t_1,x)$ дифференцируем:

$$\delta_{\mathcal{J}_k}\big((t_0,t_1,x)\,,\,(s_0,s_1,h)\big) \;=\; \int_{t_0}^{t_1}\Big((f_k)_x\,h + (f_k)_{\dot{x}}\,\dot{h}\Big)\,dt \;+\; (\psi_k)_{x_0}\,h(t_0) \;+\; (\psi_k)_{x_1}\,h(t_1) \;+$$

$$\Big(-f_k\big(t_0,\hat{x}(t_0),\dot{\hat{x}}(t_0)\big) + (\psi_k)_{t_0} + (\psi_k)_{x_0}\,\dot{\hat{x}}_0\Big)\,s_0 \;+\; \Big(f_k\big(t_1,\hat{x}(t_1),\dot{\hat{x}}(t_1)\big) + (\psi_k)_{t_1} + (\psi_k)_{x_1}\,\dot{\hat{x}}_1\Big)\,s_1$$

при $k = 0,\ldots,m$, $x,h \in C^1[t_0,t_1]$, $s_0,s_1 \in \mathbb{R}$. Воспользовавшись теоремой Лагранжа, получаем, что существуют НЕРОН множители $\lambda_0,\ldots,\lambda_m$, для которых при $x = \hat{x}, t_0 = \hat{t}_0, t_1 = \hat{t}_1$ имеем $\sum_{k=0}^m \lambda_k\,\mathcal{J}_k'(\hat{t}_0,\hat{t}_1,\hat{x}) = 0$. Сначала рассмотрим вариации с фиксированными концами $t_0 = \hat{t}_0, t_1 = \hat{t}_1$, т.е., $s_0 = s_1 = 0$. В этом случае получаем

$$\int_{t_0}^{t_1}\Big(\hat{L}_x\,h + \hat{L}_{\dot{x}}\,\dot{h}\Big)\,dt \;+\; \hat{l}_{x_0}\,h(t_0) \;+\; \hat{l}_{x_1}\,h(t_1) \quad=\quad 0$$

для всех $h \in C^1[t_0,t_1]$. Таким образом, ситуация дословно совпадает с разобранной ранее задачей Больца. Так же как в задаче Больца, получаем, что выполнены уравнения Эйлера-Лагранжа и условия трансверсальности. Теперь рассматриваем произвольные вариации, с произвольными s_0,s_1. Имеем $0 = \sum_{k=0}^m \lambda_k\,\mathcal{J}_k'(\hat{t}_0,\hat{t}_1,\hat{x})[s_0,s_1,h] = \mathcal{L}_{t_0}s_0 + \mathcal{L}_{t_1}s_1$ при любых s_0,s_1 (производные – по Гато). Следовательно, $\mathcal{L}_{t_0} = \mathcal{L}_{t_1} = 0$, что завершает доказательство.

□

Мы изучили несколько типов задач вариационного исчисления – простейшую задачу, задачу Больца, изопериметрическую задачу и задачу с подвижными концами. Во всех задачах мы проходили лишь необходимые условия локального минимума. Каковы могут быть достаточные условия ? Мы знаем пока условия только для выпуклых задач. По опыту работы с конечномерными задачами мы знаем, что достаточные условия локального минимума могут формулироваться в терминах второй производной функции. Это – условия второго порядка. Подобные условия существуют и у задач вариационного исчисления. Мы сформулируем и докажем их только для простейшей задачи. Но и это мы сделаем не сразу, а после предварительной работы. Вначале мы докажем формулу вариации функционала с подвижными концами, затем определим понятие сильного минимума, после чего докажем теорему Вейерштрасса-Эрдмана.

Рассмотрим интегральный функционал в простейшей задаче. Предположим, что функция x непрерывно зависит от некоторого параметра α, пробегающего интервал $(-\alpha_0, \alpha_0)$. Таким образом, имеем систему функций $x(t,\alpha)$, каждая из них – из отрезка $[t_0(\alpha), t_1(\alpha)]$. Производные по переменной t будем, как всегда, обозначать точкой, а по параметру α – штрихом. Будем предполагать, что функции $x, \dot x, x_\alpha, \dot x_\alpha$ непрерывны, а функции $t_0(\alpha), t_1(\alpha)$ непрерывно-дифференцируемы. Положим

$$\mathcal{J}(\alpha) \quad = \quad \int_{t_0(\alpha)}^{t_1(\alpha)} L\big(t, x(t,\alpha), \dot x(t,\alpha)\big)\, dt \tag{8.18}$$

Предполагаем также, что $L, L_x, L_{\dot x}$ непрерывны. Положим $x_i(\alpha) = x(t_i(\alpha), \alpha)\,,\ i = 0,1$. Обозначим через $\hat x, \hat t_i, \hat x_i$ соответствующие функции и величины при $\alpha = 0$. Положим, наконец, $h(t) = x_\alpha(t,0)$. Таким образом, $\dot h(t) = \dot x_\alpha(t,0)$. Предположим, что $\hat x$ – экстремаль. Имеем

$$\mathcal{J}'(0) \;=\; \int_{\hat t_0}^{\hat t_1} \Big(\hat L_{\,x}(t) h(t) \;+\; \hat L_{\,\dot x}(t)\dot h(t) \Big)\, dt \;+\; \hat L(\hat t_1)t_1'(0) \;-\; \hat L(\hat t_0)t_0'(0).$$

Проинтегрировав по частям и воспользовавшись уравнением Эйлера-Лагранжа, приводим данное выражение к виду

$$\hat L_{\dot x}(\hat t_1)h(\hat t_1) \;-\; \hat L_{\dot x}(\hat t_0)h(\hat t_0) \;+\; \hat L(\hat t_1)t_1'(0) \;-\; \hat L(\hat t_0)t_0'(0)\,.$$

Заметим, что $x_i'(0) \;=\; \frac{d}{d\alpha}\, x(t_i(\alpha),\alpha)\,\big|_{\alpha=0} = \; \dot{\hat x}(\hat t_i)t_i'(0) \;+\; h(\hat t_i)\,,\ i = 0,1$. Таким образом, $h(\hat t_i) \;=\; x_i'(0) \;-\; \dot{\hat x}(\hat t_i)t_i'(0)$. Следовательно,

$$\mathcal{J}'(0) \;=\; \hat L_{\dot x}(\hat t_1)x_1'(0) - \hat L_{\dot x}(\hat t_1)\dot{\hat x}(\hat t_1)t_1'(0) - \hat L_{\dot x}(\hat t_0)x_0'(0) + \hat L_{\dot x}(\hat t_0)\dot{\hat x}(\hat t_0)t_0'(0) + \hat L(\hat t_1)t_1'(0) - \hat L(\hat t_0)t_0'(0) \;=$$

$$=\; \hat L_{\dot x}(\hat t_1)x_1'(0) \;-\; \hat L_{\dot x}(\hat t_0)x_0'(0) \;-\; \Big(\hat L_{\dot x}(\hat t_1)\dot{\hat x}(\hat t_1) \;-\; \hat L(\hat t_1) \Big)\, t_1'(0) \;+\; \Big(\hat L_{\dot x}(\hat t_0)\dot{\hat x}(\hat t_0) \;-\; \hat L(\hat t_0) \Big)\, t_0'(0) \;=$$

$$=\; \hat p(\hat t_i)x_i'(0)\,\Big|_{i=0}^{1} \;-\; \hat H(\hat t_i)t_i'(0)\,\Big|_{i=0}^{1}$$

Итак, мы доказали следующее утверждение о вариации интегрального функционала:

Теорема 8.2 *Если $\hat x \;=\; x(t,0)$ – экстремаль интегрального функционала (8.18), то при указанных условиях гладкости имеем*

$$d\,\mathcal{J} \;=\; \hat p\; dx\,\Big|_{\hat t_0}^{\hat t_1} \;-\; \hat H\, dt\,\Big|_{\hat t_0}^{\hat t_1}$$

9 Лекция

Сильный и слабый минимум. Теорема Вейерштрасса-Эрдмана.

Пространство кусочно непрерывно дифференцируемых функций. Сильный и слабый минимум. Теорема Вейерштрасса-Эрдмана. Лемма о скруглении углов.

Мы возвращаемся к простейшей задаче вариационного исчисления. Наша цель – получение условий второго порядка (необходимых и достаточных) для локального минимума. Для этого нужно перейти к более широкому пространству функций.

Определение 9.1 *Класс функций $PC^1\big([t_0,t_1]\,,\mathbb{R}^n\big)$ состоит из непрерывных на отрезке $[t_0,t_1]$ функций, имеющих непрерывную производную $\dot{x}(t)$ в каждой точке отрезка, за исключением, быть может конечного числа точек, в которых, равно как и в концах отрезка, производная имеет односторонние пределы.*

Таким образом, PC^1 – это класс кусочно непрерывно-дифференцируемых функций (*piecewise-continuous*, отсюда и обозначение). Функции этого класса непрерывны, а их производные имеют лишь конечное число разрывов первого рода. В частности, $C^1\subset PC^1\subset C$. Так, функция $x(t)=|t|$ принадлежит $PC^1[-1,1]$, но не $C^1[-1,1]$, а функция $x(t)=\sqrt{1-t^2}$ принадлежит $C[-1,1]$, но не $PC^1[-1,1]$, поскольку её производная стремится к ∞ в концах отрезка. Заметим, что PC^1 – линейное нормированное пространство (норма – такая же, как в пространстве C), но это пространство не полно (не любая фундаментальная последовательность имеет предел), т.е., не является банаховым, в отличие от пространств C или C^1. Рассмотрим теперь классическую простейшую задачу вариационного исчисления, в которой мы расширим класс допустимых функций. Теперь мы будем искать экстремум не в пространстве C^1, а в классе PC^1. Таким образом, допустимой является любая функция $x\in PC^1[t_0,t_1]$ такая, что $x(t_0)=x_0, x(t_1)=x_1$.

Определение 9.2 *Функция $\hat{x}\in PC^1$ доставляет сильный минимум в простейшей задаче вариационного исчисления, если она допустима и существует $\varepsilon>0$ такой, что $\mathcal{J}(x)\geq\mathcal{J}(\hat{x})$ для любой допустимой функции $x\in PC^1$, удовлетворяющей условию $\big\|\,x-\hat{x}\big\|_C<\varepsilon$.*

Предложение 9.3 *Если функция $\hat{x}\in C^1[t_0,t_1]$ доставляет сильный экстремум, то она доставляет и слабый.*

Доказательство. Так как $\|\cdot\|_{C^1}\geq\|\cdot\|_C$, то если неравенство $\mathcal{J}(x)\geq\mathcal{J}(\hat{x})$ выполнено для любой допустимой PC^1-функции, для которой $\big\|\,x-\hat{x}\big\|_C<\varepsilon$, то оно выполнено и для любой C^1-функции, для которой $\big\|\,x-\hat{x}\big\|_{C^1}<\varepsilon$.

□

Обратное, вообще говоря не верно, как показывает следующий пример.

Пример 9.4 *(Слабый экстремум, не являющийся сильным).* В простейшей задаче $\mathcal{J}(x)=\int_0^1\dot{x}^3\,dt\to\min\,,\ x(0)=0,x(1)=1$ уравнение Эйлера-Лагранжа приводит к условию $L_{\dot{x}}=3\dot{x}^2=\text{const}$, откуда $\hat{x}(t)=t$. Имеем $\mathcal{J}(\hat{x}+h)-\mathcal{J}(\hat{x})=\int_0^1\dot{h}^2\,(3+\dot{h})\,dt$. Последнее выражение неотрицательно, если $\|h\|_{C^1} < 3$, поэтому экстремаль доставляет слабый локальный минимум $\hat{x}$. Но этот минимум не является сильным. В самом деле, положим $\dot{h}(t) = -n$ при $t \in \left[0, \frac{1}{2n^2}\right]$, $\dot{h}(t) = \frac{1}{n}$ при $t \in \left[\frac{1}{2}, 1\right]$ и $\dot{h}(t) = 0$ в остальных точках отрезка $[0,1]$. Так как $\int_0^1 \dot{h}(\tau)\,d\tau = 0$, то вариация $h(t) = \int_0^t \dot{h}(\tau)\,d\tau$ является допустимой ($h \in PC^1$ и $h(0) = h(1) = 0$) при каждом n. Кроме того, $\left\|h\right\|_C = \frac{1}{2n}$, а значит $\left\|h\right\|_C \to 0$ при $n \to \infty$. Однако, $\mathcal{J}(\hat{x} + h) - \mathcal{J}(\hat{x}) \to -\infty$ при $n \to \infty$, поэтому $\hat{x}$ не даёт сильного локального минимума.

Одно из преимуществ сильного минимума в сравнении со слабым состоит в том, что сильный минимум остаётся таковым при сужении задачи с отрезка $[t_0, t_1]$ на любой подотрезок $[s_0, s_1] \subset [t_0, t_1]$.

Предложение 9.5 *Если функция $\hat{x}$ доставляет сильный минимум в простейшей задаче, то для любого отрезка $[s_0, s_1] \subset [t_0, t_1]$ та же функция доставляет сильный минимум в задаче $\mathcal{J}_s(x) = \int_{s_0}^{s_1} x(t)\,dt \to \min\,,\ x(s_0) = \hat{x}(s_0)\,,\, x(s_1) = \hat{x}(s_1)$.*

Таким образом, функция $\hat{x}$, экстремальная в сильном смысле на отрезке $[t_0, t_1]$, экстремальна локально на любом его подотрезке.

`Доказательство.` Если $\hat{x}$ не является сильным минимумом для задачи на отрезке $[s_0, s_1]$, то найдётся такое $\varepsilon > 0$, что существуют допустимые функции $x \in PC^1[s_0, s_1]$, сколь угодно близкие к $\hat{x}$, для которых $\mathcal{J}_s(x) - \mathcal{J}_s(\hat{x}) \le -\varepsilon$. Продолжив каждую из функций x на весь отрезок $[t_0, t_1]$ значениями функции $\hat{x}$, получим допустимые функции $x \in PC^1[t_0, t_1]$, сколь угодно близкие к $\hat{x}$, для которых $\mathcal{J}(x) - \mathcal{J}(\hat{x}) \le -\varepsilon$, т.е., $\hat{x}$ не доставляет сильный минимум и в исходной задаче.

□

Заметим, что слабый минимум, вообще говоря, может не выдерживать локализации. Соответственно, и доказательство предложения 9.5 не проходит для слабых минимумов. Причина заключается в том, что при продолжении функции x с отрезка $[s_0, s_1]$ на весь отрезок $[t_0, t_1]$ могут появиться разрывы производной, т.е., функция выйдет из класса C^1.

Возникает естественный вопрос, как понимается уравнение Эйлера-Лагранжа для сильного минимума, а именно – в точках разрыва производной $\dot{x}$.

Предложение 9.6 *Если $\hat{x}$ – сильный экстремум, то в точках разрыва производной $\dot{\hat{x}}$ имеет место раздвоение уравнения Эйлера-Лагранжа:*

$$\tfrac{d}{dt}\,\hat{L}_{\dot{x}}\,(t-0) \;=\; L_x(t-0)\,,$$

$$\tfrac{d}{dt}\,\hat{L}_{\dot{x}}\,(t+0) \;=\; L_x(t+0)$$

`Доказательство.` Пусть t – точка разрыва. Считаем, что других точек разрыва производная не имеет, иначе перейдём к подотрезку. Для любого $s < t$ рассмотрим задачу на отрезке $[t_0, s]$. Согласно предложению 9.5 функция $\hat{x}$ доставляет сильный минимум на этом отрезке, а значит – и слабый минимум (предложение 9.3), поэтому она удовлетворяет уравнению Эйлера-Лагранжа. При $s \to t - 0$ получаем требуемое. Для предела справа рассуждение то же.

