Владикавказский математический журнал
Октябрь–декабрь, 2005, Том 7, Выпуск 4

УДК **517.5**

ГЛАДКО-АППРОКСИМАТИВНО-ВЫПУКЛЫЙ ПРИНЦИП И ЕГО ПРИЛОЖЕНИЯ

В. М. Тихомиров

Дорогому Семену Самсоновичу
к его шестидесятилетию

Доказывается принцип Лагранжа для гладкой аппроксимативно-выпуклой задачи, затем этот принцип прилагается к разнообразным классам экстремальных задач. В приложении доказываются те факты анализа (за исключением теоремы отделимости, изложенной во многих учебниках), на которых основывается принцип Лагранжа. Читателю предоставляется возможность освоить результаты важной главы теории экстремума от истоков до наших дней, фактически не прибегая к дополнительным источникам.

Введение

Ключ к решению весьма многих конкретных задач на максимум и минимум дает универсальный прием, называемый нами *принципом Лагранжа*. Согласно этому принципу, если задачу удается формализовать, как задачу с равенствами и двумя типами переменных, по одному из которых отображение является гладким, а по другому имеет выпуклый (или «почти выпуклый» — «аппроксимативно-выпуклый») образ, то для получения необходимых условий экстремума нужно составить функцию Лагранжа и поступать с ней (цитируя Лагранжа) «как если бы переменные были независимы». Иначе говоря, по гладким переменным надо применить теорему Ферма, а по выпуклым — условие минимума. Многие классы экстремальных задач (включая гладкие и выпуклые задачи математического программирования, задачи классического вариационного исчисления и оптимального управления) допускают гладко-выпуклую или гладко-аппроксимативно-выпуклую формализацию, так что необходимые условия экстремума для всех перечисленных классов задач непосредственно следуют из доказываемого ниже принципа Лагранжа для гладкой аппроксимативно-выпуклой задачи. Само доказательство опирается в существенном лишь на два основополагающих факта анализа — обобщенную теорему о неявной функции (см. Приложение) и теорему отделимости выпуклых множеств.

Структура работы такова: сначала доказывается принцип Лагранжа для гладкой аппроксимативно-выпуклой задачи, затем этот принцип прилагается к разнообразным классам экстремальных задач, а в Приложении доказываются те факты анализа (за исключением теоремы отделимости, хорошо изложенной во многих учебниках), на которых основывается принцип Лагранжа. Так что читателю предоставляется возможность освоить результаты важной главы теории экстремума от истоков до наших дней, на небольшом числе страниц, фактически не прибегая к дополнительным источникам.

1. Принцип Лагранжа для гладко-аппроксимативно-выпуклой задачи

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

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

Функцию Лагранжа задачи (P) определим равенством

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

где y^* — элемент пространства Y^*, сопряженного с Y, $\langle y^*, y\rangle$ — действие функционала y^* на элемент $y \in Y$, $\lambda = (\lambda_0, y^*)$ — набор множителей Лагранжа.

Говорят, что функция $f_0 : U \to \mathbb{R}$ дифференцируема в точке $\widehat{x}$ и пишут $f_0 \in D^1(\widehat{x})$, если существует такой элемент x^* сопряженного пространства X^*, что $f_0(\widehat{x} + x) = f_0(\widehat{x}) + \langle x^*, x\rangle + o(x)$ для всех x, удовлетворяющих условию $\widehat{x} + x \in U$. Элемент x^*, определяемый единственным образом, называется *производной* f_0 *в точке* $\widehat{x}$ и обозначается $f_0'(\widehat{x})$. Говорят, что отображение $F : U \to Y$ дифференцируемо в точке $\widehat{x}$ и пишут $F \in D^1(\widehat{x})$, если существует такой линейный непрерывный оператор $\Lambda : X \to Y$, что $F(\widehat{x} + x) = F(\widehat{x}) + \Lambda x + r(x)$ для всех x, удовлетворяющих условию $\widehat{x} + x \in U$, и при этом $\|r(x)\|_Y / \|x\|_X \to 0$ при $x \to 0$. Оператор Λ, определяемый единственным образом, называется *производной* F *в точке* $\widehat{x}$ и обозначается $F'(\widehat{x})$. Говорят, что отображение $F : U \to Y$ строго дифференцируемо в точке $\widehat{x}$ и пишут $F \in SD^1(\widehat{x})$, если существует такой линейный непрерывный оператор $\Lambda : X \to Y$, что для любого $\varepsilon > 0$ найдется число $\delta > 0$, для которого $\|F(x_2) - F(x_1) - \Lambda(x_2 - x_1)\|_Y \leqslant \varepsilon\|x_2 - x_1\|_X$, если $\|x_i - \widehat{x}\|_X < \delta$, $i = 1,2$. Очевидно, что из строгой дифференцируемости F в $\widehat{x}$ следует дифференцируемость F в этой точке и $\Lambda = F'(\widehat{x})$.

Скажем, что отображение $F : U \times \mathcal{U} \to Y$ *строго дифференцируемо по x равномерно по u в точке* $(\widehat{x}, \widehat{u})$, если отображение $x \mapsto F(x, \widehat{u})$ дифференцируемо в точке $\widehat{x}$ и при этом для всякого $\varepsilon > 0$ найдутся такие число $\delta > 0$ и окрестность W точки $\widehat{u}$, что $\|F(x_2,u) - F(x_1,u) - F_x'(\widehat{x},\widehat{u})(x_2 - x_1)\|_Y \leqslant \varepsilon\|x_2 - x_1\|_X$, если только $\|x_i - \widehat{x}\|_X < \delta$, $i = 1,2$ и $u \in W$.

Будем говорить, что в точке (x,u) осуществим (ε,α)-*микс* («*перемешивание*») *управления u с управлением v* (пишем $\mathrm{mix}\,(\alpha,\varepsilon,u,v)$), если найдется такое семейство управлений $\{u_{\varepsilon,\alpha,u,v}\}$, $\varepsilon > 0$, $\alpha \in [0,1]$, что $F(x, u_{\varepsilon,\alpha,u,v})$ стремится к $(1-\alpha)F(x,u) + \alpha F(x,v)$ при $\varepsilon \to 0$, а само $u_{\varepsilon,\alpha,u,v}$ при $\alpha \to 0$ стремится к u (равномерно по ε).

Это определение мотивировано следующим важным примером микса $u_{\varepsilon,\alpha,u,v}(\cdot)$ управлений $u(\cdot)$ и $v(\cdot)$ — кусочно непрерывных или измеримых вектор-функций, определенных на отрезке $[t_0,t_1]$: отрезок делится на одинаковые отрезки длины ε (кроме последнего, длина которого $\leqslant \varepsilon$), а затем каждый отрезок делится на две части, длины которых пропорциональны $1-\alpha$ и α, и на первом отрезке деления $u_{\varepsilon,\alpha,u,v}(\cdot)$ берется равным $u(\cdot)$, на втором — $v(\cdot)$.

Скажем, что отображение F в задаче (P) *гладко-аппроксимативно-выпукло* в точке $(\widehat{x},\widehat{u})$, если отображение F строго дифференцируемо в точке $\widehat{x}$ равномерно по u, и для любых $x \in U$, $u, v \in W$ в точке (x,u) возможно осуществить микс управления u с любым управлением $v \in \mathcal{U}$.

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

Теорема (Принцип Лагранжа для гладко-аппроксимативно-выпуклой задачи). *Пусть X и Y — банаховы пространства, $\mathcal{U}$ — топологическое пространство, $f_0 \in D^1(\widehat{x})$, отображение F гладко-аппроксимативно-выпукло, а $F_x(\widehat{x},\widehat{u})X$ — замкнутое подпространство Y конечной коразмерности. Тогда, если $(\widehat{x},\widehat{u})$ — сильный локальный минимум задачи (P), то в точке $(\widehat{x},\widehat{u})$ верен принцип Лагранжа, т. е. найдется ненулевой набор множителей Лагранжа $\lambda = (\lambda_0, y^*)$ такой, что удовлетворяются условия стационарности:*

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

и условие минимума:

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

Отображение F (в задаче (P)) назовем *гладко-выпуклым* в точке $(\widehat{x},\widehat{u})$, если отображение $x \mapsto F(x,\widehat{u})$ строго дифференцируемо в точке $\widehat{x}$, а подмножества $F(x,\mathcal{U})$ (образ $\mathcal{U}$ при отображении $u \mapsto F_x(x,u)$) выпуклы для всех $x \in U$. Очевидно, что гладко-аппроксимативно-выпуклое отображение гладко-выпукло, и потому верно

Следствие (Принцип Лагранжа для гладко-выпуклой задачи). *Пусть X и Y — банаховы пространства, $\mathcal{U}$ — топологическое пространство, $f_0 \in D^1(\widehat{x})$, а отображение F гладко-выпукло и при этом подпространство $F_x(\widehat{x},\widehat{u})X$ замкнуто и имеет конечную коразмерность. Тогда, если $(\widehat{x},\widehat{u})$ — сильный локальный минимум задачи (P), то в точке $(\widehat{x},\widehat{u})$ верен принцип Лагранжа, т. е. найдется ненулевой набор множителей Лагранжа $\lambda = (\lambda_0, y^*)$ такой, что удовлетворяются соотношения* (1) *и* (2).

Замечание. Из приводимого далее доказательства следствия можно усмотреть, что топология $\mathcal{U}$ не используется, так что можно в формулировке следствия считать, что $\mathcal{U}$ — просто некоторое множество.

