УДК 517.5

ОСНОВНЫЕ ПРИНЦИПЫ ТЕОРИИ ЭКСТРЕМУМА

© В. М. Тихомиров

Ключевые слова: гладкость; выпуклость; принцип Лагранжа; выпуклая двойственность; компактность.

Статья посвящена некоторым принципам теории экстремума: Ферма и Лагранжа в необходимых условиях, Гамильтона в достаточных условиях, двойственности в теории выпуклости, а также некоторым общим идеям в теории существования и конструировании алгоритмов оптимизации.

В статье приводятся основные положения моего доклада на заседании международной конференции «Колмогоровские чтения. ОПУ-2013», посвященном памяти А.И. Булгакова.

В этом докладе мне хотелось бы подвести некоторые итоги моей деятельности в теории экстремума.

Основные разделы теории:

- теория гладких задач;
- вариационное исчисление;
- теория выпуклых и ляпуновских задач;
- теория оптимального управления.

Роль теории экстремума в математике и ее приложениях

Вариационное исчисление лежит в основании математического естествознания, теория оптимального управления — в основании управления динамическими системами, а теория выпуклости — в основании математической экономики.

Структура теории

Основными главами теории экстремума являются:

- *база теории*, состоящая из *исчислений* (дифференциального и выпуклого исчисления, а также исчисления дифференциальных уравнений)

- *необходимые условия экстремума*

-*устойчивость решений и достаточные условия экстремума*

-*теория существования решений*

-*алгоритмы оптимизации.*

Зародыши основных идей доклада содержались в нашей совместной с А.Д. Иоффе монографии [1], написанной сорок лет тому назад. В этой монографии была сделана попытка осмыслить всю историю экстремума с неких общих позиций.

Необходимые условия экстремума. Принцип Лагранжа

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

Эта идея была аналитически оформлена Ферма, Ньютоном и Лейбницем. Назовем ее **принципом Ферма**.

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.-Л. Лагранж*

Необходимые условия экстремума (т. е. минимума или максимума) со времен Ферма и Лагранжа и до наших дней в задачах, где воедино слиты гладкая и выпуклая структуры и присутствует некая «невырожденность», соответствуют одному общему принципу **принципу Лагранжа снятия ограничений.**

Принцип Лагранжа заключается в том, что в задачах с ограничениями с гладкой и/или выпуклой структурой для выписывания необходимого условия экстремума, следует составить т. н. *функцию Лагранжа*, включающую в себя ограничения, и применить к ней *принцип Ферма* необходимые условиям экстремума в более простой задаче в задаче на экстремум функции Лагранжа, «как если бы переменные были независимы».

Роль выпуклости в теории экстремума во многом связана с *феноменом выпуклости*, вызванной интегрированием. В конечномерном случае знаменательным подтверждением этого феномена может служить теорема А. А. Ляпунова [2]. Вот ее формулировка.

Пусть Δ *это сегмент на прямой* $\mathbb{R}$, $\overline{u}(\cdot)$ $(u_1(\cdot),\dots,u_N(\cdot))$ *интегрируемая вектор-функция,* $\mathcal{A}$ σ*-алгебра измеримых по Лебегу множеств. Тогда образ отображения* $A \mapsto$ $\mapsto \int_A \overline{u}(t)dt$ *(т. е. множество* M $\{x \in \mathbb{R}^n \mid \int_A \overline{u}(t)dt,\ A \in \mathcal{A}\}$*)* *выпуклый компакт в* $\mathbb{R}^N$.

Это лишь один из фактов, подтверждающих *феномен выпуклости*, который несколько вольно можно выразить так: *интегральные операторы, определенные на пространствах с непрерывной мерой, имеют выпуклые или «почти выпуклые» образы* (такой «почти выпуклостью» обладают отображения классического вариационного исчисления и оптимального управления).

В конце тридцатых и в сороковые годы началось интенсивное развитие ветви, занимающей промежуточное положение между математическим анализом и геометрией, получившей название *выпуклого анализа*. Это развитие было во многом стимулировано рождением линейного программирования, возникшего в трудах Канторовича Л.В. по математической экономике, начавшихся с работы Канторович Л.В. Математические методы организации и планирования производства. Л: Изд-во ЛГУ, 1939. Возрождение выпуклого анализа (некоторые начальные разделы его возникли еще в XIX веке), во многом связано с появлением работы В. Фенхеля 1951 г. [3].

Придадим высказанной выше эвристической идее, названной принципом Лагранжа, форму математического результата.

Принцип Лагранжа для гладко-выпуклых задач

Требования гладкости в формулируемой далее теореме выражаются через два понятия дифференциального исчисления (относящегося к **базе теории**).

