МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

имени М.В. ЛОМОНОСОВА

МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ

КАФЕДРА ОБЩИХ ПРОБЛЕМ УПРАВЛЕНИЯ

А. С. Кочуров, В. М. Тихомиров

Краткий курс теории экстремальных задач

МОСКВА – 2013

УДК 517. 97(07)
ББК 22.161.8 я73

А. С. Кочуров, В. М. Тихомиров
Краткий курс теории экстремальных задач: Учеб. пособие — М.: Мехмат МГУ, 2013. – 48 с.

Цель пособия — изложить основные идеи, принципы и результаты теории экстремума (с полными доказательствами) от истоков до нашего времени. Здесь представлены основные в данном круге вопросов теоремы Ферма, Эйлера, Лагранжа, Лежандра, Гамильтона, Якоби, Вейерштрасса, а также результаты, полученные в этом веке — теорема существования Тонелли, теорема двойственности линейного программирования, восходящая к Канторовичу, и принцип максимума Понтрягина.

ISBN 978-5-93838-040-0

Оглавление

Введение

Экстремальные задачи (задачи на максимум и минимум), возникающие в математике, естествознании, практической деятельности, обычно ставятся без формул, используя термины той области знаний, в которой они возникли. Для того, чтобы было возможно применять для решения таких задач математические средства, необходимо произвести перевод постановки задачи на язык математики. Такой перевод называется *формализацией.*

Формализовать экстремальную задачу означает описать *функционал* $f : X \to \overline{\mathbb{R}}$, $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$, который требуется минимизировать или максимизировать, и описать *ограничение* $C \subset X$, на котором нужно найти экстремум f. Далее для формализованной записи задачи о минимизации (максимизации) f при ограничении $x \in C$ применяется запись:

$$f(x) \to \min(\max),\ x \in C. \qquad (P)$$

Если $C = X$, задачу (P) называют *задачей без ограничений.* Точки $x \in C$ называются *допустимыми*; допустимую точку $\widehat{x}$ называют *решением задачи* или *точкой абсолютного минимума (максимума)* (P), если $f(x) \geq f(\widehat{x})$ $(f(x) \leq f(\widehat{x}))$ для всех $x \in C$. Нахождению решения задачи обычно предшествует поиск *локальных* минимумов или максимумов. При этом X наделяется некоторой топологией (т.е. в нём определяется система открытых множеств) и ищется решение задачи $f(x) \to \min(\max),\ x \in C \cap V$, где V открыто в X (именно такие решения называют локальными экстремумами (P)).

В учебнике приведены доказательства результатов, истоки которых содержались в трудах Ферма, Эйлера, Лагранжа, Лежандра, Гамильтона, Якоби, Вейерштрасса, Гильберта, Минковского, Понтрягина (см. [1]–[11]).

Вот основные теории, из которых складывается курс, и принципы, на которых эти теории основываются.

Теорий три: *необходимых условий, достаточных условий и существования.* В учебнике областью исследования для них служат три класса задач: *гладкие* (включая задачи вариационного исчисления), *выпуклые* (каковы, к примеру, задачи линейного программирования) и *гладко–скрыто–выпуклые* (к ним относятся задачи оптимального управления).

Основной принцип теории необходимых условий — *принцип Лагранжа* о свед́ении задач с ограничениями к задачам без ограничений с помощью функции Лагранжа. Достаточные условия основываются на *идее Гамильтона* о том, что следует рассматривать семейства экстремалей и изучать S-функции задачи, построенные по этим семействам. Результаты по теории существования основываются в основном на *принципе компактности*, восходящем к Вейерштрассу, согласно которому непрерывная (или полунепрерывная снизу) функция на любом компактном множестве достигает наименьшего значения. В теории выпуклости фундаментальную роль играет *принцип двойственности Минковского* о том, что выпуклые объекты допускают двойное описание — в основном и двойственном пространствах.

Теории, построенные в учебнике, дают возможность исследовать разнообразные экстремальные проблемы; на его основе составлен Задачник, в котором собрано около двухсот задач. Среди них свыше сотни связаны с именами выдающихся математиков, решавших проблемы своими индивидуальными приемами. В Задачнике все эти проблемы решаются стандартно, пользуясь принципом Лагранжа. Это служит веским основанием тому, что при необходимости решить конкретную задачу из рассматриваемых классов, надо попытаться сделать это методом Лагранжа, а если получаемые уравнения недоступны для решения на руках, надо применить компьютеры.

Предварительные сведения

Для описания пространства X в задаче (P) и формулировки результатов в нашем учебнике используются, в основном, пространства $\mathbb{R}^n$ векторов-столбцов с n координатами, двойственное к нему пространство $(\mathbb{R}^n)^*$ векторов-строк с n координатами, а также пространство $C([t_0,t_1],\mathbb{R}^n)$ с равномерной нормой. Внутреннее произведение $\sum_{i=1}^n y_i x_i$ векторов $y,x\in\mathbb{R}^n$ (или $y,x\in(\mathbb{R}^n)^*$) обозначается $\langle y,x\rangle$, действие элемента $z\in(\mathbb{R}^n)^*$ на вектор $x\in\mathbb{R}^n$ обозначается $z\cdot x := z(x) := \sum_{i=1}^n z_i x_i = \langle z^T, x\rangle$ или просто

$z\,x$ (и это соответствует правилу умножения матриц).

Пространство непрерывно дифференцируемых функций $x(\cdot)$, заданных на отрезке $[t_0,t_1]$ и принимающих значения в $\mathbb{R}^n$, будем обозначать $C^1([t_0,t_1],\mathbb{R}^n)$; норму в нём определим как $\|x(\cdot)\|_{C^1} = \max\{\|x(\cdot)\|_C,\ \|\dot{x}(\cdot)\|_C\}$. Пусть $PC([t_0,t_1],U)$ и $C([t_0,t_1],U)$ — семейства всех *кусочно-непрерывных* и *непрерывных* на $[t_0,t_1]$ функций со значениями в $U\subset\mathbb{R}^n$; $PC^1([t_0,t_1],U)$ — подмножество $C([t_0,t_1],U)$, состоящее из функций с кусочно-непрерывной производной; расстояние между элементами в $PC^1([t_0,t_1],U)$ определяется как расстояние между элементами из $C([t_0,t_1],\mathbb{R}^n)$. Пусть $C_0^1([t_0,t_1],\mathbb{R}^n)$ и $PC_0^1([t_0,t_1],U)$ — подмножества $C^1([t_0,t_1],\mathbb{R}^n)$ и $PC^1([t_0,t_1],U)$, состоящие из функций, обращающихся в ноль в точках t_0 и t_1. Для краткости положим $C^1([t_0,t_1],\mathbb{R}) = C^1([t_0,t_1])$, $C_0^1([t_0,t_1],\mathbb{R}) = C_0^1([t_0,t_1])$ и $PC^1([t_0,t_1],\mathbb{R}) = PC^1([t_0,t_1])$, (**[Z], [KF]**).

Через $L(\mathbb{R}^n;\mathbb{R}^m)$ обозначаем класс всех линейных отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$; *аффинным отображением* из $\mathbb{R}^n$ в $\mathbb{R}^m$ называем сумму постоянного и линейного отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$. Пусть V окрестность точки $\widehat{x}$ в $\mathbb{R}^n$ и $F: V\to\mathbb{R}^m$ отображение из V в $\mathbb{R}^m$. Говорят, что F *дифференцируемо в точке* $\widehat{x}$, если существует отображение $A\in L(\mathbb{R}^n;\mathbb{R}^m)$ такое, что для всякого $\varepsilon>0$ найдётся $\delta>0$, при котором $|F(x)-F(\widehat{x})-A(x-\widehat{x})|<\varepsilon|x-\widehat{x}|$, если только $|x-\widehat{x}|<\delta$; в этом случае отображение A называют производной $F(\cdot)$ в точке $\widehat{x}$ и обозначают $F'(\widehat{x})$. Функцию $F(\cdot)$ называют дважды дифференцируемой в точке $\widehat{x}$, если в каждой точке x некоторой окрестности $\widehat{x}$ определена первая производная $F'(x)$ и отображение $x\mapsto F'(x)$ дифференцируемо в точке $\widehat{x}$. Если $F: V\to\mathbb{R}^m$ дифференцируемо в $\widehat{x}\in V$, то производная $F'(\widehat{x})$ задаётся матрицей $\left(\frac{\partial F_i}{\partial x_j}\right)_{1\le i\le m,\ 1\le j\le n}$ (якобианом) из частных производных функции $F(\cdot)$ в точке $\widehat{x}$; в частности, если $F(\cdot)$ действует из V в $\mathbb{R}$, то производная $F(\cdot)$ — это вектор-строка из её частных производных. Если $F: V\to\mathbb{R}^m$ дважды дифференцируемо в $\widehat{x}\in V$, то вторая производная $F''(\widehat{x})$ определяется набором m-столбцов $\frac{\partial^2 F(\widehat{x})}{\partial x_j\partial x_k}$, $1\le j,k\le n$. Если для $F: V\to\mathbb{R}^m$ существует предел $\lim\limits_{\alpha\to 0}\frac{F(\widehat{x}+\alpha h)-F(\widehat{x})}{\alpha}$, $h\in\mathbb{R}^n$, $\alpha\in\mathbb{R}$, то он называется *производной по направлению* h функции $F(\cdot)$ в точке $\widehat{x}$ и обозначается $F'(\widehat{x}, h)$, (**[Z]**).

Свойства производной:

1) если $F(\cdot)$ дифференцируема в $\widehat{x}$, то для любого $h \in \mathbb{R}^n$ существует $F'(\widehat{x}, h)$ и выполняется равенство $F'(\widehat{x}, h) = F'(\widehat{x})[h]$,

2) если $F = \binom{f_1}{f_2}$ и функции $f_1(\cdot)$, $f_2(\cdot)$ дифференцируемы в точке $\widehat{x}$, то $F(\cdot)$ также дифференцируема в $\widehat{x}$ и $F'(\widehat{x})[\cdot] = \binom{f_1'(\widehat{x})}{f_2'(\widehat{x})}[\cdot]$,

3) если $F = f \circ g$, функция $g(\cdot)$ дифференцируема в точке $\widehat{x}$, а функция $f(\cdot)$ дифференцируема в точке $g(\widehat{x})$, то $F(\cdot)$ дифференцируема в $\widehat{x}$ и $F'(\widehat{x})[\cdot] = \big(f'(g(x)) \circ g'(x)\big)[\cdot]$ (теорема о суперпозиции),

4) если $F(\cdot)$ дважды дифференцируема в $\widehat{x}$, то $F(\widehat{x} + x) = F(\widehat{x}) + F'(\widehat{x})[x] + \frac{1}{2}F''(\widehat{x})[x, x] + \bar{o}(|x|^2) := F(\widehat{x}) + \sum\limits_{j=1}^{n} \frac{\partial F(\widehat{x})}{\partial x_i} x_j + \frac{1}{2} \sum\limits_{j,k=1}^{n} \frac{\partial^2 F(\widehat{x})}{\partial x_j \partial x_j} x_j x_k + \bar{o}(|x|^2)$, (**[Z]**).

Функция $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ называется *выпуклой*, если её *надграфик* $\operatorname{epi} f = \{\binom{x}{z} : \ x \in \mathbb{R}^n, \ z \geq f(x)\}$ — выпуклое множество в $\mathbb{R}^{n+1}$; выпуклая функция $f(\cdot)$ называется *замкнутой*, если её надграфик — замкнутое множество. Функция $y \mapsto f^*(y) = \sup(y \cdot x - f(x))$, $y \in (\mathbb{R}^n)^*$ называется *преобразованием Юнга–Фенхеля* функции $f(\cdot)$ (либо *первой сопряжённой* к $f(\cdot)$ функцией); *второй сопряжённой* к $f(\cdot)$ функцией называют отображение $x \mapsto f^{**}(x) = \sup(y \cdot x - f^*(y))$, $x \in \mathbb{R}^n$. Для выпуклой функции $f(\cdot)$ понятие производной в точке $\widehat{x}$ заменяется понятием *субдифференциала* $\partial f(\widehat{x})$: так называется множество векторов $y \in (\mathbb{R}^n)^*$, для которых $f(x) - f(\widehat{x}) \geq y \cdot (x - \widehat{x})$, (**[Z], [KF]**).

Теорема об отделимости выпуклых множеств в $\mathbb{R}^n$. Если A и B — непустые, выпуклые и непересекающиеся подмножества в $\mathbb{R}^n$, то они *отделимы* некоторым ненулевым линейным функционалом $y \in \mathbb{R}^{n*}$: $\inf\limits_{x \in A} y \cdot x \geq \sup\limits_{x \in B} y \cdot x$. Если же дополнительно множество $A = \{x_0\}$ одноточечно, а множество B замкнуто, то A и B *строго отделимы* некоторым линейным функционалом $y \in (\mathbb{R}^n)^*$: $y \cdot x_0 > \sup\limits_{x \in B} y \cdot x$, (**[KF]**).

Рангом матрицы называют максимальное число её линейно независимых столбцов. *Теорема о ранге матрицы* утверждает, что *ранг матрицы есть максимальный порядок отличного от нуля минора этой матрицы.*

Пусть X_1, X_2 — векторные пространства. Тогда произведение $X_1 \times X_2$ этих пространств состоит из всевозможных пар $\binom{x_1}{x_2}$, где $x_1 \in X_1$, $x_2 \in X_2$; если пространства X_1 и X_2 нормированы, то их произведение является нормированным пространством с нормой $\|\binom{x_1}{x_2}\|_{X_1 \times X_2} = \max\{\|x_1\|_{X_1}, \|x_2\|_{X_2}\}$; если в пространствах X_1 и X_2 задана топология, то топология в их произведении порождается семейством множеств $U_1 \times U_2$, где U_1 открыто в X_1, а U_2 открыто в X_2. *Теорема (об общем виде линейного функционала на произведении пространств)*: если X_1, X_2 — векторные пространства, z — линейный функционал на произведении $X_1 \times X_2$ этих пространств, то существуют линейный функционал x_1^* на X_1 и линейный функционал x_2^* на X_2 такие, что $z = (x_1^*, x_2^*)$, т.е. $z\binom{x_1}{x_2} = x_1^*(x_1) + x_2^*(x_2)$ для всех $\binom{x_1}{x_2} \in X_1 \times X_2$. Чтобы указать вектор в произведении пространств X_1 и X_2 по заданным компонентам $x_1 \in X_1$ и $x_2 \in X_2$ употребляется конструкция $\binom{x_1}{x_2}$, (**[Z], [KF]**).

Теорема о неявной функции. Пусть $U \subset \mathbb{R}^k \times \mathbb{R}^s$ – окрестность точки $\binom{\widehat{x}}{\widehat{y}}$, $\psi \in C^1(U, \mathbb{R}^s)$, матрица $\frac{\partial \psi(\widehat{x}, \widehat{y})}{\partial y}$ невырождена. Тогда найдутся окрестность $U_1 \subset U$ точки $\binom{\widehat{x}}{\widehat{y}}$ и окрестность $V \subset \mathbb{R}^k \times \mathbb{R}^s$ точки $\binom{\widehat{x}}{\psi(\widehat{x}, \widehat{y})}$ такие, что отображение $\varphi : \binom{x}{y} \mapsto \binom{x}{\psi(x,y)}$ является диффеоморфизмом между U_1 и V. При этом $(\varphi|_{U_1})^{-1} \in C^1(V, U_1)$. Если $k = 0$, то теорему о неявной функции называем теоремой об обратном отображении, (**[Z]**).

Теорема о дифференцировании собственного интеграла по параметру. Если функция $\binom{x}{\vartheta} \mapsto f(x, \vartheta)$ и её частная производная $\frac{\partial f}{\partial \vartheta}$ непрерывны в прямоугольнике $[a, b] \times [\alpha, \beta]$, то функция $\varphi(\cdot)$, определяемая равенством $\varphi(\vartheta) = \int_a^b f(x, \vartheta)\, dx$, дифференцируема на отрезке $[\alpha, \beta]$ и $\varphi'(\vartheta) = \int_a^b \frac{\partial f(x,\vartheta)}{\partial \vartheta}\, dx$, (**[Z]**).

Теорема о существовании и единственности решения задачи Коши линейного дифференциального уравнения. Пусть $[t_0, t_1] \subset \mathbb{R}$, $A : [t_0, t_1] \to L(\mathbb{R}^n, \mathbb{R}^n)$, $b : [t_0, t_1] \to \mathbb{R}^n$ интегрируемы на $[t_0, t_1]$.

Тогда задача Коши для линейного уравнения $\dot{x} = A(t)x + b(t)$ с начальным условием $x(t_0) = \xi$ имеет единственное решение и это решение продолжается на весь отрезок $[t_0, t_1]$, (**[ATF]**).

Пусть U открыто в $\mathbb{R}^{1+n}$, функция $f : U \to \mathbb{R}^n$, зависящая от переменных $t \in \mathbb{R}$, $x \in \mathbb{R}^n$, и её частная производная $\frac{\partial f}{\partial x}$ непрерывны в U. Тогда

1) для любой точки $(t_0, x_0) \in U$ в некоторой её окрестности $(\alpha, \beta) \times V \subset U$ однозначно определено отображение $\phi : (\alpha, \beta) \times V \to C([\alpha, \beta], \mathbb{R}^n)$, сопоставляющее каждой паре $\binom{\tau}{\xi}$ решение задачи Коши $\dot{x} = f(t, x)$, $x(\tau) = \xi$, ограниченное на отрезок $[\alpha, \beta]$. Отображение ϕ непрерывно. Это утверждение называем *локальной теоремой существования, единственности и непрерывной зависимости от начальных данных для задачи Коши*,

2) если $\widehat{x} \in C^1([t_0, t_1], \mathbb{R}^n)$ — решение дифференциального уравнения $\dot{x} = f(t, x)$ и кривая $\{\binom{t}{\widehat{x}(t)} \mid t \in [t_0, t_1]\}$ лежит в U, то в некоторой окрестности $(\alpha, \beta) \times V \subset U$ точки $\binom{t_0}{\widehat{x}(t_0)}$ однозначно определено отображение $\phi : (\alpha, \beta) \times V \to C([t_0, t_1], \mathbb{R}^n)$, сопоставляющее паре $\binom{\tau}{\xi}$ решение задачи Коши $\dot{x} = f(t, x)$, $x(\tau) = \xi$, ограниченное на отрезок $[t_0, t_1]$. При этом

отображение ϕ непрерывно (это утверждение называем *глобальной теоремой существования, единственности и непрерывной зависимости от начальных данных для задачи Коши*),

отображение ϕ, рассматриваемое как отображение из $(\alpha, \beta) \times V \times [t_0, t_1]$ в $\mathbb{R}^n$, непрерывно дифференцируемо по $\xi \in V$, функция $y(\cdot) = \frac{\partial \phi}{\partial \xi}\big|_{\tau = t_0, \xi = \widehat{x}(t_0)}$ принадлежит $L(\mathbb{R}^n, C^1([t_0, t_1], \mathbb{R}^n))$ и выполняется равенство (называемое ещё уравнением в вариациях) $\dot{y}(t) = f_x(t, \widehat{x}(t)) \cdot y(t)$ (это утверждение называем *теоремой о дифференцируемой зависимости от начальных данных для задачи Коши*), (**[P]**).

