МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего
профессионального образования
**“МАТИ — Российский государственный
технологический университет
имени К. Э. Циолковского”**

Кафедра “Высшая математика”

К. Ю. Осипенко

ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ И ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

Москва 2014

Содержание

1. Введение

Многие причины побуждают ставить и решать экстремальные задачи, т. е. задачи на максимум и минимум. Интерес к ним проявился уже на заре развития математики и основными стимулами были любознательность и стремление к совершенству.

Среди наиболее ранних, точно решенных задач — так называемая *изопериметрическая задача* — задача о форме кривой заданной длины, охватывающей наибольшую площадь (ответ в ней приводил в своих сочинениях еще Аристотель — IV в. до н. э.) и задача о форме поверхности заданной площади, охватывающей наибольший объем. Ответы на эти задачи для мыслителей Древней Греции были символами совершенства человеческого разума. Крупнейшие их представители: Евклид, Архимед и Аполлоний ставили и решали различные геометрические задачи на экстремум. Задача *о параллелограмме наибольшей площади, который можно вписать в треугольник* приводится в “Началах” Евклида (III в. до н. э.); задача *о шаровом сегменте максимального объема при заданной площади шаровой части поверхности этого сегмента* содержится в сочинениях Архимеда (тоже III в. до н. э.); задача *о минимальном расстоянии от точки плоскости до эллипса и о нормалях к эллипсу из произвольной точки плоскости* была поставлена и решена Аполлонием (III–II в. до н. э.) в его знаменитых “Кониках”.

Долгое время каждая задача решалась индивидуально, посвоему. Первый шаг к исследованию экстремальных задач был сделан П. Ферма в 1638 году, который доказал (в современных терминах), что производная функции в точке ее локального экстремума равна нулю (хотя понимание этого явления можно извлечь и из более ранних высказываний И. Кеплера). Данное событие обычно считают началом становления теории экстремума.

Затем от рассмотрения задач на максимум и минимум для функций одного переменного перешли к рассмотрению экстремальных задач, где переменные — сами функции, т. е. элементы бесконечномерных пространств. Эти задачи породили новое направление в математике, получившее название *вариационного исчисления.* Рождение вариационного исчисления часто связывают с задачей о брахистохроне, поставленной И. Бернулли в 1696 году. Это задача о форме кривой наискорейшего ската, т. е. о форме кривой, соединяющей две точки в вертикальной плоскости, вдоль которой тело под действием силы тяжести без трения проходит путь от одной точки до другой за кратчайшее время (постановка, по-видимому, была навеяна более ранними размышлениями Галилея на эту тему).

Основным мотивом для развития вариационного исчисления явилось то, что многие законы природы, как выяснилось, имеют экстремальный характер, т. е. они неким загадочным образом являются решениями задач на максимум и минимум. Л. Эйлер по этому поводу высказался так: “В мире не происходит ничего, в чем не был бы виден смысл какого-нибудь максимума или минимума”.

2. Экстремальные задачи, их формализация

Задачи на экстремум изначально ставятся, как правило, на языке той области знаний, из которой они происходят. Для того, чтобы эти задачи исследовать математическими средствами, необходимо перевести их на математический язык, т. е. *формализовать*. Этот процесс заключается в описании минимизируемого или максимизируемого функционала f вместе со своей областью определения X и множеством ограничений $C \subset X$. Формализованная экстремальная задача записывается так

$$f(x) \to \min(\max), \quad x \in C, \tag{1}$$

и заключается в нахождении таких точек $x \in C$, в которых функционал f достигает своего минимума (максимума) на C. Такие точки называются *глобальными* или *абсолютными минимумами (максимумами)* в задаче (1) или ее *решениями*. Если нас интересуют и точки минимума, и точки максимума, то вместо min(max) пишем extr и говорим о задаче на экстремум функционала f.

Отметим еще, что если $\widehat{x}$ — решение задачи (1) на минимум (максимум), то ясно, что $\widehat{x}$ — решение аналогичной задачи на максимум (минимум) с функционалом $-f$ вместо f.

Точки из множества ограничений C называются *допустимыми* в задаче (1). Если $C = X$, то задача (1) называется задачей *без ограничений*.

При решении многих конкретных задач нашей целью будет нахождение глобальных экстремумов, но для этого предварительно приходится исследовать задачу на наличие локальных экстремумов (т. е. локальных минимумов и максимумов). Если в X определено понятие “окрестности точки”, то точка $\widehat{x} \in C$ называется *локальным минимумом* (*максимумом*) в задаче (1), если существует такая ее окрестность U, что $f(x) \geq f(\widehat{x})$ $(f(x) \leq f(\widehat{x}))$ для всех допустимых $x \in U$ (т. е. для всех $x \in C \cap U$).

3. Производная отображения. Теорема о среднем

Как уже отмечалось первым шагом к исследованию экстремальных задач был результат П. Ферма о равенстве нулю производной в точке локального экстремума функции. Для того чтобы сформулировать этот результат в достаточно общем виде, дадим определение производной отображения.

Пусть X и Y — линейные нормированные пространства. Совокупность всех линейных непрерывных операторов $\Lambda\colon X \to Y$ обозначим через $\mathcal{L}(X,Y)$. Это нормированное пространство с нормой

$$\|\Lambda\| = \sup_{\|x\|_X \le 1} \|\Lambda x\|_Y.$$

В случае, когда $Y = \mathbb{R}$, $\mathcal{L}(X,\mathbb{R})$ совпадает с множеством всех линейных непрерывных функционалов на X и называется *сопряженным пространством* к X. Сопряженное пространство к X обозначается X^*.

Одним из основных примеров рассматриваемых в дальнейшем нормированных пространств будет являться пространство $\mathbb{R}^d$, под которым мы будем понимать совокупность всех упорядоченных наборов $x = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}$ из d действительных чисел (если $d = 1$, то это просто множество действительных чисел, и мы пишем $\mathbb{R}$ вместо $\mathbb{R}^1$). Элементы $x \in \mathbb{R}^d$ называются *векторами* или *вектор-столбцами*, а числа x_j, $j = 1, \ldots, d$, — *координатами вектора* x. Для экономии места, элементы $\mathbb{R}^d$ будем записывать так $x = (x_1, \ldots, x_d)^T$, где символ ^{T} обозначает транспонирование строки в столбец (в общем случае — транспонирование матрицы). В $\mathbb{R}^d$ естественным образом вводится операция (покоординатного) сложения векторов и операция (покоординатного) умножения вектора на число, превращающие это множество в вещественное линейное пространство.

Пусть $x = (x_1, \ldots, x_d)^T \in \mathbb{R}^d$. Величина

$$|x| = \sqrt{x_1^2 + \ldots + x_d^2}$$

называется *длиной* или *модулем* вектора x. Положив для $x \in \mathbb{R}^d$

$$\|x\|_{\mathbb{R}^d} = |x|$$

(такая норма называется *евклидовой нормой*), получим линейное нормированное пространство $\mathbb{R}^d$.

Пусть $a = (a_1, \ldots, a_d)$ — вектор-строка из d действительных чисел. Для каждого $x = (x_1, \ldots, x_d)^T \in \mathbb{R}^d$ положим

$$a \cdot x = \sum_{j=1}^{d} a_j x_j.$$

Это матричное произведение вектор-строки a на вектор-столбец x, которое иногда называют *внутренним произведением.* Ясно, что отображение $x \mapsto a \cdot x$ есть линейный функционал на $\mathbb{R}^d$. Легко понять, что и любой линейный функционал l на $\mathbb{R}^d$ задается подобным образом с $a = (l(e_1), \ldots, l(e_d))$, где

$$e_1 = (1, 0, \ldots, 0)^T, \ldots, e_d = (0, \ldots, 0, 1)^T$$

— *стандартный базис* в $\mathbb{R}^d$. Таким образом, сопряженное пространство $(\mathbb{R}^d)^*$ можно отождествить с множеством, элементами которого являются наборы из d действительных чисел, но расположенные в строку (с аналогичными операциями сложения и умножения на числа).