$\lhd$ При доказательстве используются обобщенная теорема о накрытии — вариант теоремы об обратном отображении (и ее следствие — теорема Люстерника о касательном пространстве), лемма об аннуляторе ядра регулярного оператора и теоремы отделимости (конечномерная и бесконечномерная). Доказательства этих фактов (кроме теорем отделимости) приведены в конце работы.

Переходим к самому доказательству. Обозначим $F_x(\widehat{x},\widehat{u})$ через Λ и ΛX через Y_1. Сначала докажем теорему в простейшем — регулярном случае, когда $\Lambda X = Y$. Пусть $x \in \operatorname{Ker}\Lambda$. Тогда по теореме Люстерника существует отображение $t \mapsto r(t)$ окрестности нуля из $\mathbb{R}$ в Y такое, что $F(\widehat{x}+tx+r(t),\widehat{u}) = 0$ и $r(t) = o(t)$ при $t \to 0$. Значит, в силу того, что $(\widehat{x},\widehat{u})$ — сильный минимум, при малом t выполняется неравенство $f_0(\widehat{x}+tx+r(t)) \geqslant f_0(\widehat{x})$, откуда следует, что $\langle f_0'(\widehat{x}), x\rangle = 0$, и значит, по теореме об аннуляторе ядра регулярного оператора найдется элемент $y^* \in Y^*$ такой, что

$$f_0'(\widehat{x}) + \Lambda^* y^* = 0. \tag{i}$$

Пусть теперь $v \in \mathcal{U}$. Из-за регулярности отображения $x \to F(x,\widehat{u})$, найдется элемент $x(v)$ такой, что

$$\Lambda x(v) + F(\widehat{x},v) = 0. \tag{ii}$$

Определим отображение $\Phi : X \times \mathbb{R} \to Y$ по формуле:

$$\Phi(x,\alpha) = (1-\alpha)F(\widehat{x}+x,\widehat{u}) + \alpha F(\widehat{x}+x,v).$$

Из условия гладкости в теореме функция Φ строго дифференцируема в точке $(0,0)$ и $\Phi'(0,0)[(x,\alpha)] = \Lambda x + \alpha F(\widehat{x}, v)$. Мы видим, что элемент $(x(v),1) \in \operatorname{Ker}\Phi'(0,0)$ и значит (снова по теореме Люстерника) найдутся такие $\rho(t) = o(t)$ и $\alpha(t) = o(t)$, что

$$\begin{aligned} 0 &= \Phi(\widehat{x} + tx(v) + \rho(t), t + \alpha(t)) \\ &= (1 - t - \alpha(t))F(\widehat{x} + tx(v) + \rho(t), \widehat{u}) + (t + \alpha(t))F(\widehat{x} + tx(v) + \rho(t), v). \end{aligned}$$

Теперь завершим доказательство следствия в регулярном случае, а затем вернемся к общей ситуации. По условию следствия множество $F(\widehat{x} + tx(v) + \rho(t), \mathcal{U})$ выпукло, из чего следует существование такого элемента $u(t)$, что $F(\widehat{x} + tx(v) + \rho(t), u(t)) = 0$. Следовательно, пара $(\widehat{x} + tx(v) + \rho(t), u(t))$ допустима и значит, $f_0(\widehat{x} + tx(v) + \rho(t)) \geqslant f_0(\widehat{x})$ (в силу того, что $(\widehat{x}, \widehat{u})$ — локальный минимум), т. е.

$$\langle f_0'(\widehat{x}), x(v)\rangle \geqslant 0, \tag{iii}$$

и тогда

$$0 \overset{(iii)}{\leqslant} \langle f_0'(\widehat{x}), x(v)\rangle \overset{(i)}{=} -\langle \Lambda^* y^*, x(v)\rangle \overset{\mathrm{Id}}{=} -\langle y^*, \Lambda x(v)\rangle \overset{(ii)}{=} \langle y^*, F(\widehat{x}, v)\rangle. \tag{iv}$$

Это означает, что в регулярном случае следствие доказано (ибо (i) — это условие стационарности, а (iv) — принцип минимума). Вернемся к общему случаю.

Построим в точке $(\widehat{x} + tx(v) + \rho(t), \widehat{u})$ микс $u_{\varepsilon,t+\alpha(t),\widehat{u},v}$ управления $\widehat{u}(\cdot)$ с управлением $v(\cdot)$ и подберем ε и t столь малыми, чтобы $\{u_{\varepsilon,t+\alpha(t),\widehat{u},v}\} \in W$, а $\|F(\widehat{x} + tx(v) + \rho(t), u_{\varepsilon,t+\alpha(t),\widehat{u},v}\|_Y \leqslant \rho(t)$. Применив теперь обобщенную теорему о накрытии, получаем, что

$$F(\widehat{x} + tx(v) + \rho_1(t), u_{\varepsilon,t+\alpha(t),u,v}) = 0,$$

причем

$$\|\rho_1(t)\|_X = \|\rho(t) + \varphi(\widehat{x} + tx(v) + \rho(t), 0, u_{\varepsilon,t+\alpha(t),v})\|_X \leqslant (1 + C)\rho(t) = o(t).$$

Таким образом, $\widehat{x} + tx(v) + \rho_1(t)$ — допустимый элемент, и значит, $f_0(\widehat{x} + tx(v) + \rho_1(t)) \geqslant f_0(\widehat{x})$ и, следовательно, $\langle f_0'(\widehat{x}), x(v)\rangle \geqslant 0$, и далее идет выкладка (iv), завершающая доказательство теоремы в регулярном случае.

Доказательство в общем случае (когда $\Lambda X \neq Y$) есть модификация проведенного рассуждения. Обозначим фактор-пространство Y/Y_1 через Z и пусть $\pi : Y \to Z$ — каноническое отображение. Согласно условию, Z — конечномерное пространство. Отметим, что из гладкой аппроксимативной выпуклости следует, что множество $A = \operatorname{cl} F(\widehat{x}, \mathcal{U})$ (замыкание множества $F(\widehat{x}, \mathcal{U}) = \{y \in Y : y = F(x,u),\ u \in \mathcal{U}\}$) выпукло.

Имеются две возможности: либо внутренность πA не содержит 0_Z (вырожденный случай), либо $\operatorname{int} \pi A$ содержит 0_Z (невырожденный случай). (Отметим, что πA содержит 0_Z, ибо $F(\widehat{x}, \widehat{u}) = 0$.)

Рассмотрим сначала вырожденный случай. Тогда по конечномерной теореме отделимости существует такой нетривиальный элемент $z^* \in Z^*$, что

$$\langle z^*, z\rangle \geqslant 0 \quad \text{для любого} \quad z \in \pi(F(\widehat{x}, \mathcal{U})). \tag{v}$$

Положим $y^* = \pi^* z^*$. Тогда $y^* \neq 0$, так как оператор π сюръективен, а $z^* \neq 0$. Положим $\lambda = (0, y^*)$. Тогда для любого $u \in \mathcal{U}$ будем иметь:

$$\langle y^*, F(\widehat{x}, u)\rangle \overset{\mathrm{Id}}{=} \langle \pi^* z^*, F(\widehat{x}, u)\rangle \overset{\mathrm{Id}}{=} \langle z^*, \pi(F(\widehat{x}, u))\rangle \overset{(v)}{\geqslant} 0, \tag{vi}$$

$$\mathcal{L}_x\big((\widehat{x},\widehat{u}),\lambda\big) = \langle y^*, \Lambda x\rangle \overset{\mathrm{Id}}{=} \langle \pi^* z^*, \Lambda x\rangle \overset{\mathrm{Id}}{=} \langle z^*, \pi(\Lambda x)\rangle \overset{\mathrm{def} z^*}{=} 0. \qquad (vii)$$

Соотношение (vi) — условие минимума, а (vii) — условие стационарности для лагранжиана с $\lambda = (0, y^*)$. Принцип Лагранжа в вырожденном случае доказан.

Осталось рассмотреть невырожденный случай. Выберем конечное множество векторов $\{z_j\}_{j=1}^m$ из $\pi(F(\widehat{x},\mathcal{U}))$, выпуклая коническая оболочка которых равна Z. Пусть $z_j = \pi(F(\widehat{x}, v_j))$ для некоторых $v_j \in \mathcal{U}$, $1 \leqslant j \leqslant m$. Представляя вектор 0_Z как линейную комбинацию векторов z_j, $1 \leqslant j \leqslant m$, с положительными коэффициентами $\sum_{j=1}^m \bar\beta_j z_j = 0$, будем иметь $\pi(\sum_{j=1}^m \bar\beta_j F(\widehat{x}, v_j)) = 0$. Значит, существует элемент $\bar\xi \in X$, для которого $\Lambda\bar\xi + \sum_{j=1}^m \bar\beta_j F(\widehat{x}, v_j) = 0$. Выберем теперь произвольный элемент $v \in Y_1 \cap F(\widehat{x},\mathcal{U})$. Тогда существует $x_v \in X$ такой, что $\Lambda x_v + F(\widehat{x}, v) = 0$. Кроме того, выберем число $\delta > 0$ и определим отображение $\Phi_\delta : X \times \mathbb{R} \times \mathbb{R}^m \to Y$ по формуле

$$\Phi_\delta(x,\alpha,\beta) = \left(1 - \alpha\left(1 + \delta\sum_{j=1}^m \bar\beta_j\right) - \sum_{j=1}^m \beta_j\right) F(\widehat{x}+x, \widehat{u}) \\ + \alpha\left(F(\widehat{x}+x, v) + \delta\sum_{j=1}^m \bar\beta_j F(\widehat{x}+x, v_j)\right) + \sum_{j=1}^m \beta_j F(\widehat{x}+x, v_j).$$