Пусть X и Y нормированные пространства, $V \subset X$ окрестность точки $\widehat{x} \in X$ и $F: V \to Y$. Говорят, что отображение F *дифференцируемо (по Фреше)* [4] *в точке* $\widehat{x}$ (и пишут $F \in D^1(\widehat{x})$), если существует линейный непрерывный оператор $\Lambda: X \to Y$ такой, что для любого $\varepsilon > 0$ найдется число $\delta > 0$, для которого

$$\|x - \widehat{x}\|_X < \delta \ \Rightarrow\ \|F(x) - F(\widehat{x}) - \Lambda(x - \widehat{x})\|_Y < \varepsilon \|x - \widehat{x}\|_X.$$

Оператор Λ, входящий в эту формулу, называется *производной отображения* F *в точке* $\widehat{x}$. Его обозначают $F'(\widehat{x})$.

Говорят, что отображение F *строго дифференцируемо в точке* $\widehat{x}$ (и пишут $F \in SD^1(\widehat{x})$), если $F \in D^1(\widehat{x})$ и для любого $\varepsilon > 0$ найдется число $\delta > 0$, для которого из $\|\xi - \widehat{x}\|_X < \delta$ и $\|x - \widehat{x}\|_X < \delta$ следует неравенство $\|F(\xi) - F(x) - \Lambda(\xi - x)\|_Y < \varepsilon \|\xi - x\|_X$ (это определение принадлежит Е. Личу: Leach E. Proc. AMS. V. 12 (1961). P. 694-697).

Теперь можно переходить к формулировке теоремы.

Теория необходимых условий опирается далее у нас на два результата: теорему Люстерника [5] (т. е. в конечном счете теорему о неявной функции) из дифференциального исчисления и теорему отделимости из теории выпуклости. Теорема отделимости порождает **принцип двойственности** выпуклых объектов, согласно которому выпуклые множества функции и задачи *допускают двойное описание* (в основном и двойственном пространстве).

Пусть X и Y — нормированные пространства, $\mathcal{U}$ — некоторое множество, $V \subset X$ — окрестность точки $\widehat{x}$, $f_0 : V \to \mathbb{R}$ и $F : V \times \mathcal{U} \to Y$. Рассмотрим задачу:

$$f_0(x) \to \min; \quad F(x, u) = 0. \tag{P}$$

Далее на f_0 и F накладываются требования гладкости по x и выпуклости по u, и потому задачу (P) с такими условиями мы называем *гладко-выпуклой.*

Необходимое условие экстремума в задаче без ограничений

Если ограничение $F = 0$ в задаче отсутствует, задача (P) называется *задачей без ограничений.* Имеет место следующее предложение

Если в задаче (P) без ограничений точка $\widehat{x}$ доставляет локальный экстремум и в этой точке f_0 дифференцируема, то $f_0'(\widehat{x}) = 0$.

Этот результат часто называют *теоремой Ферма.*

Первое общее правило исследования задач на экстремум содержится в письме Ферма Декарту (1638) [6]. Оно было переведено на язык математического анализа Ньютоном и Лейбницем.

Суть дела замечательно выразили Кеплер, Ферма и Ньютон.

Кеплер (вычислитель): *Вблизи максимума и минимума изменения несущественны* [7].

Ферма (аналитик): *В точке экстремума линейная аппроксимация нулевая* [6].

Ньютон (естествоиспытатель): *Когда величина является максимальной или минимальной, в этот момент она не течет ни назад ни вперед.* (Ньютон И. Метод флюксий и бесконечных рядов. Ньютон И. Математические работы. М.: УРСС, 2012. С. 73.).

Необходимое условие в задачах с ограничениями

Пара $(\widehat{x}, \widehat{u})$ называется *(сильным) локальным минимумом* в задаче (P), если существует такое число $\varepsilon > 0$, что $f_0(x) \geq f_0(\widehat{x})$ для всякой пары (x, u), для которой $F(x, u) = 0$ и $\|x - \widehat{x}\|_X < \varepsilon$.

Функцию Лагранжа задачи (P) определим равенством

$$\mathcal{L}((x, u), \overline{\lambda}) = \lambda_0 f_0(x) + \langle \lambda, F(x, u) \rangle,$$

где $\overline{\lambda} = (\lambda_0, \lambda) \in \mathbb{R} \times Y^*$ (Y^* — пространство, сопряженное к Y) — *набор множителей Лагранжа.* Имеет место следующий результат:

Принцип Лагранжа для гладко-выпуклых задач [1]. *Пусть выполнены следующие условия: a) X и Y — банаховы пространства (условие полноты), b) $f_0 \in D^1(\widehat{x})$, а отображения $x \mapsto F(x, u)$ принадлежат $SD^1(\widehat{x})$ $\forall u \in \mathcal{U}$ (условия гладкости), c) множества $F(x, \mathcal{U})$ выпуклы в Y $\forall x \in V$ (условие выпуклости), d) подпространство $F_x(\widehat{x}, \widehat{u})X$ замкнуто в Y и имеет конечную коразмерность (условие слабой регулярности). Тогда, если пара $(\widehat{x}, \widehat{u})$ доставляет локальный минимум в задаче (P), то существует такой ненулевой набор множителей Лагранжа $\overline{\lambda} = (\lambda_0, \lambda)$, что для задачи (P) выполнены условие стационарности:*

$$\mathcal{L}_x((\widehat{x}, \widehat{u}), \overline{\lambda}) = 0 \Leftrightarrow \lambda_0 f_0'(\widehat{x}) + F_x(\widehat{x}, \widehat{u}))^* y^* = 0, \tag{1}$$