Пусть X и Y — нормированные пространства, а U — открытое подмножество X. Отображение $F\colon U \to Y$ называется дифференцируемым в точке $\widehat{x} \in U$, если найдется такой оператор $\Lambda \in \mathcal{L}(X, Y)$, что для всех $h \in X$, для которых $\widehat{x} + h \in U$ справедливо представление

$$F(\widehat{x} + h) = F(\widehat{x}) + \Lambda h + r(h), \tag{2}$$

где $r(h) = o(\|h\|_X)$ ($\|r(h)\|_Y / \|h\|_X \to 0$ при $h \to 0$). Линейный оператор Λ называется производной отображения F в точке $\widehat{x}$ и обозначается $F'(\widehat{x})$. Нетрудно показать единственность оператора Λ, удовлетворяющего равенству (2).

Если отображение F дифференцируемо в каждой точке U, то определено отображение $F'\colon U \to \mathcal{L}(X, Y)$, сопоставляющее $x \in U$ производную $F'(x)$. Если это отображение непрерывно в $\widehat{x} \in U$ (на U), то говорят, что отображение F *непрерывно дифференцируемо* в $\widehat{x}$ (на U).

Если X, Y — нормированные пространства, то норму в пространстве $X \times Y$ можно, например, определить следующим образом:

$$\|(x, y)\|_{X \times Y} = \max\{\|x\|_X, \|y\|_Y\}.$$

Пусть W — окрестность точки $(\widehat{x}, \widehat{y}) \in X \times Y$ и $F\colon W \to Z$, где Z — нормированное пространство. Если отображение $x \to F(x, \widehat{y})$ (определенное на проекции W на X) дифференцируемо в точке $\widehat{x}$, то соответствующую производную называют *частной производной отображения F по x в точке* $(\widehat{x}, \widehat{y})$ и обозначают $F_x(\widehat{x}, \widehat{y})$. Аналогично, частную производную F по y в точке $(\widehat{x}, \widehat{y})$ обозначают $F_y(\widehat{x}, \widehat{y})$.

Рассмотрим некоторые примеры.

Пример 3.1. Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^d$. Тогда функция $f\colon U \to \mathbb{R}$ дифференцируема в точке $\widehat{x}$, если существует линейный функционал на $\mathbb{R}^d$, т. е. вектор $a = (a_1, \dots, a_d) \in (\mathbb{R}^d)^*$ такой, что для всех $h \in \mathbb{R}^d$, для которых $\widehat{x} + h \in U$ справедливо представление

$$f(\widehat{x} + h) = f(\widehat{x}) + a \cdot h + r(h),$$

где $|r(h)|/|h| \to 0$ при $h \to 0$, т. е. $r(h) = o(|h|)$.

Отсюда легко следует (беря в качестве h векторы $(h_1, 0, \dots, 0)^T$, …, $(0, \dots, 0, h_d)^T$), что a_j есть частная производная функции f по переменной x_j в точке $\widehat{x}$, $j = 1, \dots d$. Таким образом,

$$f'(\widehat{x}) = \left(\frac{\partial f}{\partial x_1}(\widehat{x}), \dots, \frac{\partial f}{\partial x_d}(\widehat{x})\right).$$

В классическом анализе обозначают $h = (dx_1, \dots, dx_d)^T$ и тогда

$$f'(\widehat{x}) \cdot h = \frac{\partial f}{\partial x_1}(\widehat{x})dx_1 + \ldots + \frac{\partial f}{\partial x_d}(\widehat{x})dx_d.$$

Это выражение называется дифференциалом f и обозначается $df(\widehat{x})$.

Пример 3.2. Пусть теперь $U \subset \mathbb{R}^{d_1}$ и задано отображение $F\colon U \to \mathbb{R}^{d_2}$. Линейный оператор $\Lambda \in \mathcal{L}(\mathbb{R}^{d_1}, \mathbb{R}^{d_2})$ будем отождествлять с его матрицей в стандартных базисах

$$e_1 = (1, 0, \dots, 0)^T, \dots, e_{d_1} = (0, \dots, 0, 1)^T$$

и

$$e'_1 = (1, 0, \dots, 0)^T, \dots, e'_{d_2} = (0, \dots, 0, 1)^T$$

в $\mathbb{R}^{d_1}$ и $\mathbb{R}^{d_2}$ соответственно, т. е. если $\Lambda e_j = \sum_{i=1}^{d_2} a_{ij}e'_i$, $j = 1, \dots, d_1$, то матрицей оператора Λ называется матрица (мы ее обозначаем той же буквой) $\Lambda = (a_{ij})_{1 \le i \le d_1, 1 \le j \le d_2}$ размера $d_2 \times d_1$. В этом случае Λx — произведение матрицы Λ на вектор x. Из определения дифференцируемости вытекает, что отображение F дифференцируемо в точке $\widehat{x} \in U$, если существует линейный оператор $\Lambda \in \mathcal{L}(\mathbb{R}^{d_1}, \mathbb{R}^{d_2})$, т. е. матрица Λ размера $d_2 \times d_1$ такая, что для всех $h \in \mathbb{R}^{d_1}$, для которых $\widehat{x} + h \in U$, справедливо представление

$$F(\widehat{x} + h) = F(\widehat{x}) + \Lambda h + r(h),$$

где $|r(h)|/|h| \to 0$ при $h \to 0$, т. е. $r(h) = o(|h|)$. Матрица Λ называется в этом случае производной отображения F в точке $\widehat{x}$ и обозначается $F'(\widehat{x})$.

Отображение $F\colon U \to \mathbb{R}^{d_2}$ можно записать в виде $F(x) = (f_1(x), \dots, f_{d_1}(x))^T$, где $f_j\colon U \to \mathbb{R}$, $j = 1, \dots, d_1$ ($f_j(x)$ — это j-ая координата вектора $F(x)$ в стандартном базисе в $\mathbb{R}^{d_2}$). Легко проверить, что F дифференцируемо в точке $\widehat{x}$ тогда и только тогда, когда функции f_j, $j = 1, \dots, d_1$, дифференцируемы в $\widehat{x}$. При этом строки матрицы $F'(\widehat{x})$ являются векторами $f'_1(\widehat{x}), \dots, f'_{d_2}(\widehat{x})$. Производную $F'(\widehat{x})$ называют *матрицей Якоби* отображения F в точке $\widehat{x}$. Тем самым

$$F'(\widehat{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\widehat{x}) & \frac{\partial f_1}{\partial x_2}(\widehat{x}) & \dots & \frac{\partial f_1}{\partial x_{d_1}}(\widehat{x}) \\ \frac{\partial f_2}{\partial x_1}(\widehat{x}) & \frac{\partial f_2}{\partial x_2}(\widehat{x}) & \dots & \frac{\partial f_2}{\partial x_{d_1}}(\widehat{x}) \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_{d_2}}{\partial x_1}(\widehat{x}) & \frac{\partial f_{d_2}}{\partial x_2}(\widehat{x}) & \dots & \frac{\partial f_{d_2}}{\partial x_{d_1}}(\widehat{x}) \end{pmatrix}.$$

Если $x, y \in X$, то множество $[x, y] = \{z \in X : z = (1 - \lambda)x + \lambda y,\ 0 \le \lambda \le 1\}$ называется *отрезком*, соединяющий точки x и y.