Заметим, что это выражение представляет собой линейную комбинацию векторов $F(\widehat{x}+x, \widehat{u})$, $F(\widehat{x}+x, v)$, $\{F(\widehat{x}+x, v_j)\}_{j=1}^m$ с суммой коэффициентов, равной единице. Более того, эти коэффициенты неотрицательны, если при $\alpha > 0$ числа $\beta_1, \ldots, \beta_m$ достаточно малы. При этом Φ_δ строго дифференцируемо в точке $(0,0,0)$ в силу строгой дифференцируемости отображения F. Производная Φ_δ в $(0,0,0)$, как нетрудно понять, дается формулой:

$$\Phi'_\delta((0,0,0))[(x,\alpha,\beta)] = \Lambda x + \alpha(F(\widehat{x}, v) - \delta\Lambda\bar\xi) + \sum_{j=1}^m \beta_j F(\widehat{x}, v_j).$$

Мы видим, что Φ_δ сюръективное отображение и $(x_v + \delta\bar\xi, 1, 0) \in \operatorname{Ker}\Phi'_\delta(0,0,0)$. Применим теорему Люстерника о касательном пространстве. Из нее следует, что

$$\Phi_\delta(t(x_v + \delta\bar\xi) + r_{v\varepsilon}(t), t + \rho_{v\varepsilon}(t), \beta_{v\varepsilon}(t)) = 0$$

для некоторых $o(t)$-отображений $r_{v\varepsilon}$, $\rho_{v\varepsilon}$, $\beta_{v\varepsilon}$. Теперь, совершив в точке $(\widehat{x}+tx(v)+\rho(t), \widehat{u})$ микс $\operatorname{mix}(\varepsilon, (t + \rho_{v\delta}(t), \beta_{v\delta}(t)), v, v_1, \ldots, v_m) = u_{\varepsilon\delta t}$, выбрав ε, δ и t достаточно малыми и применив обобщенную теорему о накрытии, приходим к тому, что $F(\widehat{x} + t(x_v + \delta\bar\xi) + \overline{r}_{v\delta}(t), u_{\varepsilon\delta t}) = 0$, где $\overline{r}_{v\delta}(t) = o(t)$. Итак, построен допустимый элемент в задаче (P), причем при $t \to 0$ вектор $x(t) = \widehat{x} + tx_v + \delta\bar\xi + r_{v\varepsilon\delta}(t)$ сходится к $\widehat{x}$. Значит, из локальной минимальности $(\widehat{x},\widehat{u})$ получаем, что $\langle f'_0(\widehat{x}), x_v\rangle \geqslant 0$. Доказана следующая импликация:

$$\Lambda x_v + F(\widehat{x}, v) = 0 \ \Rightarrow\ \langle f'_0(\widehat{x}), x_v\rangle \geqslant 0.$$

Возьмем теперь $v = \widehat{u}$. Тогда $f'_0(\widehat{x}) \in (\operatorname{Ker}\Lambda)^\perp$. По теореме о ядре регулярного оператора $f'_0(x) + \Lambda^* y_0^* = 0$ для некоторого $y_0^* \in Y_1^*$. Возьмем произвольный элемент $v \in Y_1 \cap \Lambda X$. Тогда получим

$$\langle y_1^*, F(\widehat{x}, v)\rangle = -\langle y_0^*, \Lambda x_v\rangle = -\langle \Lambda^* y_0^*, x_v\rangle = \langle f'_0(\widehat{x}), x_v\rangle \geqslant 0.$$

Из определения конечной коразмерности Y_1 следует, что существуют линейно независимые элементы $\{y_i^*\}_{i=1}^n$ такие, что подпространство Y_1 определяется равенствами

$\langle y_i^*, y\rangle = 0$, $1 \leqslant i \leqslant n$. Пусть $Y_0 = \{y : \langle y_i^*, y\rangle = 0,\ 0 \leqslant i \leqslant n\}$ и π_0 — каноническое отображение из X в Y_0. Рассмотрим полупространство $\Pi = \{y \in Y_1 : \langle y_0^*, y\rangle < 0\}$. Тогда $\pi_0(\Pi) = \{(\zeta, \dots, 0) : \zeta < 0\}$, $\pi_0(F(\widehat{x}, \mathcal{U}) = \{(\eta_0, \eta_1, \dots, \eta_n) : \eta_0 \geqslant 0\}$. Эти множества отделимы (из-за того, что они не пересекаются), и согласно теореме отделимости, найдется вектор $(\alpha_0, \alpha_1, \dots, \alpha_n)$, $\alpha_0 = 1$, такой, что $\sum_{i=0}^n \alpha_i \eta_i \geqslant 0$ для любого $\eta \in \pi_0(F(\widehat{x}, \mathcal{U})$. Положим $\langle y^*, y\rangle = \sum_{i=0}^n \alpha_i \langle y_i^*, y\rangle$. Так как вектор $(1, y^*)$ удовлетворяет требованиям теоремы, то доказательство закончено. $\rhd$

2. Приложения к конкретным классам задач

2.1. Принцип Лагранжа в математическом программировании. При выводе первых четырех теорем используется лишь следствие основной теоремы — гладко-выпуклый принцип.

Теорема 1. *Пусть X и Y_1 — банаховы пространства, V — окрестность точки $\widehat{x} \in X$, $f_i : V \to \mathbb{R}$, $0 \leqslant i \leqslant m$, $\mathcal{F} : V \to Y_1$ и при этом выполнены требования гладкости: $f_i \in SD^1(\widehat{x})$, $\mathcal{F} \in SD^1(\widehat{x}, Y_1)$ и слабой регулярности: предполагается, что образ X при отображении $\mathcal{F}'(\widehat{x})$ замкнут. Тогда, если $\widehat{x}$ является локальным минимумом задачи математического программирования:*

$$f_0(x) \to \min; \quad f_i(x) \leqslant 0 \quad (1 \leqslant i \leqslant m), \quad \mathcal{F}(x) = 0, \tag{P_1}$$

то для точки $\widehat{x}$ верен принцип Лагранжа, т. е. для функции Лагранжа $\mathcal{L}(x, \lambda) = \sum_{i=0}^m \mu_i f_i(x) + \langle y^*, \mathcal{F}(x)\rangle$ существует не равный нулю набор множителей Лагранжа $\lambda = (\mu_0, \mu_1, \dots, \mu_m, y^*) \in \mathbb{R}^{m+1} \times Y_1^*$, обладающих свойствами неотрицательности ($\mu_i \geqslant 0$, $0 \leqslant i \leqslant m$), дополняющей нежесткости ($\mu_i f_i(\widehat{x}) = 0$, $1 \leqslant i \leqslant m$) и стационарности:

$$\mathcal{L}_x(\widehat{x}, \lambda) = 0 \ \Leftrightarrow\ \sum_{i=0}^n \mu_i f_i'(\widehat{x}) + (\mathcal{F}'(\widehat{x}))^* y^* = 0.$$

$\lhd$ При доказательстве, помимо гладко-выпуклого принципа и леммы о нетривиальности аннулятора, используется еще и лемма о замкнутости образа (см. Приложение).

Переходим к доказательству теоремы. Возможно одно из двух: или $\mathcal{F}'(\widehat{x})X \neq Y_1$ (вырожденный случай) или оператор $\mathcal{F}'(\widehat{x})$ сюръективен. В вырожденном случае из леммы об аннуляторе следует, что существует $y^* \in (\operatorname{Ker} \mathcal{F}'(\widehat{x}))^{\perp}$, $y^* \neq 0$. Положив $\lambda = (0, \dots, 0, y^*)$, приходим к утверждению теоремы. В невырожденном случае следует положить $Y = Y_1 \times \mathbb{R}^m$, $F(x, u) = (f_1(x) + u_1, \dots, f_m(x) + u_m)$, $u = (u_1, \dots, u_m)$, $\mathcal{U} = \mathbb{R}^m_+$ и применить гладко-выпуклый принцип (воспользовавшись леммой о замкнутости). Из условия стационарности в гладко-выпуклом принципе получим условие стационарности в теореме 1, а из принципа минимума в гладко-выпуклом принципе получим условия неотрицательности и дополняющей нежесткости. $\rhd$

2.2. Принцип Лагранжа в выпуклом программировании.

Теорема 2. *Если $\widehat{w}$ — решение (абсолютный минимум) задачи выпуклого программирования*

$$f_0(w) \to \min; \quad f_i(w) \leqslant 0 \quad (i = 1, \dots, m), \quad x \in A, \tag{P_2}$$

где f_i — выпуклые функции на векторном пространстве W и A — выпуклое подмножество W, то найдется ненулевой набор множителей Лагранжа $l = (l_0, l_1, \dots, l_m) \in (\mathbb{R}^{m+1})'$ такой, что выполняются:

a) *условие неотрицательности:* $l_i \geqslant 0,\ 0 \leqslant i \leqslant m$;
b) *условие дополняющей нежесткости:* $l_i f_i(\widehat{x}) = 0,\ i = 1, \dots, m$;
c) *условие минимума (для функции Лагранжа* $\mathcal{L}(w,l) = \sum_{i=0}^m l_i f_i(w)$*)*

$$\min_{w \in A} \mathcal{L}(w,l) = \mathcal{L}(\widehat{w}, l).$$

Если для допустимой точки $\widehat{w}$ *выполнены условия* a)–c) *и* $l_0 \neq 0$, *то* $\widehat{w}$ *— решение задачи* (P_2).