и принцип минимума

$$\min_u \mathcal{L}((\widehat{x}, u), \overline{\lambda}) = \mathcal{L}((\widehat{x}, \widehat{u}), \lambda) \Leftrightarrow \min_u \langle y^*, F(\widehat{x}, u) \rangle = 0. \tag{2}$$

Соотношения (1) и (2) находятся в соответствии с принципом Лагранжа, о котором говорилось выше: условие стационарности — это не что иное, как применение теоремы Ферма к задаче $\mathcal{L}((x,\widehat{u})\overline{\lambda}) \to \min$, а принцип минимума есть просто тавтологическое условие минимальности $y \mapsto \langle\lambda, y\rangle$ при $y \in F(\widehat{x}, u)$, $u \in \mathcal{U}$.

Доказательство (в случае, когда: $F_x(\widehat{x},\widehat{u})X = Y$). Пусть $u \in \mathcal{U}$ и ξ таковы, что

$$F_x(\widehat{x},\widehat{u})\xi + F(\widehat{x},u) = 0. \qquad (i)$$

Положим $\Phi(x,\alpha) = (1-\alpha)F(\widehat{x}+x,\widehat{u}) + \alpha F(\widehat{x}+x,u)$. Дифференцируя (с использованием условия гладкости), получаем:

$$\Phi'(0,0)[x,\alpha] = F_x(\widehat{x},\widehat{u})x + \alpha F(\widehat{x},u). \qquad (ii)$$

Из (i) и (ii) следует, что $(\xi, 1) \in \mathrm{Ker}\Phi'(0,0)$, и тогда, применив **теорему Люстерника**, получаем, что

$$\Phi(t\xi + r(t), t + \rho(t)) = 0, \ \ |t| \le \varepsilon, \ \ \|r(t)\|_X = o(t), \ \ |\rho(t)| = o(t).$$

Это соотношение можно переписать так:

$$(1-\alpha(t))F(\widehat{x} + x(t),\widehat{u}) + \alpha(t)F(\widehat{x}+x(t),u) = 0, \qquad (iii)$$

где $\alpha(t) = t + \rho(t)$, $x(t) = \widehat{x} + t\xi + r(t)$. Используя условие выпуклости, найдем такой элемент $u(t)$, для которого

$$F(x(t),u(t)) = (1-\alpha(t))F(x(t),\widehat{u}) + \alpha(t)F(x(t),u) = 0, \ |t| \le \varepsilon. \qquad (iv)$$

Это означает, что $(x(t),u(t))$ — допустимый элемент, и значит,

$$f(x(t)) = f(\widehat{x} + t\xi + \rho(t)) \ge f(\widehat{x}) \Rightarrow \langle f'(\widehat{x}), \xi\rangle \ge 0. \qquad (v)$$

Пусть теперь $\xi \in \mathrm{Ker}F_x(\widehat{x},\widehat{u})$. Тогда, положив в (i) $u = \widehat{u}$, получим, что (i) выполнено, и значит, в силу (v), $\langle f'(\widehat{x}),\xi\rangle \ge 0$. Но $-\xi$ также принадлежит $\mathrm{Ker}F_x(\widehat{x},\widehat{u})$, откуда $\langle f'(\widehat{x}),\xi\rangle \le 0$ и значит, $\langle f'(\widehat{x}),\xi\rangle = 0$, т. е. $f'(\widehat{x}) \in (\mathrm{Ker}F_x(\widehat{x},\widehat{u}))^{\perp}$. По следствию **теоремы отделимости** (о ядре регулярного оператора), существует такой элемент $\lambda \in Y^*$, что $f'(\widehat{x}) + F_x^*(\widehat{x},\widehat{u})\lambda = 0$, а это — не что иное, как условие стационарности для функции Лагранжа. Пусть теперь $v \in \mathcal{U}$ и $\xi(v)$ таковы, что $F_x(\widehat{x},\widehat{u})\xi(v) + F(\widehat{x},v) = 0$. Действуя линейным функционалом $f'(\widehat{x}) + F_x^*(\widehat{x},\widehat{u})\lambda$ на $\xi(v)$ и используя (v) и условие стационарности, приходим к условию минимума:

$$\langle\lambda, F(x,v)\rangle = -\langle\lambda, F_x(\widehat{x},\widehat{u})\xi(v)\rangle = -\langle F_x^*(\widehat{x},\widehat{u})\lambda, \xi(v)\rangle = \langle f'(\widehat{x}), \xi(v)\rangle \ge 0,$$