Пусть U открыто в $\mathbb{R}^{1+n+k}$, функция $f : U \to \mathbb{R}^n$, зависящая от переменных $t \in \mathbb{R}$, $x \in \mathbb{R}^n$ и параметра $\mu \in \mathbb{R}^k$, а также её частная производная $\frac{\partial f}{\partial x}$ непрерывны в U. Если $\widehat{x} \in C^1([t_0, t_1], \mathbb{R}^n)$ удовлетворяет дифференциальному уравнению $\dot{x} = f(t, x, \mu)$ при $\mu = \mu_0$, а кривая $\left\{\begin{pmatrix} t \\ \widehat{x}(t) \\ \mu_0 \end{pmatrix} \middle| \, t \in [t_0, t_1]\right\}$ лежит в U, то в некоторой окрестности V точки μ_0 однозначно определено отображение $\phi :$

$V \to C([t_0, t_1], \mathbb{R}^n)$, сопоставляющее параметру μ решение задачи Коши $\dot{x} = f(t, x, \mu)$, $x(t_0) = \widehat{x}(t_0)$, ограниченное на отрезок $[t_0, t_1]$. При этом

отображение ϕ непрерывно (это утверждение называем *теоремой о непрерывной зависимости от параметра для задачи Коши*),

если дополнительно в U существует непрерывная производная $\frac{\partial f}{\partial \mu}$, то отображение $\phi(\cdot)$, рассматриваемое как отображение из $V \times [t_0, t_1]$ в $\mathbb{R}^n$ непрерывно дифференцируемо по параметру $\mu \in V$, функция $y(\cdot) = \frac{\partial \phi}{\partial \mu}\big|_{\mu=\mu_0}$ принадлежит $L(\mathbb{R}^k, C^1([t_0, t_1], \mathbb{R}^n))$ и выполняется равенство $\dot{y}(t) = f_x(t, \widehat{x}(t), \mu_0) \cdot y(t) + f_\mu(t, \widehat{x}(t), \mu_0)$ (уравнение в вариациях). Это утверждение называем *теоремой о дифференцируемой зависимости от параметра для задачи Коши*, (**[P]**).

Лемма о скруглении углов. Пусть $L \in C(\mathbb{R}^{1+2n}, \mathbb{R})$ зависит от $t \in [t_0, t_1]$ и $x, \dot{x} \in \mathbb{R}^n$, функционал $\mathcal{J} : PC^1([t_0, t_1], \mathbb{R}^n) \to \mathbb{R}$ задаётся формулой $\mathcal{J}(x(\cdot)) = \int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt$, $\widehat{x} \in PC^1([t_0, t_1], \mathbb{R}^n)$. Тогда найдётся последовательность $x_k \in C^1([t_0, t_1], \mathbb{R}^n)$, $k \in \mathbb{N}$, сходящаяся к $\widehat{x}(\cdot)$ в пространстве $C([t_0, t_1], \mathbb{R}^n)$, удовлетворяющая условиям $x_k(t_0) = \widehat{x}(t_0)$, $x_k(t_1) = \widehat{x}(t_1)$ при любом $k \in \mathbb{N}$ и такая, что последовательность $\mathcal{J}(x_k(\cdot))$, $k \in \mathbb{N}$, сходится к $\mathcal{J}(\widehat{x}(\cdot))$, (**[ATF]**).

Доказательства предварительных сведений можно найти в книгах

[Z] В.А. Зорич, Математический анализ, ч. 1-2, изд. 4, Москва, МЦНМО, 2002

[KF] А.Н. Колмогоров, С.В. Фомин, Элементы теории функций и функционального анализа, изд. 5, Москва, "Наука", 1981

[P] Л.С. Понтрягин, Обыкновенные дифференциальные уравнения, Москва, "Наука", 1983

[ATF] В.М. Алексеев, В.М. Тихомиров, С.В. Фомин, Оптимальное управление, Москва, "Наука", 1979

С большей частью предварительных сведений можно ознакомиться по книге **[ATF]**.

1. Теорема Ферма и правило множителей Лагранжа для гладких задач

Необходимое условие экстремума для гладкой задачи без ограничений

Рассмотрим задачу без ограничений

$$f(x) \to \text{extr} \ \ (f : \mathbb{R}^n \to \mathbb{R}). \qquad (P_{1_0})$$

Предложение 1. *Пусть в задаче (P_{1_0}) функция f является дифференцируемой в $\widehat{x}$. Тогда, если $\widehat{x}$ является локальным экстремумом в (P_{1_0}), то выполнено условие стационарности:*

$$f'(\widehat{x}) = 0 \ (\Leftrightarrow \frac{\partial f(\widehat{x})}{\partial x_1} = \ldots = \frac{\partial f(\widehat{x})}{\partial x_n} = 0.) \qquad (1_0)$$

Этот результат называют *теоремой Ферма.* Он восходит к Ферма (1638) (см. [1]).

Доказательство. Если допустить, что, скажем, $\frac{\partial f(\widehat{x})}{\partial x_1} \neq 0$, то по определению производной функции одного переменного $f(\widehat{x}_1 + \vartheta, \widehat{x}_2, \ldots, \widehat{x}_n) = f(\widehat{x}) + \vartheta \frac{\partial f(\widehat{x})}{\partial x_1} + o(\vartheta)$, $\vartheta \in \mathbb{R}$, и тогда, выбрав ϑ достаточно малым и соответствующего знака, убеждаемся, что f не имеет в $\widehat{x}$ ни локального минимума, ни локального максимума. Противоречие доказывает предложение 1. $\square$

Необходимое условие экстремума в гладкой задаче с ограничениями в виде равенств

Рассмотрим задачу с равенствами

$$f_0(x) \to \text{extr}, \ f_i(x) = 0, \ 1 \leq i \leq m, \qquad (P_1)$$

$f_i : \mathbb{R}^n \to \mathbb{R}$, $0 \leq i \leq m$. Функция Лагранжа этой задачи имеет вид: $\mathcal{L}(x, \overline{\lambda}) = \sum_{i=0}^{m} \lambda_i f_i(x)$.

Теорема 1. *Пусть в задаче (P_1) функции f_i, $0 \leq i \leq m$, непрерывно дифференцируемы в окрестности $\widehat{x}$. Тогда необходимое условие локального экстремума в задаче (P_1) в точке $\widehat{x}$ соответствует принципу Лагранжа.* А именно, найдётся вектор $\overline{\lambda} = (\lambda_0, \ldots, \lambda_m) \neq 0$ множителей Лагранжа, такой что для функции Лагранжа (в соответствии с предложением 1) выполнено условие стационарности:

$$\mathcal{L}_x(\widehat{x}, \overline{\lambda}) = 0 \Leftrightarrow \sum_{i=0}^m \lambda_i f_i'(\widehat{x}) = 0. \tag{1}$$

Этот результат называют *правилом множителей Лагранжа* (Лагранж, 1797, [3]).

Доказательство. Соотношение (1) означает, что векторы $\{f_i'(\widehat{x})\}_{i=0}^m$ линейно зависимы. Покажем, что их линейная независимость ведёт к противоречию. Действительно, если эти вектора линейно независимы, то, по теореме о ранге матрицы, $m+1 \leq n$ и один из миноров матрицы $(\frac{\partial f_i(\widehat{x})}{\partial x_j})_{0 \leq i \leq m, 1 \leq j \leq n}$ порядка $m+1$ отличен от нуля. Будем считать, например, что матрица $(\frac{\partial f_i(\widehat{x})}{\partial x_j})_{0 \leq i \leq m, 1 \leq j \leq m+1}$ невырождена. Тогда отображение $\Psi(x_1, \ldots, x_{m+1}) = \mathcal{F}(x_1, \ldots, x_{m+1}, \widehat{x}_{m+2}, \ldots, \widehat{x}_n)$, где $\mathcal{F} = (f_0, \ldots, f_m)^T$, удовлетворяет теореме об обратной функции и, значит, в любой окрестности точки $\widehat{x}$ при любом малом по модулю действительном α разрешима система уравнений $f_0(x) = f_0(\widehat{x}) + \alpha, f_1(x) = \ldots = f_m(x) = 0$, т. е. $\widehat{x}$ – не локальный экстремум. $\square$

2. Уравнение Эйлера для простейшей задачи и необходимое условие экстремума для изопериметрической задачи вариационного исчисления

Необходимое условие экстремума в простейшей задаче

Рассмотрим задачу

$$\mathcal{J}(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \to \text{extr}, \ x(t_j) = x_j, \, j = 0, 1, \, (P_{2_0})$$

$L : \mathbb{R}^3 \to \mathbb{R}$ (её называют *простейшей задачей* вариационного исчисления).

Предложение 2. *Пусть в задаче* (P_{2_0}) *функция* L *непрерывно-дифференцируема в окрестности расширенного графика* $\Gamma_{\widehat{x}(\cdot)} = \{(t, \widehat{x}(t), \dot{\widehat{x}}(t))^T \in \mathbb{R}^3 \mid t \in [t_0, t_1]\}$ *непрерывно-дифференцируемой функции* $\widehat{x}(\cdot)$*. Тогда, если* $\widehat{x}(\cdot)$ *доставляет локальный экстремум задаче* (P_{2_0}) *в пространстве* $C^1([t_0, t_1])$*, то функция* $t \mapsto \widehat{L}_{\dot{x}}(t)$ *непрерывно-дифференци-*

руема и выполнено следующее уравнение (уравнение Эйлера):

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0, \qquad (2_0)$$

где $\widehat{L}_{\dot{x}}(t) = L_{\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$, $\widehat{L}_x(t) = L_x(t, \widehat{x}(t), \dot{\widehat{x}}(t))$ (Эйлер, 1744, см. [2]).

Доказательство. Пусть $v \in C_0^1([t_0, t_1])$. Тогда функция $\widehat{x}(\cdot) + \vartheta v(\cdot)$ допустима при любом $\vartheta \in \mathbb{R}$. Положим $\psi(\vartheta, v(\cdot)) = \mathcal{J}(\widehat{x}(\cdot) + \vartheta v(\cdot))$. По одномерной теореме Ферма, применённой к отображению $\psi(\vartheta, v(\cdot))$ в точке $\vartheta = 0$, получим (пользуясь теоремой о дифференцировании интеграла по параметру): $0 = \frac{d\psi(\vartheta, v(\cdot))}{d\vartheta}|_{\vartheta=0} = \int_{t_0}^{t_1} (\widehat{L}_{\dot{x}}(t)\dot{v}(t) + \widehat{L}_x(t)v(t))dt$. После интегрирования по частям получаем: $0 = \int_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}}(t) + \int_t^{t_1} \widehat{L}_x(\tau)d\tau\big)\dot{v}(t)dt = \int_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}}(t) + \int_t^{t_1} \widehat{L}_x(\tau)d\tau - c\big)\dot{v}(t)dt$ при любом значении $c \in \mathbb{R}$. Если выбрать функцию $v(\cdot)$ и константу c так, чтобы $\dot{v}(t) = (\widehat{L}_{\dot{x}}(t) + \int_t^{t_1} \widehat{L}_x(\tau)d\tau - c)$, $v(t_0) = 0$ и $\int_{t_0}^{t_1} \dot{v}(t)dt = 0$, то получим, что $v(\cdot)$ допустима и $\int_{t_0}^{t_1} (\widehat{L}_{\dot{x}}(t) + \int\limits_t^{t_1} \widehat{L}_x(\tau)d\tau - c)^2 dt = 0$, откуда следует, что $\widehat{L}_{\dot{x}}(t) + \int\limits_t^{t_1} \widehat{L}_x(\tau)d\tau - c \equiv 0$ и непрерывная дифференцируемость $\widehat{L}_{\dot{x}}(t)$. После дифференцирования приходим к уравнению Эйлера. □

Замечание. По такой же схеме выводятся необходимые условия экстремума для простейшей векторной задачи, когда аргументы $x, \dot{x}$ функции L принадлежат $\mathbb{R}^n$ (см. также теорему 3).

Необходимое условие экстремума в изопериметрической задаче вариационного исчисления

Рассмотрим *изопериметрическую задачу* вариационного исчисления

$$\mathcal{J}_0(x(\cdot)) \to \text{extr}, \ \mathcal{J}_i(x(\cdot)) = 0, \ 1 \le i \le m, \ x(t_j) = x_j, j = 0, 1, \qquad (P_2)$$

где $\mathcal{J}_i(x(\cdot)) = \int_{t_0}^{t_1} L_i(t, x(t), \dot{x}(t))dt$, $L_i : \mathbb{R}^3 \to \mathbb{R}$, $0 \le i \le m$. Функция Лагранжа этой задачи имеет вид: $\mathcal{L}(x, \overline{\lambda}) = \sum_{i=0}^m \lambda_i \mathcal{J}_i(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt$ и $L = \sum_{i=0}^m \lambda_i L_i$.

Теорема 2. *Пусть в задаче* (P_2) *функции* L_i, $0 \le i \le m$, *непрерывно дифференцируемы в окрестности расширенного графика* $\Gamma_{\widehat{x}(\cdot)} = \{(t, \widehat{x}(t), \dot{\widehat{x}}(t))^T \in \mathbb{R}^3 \mid t \in [t_0, t_1]\}$ *непрерывно-дифференцируемой функции* $\widehat{x}(\cdot)$. *Тогда, если функция* $\widehat{x}(\cdot)$ *доставляет локальный экстремум задаче* (P_2) *в пространстве* $C^1([t_0, t_1])$, *то необходимое условие локального экстремума в точке* $\widehat{x}(\cdot)$ *соответствует принципу Лагранжа.* А именно, найдётся вектор $\overline{\lambda} = (\lambda_0, \ldots, \lambda_m) \neq 0$ множителей Лагранжа, что для функции Лагранжа (в соответствии с предложением 2) выполнено уравнение Эйлера:

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0, \qquad (2)$$

где $\widehat{L}_{\dot{x}}(t) = L_{\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$, $\widehat{L}_x(t) = L_x(t, \widehat{x}(t), \dot{\widehat{x}}(t))$ (Эйлер, 1744, см. [2]).

Доказательство. Для линейного отображения $\Lambda(\cdot)$ из пространства $C_0^1([t_0, t_1])$ в $\mathbb{R}^{m+1}$,

$$\Lambda v(\cdot) = (\mathcal{J}_0'(\widehat{x}(\cdot))[v(\cdot)], \ldots, \mathcal{J}_m'(\widehat{x}(\cdot))[v(\cdot)])^T, \ v(\cdot) \in C_0^1([t_0, t_1]),$$

$$\mathcal{J}_i'(\widehat{x}(\cdot))[v(\cdot)] = \int_{t_0}^{t_1} ((\widehat{L}_i)_{\dot{x}}(t) \cdot \dot{v}(t) + (\widehat{L}_i)_x(t) \cdot v(t))dt\,,$$

возможно одно из двух: образ оператора Λ есть собственное подпространство $\mathbb{R}^{m+1}$ либо всё пространство $\mathbb{R}^{m+1}$. В первом случае найдётся ненулевой вектор $y \in \mathbb{R}^{m+1}$, ортогональный этому подпространству; пусть $\overline{\lambda} = (\lambda_0, \ldots, \lambda_m) = y^T$. Тогда $\overline{\lambda} \cdot \Lambda v(\cdot) = \langle y, \Lambda v(\cdot)\rangle = \int_{t_0}^{t_1} (\widehat{L}_{\dot{x}}(t) \cdot \dot{v}(t) + \widehat{L}_x(t) \cdot v(t))dt = 0 \ \forall v(\cdot) \in C_0^1([t_0, t_1])$. Отсюда (как и в предложении 2) выводится уравнение Эйлера (2). Во втором случае найдётся подпространство $M \subset C_0^1([t_0, t_1])$, $\dim M = m + 1$, такое что $\Lambda(M) = \mathbb{R}^{m+1}$. Это означает, что $\Lambda\big|_M$ — изоморфизм между M и $\mathbb{R}^{m+1}$. Функция $\widehat{x}(\cdot) + v(\cdot)$ допустима в задаче (P_2) при любом $v(\cdot) \in M$. Пусть $\vartheta \in \mathbb{R}^{m+1}$, $\psi(\vartheta) = (\mathcal{J}_0(\widehat{x}(\cdot) + v(\cdot)), \ldots, \mathcal{J}_m(\widehat{x}(\cdot) + v(\cdot)))^T$, где $v(\cdot) = (\Lambda\big|_M)^{-1}(\vartheta)$. По теореме о дифференцировании интеграла по параметру (и определению производной по направлению) $\psi'(0)[\vartheta] = \Lambda v(\cdot) = \vartheta$, т.е. $\psi'(0)$ — тождественный оператор; по теореме об обратной функции, применённой к $\psi(\cdot)$ в точке $\vartheta = 0$, в любой окрестности $\widehat{x}(\cdot)$ при любом достаточно малом по модулю ε возможно разрешить систему уравнений $\mathcal{J}_0(x(\cdot)) = \mathcal{J}_0(\widehat{x}(\cdot))+\varepsilon$, $\mathcal{J}_1(x(\cdot)) = \ldots = \mathcal{J}_m(x(\cdot)) = 0$, а это противоречит тому, что $\widehat{x}(\cdot)$ — локальный экстремум задачи. $\square$

3. Уравнение Эйлера и условия трансверсальности в задаче Больца и необходимое условие экстремума для задачи Лагранжа вариационного исчисления

Необходимое условие экстремума в задаче Больца

Рассмотрим задачу вариационного исчисления без ограничений (*задача Больца*):

$\mathcal{B}(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt + \ell(x(t_0), x(t_1)) \to \mathrm{extr}$, (P_{3_0})
где $L : \mathbb{R}^{2n+1} \to \mathbb{R}$, $l : \mathbb{R}^{2n} \to \mathbb{R}$.