□

Поскольку у экстремали сильного минимума $\hat{x}(t)$ производная может терпеть разрывы, функция $\hat{L}(t) = L(t, \hat{x}, \dot{\hat{x}})$ также может иметь разрывы в тех же точках. Оказывается, однако, что функции импульса $\hat{p}(t) = \hat{L}_{\dot{x}}(t)$ и энергии $\hat{H}(t) = \dot{\hat{x}}\hat{L}_{\dot{x}}(t) - \hat{L}(t)$ всегда непрерывны

Теорема 9.7 *(Вейерштрасса-Эрдмана) Если функции L, L_x и $L_{\dot x}$ непрерывны, а функция $\hat{x} \in PC^1[t_0,t_1]$ доставляет сильный экстремум в простейшей задаче, то функции $\hat{p}(t)$ и $\hat{H}(t)$ непрерывны на $[t_0,t_1]$.*

Доказательство. В точках непрерывности функции $\hat{x}$ доказывать нечего. При доказательстве того, что $\hat{H} \in C[t_0,t_1]$ будем предполагать, что $\hat{x}(t)$ принадлежит C^2 в окрестности рассматриваемой точки. В силу предложения 9.5 достаточно рассмотреть случай единственной точки τ разрыва производной (иначе перейдём к подотрезку). Рассмотрим два семейства функций, зависящее от параметра α: $x^{(l)}(t,\alpha)$ – левое семейство и $x^{(r)}(t,\alpha)$ – правое семейство. Пусть $t_0^{(l)}(\alpha) = t_0\,,\, t_1^{(l)}(\alpha) = \tau+\alpha\,,\, x_0^{(l)}(\alpha) = x_0\,,\, x_1^{(l)}(\alpha) = \hat{x}(\tau)\,,\, x^{(l)}(t,0) = \hat{x}(t)\,,\, t \in [t_0,\tau]$. Аналогично для правого семейства: $t_0^{(r)}(\alpha) = \tau+\alpha\,; t_1^{(r)}(\alpha) = t_1\,,\, x_0^{(r)}(\alpha) = \hat{x}(\tau)\,,\, x_1^{(r)}(\alpha) = x_1\,,\, x^{(r)}(t,0) = \hat{x}(t)\,,\, t \in [\tau,t_1]$. Потребуем, чтобы у функций обеих семейств были непрерывные частные и смешанные производные, а значит можно воспользоваться теоремой 8.2 о вариации интегрального функционала. Рассмотрим два интегральных функционала:

$$\mathcal{J}_{(l)}(\alpha) = \int_{t_0^{(l)}(\alpha)}^{t_1^{(l)}(\alpha)} L\big(\,t\,,\, x^{(l)}(t,\alpha)\,,\, \dot{x}^{(l)}(t,\alpha)\,\big)\,dt\,; \mathcal{J}_{(r)}(\alpha) = \int_{t_0^{(r)}(\alpha)}^{t_1^{(r)}(\alpha)} L\big(\,t\,,\, x^{(r)}(t,\alpha)\,,\, \dot{x}^{(r)}(t,\alpha)\,\big)\,dt\,.$$

Соответствующие абсциссы и ординаты совпадают, поэтому семейства левых и правых кривых склеиваются, и мы можем рассмотреть новое семейство функций:

$$x(t,\alpha) \quad = \quad \begin{cases} x^{(l)}(t,\alpha)\,, \ \text{при}\ t \le \tau+\alpha\,; \\ x^{(r)}(t,\alpha)\,, \ \text{при}\ t \ge \tau+\alpha\,. \end{cases}$$

При любом α функция $x(t,\alpha)$ принадлежит $PC^1[t_0,t_1]$ и $\|x(t,\alpha) - \hat{x}(t)\|_C \to 0$ при $\alpha \to 0$. Положим

$$F(\alpha) = \mathcal{J}\big(x(\,\cdot\,,\alpha)\big) = \mathcal{J}_{(l)}(\alpha) + \mathcal{J}_{(r)}(\alpha) \quad = \quad \int_{t_0}^{t_1} L\big(\,t\,,\, x(t,\alpha)\,,\, \dot{x}(t,\alpha)\,\big)\,dt$$

Если $\hat{x} = x(t,0)$ доставляет сильный минимум, то функция $F(\alpha)$ достигает минимума в точке $\alpha = 0$, поэтому $F'(0) = 0$, а значит $\mathcal{J}_{(l)}{}'(0) + \mathcal{J}_{(r)}{}'(0) = 0$. Применяем теорему 8.2 о вариации интегрального функционала. Имеем $d\,\mathcal{J} = \hat{p}\,d\,x\,\big|_{\hat{t}_0}^{\hat{t}_1} - \hat{H}\,d\,t\,\big|_{\hat{t}_0}^{\hat{t}_1}$. Заметим, что для функционала $\mathcal{J} = \mathcal{J}_{(l)}$ граничные значения $x(t_0,\alpha)$ и $x(t_1,\alpha)$ не зависят от α, поэтому $d\,x\,\big|_{\hat{t}_0}^{\hat{t}_1} = 0$. Далее, $d\,t_0 = 0\,,\, d\,t_1 = d\,\alpha$. Следовательно, $d\,\mathcal{J}_{(l)} = -\,\hat{H}(\tau-0)\,d\,\alpha$. Аналогично $d\,\mathcal{J}_{(r)} = \hat{H}(\tau+0)\,d\,\alpha$. Так как $\mathcal{J}'_{(l)}(0) + \mathcal{J}'_{(r)}(0) = 0$, то $-\,\hat{H}(\tau-0) + \hat{H}(\tau+0) = 0$. Таким образом, функция $\hat{H}(t)$ непрерывна в точке $t = \tau$. Для завершения доказательства осталось лишь предъявить левое и правое семейства функций с требуемыми свойствами. Вот они:

$$x^{(l)}(t,\alpha) = \hat{x}\Big(\,t - \alpha\,\frac{t-t_0}{\tau+\alpha-t_0}\Big)\,; \quad x^{(r)}(t,\alpha) = \hat{x}\Big(\,t - \alpha\,\frac{t-t_1}{\tau+\alpha-t_1}\Big)\,.$$

Для условий гладкости необходимо, чтобы $x \in C^2$ в окрестности точки τ.

Теперь докажем, что $\hat{p}(\tau-0) = \hat{p}(\tau+0)$. Снова построим два семейства кривых. Для левого семейства $t_0^{(l)}(\alpha) = t_0\,,\, t_1^{(l)}(\alpha) = \tau\,,\, x_0^{(l)}(\alpha) = x_0\,,\, x_1^{(l)}(\alpha) = \hat{x}(\tau) + \alpha\,,\, x^{(l)}(t,0) = \hat{x}(t)\,,\, t \in [t_0,\tau]$. Аналогично для правого семейства: $t_0^{(r)}(\alpha) = \tau\,; t_1^{(r)}(\alpha) = t_1\,,\, x_0^{(r)}(\alpha) =$

$\hat{x}(\tau) + \alpha\,,\, x_1^{(r)}(\alpha) \;=\; x_1\,,\, x^{(r)}(t,0) \;=\; \hat{x}(t)\,,\, t \in [\tau, t_1]$. Вновь воспользовавшись вариацией интегрального функционала, получаем

$$0 \;=\; F'(0) \;=\; \hat{p}(\tau - 0) \;-\; \hat{p}(\tau + 0),$$

откуда $\hat{p}(\tau - 0) \;=\; \hat{p}(\tau + 0)$. В явном виде эти семейства могут быть заданы так:

$$x^{(l)}(t) \;=\; \hat{x}(t) \;+\; \alpha\,\frac{t - t_0}{\tau - t_0}\,; \quad x^{(r)}(t) \;=\; \hat{x}(t) \;+\; \alpha\,\frac{t - t_1}{\tau - t_1}$$

□

Мы завершаем эту лекцию ответом на следующий, вполне естественный, вопрос: почему мы определили сильный локальный минимум, но не определили сильный глобальный минимум, т.е., минимум в пространстве PC^1 ? Ответ: он всегда совпадает с глобальным минимумом в C^1.

Лемма 9.8 *(о скруглении углов). В простейшей задаче при любом $\varepsilon > 0$ для любой допустимой функции $x \in PC^1[t_0, t_1]$ существует допустимая функция $\tilde{x} \in C^1[t_0, t_1]$ такая, что $\big\|x - \tilde{x}\big\|_{C[t_0,t_1]} < \varepsilon$ и $\big|\mathcal{J}(\tilde{x}) \,-\, \mathcal{J}(x)\big| < \varepsilon$.*

Таким образом, можно так “скруглить углы” у функции x (сгладить её в точках разрыва производной), что функция останется допустимой (её значения в точках t_0 и t_1 не изменятся), а значение функционала $\mathcal{J}$ изменится мало.

`Доказательство.` Возьмем константу M, превосходящую величины $\max_{t\in[t_0,t_1]} |x(t)|$ и $\sup_{t\in[t_0,t_1]} |\dot{x}(t)|$ (второй супремум взят по тем точкам t, в которых производная существует). Каждую точку разрыва производной $\tau \in (t_0, t_1)$ окружим отрезком $[\tau - \delta, \tau + \delta]$, при этом δ выберем настолько малым, чтобы эти отрезки не пересекались. Теперь изменим функцию x на каждом из этих отрезков, а в остальных точках оставим неизменной, так получим функцию $\tilde{x}$. Сделаем мы это так. На отрезке $[\tau - \delta, \tau + \delta]$ возьмём непрерывную функцию $y(t)$, которая на концах этого отрезка совпадает с $\dot{x}(t)$, в каждой точке отрезка не превосходит по модулю $2M$, и кроме того $\int_{\tau-\delta}^{\tau+\delta} y(t)\, dt \;=\; \int_{\tau-\delta}^{\tau+\delta} \dot{x}(t)\, dt$ (доказательство существования такой функции будет лёгким упражнением для читателя). Положим $\tilde{x}(t) \,=\, x(\tau - \delta) \,+\, \int_{\tau-\delta}^{t} y(s)\, ds$. Ясно, что $\tilde{x}(\tau \pm \delta) \,=\, x(\tau \pm \delta)$. Так поступим с каждым отрезком, окружающим точки разрыва производной $\dot{x}$. Получим функцию $\tilde{x} \in C^1[t_0, t_1]$. Тогда $\big|\dot{\tilde{x}}(t) \,-\, \dot{x}(t)\big| \,\le\, 3M$ в каждой точке t, и поэтому $\big|\tilde{x}(t) \,-\, x(t)\big| \,\le\, 6M\delta$ в каждой точке t. Таким образом, $\big\|\tilde{x} \,-\, x\big\|_{C[t_0,t_1]} \,\le\, 6M\delta$. Выбором δ можно сделать эту разность сколь угодно малой. Далее, $|\dot{\tilde{x}}(t)| \,\le\, 2M$ в любой точке t. Поэтому $\big|\, L\big(t, \tilde{x}(t), \dot{\tilde{x}}(t)\big) \,\big|$ не превосходит максимума функции $|L|$ на множестве $[t_0, t_1] \times [-2M, 2M] \times [-2M, 2M]$. Этот максимум достигается в силу компактности множества и непрерывности L. Обозначим его через N. Тогда $\big|\, \int_{\tau-\delta}^{\tau+\delta} L(t, \tilde{x}, \dot{\tilde{x}})\, dt - \int_{\tau-\delta}^{\tau+\delta} L(t, x, \dot{x})\, dt \,\big| \,\le\, 2\delta N$. Следовательно, $\big|\mathcal{J}(\tilde{x}) \,-\, \mathcal{J}(x)\big| \,\le\, 2\delta N k$, где k – число точек разрыва производной функции x. Выбирая δ достаточно малым, получаем $2\delta N k < \varepsilon$, что завершает доказательство.

□

Следствие 9.9 *Для простейшей задачи* $\inf\limits_{x\,\in\, PC^1[t_0,t_1]} \mathcal{J}(x) \;=\; \inf\limits_{x\,\in\, C^1[t_0,t_1]} \mathcal{J}(x)$*, где оба инфимума берутся по допустимым функциям, т.е., удовлетворяющим граничным условиям $x(t_0) = x_0, x(t_1) = x_1$.*

Замечание 9.10 Таким образом, значение абсолютного минимума в простейшей задаче, в отличие от локального минимума (пример 9.4), всегда одинаково в пространствах C^1 и PC^1. Возникает вопрос, зачем в таком случае расширять пространство функций и вводить понятие сильного минимума ? Дело в том, что во многих задачах сильный минимум достигается (а значит, может быть найден с помощью уравнения Эйлера-Лагранжа), а слабый – нет.

10 Лекция

Вторая производная по Фреше. Условие Лежандра.

Вторая производная по Фреше. Вторая вариация интегрального функционала.

Условие Лежандра – необходимое условие второго порядка на слабый минимум.

Для получения условий второго порядка в простейшей задаче, мы вначале определим вторую производную по Фреше. Пусть X, Y – нормированные пространства, $F : X \to Y$ – некоторое отображение. Обозначим также через $\mathcal{L}(X,Y)$ пространство линейных непрерывных операторов, действующих из X в Y.

Определение 10.1 *Отображение $F : X \to Y$ является дважды дифференцируемым в точке x, если оно дифференцируемо в окрестности этой точки и отображение $x \to F'(x)$ (из пространства X в $\mathcal{L}(X,Y)$) является дифференцируемым. Производная этого отображения называется второй производной функции F в точке x.*

По определению, таким образом, имеем

$$F'(x+h_1)[h_2] \;=\; F'(x)[h_2] \;+\; \big(F''(x)[h_1]\big)[h_2] \;+\; o(\|h_1\| \cdot \|h_2\|)\,.$$

Если функция дважды дифференцируема, то $\big(F''(x)[h_1]\big)[h_2] \;=\; \big(F''(x)[h_2]\big)[h_1]$ для любых $h_1, h_2 \in X$. Мы оставим этот факт без доказательства. Таким образом, вторая производная – это непрерывная симметричная билинейная форма на X: $F''[h_1,h_2] \;=\; \big(F''(x)[h_1]\big)[h_2]$. Следующее предложение мы также не будем доказывать.

Предложение 10.2 *Если отображение F дважды дифференцируемо в точке x, то* $F(x+h) \,=\, F(x) \,+\, F'(x)[h] \,+\, \frac{1}{2}\,F''(x)[h,h] \,+\, o(\|h\|^2)$.

Симметрическая билинейная форма $Q : X \times X \to \mathbb{R}$ и соответствующая ей квадратичная форма $Q(x,x)$ называется неотрицательно определённой (в других терминах – положительно полуопределённой, обозначение $Q \geq 0$), если $Q(x,x) \,\geq\, 0$ для всех $x \in X$. Если же $Q(x,x) \,>\, 0$ для всех $x \in X, x \neq 0$, то форма называется положительно определённой (обозначение $Q > 0$).

Теорема 10.3 *Если $\hat{x}$ – точка локального минимума функции $F : X \to \mathbb{R}$, и в этой точке F дважды дифференцируема, то $F'(\hat{x}) = 0$ и $F''(\hat{x}) \geq 0$. Обратно, если $F'(\hat{x}) = 0$ и $F''(\hat{x})[h,h] \geq \alpha \|h\|^2$, то точка $\hat{x}$ доставляет локальный минимум.*

Доказательство. Если $\hat{x} \in \mathrm{locmin}$, то $F'(\hat{x}) = 0$. Если при этом $F''(\hat{x})[h,h] < 0$ для некоторого $h \in X$, то $F''(\hat{x})[th, th] = t^2 F''(\hat{x})[h,h] < 0$ при любом $t \neq 0$. При $t \to 0$ применяем предложение 10.2 и получаем, что $F(\hat{x} + th) < F(\hat{x})$ при малых t. Доказательство достаточности немедленно вытекает из предложения 10.2.

□

Замечание 10.4 В пограничном случае, когда $F'(\hat{x}) = 0$ и $F''(\hat{x}) \geq 0$, т.е., когда необходимые условия выполнены, а достаточные – нет, могут быть разные ситуации. Так, для функции $F : \mathbb{R} \to \mathbb{R}$, $F(x) = x^3$ точка $\hat{x} = 0$ не даёт локального минимума.

Рассмотрим теперь простейшую задачу и соответствующий функционал $\mathcal{J}(x) = \int_{t_0}^{t_1} L(t, x, \dot{x})\, dt$. Предполагая, что вторые частные производные $L_{x\,x}$, $L_{x\,\dot{x}}$, $L_{\dot{x}\,\dot{x}}$ существуют и непрерывны, получаем

$$\mathcal{J}''(x)[h,h] \quad = \quad \int_{t_0}^{t_1} \Bigl(L_{x\,x}\, h^2 \; + \; 2\, L_{x\,\dot{x}}\, h\, \dot{h} \; + \; L_{\dot{x}\,\dot{x}}\, \dot{h}^2 \Bigr)\, dt\,. \tag{10.19}$$

Наша первая задача – получить необходимые условия неотрицательной определённости данной квадратичной формы в пространстве $C^1_0[t_0, t_1]$. Они, в силу теоремы 10.3, будут необходимыми условиями слабого локального минимума в простейшей задаче.

Определение 10.5 *Экстремаль $\hat{x}$ в простейшей задаче удовлетворяет условию Лежандра (усиленному условию Лежандра), если $\hat{L}_{\dot{x}\,\dot{x}}(t) \geq 0$ $(\hat{L}_{\dot{x}\,\dot{x}}(t) > 0)$ для всех $t \in [t_0, t_1]$.*

Адриен Мари Лежандр (1752-1833) – французский математик.

Теорема 10.6 *Если в простейшей задаче $\mathcal{J}''(\hat{x}) \geq 0$ на пространстве $C^1_0[t_0, t_1]$, то выполнено условие Лежандра.*

Доказательство. От противного, пусть $\hat{L}_{\dot{x}\,\dot{x}}(\tau) < 0$ при некотором $\tau \in [t_0, t_1]$. В силу непрерывности можно считать, что $\tau \in [t_0 + \delta, t_1 - \delta]$ и что $\hat{L}_{\dot{x}\,\dot{x}}(t) < -\varepsilon$ на отрезке $[\tau - \delta, \tau + \delta]$. Положим

$$\bar{h}(t) \quad = \quad \begin{cases} \delta \, - \, |\, t \, - \, \tau\,| & \text{при} \quad |\, t \, - \, \tau\,| \, < \, \delta\,; \\ 0 & \text{при} \quad |\, t \, - \, \tau\,| \, \geq \, \delta\,. \end{cases}$$

Тогда

$$\mathcal{J}''(\hat{x})[\bar{h}, \bar{h}] \quad = \quad \int_{\tau-\delta}^{\tau+\delta} \Bigl(L_{x\,x}\, \bar{h}^2 \; + \; 2\, L_{x\,\dot{x}}\, \bar{h}\, \dot{\bar{h}} \; + \; L_{\dot{x}\,\dot{x}}\, \dot{\bar{h}}^2 \Bigr)\, dt$$

В подынтегральной сумме первое слагаемое равно $O(\delta^2)$, второе – $O(\delta)$, а третье по модулю не меньше ε. При малых δ данный интеграл отрицателен. Это еще не конец доказательства, поскольку $\bar{h} \notin C^1_0[t_0, t_1]$. Воспользовавшись теперь леммой о скруглении углов (лемма 9.8), получаем, что $\inf_{h \in C^1_0[t_0,t_1]} \mathcal{J}''(\hat{x})[h,h]\, dt \; < \; 0\,.$

□

Следствие 10.7 *Если в простейшей задаче вторые производные $L_{x\,x}$, $L_{x\,\dot{x}}$, $L_{\dot{x}\,\dot{x}}$ существуют и непрерывны, а $\hat{x}$ – слабый локальный минимум, то на нём выполнено условие Лежандра.*

11 Лекция

Условие Якоби. Необходимые и достаточные условия второго порядка.