т. е. $f(\widehat{x}) + \langle\lambda, F(\widehat{x},v)\rangle \ge f(\widehat{x})$, и значит, в силу того, что $F(\widehat{x},\widehat{u}) = 0$, получаем, что $\mathcal{L}(\widehat{x},v,\overline{\lambda}) \ge \mathcal{L}(\widehat{x},\widehat{u},\overline{\lambda})$. □

Принцип Лагранжа для конкретных классов задач

Следствие 1 (принцип Лагранжа для гладких задач). *Если в задаче*

$$f_0(x) \to \min, \ \ f_i(x) \le 0, \ \ 1 \le i \le m, \ \ \mathcal{F}(x) = 0 \qquad (P_1)$$

функции и отображение определены в окрестности некоторой точки $\widehat{x}$ *банахова пространства* X, *являются строго дифференцируемыми в этой точке, и при этом отображение* $\mathcal{F}$ *определено в той же окрестности, отображает ее в банахово пространство* Y *и при этом*

$\mathcal{F}'(\widehat{x})X$ является замкнутым подпространством в Y, тогда, если в точке $\widehat{x}$ достигается локальный минимум, то найдется набор множителей Лагранжа $\overline{\lambda} = (\lambda_0, \dots, \lambda_m, \lambda) \in \mathbb{R}^m \times \times Y^*$, такой, что для функции Лагранжа $\mathcal{L}(x, \overline{\lambda}) = \sum\limits_{i=0}^{m} \lambda_i f_i(x) + \langle \lambda, \mathcal{F}(x) \rangle$ выполнены условия стационарности: $\mathcal{L}_x(\widehat{x}, \overline{\lambda}) = 0$, неотрицательности: $\lambda_i \geq 0,\ i \geq 0$ и дополняющей нежесткости: $\lambda_i f_i(\widehat{x}) = 0,\ i \geq 1$.

Для доказательства следует применить гладко-выпуклый принцип к задаче (P), в которой $F(x, u) = (f_1(x) + u_1, \dots, f_m(x) + u_m, \mathcal{F}(x))$, $\mathcal{U} = \mathbb{R}^m_+$.

При этом сам результат можно сформулировать так: *необходимое условие минимума в задаче (P_1) находится в соответствии с принципом Лагранжа.*

Отметим, что для доказательства следствия 1 достаточно использовать лишь доказанный ослабленный вариант гладко-выпуклого принципа: если $\mathcal{F}'(\widehat{x})X$, то надо применить **теорему отделимости**, а в регулярном случае надо последовательно присоединять к $\mathcal{F}$ (применяя **теорему о неявной функции**) или отбрасывать функции $f_i,\ i \geq 1$.

Следствие 2 (принцип Лагранжа для выпуклых задач). *Если в задаче*

$$f_0(x) \to \min, \quad f_i(x) \leq 0, \ 1 \leq i \leq m', \quad f_i(x) = 0, \ i > m' \qquad (P_2)$$

функции f_i при $i \leq m$ выпуклы и их значения точке $\widehat{x}$ конечны, а функции $f_i,\ i > m'$ аффинны, и при этом в $\widehat{x}$ достигается абсолютный минимум в задаче (P_2), то найдется набор множителей Лагранжа $\overline{\lambda} = (\lambda_0, \dots, \lambda_m) \in \mathbb{R}^m$, такой, что для функции Лагранжа $\mathcal{L}(x, \overline{\lambda}) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ выполнены условия минимума: $\min\limits_{x \in A} L(x, \overline{\lambda}) = L(\widehat{x}, \overline{\lambda})$, неотрицательности: $\lambda_i \geq \geq 0,\ 0 \leq i \leq m'$ и дополняющей нежесткости: $\lambda_i f_i(\widehat{x}) = 0,\ 1 \leq i \leq m'$.

Для доказательства следует применить гладко-выпуклый принцип к задаче (P), в которой $F(x, u) = (f_1(x) + u_1, \dots, f_m(x) + u_m, \mathcal{F}(x))$, $\mathcal{U} = \mathbb{R}^m_+ \times A$.

При этом сам результат можно сформулировать так: *необходимое условие минимума в задаче (P_2) находится в соответствии с принципом Лагранжа.*

Принцип исследования гладких конечномерных задач с равенствами был впервые высказан Лагранжем (в приводившейся выше цитате) в 1797 г. [8], и получил название *правила множителей Лагранжа.* Правило множителей Лагранжа в банаховом случае для задач с равенствами было доказано Л.А. Люстерником в 1934 г. [5], а общий результат (следствие 1) был впервые получен А.Я. Дубовицким и А.А. Милютиным в 1965 г. [9].

Следствие 2 было доказано Карушем в 1939 г. [10], но осталось незамеченным. Оно было передоказано и вошло в обиход развивавшейся тогда теории выпуклости Куном и Таккером в 1951 г. [11].