Теорема 1 (о среднем). *Пусть X, Y — нормированные пространства, U — открытое множество в X, отображение $F\colon U \to Y$ дифференцируемо на U и $[x,y] \subset U$. Тогда*

$$\|F(y) - F(x)\|_Y \leq \sup_{z\in[x,y]} \|F'(z)\|\|y-x\|_X.$$

Доказательство. Если $F(y) = F(x)$, то утверждение теоремы очевидно. Пусть $F(y) \neq F(x)$. Тогда по следствию из теоремы Хана–Банаха (см. [2, стр. 195]) найдется элемент $y^* \in Y^*$ такой, что $\|y^*\|_{Y^*} = 1$ и

$$\langle y^*, F(y) - F(x)\rangle = \|F(y) - F(x)\|_Y. \tag{3}$$

Рассмотрим функцию

$$\varphi(t) = \langle y^*, F(x + t(y-x))\rangle, \quad 0 \leq t \leq 1.$$

В силу дифференцируемости F при достаточно малых Δt имеем

$$\begin{aligned}\varphi(t+\Delta t) - \varphi(t) &= \langle y^*, F(x + (t+\Delta t)(y-x)) - F(x + t(y-x))\rangle \\ &= \langle y^*, F'(x + t(y-x))(y-x)\rangle \Delta t + o(\Delta t).\end{aligned}$$

Тем самым функция φ дифференцируема в интервале $(0,1)$ и

$$\varphi'(t) = \langle y^*, F'(x + t(y-x))(y-x)\rangle.$$

Следовательно, по теореме Лагранжа существует такое $0 < \theta < 1$, что $\varphi(1) - \varphi(0) = \varphi'(\theta)$, или

$$\langle y^*, F(y) - F(x)\rangle = \langle y^*, F'(x + \theta(y-x))(y-x)\rangle.$$

Отсюда, учитывая (3),

$$\|F(y) - F(x)\|_Y \leq \|F'(x+\theta(y-x))\|\|y-x\|_X \leq \sup_{z\in[x,y]} \|F'(z)\|\|y-x\|_X.$$

$\square$

Пусть $\Lambda \in \mathcal{L}(X,Y)$. Применяя теорему о среднем к отображению $G(x) = F(x) - \Lambda x$, получаем

$$\|F(y) - F(x) - \Lambda(y-x)\|_Y \leq \sup_{z\in[x,y]} \|F'(z) - \Lambda\|\|y-x\|_X. \tag{4}$$

4. Теорема Ферма для гладких задач без ограничений

В этом пункте будет доказан изначальный результат теории экстремума — теорема Ферма (необходимое условие экстремума для гладких задач без ограничений).

Пусть U — открытое подмножество нормированного пространства X и $f : U \to \mathbb{R}$. Рассмотрим задачу

$$f(x) \to \mathrm{extr}, \quad x \in U. \tag{5}$$

Теорема 2 (Ферма)**.** *Если* $\widehat{x}$ *— локальный экстремум в задаче* (5) *и функция* f *дифференцируема в* $\widehat{x}$*, то*

$$f'(\widehat{x}) = 0. \tag{6}$$

Доказательство. Допустим, что линейный функционал $f'(\widehat{x})$ отличен от нуля. Тогда найдется элемент $x \in X$ такой, что $f'(\widehat{x}) \cdot x > 0$. В силу открытости U существует $\delta > 0$ такое, что $\widehat{x} + tx \in U$ для всех $|t| < \delta$. Дифференцируемость f в $\widehat{x}$ означает, что при таких t

$$f(\widehat{x} + tx) = f(\widehat{x}) + t(f'(\widehat{x}) \cdot x + o(t)/t).$$

Пусть $\delta_0 \leq \delta$ таково, что

$$|o(t)/t| < f'(\widehat{x}) \cdot x/2.$$

Тогда $f(\widehat{x} + tx) > f(\widehat{x})$ для всех $t \in (0, \delta_0)$ и $f(\widehat{x} + tx) < f(\widehat{x})$ для всех $t \in (-\delta_0, 0)$. Получили противоречие с тем, что $\widehat{x}$ — локальный экстремум. □

5. Вторая производная отображения. Формула Тейлора

Пусть X и Y — нормированные пространства, а U — открытое подмножество X. Если отображение $F\colon U \to Y$ дифференцируемо в каждой точке U, то определено отображение $F'\colon U \to \mathcal{L}(X, Y)$, сопоставляющее $x \in U$ производную $F'(x)$.

Дадим определение второй производной. Пусть отображение $F'\colon U \to \mathcal{L}(X, Y)$ дифференцируемо в точке $\widehat{x}$. Тогда говорят, что F дважды дифференцируема в $\widehat{x}$ и соответствующую (вторую) производную обозначают $F''(\widehat{x})$. Тем самым для всех $h \in X$, для которых $\widehat{x} + h \in U$, имеет место равенство

$$F'(\widehat{x} + h) = F'(\widehat{x}) + F''(\widehat{x})h + o(\|h\|_X).$$

Ясно, что $F''(\widehat{x}) \in \mathcal{L}(X, \mathcal{L}(X, Y))$.

Пространство $\mathcal{L}(X, \mathcal{L}(X, Y))$ изометрически изоморфно пространству $\mathcal{L}^2(X, Y)$ всех непрерывных билинейных отображений $B\colon X \times X \to Y$ с нормой

$$\|B\| = \sup_{\|x_1\|_X \leq 1,\ \|x_2\|_X \leq 1} \|B[x_1, x_2]\|_Y.$$

Этот изоморфизм осуществляется сопоставлением каждому отображению $\Lambda \in \mathcal{L}(X, \mathcal{L}(X, Y))$ отображения $B\colon X \times X \to Y$, действующего по правилу: $B[x_1, x_2] = \Lambda x_1[x_2]$ (действие оператора Λx_1 на элементе x_2). Очевидно, что B — билинейное отображение.

Пример 5.1. Найдем первую и вторую производную функции $F(h) = B[h, h]$, где B — билинейное отображение. Имеем

$$F(h + \Delta h) - F(h) = B[\Delta h, h] + B[h, \Delta h] + B[\Delta h, \Delta h].$$

Так как $\|B[\Delta h, \Delta h]\|_Y \leq \|B\| \|\Delta h\|_X^2$, то

$$F'(h)\xi = B[\xi, h] + B[h, \xi].$$

Далее, получаем

$$F'(h+\Delta h)\xi - F'(h)\xi = B[\xi,\Delta h] + B[\Delta h,\xi].$$

Таким образом, при всех $h,\xi,\eta \in X$

$$F''(h)[\xi,\eta] = B[\xi,\eta] + B[\eta,\xi]. \tag{7}$$

Пример 5.2. Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^d$ и функция $f\colon U \to \mathbb{R}$ дважды дифференцируема во всех точках из U. Тогда, как было показано в примере 3.1,

$$f'(x) = \left(\frac{\partial f}{\partial x_1}(x),\dots,\frac{\partial f}{\partial x_d}(x)\right).$$

аналогично примеру 3.2 (здесь $f'\colon U \to (\mathbb{R}^d)^*$) получаем, что если функция f' дифференцируема в точке $\widehat{x}$, то функции $\dfrac{\partial f}{\partial x_j}$, $j=1,\dots,d$, дифференцируемы в точке $\widehat{x}$ и

$$f''(\widehat{x}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(\widehat{x}) & \dfrac{\partial^2 f}{\partial x_1\partial x_2}(\widehat{x}) & \dots & \dfrac{\partial^2 f}{\partial x_1\partial x_d}(\widehat{x}) \\ \dfrac{\partial^2 f}{\partial x_2\partial x_1}(\widehat{x}) & \dfrac{\partial^2 f}{\partial x_2^2}(\widehat{x}) & \dots & \dfrac{\partial^2 f}{\partial x_2\partial x_d}(\widehat{x}) \\ \dots & \dots & \dots & \dots \\ \dfrac{\partial^2 f}{\partial x_d\partial x_1}(\widehat{x}) & \dfrac{\partial^2 f}{\partial x_d\partial x_2}(\widehat{x}) & \dots & \dfrac{\partial^2 f}{\partial x_d^2}(\widehat{x}) \end{pmatrix}. \tag{8}$$

Эту матрицу называют *матрицей Гесса* или *гессианом функции f в точке $\widehat{x}$.*

Известно (см., например, [1, Теорема 14.13]), что у дважды дифференцируемой функции смешанные производные совпадают

$$\frac{\partial^2 f}{\partial x_j\partial x_k} = \frac{\partial^2 f}{\partial x_k\partial x_j}.$$

Тем самым для дважды дифференцируемой функции матрица Гесса симметричная.

Мы докажем более общий результат. Пусть X и Y — нормированные пространства, а U — открытое подмножество X.