◁ Произведем редукцию задачи (P_2) к задаче (P) для пары (f, F), положив

$$X = \mathbb{R}, \quad \mathcal{U} = A \times \mathbb{R}_+^{m+1}, \quad u = (w, \alpha), \quad w \in A, \quad \alpha = (\alpha_0, \dots, \alpha_m)^T,$$
$$f(x) = x, \quad F(x,u) = (\varphi_0(x,u), \varphi_1(u), \dots, \varphi_m(u)),$$
$$\varphi_0(x,u) = f_0(w) - x + \alpha_0, \quad \varphi_i(u) = f_i(w) + \alpha_i, \quad 1 \leqslant i \leqslant m.$$

Выполнение условий принципа Лагранжа для гладко-выпуклых задач проверяется (в силу условий выпуклости) тривиально. Применение теоремы 2 приводит к требуемому результату. ▷

2.3. Принцип Лагранжа для ляпуновских задач. Пусть Δ — промежуток числовой прямой (конечный или бесконечный), U — топологическое пространство, $f_i : \Delta \times U \to \mathbb{R}$, $0 \leqslant i \leqslant m$, — непрерывные функции, Z — векторное пространство, $g_i : Z \to \mathbb{R}$ — функции, выпуклые для $0 \leqslant i \leqslant m'$ и аффинные для $m'+1 \leqslant i \leqslant m$, A — выпуклое множество, $\mathcal{V}$ — совокупность измеримых отображений из Δ в U. Элемент $\mathcal{V} \times A$ обозначим w и положим $\mathcal{G}_i(w) = \int_\Delta f_i(t, u(t)dt + g_i(z)$. Задачу

$$\mathcal{G}_0(w) \to \min; \quad \mathcal{G}_i(w) \leqslant 0 \quad (1 \leqslant i \leqslant m'), \quad \mathcal{G}_i(w) = 0 \quad (m'+1 \leqslant i \leqslant m), \quad z \in A \quad (P_3)$$

назовем *ляпуновской*.

Теорема 3. *Пусть в задаче* (P_3) *все функции* f_i *непрерывны, а функции* g_i *выпуклы при* $0 \leqslant i \leqslant m'$ *и аффинны при* $m'+1 \leqslant i \leqslant m$. *Тогда, если* $\widehat{w} = (\widehat{u}(\cdot), \widehat{z})$ *доставляет абсолютный минимум в* (P_3), *то в этой точке выполнен принцип Лагранжа, т. е. для функции Лагранжа* $\mathcal{L}(w, \lambda) = \sum_{i=0}^m \mu_i \mathcal{G}_i(w)$ $(\lambda = (\mu_0, \dots, \mu_m))$, *выполнены условия* a) *неотрицательности,* b) *дополняющей нежесткости по неравенствам,* c) *минимума:* $\min_{w \in \mathcal{U} \times A} \mathcal{L}(w, \lambda) = \mathcal{L}(\widehat{w}, \lambda)$.

◁ При доказательстве, помимо гладко-выпуклого принципа, применяется теорема Ляпунова [1, п. 4.3.2]. Произведем редукцию задачи (P_3) к (P), положив $X = \mathbb{R}, \mathcal{U} = \mathcal{V} \times A \times \mathbb{R}_+$, $u = (u(\cdot), z, \alpha)$, $\alpha = (\alpha_0, \dots, \alpha_m)$, $f_0(x) = x$, $F(x,u) = (\varphi_0(x,u), \varphi_1(u), \dots, \varphi_m(u))$, $\varphi_0(x,u) = \mathcal{G}_0(u(\cdot), z) - x + \alpha_0$, $\varphi_i(u) = \mathcal{G}_i(u(\cdot), z) + \alpha_i$. Выполнение условий принципа Лагранжа для гладко-выпуклых задач проверяется (с учетом теоремы Ляпунова) тривиально. Применение теоремы приводит к требуемому результату. ▷

2.4. Принцип Лагранжа для задачи Лагранжа вариационного исчисления. Пусть $\Delta = [t_0, t_1]$ — фиксированный отрезок, $z = (x(\cdot), u(\cdot)) \in C^1(\Delta, \mathbb{R}^n) \times C(\Delta, \mathbb{R}^r)$. Определим функционалы

$$\mathcal{B}_i(z) = \int_\Delta f_i(t, x(t), u(t))\, dt + \psi_i(x(t_0), x(t_1)), \quad i = 0, 1, \dots, m,$$

где $f_i : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$, $f_i = f_i(t,x,u)$, $\psi_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\psi_i = \psi_i(\xi_0, \xi_1)$, $\xi_i \in \mathbb{R}^n$, $i = 0, 1$.

Задачей Лагранжа (с фиксированным временем) называется следующая экстремальная задача:

$$\begin{aligned} &\mathcal{B}_0(z) \to \min; \\ &\mathcal{B}_i(z) \leqslant 0 \quad (i = 1, \dots, m'), \\ &\mathcal{B}_i(z) = 0 \quad (i = m'+1, \dots, m), \\ &\dot{x}(t) - \varphi(t, x(t), u(t)) = 0. \end{aligned} \tag{P_4}$$

Функционалы $\mathcal{B}_i(z)$ называются *функционалами Больца,* а соотношение $\dot{x}(t) - \varphi(t, x(t), u(t)) = 0$ называется *дифференциальной связью.* Эту задачу будем рассматривать в банаховом пространстве $Z = C^1(\Delta, \mathbb{R}^n) \times C(\Delta, \mathbb{R}^r)$; локальный экстремум в этом пространстве называется *слабым.*

Теорема 4. *Пусть в задаче Лагранжа все функции ψ_i непрерывно дифференцируемы в окрестности точки $(\widehat{x}(t_0), \widehat{x}(t_1))$, все функции f_i и отображение φ непрерывны вместе со своими производными f_{ix}, f_{iu}, φ_x, φ_u в окрестности множества $\{(t, x, u) : t \in [t_0, t_1],\ x = \widehat{x}(t),\ u = \widehat{u}(t)\}$ (условия гладкости). Тогда, если допустимая точка $\widehat{z} = (\widehat{x}(\cdot), \widehat{u}(\cdot)) \in Z$ доставляет слабый минимум в задаче (P_4), то в этой точке выполнен принцип Лагранжа, т. е. для функции Лагранжа*

$$\mathcal{L} = \int_\Delta L(t, x(t), \dot{x}(t), u(t))\, dt + l(x(t_0), x(t_1)),$$

где $L = \sum_{i=0}^m \mu_i f_i + p(t)(\dot{x} - \varphi)$, $l = \sum_{i=0}^m \mu_i \psi_i$, выполнены условия:

a) *неотрицательности* $(\mu_i \geqslant 0,\ 0 \leqslant i \leqslant m)$,

b) *дополняющей нежесткости* $(\mu_i \mathcal{B}_i(\widehat{z}) = 0,\ 1 \leqslant i \leqslant m)$,

c) *уравнения Эйлера по x и по u* $(g = \sum_{i=0}^m u_i f_i)$:

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \ \Leftrightarrow\ -\dot{p} = p\widehat{\varphi}_x(t) - \lambda_0 \widehat{g}_x(t) = 0,$$

$$\widehat{L}_u(t) = 0 \ \Leftrightarrow\ p(t)\widehat{\varphi}_u(t) = \lambda_0 \widehat{g}_u(t).$$

$\lhd$ В доказательстве, помимо теоремы 1 и леммы о замкнутости, будет использована теорема (глобальная) существования и единственности решения задачи Коши для линейного уравнения. Полагая $F(z)(\cdot) = \dot{x}(\cdot) - \varphi(\cdot, x(\cdot), u(\cdot))$, $\mathcal{B}(z) = (\mathcal{B}_1(z), \dots, \mathcal{B}_m(z))$, $\mathcal{F} = (F, \mathcal{B}) : \bar{Z} \to C(\Delta, \mathbb{R}^n) \times \mathbb{R}^m$, $f_0 = \mathcal{B}_0$, представим задачу в виде

$$f_0(z) \to \min; \quad \mathcal{F}(z) = 0. \tag{i}$$

Если выполнены условия гладкости теоремы, то все функционалы Больца и отображение F дифференциальной связи непрерывно дифференцируемы в Z в окрестности точки $\widehat{z}$ (а значит, строго дифференцируемы в этой точке). При этом $\mathcal{F}'(\widehat{z})z = (F'(\widehat{z})z, \mathcal{B}'(\widehat{z})z)$ для всех $z = (x(\cdot), u(\cdot)) \in Z$, а $F'(\widehat{z})z = \dot{x}(\cdot) - A(\cdot)x(\cdot) - B(\cdot)u(\cdot)$, где $A(t) = \widehat{\varphi}_x(t) := \varphi_x(t, \widehat{x}(t), \widehat{u}(t))$, $B(t) = \widehat{\varphi}_u(t) = \varphi_u(t, \widehat{x}(t), \widehat{u}(t))$. По теореме существования для линейных систем отображение $F'(\widehat{z})$ сюръективно (даже уравнение $\dot{x}(\cdot) - A(t)x(t) = y(t)$ разрешимо для любой вектор-функции $y(\cdot) \in C(\Delta, \mathbb{R}^n)$). Отображение $\mathcal{B}'(\widehat{z})$ конечномерно, следовательно, по лемме о замкнутости образа выполнен гладко-выпуклый принцип, и значит, в силу этой теоремы принцип Лагранжа для задачи

(i) верен. Функция Лагранжа этой задачи имеет вид:

$$\mathcal{L}(x(\cdot), u(\cdot)), \mu, \nu) = \sum_{i=0}^{m} \mu_i \mathcal{B}_i(z) + \langle \nu, F(z) \rangle$$

$$= \int\limits_{t_0}^{t_1} L(t, x(t), u(t), \mu)\, dt + l(x(t_0), x(t_1), \mu) + \langle \nu, \dot{x}(\cdot) - \varphi(\cdot, x(\cdot), u(\cdot)) \rangle,$$

где $L = \sum_{i=0}^{m} \mu_i f_i$, $l = \sum_{i=0}^{m} \mu_i \psi_i$, $\mu = (\mu_0, \dots, \mu_m) \in (\mathbb{R}^{m+1})'$, а ν — линейный функционал на пространстве $C(\Delta, \mathbb{R}^n)$.