Принцип Лагранжа для задач вариационного исчисления и оптимального управления

Пусть $\Delta = [t_0, t_1]$, $-\infty < t_0 < t_1 < \infty$, $U \subset \mathbb{R}^r$, $L_i : \Delta \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$, $0 \leq i \leq m$, $l_i : \mathbb{R}^n \times \mathbb{R}^n \to \to \mathbb{R}$, $0 \leq i \leq m$, $\varphi : \Delta \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$, $\Xi = \{\xi = (x(\cdot), u(\cdot))\}$, определим

$$f_i(\xi) = \int_\Delta L_i(t, x(t), u(t))\, dt + l_i(x(t_0), x(t_1)), \quad 0 \leq i \leq m.$$

Задачу

$$f_0(\xi) \to \min, \ f_i(\xi) = 0, \ 1 \leq i \leq m, \ \dot{x} = \varphi(t, x, u), \ u(t) \in U, \qquad (P_3')$$

назовем *задачей оптимального управления в понтрягинской форме.* Задачу

$$g_0(u(\cdot)) = \int_\Delta h(t, u(t))dt \to \min, \quad u(t) \in U$$

назовем *элементарной задачей оптимального управления (или элементарной ляпуновской задачей).*

Задачу оптимального управления в понтрягинской форме будем рассматривать на множестве $PC^1(\Delta, \mathbb{R}^n) \times PC(\Delta, \mathbb{R}^n)$ кусочно-непрерывно дифференцируемых $x(\cdot)$ и кусочно-непрерывных $u(\cdot)$. Говорят, что пара $(\widehat{x}(\cdot), \widehat{u}(\cdot)) \in PC^1(\Delta, \mathbb{R}^n) \times PC(\Delta, \mathbb{R}^n)$ является *оптимальным процессом в задаче* (P_3) или *доставляет задаче* (P_3) *сильный локальный минимум,* если найдется число $\varepsilon > 0$ такое, что для любой допустимой пары $(x(\cdot), u(\cdot))$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C(\Delta, \mathbb{R}^n)} < \varepsilon$, выполнено неравенство $f_0(x(\cdot), u(\cdot)) \geq f_0(\widehat{x}(\cdot), \widehat{u}(\cdot))$.

Если $U = \mathbb{R}^r$, задачу (P_3) называют *задачей Лагранжа классического вариационного исчисления.* Задачу

$$f_0(\xi) = \int\limits_{\Delta} L_0(t, x(t), u(t))dt + l_0(x(t_0), x(t_1)) \to \min$$

называют *элементарной задачей вариационного исчисления* (или *задачей Больца*).

Сформулируем условия минимума в элементарных задачах.

a) Пусть $\widehat{u}(\cdot)$ — *кусочно-непрерывная функция на* Δ *со значениями в* $\mathbb{R}^r$, *а функция* $h(\cdot,\cdot)$ *непрерывна на* $\Delta \times U$. *Тогда для того чтобы функция* $\widehat{u}(\cdot)$ *доставляла абсолютный минимум в элементарной задаче оптимального управления, необходимо и достаточно, чтобы в любой точке* t *непрерывности* $\widehat{u}(\cdot)$ *выполнялось условие минимума* $h(t, u) \geq h(t, \widehat{u}(t))$ $\forall u \in U$.

b) Пусть в задаче Больца интегрант L *является непрерывной функцией по совокупности переменных и непрерывно-дифференцируемой функцией по* x *и* $\dot{x}$ *в окрестности графика* $\{(t, \widehat{x}(t), \dot{\widehat{x}}(t), | t \in \Delta\}$ *(где* $\widehat{x}(\cdot) \in C^1(\Delta)$ *), а функция* l *непрерывно дифференцируема в окрестности точки* $(\widehat{x}(t_0), \widehat{x}(t_1))$. *Тогда если функция* $\widehat{x}(\cdot)$ *доставляет слабый минимум задаче Больца, то* $\widehat{L}_{\dot{x}}(\cdot) \in C^1([t_0, t_1])$ *и выполняется уравнение Эйлера*

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

и условия трансверсальности

$$\widehat{L}_{\dot{x}}(t_i) = (-1)^i \widehat{l}_{x(t_i)}, \quad i = 0, 1,$$

где $\widehat{L}_{\dot{x}}(t) = L_{\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$, $\widehat{L}_x(t) = L_x(t, \widehat{x}(t), \dot{\widehat{x}}(t))$, $\widehat{l}_{x(t_i)} = l_{x(t_i)}(\widehat{x}(t_0), \widehat{x}(t_1))$, $i = 0, 1$.

Функцию Лагранжа задачи (P_3) определим равенством

$$\mathcal{L}((x(\cdot), u(\cdot)), \overline{\lambda}) = \sum_{i=0}^{m} \lambda_i f_i((x(\cdot), u(\cdot))) + \int_{t_0}^{t_1} p(t) \cdot (\dot{x}(t) - \varphi(t, x(t), u(t))dt =$$

$$= \int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t), u(t))dt + l(x(t_0), x(t_1)),$$

где $L(t, x, \dot{x}, u) = \widetilde{L}(t, x, u) + p(t) \cdot (\dot{x} - \varphi(t, x, u)$, $\widetilde{L}(t, x, u) = \sum\limits_{i=0}^{m} \lambda_i f_i(t, x, u)$, а $l(x(t_0), x(t_1)) = \sum\limits_{i=0}^{m} \lambda_i l_i(x(t_0), x(t_1))$ (ограничение $u(t) \in U$ в функцию Лагранжа не вносится).

