

ИСТОРИЯ МАТЕМАТИКИ

УДК 519.6

ТЕОРИЯ ЭКСТРЕМУМА ОТ ФЕРМА ДО НАШИХ ДНЕЙ*

В. М. Тихомиров

Московский государственный университет им. М. В. Ломоносова,
Россия, 119899, г. Москва, Воробьевы горы, 1;
тел.: (095) 9395632; e-mail: Tikh@tikhomir.mcsme.ru

В статье дается краткий экскурс — от истоков до наших дней — в теорию экстремума. Обсуждаются необходимые и достаточные условия экстремума, а также принципы существования решений и алгоритмов их поиска.

Ключевые слова: теория экстремума, необходимые и достаточные условия экстремума, существование решений, алгоритм поиска.

В мире не происходит ничего, в чем не был бы виден смысл какого-нибудь максимума или минимума.

Л. Эйлер

Необходимость решать задачи на максимум и минимум вызывается стремлением людей по возможности наилучшим образом распорядиться своими ресурсами, желанием постичь законы природы, а также заложенной в природе человека любознательностью. Задачи на максимум и минимум называются экстремальными задачами (термин “экстремум” объединяет понятия “максимум” и “минимум”).

Формализованная экстремальная задача состоит из функционала $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ (где X — область определения функционала) и ограничения $C \subset X$. Такую задачу будем записывать в виде

$$f(x) \rightarrow \text{extr}, \quad x \in C. \quad (1)$$

Экстремальные задачи стали решать в глубокой древности, и стимулом служила упомянутая выше любознательность, прикладного значения эти решения не имели.

Практические нужды стимулировали создание общих методов исследования задач на экстремум. Первый общий прием был высказан Ферма в 1638 году. С того момента начала формироваться теория экстремальных задач. Эта теория складывается из следующих основных разделов:

- **Необходимые условия экстремума.**
- **Возмущения задач и достаточные условия.**
- **Расширения задач и существование решений.**
- **Алгоритмы решения задач на экстремум.**

Кратко коснемся всех этих тем.

* Статья подготовлена по материалам Второй Международной конференции “Функциональные пространства. Дифференциальные операторы. Проблемы математического образования”, приуроченной к 80-летию члена-корреспондента РАН профессора Л. Д. Кудрявцева.

I. НЕОБХОДИМЫЕ УСЛОВИЯ ЭКСТРЕМУМА

1.1. Задачи без ограничений. Теорема и принцип Ферма

Когда величина является максимальной или минимальной, в этот момент она не течет ни вперед, ни назад.

И. Ньютон

Пусть X — конечномерное или банахово пространство. Задача

$$f(x) \rightarrow \text{extr} \quad (2)$$

— это задача без ограничений.

Теорема (необходимое условие в задаче без ограничений). *Если в задаче (2) функция f непрерывно дифференцируема, тогда в точке локального экстремума производная функции f равна нулю.*

Первый общий прием для решения задач (2) (о нем упоминалось) был описан Ферма до рождения анализа. На нашем современном языке он звучит так: *в точке экстремума главная линейная часть приращения функции равна нулю.* В терминах производной выразили эту мысль создатели анализа — Ньютон и Лейбниц, и тем не менее сейчас этот результат связывают с именем Пьера Ферма (1601–1665).

Для Ньютона производная — это скорость, и он объяснил суть сформулированной теоремы в словах, приведенных в качестве эпиграфа к этому разделу; для Лейбница производная — тангенс угла наклона касательной, и геометрическая суть теоремы для него состояла в том, что *касательная в точке локального экстремума должна быть горизонтальна.*

Ферма в 1662 году высказал и первую натурфилософскую концепцию, согласно которой свет “избирает” такую траекторию (от точки к точке в неоднородной среде), чтобы минимизировать время.