Уравнение Якоби. Сопряжённые точки. Условие Якоби.

Достаточные условия второго порядка на слабый минимум. Случай квадратичного функционала.

Итак, мы получили одно условие второго порядка на слабый минимум – условие Лежандра. Для формулировки дальнейших условий мы вначале повторим следующее (ошибочное) рассуждение принадлежащее Лагранжу. В нём "доказывается", что усиленное условие Лежандра *достаточно* для неотрицательной определённости второй производной.

Заметим вначале, что $\int_{t_0}^{t_1} 2B\,h\,\dot h\,dt \;=\; -\int_{t_0}^{t_1} \dot B\,h^2\,dt$ (проинтегрировали по частям) для любой функции $B \in C^1[t_0,t_1]$ и любой $h \in C_0^1[t_0,t_1]$. Проделав это со вторым слагаемым в интеграле $\mathcal{J}''(x)[h,h] \;=\; \int_{t_0}^{t_1}\Big(L_{x\,x}\,h^2 \,+\, 2\,L_{x\,\dot x}\,h\,\dot h \,+\, L_{\dot x\,\dot x}\,\dot h^2\Big)\,dt$, получим

$$\mathcal{J}''(\hat x)[h,h] \quad = \quad \int_{t_0}^{t_1} p(t)\,\dot h^2(t) \,+\, q(t)\,h^2(t)\,dt\,,$$

где p,q – некоторые функции. Пусть выполнено усиленное условие Лежандра: $p(t) > 0$ для всех $t \in [t_0,t_1]$. Прибавим под интегралом выражение $\frac{d}{dt}\big[\,\omega(t)\,h^2(t)\big]$. Оно не изменит значение интеграла, поскольку $h \in C_0^1[t_0,t_1]$. Подберем функцию $\omega(t)$ так, чтобы подынтегральное выражение $p\,\dot h^2 \,+\, 2\omega\,\dot h\,h \,+\, (\dot\omega + q)\,h^2$ стало полным квадратом. Это означает, что

$$\dot\omega \,+\, q \;=\; p^{-1}\,\omega^2. \tag{11.20}$$

Это – дифференциальное уравнение Рикатти. Решив это уравнение и получив функцию $\omega(t)$, имеем

$$\mathcal{J}''(\hat x)[h,h] \quad = \quad \int_{t_0}^{t_1}\Big(p^{1/2}\,\dot h \,+\, p^{-1/2}\,\omega\,h\Big)^2\,dt \;\ge\; 0. \tag{11.21}$$

Таким образом, усиленное условие Лежандра достаточно для неотрицательной определённости второй производной. Однако, это неверно, и вот соответствующий контрпример:

Пример 11.1 *(Гармонический осциллятор).* Пусть $\mathcal{J}(x) \;=\; \int_0^{2\pi}\big(\,\dot x^2 - x^2\,\big)\,dt$. Поскольку $L_{\dot x\,\dot x} \;=\; 2 > 0$, усиленное условие Лежандра выполнено. Тем не менее, вторая производная $\mathcal{J}''(x)[h] \;=\; 2\int_0^{2\pi}\big(\,\dot h^2 \,-\, h^2\,\big)\,dt$ равна $-\frac{3\pi}{2}$ при $h \;=\; \sin\frac{t}{2} \;\in\; C_0^1[0\,,\,2\pi]$. Следовательно, квадратичный функционал $\mathcal{J}$ не является неотрицательно-определённым, несмотря на то, что удовлетворяет усиленному условию Лежандра.

В чём же дело ? Ответ: уравнение Рикатти (11.20) не обязано иметь решение $\omega(t)$, непрерывное на всём отрезке. Решение существует лишь локально, и при продолжении может уйти на бесконечность за конечное время.

Пример 11.2 Для функционала из предыдущего примера $\mathcal{J}'' \;=\; 2\int_0^{2\pi}\big(\,\dot h^2 \,-\, h^2\,\big)\,dt$, и уравнение Рикатти (11.20) имеет вид $\dot\omega \,-\, 1 \;=\; \omega^2$. Отсюда $\operatorname{arctg}\omega \;=\; t + C$, т.е. $\omega \;=\; \operatorname{tg}(t + C)$. Поскольку длина отрезка больше π, функция $\omega(t)$ обязательно имеет на нём особую точку.

Тем не менее, рассуждения Лагранжа можно исправить, что и было сделано впоследствии Карлом Густавом Якоби (1804-1851), немецким математиком, родным братом известного физика. Для дальнейших рассуждений мы перейдём к более общим обозначениям. Рассмотрим квадратичный функционал вида

$$\mathcal{K}(h) \,=\, \int_{t_0}^{t_1} \Big(Ah^2 \,+ 2\,B\,h\,\dot h \,+\, C\,\dot h^2 \Big)\, dt\,,$$

где $A, B, C \,\in\, C^1[t_0, t_1]$. В частности, тот же вид имеет $\mathcal{J}''$. *Уравнением Якоби* называется уравнение Эйлера-Лагранжа на этот функционал:

$$\frac{d}{dt}\Big(B\,h + C\,\dot h \Big) \quad = \quad A\,h \,+\, B\,\dot h\,. \tag{11.22}$$

Так, для второй производной $\mathcal{J}''$ интегрального функционала простейшей задачи уравнение Якоби выглядит так:

$$\frac{d}{dt}\Big(\hat L_{x\,\dot x}\,h + \hat L_{\dot x\,\dot x}\,\dot h \Big) \quad = \quad \hat L_{x\,x}\,h + \,\hat L_{x\,\dot x}\,\dot h\,. \tag{11.23}$$

Уравнение Якоби – однородное дифференциальное уравнение второго порядка. Если $h(t)$ – его нетривиальное (т.е., ненулевое) решение с начальным условием $h(t_0) = 0$, то и $\lambda\, h(t)$ – тоже решение. По теореме существования и единственности решения дифференциального уравнения, h – единственное, с точностью до умножения на константу, решение уравнение Якоби с начальным условием $h(t_0) = 0$.

Определение 11.3 *Точка $\tau \in (t_0, t_1]$ называется сопряжённой к точке t_0, если $h(\tau) = 0$, где h – решение уравнения Якоби.*

Рис. 4:

Будем говорить, что выполнено *условие Якоби* (*усиленное условие Якоби*), если интервал (t_0, t_1) (соответственно, полуинтервал $(t_0, t_1]$) не содержит сопряжённых точек.

Теорема 11.4 *Если $\mathcal{K} \,\ge\, 0$ на пространстве $C_0^1[t_0, t_1]$ и выполнено усиленное условие Лежандра ($C(t) > 0\,,\ t \in [t_0, t_1]$), то выполнено условие Якоби.*

Следствие 11.5 *Если экстремаль $\hat x$ доставляет слабый локальный минимум в простейшей задаче и выполнено усиленное условие Лежандра, то выполнено условие Якоби.*

`Доказательство теоремы 11.4.` От противного: пусть $\mathcal{K} \,\ge\, 0$ на пространстве $C_0^1[t_0, t_1]$, но решение уравнения Якоби $\hat h$ обращается в ноль в некоторой точке $\tau \in (t_0, t_1)$. Так как $\mathcal{K}(0) \,=\, 0$, то $\min\limits_{h \in C_0^1[t_0,t_1]} \mathcal{K}(h) \,=\, 0$, и по лемме о скруглении углов (лемма 9.8) имеем $\min\limits_{h \in PC_0^1[t_0,t_1]} \mathcal{K}(h) \,=\, 0$. рассмотрим простейшую задачу

$$\mathcal{K}_\tau(h) \,=\, \int_{t_0}^{\tau} \big(Ah^2 \,+ 2\,B\,h\,\dot h \,+\, C\,\dot h^2 \big)\, dt \,\to\, \min\,, \qquad h(t_0) \,=\, h(\tau) \,=\, 0.$$

Для любого λ функция $\lambda\hat{h}$ – решение уравнения Якоби. Положив $f(\lambda) = \mathcal{K}_\tau(\lambda\hat{h})$, получаем

$$f'(\lambda) \;=\; \lim_{\alpha\to 0} \frac{\mathcal{K}_\tau(\lambda\hat{h} + \alpha\hat{h}) - \mathcal{K}_\tau(\lambda\hat{h})}{\alpha} \;=\; \delta_{\mathcal{K}_\tau}\left(\lambda\hat{h}\,,\,\hat{h}\right) \;=\; 0\,,$$

так как $\lambda\hat{h}$ удовлетворяет уравнению Эйлера-Лагранжа на функционал $\mathcal{K}$. Итак, $f(\lambda) \equiv \mathrm{const}$ и $f(0) = 0$, откуда $f(\lambda) \equiv 0$, а значит $\mathcal{K}_\tau(\hat{h}) = 0$. Следовательно, для функции

$$\tilde{h}(t) \quad = \quad \begin{cases} \hat{h}\,(t) & \text{при} \quad t \in [t_0, \tau]; \\ 0 & \text{при} \quad t \in (\tau, t_1]\,. \end{cases}$$

имеем $\mathcal{K}(\tilde{h}) = 0$. Но так как $\min\limits_{h\in PC_0^1[t_0,t_1]} \mathcal{K}(h) = 0$, то функция $\tilde{h}$ доставляет сильный ло-

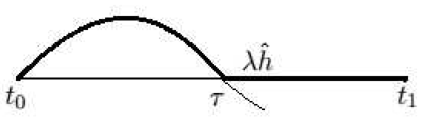

Рис. 5: Функция $\tilde{h}(t)$

кальный минимум функционала $\mathcal{K}$. Теперь воспользуемся теоремой Вейерштрасса-Эрдмана (теорема 9.7), причём только её частью – непрерывностью импульса на экстремали, дающей сильный экстремум. Обозначив через $\tilde{p}(t)$ импульс функционала $\mathcal{K}$ на экстремали $\tilde{h}$, получаем

$$\tilde{p}\,(\tau - 0) \;=\; \tilde{p}\,(\tau + 0)\,.$$

Так как $\tilde{p}(t) = 2B(t)\,\tilde{h}(t) + 2C(t)\,\dot{\tilde{h}}(t)$, то $\tilde{p}(\tau + 0) = 0$ (поскольку $\tilde{h}(\tau+0) = \dot{\tilde{h}}(\tau+0) = 0$). Значит $2B(\tau-0)\,\tilde{h}(\tau-0) + 2C(\tau-0)\,\dot{\tilde{h}}(\tau-0) = 0$. Далее, $\tilde{h}(\tau-0) = \tilde{h}(\tau) = 0$, откуда $2C(\tau)\,\dot{\tilde{h}}(\tau-0)$. По условию Лежандра $2C(\tau) > 0$, откуда $\dot{\tilde{h}}(\tau-0) = 0$, а значит $\dot{\hat{h}}(\tau) = 0$. Последнее невозможно в силу теоремы существования и единственности решения дифференциального уравнения (уравнение Якоби – второго порядка, и оно имеет только нулевое решение с граничными условиями $\dot{h}(\tau) = h(\tau) = 0$).

□

Итак, необходимым условием слабого минимума в простейшей задаче является выполнение условия Лежандра, а при выполнении усиленного условия Лежандра – условие Якоби. Перейдём теперь к достаточным условиям слабого минимума.

Теорема 11.6 *Если в простейшей задаче на экстремали $\hat{x}$ выполнены усиленные условия Лежандра и Якоби, то $\hat{x}$ доставляет слабый минимум.*

Для доказательства мы воспользуемся формулой (11.21), где функция $\omega(t)$ ищется из уравнения Рикатти (11.20). Для простейшей задачи формула выглядит так:

$$\mathcal{J}''(\hat{x})[h,h] \quad = \quad \int_{t_0}^{t_1} \left(C^{1/2}\,\dot{h} \,+\, C^{-1/2}\left(B + \omega\right) h \right)^2 dt\,, \tag{11.24}$$

а уравнение Рикатти:

$$\dot{\omega} \;=\; \frac{\left(B + \omega\right)^2}{C} \;-\; A\,, \tag{11.25}$$

где $A = \hat{L}_{x\,x}$, $B = \hat{L}_{x\,\dot{x}}$, $C = \hat{L}_{\dot{x}\,\dot{x}}$. Вначале установим два вспомогательных утверждения.

Лемма 11.7 *Если выполнены усиленные условия Лежандра и Якоби, то существует решение уравнения Якоби, положительное всюду на отрезке* $[t_0, t_1]$.

Доказательство. Как мы знаем, существует решение h_0 которое положительно на полуинтервале $(t_0, t_1]$ и $h_0(t_0) = 0$. С другой стороны, существует решение h_1 с начальными условиями $h_1(t_0) = 1, \dot{h}_1(t_0) = 0$. Из непрерывности этих решений следует что функция $h_0 + \alpha h_1$ положительна при любом малом $\alpha > 0$.

□

Лемма 11.8 *Если h – решение уравнения Якоби, положительное всюду на отрезке* $[t_0, t_1]$, *то* $\omega = -B - C\frac{\dot{h}}{h}$ *– решение уравнения Рикатти (11.25).*

Доказательство. Имеем

$$\omega h = -Bh - C\dot{h}.$$

Далее,

$$\frac{d}{dt}\big(\omega h\big) = \dot{\omega} h + \omega \dot{h}.$$

Из уравнения Якоби следует, что

$$\frac{d}{dt}\big(\omega h\big) = \frac{d}{dt}\big(-Bh - C\dot{h}\big) = -Ah - B\dot{h}$$

Приравнивая полученные выражения для $\frac{d}{dt}\big(\omega h\big)$, получаем

$$\big(\dot{\omega} + A\big) h = \big(-\omega - B\big)\dot{h}$$

Выражая из этого равенства $\dot{\omega}$ и подставляя $\frac{\dot{h}}{h} = -\frac{\omega+B}{C}$, получаем, что ω удовлетворяет уравнению Рикатти:

$$\dot{\omega} = -(\omega + B)\frac{\dot{h}}{h} - A = \frac{(\omega+B)^2}{C} - A.$$

Применив теперь формулу (11.24), получаем

Следствие 11.9 *Если выполнены усиленные условия Лежандра и Якоби, то* $\mathcal{J}'' \geq 0$ *на пространстве* $C_0^1[t_0, t_1]$.

Теперь мы в состоянии доказать теорему 11.6.

Доказательство теоремы 11.6. Разложим функцию $L(t, x, \dot{x})$ по формуле Тейлора до второго члена в точке $(t, \hat{x}, \dot{\hat{x}})$. Получим

$$\begin{aligned} L(t, \hat{x} + h, \dot{\hat{x}} + \dot{h}) &= L(t, \hat{x}, \dot{\hat{x}}) + \hat{L}_x(t)\, h(t) + \hat{L}_{\dot{x}}(t)\, \dot{h}(t) + \\ &+ \hat{L}_{xx}(t)\, h^2(t) + 2\hat{L}_{x\dot{x}}(t)\, h(t)\, \dot{h}(t) + \hat{L}_{\dot{x}\dot{x}}(t)\, \dot{h}^2(t) + o\big(h^2(t) + \dot{h}^2(t)\big), \end{aligned}$$

причем "о малое" – равномерно по t при $h^2(t) + \dot{h}^2(t) \to 0$. Следовательно, при малых значениях $\|h\|_{C^1[t_0,t_1]}$ имеем

$$\begin{aligned} L(t, \hat{x} + h, \dot{\hat{x}} + \dot{h}) &\geq L(t, \hat{x}, \dot{\hat{x}}) + \hat{L}_x(t)\, h(t) + \hat{L}_{\dot{x}}(t)\, \dot{h}(t) + \\ &+ \hat{L}_{xx}(t)\, h^2(t) + 2\hat{L}_{x\dot{x}}(t)\, h(t)\, \dot{h}(t) + \hat{L}_{\dot{x}\dot{x}}(t)\, \dot{h}^2(t) - \delta\big(h^2(t) + \dot{h}^2(t)\big), \end{aligned}$$

где $\delta > 0$ – сколь угодно малое число. Проинтегрировав данное равенство по отрезку $[t_0, t_1]$, и учитывая, что $\int_{t_0}^{t_1} \Big(\hat{L}_x(t)\, h(t) \;+\; \hat{L}_{\dot x}(t)\, \dot h(t) \Big)\, dt \;=\; 0$ (поскольку $\hat x$ удовлетворяет уравнению Эйлера-Лагранжа), а также обозначая $\hat{L}_{x\,x}(t) \;=\; A(t)\,,\; \hat{L}_{x\,\dot x}(t) \;=\; B(t)\,,\; \hat{L}_{\dot x\,\dot x}(t) \;=\; C(t)$, получаем

$$\mathcal{J}\left(\hat x + h\right) \;\geq\; \mathcal{J}\left(\hat x\right) \;+\; \int_{t_0}^{t_1} \Big(\big(\, A(t) \,-\, \delta \,\big)\, h^2(t) \;+\; 2\, B(t)\, h(t)\, \dot h(t) \;+\; \big(\, C(t) \,-\, \delta \,\big)\, \dot h^2(t) \Big)\, dt\,.$$

Если δ достаточно мало, то второе слагаемое (квадратичный функционал, который мы обозначим через $\mathcal{K}_\delta$) удовлетворяет усиленному условию Лежандра ($C(t) \,-\, \delta \,>\, 0$ при $t[t_0, t_1]$). Кроме того, в силу непрерывной зависимости решений линейного дифференциального уравнения от его коэффициентов, квадратичный функционал $\mathcal{K}_\delta$ при малых δ также удовлетворяет усиленному условию Якоби. Применяя леммы 11.7 и 11.8, получаем, что уравнение Рикатти, соответствующее функционалу $\mathcal{K}_\delta$, имеет непрерывное решение на отрезке $[t_0, t_1]$, из чего, пользуясь формулой (11.24), заключаем, что $\mathcal{K}_\delta \;\geq\; 0$.