Следствие 3 (принцип Лагранжа для задачи Лагранжа и задачи оптимального управления). *Если в задаче* (P_3) *выполнены требования гладкости, то необходимое условие минимума в них находится в соответствии с принципом Лагранжа.*

В задаче оптимального управления условия гладкости на L_i и φ состоят в том, что эти функции должны быть непрерывны по совокупности переменных и непрерывно дифференцируемы по x, а в задаче Лагранжа они должны быть непрерывно-дифференцируемы и по x и по u в окрестности графика $\{(t, \widehat{x}(t), \dot{\widehat{x}}(t), \mid t \in \Delta\}$, а l_i должны быть непрерывно дифференцируемы в окрестности точки $(\widehat{x}(t_0), \widehat{x}(t_1))$.

Теперь остается лишь расшифровать формулировку следствия 3. Согласно принципу Лагранжа, надо рассмотреть две задачи:

$\mathcal{L}((x(\cdot), \widehat{u}(\cdot), \overline{\lambda}) \to \min$ и $\mathcal{L}((\widehat{x}(\cdot), u(\cdot), \overline{\lambda}) \to \min$ и написать для них необходимые условия минимума, «как если бы переменные были независимы».

Таким образом, если в точке $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ достигается соответствующий локальный минимум в задаче (P_3), то найдется отличный от нуля набор множителей Лагранжа $\overline{\lambda} = (\lambda_0, \ldots$ $\ldots, \lambda_m, p(\cdot))$, такой, что удовлетворяются условия неотрицательности ($\lambda_i \geq 0$, $0 \leq i \leq m'$), дополняющей нежесткости ($\lambda_i f_i(\widehat{\xi}) = 0$, $1 \leq i \leq m'$) и, в соответствии с необходимыми условиями для задачи Больца, уравнения Эйлера–Лагранжа по $x(\cdot)$:

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \;\Rightarrow\; -\dot{p}(t) = p(t)\widehat{\varphi}_x(t) - \widehat{L}_x(t) = 0 \qquad (2a)$$

и условия трансверсальности

$$\widehat{L}_{\dot{x}}(t_i) = \widehat{l}_{\xi_i}, \quad i = 0, 1 \qquad (2b)$$

Необходимое условие во второй задаче о минимизации функции Лагранжа по $u(\cdot)$ для задачи оптимального управления, в соответствии с необходимым условием для элементарной задачи оптимального управления, приводит к условию минимума:

$$\min_{u \in U} L(t, \widehat{x}(t), \dot{\widehat{x}}(t), u) = \widehat{L}(t), \qquad (3)$$

а в задаче Лагранжа задача о минимизации функции Лагранжа по $u(\cdot)$ является задачей Больца и, в соответствии с необходимым условием для нее, приходим к уравнению Эйлера по u:

$$\widehat{L}_u(t) = 0. \qquad (3')$$

Аналог теоремы Ферма в вариационном исчислении (уравнение Эйлера) впервые получил Эйлер [12]. Правило множителей Лагранжа к задачам вариационного исчисления Лагранж применял с самого начала своего творчества (см. например, [13]). Так возникли в вариационном исчислении уравнения Эйлера–Лагранжа (2a) и (3$'$). Условия трансверсальности (2b) возникли с появлением других видов формализации задач вариационного исчисления (задач Майера, Больца и др.). Все это нашло отражение в первом фундаментальном учебнике по вариационному исчислению [14]. Бурное развитие вариационного исчисления в течение двух веков, начиная с работы Эйлера [12] до середины XX века, во многом было стимулировано тем, что многие законы природы являются следствием *вариационных принципов*. В середине XX века появилось *оптимальное управление*, включающее в себя вариационное исчисление, как часть. Родоначальником новой теории был Понтрягин Л.С., который со своими учениками Болтянским В.Г., Гамкрелидзе Р.В. и Мищенко Е.Ф. в 1961 г. выпустил монографию [15]. В ней уравнения Эйлера–Лагранжа были дополнены условием (3), получившим, совместно с (2a), название *принципа максимума Понтрягина*.

Некоторые комментарии

1. О доказательствах. Вывод уравнений Эйлера–Лагранжа для задачи Лагранжа вариационного исчисления не требует новых идей и методов: при достаточной гладкости условий в задаче Лагранжа она становится гладкой задачей и справедливость принципа Лагранжа вытекает из следствия 1. Задачи оптимального управления «линейные по фазе» редуцируются к ляпуновским (т. е., по сути дела, к выпуклым задачам) и принцип Лагранжа для них вытекает из следствия 2. Для вывода принципа максимума Понтрягина нужно чуть модифицировать доказанный выше гладко-выпуклый принцип, в котором условие регулярности сочетается не с выпуклостью, а с «почти выпуклостью».