Эта идея была подхвачена многими учеными в 18 столетии, и ей был придан всеобщий смысл, который можно сформулировать так: “*Мир управляется вариационными принципами.*”

Следующее великое имя в теории экстремума — это Леонард Эйлер (1707–1783).

1.2. Простейшая задача вариационного исчисления. Уравнение Эйлера

Пусть $L = L(t, x, y)$ — функция трех переменных. Задачу

$$\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt \rightarrow \text{extr}, \quad x(t_0) = x_0, \quad x(t_1) = x_1 \quad (3)$$

называют *простейшей задачей вариационного исчисления*. Функция L называется *интегрантом* задачи. Имеет место следующая

Теорема (необходимое условие в простейшей задаче). Если в задаче (3) интеграл L непрерывно дифференцируем, а функция $\hat{x}(\cdot)$ доставляет локальный экстремум задаче (3) (в пространстве непрерывно дифференцируемых функций), то выполнено уравнение Эйлера:

$$-\frac{d}{dt} L_{\dot{x}} + L_x = 0 \Leftrightarrow -\frac{d}{dt} L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) + L_x(t, \hat{x}(t), \dot{\hat{x}}(t)) = 0. \quad (4)$$

Считается, что вариационное исчисление ведет свою историю от 1696 года, когда Иоганн Бернулли (1667–1726) поставил задачу о брахистохроне (кривой наискорейшего спуска). Впоследствии он предложил своему ученику Леонарду Эйлеру попытаться найти какие-то общие методы решения задач, сходных с брахистохроной. Эйлер (1732) вывел уравнение (4) и решил с его помощью (и с помощью многочисленных обобщений этого уравнения) множество конкретных задач из геометрии и механики.

1.3. Гладкие задачи с ограничениями типа равенств.

Правило множителей Лагранжа

Можно высказать следующий общий принцип.

Если ищется максимум или минимум некоторой функции многих переменных при условии, что между этими переменными имеется связь, задаваемая одним или несколькими уравнениями, нужно прибавить к функции, о которой говорилось, функции, задающие уравнения связи, умноженные на неопределенные множители, и искать затем максимум и минимум построенной суммы, как если бы переменные были независимы. Полученные уравнения, присоединенные к уравнениям связи, послужат для определения всех неизвестных.

Ж. Лагранж (1797)

Пусть снова X — конечномерное или банахово пространство, f_i , $0 \leq i \leq m$ — гладкие функции на X . Задача

$$f_0(x) \rightarrow \text{extr}, \quad f_i(x) = 0, \quad 1 \leq i \leq m \quad (5)$$

называется *гладкой задачей с ограничениями типа равенств*. Общий рецепт для решения задач типа (5) был предложен Лагранжем (1736–1813). Он описал его (в 1797 г.) в словах, приведенных в качестве эпиграфа. Сформулируем некоторую модификацию метода Лагранжа.

Теорема (необходимое условие в задаче с ограничениями типа равенств). Если в задаче (5) все функции f_i дифференцируемы, то существуют множители $\{\lambda_i\}_{i=0}^m$, не равные одновременно нулю и такие, что в точке локального экстремума задачи производная по x функции $\mathcal{L} = \sum_{i=0}^m \lambda_i f_i(x)$ равна нулю.

Функция $\mathcal{L} = \sum_{i=0}^m \lambda_i f_i(x)$ называется *функцией Лагранжа*, множители $\{\lambda_i\}_{i=0}^m$ — *множителями Лагранжа*. (Отличие от того, что было предложено Лагранжем, в том, что и функция, экстремум которой находится, также умножается на неопределенный множитель. Это позволяет избежать лишних предположений в теореме.)