□

В случае, когда функционал $\mathcal{J}$ является суммой квадратичного и линейного функционала, т.е., имеет вид

$$\mathcal{J}(x) \;=\; \mathcal{J}_{\text{К}}(x) + \mathcal{J}_{\text{Л}}(x) \;=\; \int_{t_0}^{t_1} \Big[\, A(t)x^2(t) + 2\, B(t)x(t)\dot x(t) + C(t)\, \dot x^2 \,\Big] + \Big[\, 2\, D(t)\, x(t) + 2\, E(t)\, \dot x(t) \,\Big]\, dt\,,$$

условия Лежандра и Якоби позволяют провести полный (вернее, почти полный) анализ простейшей задачи на абсолютный минимум. Обозначим через $S_{\min}$ наименьшее значение $\mathcal{J}(x)$ среди всех допустимых C^1-функций (а значит, в силу леммы о срезании углов, и среди PC^1-функций).

Теорема 11.10 *Если $\mathcal{J}(x)$ – "квадратичный плюс линейный" функционал, и $\hat x$ – экстремаль, то*

а) если не выполнено условие Лежандра, то $S_{\min} \;=\; -\infty$;

б) если выполнено усиленное условие Лежандра, но не выполнено условие Якоби, то $S_{\min} = -\infty$;

в) если выполнены усиленные условия Лежандра и Якоби, то $\hat x \in \text{absmin}$ и $S_{\min} \;=\; \mathcal{J}(\hat x)$.

Доказательство. Применив разложение в ряд Тейлора (предложение 10.2) для функционала $\mathcal{J}$ в точке $\hat x$, и учитывая, что $\mathcal{J}'(\hat x) \;=\; 0$, получаем $\mathcal{J}(\hat x + h) \;=\; \mathcal{J}(\hat x) \;+\; \frac{1}{2}\, \mathcal{J}''(\hat x)[h, h] \;=\; \mathcal{J}(\hat x) \;+\; \frac{1}{2}\, \mathcal{J}_{\text{К}}(h)$. В случаях (а) и (б), в силу теорем 10.6 и 11.4, функционал $\mathcal{J}_{\text{К}}(h)$ не является неотрицательно определённым, а значит, $\mathcal{J}_{\text{К}}(h_0) \;=\; -b \;<\; 0$ при некотором $h_0 \;\in\; C_0^1[t_0, t_1]$. Тогда $\mathcal{J}(\hat x + \lambda h_0) \;=\; -\frac{1}{2}\, b\lambda^2 \;\to\; -\infty$ при $\lambda \to \infty$. В случае (с) из теоремы 11.6 заключаем, что $\mathcal{J}''(\hat x) \;\geq\; 0$, а значит $\mathcal{J}(\hat x + h) \;\geq\; \mathcal{J}(\hat x)$ при всех $h \in C_0^1[t_0, t_1]$.

□

12 Лекция

Функция Вейерштрасса. Необходимое условие сильного минимума. Поле экстремалей.

Игольчатые вариации. Функция Вейерштрасса. Условие Вейерштрасса – необходимое условие сильного минимума. Элементы теории поля. Центральное поле экстремалей.

Для исследования экстремали на сильный минимум рассмотрим так называемые *игольчатые вариации.* Возьмём произвольную точку $\tau \in (t_0, t_1)$, произвольное число $v \in \mathbb{R}$ и малый параметр α. Положим $u = \dot{\hat{x}}(\tau)$ и построим следующую функцию.

$$x(t,\alpha) \quad = \quad \begin{cases} \hat{x}(t)\,, & t \in [\,t_0\,,\,\tau-\alpha\,]\ ; \\ \hat{x}(t) + (v-u)\,(t-\tau+\alpha)\,, & t \in [\,\tau-\alpha\,,\,\tau\,]\ ; \\ \hat{x}(t) + \frac{(v-u)\alpha}{t_1-\tau}\,(t_1 - t)\,, & t \in [\,\tau\,,\,t_1\,]\ . \end{cases} \tag{12.26}$$

Это – игольчатая вариация. Заметим, что $x(t,0) = \hat{x}(t)$ и $x(t,\alpha) \in PC^1[t_0,t_1]$ при любом α. Рассмотрим разность

$$\mathcal{J}\big(\,x(\cdot,\alpha)\big) \ - \ \mathcal{J}\big(\,\hat{x}(\cdot)\big) \quad = \quad \int_{t_0}^{\tau-\alpha} + \int_{\tau-\alpha}^{\tau} + \int_{\tau}^{t_1}\ .$$

Обозначим данные интегралы через I_0, I_1, I_2 и оценим их по отдельности. Ясно, что $I_0 = 0$, поскольку $x(t,\alpha) = \hat{x}(t)$ на отрезке $[t_0, \tau-\alpha]$. Далее, по формуле вариации интегрального функционала с подвижными концами (теорема 8.2), имеем при $\alpha = 0$:

$$\frac{d\,I_2}{d\,\alpha} \quad = \quad p(t)\,\frac{dx}{d\alpha}\Big|_{t=\tau}^{t=t_1} \ - \ H(t)\,\frac{dt}{d\alpha}\Big|_{t=\tau}^{t=t_1} \quad = \quad p(t)\,\frac{dx}{d\alpha}\Big|_{t=\tau}^{t=t_1} \quad = \quad -\,p(\tau)\,(v-u),$$

следовательно, $I_2 = -\,\hat{L}_{\dot{x}}(\tau)\,(v-u)\,\alpha + o(\alpha)$, при $\alpha \to 0$. Теперь оценим I_1. Воспользовавшись теоремой о среднем, имеем

$$I_1 \ = \ \alpha\Big(\,L\,\big(\,\tilde{t}, x(\tilde{t},\alpha)\,,\,\dot{x}(\tilde{t},\alpha)\,\big) \ - \ L\,\big(\,\tilde{t}\,,\,\hat{x}(\tilde{t})\,,\,\dot{\hat{x}}(\tilde{t})\,\big)\,\Big)\,,$$

для некоторого $\tilde{t} \in (\tau-\alpha, \tau)$. Заметим, что $L\,\big(\,\tilde{t}\,,\,\hat{x}(\tilde{t})\,,\,\dot{\hat{x}}(\tilde{t})\,\big) \ \to \ L\,\big(\,\tau\,,\,\hat{x}(\tau)\,,\,u\,\big)$ при $\alpha \to 0$ и $L\,\big(\,\tilde{t}\,,\,x(\tilde{t},\alpha)\,,\,\dot{x}(\tilde{t},\alpha)\,\big) \ \to \ L\,\big(\,\tau\,,\,\hat{x}(\tau)\,,\,v\,\big)$ при $\alpha \to 0$. Следовательно,

$$I_1 \ = \ \alpha\Big(\,L\,\big(\,\tau\,,\,\hat{x}(\tau)\,,\,v\,\big) \ - \ L\,\big(\,\tau\,,\,\hat{x}(\tau)\,,\,u\,\big)\Big) \ + \ o(\alpha)\,, \ \text{при}\ \alpha \ \to \ 0\,.$$

Объединяя выражения для I_0, I_1 и I_2, получаем:

$$\mathcal{J}\big(\,x(\cdot,\alpha)\big) - \mathcal{J}\big(\,\hat{x}(\cdot)\big) \quad = \quad \alpha\Big(\,L\,\big(\,\tau\,,\,\hat{x}(\tau)\,,\,v\,\big) - L\,\big(\,\tau\,,\,\hat{x}(\tau)\,,\,u\,\big) - (v-u)\,\hat{L}_{\dot{x}}(\tau)\Big) + o(\alpha)\,, \quad \alpha \to 0\,. \tag{12.27}$$

Определение 12.1 *Функция Вейерштасса:*
$\mathcal{E}\,\big(t,x,u,v\big) \ = \ L(t,x,v) \ - \ L(t,x,u) \ - \ (v-u)\,L_u(t,x,u).$

Полагая в равенстве (12.27) $\alpha \to +0$, получаем, что если $\hat{x}$ – сильный локальный минимум, то $\mathcal{E}\left(\tau, \hat{x}(\tau), \dot{\hat{x}}(\tau), v\right) \ \geq\ 0$ для любых $\tau \in (t_0, t_1)$, $v \in \mathbb{R}$. Из условий гладкости функций L и $L_{\dot{x}}$, следует, что данное неравенство выполнено и на концах отрезка $\tau = t_0, \tau = t_1$. Более того, это условие выполнено и в случае, когда $x \notin C^1$, а лишь $x \in PC^1$. Для доказательства достаточно ограничить простейшую задачу на отрезок, который не содержит точек разрыва производной экстремали $\hat{x}$ кроме, может быть, точки τ (это возможно, поскольку экстремаль доставляет сильный минимум, предложение 9.5).

Будем говорить, что на экстремали $\hat{x}$ выполнено *условие Вейерштрасса*, если

$$\mathcal{E}\left(t, \hat{x}(t), \dot{\hat{x}}(t), v\right) \ \geq\ 0$$

для любого $v \in \mathbb{R}$ в любой точке $t \in [t_0, t_1]$. Таким образом, нами доказана

Теорема 12.2 *Если экстремаль $\hat{x} \in PC^1[t_0, t_1]$ доставляет сильный минимум, то выполнено условие Вейершрасса.*

Условие Вейерштрасса означает, что при любом фиксированном t график функции $f(v) = L(t, \hat{x}(t), v)$ целиком лежит над касательной, проведенной в точке $v = \dot{\hat{x}}(t)$. Таким образом, это – глобальное условие. Оно очевидно, выполнено, если функция $L(t, x, \dot{x})$ *квазирегулярна*, т.е., при каждом фиксированном t функция $f(v) = L(t, \hat{x}(t), v)$ выпукла.

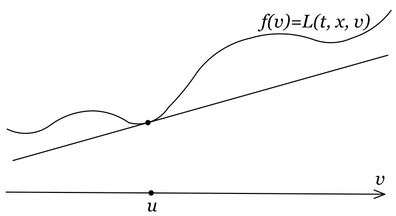

Рис. 6: Условие Вейерштрасса

Предложение 12.3 *Квазирегулярность интегранта L достаточна для выполнения условия Вейерштрасса.*

Далее, заметим, что (положив $u = \dot{\hat{x}}(t)$), $\mathcal{E}(t, \hat{x}, u, u) = 0$ и $\mathcal{E}_v(t, \hat{x}, u, v)\big|_{v=u} = L_v(t, \hat{x}, v)\big|_{v=u} - L_v(t, \hat{x}, u) = 0$. Но поскольку $\mathcal{E}(t, \hat{x}, u, v) \geq 0$ при всех v, отсюда следует, что $L_{vv}(t, \hat{x}, v)\big|_{v=u} = \mathcal{E}_{vv}(t, \hat{x}, u, v)\big|_{v=u} \geq 0$. Поэтому, $\hat{L}_{\dot{x}\dot{x}}(t) \geq 0$. Таким образом, доказано

Предложение 12.4 *Если интегрант L дважды дифференцируем по $\hat{x}$, то условие Вейерштрасса влечёт условие Лежандра.*

Упражнение 12.5 Верно ли, что усиленное условие Лежандра влечет условие Вейерштрасса ?

Функция Вейершрасса позволяет также получить достаточные условия сильного минимума, и даже достаточные условия глобального минимума. Для этого нам сначала нужно будет ознакомиться с элементами теории поля.

Рассмотрим некоторое семейство экстремалей (т.е., решений уравнения Эйлера-Лагранжа) $x(t, \alpha)$, где параметр α пробегает окрестность нуля. При этом мы полагаем $x(t, 0) = \hat{x}(t)$. Предположим, что функции $x, \dot{x}, x_\alpha$ и $\dot{x}_\alpha$ непрерывны. Пусть V – некоторая открытая односвязная окрестность графика экстремали $\hat{x}(t)$.

Определение 12.6 *Экстремаль $\hat{x}$ окружена полем экстремалей, если существует семейство функций $x(t,\alpha)$ такое, что*

1) $x(t,0) = \hat{x}(t)$;

2) при любом α функция $x(t,\alpha)$ – решение уравнения Эйлера-Лагранжа (не обязательно удовлетворяющая краевым условиям данной простейшей задачи).

3) α пробегает окрестность нуля и выполнены условия гладкости: $x, \dot{x}, x_\alpha$ и $\dot{x}_\alpha$ непрерывны;

4) через каждую точку $(t,x) \in V$ проходит единственная функция данного семейства. Это означает, что определена функция $\alpha : V \to \mathbb{R}$, $\alpha = \alpha(t,x)$.

Определение 12.7 *Поле экстремалей называется центральным (ЦПЭ), если все функции семейства $x(t,\alpha)$ проходят через некоторую точку (t^*, x^*).*

Рис. 7:

Таким образом, для центрального поля экстремалей $x(t^*,\alpha) = x^*$ при любом α. При этом всегда предполагается, что $t^* \notin [t_0, t_1]$. Иначе, если $t^* \in [t_0,t_1]$, то поскольку $x^* = x(t^*,0) = \hat{x}(t^*)$, получаем $(t^*,x^*) \in V$, что для поля невозможно, поскольку через эту точку проходит бесконечно много экстремалей.

Пример 12.8 Для задачи $\int_0^1 \dot{x}^2\,dt \to \min$, $x(0) = x(1) = 0$, уравнение Эйлера-Лагранжа: $\ddot{x} = 0$, и все экстремали – суть линейные функции $x(t) = at+b$. Есть много способов окружить экстремаль $\hat{x} = 0$ полем. При этом в качестве окрестности V мы берем вертикальную полосу $V = \big\{(t,x) \in \mathbb{R}^2 \bigm| -\varepsilon < t < 1+\varepsilon\big\}$. Например, $x(t,\alpha) = \alpha(t+1)$, $\alpha \in \mathbb{R}$ – центральное поле. Все его экстремали проходят через точку $(-1,0)$. Поле $x(t,\alpha) = \alpha$ так же окружает нашу экстремаль. Оно не является центральным. Наконец, семейство $x(t,\alpha) = \alpha t$ не является полем, поскольку через точку $(0,0)$ проходит много экстремалей.

Далее везде будем требовать, чтобы $\alpha(t,x) \in C^1(V)$ и будем для определённости полагать, что $t^* < t_0$.

Определение 12.9 *Для произвольного поля экстремалей функция $u(\tau,\xi) = \dot{x}\big(\tau\,,\,\alpha(\tau,\xi)\big)$ называется функцией наклона поля.*

Итак, через данную точку (τ,ξ) мы проводим (единственную!) экстремаль нашего семейства и вычисляем её производную в точке τ. Получаем значение функции $u(\tau,\xi)$ наклона поля.

Пример 12.10 Для задачи о гармоническом осциляторе:

$\int_\varepsilon^{\pi-\varepsilon} \dot{x}^2 - x^2\,dt \to \min$, $x(\varepsilon) = x(\pi-\varepsilon) = 0$, все экстремали имеют вид $x(t) = a\sin t + b\cos t$. Из них можно составить поле, окружающее экстремаль $\hat{x} = 0$. Например, такое: $x(t,\alpha) =$

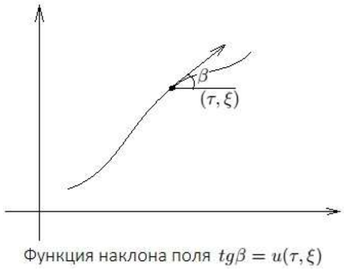

Рис. 8:

$\alpha \sin t$. Оно, более того, центральное: все экстремали проходят через точку $(0,0)$. Вычислим функцию наклона этого поля. Для данной точки $(\tilde{t},\tilde{x})$ имеем $\alpha \sin \tilde{t} = \tilde{x}$, откуда $\alpha = \frac{\tilde{x}}{\sin \tilde{t}}$. Итак, через произвольную точку $(\tilde{t},\tilde{x})$ проходит экстремаль $x(t,\alpha) = \frac{\tilde{x}}{\sin \tilde{t}} \sin t$. Тогда $u(\tilde{t},\tilde{x}) = \dot{x}(\tilde{t},\alpha) = \tilde{x} \operatorname{ctg} \tilde{t}$. Итак, $u(t,x) = x \operatorname{ctg} t$.