Применим теорему 1 (следствие гладко-выпуклого принципа). Согласно ей существует нетривиальный набор множителей Лагранжа (μ, ν) такой, что выполнены условия неотрицательности, дополняющей нежесткости и стационарности: $\widehat{\mathcal{L}}_{x(\cdot)} = 0$, $\widehat{\mathcal{L}}_{u(\cdot)} = 0$. Для любого $x(\cdot)$ имеем:

$$\langle \widehat{\mathcal{L}}_{x(\cdot)}, x(\cdot) \rangle = \int_{t_0}^{t_1} \widehat{L}_x(t) x(t)\, dt + \widehat{l}_{\xi_0} x(t_0) + \widehat{l}_{\xi_1} x(t_1) + \langle \nu, \dot{x}(\cdot) - A(\cdot) x(\cdot) \rangle = 0. \qquad (ii)$$

Проанализируем это условие. Пусть $x_0 \in \mathbb{R}^n$ и $y(\cdot) \in C(\Delta, \mathbb{R}^n)$ выбраны произвольно, $x(\cdot)$ найдено из условий: $\dot{x}(t) = A(t)x(t) + y(t)$, $x(t_0) = x_0$, а $p(\cdot)$ из условий: $\dot{p}(t) = -A^*(t)p(t) + \widehat{L}_x(t)$, $p(t_1) = -\widehat{l}_{\xi_1}$. Тогда

$$p(t_1)x(t_1) - p(t_0)x(t_0) = \int_{t_0}^{t_1} \frac{d}{dt} p(t)x(t)$$

$$= \int_{t_0}^{t_1} (\dot{p}(t)x(t) + p(t)\dot{x}(t))\, dt = \int_{t_0}^{t_1} (\widehat{L}_x(t)x(t) + p(t)y(t))\, dt,$$

откуда

$$\int_{t_0}^{t_1} \widehat{L}_x(t)x(t)\, dt = -\int_{t_0}^{t_1} p(t)y(t)\, dt + p(t_1)x(t_1) - p(t_0)x_0.$$

Подставляя это выражение в (i), и учитывая уравнение $y(t) = \dot{x}(t) - A(t)x(t)$, а также граничные значения, будем иметь

$$-\int_{t_0}^{t_1} p(t)y(t)\, dt + \widehat{l}_{\xi_0} - p(t_0)x_0 + \langle \nu, y(\cdot) \rangle = 0, \quad (x_0, y(\cdot)) \in \mathbb{R}^n \times C(\Delta, \mathbb{R}^n),$$

откуда $\langle \nu, y(\cdot) \rangle = \int_{t_0}^{t_1} p(t) \cdot y(t)\, dt$, $p(t_0) = \widehat{l}_{\xi_0}$. Наконец, из найденного выражения для ν и из стационарности по $u(\cdot)$ получаем:

$$\langle \widehat{\mathcal{L}}_{u(\cdot)}, u(\cdot) \rangle = \int_{t_0}^{t_1} \widehat{L}_u(t) \cdot u(t)\, dt - \langle \nu, B(t)u(t) \rangle$$

$$= \int_{t_0}^{t_1} (\widehat{L}_u(t) - B^*(t)p(t))u(t)\, dt = 0, \quad u(\cdot) \in C(\Delta, \mathbb{R}^r),$$

откуда следует, что $B^*(t)p(t) = \widehat{L}_u(t)$.

В заключение заметим, что если $\mu = 0$, то $\widehat{L}_x(\cdot) = 0$ и $\widehat{l}_{\xi_0} = 0$, а тогда и $p(\cdot) = 0$. Поэтому нетривиальность набора множителей Лагранжа эквивалентна условию $\mu \neq 0$. $\triangleright$

2.5. Принцип Лагранжа для задач оптимального управления. Ограничимся здесь (чтобы избежать несущественных технических осложнений) частным случаем задачи оптимального управления, а именно, простейшей задачей оптимального управления, близкой к той, что рассматривалась в монографии Понтрягина и его учеников:

$$\begin{aligned} &J(x(\cdot)) = \int_{t_0}^{t_1} f(t, x(t))\,dt \to \min; \\ &\dot{x} = \varphi(t, x, u(t)), \quad x(t_0) = x_0, \quad x(t_1) = x_1, \quad u(t) \in U. \end{aligned} \tag{P_5}$$

На этом примере вполне можно продемонстрировать применимость гладко-аппроксимативно-выпуклого принципа к задачам оптимального управления.

В задаче (P_5) $\varphi : V \times U \to \mathbb{R}^n$ и $f : V \to \mathbb{R}$, где V — открытое множество в $\mathbb{R} \times \mathbb{R}^n$ и U — компакт в $\mathbb{R}^r$. Функцию f и отображение φ будем предполагать непрерывными. Для простоты будем рассматривать кусочно непрерывные управления (переход к измеримым управлениям требует преодоления лишь некоторых технических трудностей). Пара $(x(\cdot), u(\cdot))$ допустима в задаче (P_5), если $x(\cdot)$ — кусочно непрерывно дифференцируемая функция, $u(\cdot)$ — кусочно непрерывна, в точках непрерывности $u(\cdot)$ выполняется равенство $\dot{x}(t) = \varphi(t, x(t), u(t))$ и включение $u(t) \in U$ и, кроме того, $(t, x(t)) \in V$ для всех $t \in [t_0, t_1]$. Допустимый процесс $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ называется *оптимальным*, если существует такое число $\varepsilon > 0$, что для всякого допустимого процесса $(x(\cdot), u(\cdot))$, для которого $\|x(\cdot) - \widehat{x}(\cdot)\|_{C([t_0,t_1])} < \varepsilon$, выполнено неравенство $J(x(\cdot)) \geqslant J(\widehat{x}(\cdot))$.

Теорема 5. *Пусть в задаче (P_5) функция f и отображение φ непрерывны на V вместе со своими частными производными f_x и φ_x (условие гладкости). Тогда, если $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ — оптимальный процесс, то выполнен принцип Лагранжа, т. е. найдутся такое число $\lambda_0 \geqslant 0$ и кусочно непрерывно дифференцируемая функция $p(\cdot)$, не равные нулю одновременно, что для функции Лагранжа $\mathcal{L} = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\,dt$, где $L(t, x, \dot{x}, u) = \lambda_0 f(t, x) + p(\dot{x} - \varphi(t, x, u))$, выполнено уравнение Эйлера по x:*

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \quad \Leftrightarrow \quad -\dot{p} = p\widehat{\varphi}_x(t) - \lambda_0 \widehat{f}_x(t) = 0$$

и условие минимума по u:

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

в точках непрерывности $\widehat{u}(\cdot)$.

$\lhd$ Пусть $X = Y = C([t_0, t_1]) \times \mathbb{R}^r$, $\mathcal{U}$ состоит из таких кусочно непрерывных управлений, что в точках непрерывности $u(t) \in U$. Это множество будем рассматривать в пространстве $L_1([t_0, t_1], \mathbb{R}^r)$. Тогда исходную задачу можно записать в виде

$$f_0(x(\cdot)) \to \min; \quad \mathcal{F}(x(\cdot), u(\cdot)) = 0,$$

где $f_0(x(\cdot)) = J(x(\cdot))$, $\mathcal{F}(x(\cdot), u(\cdot)) = (x(t) - x_0 - \int_{t_0}^{t} \varphi(\tau, x(\tau), u(\tau))\,d\tau, x(t_1) - x_1)$. Покажем, что все условия теоремы о гладко-аппроксимативно-выпуклом принципе удовлетворяются:

a) *пространства X и Y банаховы и топология в* $\mathrm{cl}\,\mathcal{U}$ *была введена*;

b) *дифференцируемость* отображений, более общих чем f_0, доказана в предыдущем пункте, где рассматривалась задача Лагранжа;

c) *непрерывность* отображения $u(\cdot) \to \int_{t_0}^{t} \varphi(\tau, x(\tau), u(\tau))\,d\tau$ в $L_1([t_0, t_1], \mathbb{R}^r)$

и

d) *строгая дифференцируемость* отображения $\mathcal{F}$ равномерная по $u(\cdot)$ легко проверяется, если учесть, что близость кусочно непрерывных функций в L_1 влечет равномерно малую близость этих функций, за исключением множества малой меры;

e) *замкнутость и конечная коразмерность* были доказаны при рассмотрении задачи Лагранжа;

f) *свойство микса*, построенного в начале статьи, проверяется без труда.