Предложение 3. *Пусть в задаче* (P_{3_0}) *функция* l *непрерывно дифференцируема в окрестности точки* $\binom{\widehat{x}(t_0)}{\widehat{x}(t_1)}$, *функция* L *непрерывно дифференцируема в окрестности расширенного графика* $\left\{\begin{pmatrix} t \\ \widehat{x}(t) \\ \dot{\widehat{x}}(t) \end{pmatrix} \in \mathbb{R}^{2n+1} \mid t \in [t_0, t_1]\right\}$ *функции* $\widehat{x}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$. *Тогда, если* $\widehat{x}(\cdot)$ *доставляет локальный экстремум задаче* (P_{3_0}) *в* $C^1([t_0, t_1], \mathbb{R}^n)$, *то отображение* $t \mapsto \widehat{L}_{\dot{x}}(t)$ *непрерывно-дифференцируемо и выполнены* a) *уравнение Эйлера и* b) *условие трансверсальности*:

a) $-\frac{d}{dt}\widehat{L}_{\dot{x}}(t)+\widehat{L}_x(t) = 0$, *b)* $\widehat{L}_{\dot{x}}(t_i) = (-1)^i \widehat{\ell}_{x(t_i)}$, $i = 0, 1$, (3_0)
где $\widehat{L}_{\dot{x}}(t) = L_{\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$, $\widehat{L}_x(t) = L_x(t, \widehat{x}(t), \dot{\widehat{x}}(t))$, $\widehat{\ell}_{x(t_i)} = \frac{\partial \ell(\widehat{x}(t_0), \widehat{x}(t_1))}{\partial(x(t_i))}$, $i = 0, 1$.

Доказательство (при $n = 1$). Выбрав $v(\cdot) \in C^1([t_0, t_1])$, полагаем $\psi(\vartheta, v(\cdot)) = \mathcal{B}(\widehat{x}(\cdot) + \vartheta v(\cdot))$. Условия гладкости в предложении 3 и теоремы о дифференцировании интеграла по параметру позволяют дифференцировать эту функцию по ϑ в точке 0. Применив предложение 1, получаем:

$0 = \frac{d\psi(\vartheta, v(\cdot))}{d\vartheta}\Big|_{\vartheta=0} = \int\limits_{t_0}^{t_1}(\widehat{L}_{\dot{x}}(t)\dot{v}(t) + \widehat{L}_x(t)v(t))dt + \widehat{\ell}_{x(t_0)}v(t_0) + \widehat{\ell}_{x(t_1)}v(t_1).$ (i)

Решив задачу Коши: $\dot{p} = \widehat{L}_x(t),\ p(t_1) = -\widehat{\ell}_{x(t_1)}$, т. е. положив $p(t) = -(\widehat{\ell}_{x(t_1)} + \int_t^{t_1} \widehat{L}_x(\tau)d\tau)$, подставляем $\dot{p}(t)$ вместо $\widehat{L}_x(t)$ в (i), интегрируем по частям и, учитывая, что $p(t_1) = -\widehat{\ell}_{x(t_1)}$, получаем:

$$\int\limits_{t_0}^{t_1}(\widehat{L}_{\dot{x}}(t) - p(t))\dot{v}(t)dt + (\widehat{\ell}_{x(t_0)} - p(t_0))v(t_0) = 0 \qquad (ii)$$

при любом $v(\cdot) \in C^1([t_0, t_1])$. А положив $v(t) = \widehat{\ell}_{x(t_0)} - p(t_0) + \int_{t_0}^{t}(\widehat{L}_{\dot{x}}(\tau) - p(\tau))d\tau$ и подставив это выражение в (ii), получаем: $\int\limits_{t_0}^{t_1}(\widehat{L}_{\dot{x}}(t) - p(t))^2dt + (\widehat{\ell}_{x(t_0)} - p(t_0))^2 = 0$, откуда приходим к соотношениям $p(t) = \widehat{L}_{\dot{x}}(t)$, $\dot{p}(t) = \widehat{L}_x(t)$, $p(t_1) = -\widehat{\ell}_{x(t_1)}$, $p(t_0) = \widehat{\ell}_{x(t_0)}$, которые равнозначны соотношениям (3_0). Случай $n > 1$ рассматривается аналогично. □

Замечание. *Доказательство предложения 3, которое может показаться вычурным и немотивированным, пролагает путь к доказательству общего результата, приводимого ниже.*

Необходимое условие экстремума в задаче Лагранжа в понтрягинской форме

Рассмотрим задачу

$$\mathcal{J}(x(\cdot), u(\cdot)) \to \mathrm{extr},\ \dot{x} = \varphi(t, x(t), u(t)), x(t_i) = x_i,\ i = 0, 1, \qquad (P_3)$$

в которой $\mathcal{J}(x(\cdot), u(\cdot)) = \int\limits_{t_0}^{t_1} f(t, x(t), u(t))dt$, $f : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r \to \mathbb{R}$, $\varphi : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r \to \mathbb{R}^n$. Будем называть её *задачей Лагранжа в понтрягинской форме*. Функция Лагранжа этой задачи имеет вид: $\mathcal{L}(x(\cdot), u(\cdot), \overline{\lambda}) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t), u(t))dt$, где $L(t, x, \dot{x}, u) = \lambda_0 f(t, x, u) + p(t)\cdot(\dot{x} - \varphi(t, x, u))$, $\overline{\lambda} = (\lambda_0, p(\cdot)) \in \mathbb{R}\times C^1([t_0, t_1], (\mathbb{R}^n)^*)$.

Теорема 3. *Пусть в задаче* (P_3) *функции* f *и* φ *непрерывно дифференцируемы в окрестности кривой* $\left\{ \begin{pmatrix} t \\ \widehat{x}(t) \\ \widehat{u}(t) \end{pmatrix} \in \mathbb{R}^{1+n+r} \mid t \in [t_0, t_1] \right\}$. *Тогда необходимое условие локального экстремума в точке* $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)}$ *из пространства* $C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$ *в задаче* (P_3) *соответствует принципу Лагранжа.* А именно, найдётся ненулевой вектор $\overline{\lambda} = (\lambda_0, p(\cdot))$ множителей Лагранжа, что для функции Лагранжа выполнены (в соответствии с условием экстремума для простейшей задачи) a) уравнение Эйлера по $x(\cdot)$ и a') уравнение Эйлера по $u(\cdot)$:

$$a) -\dot{p}(t) = p(t) \cdot \widehat{\varphi}_x(t) - \lambda_0 \widehat{f}_x(t)\,,\ a')\ 0 = p(t) \cdot \widehat{\varphi}_u(t) - \lambda_0 \widehat{f}_u(t) \quad (3)$$

$\big($здесь $\widehat{f}_x(t) = f_x(t, \widehat{x}(t), \widehat{u}(t)) \in (\mathbb{R}^n)^*$ при любом $t \in [t_0, t_1]$. Аналогично, при $t \in [t_0, t_1]$ $\widehat{f}_u(t) = f_u(t, \widehat{x}(t), \widehat{u}(t)) \in (\mathbb{R}^r)^*$, $\widehat{\varphi}_x(t) = \varphi_x(t, \widehat{x}(t), \widehat{u}(t)) \in L(\mathbb{R}^n, \mathbb{R}^n)$, $\widehat{\varphi}_u(t) = \varphi_u(t, \widehat{x}(t), \widehat{u}(t)) \in L(\mathbb{R}^r, \mathbb{R}^n)\big)$. Соотношение a) называется *сопряжённым уравнением по* x, соотношение a') называется *сопряжённым уравнением по* u.

Доказательство. Пусть пара $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)}$ доставляет локальный экстремум задаче (P_3). Рассмотрим любую функцию $v(\cdot) \in C([t_0, t_1], \mathbb{R}^r)$ и положим $u_\vartheta(t) = u_\vartheta(t, v(\cdot)) = \widehat{u}(t) + \vartheta v(t)$, $\vartheta \in \mathbb{R}$. Функцию $u_\vartheta(\cdot)$ называют *вариацией* функции $\widehat{u}(\cdot)$. Если ϑ достаточно мало по модулю, то из теоремы о дифференцируемой зависимости решения задачи Коши от параметра следует, что на $[t_0, t_1]$ существует единственное решение дифференциального уравнения: $\dot{x} = \varphi(t, x, u_\vartheta(t))$ с начальным условием $x(t_0) = x_0$. При $t \in [t_0, t_1]$ его значения будем обозначать через $x_\vartheta(t) = x_\vartheta(t, v(\cdot))$. По теореме о дифференцируемой зависимости решения задачи Коши от параметров для $t \in [t_0, t_1]$ существует $\frac{d}{d\vartheta} x_\vartheta(t, v(\cdot))|_{\vartheta=0} = y(t, v(\cdot))$, причём функция $y(\cdot, v)$ является решением задачи Коши для линейного неоднородного уравнения $\dot{y}(t) = \widehat{\varphi}_x(t)\, y(t) + \widehat{\varphi}_u(t)\, v(t)$ с начальным условием $y(t_0) = 0$. Из теоремы о дифференцировании интеграла по параметру $\frac{d}{d\vartheta}\mathcal{J}(x_\vartheta(\cdot),u_\vartheta(\cdot))|_{\vartheta=0} = \int_{t_0}^{t_1}(\widehat{f}_x(t)\cdot y(t,v(\cdot)) + \widehat{f}_u(t)\cdot v(t))dt$. Обозначим это число $y_0(v(\cdot))$. Функции $v(\cdot)$ сопоставим вектор $\eta(v(\cdot)) = \binom{y_0(v(\cdot))}{y(t_1,v(\cdot))} \in \mathbb{R}^{n+1}$. Возможно одно из двух: линейная оболочка $\{\eta(v(\cdot))\}_{v(\cdot)\in C([t_0,t_1],\mathbb{R}^n)}$ не совпадает или совпадает с $\mathbb{R}^{n+1}$. В первом случае по теореме отделимости найдём ненулевой элемент $(\lambda_0,\lambda) \in \mathbb{R}\times(\mathbb{R}^n)^*$, для которого $\lambda_0 y_0(v(\cdot)) + \lambda\cdot y(t_1,v(\cdot)) = 0\ \forall\, v(\cdot)\in C([t_0,t_1],\mathbb{R}^n)$. Решим задачу Коши для линейного уравнения $-\dot p = p\,\widehat{\varphi}_x(t) - \lambda_0\widehat{f}_x(t),\ p(t_1) = -\lambda$. Тогда, из определения (λ_0,λ) и отображений $p(\cdot)$, $y(\cdot,v)$ получим: $0 = \int_{t_0}^{t_1}(\dot p(t) + p(t)\,\widehat{\varphi}_x(t) - \lambda_0\widehat{f}_x(t))\cdot y(t,v(\cdot))dt = -\lambda\cdot y(t_1,v(\cdot)) + \int_{t_0}^{t_1}(p(t)\cdot\big(-\widehat{\varphi}_x(t)y(t,v(\cdot)) - \widehat{\varphi}_u(t)v(t)\big) + \big(p(t)\widehat{\varphi}_x(t) - \lambda_0\widehat{f}_x(t)\big)\cdot y(t,v(\cdot)))dt = \int_{t_0}^{t_1}(\lambda_0\widehat{f}_u(t) - p(t)\widehat{\varphi}_u(t))\cdot v(t)dt$. Ввиду произвольности $v(\cdot)$ приходим ко второму сопряжённому уравнению.

Во втором случае выберем $\{v_1(\cdot),\dots,v_{n+1}(\cdot)\}$ так, чтобы система векторов $\{\eta(v_i(\cdot))\}_{i=1}^{n+1}$ образовывала бы базис в $\mathbb{R}^{n+1}$. Рассмотрим вариацию $u_{\overline{\vartheta}}(\cdot) = \widehat{u}(\cdot) + \sum_{i=1}^{n+1}\vartheta_i v_i(\cdot)$, $\overline{\vartheta} = (\vartheta_i) \in \mathbb{R}^{n+1}$, и найдем $x_{\overline{\vartheta}}(\cdot)$, являющимся решением задачи Коши $\dot x = \varphi(t,x,u_{\overline{\vartheta}}(t))$, $x(t_0) = x_0$. Теорема о непрерывной зависимости решения от параметров позволяет сделать это, если длина вектора $\overline{\vartheta}$ достаточно мала, а теорема о дифференцируемой зависимости решения от параметров приводит к тому, что отображение $g(\overline{\vartheta}) = \binom{\mathcal{J}(x_{\overline{\vartheta}}(\cdot),u_{\overline{\vartheta}}(\cdot))}{x_{\overline{\vartheta}}(t_1)}$ является непрерывно-дифференцируемым в окрестности нуля и его производная при $\overline{\vartheta} = 0$ задаёт изоморфизм $\mathbb{R}^{n+1}$ на себя. Поэтому, по теореме об обратном отображении, применённой к $g(\cdot)$ в точке $\overline{\vartheta} = 0$, в любой окрестности пары $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)}$ при любом достаточно малом по модулю $\sigma\in\mathbb{R}$ возможно решить систему уравнений $\mathcal{J}(x_{\overline{\vartheta}(\sigma)}(\cdot),u_{\overline{\vartheta}(\sigma)}(\cdot)) = \mathcal{J}(\widehat{x}(\cdot),\widehat{u}(\cdot)) + \sigma$, $x_{\overline{\vartheta}(\sigma)}(t_1) = x_1$, что противоречит локальной экстремальности $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)}$. Значит, вторая возможность ведёт к противоречию. Теорема доказана. $\square$

4. Критерий минимума для простейшей задачи и принцип максимума Понтрягина для общей (понтрягинской) задачи оптимального управления

Критерий минимума для простейшей задачи оптимального управления

Рассмотрим задачу:

$$\mathcal{I}(u(\cdot))=\int_{t_0}^{t_1} L(t,u(t))dt \to \min,\ u(t)\in U\ (L:\mathbb{R}\times\mathbb{R}^r\to\mathbb{R}).(P_{4_0})$$

Будем называть её *простейшей задачей оптимального управления.*

Предложение 4. *Пусть в задаче* (P_{4_0}) *функция* L *непрерывна на* $\mathbb{R}\times\mathbb{R}^r$, $U\subset\mathbb{R}^r$. *Тогда* $\widehat{u}(\cdot)\in PC([t_0,t_1],U)$ *является решением задачи* (P_{4_0}) *в том и только в том случае, если в любой точке непрерывности* $\widehat{u}(\cdot)$ *выполнено условие минимума:*

c) $L(t,u)\geq L(t,\widehat{u}(t))\ \forall u\in U.$ $\qquad (4_0)$

Доказательство. Если $\tau\in[t_0,t_1]$ – точка непрерывности $\widehat{u}(\cdot)$ и величина $L(\tau,u)$ не достигает минимума по $u\in U$ в точке $\widehat{u}(\tau)\in U$, то в малой окрестности точки τ, при некотором $\overline{u}\in U$ значение $L(t,\overline{u})$ меньше чем $L(t,\widehat{u}(t))$; тогда, заменив $\widehat{u}(\cdot)$ на этой окрестности значением $\overline{u}$, приходим к противоречию с тем, что $\widehat{u}(\cdot)$ является решением задачи. $\square$

Принцип максимума Понтрягина для задачи оптимального управления в понтрягинской форме

Пусть $\Xi=PC^1([t_0,t_1],\mathbb{R}^n)\times PC([t_0,t_1],\mathbb{R}^r)$, $\mathcal{J}(x(\cdot),u(\cdot))=\int_{t_0}^{t_1} f(t,x(t),u(t))dt$, где $\binom{x(\cdot)}{u(\cdot)}\in\Sigma$, $f:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r\to\mathbb{R}$, $t_0,t_1\in\mathbb{R}$. Рассмотрим задачу (которую назовём *задачей оптимального управления в понтрягинской форме*)

$$\mathcal{J}(x(\cdot),u(\cdot))\to\min,\ \dot{x}=\varphi(t,x(t),u(t)),\ x(t_0)=x_0, x(t_1)=x_1,$$
$$u(t)\in U,\ t\in[t_0,t_1], \qquad (P_4)$$

в которой $\varphi:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r\to\mathbb{R}^n$, $U\subset\mathbb{R}^r$. Её функция Лагранжа имеет вид: $\mathcal{L}(x(\cdot),u(\cdot),\overline{\lambda})=\int_{t_0}^{t_1} L(t,x(t),\dot{x}(t),u(t))dt$, где $L(t,x,\dot{x},u)=\lambda_0 f(t,x,u)+p(t)\cdot(\dot{x}-\varphi(t,x,u))$, $\overline{\lambda}=(\lambda_0,p(\cdot))\in$

$\mathbb{R}_+ \times PC^1([t_0,t_1],(\mathbb{R}^n)^*)$. Скажем, что $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)} \in \Sigma$ доставляет сильный локальный минимум задаче (P_4), если пара $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)}$ допустима и для некоторого $\delta > 0$ при всех допустимых $\binom{x(\cdot)}{u(\cdot)}$, удовлетворяющих условию $\|x(\cdot)-\widehat{x}(\cdot)\|_{C[t_0,t_1]} < \delta$, верно неравенство $\mathcal{J}(x(\cdot),u(\cdot)) \geq \mathcal{J}(\widehat{x}(\cdot),\widehat{u}(\cdot))$.

Теорема 4. *Пусть в задаче (P_4) пара $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)} \in \Sigma$, G — окрестность графика $\{\binom{t}{\widehat{x}(t)} \mid t \in [t_0,t_1]\}$ функции $\widehat{x}(\cdot)$, $f \in C(G\times U, \mathbb{R})$, $\varphi \in C(G\times U, \mathbb{R}^n)$, причём f_x, φ_x также непрерывны на $G\times U$. Тогда, если $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)}$ доставляет сильный локальный минимум задаче (P_4), то необходимое условие минимума в этой задаче в точке $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)}$ находится в соответствии с принципом Лагранжа.* А именно, найдётся ненулевой вектор $\overline{\lambda} = (\lambda_0, p(\cdot)) \in \mathbb{R}_+ \times PC^1([t_0,t_1],(\mathbb{R}^n)^*)$ множителей Лагранжа, при котором для функции Лагранжа в точках непрерывности *управления* $\widehat{u}(\cdot)$ выполнены (в соответствии с предложениями 2 и 4) a) уравнение Эйлера по $x(\cdot)$ и c) условие минимума по $u(\cdot)$:

$$a)\ -\dot{p}(t)=p(t)\widehat{\varphi}_x(t)-\lambda_0\widehat{f}_x(t),\ \ c)\ \min_{u\in U} L(t,\widehat{x}(t),\dot{\widehat{x}}(t),u)=\widehat{L}(t), \quad (4)$$

где $\widehat{L}(t) = L(t,\widehat{x}(t),\dot{\widehat{x}}(t),\widehat{u}(t))$, $\widehat{\varphi}_x(t) = \varphi_x(t,\widehat{x}(t),\widehat{u}(t))$, $\widehat{f}_x(t) = f_x(t,\widehat{x}(t),\widehat{u}(t))$, $t \in [t_0,t_1]$. В задаче со свободным правым концом (когда условие $x(t_1) = x_1$ отсутствует, к условиям a) и c) добавляется условие b) $p(t_1) = 0$ трансверсальности и $\lambda_0 = 1$.