13 Лекция

Основная формула Вейерштрасса. Достаточные условия сильного минимума и глобального минимума.

Основная формула Вейерштрасса. Достаточные условия глобального минимума. Задача о брахистохроне.

Определение 13.1 *Для центрального поля экстремалей S-функция определяется следующим образом* $\mathcal{S}(\tau,\xi) = \int\limits_{t^*}^{\tau} L\big(\, t, x(t,\alpha), \dot{x}(t,\alpha) \big)\, dt$*, где* $\alpha = \alpha(\tau,\xi)$.

Таким образом, через точку (τ,ξ) проводится экстремаль $x(t,\alpha)$ (которая, как и все экстремали поля, проходит через точку (t^*,x^*)) и вдоль этой экстремали вычисляется интеграл от $L(t,x,\dot{x})$. Найдём дифференциал S-функции. Фиксируем τ,ξ,δ,h. Рассмотрим семейство экстремалей $X(t,\beta) = x\left(t, \alpha(\tau+\beta\delta\,,\, \xi+\beta h)\right)$, где β пробегает окрестность нуля. Положим также

$$t_0(\beta) = t^*\,,\ t_1(\beta) = \tau + \beta\delta\,;$$
$$x_0(\beta) = x^*\,,\ x_1(\beta) = \xi + \beta h\,,$$

$$\mathcal{J}(\beta) = \int_{t_0(\beta)}^{t_1(\beta)} L\left(t, X(t,\beta), \dot{X}(t,\beta)\right) dt\,.$$

Пользуемся теоремой о дифференцировании интегрального функционала с подвижными концами (теорема 8.2), имеем:

$$\mathcal{J}'(0) \;=\; p\,\frac{d\,x}{d\,\beta}\Big|_{t_0}^{t_1} - H\,\frac{d\,t}{d\,\beta}\Big|_{t_0}^{t_1} \;=\; L_{\dot{x}}\left(\tau,\xi,u(\tau,\xi)\right)h - L_{\dot{x}}\left(\tau,\xi,u(\tau,\xi)\right)u(\tau,\xi)\,\delta + L\left(\tau,\xi,u(\tau,\xi)\right)\delta\,.$$

Таким образом,

$$d\,\mathcal{S} \;=\; L_{\dot{x}}\left(\tau,\xi,u(\tau,\xi)\right)d\,\xi \;-\; L_{\dot{x}}\left(\tau,\xi,u(\tau,\xi)\right)u(\tau,\xi)\,d\,\tau \;+\; L\left(\tau,\xi,u(\tau,\xi)\right)d\,\tau\,.$$

Итак, $\mathcal{S}_\tau \;=\; -\,H\left(\tau,\xi,\mathcal{S}_\xi\right)$, где $H(\tau\,,\,\xi\,,\,p) \;=\; p\,u(\tau,\xi) \;-\; L\big(\tau,\xi,u(\tau,\xi)\big)$. Получили *уравнение Гамильтона-Якоби* на S-функцию. Теперь мы в состоянии доказать следующее основное утверждение:

Теорема 13.2 *Если экстремаль $\hat{x}$ можно окружить ЦПЭ в некоторой окрестности V, то для произвольной допустимой функции x, график которой принадлежит V, выполнено равенство*

$$\mathcal{J}(x) \;-\mathcal{J}(\hat{x}) \quad=\quad \int\limits_{t_0}^{t_1} \mathcal{E}\,\Big(\,t\,,x(t)\,,\,u\big(t,x(t)\big)\,,\,\dot{x}(t)\,\Big)\,d\,t\,. \qquad (13.28)$$

Формула (13.28) называется *основной формулой Вейерштрасса.*

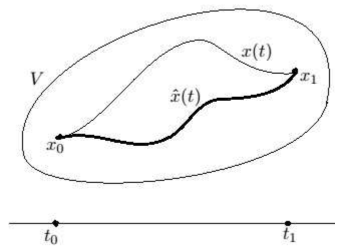

Рис. 9:

`Доказательство.`

$$\mathcal{J}(\hat{x}) \;=\; \int_{t_0}^{t_1} L\left(\,t,x(t,0),\dot{x}(t,0)\,\right)dt \;=\; \int_{t^*}^{t_1} L\left(\,t,x(t,0),\dot{x}(t,0)\,\right)dt - \int_{t^*}^{t_0} L\left(\,t,x(t,0),\dot{x}(t,0)\,\right)dt \;=$$

$$\mathcal{S}(t_1,x_1) \;-\; \mathcal{S}(t_0,x_0) \;=\; \int_{t_0}^{t_1} d\,\mathcal{S}\left(t,x(t)\,\right) \;=\; \left\{\text{поскольку}\,\xi = x(t),\; d\xi \;=\; \dot{x}(t)\,dt\,\right\} \;=$$

$$\int_{t_0}^{t_1} -\Big(\,L_{\dot{x}}\big(t,x(t),u(t,x(t))\big)\,u(t,x(t))\,dt - L\left(t,x(t),u(t,x(t))\right)\Big)\,d\,t + L_{\dot{x}}\big(t,x(t),u(t,x(t)\big)\,\dot{x}(t)\,dt \;=$$

$$\int_{t_0}^{t_1}\Big(L\left(t,x(t),u(t,x(t))\right) \;+\; L_{\dot{x}}\big(t,x(t),u(t,x(t))\big)\left(\dot{x}(t) \;-\; u(t,x(t))\,\right)\Big)\,dt\,.$$

Таким образом, $\mathcal{J}(x) \;-\mathcal{J}(\hat{x}) \quad=$

$$\int\limits_{t_0}^{t_1}\Big(\,L\left(t,x(t),\dot{x}(t)\right) \;-\; L\left(t,x(t),u(t,x(t))\right) \;-\; L_{\dot{x}}\big(t,x(t),u(t,x(t))\big)\left(\dot{x}(t) \;-\; u(t,x(t))\,\right)\Big)\,dt \;=$$

$$\int_{t_0}^{t_1} \mathcal{E}\left(t\,, x(t)\,,\, u\big(t, x(t)\big)\,,\, \dot{x}(t) \right) dt\,.$$

□

Будем говорить, что на экстремали $\hat{x}$, окружённой ЦПЭ, выполнено *усиленное условие Вейерштрасса*, если найдётся $\varepsilon > 0$ такое, что при любом $t \in [t_0, t_1]$, при любых x, u, для которых $|x - \hat{x}(t)| < \varepsilon$, и $|u - \dot{\hat{x}}(t)| < \varepsilon$, и при любом $v \in \mathbb{R}$ имеем $\mathcal{E}\left(t, x, u, v\right) \geq 0$.

Теорема 13.3 *Если экстремаль $\hat{x} \in PC^1[t_0, t_1]$ можно окружить ЦПЭ и выполнено усиленное условие Вейерштрасса, то $\hat{x}$ – сильный минимум.*

`Доказательство.` В случае $x \in C^1$ достаточно воспользоваться непрерывностью функции $u(t, x)$ и формулой (13.28). На общий случай теорема распространяется с помощью леммы о скруглении углов.

□

Следствие 13.4 *Если экстремаль $\hat{x} \in PC^1[t_0, t_1]$ можно окружить ЦПЭ и интегрант L квазирегулярен, то $\hat{x}$ – сильный минимум.*

Следующую теорему мы примем без доказательства.

Теорема 13.5 *Если $L \in C^3$, экстремаль $\hat{x} \in PC^1[t_0, t_1]$ удовлетворяет усиленным условиям Лежандра и Якоби, то она может быть окружена ЦПЭ.*

Следствие 13.6 *Если $L \in C^3$, экстремаль $\hat{x} \in PC^1[t_0, t_1]$ удовлетворяет усиленным условиям Лежандра и Якоби, и выполнено усиленное условие Вейерштрасса, то $\hat{x}$ – сильный минимум.*

Следствие 13.7 *Если интегрант $L \in C^3$ квазирегулярен, экстремаль $\hat{x} \in PC^1[t_0, t_1]$ удовлетворяет усиленным условиям Лежандра и Якоби, то $\hat{x}$ – сильный минимум.*

Займёмся теперь условиями глобального минимума, которые следуют из основной формулы Вейерштрасса. Мы будем формулировать их для всей полосы $t_0 \leq t \leq t_1$, хотя то же верно и для любой односвязной окрестности V графика экстремали (в этом случае речь идет о минимуме среди всех допустимых функций, график которых лежит в данной окрестности). Применив основную формулу Вейерштрасса, получаем

Теорема 13.8 *Если экстремаль $\hat{x} \in PC^1[t_0, t_1]$ окружена ЦПЭ, и $\mathcal{E}\big(t, x(t), u\big(t, x(t)\big), v\big) \geq 0$ при всех $v \in \mathbb{R}$ для любой допустимой функции $x(t)$, в любой точке $t \in [t_0, t_1]$, то $\hat{x}$ – глобальный минимум.*

Интегрант L удовлетворяет *глобальному условию Вейерштрасса*, если при любых $t \in [t_0, t_1]$ и $x, u, v \in \mathbb{R}$ имеем $\mathcal{E}\left(t, x, u, v\right) \geq 0$.

Следствие 13.9 *Если экстремаль $\hat{x} \in PC^1[t_0, t_1]$ окружена ЦПЭ, и функция $L(t, x, v)$ удовлетворяет глобальному условию Вейерштрасса, $\hat{x}$ – глобальный минимум.*

Следствие 13.10 *Если экстремаль $\hat{x} \in PC^1[t_0, t_1]$ окружена ЦПЭ, и при этом функция $L(t, x, v)$ при любых t, x выпукла по v, то $\hat{x}$ – глобальный минимум.*

Итак, применять функцию Вейерштрасса для доказательства того, что экстремаль даёт сильный минимум или глобальный минимум (т.е., в достаточных условиях) можно, только убедившись в том, что её можно окружить ЦПЭ. Это не удивительно, поскольку все наши доказательства базируются на основной формуле Вейерштрасса, которая выполнена при наличии ЦПЭ. Есть, однако, одно замечательное исключение. Это – случай интегранта, не зависящего явно от переменной x.

Предложение 13.11 *Если в простейшей задаче интегрант не зависит явно от x (т.е., $L = L(t, \dot{x})$), а экстремаль $\hat{x} \in PC^1[t_0, t_1]$ удовлетворяет условию Вейерштрасса, то она доставляет глобальный минимум*

Итак, условие, которое в общем случае лишь необходимо для сильного локального минимума (теорема 12.2), в случае интегранта, не зависящего явно от x, становится достаточным для абсолютного минимума. Заметим, что даже существования поля экстремалей для этого не требуется. Таким образом, в этом случае сильный локальный минимум всегда является глобальным.

`Доказательство предложения 13.11.` Если $L = L(t, \dot{x})$, то уравнение Эйлера-Лагранжа принимает вид $\hat{L}_{\dot{x}}(t) = c$. Для любой допустимой функции x, используя условие Вейерштасса, получаем

$$\mathcal{J}(x) - \mathcal{J}(\hat{x}) = \int_{t_0}^{t_1} \Big(L(t, \dot{x}) - L(t, \dot{\hat{x}}(t)) \Big)\, dt \ \geq \ \int_{t_0}^{t_1} \big(\dot{x}(t) - \dot{\hat{x}}(t) \big)\, L_{\dot{x}}(t, \dot{\hat{x}}(t))\, dt =$$

$$\int_{t_0}^{t_1} \big(\dot{x}(t) - \dot{\hat{x}}(t) \big)\, c\, dt = \big(x(t) - \hat{x}(t) \big)\, c \,\Big|_{t_0}^{t_1} = \big(x(t_1) - \hat{x}(t_1) \big)\, c - \big(x(t_0) - \hat{x}(t_0) \big)\, c = 0 .$$

$\square$

Пример 13.12 Рассмотрим вновь задачу из примера 12.10. Для нее имеем $u(t,x) = x \operatorname{ctg} t$, и следовательно

$$\mathcal{E}(t, x, u, v) = (v^2 - x^2) - (u^2 - x^2) - (v - u)\, 2\, u = (v - u)^2 = (v - x \operatorname{ctg} t)^2 .$$

В силу основной формулы Вейерштрасса $\mathcal{J}(x) - \mathcal{J}(\hat{x}) = \int_{\varepsilon}^{\pi - \varepsilon} (\dot{x} - x \operatorname{ctg} t)^2\, dt$. Следовательно, $\hat{x} = 0$ – абсолютный минимум. Это, впрочем следует уже из теоремы 11.10, поскольку выполнены усиленные условия Лежандра и Якоби. Здесь, однако, интересно заметить, что та же интегральная формула верна и в случае $\varepsilon = 0$ (что легко проверить непосредственно). Следовательно, и на отрезке $[0, \pi]$ экстремаль $\hat{x} = 0$ даёт глобальный минимум, хотя в этом случае не выполнено усиленное условие Якоби, и экстремаль нельзя окружить полем.

Пример 13.13 *(Задача о брахистохроне).* Вариационное исчисление появилось в 1696 году. Столь точная дата известна потому, что именно в этом году швейцарский математик Иоганн Бернулли опубликовал в журнале "Acta Eruditorum"первую задачу, которая относится к вариационному исчислению – задачу о брахистохроне:

Две данные точки вертикальной плоскости соединить линией, по которой тело, двигающееся под действием силы тяжести с нулевой начальной скоростью без трения, проходит от одной точки к другой за наименьшее время.

Бернулли пригласил к решению задачи своих современников, обещая "воздать хвалу тому, кто справится с её решением". За короткое время он получил три письма с решениями:

от Лейбница. Лопиталя, от собственного брата – Якоба Бернулли, и одно анонимное письмо (автор которого, впрочем, сразу был опознан по оригинальной и свежей идее, “Как по когтям узнают льва” – это Исаак Ньютон). Мы приведем решение, основанное на идее Ньютона.

Введем на плоскости систему координат (t, x), начальную точку поместим в начало координат, а конечную – в точку $(a, -b)$. Нужно найти кривую $x(t)$, по которой тело спускается от начальной точки к конечной за наименьшее время. По закону сохранения энергии скорость тела в точке (t, x) равна $\sqrt{-2gx}$, длина бесконечно малого участка кривой равна $\sqrt{1+\dot{x}^2(t)}\,dt$. Таким образом, получаем простейшую задачу вариационного исчисления:

$$\begin{cases} \mathcal{J}(x) \; = \; \int\limits_0^a \frac{\sqrt{1+\dot{x}^2(t)}}{\sqrt{-2gx}}\,dt \; \to \; \min\,, \\ \\ x(0) \,=\, 0\,, \quad x(a) \,=\, -b\,. \end{cases} \tag{13.29}$$

Поскольку интегрант не зависит явно от времени, вместо уравнений Эйлера-Лагранжа можно воспользоваться интегралом энергии:

$$\dot{x}\,L_{\dot{x}} \,-\, L \quad = \quad \frac{\dot{x}^2}{\sqrt{-2gx}\,\sqrt{1\,+\,\dot{x}^2}} \,-\, \frac{\sqrt{1+\dot{x}^2(t)}}{\sqrt{-2gx}} \quad \equiv \quad \text{const},$$

Выражая из этого равенства $\dot{x}$ и учитывая, что $x(t) < 0$, получаем

$$-\frac{\dot{x}}{\sqrt{-\frac{C}{x}\,-\,1}} \quad = \quad 1\,,$$

или

$$-\frac{d\,x}{\sqrt{-\frac{C}{x}\,-\,1}} \quad = \quad d\,t\,.$$

После замены $x \,=\, -\,C\,\sin^2\frac{\tau}{2} \,=\, -\frac{C}{2}\left[1\,-\,\cos\tau\right]$, имеем

$$t \,=\, \int \frac{2\,C\,\sin\frac{\tau}{2}\,\cos\frac{\tau}{2}\,\frac{1}{2}\,d\,\tau}{\sqrt{\frac{1}{\sin^2\frac{\tau}{2}}\,-\,1}} \quad = \quad \int C\,\sin^2\frac{\tau}{2}\,d\,\tau \quad = \quad \frac{C}{2}\Big[\tau\,-\,\sin\tau\Big].$$

Получаем уравнение циклоиды:

$$\begin{cases} t \; = \; \frac{C}{2}\Big[\tau\,-\,\sin\tau\Big] \\ x \; = \; -\frac{C}{2}\Big[1\,-\,\cos\tau\Big]\,. \end{cases}$$

Эту кривую описывает точка окружности диаметра C, когда окружность катится по оси OX. Отсюда, в частности, следует, что все циклоиды подобны, коэффициент их подобия – отношение радиусов окружностей.

Упражнение 13.14 Докажите, что дуга циклоиды и ось абсцисс ограничивают выпуклую фигуру.

К сожалению, строго говоря, экстремалью в нашей задаче она не является, поскольку $x(t) \;\sim\; -\,t^{2/3}$ при $t \to 0$, а значит производная функции $x(t)$ стремится к $-\infty$ при $t \to 0$, т.е., $\hat{x} \notin PC^1[0, a]$. Поэтому, она не удовлетворяет определению 6.1, условия первого порядка позволяют искать локальный минимум только для функций из C^1.