Применение принципа Лагранжа проходит по той же схеме, что и для задачи Лагранжа. $\triangleright$

Библиографический комментарий. Автором с середины шестидесятых годов руководила мысль выделить истинную суть необходимых условий экстремума во всей совокупности задач теории экстремума. Эта суть виделась мне в сочетании гладкости и выпуклости. Гладко-выпуклый принцип был впервые опубликован в книге Иоффе и Тихомирова [1] (см. также [2]). Там же был доказан некий вариант гладкого аппроксимативно-выпуклого принципа, приспособленного к задачам оптимального управления, но он оказался слишком громоздок. В той форме, которая излагается здесь, гладкий аппроксимативно-выпуклый принцип публикуется впервые. Идеология принципа Лагранжа разрабатывалась также в [3] и [4]. Прокомментируем следствия из гладко-аппроксимативно-выпуклого принципа.

Теорема 1 обобщает *правило множителей Лагранжа*, когда рассматривается гладкая конечномерная задача с ограничениями типа равенств (сам результат был сформулирован Лагранжем в [5] (1797 г.)), а также конечномерный вариант, с равенствами и неравенствами, доказанный впервые, по-видимому, Джоном в [6] (1948). В приведенной выше форме теорема была впервые доказана Дубовицким и Милютиным в [7].

Теорему 2 называют теоремой Каруша — Куна — Таккера. Каруш опубликовал доказательство этой теоремы в [8] (1939), но эта публикация осталась незамеченной. Теорема была передоказана Куном и Таккером в [9] (1951).

Я затрудняюсь точно сказать, кто впервые доказал теорему 3.

Теорема 4 имеет очень долгую историю. Ее изначальный вариант — уравнение Эйлера для простейшей задачи — был получен Эйлером (это уравнение заведомо содержится в [10]). Затем рассматривались изопериметрические задачи (их впервые рассматривал сам Эйлер), задачи со старшими производными (их также изучал Эйлер, а более детально — Пуассон), задачи с подвижными концами (их подробно исследовал Кнезер) и многие др. Большинство задач, попавших в многочисленные учебники по вариационному исчислению, являются следствием рассмотренной нами задачи Лагранжа. Долгое время правило множителей Лагранжа применялось к таким задачам без обоснования. Первым, за кем было признано обоснование правила множителей для задачи, сходной с задачей Лагранжа, был Майер [11] (1886), затем эту же теорему доказывали Турксма, Гильберт [12] и др. Важную роль сыграл учебник Кнезера (1900). В 1902 году Хан писал: «Метод множителей, носящий имя Лагранжа, долгое время был лишен доказательства. Первым доказательством мы обязаны Майеру. Наконец, в учебнике Кнезера доказательство Майера получило такое изложение, которое способно устоять перед всякой критикой».

Класс задач оптимального управления был определен Понтрягиным в середине пятидесятых годов прошлого столетия. Первое доказательство необходимого условия для достаточно общей задачи оптимального управления, получившее название принципа максимума Понтрягина, было получено Болтянским в 1960 году.

ПРИЛОЖЕНИЕ

Г. Г. Магарил-Ильяев, В. М. Тихомиров

Теорема (отделимости). *Пусть A и B — непустые выпуклые подмножества нормированного пространства X, причем внутренность одного из них, скажем, A непуста и $\operatorname{int} A \cap B = \varnothing$. Тогда найдется ненулевой элемент $x^* \in X^*$ такой,* что $\sup_{x\in A}\langle x^*, x\rangle \leqslant \inf_{x\in B}\langle x^*, x\rangle$.

Лишь это утверждение мы оставляем без доказательства (доказательство см., например, в [14]).

Пусть $X_0 \subset X$. Множество $\{x^* \in X^* : \langle x^*, x\rangle = 0,\ x \in X_0\}$ называют *аннулятором* X_0 и обозначают $X_0^{\perp}$.

Следствие 1 (Лемма о нетривиальности аннулятора). *Пусть X — нормированное пространство и X_0 — замкнутое подпространство X, не совпадающее с X. Тогда аннулятор X_0 содержит ненулевой элемент.*

$\triangleleft$ По условию найдется $\widehat{x} \in X \setminus X_0$. Функционал, отделяющий шар с центром в $\widehat{x}$, не пересекающийся с X_0, будет искомым. $\triangleright$

Далее доказываются несколько утверждений о существовании обратных отображений и неявных функций. Основной (и по сути дела единственный) инструмент доказательства — это метод Ньютона последовательных приближений. Первое из этих утверждений равносильно теоремам Банаха об открытом отображении и обратном операторе.

Ниже будем пользоваться следующими обозначениями. Если X — нормированное пространство, то $U_\delta(x;X)$ и $B_\delta(x;X)$ обозначают соответственно открытый и замкнутый шары в точке x радиуса δ; $\operatorname{cl} A$ обозначает замыкание множества A.

Лемма (о правом обратном). *Пусть X, Y — банаховы пространства и $\Lambda : X \to Y$ — линейный непрерывный сюръективный оператор. Тогда существуют отображение $R : Y \to X$ и константа $\gamma > 0$ такие, что $\Lambda(R(y)) = y$ и $\|R(y)\|_X \leqslant \gamma\|y\|_Y$ для всех $y \in Y$.*

$\triangleleft$ Сначала воспользуемся теоремой Бэра (см. [14]) для построения «почти» обратного к Λ. Так как Λ — сюръективный оператор, то $\bigcup_{n\in\mathbb{N}} \Lambda(U_n(0;X)) = Y$. По теореме Бэра существуют такие $N \in \mathbb{N}$, $y_0 \in Y$ и $r > 0$, что $B_r(y_0;Y) \subset \operatorname{cl}(\Lambda(U_N(0;X)))$. Поскольку множество $\Lambda(U_N(0;X))$ выпукло и центрально симметрично, то нетрудно проверить, что этим свойством обладает и его замыкание. Отсюда легко следует, что $B_r(0;Y) \subset \operatorname{cl}(\Lambda(U_N(0;X)))$.

Пусть $y \in Y$, $y \neq 0$. Тогда $(r/\|y\|_Y)y \in B_r(0;Y)$ и поэтому найдется такой элемент $x(y) \in U_N(0;X)$, что
$$\left\| \frac{r}{\|y\|_Y} y - \Lambda x(y) \right\|_Y \leqslant \frac{r}{2}. \tag{1}$$

Положим $R_1(y) = (\|y\|_Y/r)x(y)$ и $R_1(0) = 0$. Тогда
$$\|R_1(y)\|_X \leqslant \frac{\|y\|_Y}{r} N = \gamma_1 \|y\|_Y \tag{2}$$
и (с учетом (1))
$$\|y - \Lambda R_1(y)\|_Y = \frac{\|y\|_Y}{r} \left\| \frac{r}{\|y\|_Y} y - \Lambda x(y) \right\|_Y \leqslant \frac{\|y\|_Y}{2}. \tag{3}$$

Для построения правого обратного к Λ используем метод Ньютона. Пусть $y \in Y$. Рассмотрим последовательность
$$x_n = x_{n-1} + R_1(y - \Lambda x_{n-1}), \quad n \in \mathbb{N}, \quad x_0 = 0 \tag{4}$$
и покажем, что она фундаментальна. Действительно, используя последовательно (4) (после применения к обеим частям этого равенства оператора Λ) и (3), а затем итерируя процедуру, будем иметь
$$\|y - \Lambda x_n\|_Y = \|y - \Lambda x_{n-1} - \Lambda R_1(y - \Lambda x_{n-1})\|_Y \leqslant \frac{1}{2}\|y - \Lambda x_{n-1}\|_Y \leqslant \ldots \leqslant \frac{1}{2^n}\|y\|_Y. \tag{5}$$

Тогда в силу (4), (2) и (5)
$$\|x_{n+1} - x_n\|_X = \|R_1(y - \Lambda x_n)\|_X \leqslant \gamma_1 \|y - \Lambda x_n\|_Y \leqslant \frac{\gamma_1\|y\|_Y}{2^n} \tag{6}$$
и, следовательно, для любых $n, m \in \mathbb{N}$
$$\begin{aligned} \|x_{n+m} - x_n\|_X &\leqslant \|x_{n+m} - x_{n+m-1}\|_X + \ldots + \|x_{n+1} - x_n\|_X \\ &\leqslant \left(\frac{1}{2^{n+m-1}} + \ldots + \frac{1}{2^n} \right) \gamma_1 \|y\|_Y \leqslant \frac{2\gamma_1\|y\|_Y}{2^n} \to 0 \end{aligned}$$
при $n \to \infty$, т. е. последовательность $\{x_n\}$ фундаментальна.