Задачу со свободным правым концом можно свести к задаче

$$g(x(t_1)) \to \min,\ \dot{x} = \varphi(t,x(t),u(t)),\ x(t_0) = x_0, \quad u(t) \in U,\ t \in [t_0,t_1], \qquad (P_4')$$

в которой $\varphi : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r \to \mathbb{R}^n$, $U \subset \mathbb{R}^r$, $g : \mathbb{R}^n \to \mathbb{R}$. Её функцию Лагранжа запишем в виде: $\mathcal{L}(x(\cdot),u(\cdot),\overline{\lambda}) = \int_{t_0}^{t_1} L(t,x(t),\dot{x}(t),u(t))dt + g(x(t_1))$, где $L(t,x,\dot{x},u) = p(t)\cdot(\dot{x} - \varphi(t,x,u))$, $\overline{\lambda} = (1,p(\cdot)) \in \mathbb{R}_+ \times PC^1([t_0,t_1],(\mathbb{R}^n)^*)$. Теорема 4 для задачи (P_4') приобретает вид

Теорема 4′. *Пусть в задаче (P_4') пара $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)} \in \Sigma$, G*

— окрестность графика $\{\binom{t}{\widehat{x}(t)} \mid t \in [t_0, t_1]\}$ *функции* $\widehat{x}(\cdot)$, $f \in C(G \times U, \mathbb{R})$, $\varphi \in C(G \times U, \mathbb{R}^n)$, *причём* f_x, φ_x *также непрерывны на* $G \times U$, *пусть функция* $g(\cdot)$ *непрерывно дифференцируема в окрестности точки* $\widehat{x}(t_1)$. *Тогда, если* $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)}$ *доставляет сильный локальный минимум задаче* (P_4'), *то необходимое условие минимума в этой задаче в точке* $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)}$ *соответствует принципу Лагранжа.* А именно, найдётся вектор $\overline{\lambda} = (1, p(\cdot)) \in \mathbb{R}_+ \times PC^1([t_0, t_1], (\mathbb{R}^n)^*)$ множителей Лагранжа, при котором для функции Лагранжа в точках непрерывности *управления* $\widehat{u}(\cdot)$ выполнены (в соответствии с предложениями 2 и 4) a) уравнение Эйлера по $x(\cdot)$, c) условие минимума по $u(\cdot)$ и b) условие трансверсальности в правом конце:

$$a)\ -\dot{p}(t) = p(t)\widehat{\varphi}_x(t), \ \ c) \min_{u \in U} L(t, \widehat{x}(t), \dot{\widehat{x}}(t), u) = \widehat{L}(t),$$

$$b)\ p(t_1) = -g_x(\widehat{x}(t_1)), \tag{4'}$$

и $\widehat{L}(t) = L(t, \widehat{x}(t), \dot{\widehat{x}}(t), \widehat{u}(t))$, $\widehat{\varphi}_x(t) = \varphi_x(t, \widehat{x}(t), \widehat{u}(t))$, $t \in [t_0, t_1]$.

Доказательство теоремы 4$'$. Достаточно доказать теорему для точек $\tau \in (t_0, t_1)$, в которых $\widehat{u}(\cdot)$ непрерывна. Пусть $\tau_1 < \dots < \tau_m < t_1$ – все точки разрыва $\widehat{u}(\cdot)$, расположенные правее τ, $\tau_0 = \tau$, $\tau_{m+1} = t_1$ ($m = 0$, если точек разрыва нет). Для $\alpha \in \mathbb{R}_+$, $\tau \in (t_0, t_1)$, $v \in U$ положим $u_\alpha(t) = u_\alpha(t, \tau, v) = \widehat{u}(t)$, если $t \notin [\tau - \alpha, \tau]$ и $u_\alpha(t) = v$, если $t \in [\tau - \alpha, \tau]$. Функцию $u_\alpha(\cdot, \tau, v)$ называют *игольчатой вариацией.* Если $\alpha > 0$ достаточно мало, то $u_\alpha(\cdot)$ допустима в задаче.

По локальной теореме существования (и единственности) решения задачи Коши, применённой в точке $\binom{\tau}{\widehat{x}(\tau)}$ к задаче $\dot{x}(t) = \varphi(t, x(t), v)$ с начальным условием $x(\tau - \alpha) = \widehat{x}(\tau - \alpha)$, и по глобальной теореме о непрерывной зависимости решения задачи Коши от начальных данных, применённой при каждом $k \in \{0, \dots, m\}$ к уравнению $\dot{x} = \varphi(t, x, \widehat{u}(t))$ для $\widehat{x}(\cdot) : [\tau_k, \tau_{k+1}] \to \mathbb{R}^{n+1}$, получаем, что в классе $PC^1([t_0, t_1], \mathbb{R}^{n+1})$ существует решение задачи Коши $\dot{x} = \varphi(t, x, u_\alpha(t)),\ x(t_0) = x_0$, стремящееся к $\widehat{x}(\cdot)$ при $\alpha \to +0$ в метрике пространства $C[t_0, t_1]$. Обозначим это решение $x_\alpha(\cdot) = x_\alpha(\cdot, \tau, v)$.

Так как $x_\alpha(\cdot)$ равномерно стремится к $\widehat{x}(\cdot)$ при $\alpha \to +0$, функция $\varphi(\cdot)$ непрерывна, $x_\alpha(\tau) = \widehat{x}(\tau-\alpha) + \int\limits_{\tau-\alpha}^{\tau} \varphi(s, \widehat{x}(s), v)ds + \int\limits_{\tau-\alpha}^{\tau} (\varphi(s, x_\alpha(s), v) - \varphi(s, \widehat{x}(s), v))ds$, то

$$\frac{d\, x_\alpha(\tau)}{d\alpha}\Big|_{\alpha=+0} = \lim_{\alpha\to+0} \frac{x_\alpha(\tau) - \widehat{x}(\tau)}{\alpha} = \varphi(\tau, \widehat{x}(\tau), v) - \widehat{\varphi}(\tau). \quad (i)$$

Обозначим $\Delta_{\tau,v}\varphi := \varphi(\tau, \widehat{x}(\tau), v) - \widehat{\varphi}(\tau)$. По теореме о дифференцируемой зависимости решения от начальных данных на каждом отрезке $[\tau_k, \tau_{k+1}]$, $k = 0, \dots, m$, существует непрерывная производная

$$\frac{dx_\alpha(t)}{dx_\alpha(\tau_k)}\Big|_{\alpha=+0};$$

отсюда, из (i) и теоремы о суперпозиции следует, что на каждом отрезке $[\tau_k, \tau_{k+1}]$, $k = 0, \dots, m$, существует непрерывная производная $y(t) := \frac{dx_\alpha(t)}{d\alpha}\big|_{\alpha=+0}$. Из того, что на $[\tau, t_1]$ значение $y(t)$ определено однозначно и $y(\cdot)$ непрерывна на каждом из отрезков $[\tau_k, \tau_{k+1}]$, $k = 0, \dots, m$, следует, что она непрерывна на объединении этих отрезков, т.е. на $[\tau, t_1]$. По теореме о дифференцируемой зависимости решения задачи Коши от начальных данных, получаем, что $y(\cdot)$, ограниченная на любой из отрезков $[\tau_k, \tau_{k+1}]$, $k = 0, \dots, m$, удовлетворяет на нём уравнению в вариациях $\dot{y} = \widehat{\varphi}_x(t)y$, т.е. $y(\cdot) \in PC^1([\tau, t_1], \mathbb{R}^n)$ является решением задачи Коши линейного уравнения:

$$\dot{y} = \widehat{\varphi}_x(t)y,\ y(\tau) = \Delta_{\tau,v}\varphi,\ \widehat{\varphi}_x(t) \in L(\mathbb{R}^n, \mathbb{R}^n),\ t \in [\tau, t_1], \quad (ii)$$

$\Delta_{\tau,v}\varphi = \varphi(\tau, \widehat{x}(\tau), v) - \widehat{\varphi}(\tau)$.

Пусть теперь $p(\cdot) \in PC^1([t_0, t_1], (\mathbb{R}^n)^*)$ удовлетворяет условию (4) a) и условию b) трансверсальности на правом конце (существование и единственность такой функции следует из теоремы существования для линейных систем):

$$a)\ -\dot p(t) = p(t)\widehat{\varphi}_x(t),\ b)\ p(t_1) = -g_x(\widehat{x}(t_1)), \qquad (iii)$$

так что остается доказать, что выполнено условие минимума c), для чего достаточно учесть, что $\frac{d}{d\alpha}g(x_\alpha(t_1))|_{\alpha=+0} \geq 0$, поскольку $\binom{\widehat{x}(\cdot)}{\widehat{u}(\cdot)}$ — решение задачи, а пара $\binom{x_\alpha(\cdot)}{u_\alpha(\cdot)}$ допустима:

$$\frac{d}{dt}(p(t)\cdot y(t)) \overset{\mathrm{Id}}{=} \dot p(t)\cdot y(t) + p(t)\cdot \dot y(t) \overset{(ii),(iii)}{=} 0, \qquad (iv)$$

откуда, если воспользоваться граничными условиями для $p(\cdot)$ в точке t_1 и $y(\cdot)$ в точке τ, получим:

$0 \leq \frac{d}{d\alpha}g(x_\alpha(t_1))|_{\alpha=+0} = g_x(\widehat{x}(t_1))y(t_1) \overset{(iv)}{=} \int\limits_\tau^{t_1} \frac{d}{dt}(p(t)\cdot y(t))\,dt +$ $g_x(\widehat{x}(t_1))y(t_1) = -p(\tau)\Delta_{\tau,v}\varphi$, а в этом и состоит условие минимума. □

5. Критерий минимума для выпуклой задачи без ограничений и теорема Каруша–Куна–Таккера (правило множителей Лагранжа) для выпуклых задач

Критерий минимума для выпуклой задачи без ограничений

Рассмотрим задачу

$$f(u) \to \min,\ u \in U, \qquad (P_{5_0})$$

где $f : \mathbb{R}^n \to \mathbb{R}$ — выпуклая функция, а $U \subset \mathbb{R}^n$ — выпуклое множество.

Предложение 5. *Элемент $\widehat{u}$ является решением задачи (P_{5_0}) тогда и только тогда, когда выполнено условие минимума:*

$$c)\ f_0(u) \geq f_0(\widehat{u}),\ \forall u \in U, \qquad (5_0)$$

а если $U = \mathbb{R}^n$, то верен критерий минимума – условие стационарности α):

$$\alpha)\ 0 \in \partial f_0(\widehat{u}). \qquad (5_0')$$

Последний результат будем называть *теоремой Ферма для выпуклых функций.*

Доказательство обоих утверждений совпадает с определением соответствующих понятий. □

Правило множителей Лагранжа для выпуклых задач

Пусть $\mathcal{U}$ — пространство $\mathbb{R}^n$ или другое векторное пространство, f_i при $i = 0, \dots, m'$ — выпуклые, а при $i = m'+1 \dots, m$ — аффинные функции на $\mathcal{U}$, U — выпуклое подмножество $\mathcal{U}$. Рассмотрим задачу

$$f_0(u) \to \min,\ f_i(u) \le 0,\ 1 \le i \le m',\ f_i(u) = 0,\ m'+1 \le i \le m, \quad u \in U. \qquad (P_5)$$

(такие задачи будем называть *выпуклыми*). Её функция Лагранжа имеет вид: $\mathcal{L}(u, \overline{\lambda}) = \sum_{i=0}^m \lambda_i f_i(u)$ и $\overline{\lambda} = (\lambda_0, \dots, \lambda_m)$.

Теорема 5. *Если $\widehat{u}$ доставляет минимум (абсолютный) задаче (P_5), то необходимые условия минимума в этой задаче соответствуют принципу Лагранжа.* А именно, найдётся вектор $\overline{\lambda} = (\lambda_0, \dots, \lambda_m) \ne 0$ множителей Лагранжа, при котором для функции Лагранжа выполнены (в соответствии с предложением 5)

c) (условие минимума по u): $\min\limits_{u \in U} \mathcal{L}(u, \overline{\lambda}) = \mathcal{L}(\widehat{u}, \overline{\lambda})$,

β) условие дополняющей нежесткости: $\lambda_i f_i(\widehat{u}) = 0$, $1 \le i \le m'$,

γ) условие неотрицательности: $\lambda_i \ge 0$, $0 \le i \le m'$.

В случае, если при некоторых $\widehat{u} \in U$ и $\overline{\lambda} = (\lambda_0, \dots, \lambda_m)$ для функции Лагранжа $\mathcal{L}(u, \overline{\lambda})$ выполняются условия c), β), γ) и $\lambda_0 = 1$, то $\widehat{u}$ — решение задачи (P_5).

Доказательство. Не ограничивая общности, считаем, что $f(\widehat{u}) = 0$. Множество $\mathcal{A} = \{\alpha = (\alpha_0, \dots, \alpha_{m'}, 0, \dots, 0)^T \in \mathbb{R}^{m+1} \mid \exists\, u \in U : \alpha_k \ge f_k(u),\ 0 \le k \le m',\ 0 = f_k(u),\ m'+1 \le k \le m\}$ выпукло и непусто, открытый луч $\mathcal{C} = \{(\alpha_0, 0, \dots, 0) \mid \alpha_0 < 0\}$ не пересекается с $\mathcal{A}$ (ибо иначе точка $\widehat{u}$ не была бы решением задачи). Согласно теореме отделимости найдётся не равный нулю вектор $\overline{\lambda} = (\lambda_0, \dots, \lambda_m)$ такой, что

$$\inf_{\alpha \in \mathcal{A}} \langle \lambda, \alpha \rangle \ge \sup_{\alpha \in \mathcal{C}} \langle \lambda, \alpha \rangle = \sup_{\alpha_0 < 0} \lambda_0 \alpha_0. \qquad (i)$$

Из (i) следует, что $\lambda_0 \ge 0$, и значит, $\sup\limits_{c_0 < 0} \lambda_0 c_0 = 0$. Подставив в это неравенство элемент из $\mathcal{A}$, у которого k-я координата, $1 \le k \le m'$, равна единице, а остальные — нулю, приходим к условию γ) неотрицательности. Выбрав для подстановки в (i) элемент из $\mathcal{A}$, у которого k-я координата,

$1 \leq k \leq m'$, равна $f_k(\widehat{u})$, а остальные — нулевые, приходим к условию $\beta)$ дополняющей нежесткости. Подставив в (i) элемент $(f_0(u), \ldots, f_m(u))^T$, $u \in U$, и пользуясь тем, что $\sum_{i=0}^m \lambda_i f_i(\widehat{u}) = 0$, приходим к условию $c)$ минимума. Необходимость доказана.

Обратно. Если для $\widehat{u} \in U$ и $\overline{\lambda} = (\lambda_0, \ldots, \lambda_m)$, $\lambda_0 = 1$, выполнены условия $c)$, $\beta)$, $\gamma)$ и $u \in U$ произвольно, то: $f_0(u) \overset{\gamma)}{\geq} f_0(u) + \sum_{k=1}^m \lambda_k f_k(u) \overset{c)}{\geq} f_0(\widehat{u}) + \sum_{k=1}^m \lambda_k f_k(\widehat{u}) \overset{\beta)}{=} f_0(\widehat{u})$. Достаточность доказана. $\square$

В следующих трех пунктах, в частном случае простейшей задачи вариационного исчисления, обсуждается ещё один важный принцип теории экстремума — принцип глобального снятия ограничений, построенный на основе "возмущения" экстремальных задач. (В основе его лежит идея Гамильтона о том, что для решения задач вариационного исчисления полезно рассматривать не одну экстремаль, а целое семейство экстремалей, охватывающих заданную). Полное снятие ограничений основывается на построении полей экстремалей, т. е. семейств экстремалей, зависящих от параметра.

6. Условия минимума второго порядка для гладкой задачи без ограничений и с ограничениями типа равенств, условия Лежандра и Якоби для простейшей задачи вариационного исчисления

Необходимое условие и достаточное условие минимума для гладкой задачи без ограничений и с ограничениями типа равенств

Рассмотрим задачу без ограничений
$$f(x) \to \min \quad (f : \mathbb{R}^n \to \mathbb{R}) \qquad (P_{1_0})$$
и задачу с ограничениями типа равенств
$$f_0(x) \to \min,\ f_i(x) = 0,\ 1 \leq i \leq m, \qquad (P_1)$$
где $f_i : \mathbb{R}^n \to \mathbb{R}$, $0 \leq i \leq m$ (напомним, что функция Лагранжа для (P_1) имеет вид: $\mathcal{L}(x, \overline{\lambda}) = \sum_{i=0}^m \lambda_i f_i(x)$). Обозначим

$F(x){:=}\begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$, $x \in \mathbb{R}^n$.