Теорема 3 (о смешанных производных)**.** *Если отображение $F\colon U \to Y$ дважды дифференцируемо в точке $\widehat{x} \in U$, то для всех $\xi,\eta \in X$*

$$F''(\widehat{x})[\xi,\eta] = F''(\widehat{x})[\eta,\xi].$$

Доказательство. Из определения второй производной для x достаточно близких к $\widehat{x}$

$$F'(x) - F'(\widehat{x}) = F''(\widehat{x})(x-\widehat{x}) + \alpha(x)\|x-\widehat{x}\|_X,$$

где $\alpha \in \mathcal{L}(X, Y)$ и $\alpha(x) \to 0$ при $x \to \widehat{x}$. При достаточно малых η положим

$$\varphi(x) = F(x+\eta) - F(x).$$

Имеем

$$\begin{aligned} \varphi'(x) &= F'(x+\eta) - F'(\widehat{x}) - (F'(x) - F'(\widehat{x})) = \\ &F''(\widehat{x})(x+\eta-\widehat{x}) + \alpha(x+\eta)\|x+\eta-\widehat{x}\|_X - F''(\widehat{x})(x-\widehat{x}) - \alpha(x)\|x-\widehat{x}\|_X \\ &= F''(\widehat{x})\eta + \alpha(x+\eta)\|x+\eta-\widehat{x}\|_X - \alpha(x)\|x-\widehat{x}\|_X. \end{aligned}$$

При $x = \widehat{x}$ получаем

$$\varphi'(\widehat{x}) = F''(\widehat{x})\eta + \alpha(\widehat{x}+\eta)\|\eta\|_X. \tag{9}$$

Таким образом,

$$\varphi'(x) - \varphi'(\widehat{x}) = \alpha(x+\eta)\|x+\eta-\widehat{x}\|_X - \alpha(x)\|x-\widehat{x}\|_X - \alpha(\widehat{x}+\eta)\|\eta\|_X.$$

Для произвольного $\varepsilon > 0$ найдется $\delta > 0$ такое, что при $\|x-\widehat{x}\|_X < \delta$ выполняется неравенство $\|\alpha(x)\| < \varepsilon$. Поэтому, если $\|x-\widehat{x}\|_X < \delta/2$ и $\|\eta\|_X < \delta/2$, то $\|\alpha(x+\eta)\| < \varepsilon$, $\|\alpha(x)\| < \varepsilon$ и $\|\alpha(\widehat{x}+\eta)\| < \varepsilon$. Следовательно,

$$\begin{aligned} \|\varphi'(x) - \varphi'(\widehat{x})\| &\le \varepsilon(\|x-\widehat{x}\|_X + \|\eta\|_X) + \varepsilon\|x-\widehat{x}\|_X + \varepsilon\|\eta\|_X \\ &= 2\varepsilon(\|x-\widehat{x}\|_X + \|\eta\|_X). \end{aligned} \tag{10}$$

Для достаточно малых η и ξ положим

$$\Delta(\eta, \xi) = F(\widehat{x}+\xi+\eta) - F(\widehat{x}+\xi) - F(\widehat{x}+\eta) + F(\widehat{x}).$$

Учитывая (9), имеем

$$\begin{aligned} \|\Delta(\eta,\xi) - F''(\widehat{x})[\eta,\xi]\|_Y &= \|\varphi(\widehat{x}+\xi) - \varphi(\widehat{x}) - F''(\widehat{x})[\eta,\xi]\|_Y \\ &= \|\varphi(\widehat{x}+\xi) - \varphi(\widehat{x}) - \varphi'(\widehat{x})\xi - \alpha(\widehat{x}+\eta)\xi\|\eta\|_X\|_Y. \end{aligned}$$

Используя (4), получаем

$$\begin{aligned} \|\Delta(\eta,\xi) - F''(\widehat{x})[\eta,\xi]\|_Y \le \sup_{x\in[\widehat{x},\widehat{x}+\xi]} &\|\varphi'(x) - \varphi'(\widehat{x})\|\|\xi\|_X \\ &+ \|\alpha(\widehat{x}+\eta)\|\|\xi\|_X\|\eta\|_X. \end{aligned}$$

Отсюда, с учетом (10), при $\|\xi\|_X < \delta/2$ и $\|\eta\|_X < \delta/2$ получаем

$$\begin{aligned} \|\Delta(\eta,\xi) - F''(\widehat{x})[\eta,\xi]\|_Y &\le 2\varepsilon(\|\xi\|_X + \|\eta\|_X)\|\xi\|_X + \varepsilon\|\xi\|_X\|\eta\|_X \\ &\le 3\varepsilon(\|\xi\|_X + \|\eta\|_X)\|\xi\|_X. \end{aligned}$$

Пусть теперь ξ и η — произвольные элементы X. Для достаточно малого $t \in \mathbb{R}$ будем иметь

$$\|\Delta(t\eta, t\xi) - t^2F''(\widehat{x})[\eta,\xi]\|_Y \le 3t^2\varepsilon(\|\xi\|_X + \|\eta\|_X)\|\xi\|_X.$$

Поскольку $\Delta(t\eta, t\xi) = \Delta(t\xi, t\eta)$, то

$$\|t^2F''(\widehat{x})[\eta,\xi] - t^2F''(\widehat{x})[\xi,\eta]\|_Y \le 3t^2\varepsilon(\|\xi\|_X + \|\eta\|_X)^2.$$

После сокращения на t^2 в силу произвольности ε приходим к доказываемому равенству.

$\square$

Теорема 4 (Формула Тейлора)**.** *Пусть X и Y — нормированные пространства, U — окрестность точки $\widehat{x} \in X$. Если отображение $F\colon U \to Y$ дважды дифференцируемо в точке $\widehat{x}$, то имеет место формула Тейлора*

$$F(\widehat{x}+h) = F(\widehat{x}) + F'(\widehat{x})h + \frac{1}{2}F''(\widehat{x})[h,h] + r(h),$$

где $r(h) = o(\|h\|_X^2)$, т. е. $\|r(h)\|_Y/\|h\|_X^2 \to 0$ при $h \to 0$.

Доказательство. Очевидно, что $r(0) = 0$. Из примера 5.1 следует, что $r'(0) = 0$. Таким образом, из (4) получаем

$$\|r(h)\|_Y = \|r(h) - r(0) - r'(0)h\|_Y \leq \sup_{h_1 \in [0,h]} \|r'(h_1) - r'(0)\| \|h\|_X.$$

Из того же примера 5.1 (см. (7)) и теоремы 3 вытекает, что $r''(0) = 0$. Поэтому

$$r'(h_1) - r'(0) = r''(0)h_1 + o(\|h_1\|_X) = \alpha(h_1)\|h_1\|_X,$$

где $\alpha(h_1) \to 0$ при $h_1 \to 0$. В силу того, что $\|h_1\|_X \leq \|h\|_X$, имеем

$$\|r'(h_1) - r'(0)\| \leq \beta(h)\|h\|_X,$$

где $\beta(h) \to 0$ при $h \to 0$. Таким образом,

$$\|r(h)\|_Y \leq \beta(h)\|h\|_X^2.$$

$\square$

6. Необходимые и достаточные условия экстремума второго порядка

Пусть U — открытое подмножество нормированного пространства X и $f : U \to \mathbb{R}$.

Теорема 5 (необходимые условия экстремума второго порядка в задаче без ограничений)**.** *Если $\widehat{x}$ — локальный минимум (максимум) в задаче* (5) *и функция f дважды дифференцируема в $\widehat{x}$, то $f'(\widehat{x}) = 0$ и для любого $h \in X$ выполняется неравенство*

$$f''(\widehat{x})[h,h] \geq 0 \quad (\leq 0).$$

Доказательство. Пусть, для определенности, $\widehat{x}$ — локальный минимум функции f. По теорема Ферма $f'(\widehat{x}) = 0$ и тогда по формуле Тейлора для любого $h \in X$ и достаточно малых $t \in \mathbb{R}$ имеем

$$0 \leq f(\widehat{x}+th) - f(\widehat{x}) = \frac{1}{2}t^2 f''(\widehat{x})[h,h] + o(t^2).$$

Отсюда (деля на t^2 и устремляя t к нулю) получаем требуемое неравенство. Для локального максимума рассуждения аналогичны. $\square$

Если $X = \mathbb{R}^d$, то $f''(\widehat{x})$ — матрица Гесса (8). Тем самым из доказанной теоремы вытекает

Следствие 1. *Если $X = \mathbb{R}^d$, $\widehat{x}$ — локальный минимум (максимум) в задаче* (5) *и функция f дважды дифференцируема в $\widehat{x}$, то $f'(\widehat{x}) = 0$ и для любого $h \in \mathbb{R}^d$ выполняется неравенство*

$$h^T f''(\widehat{x})h \geq 0 \quad (\leq 0).$$

Перейдем теперь к достаточным условиям экстремума.