Так что рецепт Лагранжа можно сформулировать еще и так: *для решения задачи (5) составь функцию Лагранжа, примени теорему Ферма и решишь полученные уравнения. Этот рецепт называют правилом множителей Лагранжа.*

1.4. Необходимые условия экстремума от Ферма до наших дней

Необходимые условия для разных классов задач на экстремум доказывали многие математики. Назовем некоторые наиболее известные имена: Ферма, Эйлер, Лагранж, Лежандр, Гаусс, Остроградский, Якоби, Пуассон, Кнезер, Вейерштрасс, Больца, Макшейн, Блисс, Люстерник, Грейвс, Кун, Таккер, Понтрягин. . . Список этот, разумеется, далеко не полон. Большинство же необходимых условий экстремума составляются в соответствии с тем же *принципом Лагранжа*, который сформулирован в предыдущем разделе (правда, чуть более широко понимаемым). Этот принцип можно описать так: *составь функцию Лагранжа и примени необходимое условие в задаче без ограничений.*

Приведем два примера.

Пример 1. Задача Лагранжа вариационного исчисления

Рассмотрим задачу:

$$\int_{t_0}^{t_1} f(t, x(t), u(t)) dt \rightarrow \text{extr}, \quad \dot{x} = \varphi(t, x, u), \quad x(t_0) = x_0, \quad x(t_1) = x_1, \quad (6)$$

в которой $x \in \mathbb{R}^n$, $u \in \mathbb{R}^r$, f — функция $n + r + 1$ переменного, φ — n -мерная вектор-функция тех же переменных. Переменные x называют *фазовыми*, u — управлениями. Задачи (6) называются *задачами Лагранжа вариационного исчисления*. Приложим к ним рецепт Лагранжа.

Функция Лагранжа здесь имеет вид (так её составляют со времен Лагранжа):

$$\mathcal{L} = \int_{t_0}^{t_1} L(t, x, \dot{x}, u) dt, \quad L = \lambda_0 f(t, x, u) + p(t) \cdot (\dot{x} - \varphi(t, x, u)).$$

Применить необходимое условие в задаче на экстремум функции Лагранжа без ограничений — значит составить уравнения Эйлера по x и по u . В итоге приходим к уравнениям

$$-\frac{d}{dt} L_{\dot{x}} + L_x = 0, \quad L_u = 0, \quad (7)$$

которые в разных вариантах составлялись, начиная с брахистохроны (1696) по сороковые годы прошлого века (итоговая книга по этой тематике была опубликована в 1939 году американским математиком Г. Блиссом).

Если предположить, что f и φ достаточно гладкие, то можно доказать теорему, согласно которой *если пара $(\hat{x}(\cdot), \hat{u}(\cdot))$ доставляет локальный экстремум в задаче (6) в $C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$ (учитывается близость не только фазовых координат, но и управлений; такой экстремум называют *слабым*), то удовлетворяются уравнения (7).*

Пример 2. Задача оптимального управления

Пусть теперь функция f и отображение φ определены на произведении $[t_0, t_1] \times \mathbb{R}^n \times U$, где U — подмножество $\mathbb{R}^r$. При этом предположим, что f и φ непрерывны

по всем переменным и гладки по x . Рассмотрим ту же по форме задачу, что и (6), но дадим ей свой номер:

$$\int_{t_0}^{t_1} f(t, x(t), u(t)) dt \rightarrow \text{extr}, \quad \dot{x} = \varphi(t, x, u), \quad x(t_0) = x_0, \quad x(t_1) = x_1. \quad (8)$$

Задачу (8) называют *задачей оптимального управления*. Такие задачи стали исследовать на семинаре Л. С. Понтрягина (1908–1988) в пятидесятые годы прошлого века. Рассматривается случай сильного экстремума, для которого требуется лишь близость фазовых координат, а близость управлений не предполагается.

Множество U в (8) может быть произвольным, например, состоящим из конечного числа точек. Методы гладкого анализа, которые приводят к уравнениям (7), здесь не применимы. Но принцип Лагранжа сохраняет свою силу.