Предложение 6. Пусть в задачах (P_{1_0}) и (P_1) функции f, f_i, $0 \le i \le m$, дважды дифференцируемы в $\widehat{x} \in \mathbb{R}^n$. Тогда

a) если $\widehat{x}$ является локальным минимумом для (P_{1_0}), то выполнены условие стационарности и неотрицательности квадратичной формы $f''(\widehat{x})[\cdot,\cdot]$:

$$f'(\widehat{x}) = 0,\ f''(\widehat{x})[x,x] \ge 0 \ \ \forall x \in \mathbb{R}^n; \qquad (6_{0a})$$

b) если выполнены условия

$$f'(\widehat{x}) = 0,\ f''(\widehat{x})[x,x] > 0 \ \ \forall x \in \mathbb{R}^n \setminus \{0\} \qquad (6_{0b})$$

стационарности и положительности квадратичной формы $f''(\widehat{x})[\cdot,\cdot]$, то $\widehat{x}$ является локальным минимумом f для (P_{1_0});

c) если $\widehat{x}$ — локальный минимум в (P_1), функции f_i, $0 \le i \le m$, непрерывно дифференцируемы в некоторой окрестности точки $\widehat{x}$, а образ оператора $F'(\widehat{x})$ совпадает с $\mathbb{R}^m$, то найдётся вектор $\overline{\lambda} = (1, \lambda_1, \ldots, \lambda_m)$ множителей Лагранжа, такой что для функции Лагранжа $\mathcal{L}(\cdot, \overline{\lambda})$ выполнены условие стационарности в точке $\widehat{x}$ и условие неотрицательности квадратичной формы $\mathcal{L}_{xx}(\widehat{x}, \overline{\lambda})[\cdot,\cdot]$ на ядре оператора $F'(\widehat{x})$:

$$\mathcal{L}_x(\widehat{x}, \overline{\lambda}) = 0, \quad \mathcal{L}_{xx}(\widehat{x}, \overline{\lambda})[x,x] \ge 0 \quad \forall x \in \operatorname{Ker} F'(\widehat{x}); \quad (6_{0c})$$

d) если в допустимой точке $\widehat{x}$ для функции Лагранжа $\mathcal{L}(\cdot, \overline{\lambda})$ с множителями Лагранжа $\overline{\lambda} = (1, \lambda_1, \ldots, \lambda_m)$ выполнены условие стационарности и условие положительности квадратичной формы $\mathcal{L}_{xx}(\widehat{x}, \overline{\lambda})[\cdot,\cdot]$ на ядре оператора $F'(\widehat{x})$:

$$\mathcal{L}_x(\widehat{x}, \overline{\lambda}) = 0,\ \mathcal{L}_{xx}(\widehat{x})[x,x] > 0 \ \forall x \in \operatorname{Ker} F'(\widehat{x}),\ x \ne 0, \quad (6_{0d})$$

то $\widehat{x}$ — локальный минимум в задаче (P_1).

Доказательство. *a*) Если $f(\cdot)$ дважды дифференцируема в точке $\widehat{x}$, являющейся локальным минимумом для (P_{1_0}), то при любом $x \in \mathbb{R}^n$ 1) (в соответствии с предложением 1) выполнено условие $f'(\widehat{x})[x] = 0$ и 2) $0 \le f(\widehat{x} + \vartheta x) - f(\widehat{x}) = \frac{\vartheta^2}{2} f''(\widehat{x})[x,x] + \bar{o}(\vartheta^2)$ при $\vartheta \in \mathbb{R}$ стремящемся к нулю (в соответствии с определением второй производной). Поэтому неравенство $f''(\widehat{x})[x,x] < 0$ выполняться не может, а справедливо (6_{0a}).

b) Пусть $f(\cdot)$ дважды дифференцируема в $\widehat{x}$ и при этом выполнены условия (6_{0b}). Единичная сфера пространства $\mathbb{R}^n$ является компактом, поэтому (в силу теоремы Вейерштрасса о достижении минимума непрерывной функцией, заданной на компакте) квадратичная форма $f''(\widehat{x})[x,x]$ достигает своего минимума по $x \in \mathbb{R}^n$, $|x| = 1$, в некоторой точке $\bar{x}$, $|\bar{x}| = 1$: $f''(\widehat{x})[x,x] \geq f''(\widehat{x})[\bar{x},\bar{x}] = r > 0$. Так как в точке $\widehat{x}$ выполнено условие стационарности, то $f(\widehat{x}+x) - f(\widehat{x}) - \frac{1}{2}f''(\widehat{x})[x,x] = \bar{o}(|x|^2)$ при $|x| \to 0$. Значит, можно указать малое $\delta > 0$, для которого $|f(\widehat{x}+x) - f(\widehat{x}) - \frac{1}{2}f''(\widehat{x})[x,x]| < \frac{r}{2}\,|x|^2$ при $|x| < \delta$. Поэтому,

$$f(\widehat{x}+x)-f(\widehat{x}) = f(\widehat{x}+x)-f(\widehat{x})-\frac{1}{2}f''(\widehat{x})[x,x]+\frac{1}{2}f''(\widehat{x})[x,x] >$$

$$-\frac{r}{2}\,|x|^2 + \frac{|x|^2}{2}f''(\widehat{x})[\frac{x}{|x|},\frac{x}{|x|}] > 0$$

при любом $x \in \mathbb{R}^n$, $0 < |x| < \delta$, т.е. $\widehat{x}$ — локальный минимум в (P_{1_0}).

c) Условие стационарности для набора множителей $\overline{\lambda} = (\lambda_0, \ldots, \lambda_m)$ доказано в теореме 1. Если в этом наборе $\lambda_0 = 0$, то заключение теоремы 1 означает линейную зависимость строк $(f_i)'(\widehat{x})$, $i = 1,\ldots,m$, а это противоречит условиям предложения 6 *c*). Значит, $\lambda_0 \neq 0$, и вектор $(\lambda_0,\ldots,\lambda_m)$ может быть разделён на λ_0, так что утверждение теоремы 1 останется справедливым с $\lambda_0 = 1$. Докажем неотрицательность $\mathcal{L}_{xx}(\widehat{x},\overline{\lambda})[\cdot,\cdot]$ на ядре оператора $F'(\widehat{x})$ для построенного набора $\overline{\lambda}$: пусть $\xi \in \operatorname{Ker} F'(\widehat{x})$, $L_1 \oplus L_2$ — прямая сумма одномерного подпространства L_1, порождённого вектором ξ, и подпространства L_2, для которого $\dim L_2 = m$, $F'(\widehat{x})[L_2] = \mathbb{R}^m$. Применим теорему о неявной функции к отображению $\psi : L_1 \times L_2 \to \mathbb{R}^m$, $\psi(x,y) := F(x+y+\widehat{x})$ в точке $\binom{0}{0}$, — отображение $\varphi : \binom{x}{y} \mapsto \binom{x}{\psi(x,y)}$ является диффеоморфизмом между некоторой окрестностью точки $\binom{0}{0}$ и некоторой окрестностью V точки $\binom{0}{\psi(0,0)}$, $\psi(0,0)=0$. Для "второй координаты" $\varphi_{-1}\in C^1(V)$ обратного к φ отображения выполнено тождество $\psi(x,\varphi_{-1}(x,z))=z$, $\binom{x}{z}\in V$, т.е. $z=F(x+\varphi_{-1}(x,z)+\widehat{x})$, $\binom{x}{z}\in V$, и вектор $x+\varphi_{-1}(x,0)+\widehat{x}$ допусти́м в задаче (P_1) при $\binom{x}{0}\in V$. Воспользуемся гладкостью отображения $\varphi_{-1}(\cdot)$ на V и, как следствие, оценкой $|\varphi_{-1}(x,0)|=\underline{O}(|x|)$ при $|x|\to 0$, чтобы установить более сильное свойство $|\varphi_{-1}(x,0)|=\bar{o}(|x|)$: так как $0=F(x+\varphi_{-1}(x,0)+\widehat{x})=F'(\widehat{x})[x+\varphi_{-1}(x,0)]+\bar{o}(|x+\varphi_{-1}(x,0)|)=F'(\widehat{x})[\varphi_{-1}(x,0)]+\bar{o}(|x+\varphi_{-1}(x,0)|)$, то при $|x|\to 0$ обязательно $|\varphi_{-1}(x,0)|=\bar{o}(|x|)$. Следовательно, $0\le f_0(x+\varphi_{-1}(x,0)+\widehat{x})-f_0(\widehat{x})=\mathcal{L}(x+\varphi_{-1}(x,0)+\widehat{x},\overline{\lambda})-\mathcal{L}(\widehat{x},\overline{\lambda})=\frac{1}{2}\mathcal{L}_{xx}(\widehat{x},\overline{\lambda})[x,x]+\bar{o}(|x|^2)$ при $|x|\to 0$, т.е. $\mathcal{L}_{xx}(\widehat{x},\overline{\lambda})[\xi,\xi]\ge 0$.

d) Пусть $\mathbb{R}^n=X\oplus Y$, $X=\operatorname{Ker}F'(\widehat{x})$, — какое-нибудь разложение пространства $\mathbb{R}^n$ на прямую сумму, в котором одно из слагаемых совпадает с $\operatorname{Ker}F'(\widehat{x})$; $x\in X$, $y\in Y$ таковы, что элемент $x+y+\widehat{x}$ допусти́м в (P_1). Тогда $0=F(x+y+\widehat{x})-F(\widehat{x})=F'(\widehat{x})[x+y]+\bar{o}(|x+y|)=F'(\widehat{x})[y]+\bar{o}(|x+y|)$. Значит, если $x+y+\widehat{x}$ допусти́м, $|x|=\bar{o}(1)$, $|y|=\bar{o}(1)$, то $|y|=\bar{o}(|x|)$: предположив обратное, получим, что найдутся последовательности $\{x_n\},\{y_n\}$ сходящиеся к нулю, для которых $|x_n|=\underline{O}(|y_n|)$, $y_n\ne 0$. Но тогда выполняется равенство $0=F'(\widehat{x})[y_n]+\bar{o}(|y_n|)$, из которого, в силу конечномерности $Y\subset\mathbb{R}^n$, следует, что $F'(\widehat{x})[\cdot]$ вырождается на Y – противоречие.

Поэтому, если $|z|=\bar{o}(1)$, $z+\widehat{x}$ допусти́м, $z=x+y$ — представление z, соответствующее разложению $\mathbb{R}^n$ в прямую сумму, то $f_0(x+y+\widehat{x})-f_0(\widehat{x})=\mathcal{L}(x+y+\widehat{x},\overline{\lambda})-\mathcal{L}(\widehat{x},\overline{\lambda})=\frac{1}{2}\mathcal{L}_{xx}(\widehat{x},\overline{\lambda})[x+y,x+y]+\bar{o}(|x+y|^2)=\frac{1}{2}\mathcal{L}_{xx}(\widehat{x},\overline{\lambda})[x,x]+\bar{o}(|x|^2)>0$ (пользуемся стационарностью $\mathcal{L}(\cdot,\overline{\lambda})$ в точке $\widehat{x}$; неравенство >0 для достаточно малого $|x|$ устанавливается также как и в части *b*). $\square$

Рассмотрим задачу с равенствами и неравенствами

$$f_0(x)\to\min,\ f_i(x)\le 0, 1\le i\le m,\ f_i(x)=0, 1+m\le i\le m',\quad (P_1')$$

где $f_i:\mathbb{R}^n\to\mathbb{R}$, $0\le i\le m'$. Будем использовать ещё одну постановку этой задачи:

$$f_0(x)\to\min,\ f_i(x)+u_i=0,\ u_i\ge 0,\ 1\le i\le m,$$
$$f_i(x)=0, 1+m\le i\le m'. \qquad (P_1'')$$

Функцию Лагранжа для (P_1') запишем в виде: $\mathcal{L}(x,\overline{\lambda}) = \sum\limits_{i=0}^{m}\lambda_i(f_i(x)+u_i) + \sum\limits_{i=m+1}^{m'}\lambda_i f_i(x)$. Если $m < m'$, обозначим $F(x){:=}\begin{pmatrix} f_{m+1}(x) \\ \vdots \\ f_{m'}(x)\end{pmatrix}$, если же ограничений вида равенства нет, то пусть $F(x):=0$, $F:\mathbb{R}^n\to\mathbb{R}^{m'-m}$, $x\in\mathbb{R}^n$.

Теорема 1$'$ (принцип Лагранжа для задач с равенствами и неравенствами). Пусть в задаче (P_1') функции f_i, $0\le i\le m'$, непрерывно дифференцируемы в окрестности точки $\widehat{x}\in\mathbb{R}^n$, $n\in\mathbb{N}$. Тогда, *если $\widehat{x}$ является локальным минимумом для (P_1'), то необходимые условия минимума в этой задаче соответствуют принципу Лагранжа.* А именно, найдётся вектор $\overline{\lambda}=(\lambda_0,\dots,\lambda_{m'})\ne 0$ множителей Лагранжа, при котором для функции Лагранжа выполнены (в соответствии с теоремой 1 для переменной x) условие стационарности $\mathcal{L}_x(\widehat{x},\overline{\lambda})=0$ и (в соответствии с теоремой 5 для переменной u) условие минимальности по $u\ge 0$ функции Лагранжа, условия неотрицательности и дополняющей нежёсткости. Это приводит к тому, что $\lambda_i\ge 0$, $i=0,\dots,m$ и $\lambda_i f_i(\widehat{x})=0$, $i=0,\dots,m$.

Доказательство. В вырожденном случае $F'(\widehat{x})[\mathbb{R}^n]\ne\mathbb{R}^{m'-m}$; тогда найдётся $\lambda\ne 0\in(\mathbb{R}^{m'-m})^*$ такой, что $\lambda\cdot F'(\widehat{x})=0$ и в качестве вектора множителей Лагранжа, для которого выполняются утверждения предложения, можно выбрать $\overline{\lambda}=(0,\dots,0,\lambda)\in(\mathbb{R}^{m'})^*$. Пусть $F'(\widehat{x})[\mathbb{R}^n]=\mathbb{R}^{m'-m}$. Положим $A_k=\{h\in\mathbb{R}^n:\ \langle f_i'(\widehat{x}),h\rangle<0,\ k\le i\le m,\ F'(\widehat{x})[h]=0\}$, $k=0,\dots,m$, $A_{m+1}=\{h\in\mathbb{R}^n:\ F'(\widehat{x})[h]=0\}$; тогда (из определения) $A_0\subset A_1\subset\dots\subset A_m\subset A_{m+1}\ne\emptyset$. Проверим, что множество A_0 является пустым. Действительно, пусть $\xi\in A_0\subset\operatorname{Ker}F'(\widehat{x})$, $L_1\oplus L_2$ — прямая сумма одномерного подпространства L_1, порождённого вектором ξ, и подпространства L_2, для которого $\dim L_2 = m' - m$, $F'(\widehat{x})[L_2] = \mathbb{R}^{m'-m}$. Применим теорему о неявной функции к отображению $\psi : L_1 \times L_2 \to \mathbb{R}^{m'-m}$, $\psi(x,y) := F(x+y+\widehat{x})$ в точке $\binom{0}{0}$, — отображение $\varphi : \binom{x}{y} \mapsto \binom{x}{\psi(x,y)}$ является диффеоморфизмом между некоторой окрестностью точки $\binom{0}{0}$ и некоторой окрестностью V точки $\binom{0}{\psi(0,0)}$, $\psi(0,0) = 0$. Для "второй координаты" $\varphi_{-1} \in C^1(V)$ обратного к φ отображения выполнено тождество $\psi(x, \varphi_{-1}(x,z)) = z$, $\binom{x}{z} \in V$, т.е. $z = F(x + \varphi_{-1}(x,z) + \widehat{x})$, $\binom{x}{z} \in V$, и вектор $x + \varphi_{-1}(x,0) + \widehat{x}$ допусти́м в задаче (P_1) при $\binom{x}{0} \in V$. Воспользуемся гладкостью отображения $\varphi_{-1}(\cdot)$ на V и, как следствие, оценкой $|\varphi_{-1}(x,0)| = \underline{O}(|x|)$ при $|x| \to 0$, чтобы установить что $|\varphi_{-1}(x,0)| = \bar{o}(|x|)$: так как $0 = F(x + \varphi_{-1}(x,0) + \widehat{x}) - F(\widehat{x}) = F'(\widehat{x})[x + \varphi_{-1}(x,0)] + \bar{o}(|x + \varphi_{-1}(x,0)|) = F'(\widehat{x})[\varphi_{-1}(x,0)] + \bar{o}(|x + \varphi_{-1}(x,0)|)$, то при $|x| \to 0$ обязательно $|\varphi_{-1}(x,0)| = \bar{o}(|x|)$. Следовательно, $f_i(x + \varphi_{-1}(x,0) + \widehat{x}) - f_i(\widehat{x}) = f_i'(\widehat{x})[x + \varphi_{-1}(x,0)] + \bar{o}(|x + \varphi_{-1}(x,0)|) = f_i'(\widehat{x})[x] + \bar{o}(|x|) < 0$ при всех малых по норме x, $i = 0, 1, \ldots, m$. Но это противоречит предположению о том, что $\widehat{x}$ — локальный минимум в задаче.

Таким образом, найдётся k, $0 \le k \le m$, при котором $A_k = \emptyset$ и $A_{k+1} \neq \emptyset$. Но тогда нуль является решением следующей выпуклой задачи:
$\langle f_k'(\widehat{x}), h\rangle \to \min \langle f_i'(\widehat{x}), h\rangle \le 0,\ k+1 \le i \le m,\ F'(\widehat{x})[h] = 0, (i)$

так как если $\langle f_k'(\widehat{x}), \xi\rangle < 0$, $\langle f_i'(\widehat{x}), \xi\rangle \le 0$, $k+1 \le i \le m$, $F'(\widehat{x})[\xi] = 0$ для некоторого $\xi \in \mathbb{R}^n$, то, взяв $\eta \in A_{k+1} \neq \emptyset$, обнаружим, что при $t > 0$ элемент $\xi + t\eta$ принадлежит A_k в противоречии с нашим допущением. По теореме 5 для задачи (i) найдётся ненулевой вектор $(\lambda_k, \ldots, \lambda_{m'})$ множителей Лагранжа, удовлетворяющий условиям условиям минимальности, дополняющей нежёсткости и неотрицательности. Условие минимальности для (i) превращается в равенства $\langle \sum_{i=k}^{m'} \lambda_i f_i'(\widehat{x}), h\rangle = 0$, $h \in \mathbb{R}^n$ — произвольно. Таким образом и в этом случае утверждение предложения выполнены с вектором $\overline{\lambda} = (0, \ldots, 0, \lambda_k, \ldots, \lambda_{m'}) \neq 0 \in (\mathbb{R}^{m'})^*$ множителей

Лагранжа. $\square$

Применим предложение 6 к простейшей векторной задаче

$$\mathcal{J}(x(\cdot))=\int\limits_{t_0}^{t_1} L(t,x(t),\dot x(t))dt\to\min,\ \ x(t_j)=x_j,\ \ j=0,1, \quad (P_{2_0})$$

где $L:\mathbb{R}^{2n+1}\to\mathbb{R}$. При этом будем пользоваться обозначениями и рассуждениями из доказательства предложения 2, но дополнительно к условиям этого предложения потребуем, чтобы функция L была бы дважды непрерывно дифференцируема в окрестности $\Gamma_{\widehat{x}(\cdot)}$. Тогда функция $\psi(\cdot,v)$, построенная в ходе доказательства предложения 2, дважды непрерывно дифференцируема по ϑ в окрестности нуля и для неё (в соответствии с предложением 6) выполнены необходимые условия экстремума второго порядка: $\frac{d^2\psi(\vartheta,v(\cdot)}{d\vartheta^2}|_{\vartheta=0}\geq 0$. Вычислим эту производную, пользуясь теоремой о дифференцируемости интеграла по параметру: $\int_{t_0}^{t_1}(\langle\widehat{L}_{\dot x\dot x}(t)\dot v(t),\dot v(t)\rangle + 2\langle\widehat{L}_{x\dot x}(t)\dot v(t),v(t)\rangle + \langle\widehat{L}_{xx}(t)v(t),v(t)\rangle)\,dt \ \geq\ 0$. Так как $v\in C_0^1([t_0,t_1],\mathbb{R}^n)$ произвольна, то доказано

Утверждение 1. Пусть в дополнение к условиям предложения 2 функция L дважды непрерывно дифференцируема в окрестности $\Gamma_{\widehat{x}(\cdot)}$. Тогда функция $\widehat{v}(\cdot)\equiv 0$ является решением задачи

$$\int_{t_0}^{t_1}\big(\langle\widehat{L}_{\dot x\dot x}(t)\dot v(t),\dot v(t)\rangle+2\langle\widehat{L}_{x\dot x}(t)\dot v(t),v(t)\rangle+ \\ +\langle\widehat{L}_{xx}(t)v(t),v(t)\rangle\big)\,dt\to\min\quad v(t_0)=v(t_1)=0. \quad\square$$

Задачу из утверждения 1 называют *присоединённой* к (P_{2_0}); исследуем её.