Теорема 6 (достаточные условия экстремума второго порядка в задаче без ограничений)**.** *Если в задаче* (5) *функция f дважды дифференцируема в $\widehat{x}$, $f'(\widehat{x}) = 0$ и существует $\alpha > 0$ такое, что для всех $h \in X$*

$$f''(\widehat{x})[h, h] \geq \alpha \|h\|_X^2 \quad (\leq \alpha \|h\|_X^2),$$

то $\widehat{x}$ — локальный минимум (максимум).

Доказательство. Предположим для определенности, что

$$f''(\widehat{x})[h, h] \geq \alpha \|h\|_X^2.$$

Тогда для достаточно малых h по формуле Тейлора

$$f(\widehat{x} + h) - f(\widehat{x}) = \frac{1}{2} f''(\widehat{x})[h, h] + o(\|h\|_X^2) \geq \frac{1}{2} \alpha \|h\|_X^2 + o(\|h\|_X^2) = \left(\frac{\alpha}{2} + o(1)\right) \|h\|_X^2.$$

Отсюда получаем, что $f(\widehat{x} + h) - f(\widehat{x}) \geq 0$, т. е. $\widehat{x}$ — локальный минимум. Для локального максимума доказательство аналогично. □

Следствие 2. *Если $X = \mathbb{R}^d$, функция f в задаче* (5) *дважды дифференцируема в $\widehat{x}$, $f'(\widehat{x}) = 0$ и для всех $h \in \mathbb{R}^d$, $h \neq 0$,*

$$h^T f''(\widehat{x})h > 0 \quad (< 0), \tag{11}$$

то $\widehat{x}$ — локальный минимум (максимум).

Доказательство. Пусть $h^T f''(\widehat{x})h > 0$ для любого ненулевого $h \in \mathbb{R}^d$. Функция $h^T f''(\widehat{x})h$ непрерывна на $\mathbb{R}^d$. Обозначим через α ее минимальное значение на единичной сфере

$$\mathbb{S}^{d-1} = \{\, x \in \mathbb{R}^d : |x| = 1 \,\}.$$

Ясно, что $\alpha > 0$. Учитывая, что $h/|h|$ принадлежит $\mathbb{S}^{d-1}$, будем иметь

$$\frac{h^T}{|h|} f''(\widehat{x}) \frac{h}{|h|} > \alpha.$$

Следовательно, для всех $h \in \mathbb{R}^d$

$$h^T f''(\widehat{x})h \geq \alpha |h|^2.$$

Теперь утверждение следствия непосредственно вытекает из теоремы 6. □

Условия (11) означают, что квадратичная форма с матрицей (8) положительно (отрицательно) определена. Согласно критерию Сильвестра это равносильно тому, что главные миноры этой матрицы положительны (чередуют знаки, причем первый — отрицательный).

Заметим, что в бесконечномерном пространстве условия следствия 2 уже не являются достаточными для экстремума. Приведем соответствующий пример. Пусть

$$X = l_2 = \left\{ x = (x_1, x_2, \ldots) : \|x\|_{l_2} = \left(\sum_{j=1}^{\infty} x_j^2 \right)^{1/2} < \infty \right\}.$$

Рассмотрим отображение $f\colon l_2 \to \mathbb{R}$, задаваемое равенством

$$f(x) = \sum_{j=1}^{\infty} \left(\frac{x_j^2}{j^3} - x_j^4 \right).$$

Нетрудно убедиться, что

$$f'(x)h = \sum_{j=1}^{\infty} \left(\frac{2x_j h_j}{j^3} - 4x_j^3 h_j \right), \quad f''(x)[h,h] = \sum_{j=1}^{\infty} \left(\frac{2h_j^2}{j^3} - 12x_j^2 h_j^2 \right).$$

Следовательно, $f'(0) = 0$ и при всех $h \in l_2$, $h \neq 0$,

$$f''(0)[h,h] = \sum_{j=1}^{\infty} \frac{2h_j^2}{j^3} > 0.$$

Тем не менее, $\widehat{x} = 0$ не является локальным минимумом, т. к. $f(0) = 0$,

$$f(e_k/k) = \frac{1}{k^5} - \frac{1}{k^4} < 0, \quad k = 2, 3, \ldots,$$

где e_k — k-ый базисный вектор

$$(e_k)_j = \begin{cases} 1, & j = k, \\ 0, & j \neq k, \end{cases} \quad j = 1, 2, \ldots,$$

и любая окрестность нуля при достаточно больших k содержит точки e_k/k.

7. Строгая дифференцируемость. Теорема о суперпозиции

Пусть X — нормированное пространство, $\widehat{x} \in X$ и $r > 0$. Положим

$$B_X(\widehat{x}, r) = \{ x \in X : \|x - \widehat{x}\|_X < r \}.$$

Отображение $F\colon U \to Y$, где U — открытое подмножество X, называется *строго дифференцируемым в точке* $\widehat{x} \in U$, если найдется такой оператор $\Lambda \in \mathcal{L}(X,Y)$, что для любого $\varepsilon > 0$ существует $\delta = \delta(\varepsilon) > 0$, обладающее тем свойством, что для всех

$x_1, x_2 \in B_X(\widehat{x}, \delta)$ справедливо неравенство

$$\|F(x_1) - F(x_2) - \Lambda(x_1 - x_2)\|_Y \le \varepsilon \|x_1 - x_2\|_X.$$

Отсюда следует (полагая $x_2 = \widehat{x}$), что F дифференцируемо в $\widehat{x}$ и тем самым $\Lambda = F'(\widehat{x})$.

Предложение 1. *Пусть X, Y — нормированные пространства, U — открытое подмножество X и отображение $F\colon U \to Y$ непрерывно дифференцируемо в точке $\widehat{x} \in U$. Тогда F строго дифференцируемо в $\widehat{x}$.*

Доказательство. Пусть $\varepsilon > 0$ и $\delta = \delta(\varepsilon) > 0$ такое, что $\|F'(x) - F'(\widehat{x})\| < \varepsilon$ для $x \in B_X(\widehat{x}, \delta)$. Если $x_j \in B_X(\widehat{x}, \delta)$, $j = 1, 2$, то $[x_1, x_2] \subset B_X(\widehat{x}, \delta)$ и тогда, положив $\Lambda = F'(\widehat{x})$, в силу (4), получаем

$$\|F(x_1) - F(x_2) - F'(\widehat{x})(x_1 - x_2)\|_Y \le \varepsilon \|x_1 - x_2\|_X,$$

т. е. F строго дифференцируемо в $\widehat{x}$. □

Пусть X, Y, Z — нормированные пространства, U — окрестность точки $\widehat{x} \in X$, V — окрестность точки $\widehat{y} \in Y$, $\varphi\colon U \to V$, $\varphi(\widehat{x}) = \widehat{y}$, $\psi\colon V \to Z$, $F = \psi \circ \varphi\colon U \to Z$ — суперпозиция отображений φ и ψ.