Функция Лагранжа задачи (8):

$$\mathcal{L} = \int_{t_0}^{t_1} L(t, x, \dot{x}, u) dt, \quad L = \lambda_0 f(t, x, u) + p(t) \cdot (\dot{x} - \varphi(t, x, u))$$

сохраняет ту же форму, что и в предыдущем примере. Рассмотрим две задачи, закрепив сначала $\hat{u}(\cdot)$ и минимизируя по $x(\cdot)$, а затем, закрепив $\hat{x}(\cdot)$ и минимизируя по $u(\cdot)$.

В первом случае задача минимизации относится к числу простейших задач вариационного исчисления, и необходимое условие в ней — это уравнение Эйлера по x :

$$-\frac{d}{dt} L_{\dot{x}} + L_x = 0. \quad (9)$$

Во втором случае получается задача

$$\int_{t_0}^{t_1} L(t, \hat{x}(t), \dot{\hat{x}}(t), u(t)) dt \rightarrow \min \text{ по } u(\cdot), \quad (10)$$

и при этом надо учитывать, что $u(t) \in U$. Очень легко понять, что (при самых широких допущениях) критерием минимальности $\hat{u}(\cdot)$ является “принцип минимума”:

$$\min_{u \in U} L(t, \hat{x}(t), \dot{\hat{x}}(t), u) = L(t, \hat{x}(t), \dot{\hat{x}}(t), \hat{u}(t)). \quad (11)$$

Если поменять знаки, то приходим к той форме, в которой необходимое условие для задачи (8) формулировалось в школе Понтрягина:

$$\begin{aligned} \max_{u \in U} \left(p(t) \cdot \varphi(t, \hat{x}(t), \dot{\hat{x}}(t), u) - \lambda_0 f(t, \hat{x}(t), \dot{\hat{x}}(t), u) \right) = \\ = \max_{u \in U} \left(p(t) \cdot \varphi(t, \hat{x}(t), \dot{\hat{x}}(t), \hat{u}(t)) - \lambda_0 f(t, \hat{x}(t), \dot{\hat{x}}(t), \hat{u}(t)) \right). \end{aligned} \quad (12)$$

Сочетание (9) и (12) называют *принципом максимума Понтрягина*.

Если применить принцип максимума Понтрягина ко второму дифференциалу функционала $J(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt$ простейшей задачи (функционал на кривой, на которой достигается минимум, должен быть неотрицателен, т. е. иметь минимум в нуле), то приходим к необходимым условиям (слабого минимума) Лежандра $(L_{\dot{x}\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) \geq 0)$ и Якоби (огибающая центрального пучка экстремалей, имеющих общую начальную точку, не должна пересекаться с экстремалью на полуинтервале $[t_0, t_1)$). Точка пересечения огибающей центрального поля с экстремалью называют *сопряженной точкой*.

Принцип максимума Понтрягина немедленно приводит и к необходимым условиям первого порядка для сильного минимума в простейшей задаче. Если применить принцип максимума Понтрягина к простейшей задаче (3), представив её в форме задачи оптимального управления

$$\int_{t_0}^{t_1} L(t, x(t), u(t)) dt \rightarrow \min, \quad \dot{x} = u, \quad x(t_0) = x_0, \quad x(t_1) = x_1, \quad (13)$$

то приходим к *необходимому условию Вейерштрасса* сильного минимума:

$$\mathcal{E}(t, \hat{x}(t), \dot{\hat{x}}(t), u) \geq 0 \quad \forall u \in \mathbb{R},$$

где $\mathcal{E}(t, x, y, u) = L(t, x, u) - L(t, x, y) - L_y(t, x, y)(u - y)$ — функция Вейерштрасса. К. Вейерштрасс (1815–1897) завершил теорию простейшей задачи, разработав теорию сильного минимума (дополнив необходимые условия достаточными).

Условия Лежандра, Якоби и Вейерштрасса оказались близки к достаточным в той же мере, как неотрицательность второго дифференциала (необходимое условие минимума в задаче без ограничений) и положительность его (как достаточное условие). Но само доказательство этих достаточных условий потребовало использования новой идеи.