Условия минимума для простейшей квадратичной задачи вариационного исчисления с нулевыми граничными условиями

Пусть в простейшей квадратичной задаче вариационного исчисления

$$\int_{t_0}^{t_1}\big(\langle A(t)\dot x(t),\dot x(t)\rangle+2\langle C(t)\dot x(t),x(t)\rangle+ \\ +\langle B(t)x(t),x(t)\rangle\big)\,dt\to\min\quad x(t_0)=x(t_1)=0,$$

матричные функции $A(\cdot)$, $B(\cdot)$, $C(\cdot)\ :\ [t_0,t_1]\ \to\ L(\mathbb{R}^n,\mathbb{R}^n)$

(из которых первые две симметричны) непрерывны. Тогда *условием Лежандра* называется выполнение соотношения: $\langle A(t)u, u\rangle \geq 0$ для всех $u \in \mathbb{R}^n$ и всех $t \in [t_0, t_1]$; *усиленным условием Лежандра* – выполнение соотношения: $\langle A(t)u, u\rangle > 0$ для всех $u \in \mathbb{R}^n \setminus \{0\}$ и всех $t \in [t_0, t_1]$. Если выполнено усиленное условие Лежандра, то *условием Якоби* называется отсутствие на (t_0, t_1) точек, сопряженных с точкой t_0 (т.е. точек $\tau \in (t_0, t_1)$, для которых существует нетривиальное класса $C^1([t_0, \tau], \mathbb{R}^n)$ решение уравнения Эйлера–Якоби: $-\frac{d}{dt}(A(t)\dot{x}(t) + C^T(t)x(t)) + C(t)\dot{x}(t) + B(t)x(t) = 0$, удовлетворяющее граничным условиям $x(t_0) = x(\tau) = 0$); *усиленным условием Якоби* называется отсутствие на $(t_0, t_1]$ точек, сопряженных с точкой t_0.

Теорема 6. *Пусть в простейшей квадратичной задаче вариационного исчисления*

$$\int_{t_0}^{t_1} \big(\langle A(t)\dot{x}(t), \dot{x}(t)\rangle + 2\langle C(t)\dot{x}(t), x(t)\rangle + \\ + \langle B(t)x(t), x(t)\rangle\big)\, dt \to \min \quad x(t_0) = x(t_1) = 0,$$

матричные функции $A(\cdot)$, $B(\cdot)$ и $C(\cdot)$ (из которых первые две симметричны) непрерывны на $[t_0, t_1]$. Тогда, если функция $\widehat{x}(t) \equiv 0$ является решением задачи в пространстве $C^1([t_0, t_1], \mathbb{R}^n)$, то выполнено (i) условие Лежандра, а если выполнено усиленное условие Лежандра, то выполнено условие (ii) Якоби.

Если же выполнены усиленные условия Лежандра и Якоби, то тождественный нуль является решением простейшей квадратичной задачи вариационного исчисления (её абсолютным минимумом).

Доказательство. *а*) Обозначив $\dot{x} = u$, переформулируем исходную задачу, как задачу оптимального управления:

$$\int_{t_0}^{t_1} (\langle A(t)u(t), u(t)\rangle + 2\langle C(t)u(t), x(t)\rangle + \langle B(t)x(t), x(t)\rangle)\, dt \to \min$$

$\dot{x} = u$, $x(t_0) = x(t_1) = 0$; так как $\widehat{x}(t) \equiv 0$ является абсолютным минимумом в исходной задаче, то (по лемме о скруглении углов) $\binom{\widehat{x}(\cdot)}{\dot{\widehat{x}}(\cdot)}$ доставляет сильный минимум для переформулированной задачи оптимального управления. Воспользуемся условием минимума из теоремы 4, из которого следует, что $\lambda_0 \neq 0$ и $\langle A(t)u, u\rangle \geq 0 \ \forall u \in \mathbb{R}^n$, $t \in [t_0, t_1]$. Но это и означает, что условие Лежандра выполнено. b) Допустим теперь, что сопряженная точка τ лежит в интервале (t_0, t_1) и пусть $\widehat{x}(\cdot)$ — нетривиальное решение уравнения Эйлера–Якоби: $-\frac{d}{dt}(A(t)\dot{\widehat{x}}(t) + C^T(t)\widehat{x}(t)) + C(t)\dot{\widehat{x}}(t) + B(t)\widehat{x}(t) = 0$. Умножив это равенство на $\widehat{x}(\cdot)$, проинтегрировав полученное выражение по отрезку $[t_0, \tau]$ и выполнив интегрирование по частям, получим, что $\int\limits_{t_0}^{\tau}(\langle A(t)\dot{\widehat{x}}(t), \dot{\widehat{x}}(t)\rangle + 2\langle C(t)\dot{\widehat{x}}(t), \widehat{x}(t)\rangle + \langle B(t)\widehat{x}(t), \widehat{x}(t)\rangle)dt = 0$. Таким образом, функция $\widetilde{x}(\cdot)$, равная $\widehat{x}(\cdot)$ на отрезке $[t_0, \tau]$ и продолженная нулем на отрезок $[\tau, t_1]$, также доставляет абсолютный минимум нашей квадратичной задаче. А это противоречит условию минимума из теоремы 4 для $\binom{\widetilde{x}(\cdot)}{\dot{\widetilde{x}}(\cdot)}$, согласно которому найдётся $p(\cdot) \in PC^1([t_0, t_1], (\mathbb{R}^n)^*)$ такое, что при $t > \tau$ имеет место равенство $p(t) = 0$, а при $t < \tau$ – равенство $p(t) = \big(2A(t)\dot{\widetilde{x}}(t) + 2C(t)^T\widetilde{x}(t)\big)^T$, т.к., например, $\langle C(t)u, \widehat{x}(t)\rangle = \big(C(t)^T\widehat{x}(t)\big)^T u$. Из непрерывности $p(\cdot)$, условия $\widetilde{x}(\tau) = 0$ и усиленного условия Лежандра заключаем отсюда, что $\dot{\widetilde{x}}(\tau) = 0$ также. Но, если $\dot{\widetilde{x}}(\tau) = \widetilde{x}(\tau) = 0$, то в силу теоремы единственности решения задачи Коши для линейного уравнения, которым является уравнение Якоби, $\widetilde{x}(t) \equiv 0$. Противоречие получено.

И обратно: то, что тождественный нуль является решением квадратичной задачи, если выполнены усиленные условие Лежандра и условие Якоби, следует из рассматриваемой ниже теоремы 8 и однородности квадратичной задачи. □

Следствием доказанной теоремы являются необходимые условия экстремума (минимума) первого и второго порядка для простейшей задачи вариационного исчисления (P_{2_0}) (такой экстремум принято называть *слабым*):

Утверждение 2. *Пусть в задаче* (P_{2_0}) *функция* L *два-*

жды непрерывно дифференцируема в окрестности расширенного графика $\left\{ \begin{pmatrix} t \\ \widehat{x}(t) \\ \dot{\widehat{x}}(t) \end{pmatrix} \in \mathbb{R}^{1+n+r} \mid t \in [t_0, t_1] \right\}$ *непрерывно дифференцируемой функции* $\widehat{x}(\cdot)$. *Тогда, если функция* $\widehat{x}(\cdot)$ *доставляет локальный экстремум задаче* (P_{2_0}) *в пространстве* $C^1([t_0,t_1],\mathbb{R}^n)$, *то для* $\widehat{x}(\cdot)$ *выполнено a) уравнение Эйлера* (2_0), (i) *условие Лежандра для присоединённой задачи и, если для присоединённой задачи выполнено усиленное условие Лежандра, то для неё же выполнено условие* (ii) *Якоби.*

Условия и усиленные условия Лежандра и Якоби для присоединённой к (P_{2_0}) задачи называют, соответственно, условиями и усиленными условиями Лежандра и Якоби для задачи (P_{2_0}).

Доказательство следует из предложения 2, утверждения 1 и теоремы 6. □

7. Построение поля и уравнение Гамильтона–Якоби для простейшей задачи вариационного исчисления

Пусть в простейшей задаче вариационного исчисления

$$\mathcal{J}(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt \to \text{extr} \quad x(t_j) = x_j,\ j = 0, 1, \quad (P_{2'_0})$$

интегрант $L : \mathbb{R}^3 \to \mathbb{R}$ трижды непрерывно дифференцируем в окрестности расширенного графика $\Gamma_{\widehat{x}(\cdot)} = \{(t, \widehat{x}(t), \dot{\widehat{x}}(t))^T \in \mathbb{R}^3 \mid t \in [t_0, t_1]\}$ функции $\widehat{x}(\cdot) \in C^2([t_0, t_1])$, для которой выполнены уравнение Эйлера и усиленные условия Лежандра и Якоби. Условия гладкости для L позволяют выполнить дифференцирование в первом слагаемом из уравнения Эйлера $-\frac{d}{dt}L_{\dot{x}} + L_x = 0$:

$$L_{\dot{x}\dot{x}}(t, x, \dot{x})\ddot{x} + L_{\dot{x}x}(t, x, \dot{x})\dot{x} + L_{\dot{x}t}(t, x, \dot{x}) - L_x(t, x, \dot{x}) = 0.$$

Из усиленного условия Лежандра (и непрерывности $L_{\dot{x}\dot{x}}$) следует, что в некоторой окрестности множества $\Gamma_{\widehat{x}(\cdot)}$ коэффициент $L_{\dot{x}\dot{x}}(t, x, \dot{x})$ не обращается в ноль и, значит, в этой окрестности уравнение Эйлера равносильно уравнению, разрешённому относительно старшей производной:

$$\ddot{x} = \Phi(t, x, \dot{x}) := \frac{L_x(t, x, \dot{x}) - L_{\dot{x}t}(t, x, \dot{x}) - L_{\dot{x}x}(t, x, \dot{x})\dot{x}}{L_{\dot{x}\dot{x}}(t, x, \dot{x})}\,.$$

Применяя к этому уравнению локальную теорему существования и единственности решений задачи Коши, получим, что решение $\widehat{x}(\cdot)$ уравнения Эйлера может быть продолжено за пределы отрезка $[t_0, t_1]$ на некоторый отрезок $[\bar{t}_0, \bar{t}_1]$, $\bar{t}_0 < t_0 < t_1 < \bar{t}_1$; по глобальной теореме о непрерывной зависимости решений задачи Коши от начальных данных, можно считать, что при этом для $\widehat{x}(\cdot)$ на $[\bar{t}_0, \bar{t}_1]$ также выполнены усиленные условия Лежандра и Якоби. В следующем предложении устанавливаются некоторые свойства семейства $x(\cdot, \lambda)$, $\lambda \in \mathbb{R}$, решений уравнения Эйлера

$$\ddot{x} = \Phi(t, x, \dot{x}), \quad x(\bar{t}_0) = \widehat{x}(\bar{t}_0),\ \dot{x}(\bar{t}_0) = \dot{\widehat{x}}(\bar{t}_0) + \lambda. \qquad (i)$$

Предложение 7 (о поле, окружающем экстремаль). *Пусть в задаче* $(P_{2_0'})$ *интегрант* L *трижды непрерывно дифференцируем в окрестности расширенного графика* $\Gamma_{\widehat{x}(\cdot)}$ *функции* $\widehat{x}(\cdot) \in C^2([\bar{t}_0, \bar{t}_1])$, *для которой выполнены уравнение Эйлера и усиленные условия Лежандра и Якоби. Тогда найдётся* $\lambda_0 > 0$ *такое, что для решений* (i) *определено и принадлежит* $C^1([\bar{t}_0, \bar{t}_1] \times [-\lambda_0, \lambda_0], \mathbb{R}^2)$ *отображение* $\varphi : \binom{t}{\lambda} \mapsto \binom{t}{x(t,\lambda)}$. *Это отображение осуществляет диффеоморфизм класса* C^1 *в обе стороны некоторой окрестности графика* $\{\binom{t}{0} : t \in [t_0, t_1]\}$ *на некоторую окрестность* V *графика* $\{\binom{t}{x(t,0)} : t \in [t_0, t_1]\}$.

Доказательство. Существование числа $\lambda_0 > 0$, определённость отображения $\varphi(\cdot)$ и его принадлежность классу $C^1([\bar{t}_0, \bar{t}_1] \times [-\lambda_0, \lambda_0], \mathbb{R}^2)$ следует из теоремы о дифференцируемой зависимости решений задачи Коши от начальных данных. По этой же теореме функция $h(\cdot) = \frac{\partial x(\cdot,\lambda)}{\partial \lambda}\big|_{\lambda=0}$ удовлетворяет уравнению в вариациях — уравнению Якоби $-\frac{d}{dt}(A(t)\dot{h}) + B(t)h = 0$, где $A(t) = L_{\dot{x}\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$, а

$B(t) = L_{xx}(t, \widehat{x}(t), \dot{\widehat{x}}(t)) - \frac{d}{dt} L_{\dot{x}x}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$, и краевым условиям условиям $h(\bar{t}_0) = 0$, $\dot{h}(\bar{t}_0) = 1$. Из-за усиленного условия Якоби отсюда следует, что $h(t) \neq 0$ при $t \in [t_0, t_1]$.

Применим теорему о неявной функции в каждой точке $\binom{\tau}{0}$, $\tau \in [t_0, t_1]$, к отображению $\binom{t}{\lambda} \mapsto x(t, \lambda)$, пользуясь при этом тем, что $\frac{\partial x(\tau,\lambda)}{\partial \lambda}\big|_{\lambda=0} \neq 0$: для $\tau \in [t_0, t_1]$ найдётся открытое множество $U_1(\tau)$, содержащее $\binom{\tau}{0}$, на котором отображение $\varphi : \binom{t}{\lambda} \mapsto \binom{t}{x(t,\lambda)}$ является диффеоморфизмом между $U_1(\tau)$ и открытым множеством $V(\tau)$, содержащем $\binom{\tau}{x(\tau,0)}$. Отсюда следует, что $\varphi(\cdot)$ является также диффеоморфизмом между открытыми множествами $\cup_{\tau \in [t_0,t_1]} U_1(\tau)$ и $V = \cup_{\tau \in [t_0,t_1]} V(\tau)$. По теореме о неявной функции этот диффеоморфизм непрерывно дифференцируем в обе стороны. $\square$

Так как V содержит график $\{\binom{\tau}{\widehat{x}(\tau)} \mid \tau \in [t_0, t_1]\}$ экстремали $\widehat{x}(\cdot)$, то тем самым в окрестности V графика $\widehat{x}(\cdot)$ определено отображение $\lambda \in C^1(V, \mathbb{R})$, — "вторая координата" обратного к φ отображения (из предложения 7), — для которого $x(\tau, \lambda(\tau, \xi)) \equiv \xi$ при всех $\binom{\tau}{\xi} \in V$. Для $\tau \in [t_0, t_1]$, $\lambda \in [-\lambda_0, \lambda_0]$ положим $G(\tau, \lambda) = \int\limits_{\bar{t}_0}^{\tau} L(t, x(t, \lambda), \dot{x}(t, \lambda))\, dt$. Тогда $\frac{\partial G(\tau,\lambda)}{\partial \tau} = L(\tau, x(\tau, \lambda), \dot{x}(\tau, \lambda))$. По теореме о дифференцировании интеграла по параметру и теореме о дифференцируемой зависимости решений задачи Коши от начальных данных $\frac{\partial G(\tau,\lambda)}{\partial \lambda} = \int\limits_{\bar{t}_0}^{\tau} \big(L_x(t, x(t, \lambda), \dot{x}(t, \lambda)) \cdot h(t, \lambda) + L_{\dot{x}}(t, x(t, \lambda), \dot{x}(t, \lambda)) \cdot \dot{h}(t, \lambda)\big)\, dt$, где $h(\cdot, \lambda) = \frac{\partial x(\cdot,\lambda)}{\partial \lambda}$. Так как $h(\bar{t}_0, \lambda) = 0$, а $x(\cdot, \lambda)$ удовлетворяет уравнению Эйлера, то после интегрирования по частям получим $\frac{\partial G(\tau,\lambda)}{\partial \lambda} = L_{\dot{x}}(\tau, x(\tau, \lambda), \dot{x}(\tau, \lambda)) \cdot h(\tau, \lambda)$. В итоге, $dG(\tau, \lambda) =$

$$= L_{\dot{x}}(\tau, x(\tau, \lambda), \dot{x}(\tau, \lambda)) \cdot h(\tau, \lambda)\, d\lambda + L(\tau, x(\tau, \lambda), \dot{x}(\tau, \lambda))\, d\tau.$$

Продифференцируем тождество $x(\tau, \lambda(\tau, \xi)) \equiv \xi$, $\binom{\tau}{\xi} \in V$, по переменным τ и ξ: а) $\dot{x}(\tau, \lambda(\tau, \xi)) + h(\tau, \lambda(\tau, \xi)) \cdot \lambda_\tau(\tau, \xi) = 0$, б) $h(\tau, \lambda(\tau, \xi)) \cdot \lambda_\xi(\tau, \xi) = 1$.