Теорема 7 (о суперпозиции)**.** *Если отображение ψ дифференцируемо (строго дифференцируемо) в точке $\widehat{y}$, а φ дифференцируемо (строго дифференцируемо) в точке $\widehat{x}$, то отображение F дифференцируемо (строго дифференцируемо) в точке $\widehat{x}$ и*

$$F'(\widehat{x}) = \psi'(\widehat{y}) \circ \varphi'(\widehat{x}).$$

Доказательство. Положим для краткости $L = \varphi'(\widehat{x})$ и $M = \psi'(\widehat{y})$. Будем предполагать сначала строгую дифференцируемость ψ в точке $\widehat{y}$ и φ в точке $\widehat{x}$. По определению строгой дифференцируемости для любого $\varepsilon_1 > 0$ найдутся $\delta_1 > 0$ и $\delta_2 > 0$ такие, что для всех $x_1, x_2 \in B_X(\widehat{x}, \delta_1)$ и для всех $y_1, y_2 \in B_Y(\widehat{y}, \delta_2)$ справедливы неравенства

$$\|\varphi(x_1) - \varphi(x_2) - L(x_1 - x_2)\|_Y \le \varepsilon_1 \|x_1 - x_2\|_X, \tag{12}$$

$$\|\psi(y_1) - \psi(y_2) - M(y_1 - y_2)\|_Z \le \varepsilon_1 \|y_1 - y_2\|_Y. \tag{13}$$

Для любого $\varepsilon > 0$ выберем $\varepsilon_1 > 0$ так, чтобы выполнялось неравенство

$$\varepsilon_1 \|M\| + \varepsilon_1 \|L\| + \varepsilon_1^2 < \varepsilon.$$

По так выбранному ε_1 найдем $\delta_1 > 0$ и $\delta_2 > 0$ так, чтобы имели место неравенства (12) и (13). Положим

$$\delta = \min\left(\delta_1, \frac{\delta_2}{\varepsilon_1 + \|L\|}\right).$$

Если теперь $x_1, x_2 \in B_X(\widehat{x}, \delta)$, то из (12) имеем

$$\begin{aligned} (14) \quad \|\varphi(x_1) - \varphi(x_2)\|_Y &\le \|\varphi(x_1) - \varphi(x_2) - L(x_1 - x_2)\|_Y \\ &+ \|L(x_1 - x_2)\|_Y \le \varepsilon_1 \|x_1 - x_2\|_X + \|L\| \|x_1 - x_2\|_X \\ &= (\|L\| + \varepsilon_1)\|x_1 - x_2\|_X. \end{aligned}$$

Полагая в этом неравенстве $x_1 = \widehat{x}$, а потом $x_2 = \widehat{x}$, получаем

$$\|\varphi(x_j) - \widehat{y}\|_Y < (\|L\| + \varepsilon_1)\delta \le \delta_2, \quad j = 1, 2.$$

Таким образом, для $y_j = \varphi(x_j)$, $j = 1, 2$, справедливо (13). Пользуясь (13), (12) и (14), имеем

$$\begin{aligned} &\|F(x_1) - F(x_2) - M \circ L(x_1 - x_2)\|_Z \\ &\quad \le \|\psi(\varphi(x_1)) - \psi(\varphi(x_2)) - M(\varphi(x_1) - \varphi(x_2))\|_Z \\ &+ \|M(\varphi(x_1) - \varphi(x_2)) - M \circ L(x_1 - x_2)\|_Z \le \varepsilon_1 \|\varphi(x_1) - \varphi(x_2)\|_Y \\ &+ \|M\| \|\varphi(x_1) - \varphi(x_2) - L(x_1 - x_2)\|_Z \le \varepsilon_1 (\|L\| + \varepsilon_1)\|x_1 - x_2\|_X \\ &+ \|M\| \varepsilon_1 \|x_1 - x_2\|_X = (\varepsilon_1\|M\| + \varepsilon_1\|L\| + \varepsilon_1^2)\|x_1 - x_2\|_X \le \varepsilon \|x_1 - x_2\|_X. \end{aligned}$$

Докажем теперь утверждение теоремы для случая, когда ψ дифференцируемо в точке $\widehat{y}$ и φ дифференцируемо в точке $\widehat{x}$. По определению дифференцируемости для любого $\varepsilon_1 > 0$ найдутся $\delta_1 > 0$ и $\delta_2 > 0$ такие, что для всех $\Delta x \in B_X(0, \delta_1)$ и для всех $\Delta y \in B_Y(0, \delta_2)$ справедливы неравенства

$$(15) \qquad \|\varphi(\widehat{x} + \Delta x) - \varphi(\widehat{x}) - L\Delta x\|_Y \le \varepsilon_1 \|\Delta x\|_X,$$

$$(16) \qquad \|\psi(\widehat{y} + \Delta y) - \psi(\widehat{y}) - M\Delta y\|_Y \le \varepsilon_1 \|\Delta y\|_Y.$$

Для любого $\varepsilon > 0$ выберем $\varepsilon_1 > 0$ так, чтобы выполнялось неравенство

$$\varepsilon_1\|M\| + \varepsilon_1\|L\| + \varepsilon_1^2 < \varepsilon.$$

По так выбранному ε_1 найдем $\delta_1 > 0$ и $\delta_2 > 0$ так, чтобы имели место неравенства (15) и (16). Положим

$$\delta = \min\left(\delta_1, \frac{\delta_2}{\varepsilon_1 + \|L\|}\right).$$

Если теперь $\Delta x \in B_X(0, \delta)$, то из (15) имеем

$$\begin{aligned} (17) \quad \|\varphi(\widehat{x} + \Delta x) - \varphi(\widehat{x})\|_Y &\le \|\varphi(\widehat{x} + \Delta x) - \varphi(\widehat{x}) - L\Delta x\|_Y \\ &+ \|L\Delta x\|_Y \le \varepsilon_1\|\Delta x\|_X + \|L\|\|\Delta x\|_X \\ &= (\|L\| + \varepsilon_1)\|\Delta x\|_X < (\|L\| + \varepsilon_1)\delta \le \delta_2. \end{aligned}$$

Таким образом, для $\Delta y = \varphi(\widehat{x} + \Delta x) - \varphi(\widehat{x})$ справедливо неравенство (16). Имеем

$$
\begin{aligned}
&\|F(\widehat{x} + \Delta x) - F(\widehat{x}) - M \circ L\Delta x\|_Z \\
&\qquad \le \|\psi(\varphi(\widehat{x} + \Delta x)) - \psi(\varphi(\widehat{x})) - M(\varphi(\widehat{x} + \Delta x) - \varphi(\widehat{x}))\|_Z \\
&+ \|M(\varphi(\widehat{x}+\Delta x) - \varphi(\widehat{x})) - M \circ L\Delta x\|_Z = \|\psi(\widehat{y}+\Delta y) - \psi(\widehat{y}) - M\Delta y\|_Z \\
&\qquad + \|M(\varphi(\widehat{x} + \Delta x) - \varphi(\widehat{x})) - M \circ L\Delta x\|_Z \le \varepsilon_1\|\Delta y\|_Y \\
&\qquad + \|M\|\|\varphi(\widehat{x} + \Delta x) - \varphi(\widehat{x}) - L\Delta x\|_Y \le \varepsilon_1(\|L\| + \varepsilon_1)\|\Delta x\|_X \\
&\qquad + \|M\|\varepsilon_1\|\Delta x\|_X = (\varepsilon_1\|M\| + \varepsilon_1\|L\| + \varepsilon_1^2)\|\Delta x\|_X \le \varepsilon\|\Delta x\|_X.
\end{aligned}
$$

□

8. Теорема о полном дифференциале. Производная оператора Немыцкого

Теорема 8 (о полном дифференциале)**.** *Пусть X, Y и Z — нормированные пространства, W — открытое подмножество $X \times Y$ и $F\colon W \to Z$. Если частные производные F_x и F_y непрерывны в точке $(\widehat{x}, \widehat{y}) \in W$, то F — строго дифференцируемо в $(\widehat{x}, \widehat{y})$ и*

$$F'(\widehat{x}, \widehat{y})(\xi, \eta) = F_x(\widehat{x}, \widehat{y})\xi + F_y(\widehat{x}, \widehat{y})\eta, \quad \xi \in X,\ \eta \in Y.$$

Доказательство. Для любого $\varepsilon > 0$ найдется $\delta > 0$ такое, что $U = B_X(\widehat{x}, \delta) \times B_Y(\widehat{y}, \delta) \subset W$ и для всех $(x, y) \in U$

$$\|F_x(x, y) - F_x(\widehat{x}, \widehat{y})\| < \varepsilon/2, \quad \|F_y(x, y) - F_y(\widehat{x}, \widehat{y})\| < \varepsilon/2.$$

Имеем

$$
\begin{aligned}
\Delta &= F(x_1, y_1) - F(x_2, y_2) - F_x(\widehat{x}, \widehat{y})(x_1 - x_2) - F_y(\widehat{x}, \widehat{y})(y_1 - y_2) \\
&= F(x_1, y_1) - F(x_2, y_1) - F_x(\widehat{x}, \widehat{y})(x_1 - x_2) \\
&\qquad + F(x_2, y_1) - F(x_2, y_2) - F_y(\widehat{x}, \widehat{y})(y_1 - y_2)
\end{aligned}
$$