II. ВОЗМУЩЕНИЯ ЭКСТРЕМАЛЬНЫХ ЗАДАЧ И ДОСТАТОЧНЫЕ УСЛОВИЯ

Следует сравнивать динамически возможные движения, варьируя крайние положения системы.

У. Гамильтон

Включение задачи в семейство задач, зависящее от каких-либо параметров, называется *возмущением* задачи. Первым математиком, осознавшим полезность возмущений задач при построении теории вариационного исчисления, был У. Гамильтон (1805–1865). Путеводной нитью для него была оптика, распространение света в неоднородной среде. Представляя себе мысленно пучок света, исходящего из одной точки, Гамильтон (1836) наряду с траекториями лучей света (удовлетворяющими вариационному принципу Ферма), стал рассматривать волновые фронты — линии уровня S -функции — времени достижения светом заданной точки пространства. “Варируя крайние положения системы”, он выписал уравнение в частных производных для S -функции в оптике. Годом позже К. Якоби (1804–1852) применил идеологию Гамильтона к любым задачам вариационного исчисления. Уравнение для S -функции простейшей задачи $S(\tau, \xi) = \inf\{J(x(\cdot)) \mid x(t_0) = x_0, x(\tau) = \xi\}$, имеющее вид

$$\frac{\partial S}{\partial t}(t, x) + \mathcal{H}\left(t, x, \frac{\partial S}{\partial x}(t, x)\right) = 0$$

(где $\mathcal{H}(t, x, y) = \sup\{yu - L(t, x, u) \mid u \in \mathbb{R}\}$ — преобразование Лежандра функции L), называется ныне *уравнением Гамильтона–Якоби*. Уравнение Гамильтона–Якоби позволяет выразить приращение функционала J на экстремали через функцию Вейерштрасса, и это приводит к достаточным условиям в вариационном исчислении, которые при условии регулярности интегранта (т. е. выпуклости его по последнему аргументу) состоят в усиленных условиях Лежандра и Якоби (когда неравенство в условии Лежандра строгое и сопряженной точки нет на всем отрезке $[t_0, t_1]$).

III. СУЩЕСТВОВАНИЕ РЕШЕНИЙ

Я убежден, что будет возможно доказывать теоремы существования с помощью общего принципа, чья сущность навеяна принципом Дирихле. Этот общий принцип, возможно приблизит нас к ответу на следующий вопрос: имеет ли решение каждая регулярная вариационная проблема, если самому понятию “решение” при случае придавать расширенное толкование?

Д. Гильберт

Это очень глубокое высказывание Гильберта (1862–1943) было произнесено им при формулировании двадцатой проблемы на конгрессе в Париже (1900). Общий принцип, о котором говорит Гильберт, это — безусловно *принцип компактности Вейерштрасса – Лебега – Бэра*, согласно которому *полунепрерывная снизу на компакте функция достигает своего минимума*.

Применение этого принципа проиллюстрируем на той же простейшей задаче, которая рассматривалась выше.

В одномерном случае (когда t одномерно) естественно в качестве пространства, на котором рассматривается задача (P_2) , выбрать в некотором отношении максимальное пространство, на котором сама задача может ставиться. А таковым можно считать пространство абсолютно непрерывных функций (или пространство $W_1^1([t_0, t_1])$). Регулярность интегранта гарантирует полунепрерывность снизу, а условие роста (интегрант по последнему аргументу должен расти быстрее линейной функции) и ограниченность снизу интегранта гарантируют компактность, а значит, в силу принципа компактности, и существование. Так формулируются многочисленные теоремы существования, начиная с работ Л. Тонелли двадцатых годов прошлого века.