2. О конкретных задачах. Начиная с классической изопериметрической задачи, обсуждавшейся еще Аристотелем в V веке до н. э. по наши дни, было исследовано огромное количество конкретных экстремальных задач (бездна геометрических задач на экстремум, аэродинамическая задача Ньютона, задача Бернулли о брахистохроне, задачи о неравенствах (которым посвящена монография Харди Литтлвуда Пойа, экстремальные задачи теории аппроксимации и т. д., и т. п.). Безусловным мотивом создания *теории экстремума* является разработка средств для решения конкретных задач. По мнению докладчика (который демонстрировал это во многих своих книгах), подавляющее число «явно» решенных (без помощи компьютеров) конкретных задач решаемы методом Лагранжа, который (при удачной формализации) приводит к разрешимым уравнениям. А если уравнения не решаются, то методы оптимизации при современной компьютерной технике позволят для почти всех задач получить (хотя бы приближенный) ответ.

Устойчивость решений и достаточные условия экстремума

Необходимое условие «первого порядка» для минимума в точке $\widehat{x}$ в гладкой задаче $f(x) \to$ $\to \min$, состоящее в условии стационарности $f'(\widehat{x}) \quad 0$, было дополнено Ньютоном и Лейбницем условием «второго порядка» $f >> (\widehat{x}) \geq 0$. При этом условие $f >> (\widehat{x}) > 0$ является достаточным условием минимума и гарантирует устойчивость этого минимума при малых возмущениях.

Начальные необходимые условия второго порядка в задачах вариационного исчисления были получены Лежандром (условия Лежандра) [16], но путь к достаточным условиям оказался достаточно далек. Для простейшей задачи вариационного исчисления этот вопрос был решен усилиями Гамильтона [17], Якоби [18] и Вейерштрасса [19]. Фундаментальную роль в этом круге вопросов играет следующая идея Гамильтона о том, что *следует включать отдельную экстремаль в семейство* (стали говорить — *поле*) *экстремалей.*

Посмотрим на эту идею с «общих позиций», рассмотрев задачу

$$f(x) \to \min, \quad F(x) \quad 0, \tag{1}$$

где f функция на гильбертовом пространстве H_1, а F отображение из H_1 в гильбертово пространство H_2. Пусть $\mathcal{L}(x,\eta) \quad f(x) + \langle \eta, F(x) \to$ функция Лагранжа задачи (1) (с λ_0 $= 1$). Рассмотрим «стандартное возмущение» задачи (1):

$$f(x) \to \min, \quad F(x) = y. \tag{2}$$

Теорема о поле экстремалей. *Пусть производная F' в точке $\widehat{x}$ сюръективна, существует множитель Лагранжа $\widehat{\eta}$ такой, что $\mathcal{L}_x(\widehat{x},\widehat{\eta}) \quad 0$ и при этом*

$$\mathcal{L}_{xx}(\widehat{x},\widehat{\eta})[x,x] \geq \alpha\|x\|^2 \quad \forall x \in \mathrm{Ker}F'(\widehat{x}).$$

Тогда существуют окрестность $W \subset H_1$ точки $0 \in H_1$ и отображение $y \mapsto (x(y),\eta(y))$, определенное на W, такие, что $\mathcal{L}_x(x(y),\eta(y)) = 0$, $x(0) = \widehat{x}$, $\eta(0) = \widehat{\eta}$ и $F(x(y)) = y$.

Принцип компактности в теории существования

Одним из основных принципов доказательства теорем существования решения задач на экстремум является принцип компактности, когда при выборе этого пространства функционал (или его расширение) полунепрерывен снизу, а ограничение компактно. При этом, существует убеждение, что почти каждая естественным образом поставленная задача имеет решение, возможно, в некоем естественном же обобщенном толковании этого понятия.

Вторая часть нашего тезиса фактически является пересказом мысли Гильберта, высказанной им при формулировке его 22 проблемы, в которой обсуждались задачи вариационного исчисления. «Я убежден, — писал Гильберт, — что доказательства существования можно будет провести с помощью некоего основного положения, на которое указывает принцип Дирихле и который, вероятно, приблизит нас к вопросу о том, не допускает ли всякая регулярная вариационная задача решение, если в случае необходимости самому понятию решения придать расширенное толкование» [20].

«Основное положение», о котором говорит Гильберт, по-видимому, **принцип компактности**, согласно которому *полунепрерывная на компакте функция достигает своего минимума*, с которым естественно связать имена Вейерштрасса, Лебега и Бэра.

Концепция Гильберта получила свое воплощение сначала в теории Тонелли для одномерных задач вариационного исчисления, а затем в концепции Соболева С.Л., который начал создавать шкалу пространств специально приспособленных к задачам вариационного исчисления многомерных задач.