Функцию $S : V \to \mathbb{R}$, определённую равенством $S(\tau,\xi) = G(\tau, \lambda(\tau,\xi))$, называют S-функцией задачи $(P_{2'_0})$. Обозначим $u(\tau,\xi) = \dot{x}(\tau, \lambda(\tau,\xi))$, $\binom{\tau}{\xi} \in V$ (эту функцию называют ещё *гильбертовым полем*). Тогда дифференциал функции S можно записать в виде

$$dS(\tau,\xi) = L_{\dot{x}}\big(\tau,\xi,u(\tau,\xi)\big)\cdot h(\tau,\lambda(\tau,\xi))(\lambda_\tau(\tau,\xi)\,d\tau + \lambda_\xi(\tau,\xi)\,d\xi) + L\big(\tau,\xi,u(\tau,\xi)\big)\,d\tau = \Big(L\big(\tau,\xi,u(\tau,\xi)\big) - L_{\dot{x}}\big(\tau,\xi,u(\tau,\xi)\big)u(\tau,\xi)\Big)d\tau + L_{\dot{x}}\big(\tau,\xi,u(\tau,\xi)\big)d\xi.$$

Таким образом, доказана

Теорема 7 (об уравнении Гамильтона — Якоби). *При сделанных предположениях относительно задачи $(P_{2'_0})$, построенная функция S удовлетворяет уравнению Гамильтона–Якоби* $\frac{\partial S(\tau,\xi)}{\partial \tau} + \mathcal{H}\left(\tau, \xi, \frac{\partial S(\tau,\xi)}{\partial \xi}\right) = 0$, где через $p(\tau,\xi)$ обозначена функция $L_{\dot{x}}(\tau,\xi,u(\tau,\xi))$, а через $\mathcal{H}(\tau,\xi,p)$ — функция $u(\tau,\xi)\cdot p - L(\tau,\xi,u(\tau,\xi))$. □

Замечание 1. *Имеется несколько определений, связанных с рассмотренными объектами. Если функция $\lambda(\cdot)$ определена на области V, то семейство $x(\cdot,\lambda)$, $\lambda \in [-\lambda_0, \lambda_0]$, с помощью которого $\lambda(\cdot)$ построена, называют центральным полем экстремалей (с центром в $\binom{\bar{t}_0}{\widehat{x}(\bar{t}_0)}$), покрывающим область V, функцию $u : V \to \mathbb{R}$, заданную формулой $u(\tau,\xi) = \dot{x}(\tau,\lambda(\tau,\xi))$ — наклоном поля экстремалей $x(\cdot,\lambda)$; про экстремаль $\widehat{x}(\cdot) = x(\cdot,0)$ говорят, что она включена в центральное поле экстремалей, покрывающее окрестность V графика $\widehat{x}(\cdot)$. В терминах функций $p(\tau,\xi)$ и $\mathcal{H}(\tau,\xi,p)$ из теоремы 7 выражение для дифференциала функции $S(\cdot)$ приобретает вид $dS(\tau,\xi) = -\mathcal{H}(\tau,\xi,p(\tau,\xi))d\tau + p(\tau,\xi)d\xi$.*

Замечание 2. *Если функция $u \in C^1(V)$, то в области V она удовлетворяет квазилинейному дифференциальному уравнению первого порядка*

$$L_{\dot{x}\dot{x}}(\tau,\xi,u)\cdot(u_\tau + u\cdot u_\xi) = L_x(\tau,\xi,u) - L_{\dot{x}x}(\tau,\xi,u)\cdot u - L_{\dot{x}\tau}(\tau,\xi,u)$$

и условию $u(\tau,\widehat{x}(\tau)) = \dot{\widehat{x}}(\tau)$, $\tau \in [\bar{t}_0, \bar{t}_1]$, вдоль графика экс-

тремали $\widehat{x}(\cdot)$.

8. Формула Вейерштрасса и достаточные условия минимума в простейшей задаче

Пусть $L \in C(\mathbb{R}^3, \mathbb{R})$ – интегрант простейшей задачи $(P_{2'_0})$ переменных t, x, y дифференцируем по y. Функцию $(t, x, y, u) \mapsto \mathcal{E}(t, x, y, u) = L(t, x, u) - L(t, x, y) - (u - y)L_y(t, x, y)$, $u \in \mathbb{R}$, называют *$\mathcal{E}$-функцией Вейерштрасса* (построенной по L).

Предложение 8. *Пусть интегрант L простейшей задачи $(P_{2'_0})$ и экстремаль $\widehat{x}(\cdot)$ этой задачи удовлетворяют условиям предложения 7, V — окрестность графика этой экстремали, в которой определены S-функция $S(\tau, \xi)$ задачи и её гильбертово поле $u(\tau, \xi)$, $\binom{\tau}{\xi} \in V$, $u(\tau, \widehat{x}(\tau)) = \dot{\widehat{x}}(\tau)$, $\tau \in [t_0, t_1]$. Тогда для любой допустимой функции $x(\cdot) \in PC^1([t_0, t_1])$, график которой лежит в V, имеет место формула:*

$$\mathcal{J}(x(\cdot)) - \mathcal{J}(\widehat{x}(\cdot)) = \int_{t_0}^{t_1} \mathcal{E}(t, x(t), u(t, x(t)), \dot{x}(t))dt.$$

Теорема 8 (достаточные условия минимума в простейшей задаче). *Пусть в простейшей задаче $(P_{2'_0})$ интегрант L трижды непрерывно дифференцируем и квазирегулярен в окрестности графика $\widehat{x}(\cdot)$ (это означает, что в любой точке (t, x) вблизи $(t, \widehat{x}(t))$ функция $\dot{x} \to L(t, x, \dot{x})$ выпукла) дважды непрерывно дифференцируемой экстремали $\widehat{x}(\cdot)$. Тогда, если на этой экстремали выполнены усиленные условия Лежандра и Якоби, то эта экстремаль доставляет сильный минимум задаче $(P_{2'_0})$.*

Доказательство предложения и теоремы 8. Продолжим рассуждения, начатые в пункте 7: если график $x(\cdot)$ лежит в V, то $\int\limits_{t_0}^{t_1} dS(t, \widehat{x}(t)) = \int\limits_{t_0}^{t_1} dS(t, x(t)) = \int\limits_{t_0}^{t_1}\Big(L(t, x(t), u(t, x(t))) - L_{\dot{x}}(t, x(t), u(t, x(t)))u(t, x(t))\Big)dt + L_{\dot{x}}(t, x(t), u(t, x(t)))dx(t) = \int_{t_0}^{t_1} \widehat{L}(t)dt = J(\widehat{x}(\cdot))$.

Отсюда вытекает *основная формула Вейерштрасса*: $\mathcal{J}(x(\cdot))-\mathcal{J}(\widehat{x}(\cdot))=\int\limits_{t_0}^{t_1} L(t,x(t),\dot{x}(t))dt-\int\limits_{t_0}^{t_1} dS(t,x(t)) = \int\limits_{t_0}^{t_1}(L(t,x(t),\dot{x}(t))-L(t,x(t),u(t,x(t))) - L_{\dot{x}}(t,x(t),u(t,x(t)))(\dot{x}(t)-u(t,x(t))))dt = \int_{t_0}^{t_1} \mathcal{E}(t,x(t),u(t,x(t)),\dot{x}(t))dt$.

Если же дополнительно график $x(\cdot)$ лежит в области, где интегрант квазирегулярен, то из условия о квазирегулярности теоремы следует неотрицательность функции Вейерштрасса, и значит, $J(x(\cdot)) \geq J(\widehat{x}(\cdot))$, т. е. $\widehat{x}(\cdot)$ доставляет задаче $(P_{2'_0})$ сильный локальный минимум. $\square$

Замечание 1. Аналогично выводятся достаточные условия экстремума для простейшей векторной задачи.

Замечание 2. В доказательстве теоремы 8 фактически содержится доказательство ещё одного достаточного признака экстремальности, который бывает полезен при изучении простейшей задачи: пусть интегрант L из постановки $(P_{2'_0})$ дважды непрерывно дифференцируем и квазирегулярен в области V, содержащей график $\Gamma_{\widehat{x}(\cdot)}$ непрерывно дифференцируемой экстремали $\widehat{x}(\cdot)$; пусть на множестве V определено отображение $u \in C(V,\mathbb{R}^n)$ (гильбертово поле), для которого дифференциальная форма $L_{\dot{x}}(\tau,\xi,u(\tau,\xi))d\xi + \big(L(\tau,\xi,u(\tau,\xi))-L_{\dot{x}}(\tau,\xi,u(\tau,\xi))\cdot u(\tau,\xi)\big)d\tau$ на V является полным дифференциалом и для которого $u(\tau,\widehat{x}(\tau)) = \dot{\widehat{x}}(\tau)$ при всех $\tau \in [t_0,t_1]$. Тогда $\widehat{x}$ доставляет сильный минимум задаче $(P_{2'_0})$ среди всех допустимых функций из $PC^1([t_0,t_1],\mathbb{R}^n)$, графики которых лежат в V (при доказательстве теоремы 8 в качестве гильбертова поля выступала функция наклона поля $u(\tau,\xi)=\dot{x}(\tau,\lambda(\tau,\xi))$, $\binom{\tau}{\xi} \in V$, а полным дифференциалом был дифференциал $dS(\tau,\xi)$, $\binom{\tau}{\xi} \in V$, S-функции $S(\cdot)$).

Проблемам существования в простейшей задаче посвящен следующий пункт.

9. Принцип компактности Вейерштрасса–Бэра и теорема Тонелли существования решения простейшей задачи вариационного исчисления

Действительнозначную функцию $f(\cdot)$, определённую на топологическом пространстве X, называем *полунепрерывной снизу*, если для любого $r \in \mathbb{R}$ множество $\{x \mid f(x) \le r\}$ замкнуто в X. Топологическое пространство X называем *компактным*, если из любого покрытия X открытыми множествами можно выделить конечное подпокрытие, (**[Z], [KF]**).

Непрерывную функцию $f \in C([t_0,t_1])$, $t_0, t_1 \in \mathbb{R}$, называем *абсолютно непрерывной* на $[t_0,t_1]$, если существует суммируемая на $[t_0,t_1]$ функция $g(\cdot)$, для которой $f(t) - f(t_0) = \int_{t_0}^{t} g(\tau)\,d\tau$ при любом $t \in [t_0,t_1]$. Множество всех абсолютно непрерывных на $[t_0,t_1]$ функций обозначаем $AC([t_0,t_1]) \subset C([t_0,t_1])$; оно является нормированным пространством с нормой, перенесённой из пространства $C([t_0,t_1])$. Если функция $g(\cdot)$, однозначно (с точностью до значений на множестве меры нуль из $[t_0,t_1]$) определяемая по функции $f \in AC([t_0,t_1])$, принадлежит пространству $L_p([t_0,t_1])$, $1 \le p \le \infty$, то говорим, что $f \in W_p^1([t_0,t_1])$. Линейное пространство $W_p^1([t_0,t_1])$ является банаховым с нормой $\|f\|_{W_p^1([t_0,t_1])} = |f(t_0)| + \|g\|_{L_p([t_0,t_1])}$, (**[KF]**).

Теорема (об абсолютной непрерывности интеграла). Если $g \in L_1([t_0,t_1])$, то для любого $\varepsilon > 0$ найдётся $\delta > 0$, при котором $\int_A |g(t)|\,dt < \varepsilon$ для любого измеримого подмножества $A \subset [t_0,t_1]$ меры меньше δ: $\mu(A) < \delta$, (**[KF]**).

Теорема Асколи-Арцела. Семейство функций $\{x_n\}_{n\in\mathbb{N}} \subset C([t_0,t_1])$, удовлетворяющее двум условиям: 1) найдётся $A > 0$ такое, что $\sup_{n\in\mathbb{N}} \|x_n(\cdot)\|_{C([t_0,t_1]} \le A$ (равномерная ограниченность семейства) и 2) для любого $\varepsilon > 0$ найдётся $\delta > 0$ такое, что $\sup_{n\in\mathbb{N}} |x_n(\tau_1) - x_n(\tau_2)| \le \varepsilon$, если $|\tau_1 - \tau_2| \le \delta$ (равностепенная непрерывность семейства), — является предкомпактным в $C([t_0,t_1])$ (т. е. из него можно выделить сходящуюся в $C([t_0,t_1])$ подпоследовательность), (**[Z], [KF]**).

Пусть $1 \le p \le \infty$, $p^{-1} + p'^{\,-1} = 1$. Последовательность $\{x_n(\cdot)\}_{n\in\mathbb{N}} \subset L_p([t_0,t_1])$, называется слабо сходящейся к $z \in L_p([t_0,t_1])$, если $\lim\limits_{n\to\infty} \int_{t_0}^{t_1} (x_n(t) - z(t))\xi(t)dt = 0$ для любой

$\xi(\cdot) \in L_{p'}([t_0,t_1])$. Теорема о секвенциальной слабой компактности шара для пространства $L_p([t_0,t_1])$ утверждает, что *при $1 < p < \infty$ из любой ограниченной в $L_p([t_0,t_1])$ последовательности функций можно выделить слабо сходящуюся подпоследовательность*, (**[KF]**).

Предложение 9 (принцип компактности). *Пусть X — компактное топологическое пространство и $f : X \to \mathbb{R} \cup \{+\infty\}$ — полунепрерывная снизу функция на X, не равная тождественно $+\infty$. Тогда f ограничена снизу и существует точка $\widehat{x} \in X$, в которой f достигает абсолютного минимума.*

Этот принцип называют ещё *принципом компактности Вейерштрасса–Бэра.*

Доказательство. Обозначим: $\mathcal{L}_n f := \{x \mid f(x) \leq n,\ n \in \mathbb{Z}\}$. Из определения полунепрерывности снизу следует, что $U_n := X \setminus \mathcal{L}_n f$, $n \in \mathbb{Z}$, — открытые множества в X. Ясно, что $\ldots \subset U_n \subset U_{n-1} \subset \ldots$ и что $\{U_n\}_{n\in\mathbb{Z}}$ — открытое покрытие X. Так как X — компакт, то существует $m \in \mathbb{Z}$ такое, что $U_m = X$, т.е. f ограничена снизу и потому существует нижняя грань $\mu := \inf_{x\in X} f(x)$. Если нижняя грань не достигается, положим $V_n := X \setminus \mathcal{L}_{\mu+\frac{1}{n}} f, n \in \mathbb{N}$. Из определения полунепрерывности снизу следует, что $\{V_n\}_{n\in\mathbb{N}}$ есть открытое покрытие X. Следовательно, (снова из определения компактности) существует число $s \in \mathbb{N}$ такое, что $X = V_s$, т.е. $f > \mu + 1/s$. Но это противоречит определению μ. $\square$

О том, насколько требование компактности существенно, свидетельствует следующий пример задачи минимизации на $\mathbb{R}^2$ *неотрицательного многочлена четвёртого порядка от двух переменных*, в которой *нет решения*: $x_1^2 + (x_1x_2 - 1)^2 \to \min$.

Теорема 9 (Тонелли о существовании решения в простейшей задаче вариационного исчисления).

Пусть в задаче

$$\int_{t_0}^{t_1} L(t,x(t),\dot x(t))dt \to \min,\ x(t_i)=x_i,\ i=0,1,$$

интегрант $L:\mathbb{R}^3\to\mathbb{R}$ *и его частная производная* $L_{\dot x}(\cdot)$ *непрерывны в* $\mathbb{R}^3$, *пусть* $L(\cdot)$ *является выпуклым по* $\dot x$ *(при фиксированных* t *и* x*) и удовлетворяет следующему условию роста:* $L(t,x,\dot x)\ge \alpha|\dot x|^p+\beta,\ \alpha>0,\ \beta\in\mathbb{R},\ p>1$. *Тогда в пространстве абсолютно непрерывных функций* $AC([t_0,t_1])$ *(и даже в пространстве* $W_p^1([t_0,t_1])$*) существует решение задачи.*

Этот результат был доказан Л.Тонелли в 1928 г.

Доказательство. Доказательство основывается на двух вспомогательных утверждениях:

1) *Пусть* $\{x_n(\cdot)\}_{n\in\mathbb{N}}$ — *минимизирующая последовательность. Тогда найдётся её подпоследовательность, сходящаяся в* $W_p^1([t_0,t_1])$ *(к некоторой функции* $\widehat{x}(\cdot)$*).*

Не ограничивая общности, считаем, что последовательность $\int\limits_{t_0}^{t_1} L(t,x_n(t),\dot x_n(t))dt$, $n\in\mathbb{N}$, невозрастает. Тогда $\int\limits_{t_0}^{t_1}|\dot x_n(t)|^p dt\le \frac{1}{\alpha}\Big(\int\limits_{t_0}^{t_1} L(t,x_1(t),\dot x_1(t))dt-\beta\Big) =: C$. Пользуясь неравенством Гёльдера, получим: $|x_n(\tau+h)-x_n(\tau)| = |\int_\tau^{\tau+h}\dot x_n(s)ds| \le \left(\int_{t_0}^{t_1}|\dot x_n(s)|^p dt\right)^{\frac{1}{p}} h^{\frac{1}{p'}} \le Ch^{\frac{1}{p'}},\ p^{-1}+{p'}^{-1}=1$. Отсюда и из того, что $x_n(t_0)=x_0$, $n\in\mathbb{N}$, следует равномерная ограниченность и равностепенная непрерывность семейства $\{x_n(\cdot)\}_{n\in\mathbb{N}}$. По теореме Асколи-Арцела найдётся подпоследовательность из $\{x_n(\cdot)\}_{n\in\mathbb{N}}$ равномерно сходящаяся к $\widehat{x}(\cdot)$. Не ограничивая общности, считаем, что сама $\{x_n(\cdot)\}_{n\in\mathbb{N}}$ является равномерно сходящейся. По теореме о слабой компактности шара пространства $L_p([t_0,t_1])$ некоторая подпоследовательность из семейства $\{\dot x_n(\cdot)\}_{n\in\mathbb{N}}$ слабо сходится к $\bar z(\cdot)$ в $L_p([t_0,t_1])$. Опять таки, не ограничивая общности, считаем, что сама $\{\dot x_n(\cdot)\}_{n\in\mathbb{N}}$ является слабо сходящейся в $L_p([t_0,t_1])$.