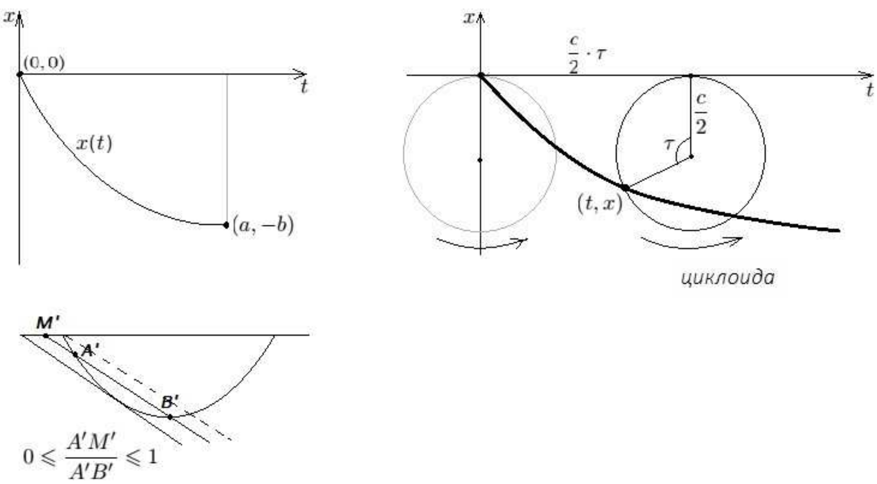

Рис. 10: Задача о брахистохроне. Циклоида

Для корректного решения задачи мы сначала рассмотрим ситуацию, когда тело имеет ненулевую начальную скорость, т.е. $x(0) = -\delta < 0$. При этом предполагается $\delta < b$. В этом случае циклоида $\hat{x}$ является гладкой кривой на $[0, a]$. Заметим, что интегрант в данной задаче $L(t, x, \dot{x}) = \frac{\sqrt{1+\dot{x}^2(t)}}{\sqrt{-2gx}}$ является выпуклой функцией от $\dot{x}$. Поэтому, если экстремаль $\hat{x}$ окружена ЦПЭ в области $V = \{(t, x) \mid x < 0\}$, то согласно следствию 13.10, она доставляет абсолютный минимум среди всех допустимых функций, принимающих неположительные значения. Из физического смысла задачи ясно, что мы ищем минимум только среди таких функций. Для доказательства существования ЦПЭ мы воспользуемся приёмом, который работает во многих аналогичных задачах. Покажем, что *через любую пару точек* A, B*, принадлежащих замыканию области* V *и имеющих разные абсциссы, проходит единственная экстремаль*. Из этого следует существование ЦПЭ. В самом деле. Во-первых, из этого следует существование единственной допустимой экстремали $\hat{x}$. Так как это – участок циклоиды, экстремаль можно продолжить до некоторой точки (t^*, x^*), где $t^* \notin [t_0, t_1]$. Теперь любую точку (t, x), $t \in [t_0, t_1]$, $x \le 0$ соединяем единственной экстремалью с точкой (t^*, x^*). Получаем ЦПЭ, окружающее $\hat{x}$.

Итак, даны точки A, B с неположительными ординатами и разными абсциссами. Пусть прямая AB пересекает ось абсцисс в некоторой точке M. Не ограничивая общности, считаем, что $MA < MB$. Возьмем полную дугу циклоиды. Она ограничивает выпуклую фигуру (упражнение 13.14). Проведём произвольную прямую, параллельную AB, пересекающую данную дугу в точках A', B', а ось абсцисс – в точке M', причём $M'A' < M'B'$ (рис. 10). При движении прямой $A'B'$ отношение $M'A'/M'B'$ непрерывно возрастает от нуля до единицы (единица соответствует касанию). Следовательно, существует единственное положение, при котором $M'A'/M'B' = MA/MB$. Тогда гомотетия, переводящая вектор $A'B'$ в AB переводит дугу циклоиды $A'B'$ в дугу циклоиды AB. Эта циклоида единственна.

Если прямая AB параллельна оси абсцисс, то берём полную дугу циклоиды и прямую, параллельную AB, пересекающую дугу в точках A', B'. Существует единственное положение прямой, при котором отношение длины отрезка $A'B'$ к расстоянию от прямой $A'B'$ до оси абсцисс равно аналогичной величине для отрезка AB. Тогда гомотетия, переводящая вектор $A'B'$ в AB переводит дугу циклоиды $A'B'$ в дугу циклоиды AB.

Упражнение 13.15 Докажите, что единственная дуга циклоиды, проходящая через точки A, B непрерывно зависит от этих точек.

Итак, если $x(0) < 0$, то циклоида даёт абсолютный минимум в задаче о брахистохроне. Осталось осуществить переход к случаю $x(0) = 0$. Пусть $\hat{x}$ – циклоида, проведённая между точками $(0,0)$ и $(a,-b)$. Для произвольного $x \in C^1[0,a]$ обозначим через s наименьшее решение уравнения $x(s) = -\delta$, где $\delta > 0$ – малое число, пусть также r – расстояние от точки $(0,0)$ до $(s,-\delta)$, а $x_\delta(t)$ – циклоида, проведённая между точками $(s,-\delta)$ и $(a,-b)$. Скорость на участке $t \in [0,s]$, по закону сохранения энергии, не превосходит $\sqrt{-2g\delta}$, поэтому время, которое тело проходит этот участок не меньше, чем $\frac{r}{\sqrt{-2g\delta}}$. А время, за которое тело проходит второй участок, не меньше, чем $\mathcal{J}(x_\delta)$ (поскольку на отрезке $[s,a]$ циклоида x_δ даёт абсолютный минимум). Таким образом,

$$\mathcal{J}(x) \quad \geq \quad \frac{r(\delta)}{\sqrt{-2g\delta}} \; + \; \mathcal{J}(x_\delta)\,. \tag{13.30}$$

для любого $\delta > 0$. Заметим, что $r(\delta) \to 0$ при $\delta \to 0$, иначе правая часть данного неравенства может принимать сколь угодно большие значения, и следовательно $\mathcal{J}(x) = +\infty$. Но если $r(\delta) \to 0$, то циклоида x_δ стремится к циклоиде $\hat{x}$ (упражнение 13.15), а значит $\mathcal{J}(x_\delta) \to \mathcal{J}(\hat{x})$. Потому, при $\delta \to 0$ неравенство (13.30) влечёт $\mathcal{J}(x) \geq \mathcal{J}(\hat{x})$.

14 Лекция

Задача Лагранжа.

Общая задача Лагранжа. Задачи вариационного исчисления как частные случаи задачи Лагранжа. Необходимые условия слабого минимума.

Мы изучили несколько типов задач вариационного исчисления – простейшую задачу, задачу Больца, изопериметрическую задачу и задачу с подвижными концами. Сейчас мы сформулируем общую задачу Лагранжа, которая включает все перечисленные задачи как частные случаи.

Общая задача Лагранжа. Дан отрезок $\Delta = [a,b]$. Рассматриваются всевозможные четверки $\xi = (x, u, t_0, t_1)$, где функция $x \in C^1(\Delta, \mathbb{R}^d)$ – *фазовая переменная*, функция $u \in C(\Delta, \mathbb{R}^r)$ – *управление*, $t_0, t_1 \in (a,b)$ – концы отрезка ($t_0 < t_1$). Такая четвёрка называется *управляемым процессом*. Задана непрерывная функция $\varphi : \Delta \times \mathbb{R}^d \times \mathbb{R}^r \to \mathbb{R}^d$, а также $n+m+1$ функционал вида

$$\mathcal{J}_k(\xi) \quad = \quad \int_{t_0}^{t_1} f_k\big(t, x, u\big)\, dt \; + \; \psi_k(t_0, t_1, x_0, x_1)\,, \qquad k \, = \, 0, \ldots, n+m\,, \tag{14.31}$$

где $f_k \in C(\Delta \times \mathbb{R}^d \times \mathbb{R}^r)$, $\psi_k \in C(\Delta^2 \times \mathbb{R}^d \times \mathbb{R}^r)$ – заданные функции, $x_i = x(t_i)$, $i = 0, 1$.

Определение 14.1 *Задача*

$$
\begin{array}{l}
\mathcal{J}_0(\xi) \quad \to \quad \min \\
\\
\textit{при условиях}: \\
\dot{x}(t) \,=\, \varphi\big(t, x(t), u(t)\big)\,,\; t \in \Delta \\
\mathcal{J}_i(\xi) \,\le\, 0\,, \qquad i \,=\, 1, \dots, n\,, \\
\mathcal{J}_j(\xi) \,=\, 0\,, \qquad j \,=\, n+1, \dots, n+m\,.
\end{array}
\tag{14.32}
$$

называется общей задачей Лагранжа.

Четверка $\hat{\xi} = (\hat{x}, \hat{u}, \hat{t}_0, \hat{t}_1)$ является оптимальной в слабом смысле (доставляет слабый локальный минимум) в задаче Лагранжа, если существует $\varepsilon > 0$ такое, что $\mathcal{J}_0(\xi) \ge \mathcal{J}_0(\hat{\xi})$ для любой допустимой (удовлетворяющей всем условиям задачи) четверки ξ, для которой $\|x - \hat{x}\|_{C^1(\Delta, \mathbb{R}^d)} < \varepsilon$, $\|u - \hat{u}\|_{C(\Delta, \mathbb{R}^r)} < \varepsilon$ и $|t_0 - \hat{t}_0| < \varepsilon$, $|t_1 - \hat{t}_1| < \varepsilon$.

Все задачи вариационного исчисления, которые мы изучали, являются частными случаями задачи (14.32). Так, например, изопериметрическая задача (7.16):

$$
\begin{array}{l}
\int\limits_{t_0}^{t_1} f_0\big(t, x, \dot{x}\big)\, dt \quad \to \quad \min \\
\int\limits_{t_0}^{t_1} f_j\big(t, x, \dot{x}\big)\, dt \;=\; a_j\,, \quad j \,=\, 1\,, \dots\,, m\,, \\
x(t_0) \;=\; x_0\,, \quad x(t_1) \;=\; x_1
\end{array}
$$

может быть записана в виде задачи Лагранжа следующим образом:

$$
\begin{array}{l}
\mathcal{J}_0 \;=\; \int\limits_{t_0}^{t_1} f_0\big(t, x, u\big)\, dt \quad \to \quad \min \\
\dot{x}(t) \;=\; u(t)\,, \quad t \in [t_0, t_1] \\
\mathcal{J}_j \;=\; \int\limits_{t_0}^{t_1} f_j\big(t, x, u\big)\, dt \;-\; a_j \;=\; 0\,, \quad j \,=\, 1\,, \dots\,, m\,, \\
\mathcal{J}_{m+1} \,=\, x(t_0) \,-\, x_0 \,=\, 0\,, \\
\mathcal{J}_{m+2} \,=\, x(t_1) \,-\, x_1 \,=\, 0\,.
\end{array}
$$

Аналогично могут быть записаны и другие задачи в вариационного исчисления, везде мы полагаем $\dot{x} = u$. Важно отметить, что задача Лагранжа не является просто задачей с подвижными концами (задача (8.17)), дополненной интегральными условиями типа равенств и неравенств $\mathcal{J}_i \le 0$, $\mathcal{J}_j = 0$. Основная общность задачи Лагранжа заключается в наличии управления u, которое мы можем выбирать по своему усмотрению. Это даёт возможность решать разнообразные задачи, включая задачи со старшими производными.

Пример 14.2 Задача

$$
\begin{array}{l}
\int_{t_0}^{t_1} L(t, x, \dot{x}, \ddot{x})\, dt \quad \to \quad \min \\
x(t_0) = a_0,\; x(t_1) = a_1, \\
\dot{x}(t_0) = v_0, \dot{x}(t_1) = v_1
\end{array}
$$

является задачей Лагранжа (с вектор-функцией x, случай $d = 2$). После переобозначения $x_1(t) = x(t)$, получаем:

$$\begin{array}{l}
\int_{t_0}^{t_1} L(t, x_1, x_2, u)\, dt \quad \to \quad \min \\
\dot{x}_1(t) \;=\; x_2(t)\,, \\
\dot{x}_2(t) \;=\; u(t)\,, \quad t \in [t_0, t_1]\,, \\
x_1(t_0) - a_0 \;=\; 0\,,\; x_1(t_1) - a_1 \;=\; 0 \\
x_2(t_0) - v_0 \;=\; 0\,,\; x_2(t_1) - v_1 \;=\; 0\,.
\end{array}$$

Для формулировки необходимых условий слабого минимума в задаче Лагранжа, нужно во-первых выписать лагранжиан:

$$\mathcal{L}(\lambda, x, u, t_0, t_1) \;=\; \int_{t_0}^{t_1} L(t, x, \dot{x}, u)\, dt \;+\; l(t_0, t_1, x_0, x_1)\,,$$

где

$$L(t, x, \dot{x}, u) \quad = \quad \bigl(\, p(t)\,,\, \dot{x} \,-\, \varphi(t, x, u)\,\bigr) \;+\; \sum_{k=0}^{n+m} \lambda_k\, f_k(t, x, u)$$

– интегрант, а

$$l(t_0, t_1, x_0, x_1) \quad = \quad \sum_{k=0}^{n+m} \lambda_k\, \psi_k(t_0, t_1, x_0, x_1)$$

– терминант. Здесь $\lambda = (\lambda_0, \ldots, \lambda_{n+m})$ – произвольные действительные множители, $p(t) = \bigl(\, p_1(t)\,,\, \ldots\,,\, p_d(t)\,\bigr)$ – непрерывно-дифференцируемая вектор-функция, $\bigl(\, p(t)\,,\, \dot{x} \,-\, \varphi(t, x, u)\,\bigr) \;=\; \sum_{k=1}^{d} p_k(t)\,\bigl(\, \dot{x}_k \,-\, \varphi_k(t, x, u)\,\bigr)$ – скалярное произведение. Функция p, таким образом, также выступает в качестве множителя. Будем говорить, что выполнено условие НЕРОН, если все множители λ вместе с функцией p не обратились в ноль одновременно. Иначе говоря, НЕРОН не выполнен, если все множители λ равны нулю, и функция $p(t)$ – тождественный ноль.

Теорема 14.3 *(необходимые условия слабого минимума в задаче Лагранжа). Предположим, что в задаче Лагранжа (14.32) функции φ и f_k и их частные производные по x и непрерывны, а функции ψ_k принадлежат C^1, $k = 0, \ldots, n + m$. Если четверка $(\hat{x}, \hat{u}, \hat{t}_0, \hat{t}_1)$ доставляет слабый локальный минимум, то существуют НЕРОН множители $p \in C^1(\Delta, \mathbb{R}^d)$, $\lambda \in \mathbb{R}^{n+m+1}$ такие, что выполнены следующие условия:*

1) условие стационарности по x (уравнение Эйлера-Лагранжа):

$$\frac{d}{dt}\,\hat{L}_{\dot{x}} \;=\; \hat{L}_{x} \quad \Leftrightarrow \quad \dot{p}(t) \;=\; -\,p(t)\,\varphi_x(t) \;+\; \sum_{k=0}^{n+m} \lambda_k\,(f_k)_x$$

(это – система из d уравнений);

2) условие трансверсальности:

$$\hat{L}_{\dot{x}}(t_i) \;=\; (-1)^i\,\hat{l}_{x_i}\,,\; i = 0, 1 \quad \Leftrightarrow \quad p(t_i) \;=\; (-1)^i \sum_{k=0}^{m} \lambda_k\,(\psi_k)_{x_i} \quad ,\; i = 0, 1;$$

3) условие стационарности по управлению:

$$\hat{L}_u(t) \;=\; 0 \quad \Leftrightarrow \quad -\big(p(t)\,,\,\varphi_u(t)\big) \;+\; \sum_{k=0}^{n+m} \lambda_k\,(f_k)_u \quad = \quad 0\,;$$

4) условие стационарности на подвижные концы (выписываются только для подвижных концов):

$$\mathcal{L}_{t_i} \;=\; 0 \quad \Leftrightarrow \quad \sum_{k=0}^{n+m} \lambda_k\, f(\hat{t}_i, \hat{x}, \hat{u}) \;=\; (-1)^i \sum_{k=0}^{n+m} \lambda_k \Big((\psi_k)_{t_i} \;+\; (\psi_k)_{x_i}\, \dot{\hat{x}}_i \Big), \quad i \;=\; 0,1\,.$$

5) условие неотрицательности: $\lambda_i \;\ge\; 0,\;\; i \;=\; 0,\ldots,n;$

6) условие дополняющей нежёсткости: $\lambda_i\, \mathcal{J}_i(\hat{\xi}) \;=\; 0,\;\; i \;=\; 1,\ldots,n.$

Чтобы не перегружать изложение, мы докажем теорему только для случая *задачи Лагранжа с фиксированными концами*:

$$\begin{array}{l} \int_{t_0}^{t_1} f(t,x,u)\,dt \quad \to \quad \min \\ \dot{x}(t) \;=\; \varphi\big(t, x(t), u(t)\big) \\ x(t_0) \;=\; x_0\,,\, x(t_1) \;=\; x_1\,, \end{array} \tag{14.33}$$

где $x \;\in\; C^1([t_0,t_1],\mathbb{R}^d)\,,\;\; u \;\in\; C([t_0,t_1],\mathbb{R}^r)$. Таким образом, мы убрали терминальную часть, а также все ограничения типа равенств и неравенств, за исключением граничных условий. Распространение теоремы на общий случай осуществляется несложно, с помощью теоремы 3.1, но мы не будем этого делать. Из всех условий на слабый минимум, которые даёт теорема 14.3, здесь останется только первое и третье, которые теперь выглядят так:

$$\begin{array}{l} \dot{p}\,(t) \quad = \quad -\,p\,(t)\,\varphi_x(t) \;+\; \lambda_0\,\hat{f}_x(t)\,; \\ -\big(p\,(t)\,,\,\varphi_u(t)\big) \;+\; \lambda_0\,\hat{f}_u(t) \;=\; 0\,. \end{array} \tag{14.34}$$

Мы не стали выписывать условия трансверсальности, которые в данной задаче не несут смысловой нагрузки: каждое из них однозначно определяет свой множитель λ_i, который не принимает участия в дальнейшем решении. Надо, таким образом, доказать, что если $\hat{\xi}$ – слабый локальный минимум, то найдётся функция p и число λ_0, не обращающиеся в ноль одновременно, при которых выполнены условия (14.34). Нам понадобятся три вспомогательных факта (два – из дифференциальных уравнений, один – из общей топологии), которые мы примем без доказательства.