Некоторые принципы оптимизационных алгоритмов

Ввиду нехватки места — лишь перечисление некоторых методов численного решения задач на экстремум:

- методы «рационального» спуска решения гладких задач (например, метод градиента)

- прямые методы решения задач вариационного исчисления (например, методы Бубнова и Галеркина)

- методы «отсечения» решения выпуклых задач (например, методы А. Левина, эллипсоидов, Немировского—Нестерова)

Заключение. В докладе я постарался представить то, что отношу к основным идеям и принципам теории экстремума (которые полезно знать всем). Их так немного, что я могу их здесь всех назвать. Это:

- принцип Ферма для задач без ограничений;
- принцип Лагранжа для задач с ограничениями;
- идея Гамильтона о полях экстремалей;
- принцип двойственности в выпуклом анализе;
- принцип компактности в теории существования;
- идеи рационального спуска, отсечений и штрафов в методах оптимизации.

Все это (в принципе) было известно и сорок лет тому назад. О принципах и идеях родившихся в последние годы (о которых полезно знать всем), должны рассказывать другие.

ЛИТЕРАТУРА

1. *Иоффе А.Д., Тихомиров В.М.* Теория экстремальных задач. М.: Наука, 1974.
2. *Ляпунов А.А.* О вполне аддитивных функциях // Изв. АН СССР. Сер. матем. 1940. Т. 4. № 6. С. 465-478.
3. *Fenchel W.* Convex Cones, Sets and Functions. Princeton Univ. Press. 1951.
4. *Frèchet V.* Sur la notion de differentielle // Nouvelle annale de mathematique. 1912. Ser. 4. V. 12. P. 845.
5. *Люстерник Л.А.* Об условных экстремумах функционалов // Матем. сб. 1934. Т.41. №3. С. 390-401.
6. *P. de Fermat* Oeuvres de Fermat. V. 1. P. Gauthier-Villars. 1891. Рус. пер. Ферма П. Метод отыскания наибольших и наименьших значений. Декарт Р. Геометрия. Москва; Ленинград: Техтеорлит, 1938.
7. *Кеплер И.* Новая стереометрия винных бочек. Москва; Ленинград: ГТТИ, 1935. 360 с.
8. *Lagrange J.L.* Théorie des fonctions analytiques. P. 1797.
9. *Дубовицкий А.Я., Милютин А.А.* Задачи на экстремум при наличии ограничений // ЖВМ и МФ. 1965. Т. 5. № 3. С. 395-453
10. *Karush W.E.* Minima of functions of several variables with inequalities as side conditions. Univ. of Chicago Press, 1939.
11. *Kuhn H.W., Tucker A.W.* Nonlinear programming // Berkley. Univ. of California Press. 1951. P. 481-482.

12. *Euler L.* Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes sive solutio problematis isoperimetrici latssimo sencu accepti. Lausanne, 1744. Рус. пер. Эйлер Л. Метод нахождения кривых линий... Москва; Ленинград: Гостехтеориздат, 1934.

13. *Lagrange J.L.* Essai d'une nouvelle méthode pour determiner les maxima et les minima periales // Petropolitanae, 1766. V. 10. P. 51-93.

14. *Kneser A.* Lehrbuch der Variationsrechnung. Berlin, 1900.

15. *Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф.* Математическая теория оптимальных поцессов. М.: Физматлит, 1961.

16. *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. P. 1788. P. 7-37.

17. *Hamilton W.R.* Second Essay on a General Method in Dynamics. Philos. Trans., 1835.

18. *Jacobi C.G. J.* Zur Theorie der Variations-Rechnung und der Differential-Gleichungen // Krelle's Journall. 1837. Bd. 17.

19. *Weierstrass K.* Mathematische Werke. Bd. 7. Vorlesungemn uber Variationsrehtung. Berlin; Leipzig: Akad. Verlag., 1927.

20. *Гильберт Д.* Избранные труды. О вариационном исчислении. 1906. Т. 2. с. 351-370. Анализ. Физика. Проблемы. Personalia. М.: Изд-во «Факториал», 1998.

Поступила в редакцию 21 ноября 2013 г.

Tikhomirfov V.M.

GENERAL PRINCIPLES OF THEORY OF EXTREMUM

The article is devoted to general principles of the theory of extremum: Fermat and Lagrange principes in necessary conditions, Hamiltonian principles in sufficient conditions, duality principles in convexity and some general ideas in the theory of existence and algorithms of optimization.

Key words: smoothness; convexity; Lagrange's principle; convex duality; compactness.

Тихомиров Владимир Михайлович, Московский государственный университет, г. Москва, Российская Федерация, доктор физико-математических наук, профессор, e-mail: tikhomir@mccme.ru

Tikhomirov Vladimir Mikhaylovich, Moscow State University, Moscow, Russian Federation, Doctor of Physics and Mathematics, Professor, e-mail: tikhomir@mccme.ru