Положим $R(y) = \lim_{n\to\infty} x_n$. Переходя к пределу в (5), получаем, что $\Lambda(R(y)) = y$. Далее, используя (6), имеем
$$\|x_n\|_X \leqslant \|x_n - x_{n-1}\|_X + \ldots + \|x_1\|_X \leqslant \left(\frac{1}{2^{n-1}} + \ldots + 1 \right) \gamma_1 \|y\|_Y \leqslant 2\gamma_1 \|y\|_Y.$$

Переходя здесь к пределу, получаем, что $\|R(y)\|_X \leqslant \gamma\|y\|_Y$, где $\gamma = 2\gamma_1$. $\triangleright$

Следствие 2 (Теорема Банаха об открытом отображении). *Пусть X, Y — банаховы пространства, Λ — линейный непрерывный сюръективный оператор. Тогда Λ — открытое отображение.*

$\triangleleft$ Пусть G — открытое подмножество X. Покажем, что $\Lambda(G)$ открыто. Действительно, пусть $y_0 \in Y$ и $x_0 \in G$ такое, что $\Lambda x_0 = y_0$. Существует такое $r > 0$, что $U_r(x_0;X) \subset G$. Пусть R — правый обратный к

Λ с соответствующей константой γ. Возьмем $\rho > 0$ столь малым, что $\rho\gamma < r$. Тогда для всех $y \in U_\rho(y_0; Y)$ имеем $\|R(y)\|_X \leqslant \gamma\|y\|_Y < r$, т. е. шар $U_\rho(y_0; Y)$ содержится в множестве $\Lambda(G)$ и значит, оно открыто. $\rhd$

Следствие 3 (Теорема Банаха об обратном операторе). *Пусть X, Y — банаховы пространства, Λ — линейный непрерывный биективный оператор. Тогда обратный оператор к Λ непрерывен.*

$\lhd$ Пусть Λ^{-1} — обратный оператор к Λ. Линейность его доказывается элементарно. Проверим непрерывность. Согласно теореме существует правый обратный R к Λ. Тогда $\Lambda(R(y)) = \Lambda(\Lambda^{-1}(y)) = y$ или $\Lambda(R(y) - \Lambda^{-1}y) = 0$ для любого $y \in Y$. Но Λ — инъективный оператор и поэтому $\Lambda^{-1} = R$. Из оценки для R следует непрерывность Λ^{-1} в нуле и тем самым всюду. $\rhd$

Следствие 4 (Лемма о замкнутости). *Пусть X, Y и Z — банаховы пространства, $A : X \to Y$, $B : X \to Z$ — линейные непрерывные операторы, подпространство $\operatorname{Im} A$ замкнуто в Y, подпространство $B(\operatorname{Ker} A)$ замкнуто в Z и $C : X \to Y \times Z$, $Cx := (Ax, Bx)$. Тогда C — линейный непрерывный оператор и подпространство $\operatorname{Im} C$ замкнуто в $Y \times Z$.*

$\lhd$ То, что оператор C линеен и непрерывен очевидно. Обозначим $\widetilde{Y} = \operatorname{Im} A$. Это — банахово пространство и значит, существует правый обратный $R : \widetilde{Y} \to X$ к $A : X \to \widetilde{Y}$, т. е. $(A \circ R)y = y$ и $\|Ry\|_X \leqslant \gamma\|y\|_Y$ $(y \in \widetilde{Y})$. Пусть пара (y, z) принадлежит замыканию $\operatorname{Im} C$ и $\{x_n\}_{n\in\mathbb{N}}$ — такая последовательность, что $y = \lim Ax_n \in \widetilde{Y}$, $z = \lim Bx_n$. Положим $h_n := R(Ax_n - y)$. Тогда $\|h_n\|_X = \|R(Ax_n - y)\|_X \leqslant \gamma\|Ax_n - y\|_Y \to 0$ и, очевидно, $A(x_n - h_n) = y$. Поэтому $Bh_n \to 0$ и обозначив $z_n = B(x_n - h_n)$, получаем $\lim z_n = \lim B(x_n - h_n) = \lim Bx_n = z$, т. е. z принадлежит замыканию множества $\Xi = \{\xi = Bx : Ax = y\}$. Это множество, как легко понять, является сдвигом подпространства $B(\operatorname{Ker} A)$, следовательно, замкнуто. Это означает, что существует $x \in X$, для которого $Ax = y$, $Bx = z$, т. е. $(y, z) \in \operatorname{Im} C$. $\rhd$

Следствие 5 (о ядре регулярного оператора). *Пусть X и Y — банаховы пространства, $A : X \to Y$ — линейный непрерывный сюръективный оператор. Тогда аннулятор ядра этого оператора совпадает с образом сопряженного оператора:* $(\operatorname{Ker} A)^\perp = \operatorname{Im} A^*$.

$\lhd$ Включение $\operatorname{Im} A^* \subset \operatorname{Ker} A)^\perp$ проверяется без труда. Пусть $x^* \in (\operatorname{Ker} A)^\perp$. Образ оператора $M : X \to Y \times \mathbb{R}$, $Mx = (Ax, \langle x^*, x\rangle)$, замкнут по лемме о замкнутости и не совпадает с $Y \times \mathbb{R}$ (ибо $(0, 1) \neq \operatorname{Im} M$). Значит, по лемме о нетривиальности аннулятора существуют $\beta \neq 0$ и $y^* \in Y^*$ такие, что $\langle y^*, Ax\rangle + \beta\langle x^*, x\rangle = 0$ для любого x и значит, $x^* \in \operatorname{Im} A^*$. $\rhd$

Теорема (Обобщенная теорема о неявной функции). *Пусть X — топологическое пространство, Y и Z — банаховы пространства, W — окрестность точки $(\widehat{x}, \widehat{y}) \in X \times Y$ и $F : W \to Z$. Если*

1) $F(\widehat{x}, \widehat{y}) = 0$;

2) *отображение $x \to F(x, \widehat{y})$ непрерывно в точке $\widehat{x}$;*

3) *отображение $y \to F(\widehat{x}, y)$ дифференцируемо в точке $\widehat{y}$ и при этом для каждого $\varepsilon > 0$ найдутся такие окрестности $U(\varepsilon)$ и $V(\varepsilon)$ точек $\widehat{x}$ и $\widehat{y}$ соответственно, что $U(\varepsilon) \times V(\varepsilon) \subset W$ и для всех $x \in U(\varepsilon)$ и $y_1, y_2 \in V(\varepsilon)$ справедливо соотношение*

$$\|F(x, y_1) - F(x, y_2) - F_y(\widehat{x}, \widehat{y})(y_1 - y_2)\|_Z \leqslant \varepsilon\|y_1 - y_2\|_Y; \qquad (*)$$

4) $\mathrm{Im\,F_y(\widehat{x}, \widehat{y}) = Z}$,

то существуют окрестности U и V точек $\widehat{x}$ и $\widehat{y}$ соответственно, отображение $\varphi : U \to V$ и константа $K > 0$ такие, что $F(x, \varphi(x)) = 0$ и $\|\varphi(x) - \widehat{y}\|_Y \leqslant K\|F(x, \widehat{y})\|_Z$ для всех $x \in U$.

$\lhd$ Обозначим $\Lambda = F_y(\widehat{x}, \widehat{y})$. Этот оператор удовлетворяет условиям теоремы 1. Пусть R — правый обратный к Λ и $\gamma > 0$ такое, что $\|R(z)\|_Y \leqslant \gamma\|z\|_Z$ для всех $z \in Z$.

Пусть $\varepsilon = 1/2\gamma$ и соответствующие ему (согласно условию 3) теоремы) окрестности $U(\gamma)$ и $V(\gamma)$ точек $\widehat{x}$ и $\widehat{y}$. Пусть, далее, $\delta > 0$ такое, что $U_\delta(\widehat{y}; Y) \subset V(\gamma)$. В силу условия 2) можно считать, что $\|F(x, \widehat{y})\|_Z < \delta/8\gamma$ для всех $x \in U(\gamma)$.

Положим $U = U(\gamma)$, $V = U_\delta(\widehat{y}; Y)$ и пусть $x \in U$. Рассмотрим последовательность

$$y_n = y_{n-1} - R(F(x, y_{n-1})), \quad n \in \mathbb{N}, \quad y_0 = \widehat{y}. \qquad (7)$$

Покажем, что $y_n \in U_{\delta/4}(\widehat{y}; Y)$. Доказываем по индукции. Ясно, что $y_0 \in U_{\delta/4}(\widehat{y}; Y)$. Пусть $y_n \in U_{\delta/4}(\widehat{y}; Y)$, $n = 0, 1, \ldots, N$. Покажем, что и y_{N+1} из этой окрестности. Из (7) следует, что

$$\Lambda(y_N - y_{N-1}) = -F(x, y_{N-1}). \qquad (8)$$

Используя последовательно (8) (с оценкой для правого обратного), (8), $(*)$ и затем итерируя процедуру, будем иметь

$$\begin{aligned} \|y_{N+1} - y_N\|_Y \leqslant \gamma\|F(x, y_N)\|_Z &= \gamma\|F(x, y_N) - F(x, y_{N-1}) - \Lambda(y_N - y_{N-1})\|_Z \\ &\leqslant \frac{1}{2}\|y_N - y_{N-1}\|_Y \leqslant \ldots \leqslant \frac{1}{2^N}\|y_1 - \widehat{y}\|_Y. \end{aligned} \qquad (9)$$

Теперь по неравенству треугольника, (9) и (7) (снова с оценкой для правого обратного) получаем, что

$$\|y_{N+1}-\widehat{y}\|_Y \leqslant \|y_{N+1}-y_N\|_Y+\ldots+\|y_1-\widehat{y}\|_Y \leqslant \left(\frac{1}{2^N}+\ldots+1\right)\|y_1-\widehat{y}\|_Y \leqslant 2\gamma\|F(x,\widehat{y})\|_Z < 2\gamma\frac{\delta}{8\gamma}=\frac{\delta}{4},$$

т. е. $y_{N+1}\in U_{\delta/4}(\widehat{y};Y)$ и значит, $y_n\in U_{\delta/4}(\widehat{y};Y)$ для всех $n=0,1,\ldots$

Последовательность $\{y_n\}$ фундаментальна. Действительно, по неравенству треугольника и (9)

$$\|y_{n+m}-y_n\|_Y \leqslant \|y_{n+m}-y_{n+m-1}\|_Y+\ldots+\|y_{n+1}-y_n\|_Y$$
$$\leqslant \left(\frac{1}{2^{n+m-1}}+\ldots+\frac{1}{2^n}\right)\|y_1-\widehat{y}\|_Y \leqslant \frac{2}{2^n}\|y_1-\widehat{y}\|_Y \to 0$$

при $n\to\infty$.