Лемма 14.4 *Пусть* $F(t,x,\mu) : \mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^k \;\to\; \mathbb{R}^d$ *отображение, непрерывно-дифференцируемое в произведении окрестности графика решения* $\hat{x}(t)\,,\; t \in [t_0,t_1]$ *уравнения* $\dot{x} = F(t,x,\hat{\mu})\,, x(t_0) = x_0$ *и шара* $\Omega_\varepsilon \;=\; \{\mu \;\in\; \mathbb{R}^k \;\mid\; |\mu - \hat{\mu}| \;<\; \varepsilon\}$*. Тогда найдётся такое* $\delta \;>\; 0$*, что для любого* $\mu \;\in\; \Omega_\delta$ *данное уравнение имеет единственное решение* x *на отрезке* $[t_0,t_1]$*. Это решение имеет непрерывную производную по* μ*:* $y(t) \;=\; x_\mu(t,\hat{\mu})$*, которая удовлетворяет уравнению* $\dot{y} \;=\; F_x\, y \;+\; F_\mu\,,\;\; y(t_0) \;=\; 0$*.*

Лемма 14.5 *(о решении линейных систем). Пусть* $A(t)$ *и* $a(t)$ *– соответственно матрица и вектор в* $\mathbb{R}^d$*, кусочно-непрерывно зависящие от* $t \in [t_0,t_1]$*. Тогда каждое из дифференциальных уравнений* $\dot{p}(t) \;=\; p(t)\,A(t) \;+\; a(t)\,,\;\; p(t_1) = p_1$ *и* $\dot{x}(t) \;=\; A(t)\,x(t) + a(t)\,,\;\; x(t_0) = x_0$ *имеет решение на* $[t_0,t_1]$*. Если* $A(t)$ *и* $a(t)$ *– непрерывны, то данное решение единственно.*

Лемма 14.6 *(о центрированной системе). Если семейство компактов в полном метрическом пространстве таково, что любое конечное подсемейство имеет непустое пересечение, то и все семейство имеет непустое пересечение.*

`Доказательство теоремы 14.3 в случае (14.33).` Пусть $\hat{\xi} = (\hat{x}, \hat{u})$ доставляет слабый минимум в задаче (14.33). Возьмём произвольное семейство h из N функций $h_i \in C([t_0, t_1], \mathbb{R}^r)$, $i = 1, \ldots, N$ и вектор $\alpha = (\alpha_1, \ldots, \alpha_N)$. Пусть $u = \hat{u} + \sum_{i=1}^N \alpha_i h_i$ и x – решение дифференциального уравнения $\dot{x} = \varphi(t, x, u)$ при условии $x(t_0) = x_0$. В силу леммы 14.4 при всех малых α решение существует на $[t_0, t_1]$ (поскольку оно существует при $\alpha = 0$), и оно дифференцируемо по α. Пусть $g_0(\alpha) = \int_{t_0}^{t_1} f\big(t, x(t), u(t)\big)\, dt$, и $g_k(\alpha)$ – k-тая координата вектора $x(t_1)$. Таким образом (далее для краткости скалярные произведения обозначаем как обычное умножение),

$$(g_0)_{\alpha_i}(0) \;=\; \int_{t_0}^{t_1} \big(\hat{f}_x(t) y_i(t) \;+\; \hat{f}_u(t)\, h_i(t)\big)\, dt\,, \qquad i = 1, \ldots, N, \tag{14.35}$$

где y_i – решение уравнения

$$\dot{y}_i \;=\; \varphi_x\, y \;+\; \varphi_u\, h_i\,, \qquad y_i(t_0) \;=\; 0\,, \qquad i = 1, \ldots, N\,. \tag{14.36}$$

Поскольку пара $(\hat{x}, \hat{u})$ даёт слабый минимум в исходной задаче, точка $\alpha = 0$ – точка локального минимума в следующей конечномерной задаче

$$g_0(\alpha) \;\to\; \min\,, \quad g_k(\alpha) \;=\; 0\,, \qquad k \;=\; 1, \ldots, d\,.$$

Пользуясь правилом множителей Лагранжа (теорема 1.9), получаем, что существует вектор $\mu = (\mu_0, \mu_1, \ldots, \mu_d)$, лежащий на единичной сфере пространства $\mathbb{R}^{d+1}$ такой, что $\sum_{k=0}^d \mu_k\, (g_k)_\alpha(0) \;=\; 0$, или (подставляя в (14.35) и учитывая, что $(g_k)_{\alpha_i}(0) \;=\; y_i^k(t_1)$ (k-тая координата вектор-функции y_i) при $k \geq 1$, получаем:

$$\mu_0 \int_{t_0}^{t_1} \Big(\hat{f}_x(t) y(t) \;+\; \hat{f}_u(t)\, h(t)\Big)\, dt \;+\; \sum_{k=1}^d \mu_k\, y^k(t_1) \;=\; 0\,. \tag{14.37}$$

Для каждой совокупности функций $h = \{h_i\}_{i=1}^N$ существует, таким образом, вектор μ. Множество всех таких векторов μ для данной совокупности h образует компакт, лежащий на единичной сфере пространства $\mathbb{R}^d$. Любое конечное пересечение таких компактов непусто. В самом деле, оно соответствует объединению конечного числа совокупностей h, для которого также существует некоторый вектор μ. По лемме о центрированной системе (лемма 14.6) существует вектор μ, принадлежащий всем этим компактам, т.е., для него

$$\mu_0 \int_{t_0}^{t_1} \Big(\hat{f}_x(t)\, y(t) \;+\; \hat{f}_u(t)\, h(t)\Big)\, dt \;+\; \sum_{k=1}^d \mu_k\, y^k(t_1) \;=\; 0\,. \tag{14.38}$$

для любой функции $h \in C([t_0, t_1], \mathbb{R}^r)$. Причём $x(t_1) = x_1$. Пусть теперь $p \in C^1([t_0, t_1], \mathbb{R}^d)$ – решение линейного уравнения $\dot{p} \;=\; -\, p\, \varphi_x + \mu_0\, \hat{f}_x$ с условиями $p_k(t_1) \;=\; -\, \mu_k$, $k = 1, \ldots, d$ (оно существует в силу леммы 14.5). Подставляя в (14.38) вместо $\mu_0 \hat{f}_x(t)\, y(t)$ выражение $\dot{p}\, y + p\, \varphi_x\, y$ и интегрируя по частям, получаем:

$$0 \;=\; \int_{t_0}^{t_1} \big(\dot{p}\, y \;+\; p\, \varphi_x\, y \;+\; \mu_0 \hat{f}_u\, h\big)\, dt \;+\; \sum_{k=1}^d \mu_k\, y^k(t_1) \;=$$

$$p(t_1)y(t_1) \;+\; \int_{t_0}^{t_1} \big(-p\,\dot y \,+\, p\,\varphi_x\, y \,+\, \mu_0 \hat f_u\, h\big)\, dt \;+\; \sum_{k=1}^{d} \mu_k\, y^k(t_1) \;=$$

$$\int_{t_0}^{t_1} \big(-p\,\dot y \,+\, p\,\varphi_x\, y \,+\, \mu_0 \hat f_u\, h\big)\, dt \;=\; \int_{t_0}^{t_1} \big(-p\,(\varphi_x\, y \;+\; \varphi_u\, h) \,+\, p\,\varphi_x\, y \,+\, \mu_0 \hat f_u\, h\big)\, dt$$

(использовали равенство (14.36)) $= \int_{t_0}^{t_1} \big(-p\,\varphi_u\, h \,+\, \mu_0 \hat f_u\, h\big)\, dt.$

Таким образом, $0 \;=\; \int_{t_0}^{t_1} \big(-p\,\varphi_u \;+\; \mu_0 \hat f_u\big)\, h\, dt$ для любой функции $h \in C([t_0, t_1], \mathbb{R}^r)$, следовательно $-p\,\varphi_u \,+\, \mu_0 \hat f_u \,=\, 0$. Итак, функция p удовлетворяет всем условиям (14.34), что завершает доказательство.

□

15 Лекция

Принцип максимума Понтрягина.

Общая задача оптимального управления. Принцип максимума Понтрягина – необходимые условия сильного минимума. Доказательство в случае задачи со свободным концом.

Задача оптимального управления. Дан отрезок $\Delta = [a, b]$. Рассматриваются всевозможные четверки $\xi = (x, u, t_0, t_1)$, где функция $x \in PC^1(\Delta, \mathbb{R}^d)$ – *фазовая переменная*, функция $u \in PC(\Delta, \mathbb{R}^r)$ – *управление*, $t_0, t_1 \in (a, b)$ – концы отрезка ($t_0 < t_1$). Такая четвёрка называется *управляемым процессом*. Задана непрерывная функция $\varphi : \Delta \times \mathbb{R}^d \times \mathbb{R}^r \to \mathbb{R}^d$, а также $n + m + 1$ функционал вида

$$\mathcal{J}_k(\xi) \quad = \quad \int_{t_0}^{t_1} f_k\big(t, x, u\big)\, dt \;+\; \psi_k(t_0, t_1, x_0, x_1)\,, \qquad k \;=\; 0, \dots, n+m\,, \tag{15.39}$$

где $f_k \in C(\Delta \times \mathbb{R}^d \times \mathbb{R}^r)$, $\psi_k \in C(\Delta^2 \times \mathbb{R}^d \times \mathbb{R}^r)$ – заданные функции, $x_i = x(t_i)$, $i = 0, 1$. Кроме того, задано произвольное множество $U \subset \mathbb{R}^r$.

Определение 15.1 *Задача*

$$\mathcal{J}_0(\xi) \quad \to \quad \min$$

$$\begin{array}{l} \textit{при условиях}: \\ \dot x(t) \;=\; \varphi\big(t, x(t), u(t)\big)\,, \textit{ в точках непрерывности } t \in \Delta \textit{ функции } u(t) \\ \mathcal{J}_i(\xi) \,\le\, 0\,, \qquad i \,=\, 1, \dots, n\,, \\ \mathcal{J}_j(\xi) \,=\, 0\,, \qquad j \,=\, n+1, \dots, n+m\,, \\ u(t) \;\in U \textit{ при любом } t \,\in\, \Delta \end{array} \tag{15.40}$$

называется общей задачей оптимального управления.

Четверка $\hat{\xi} = (\hat{x}, \hat{u}, t_0, t_1)$ *является оптимальной в сильном смысле (доставляет сильный локальный минимум) в задаче оптимального управления, если существует* $\varepsilon > 0$ *такое, что* $\mathcal{J}_0(\xi) \geq \mathcal{J}_0(\hat{\xi})$ *для любой допустимой (удовлетворяющей всем условиям задачи) четверки* ξ*, для которой* $\|x - \hat{x}\|_{C(\Delta,\mathbb{R}^d)} < \varepsilon$ *и* $|t_0 - \hat{t}_0| < \varepsilon$*,* $|t_1 - \hat{t}_1| < \varepsilon$*.*

На первый взгляд, постановка задачи оптимального управления практически ничем не отличается от задачи Лагранжа. Однако, у этих задач есть несколько существенных различий, благодаря которым задачи оптимального управления являются значительно более общими, а метод их решения – принцип максимума Понтягина – более мощным. Перечислим и обсудим эти различия в порядке их значимости.

1) Условие $u(t) \in U$ является новым, его не было в задаче Лагранжа. Оно означает, что задана область значений функции u, и управление в любой точке может принимать значения только из этого множества. Множество U – любое, не обязательно открытое или замкнутое. Оно может совпадать со всем $\mathbb{R}^d$ (как в задаче Лагранжа), а может и состоять, например, всего из двух точек. Наличие такого условия позволяет решать задачи с ограничениями вида $|u(t)| \leq 1$ или $u(t) \geq 2$ (в первом примере $U = [-1, 1]$, во втором $U = [2, +\infty)$). Это исключительно важно для многих практических задач, в которых управление, как правило, ограниченно некоторыми естественными условиями (например, мощность двигателя, скорость, темпы инвестиций, и .д.)

2) Задача оптимального управления ищет минимум в более широком пространстве функций – PC^1 для фазовой переменной и PC для управления (вместо, соответственно, C^1 и C в задаче Лагранжа).

3) С другой стороны, в задаче Лагранжа ищется слабый минимум, а оптимальный процесс понимается локально как по x, так и по управлению u. В то время как в задаче оптимального управления ищется сильный минимум, а оптимальный процесс понимается локально только по x. Это означает, что в задаче Лагранжа мы имеем дело с более широким классом минимумов. Данное обстоятельство является недостатком задач оптимального управления. В частности, задача Лагранжа не являются частным случаем последних, как может показаться с первого взгляда.

Для решения задачи оптимального управления нужно выписать лагранжиан, который выглядит так же, как в задаче Лагранжа:

$$\mathcal{L}(\lambda, x, u, t_0, t_1) \;=\; \int_{t_0}^{t_1} L(t, x, u)\, dt \;+\; l(t_0, t_1, x_0, x_1)\,,$$

где

$$L(t, x, \dot{x}, u) \quad = \quad \big(p(t)\,,\, \dot{x} \,-\, \varphi(t, x, u)\big) \;+\; \sum_{k=0}^{n+m} \lambda_k\, f_k(t, x, u)$$

– интегрант, а

$$l(t_0, t_1, x_0, x_1) \quad = \quad \sum_{k=0}^{n+m} \lambda_k\, \psi_k(t_0, t_1, x_0, x_1)$$

– терминант. Здесь $\lambda = (\lambda_0, \ldots, \lambda_{n+m})$ – произвольные действительные множители, $p(t) = \big(p_1(t)\,,\, \ldots\,,\, p_d(t)\big) \in PC^1(\Delta, \mathbb{R}^d)$ – некоторая, изначально не известная, вектор-функция, $\big(p(t)\,,\, \dot{x} \,-\, \varphi(t, x, u)\big) \;=\; \sum_{k=1}^{d} p_k(t)\,\big(\,\dot{x}_k \,-\, \varphi_k(t, x, u)\big)$ – скалярное произведение.

Теорема 15.2 *(необходимые условия сильного минимума в задаче оптимального управления). Предположим, что в задаче (15.40) функции φ и f_k и их частные производные по x и u непрерывны, а функции ψ_k принадлежат C^1, $k = 0, \dots, n+m$. Если четверка $(\hat{x}, \hat{u}, \hat{t}_0, \hat{t}_1)$ доставляет сильный локальный минимум, то существуют НЕРОН множители $p \in C^1(\Delta, \mathbb{R}^d)$, $\lambda \in \mathbb{R}^{n+m+1}$ такие, что выполнены следующие условия:*

1) условие стационарности по x (уравнение Эйлера-Лагранжа):

$$\frac{d}{dt}\hat{L}_{\dot{x}} \;=\; \hat{L}_x \quad \Leftrightarrow \quad \dot{p}(t) \;=\; -\,p(t)\,\varphi_x(t) \;+\; \sum_{k=0}^{n+m} \lambda_k\,(f_k)_x\,;$$

2) условие трансверсальности:

$$\hat{L}_{\dot{x}}(t_i) \;=\; (-1)^i\,\hat{l}_{x_i}\,,\; i = 0,1 \quad \Leftrightarrow \quad p(t_i) \;=\; (-1)^i \sum_{k=0}^{m} \lambda_k\,(\psi_k)_{x_i} \quad ,\; i = 0,1;$$

3) условие стационарности по управлению:

$$\min_{u \in U} L\bigl(t, \hat{x}(t), u\bigr) \quad = \quad L\bigl(t, \hat{x}(t), \hat{u}(t)\bigr)$$

в каждой точке $t \in \Delta$ непрерывности управления u. Таким образом, в каждой точке непрерывности значение $\hat{u}(t)$ – есть точка абсолютного минимума лагранжиана $L(u)$.