По теореме о среднем (см. (4)) для всех $(x_1, y_1), (x_2, y_2) \in U$

$$
\begin{aligned}
&\|\Delta\|_Z \le \sup_{a \in [x_1, x_2]} \|F_x(a, y_1) - F_x(\widehat{x}, \widehat{y})\|\|x_1 - x_2\|_X \\
&+ \sup_{b \in [y_1, y_2]} \|F_x(x_2, b) - F_x(\widehat{x}, \widehat{y})\|\|y_1 - y_2\|_Y \le \frac{\varepsilon}{2}\|x_1 - x_2\|_X + \frac{\varepsilon}{2}\|y_1 - y_2\|_Y \\
&\qquad \le \varepsilon \max\{\|x_1 - x_2\|_X, \|y_1 - y_2\|_Y\}.
\end{aligned}
$$

□

Пусть G открытое подмножество $\mathbb{R} \times \mathbb{R}^{d_1}$ и $f\colon G \to \mathbb{R}^{d_2}$ — функция переменных $t \in \mathbb{R}$ и $x \in \mathbb{R}^{d_1}$, непрерывная вместе со своей частной производной f_x на G. Пусть существует функция $x \in C([t_0, t_1], \mathbb{R}^{d_1})$ такая, что $\Gamma(x) = \{\, (t, x(t)) : t \in [t_0, t_1] \,\} \subset G$. Положим

$$U = \{\, x \in C([t_0, t_1], \mathbb{R}^{d_1}) : \Gamma(x) \subset G \,\}.$$

Нетрудно убедиться, что U открыто в $C([t_0,t_1],\mathbb{R}^{d_1})$. Отображение $F\colon U\to C([t_0,t_1],\mathbb{R}^{d_2})$, определенное по правилу

$$F(x)(t)=f(t,x(t))$$

называется *оператором Немыцкого.*

Предложение 2. *Оператор Немыцкого F непрерывно дифференцируем на U и $F'(x)h(t)=f_x(t,x(t))h(t)$ для любых $x\in U$, $h\in C([t_0,t_1],\mathbb{R}^{d_1})$ и $t\in[t_0,t_1]$.*

Доказательство. Пусть $\widehat{x}\in U$. Существует $\delta_0>0$ такое, что компакт

$$K=\{\,(t,x):t\in[t_0,t_1],\ |x-\widehat{x}(t)|\le\delta_0\,\}$$

принадлежит G. Пусть $\varepsilon>0$. Функция f_x равномерно непрерывна на K и поэтому найдется $0<\delta\le\delta_0$ такое, что если $|x_1-x_2|<\delta$, то $\|f_x(t,x_1)-f_x(t,x_2)\|<\varepsilon$ для всех $(t,x_j)\in K$, $j=1,2$.

Для любого $t\in[t_0,t_1]$ отображение $g\colon B_{\mathbb{R}^{d_1}}(\widehat{x}(t),\delta)\to\mathbb{R}^{d_2}$, $g(x)=f(t,x)-f_x(t,\widehat{x}(t))x$, дифференцируемо на $B_{\mathbb{R}^{d_1}}(\widehat{x}(t),\delta)$ и его производная в точке x имеет вид $g'(x)=f_x(t,x)-f_x(t,\widehat{x}(t))$. Пусть $x_j\in B_{C([t_0,t_1],\mathbb{R}^{d_1})}(\widehat{x},\delta)$, $j=1,2$. Тогда $x_j(t)\in B_{\mathbb{R}^{d_1}}(\widehat{x}(t),\delta)$, $j=1,2$, и мы имеем по теореме о среднем, примененной к отображению g (учитывая, что если $x\in[x_1(t),x_2(t)]$, то $x\in B_{\mathbb{R}^n}(\widehat{x}(t),\delta)$)

$$\begin{aligned}&|f(t,x_1(t))-f(t,x_2(t))-f_x(t,\widehat{x}(t))(x_1(t)-x_2(t))|\le\\ &\le\sup_{x\in[x_1(t),x_2(t)]}\|f_x(t,x)-f_x(t,\widehat{x}(t))\||x_1(t)-x_2(t)|\le\varepsilon|x_1(t)-x_2(t)|.\end{aligned}$$

Поскольку это верно для любого $t\in[t_0,t_1]$, то отсюда следует, что отображение F строго дифференцируемо в $\widehat{x}$ и что $F'(\widehat{x})h(t)=f_x(t,\widehat{x}(t))h(t)$. Так как $\widehat{x}$ — произвольная функция из U, то F дифференцируемо на U.

Непрерывная дифференцируемость отображения F на U проверяется непосредственно, используя равномерную непрерывность f_x на соответствующем компакте. $\square$

Нам понадобится еще один оператор, который является некоторым обобщением оператора Немыцкого. Пусть G — открытое подмножество $\mathbb{R}\times\mathbb{R}^{d_1}\times\mathbb{R}^{d_3}$ и $f\colon G\to\mathbb{R}^{d_2}$ — функция переменных $t\in\mathbb{R}$, $x\in\mathbb{R}^{d_1}$ и $u\in\mathbb{R}^{d_3}$, непрерывная вместе со своими частными производными f_x и f_u на G. Пусть существует пара $(x,u)\in C^1([t_0,t_1],\mathbb{R}^{d_1})\times C([t_0,t_1],\mathbb{R}^{d_3})$ такая, что

$$\Gamma(x,u)=\{\,(t,x(t),u(t)):t\in[t_0,t_1]\,\}\subset G.$$

Положим

$$U=\{\,(x,u)\in C^1([t_0,t_1],\mathbb{R}^{d_1})\times C([t_0,t_1],\mathbb{R}^{d_3}):\Gamma(x,u)\subset G\,\}.$$

Легко проверить, что множество открыто в $C([t_0,t_1],\mathbb{R}^{d_1})$. Определим отображение $F\colon U\to C([t_0,t_1],\mathbb{R}^{d_2})$ по правилу

$$F(x,u)(t)=f(t,x(t),u(t)),$$

которое назовем *обобщенным оператором Немыцкого.*

Следствие 3. *Обобщенный оператор Немыцкого F непрерывно дифференцируем на U и*

$$F'(x,u)(h(t),\xi(t))=f_x(t,x(t),u(t))h(t)+f_u(t,x(t),u(t))\xi(t)$$

для любых $(x,u)\in U$, $(h,\xi)\in C^1([t_0,t_1],\mathbb{R}^{d_1})\times C([t_0,t_1],\mathbb{R}^{d_3})$ и $t\in[t_0,t_1]$.

Доказательство. Частная производная по x отображения F, согласно предложению 2, равна $f_x(t,x(t),u(t))$. Ее непрерывность на $C([t_0,t_1],\mathbb{R}^{d_1})\times C([t_0,t_1],\mathbb{R}^{d_3})$ вытекает из равномерной непрерывности $f_x(t,x,u)$ на компакте вида

$$K_1=\{\,(t,x,u): t\in[t_0,t_1],\ |x-\widehat{x}(t)|\le\delta_0,\ |u-\widehat{u}(t)|\le\delta_0\,\}.$$

Тем самым она непрерывна и на $C^1([t_0,t_1],\mathbb{R}^{d_1})\times C([t_0,t_1],\mathbb{R}^{d_3})$. Частная производная по u также непрерывна в силу тех же причин. Поэтому по теореме о полном дифференциале (теорема 8) отображение F непрерывно дифференцируемо на U и справедлива соответствующая формула для производной. $\square$

9. Лемма о правом обратном. Теорема о неявной функции. Теорема Люстерника

Лемма 1 (о правом обратном)**.** *Пусть X и Y — банаховы пространства, $\Lambda\in\mathcal{L}(X,Y)$ и $\operatorname{Im}\Lambda=Y$. Тогда существуют отображение $R\colon Y\to X$ и константа $\gamma>0$ такие, что $\Lambda R(y)=y$ и $\|R(y)\|_X\le\gamma\|y\|_Y$ для любого $y\in Y$.*