Положим $\varphi(x)=\lim_{n\to\infty}y_n$. Ясно, что $\varphi(x)\in U_{\delta/2}(\widehat{y};Y)$. Покажем, что функция $y\to F(x,y)$ непрерывна на $U_{\delta/2}(\widehat{y};Y)$. Пусть $\overline{y}\in U_{\delta/2}(\widehat{y};Y)$ и $\varepsilon>0$. Тогда, если $y\in U_{\delta_0}(\overline{y};Y)$, где $\delta_0=\min\{\delta/2,\varepsilon\gamma,\varepsilon/2\|\Lambda\|\}$, то легко видеть, что $y\in U_\delta(\widehat{y};Y)$ и мы имеем, используя $(*)$:

$$\|F(x,y)-F(x,\overline{y})\|_Z=\|F(x,y)-F(x,\overline{y})-\Lambda(y-\overline{y})+\Lambda(y-\overline{y})\|_Z \leqslant (2\gamma)^{-1}\|y-\overline{y}\|_Y+\|\Lambda\|\|y-\overline{y}\|_Y<\varepsilon/2+\varepsilon/2=\varepsilon.$$

В силу доказанного и (8)

$$\|F(x,\varphi(x))\|_Z=\lim_{n\to\infty}\|F(x,y_n)\|_Z=\lim_{n\to\infty}\|\Lambda(y_{n+1}-y_n)\|_Z=0,$$

т. е. $F(x,\varphi(x))=0$.

Далее по неравенству треугольника

$$\|y_n-\widehat{y}\|_Y \leqslant \|y_n-y_{n-1}\|_Y+\ldots+\|y_1-\widehat{y}\|_Y \leqslant 2\gamma\|F(x,\widehat{y}\|_Z.$$

Переходя здесь к пределу, получаем, что $\|\varphi(x)-\widehat{y}\|_Y \leqslant K\|F(x,\widehat{y})\|_Z$, где $K=2\gamma$. $\rhd$

Следствие 6 (Теорема о поправке). *Пусть X,Y — банаховы пространства, U — окрестность точки $\widehat{x}\in X$, отображение $F:U\to Y$ строго дифференцируемо в $\widehat{x}$ и $\mathrm{Im}F'(\widehat{x})=Y$. Тогда найдутся окрестности U_0 и V_0 точек $\widehat{x}$ и нуля в X соответственно, $U_0+V_0\subset U$, отображение $\varphi:U_0\to V_0$ и константа $K>0$ такие, что $F(x+\varphi(x))=F(\widehat{x})$ и $\|\varphi(x)\|_X\leqslant K\|F(x)-F(\widehat{x})\|_Y$ для всех $x\in U_0$.*

$\lhd$ Пусть U_1 и V_1 такие окрестности точки $\widehat{x}$ и нуля в X, что $U_1+V_1\subset U$. Рассмотрим отображение $\Phi:U_1\times V_1\to Y$, $\Phi(x,y)=F(x+y)-F(\widehat{x})$. Проверим выполнение условий теоремы. Ясно, что $\Phi(\widehat{x},0)=0$ и отображение $x\to\Phi(x,0)=F(x)-F(\widehat{x})$ непрерывно в $\widehat{x}$. Так как $\Phi_y(\widehat{x},0)=F'(\widehat{x})$, то выполняется условие 4) теоремы о неявной функции. Проверим условие 3). В силу строгой дифференцируемости F в $\widehat{x}$ для любого $\varepsilon>0$ найдутся окрестности $U_2\subset U_1$ и $V_2\subset V_1$ точки $\widehat{x}$ и нуля в X такие, что для любых $x\in U_2$ и $y_1,y_2\in V_2$ справедливо неравенство

$$\|\Phi(x,y_1)-\Phi(x,y_2)-\Phi_y(\widehat{x},0)(y_1-y_2)\|_Y\leqslant\varepsilon\|y_1-y_2\|_X,$$

т. е. выполнено условие 3) теоремы и теперь утверждение следствия сразу следует из утверждения теоремы о неявной функции. $\rhd$

Пусть M — подмножество нормированного пространства X и $\widehat{x}\in M$. Элемент $h\in X$ называется *касательным вектором к M в точке $\widehat{x}$*, если существуют $\varepsilon>0$ и отображение $r:(-\varepsilon,\varepsilon)\to X$ такие, что $\widehat{x}+th+r(t)\in M$ для всех $t\in(-\varepsilon,\varepsilon)$ и $r(t)=o(t)$ при $t\to0$.

Множество всех касательных векторов к M в точке $\widehat{x}$ обозначим $T_{\widehat{x}}M$.

Следствие 7 (Теорема Люстерника). *Пусть X,Y — банаховы пространства, U — окрестность точки $\widehat{x}\in X$, отображение $F:U\to Y$ строго дифференцируемо в $\widehat{x}$, $\mathrm{Im}\,F'(\widehat{x})=Y$ и $M=\{x\in X: F(x)=F(\widehat{x})\}$. Тогда $T_{\widehat{x}}M=\mathrm{Ker}\,F'(\widehat{x})$.*

$\lhd$ Пусть $h\in T_{\widehat{x}}M$ и $r(\cdot)$ из определения касательного вектора. В силу дифференцируемости F в $\widehat{x}$ имеем для всех достаточно малых t:

$$F(\widehat{x}+th+r(t))=F(\widehat{x})+tF'(\widehat{x})h+o(t)=F(\widehat{x}). \tag{10}$$

Отсюда легко следует, что $F'(\widehat{x})h=0$, т. е. $T_{\widehat{x}}M\subset\mathrm{Ker}\,F'(\widehat{x})$.

Докажем противоположное включение. Ясно, что выполнены условия теоремы о поправке и пусть окрестности U_0, V_0, отображение φ и константа K из этой теоремы. Пусть далее $h\in\mathrm{Ker}\,F'(\widehat{x})$ и $\varepsilon>0$ таково, что если $t\in(-\varepsilon,\varepsilon)$, то $\widehat{x}+th\in U_0$. Положим $r(t)=\varphi(\widehat{x}+th)$. Согласно теореме о поправке (и учитывая (10)) $\|r(t)\|_X\leqslant K\|F(\widehat{x}+th)-F(\widehat{x})\|_Y=K\|o(t)\|_Y=o(t)$ и $F(\widehat{x}+th+r(t)=F(\widehat{x})$, т. е. $h\in T_{\widehat{x}}M$. $\rhd$

Литература

1. *Иоффе А. Д., Тихомиров В. М.* Теория экстремальных задач.—М.: Наука, 1974.—479 с.
2. *Магарил-Ильяев Г. Г, Тихомиров В. М.* Выпуклый анализ и его приложения.—М.: УРСС, 2003.—175 с.
3. *Алексеев В. М., Тихомиров В. М., Фомин С. В.* Оптимальное управление.—М.: Наука, 1979.—429 с.
4. *Алексеев В. М., Галеев Э. М., Тихомиров В. М.* Сборник задач по оптимизации.—М.: Наука, 1984.
5. *Lagrange.* Théorie des fonctions analytique.—Paris, 1797.
6. *John F.* Extremum problems with inequalities as subsistary conditions. (In Studies and Essais. Courant Anniversary Volume) N.Y., Interscience, 1948.—P. 187–204.
7. *Дубовицкий А. Я., Милютин А. А.* Задачи на экстремум при наличии ограничений // Журн. вычисл. мат. и матем. физики.—1965.—Т. 5, № 3.—С. 395–453.
8. *Karush W. E.* Minima of functions of several variables with inequalities as side conditions. Chicago: Univ. of Chicago, 1939.
9. *Kuhn H. W., Tucker A. W.* Nonlinear programming.—Berkeley: Univ. of California Press, 1951.—P. 481–492.
10. *Euler L.* Methodus inveniendi lineas curvas proprietate maimi minmive gaudentes.—Geneva: Lausanne, 1744.
11. *Kneser A.* Lehrbuch der Variationsreichnung.—Draunschveig, 1900.
12. *Гильберт Д.* О вариационном исчислении. Избр. труды.—М.: Факториал, 1998.
13. *Понтрягин Л. С. и др.* Математическая теория оптимальных процессов.—М.: Физматгиз, 1961.
14. *Колмогоров А. Н., Фомин С. В.* Элементы теории функций и функционального анализа.—М.: Наука, 2004.—542 с.

Статья поступила 20 сентября 2005 г.

Тихомиров Владимир Михайлович, д. ф.-м. н.
г. Москва, Московский государственный
университет им. В. М. Ломоносова