В итоге получаем, что для любого $t \in [t_0, t_1]$ последовательность $x_n(t) = x_0 + \int_{t_0}^{t} \dot{x}_n(s)ds$ сходится, с одной стороны, к $\widehat{x}(t)$ а с другой стороны, к $x_0 + \int_{t_0}^{t} \bar{z}(s)ds$. Отсюда следует, что $\widehat{x}(\cdot) \in W_p^1([t_0, t_1])$, поскольку $\dot{\widehat{x}}(t) = \bar{z}(t)$ и $\bar{z}(\cdot) \in L_p([t_0, t_1]) \subset L_1([t_0, t_1])$. $\square$

2) *Функция $\widehat{x}(\cdot)$ является решением задачи.*
Опять считаем, что $\int_{t_0}^{t_1} L(t, x_n(t), \dot{x}_n(t))dt$, $n \in \mathbb{N}$, невозрастает; тогда, так как функция $L(\cdot)$ ограничена снизу константой β, то $\sup\limits_n \int_{t_0}^{t_1} |L(t, x_n(t), \dot{x}_n(t))|dt = C_1 < \infty$. Пусть $\widehat{L}(t) = L(t, \widehat{x}(t), \dot{\widehat{x}}(t))$, $\widehat{L}_{\dot{x}}(t) = L_{\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$ для $t \in [t_0, t_1]$. Выберем $A > 0$ произвольно и обозначим через Δ множество $\{t \in [t_0, t_1] \,:\, |\dot{\widehat{x}}(t)| \le A\}$. Тогда $\mathcal{J}(x_n(\cdot)) - \mathcal{J}(\widehat{x}(\cdot)) =$

$$\int\limits_{\Delta} \big(L(t, x_n(t), \dot{x}_n(t)) - \widehat{L}(t)\big)\, dt + \int\limits_{[t_0,t_1]\setminus\Delta} L(t, x_n(t), \dot{x}_n(t))\, dt - \int\limits_{[t_0,t_1]\setminus\Delta} \widehat{L}(t)\, dt\,.$$

Оценим интегралы из правой части раздельно. Для первого из них: $\int\limits_{\Delta}\big(L(t,x_n(t), \dot{x}_n(t)) - \widehat{L}(t)\big)\, dt = \int\limits_{\Delta}\big(L(t,x_n(t), \dot{x}_n(t)) - L(t,x_n(t), \dot{\widehat{x}}(t)) + L(t,x_n(t), \dot{\widehat{x}}(t)) - \widehat{L}(t)\big)dt \ge \int\limits_{\Delta}(\dot{x}_n(t) - \dot{\widehat{x}}(t)) \cdot L_{\dot{x}}(t, x_n(t), \dot{\widehat{x}}(t))dt + \int\limits_{\Delta}\big(L(t, x_n(t), \dot{\widehat{x}}(t)) - \widehat{L}(t)\big)dt = \int\limits_{\Delta}(\dot{x}_n(t) - \dot{\widehat{x}}(t)) \cdot \widehat{L}_{\dot{x}}(t)dt + \int\limits_{\Delta}(\dot{x}_n(t) - \dot{\widehat{x}}(t)) \cdot (L_{\dot{x}}(t, x_n(t), \dot{\widehat{x}}(t)) - \widehat{L}_{\dot{x}}(t))dt + \int\limits_{\Delta}\big(L(t, x_n(t), \dot{\widehat{x}}(t)) - \widehat{L}(t)\big)dt =: I_{1n} + I_{2n} + I_{3n} \to 0$, когда $n \to \infty$.
В самом деле, $I_{1n} \to 0$, поскольку $\dot{x}_n(\cdot)$ слабо сходится к $\dot{\widehat{x}}(\cdot)$ в $L_p([t_0, t_1])$, а функция $\widehat{L}_{\dot{x}}(\cdot)$, продолженная нулем с Δ на $[t_0, t_1]$, ограничена на $[t_0, t_1]$; $I_{3n} \to 0$, так как $x_n(\cdot)$ равномерно сходится к $\widehat{x}(\cdot)$, а функция $L(\cdot)$ равномерно непрерывна на множестве $[t_0, t_1] \times [-M, M] \times [-A, A]$, $M = \sup\limits_k \|x_k(\cdot)\|_{C([t_0,t_1])}$; $I_{2n} \to 0$, так как $|I_{2n}| \le \|\dot{x}_n(\cdot) - \dot{\widehat{x}}(\cdot)\|_{L_p([t_0,t_1])} \cdot \|L_{\dot{x}}(\cdot, x_n(\cdot), \dot{\widehat{x}}(\cdot)) - \widehat{L}_{\dot{x}}(\cdot)\|_{L_{p'}(\Delta)} \le 2C\|L_{\dot{x}}(\cdot, x_n(\cdot), \dot{\widehat{x}}(\cdot)) - \widehat{L}_{\dot{x}}(\cdot)\|_{L_{p'}(\Delta)} \to 0$, посколь-

ку $x_n(\cdot)$ равномерно сходится к $\widehat{x}(\cdot)$, а функция $L_{\dot{x}}(\cdot)$ равномерно непрерывна на множестве $[t_0, t_1] \times [-M, M] \times [-A, A]$. Из полученной оценки для первого интеграла следует, что

$$|\int_\Delta \widehat{L}(t)\,dt| \le \overline{\lim_n}\ |\int_\Delta L(t, x_n(t), \dot{x}_n(t))\,dt| \le C_1.$$

Так как $A > 0$ можно выбрать произвольно, то $|\int\limits_{[t_0,t_1]} \widehat{L}(t)\,dt| \le C_1$, что, как и выше для $L(\cdot, x_n(\cdot), \dot{x}_n(\cdot))$, влечёт суммируемость $\widehat{L}(\cdot)$ на $[t_0, t_1]$. Это, в свою очередь, означает, что для третьего из интегралов (вследствие абсолютной непрерывности интеграла) $\int_{[t_0,t_1]\setminus\Delta} \widehat{L}(t)\,dt = \bar{o}(1)$ при $\mathrm{mes}([t_0, t_1]\setminus\Delta)) \to 0$.

Для второго интеграла при любом $n \in \mathbb{N}$

$$\int_{[t_0,t_1]\setminus\Delta} L(t, x_n(t), \dot{x}_n(t))\,dt \ge -|\beta| \cdot \mathrm{mes}([t_0, t_1] \setminus \Delta)\,.$$

В итоге,

$$\lim_n \mathcal{J}(x_n(\cdot)) - \mathcal{J}(\widehat{x}(\cdot)) \ge -|\beta| \cdot \mathrm{mes}([t_0, t_1] \setminus \Delta) + \bar{o}(1)$$

при $\mathrm{mes}([t_0, t_1] \setminus \Delta)) \to 0$. Так как $\mathrm{mes}([t_0, t_1] \setminus \Delta)) \to 0$ при $A \to +\infty$, то это доказывает теорему. $\square$

Последний пункт посвящен важнейшему принципу выпуклости, согласно которому, допуская некоторую вольность речи, всё выпуклое имеет двойное описание. Впервые это явление на примере выпуклых множеств было описано Минковским. Здесь в теореме Фенхеля–Моро выпуклая и замкнутая функция с одной стороны описывается, как функция с выпуклым и замкнутым надграфиком, а с другой, как верхняя грань аффинных функций, лежащих под графиком заданной выпуклой замкнутой функции. В теории линейного программирования каждая задача имеет двойственную, для которой аргументами служат множители Лагранжа исходной задачи; знание решения одной из этих задач даёт исчерпывающую информацию о решении второй.

10. Двойственность в линейном программировании и теорема Фенхеля–Моро

Рассмотрим следующую экстремальную задачу:

$$\langle c, x\rangle \to \min \quad Ax \geq b,\ x \geq 0\,, \qquad (P_{10})$$

где $c, x \in \mathbb{R}^n$, $b \in \mathbb{R}^m, A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Её называют *задачей линейного программирования в нормальной форме*. Функцию Лагранжа этой задачи запишем в виде: $\mathcal{L}(x, \lambda_0, y) = \lambda_0\langle c, x\rangle + y \cdot (b - Ax)$.

Задачу

$$\langle y, b^T\rangle \to \max \quad yA \leq c^T,\ y \geq 0 \quad (y \in (\mathbb{R}^m)^*) \qquad (P^*_{10})$$

называют *двойственной* к (P_{10}). Имеет место

Предложение 10 (о двойственности в линейном программировании). *Если решение задачи (P_{10}) существует, то значения задач (P_{10}), (P^*_{10}) совпадают и, в дополнение к утверждению теоремы 5 (Каруша–Куна–Таккера) для задачи (P_{10}), компонента $\widehat{y}$ вектора множителей Лагранжа $(\lambda_0, \widehat{y})$, нормированная условием $\lambda_0 = 1$ (если $\lambda_0 \neq 0$) и выбранная произвольно (если $\lambda_0 = 0$), является решением задачи (P^*_{10}).*

Доказательство (опирается на принцип Лагранжа для выпуклых задач). Функция Лагранжа задачи (P_{10}) такова: $\mathcal{L}(x, \lambda) = \lambda_0\langle c, x\rangle + y \cdot (b - Ax)$. Пусть $\widehat{x}$ — решение задачи (P_{10}). Тогда в силу принципа Лагранжа для выпуклых задач (теоремы Каруша–Куна–Таккера) найдётся такой вектор $\overline{\lambda} = (\lambda_0, \widehat{y}) \in (\mathbb{R}^{m+1})^*$ множителей Лагранжа, что

$$\min_{x\geq 0} \mathcal{L}(x, \overline{\lambda}) = \mathcal{L}(\widehat{x}, \overline{\lambda}),\ \lambda_0 \geq 0,\ \widehat{y} \geq 0,\ \widehat{y} \cdot (b - A\widehat{x}) = 0. \qquad (i)$$

Так как вектор $(\lambda_0, \widehat{y})$ в (i) определяется с точностью до положительного множителя, то при $\lambda_0 \neq 0$, его можно выбрать так, чтобы $\lambda_0 = 1$. Следовательно, $\mathcal{L}(\widehat{x}, (1, \widehat{y})) \overset{\text{def}}{=} \langle c, \widehat{x}\rangle + \widehat{y}\cdot(b - A\widehat{x}) \overset{(i)}{=} \langle c, \widehat{x}\rangle \overset{(i)}{=} \min\limits_{x\geq 0} \mathcal{L}(x, (1, \widehat{y})) \overset{\text{def}}{=} \min\limits_{x\geq 0}\langle (c - (\widehat{y}A)^T), x\rangle + \widehat{y} \cdot b.(ii)$

Из (ii) сразу следует, что

$$(\widehat{y}A)^T \leq c \text{ и } \langle c, \widehat{x}\rangle \leq \widehat{y} \cdot b = \langle \widehat{y}, b^T\rangle. \qquad (iii)$$

Покажем, что множитель Лагранжа $\widehat{y}$ является решением задачи (P^*_{10}) и значения задач (P_{10}) и (P^*_{10}) совпадают. Имеем:

a) $\langle c,\widehat{x}\rangle = c^T\cdot\widehat{x} \overset{(iii)}{\geq} (\widehat{y}A)\cdot\widehat{x} = \langle A^T\widehat{y}^T,\widehat{x}\rangle = \widehat{y}\cdot A\widehat{x} \overset{\mathrm{Cond}(P_{10})}{\geq} \widehat{y}\cdot b \overset{(iii)}{\Rightarrow}$
$\langle c,\widehat{x}\rangle = \langle\widehat{y},b^T\rangle,$ $\qquad (iv)$

b) если y удовлетворяет ограничениям задачи (P^*_{10}), то

$\langle y,b^T\rangle = \langle y^T,b\rangle \overset{\mathrm{Cond}(P_{10})}{\leq} \langle y^T,A\widehat{x}\rangle = \langle (yA)^T,\widehat{x}\rangle \overset{\mathrm{Cond}(P^*_{10})}{\leq}$
$\langle c,\widehat{x}\rangle \overset{(iv)}{=} \langle\widehat{y},b^T\rangle.$

Если $\lambda_0=0$, то $\widehat{x}$ – решение задачи $\widehat{y}\cdot(-Ax)\to\min,\ x\geq 0$, т.е. $\widehat{x}=0$. Тогда $\mathcal{L}(\widehat{x},(0,\widehat{y}))=\widehat{y}\cdot(b-A\widehat{x}) \overset{(i)}{=} 0 \overset{(i)}{=} \min\limits_{x\geq 0}\mathcal{L}(x,(0,\widehat{y})) =$
$\min\limits_{x\geq 0}\langle -(\widehat{y}A)^T,x\rangle + \widehat{y}\cdot b.$ $\qquad (ii')$

Из (ii') следует, что $(\widehat{y}A)^T\leq 0$ и $\langle c,\widehat{x}\rangle = 0 = \widehat{y}\cdot b = \langle\widehat{y},b^T\rangle$.

То, что множитель Лагранжа $\widehat{y}$ в этом случае также является решением задачи (P^*_{10}), является повтором рассуждений из b). □

Теорема 10 (Фенхеля – Моро о двойственности выпуклых функций). *Пусть $f: X\to\mathbb{R}\cup\{+\infty\}$. Тогда равенство $f=f^{**}$ имеет место в том и только в том случае, когда f выпукла и замкнута.*

Доказательство. Доказательство складывается из двух частей a) если $f^{**}=f$, то $f(\cdot)$ выпукла и замкнута и b) если $f(\cdot)$ выпукла и замкнута, то $f^{**}=f$.

a) По определению второй сопряжённой функции она есть верхняя грань аффинных функций $x\mapsto x^*\cdot x - f^*(x^*)$, надграфики которых — замкнутые полупространства. Значит, надграфик f^{**} является пересечением замкнутых полупространств. Отсюда следует, что если $f=f^{**}$, то f выпукла и замкнута.

b) Если $f\equiv\infty$, то из определений получаем, что $f^{**}=f$. В ином случае найдётся элемент x_0, для которого $f(x_0)<\infty$. Существующая при любом $\beta>0$ гиперплоскость, строго отделяющая epi$f$ от точки $\binom{x_0}{f(x_0)-\beta}$ есть график аффинной функции $a_0(x) = x_0^* \cdot x - \alpha_0$: гиперплоскость не может быть "вертикальной" (т.е. иметь уравнение $x_0^* \cdot x = c$), ибо иначе она проходила бы через точку $\binom{x_0}{f(x_0)-\beta}$, а не отделяла бы её. Таким образом, $a_0(x) \le f(x)\ \forall x$ и при этом $a_0(x_0) > f(x_0) - \beta$.

Из определения сопряжённой функции следует неравенство Юнга: $x^* \cdot x \le f(x) + f^*(x^*)$, откуда вытекает неравенство $f^{**}(x) \le f(x)\ \forall x$. Допустим, что существует точка x_1, для которой $f^{**}(x_1) < f(x_1)$. Докажем, что существует аффинная функция $a_1(\cdot)$, график которой лежит под $\mathrm{epi}\, f$ и при этом $a_1(x_1) > f^{**}(x_1)$. Если $f(x_1) < \infty$, надо повторить предыдущее рассуждение. Осталось рассмотреть случай, когда $f(x_1) = \infty$. По теореме отделимости можно строго отделить $\binom{x_1}{f^{**}(x_1)}$ от $\mathrm{epi} f$. Если отделяющая гиперплоскость есть график аффинной функции, то мы добились своей цели. Если же эта гиперплоскость "вертикальна", т. е. задаётся функционалом x_1^* таким, что $x_1^* \cdot x_1 > c = \sup\limits_{\binom{x}{\alpha} \in \mathrm{epi} f} x_1^* \cdot x$, положим $a_\mu(x) = a_0(x) + \mu(x_1^* \cdot x - c)$. Легко понять, что при любом числе $\mu > 0$ аффинная функция $a_\mu(x) \le a_0(x) \le f(x)$, а при достаточно большом $\widehat{\mu}$ получим, что $a_{\widehat{\mu}}(x_1) > f^{**}(x_1)$. Функция $a_{\widehat{\mu}}$ и есть искомая функция a_1. Действительно, пусть $a_1(x) = a_{\widehat{\mu}}(x) = x_2^* \cdot x - \alpha_2$. Значит, выполнены неравенства $x_2^* \cdot x - \alpha_2 \le f(x)\ \forall x \ \Leftrightarrow\ \alpha_2 \ge x_2^* \cdot x - f(x) \ \Rightarrow\ \alpha_2 \ge f^*(x_2^*)$. С другой стороны, $x_2^* \cdot x_1 - \alpha_2 > f^{**}(x_1) \ \Rightarrow\ \alpha_2 < x_2^* \cdot x_1 - f^{**}(x_1)$. В итоге приходим к противоречию с неравенством Юнга: $x_2^* \cdot x_1 > f^*(x_2^*) + f^{**}(x_1)$. □

Литература

[1] P. de Fermat, Oeuvres de Fermat, vol. 1, Gauthier-Villars, Paris, 1891. Рус. пер. Ферма П. Метод отыскания наибольших и наименьших значений. В кн. Декарт. Геометрия. М.-Л.: Техтеорлит, 1938

[2] Euler L. Methodus inveniendi ... Lausanne, 1744. Рус. пер. Л. Эйлер. Метод нахождения кривых линий...М.-Л. Гостехтеориздат, 1934

[3] Lagrange J. L. Essai d'une nouvelle méthode pour determiner les maxima et les minima periales Petropolitanae, 1766, vol.10, 51-93

[4] Lagrange J. L. Théorie des fonctions analytiques, Paris, 1797.

[5] Legendre A. M. Mémoire sur la maniere de distingue les maxima des minima dans le calcul de variations/ Histoire de l'Academie Royallle des Sciences. Paris, 1786, pub 1788. 7–37

[6] Hamilton W. R. Second Essay on a General Method in Dynamics, Philos. Trans., 1835

[7] Jacobi C. G. J. Zur Theorie der Variations-Rechnung und der Differential-Gleichungen. Krelle's Journall Bd 17 (1837)

[8] Weierstrass K. Mathematische Werkw Bd. 7. Vorlesungemn uber Variationsrehtung. Berlin–Leipzig, Akad. Verlag., 1927

[9] Гильберт Д. Избранные труды. Т. II. Анализ. Физика. Проблемы. Personalia. М. Изд-ва «Факториал», 1998

a) Математические проблемы, 1901, стр. 401–436

b) О принципе Дирихле, 1901, стр. 13–34

c) О вариационном исчислении, 1906, стр. 351–370

[10] Minkowski H. Geometrie der Zahlen. Leipzig, Teubner, 1910

[11] Понтрягин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф. Математическая теория оптимальных поцессов. М.: Физматлит (1961)

Краткий курс теории экстремальных задач.
А.С. Кочуров, В.М. Тихомиров. – М.: Издательство Попечительского совета механико-математического факультета МГУ, 2013 – 48 с.

Подписано в печать 15.02.2013
Формат 60 × 909 1/16. Усл. печ. л. 3
Заказ 9. Тираж 150 экз.

Отпечатано на типографском оборудовании механико-математического факультета МГУ