Но в многомерных задачах естественного “самого широкого” пространства нет. И здесь стали строить пространства, которые разумно соответствуют интегрантам задачи. Например, в задаче Дирихле в области Ω таким пространством является, безусловно, соболевское пространство $W_2^1(\Omega)$. Так возникли *обобщенные функции*. А для решения вариационных задач прямыми методами (о которых кратко упомянем в следующем пункте), начали разрабатывать теорию вложения функциональных пространств, теорию, в которую так много внес Л. Д. Кудрявцев. Сам же стиль теорем существования остался тем же: если пространство строится по интегранту L , который гарантирует полунепрерывность и компактность, то имеет место существование решения. Во многих важных случаях полунепрерывность следует из регулярности, компактность из роста интегранта, а граничные условия должны браться из пространства, которое вкладывается в исходное (в задаче Дирихле на круге D это будет пространство $W_2^{1/2}(\partial D)$ на окружности).

IV. АЛГОРИТМЫ ОПТИМИЗАЦИИ

Таким образом, дело сводится к решению легкой задачи: даны две точки A и C и проходящая между ними горизонтальная прямая DE ; требуется найти на этой прямой точку B , чтобы путь ABC был наискорейшим.

Г. Лейбниц

(Из письма И. Бернулли 31 июля 1696 г. по поводу брахистохроны)

Эта цитата из Лейбница возвращает нас к истокам. Вариационное исчисление, как было сказано, родилось “из брахистохроны” — задачи, к решению

которой Иоганн Бернулли “приглашал” современных ему математиков. Решение (помимо самого Иоганна) представили его брат Яков, его ученик Лопиталь, а также праотцы всей современной математики Ньютон и Лейбниц. В этих решениях были зерна тех идей, которые питали и питают теорию экстремума от самых истоков до наших дней. Иоганн Бернулли решил задачу, базируясь на оптико-механической аналогии, которая вдохновила Гамильтона, а затем Якоби на построение их теории; Яков воспользовался принципом Гюйгенса, также составившим важнейший компонент теории Гамильтона – Якоби. А Лейбниц, как мы видим из приведенного эпиграфа, заложил основания “прямых методов”, заменив бесконечномерный объект — кривую — конечно-параметрическим — ломаной.

Продолжил “дело Лейбница” Эйлер, который вывел уравнение Эйлера, заменяя кривую ломаными. А в наше время редукция к конечномерным задачам — основной подход к численному решению задач на экстремум. После же редукции применяют методы целесообразного спуска (для задач минимизации) (градиентный метод и всевозможные его модификации, метод сопряженных направлений и т. д. и т. п., к этому же классу относится симплекс-метод Данцига), разнообразные методы штрафа, барьерные методы, в выпуклых задачах — методы отсечений. Остановиться здесь на всем этом подробно не представляется возможным.

В заключение приведу два высказывания.

Мир устроен просто. Очень просто. Но не более того.

А. Эйнштейн.

Это, безусловно, шутка, на самом деле Мир устроен, увы, непросто, и Эйнштейн об этом знал. Но отдельные “фрагменты” нашего мира устроены на самом деле не так уж сложно. Таков, например, мир экстремума. Это я и старался показать в своей статье.

И еще — цитата, которую я очень люблю.

Если ничто не собьет нас с пути, то мы уподобимся Зигфриду, перед которым огненный вал расступается сам собою.

Д. Гильберт.

ЛИТЕРАТУРА

1. Иоффе А. Д., Тихомиров В. М. Теория экстремальных задач. — М.: Наука, 1974.
2. Алексеев В. М., Тихомиров В. М., Фомин С. В. Оптимальное управление. — М.: Наука, 1979.

EXTREMUM THEORY FROM FERMAT TILL OUR DAYS

V. M. Tikhomirov

The author gives a brief excursus into extremum theory development (from the beginning till our days). He considers the necessary and sufficient conditions of extremum and principles of existence of solutions and search algorithm.

Keywords: extremum theory, necessary and sufficient conditions of extremum, existence of solutions, search algorithm.