4) условие стационарности на подвижные концы (выписываются только для подвижных концов):

$$\mathcal{L}_{t_i} \;=\; 0 \quad \Leftrightarrow \quad \sum_{k=0}^{n+m} \lambda_k\, f(\hat{t}_i, \hat{x}, \hat{u}) \;=\; (-1)^i \sum_{k=0}^{n+m} \lambda_k \Bigl((\psi_k)_{t_i} \;+\; (\psi_k)_{x_i}\,\dot{\hat{x}}_i \Bigr), \quad i \;=\; 0,1\,.$$

5) условие неотрицательности: $\lambda_i \;\geq\; 0,\;\; i \;=\; 0, \dots, n;$

6) условие дополняющей нежёсткости: $\lambda_i\, \mathcal{J}_i(\hat{\xi}) \;=\; 0,\;\; i \;=\; 1, \dots, n.$

Данная теорема называется *принципом максимума Понтрягина.* Как и следовало ожидать, различие с решением задачи Лагранжа – только в пункте (3), в условии стационарности по управлению. Это маленькое различие, однако, ведет к совсем другой тактике решения. Кроме того, доказательство принципа максимума значительно более сложно и основано на новых идеях. Мы докажем принцип максимума в частном случае *задачи со свободным концом*:

$$\begin{array}{l} \mathcal{J}_0(\xi) \;=\; \int_{t_0}^{t_1} f(t,x,u)\,dt \;+\; \psi\bigl(x(t_1)\bigr) \quad \to \quad \min \\ \dot{x}(t) \;=\; \varphi\bigl(t, x(t), u(t)\bigr) \quad \text{в точках непрерывности } u(t) \\ u(t) \;\in\; U\,, \qquad t \in [t_0, t_1]\,, \\ x(t_0) \;=\; x_0\,. \end{array} \tag{15.41}$$

Мы, таким образом, зафиксировали концы отрезка, убрали терминальную часть, а также все ограничения типа равенств и неравенств, за исключением граничных условий. В этом случае лагранжиан имеет вид (сразу берём $\lambda_0 = 1$):

$$\mathcal{L}\left(\xi, \lambda, p\right) \;=\; \int_{t_0}^{t_1} \Bigl(f\bigl(t, x(t), u(t)\bigr) + \bigl(p(t), \dot{x}(t) - \varphi(t, x(t), u(t))\bigr) \Bigr)\,dt + \psi\bigl(x(t_1)\bigr) + \sum_{i=0}^{d} \lambda_i\,\bigl(x_i(t_0) - (x_0)_i\bigr)\,.$$

Из всех условий на слабый минимум, которые даёт теорема 15.2, здесь останется только первое, второе и третье. :

1) уравнение Эйлера-Лагранжа: $\dot{p}(t) \;=\; -\,p(t)\,\hat{\varphi}_x(t) \;+\; \hat{f}_x(t)\,;$

2) условия трансверсальности:
$$p_i(t_0) \,=\, \lambda_i\,, \quad i \,=\, 1,\ldots,d\,, \qquad p(t_1) \,=\, -\,\hat{\psi}_{x_1}(t_1) \tag{15.42}$$

3) условие стационарности по u :
$u \,=\, \hat{u}(t)$ – точка минимума функции $\quad f\,\big(t,\hat{x}(t),u\big) \,-\, \big(p(t)\,,\,\varphi(t,\hat{x}(t),u)\,\big)\,.$

Надо, таким образом, доказать, что если $\hat{\xi}$ – сильный минимум, то найдутся функции $p_1(t),\ldots,p_d(t)$ и числа λ_i, не обращающиеся в ноль одновременно, при которых выполнены условия (15.42).

Доказательство. Определим вектор-функцию $p(t)$ условиями из пунктов 1 и 2 в равенствах (15.42):
$$\begin{cases} \dot{p}(t) \;=\; -\,p(t)\,\hat{\varphi}_x(t) \;+\; \hat{f}_x(t)\,; \\ p(t_1) \;=\; -\,\hat{\psi}_{x_1}(t_1)\,. \end{cases} \tag{15.43}$$
Если мы докажем, что для этой функции выполнено условие 3, то после этого все множители λ_i определятся из условия 2, и всё будет доказано. Решение системы существует в силу леммы 14.5. Пусть $\tau_i\,,\ i = 1,\ldots,q$ – все точки разрыва функции u, в порядке возрастания. На каждом из отрезков, на которые данные точки делят отрезок $[t_0,t_1]$ правая часть уравнения непрерывна, а значит решение p единственно. Таким образом, p определяется однозначно на отрезке $[\tau_q,t_1]$, значит – и на отрезке $[\tau_{q-1},\tau_q]$, и т.д. Итак, решение p однозначно определено на всём отрезке $[t_0,t_1]$.

Возьмём произвольную точку τ, в которой $u(t)$ непрерывно, точку $v \in U$ и малое число $\alpha > 0$. Рассмотрим следующую игольчатую вариацию:
$$u_{\tau\,,\,v\,,\,\alpha}(t) \,=\, u_{\alpha}(t) \;=\; \begin{cases} \hat{u}(t)\,,\ t \notin (\tau-\alpha,\tau]\,, \\ v\ ,\ t \in (\tau-\alpha,\tau]\,. \end{cases}$$

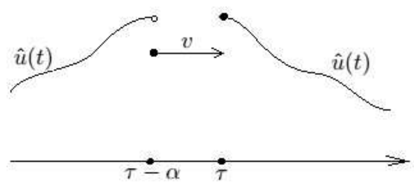

Рис. 11: Функция $u_{\tau\,,\,v\,,\,\alpha}(t)$

Пусть также $x_{\tau\,,\,v\,,\,\alpha}(t) \,=\, x_{\alpha}(t)$ – решение следующей задачи:
$$\begin{cases} \dot{x}_\alpha \;=\; \varphi(t,x_\alpha,u_\alpha)\,, \quad t\, \text{– точка непрерывности}\, u\,, t \neq \tau-\alpha, t \neq \tau\,. \\ x_\alpha(t_0) \;=\; x_0\,. \end{cases}$$

Из леммы 14.4 следует, что при достаточно малых $\alpha > 0$ решение x_α существует (на всем отрезке $[t_0,t_1]$), при этом $\|x_\alpha - \hat{x}\|_{C([t_0,t_1],\mathbb{R}^d)} \to 0$ при $\alpha \to +0$, решение x_α дифференцируемо по α на отрезке $[\tau, t_1]$. Это значит, что предел $y(t) = \lim_{\alpha\to+0} \frac{x_\alpha - \hat{x}}{\alpha}$ сходится равномерно на отрезке $[\tau, t_1]$). Функция $y(t)$ – это производная функции $x_\alpha(t)$ по параметру α. В силу леммы 14.4, она является решением дифференциального уравнения:

$$\begin{cases} \dot{y} \ = \ \hat{\varphi}_x(t)\, y(t)\,, & t \in (\tau, t_1] \\ y(\tau) \ = \ \varphi\big(\tau, \hat{x}(\tau), v\big) \ - \ \varphi\big(\tau, \hat{x}(\tau), \hat{u}(\tau)\big). \end{cases} \tag{15.44}$$

В самом деле, в силу леммы 14.4, имеем $\dot{y}(t) = \hat{\varphi}_x(t) y(t) + \hat{\varphi}_u(t) \frac{d}{d\alpha} u_\alpha(t)$, но для $t \in (\tau, t_1]$ имеем $u_\alpha(t) = \hat{u}(t)$, а значит $\frac{d}{d\alpha} u_\alpha(t) = 0$. Поясним начальное условие в точке τ. Имеем $y(\tau) = \lim_{\alpha\to 0} \frac{x_\alpha - \hat{x}}{\alpha} = \lim_{\alpha\to 0} \alpha^{-1} \int_{\tau-\alpha}^{\tau} (\dot{x}_\alpha - \dot{\hat{x}})\, dt \ = \ \lim_{\alpha\to 0} \alpha^{-1} \int_{\tau-\alpha}^{\tau} (\varphi(t, x_\alpha(t), v) - \varphi(t, \hat{x}(t), \hat{u}(t)))\, dt \ = \varphi(t, \hat{x}(\tau), v) - \varphi(t, \hat{x}(\tau), \hat{u}(\tau))$.

Сравним значения функционала $\mathcal{J}_0(\xi)$ в точках $\hat{\xi} = (\hat{x}, \hat{u})$ и $\xi_\alpha = (x_\alpha, u_\alpha)$. Начнём с интегральной части:

$$\int_{t_0}^{t_1} f\left(t, x_\alpha(t), u_\alpha(t)\right) dt \ - \ \int_{t_0}^{t_1} f\left(t, \hat{x}(t), \hat{u}(t)\right) dt \quad =$$

$$\int_{t_0}^{\tau-\alpha} \Big(f\left(t, x_\alpha(t), u_\alpha(t)\right) \ - \ f\left(t, \hat{x}(t), \hat{u}(t)\right) \Big)\, dt \ + \int_{\tau-\alpha}^{\tau} \ + \int_{\tau}^{t_1} \quad = \quad I_0 \ + \ I_1 \ + \ I_2 \ = \ I_1 \ + \ I_2\,.$$

($I_0 = 0$, поскольку на отрезке $[t_0\,,\, \tau - \alpha]$ имеем $u_\alpha = \hat{u}$ и $x_\alpha = \hat{x}$). Для вычисления I_1 воспользуемся теоремой о среднем. Для этого выберем α настолько малым, чтобы подынтегральная функция была непрерывна на $[\tau - \alpha, \tau]$. Имеем

$$I_1 \ = \ \alpha \Big(f\big(s, x_\alpha(s), u_\alpha(s)\big) \ - \ f\big(s, \hat{x}(s), \hat{u}(s)\big) \Big)\,, \qquad s \in (\tau - \alpha, \tau]\,.$$

При $\alpha \to +0$ имеем $s \to \tau$, $x_\alpha(s) \to \hat{x}(\tau)$, $u_\alpha(s) \to v$, $\hat{x}(s) \to \hat{x}(\tau)$, $\hat{u}(s) \to \hat{u}(\tau)$. Поэтому

$$I_1 \ = \ \alpha \Big(f\big(\tau, \hat{x}(\tau), v\big) \ - \ f\big(\tau, \hat{x}(\tau), \hat{u}(\tau)\big) \Big) \ + \ o(\alpha)\,, \qquad \alpha \to +0\,. \tag{15.45}$$

Теперь разберёмся с I_2.

$$f\big(t, x_\alpha(t), u_\alpha(t)\big) \ - \ f\big(t, \hat{x}(t), \hat{u}(t)\big) \quad = \quad \big(\, \hat{f}_x(t)\,,\, x_\alpha(t) \ - \ \hat{x}(t) \,\big) \ + \ o\left(|x_\alpha(t) \ - \ \hat{x}(t)|\right) \ =$$

$$\big(\, \hat{f}_x(t)\,,\, \alpha\, y(t) \ + \ o(\alpha) \,\big) \ + \ o\left(\alpha\, y(t) \ + \ o(\alpha)\right) \ = \ \alpha \left(\, \hat{f}_x(t)\,,\, y(t) \,\right) \ + \ o\left(\alpha\right)\,, \qquad \alpha \to +0\,.$$

Таким образом,

$$I_2 \quad = \quad \alpha \int_{\tau}^{t_1} \left(\, \hat{f}_x(t)\,,\, y(t) \,\right) dt \ + \ o\left(\alpha\right)\,, \qquad \alpha \to +0\,. \tag{15.46}$$

Теперь вспомним, что $p\,(t_1) = -\,\hat{\psi}_{x_1}(t_1)$ и выразим разность $\big(p(t_1), y(t_1)\big) - \big(p(\tau), y(\tau)\big)$ двумя способами. С одной стороны

$$\big(p(t_1), y(t_1)\big) \ - \ \big(p(\tau), y(\tau)\big) \ = \ -\Big(\hat{\psi}_{x_1}(t_1)\,,\, y(t_1) \Big) \ - \ \Big(p(\tau)\,,\, \varphi(\tau, \hat{x}(\tau), v) \ - \ \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau)) \Big)\,. \tag{15.47}$$

С другой стороны, применив формулу Лейбница и уравнение Эйлера-Лагранжа, получаем

$$\frac{d}{dt} \left(\, p\,,\, y \,\right) \ = \ \dot{p}\, y \ + \ p\, \dot{y} \ = \ \Big(\hat{f}_x \ - \ p\, \hat{\varphi}_x \Big)\, y \ + \ p\, \dot{y} \ = \ \Big(\hat{f}_x \ - \ p\, \hat{\varphi}_x \Big)\, y \ + \ p\, \hat{\varphi}_x\, y \ = \ \hat{f}_x\, y\,.$$

Мы воспользовались уравнением (15.44). Таким образом,

$$\big(p(t_1), y(t_1)\big) \;-\; \big(p(\tau), y(\tau)\big) \;=\; \int_{\tau}^{t_1} \big(\hat{f}_x(t), y(t)\big)\, dt\,.$$

Применив теперь равенства (15.46) и (15.47), получаем

$$I_2 \quad = \quad \alpha\,\Big(\,-\big(\hat{\psi}_{x_1}(t_1), y(t_1)\big) \;-\; \Big(p(\tau)\,,\,\varphi(\tau,\hat{x}(\tau),v) \;-\; \varphi(\tau,\hat{x}(\tau),\hat{u}(\tau))\Big)\,\Big) \;+\; o(\alpha)\,.$$

Совместив это с выражением (15.46) для интеграла I_1, мы получаем такое равенство для интегральной части:

$$\int_{t_0}^{t_1} f\,\big(t, x_\alpha(t), u_\alpha(t)\big)\, dt \;-\; \int_{t_0}^{t_1} f\,\big(t, \hat{x}(t), \hat{u}(t)\big)\, dt \quad = \quad I_1 \;+\; I_2 \quad =$$

$$\alpha\,\Big(\,f\big(\tau,\hat{x}(\tau),v\big) - f\big(\tau,\hat{x}(\tau),\hat{u}(\tau)\big) - \big(\hat{\psi}_{x_1}(t_1), y(t_1)\big) - \Big(p(\tau)\,,\,\varphi(\tau,\hat{x}(\tau),v) - \varphi(\tau,\hat{x}(\tau),\hat{u}(\tau))\Big)\,\Big) + o(\alpha)\,.$$

Разберёмся теперь с терминальной частью: $x_\alpha(t_1) - \hat{x}(t_1) \;=\; \alpha y(t_1) + o(\alpha)$. Таким образом,

$$\psi\big(x_\alpha(t_1)\big) \;-\; \psi\big(\hat{x}(t_1)\big) \;=\; \big(\,\hat{\psi}_{x_1}\,,\,\alpha y(t_1)\,\big) \;+\; o(\alpha)\,.$$

Теперь оцениваем разность $\mathcal{J}_0(\xi_\alpha) - \mathcal{J}_0(\hat{\xi})$. Поскольку $\hat{\xi}$ доставляет сильный минимум, имеем:

$$0 \;\le\; \mathcal{J}_0(\xi_\alpha) - \mathcal{J}_0(\hat{\xi}) \;=\; \alpha\,\Big(\,f\big(\tau,\hat{x}(\tau),v\big) - f\big(\tau,\hat{x}(\tau),\hat{u}(t)\big) - \big(\,\hat{\psi}_{x_1}\,,\,y(t_1)\,\big) - \big(p(\tau)\,,\,\varphi(\tau,\hat{x}(\tau),v) -$$

$$-\,\varphi(\tau,\hat{x}(\tau),\hat{u}(\tau))\big) \;+\; \big(\,\hat{\psi}_{x_1}\,,\,y(t_1)\,\big)\,\Big) \;+\; o(\alpha)\,.$$

При $\alpha \to +0$ получаем

$$f\big(\tau,\hat{x}(\tau),v\big) \;-\; f\big(\tau,\hat{x}(\tau),\hat{u}(t)\big) \;-\; \big(p(\tau)\,,\,\varphi(\tau,\hat{x}(\tau),\hat{u}(\tau))\,\big) \;+\; \big(p(\tau)\,,\,\varphi(\tau,\hat{x}(\tau),v)\,\big) \;\ge\; 0$$

при любом $v \in U$ в каждой точке τ, в которой функция $\hat{u}(t)$ непрерывна. А это и есть условие оптимальности по управлению u. Таким образом, функция p и множитель $\lambda_0 = 1$ удовлетворяют всем условиям 1 – 3 принципа максимума для случая задачи с подвижным концом.

□

Список литературы

[1] Алексеев В.М., Тихомиров В.М., Фомин С.В. *Оптимальное управление*, М. Наука, 1979.

[2] Галев Э.М., Зеликин М.И., Конягин С.В., Магарил-Ильяев Г.Г., Осмоловский Н.П., Протасов В.Ю., Тихомиров В.М., Фурсиков А.В. *Оптимальное управление*, М. МЦНМО, 2008.

[3] Зеликин М.И., *Оптимальное управление и вариационное исчисление*, М. УРСС, 2004.

[4] Конягин С.В., *Вариационное исчисление и оптимальное управление*, 2005, http://mech.math.msu.su/ kon/opu05.html.

[5] Алексеев В.М., Галеев Э.М., Тихомиров В.М., *Сборник задач по оптимизации*, М. Наука, 1984.

[6] Постников М.М., *Лекции по геометрии. Семестр III. Гладкие многообразия*, М. Наука, 1987.

[7] Бляшке В., *Круг и шар*, М. Наука, 1967.

Содержание

Предметный указатель