Доказательство. По теореме Банаха об открытом отображении [2, стр. 243] множество $\Lambda(B_X(0,1))$ открыто. Оно, очевидно, содержит ноль и тем самым содержит некоторый шар $B_Y(0,r)$, $r>0$, т. е. для каждого $z\in B_Y(0,r)$ найдется элемент $x(z)\in B_X(0,1)$ такой, что $\Lambda x(z)=z$. Положим $R(0)=0$, а если $y\ne0$, то

$$R(y)=\frac{2\|y\|_Y}{r}x\left(\frac{r}{2\|y\|_Y}y\right).$$

Тогда $\Lambda R(y)=y$ и $\|R(y)\|_X\le\gamma\|y\|_Y$, где $\gamma=2/r$. $\square$

Лемма 2 (о замкнутости образа)**.** *Пусть X и Y — банаховы пространства, $A\in\mathcal{L}(X,\mathbb{R}^d)$, $B\in\mathcal{L}(X,Y)$, $C\colon X\to\mathbb{R}^d\times Y$, $Cx=(Ax,Bx)$ и $\operatorname{Im}B=Y$. Тогда $\operatorname{Im}C$ — замкнутое подпространство в $\mathbb{R}^d\times Y$.*

Доказательство. Пусть $(y,z) \in \operatorname{cl}\operatorname{Im} C$ и пусть $\{x_n\}$ — последовательность в X такая, что $y = \lim_{n\to\infty} Ax_n$ и $z = \lim_{n\to\infty} Bx_n$. Положим $h_n = R(Bx_n - z)$, где R — правый обратный к B, тогда $B(x_n - h_n) = z$. Так как $\|h_n\|_X = \|R(Bx_n - z)\|_X \le \gamma\|Bx_n - z\|_Y$, то $h_n \to 0$ при $n \to \infty$ и поэтому $\lim_{n\to\infty} A(x_n - h_n) = y$. Таким образом, y принадлежит замыканию образа множества

$$X_z = \{\, x \in X : Bx = z \,\}$$

при отображении A. Но $A(X_z)$ — линейное многообразие в $\mathbb{R}^d$ и тем самым замкнуто. Следовательно, существует такое $\overline{x} \in X$, что $B\overline{x} = z$ и $y = A\overline{x}$, т. е. $(y,z) \in \operatorname{Im} C$. □

Теорема 9 (Обобщенная теорема о неявной функции)**.** *Пусть Σ — топологическое пространство, X и Y — банаховы пространства, U — окрестность точки $\widehat{x} \in X$, $F\colon U \times \Sigma \to Y$ и $\widehat{\sigma} \in \Sigma$. Если*

1) $F(\widehat{x},\widehat{\sigma}) = 0$;
2) *F непрерывно в точке $(\widehat{x},\widehat{\sigma})$;*
3) *F дифференцируемо по x в точке $(\widehat{x},\widehat{\sigma})$ и для каждого $\varepsilon > 0$ существуют окрестности $U(\varepsilon) \subset U$ и $V(\varepsilon)$ точек $\widehat{x}$ и $\widehat{\sigma}$ такие, что для всех $x, x' \in U(\varepsilon)$ и $\sigma \in V(\varepsilon)$ выполняется соотношение*

$$\|F(x,\sigma) - F(x',\sigma) - F_x(\widehat{x},\widehat{\sigma})(x - x')\|_Y \le \varepsilon\|x - x'\|_X;$$

4) $\operatorname{Im} F_x(\widehat{x},\widehat{\sigma}) = Y$,

то найдутся окрестности $U_0 \subset U$ и V_0 точек $\widehat{x}$ и $\widehat{\sigma}$, отображение $\varphi\colon U_0 \times V_0 \to U$ и константа $K > 0$ такие, что $F(\varphi(x,\sigma),\sigma) = 0$ и $\|\varphi(x,\sigma) - x\|_X \le K\|F(x,\sigma)\|_Y$ для всех $(x,\sigma) \in U_0 \times V_0$.

Доказательство. Обозначим для краткости $\Lambda = F_x(\widehat{x},\widehat{\sigma})$. Так как $\operatorname{Im} F_x(\widehat{x},\widehat{\sigma}) = Y$, то по лемме о правом обратном (лемма 1) существует отображение $R\colon Y \to X$ и константа $\gamma > 0$ такие, что $\Lambda R(y) = y$ и $\|R(y)\|_X \le \gamma\|y\|_Y$ для всех $y \in Y$.

Пусть $\varepsilon_0 > 0$ таково, что $\theta = \varepsilon_0\gamma < 1$ и $U(\varepsilon_0)$ и $V(\varepsilon_0)$ — окрестности точек $\widehat{x}$ и $\widehat{\sigma}$, соответствующие ε_0 (из формулировки теоремы). Пусть $\delta > 0$ такое, что $B_X(\widehat{x},\delta) \subset U(\varepsilon_0)$. Выберем окрестности U_0 и V_0 так, что $U_0 \subset B_X(\widehat{x},\delta/2)$, $V_0 \subset V(\varepsilon_0)$ и при этом $\|F(x,\sigma)\|_Y < \delta(1-\theta)/2\gamma$, если $(x,\sigma) \in U_0 \times V_0$.

Пусть $(x,\sigma) \in U_0 \times V_0$. Рассмотрим последовательность

$$x_n = x_{n-1} - R(F(x_{n-1},\sigma)), \quad n \in \mathbb{N}, \quad x_0 = x. \tag{18}$$

Докажем, что эта последовательность принадлежит $B_X(\widehat{x},\delta)$ и фундаментальна. Первое доказываем по индукции. Ясно, что $x_0 \in B_X(\widehat{x},\delta)$. Пусть $x_k \in B_X(\widehat{x},\delta)$, $1 \le k \le n$. Применяя к обеим частям (18) оператор Λ, получим

$$\Lambda(x_n - x_{n-1}) = -F(x_{n-1},\sigma). \tag{19}$$

Используя последовательно (18), оценку для правого обратного, (19), условие 3) теоремы и затем итерируя процедуру, будем иметь

$$(20)\quad \|x_{n+1}-x_n\|_X \le \gamma\|F(x_n,\sigma)\|_Y = \gamma\|F(x_n,\sigma)-F(x_{n-1},\sigma) \\ -\Lambda(x_n-x_{n-1})\|_Y \le \theta\|x_n-x_{n-1}\|_X \le \ldots \le \theta^n\|x_1-x\|_X.$$

Далее, по неравенству треугольника, (20), (18), условию 2) теоремы и согласно определению окрестностей U_0 и V_0, получаем, что

$$\begin{aligned}(21)\quad \|x_{n+1}-\widehat{x}\|_X &\le \|x_{n+1}-x\|_X+\|x-\widehat{x}\|_X \\ &\le \|x_{n+1}-x_n\|_X+\ldots+\|x_1-x\|_X+\|x-\widehat{x}\|_X \\ &\le (\theta^n+\theta^{n-1}+\ldots+1)\|x_1-x\|_X+\|x-\widehat{x}\|_X \\ &\le \frac{\gamma}{1-\theta}\|F(x,\sigma)\|_Y+\|x-\widehat{x}\|_X < \frac{\delta}{2}+\frac{\delta}{2}=\delta,\end{aligned}$$

т. е. $x_{n+1}\in B_X(\widehat{x},\delta)$ и значит, вся последовательность $\{x_n\}$ принадлежит $B_X(\widehat{x},\delta)$.

Последовательность $\{x_n\}$ фундаментальна. Действительно, используя (20) и рассуждая как в предыдущем неравенстве, будем иметь для всех $n,m\in\mathbb{N}$

$$\|x_{n+m}-x_n\|_X \le \|x_{n+m}-x_{n+m-1}\|_X+\ldots+\|x_{n+1}-x_n\|_X \le \\ \le (\theta^{n+m-1}+\ldots+\theta^n)\|x_1-x\|_X \le \frac{\theta^n}{1-\theta}\|x_1-x\|_X \le \frac{\delta}{2}\theta^n.$$

Положим $\varphi(x,\sigma)=\lim_{n\to\infty}x_n$. Из (21) следует, что $\varphi(x,\sigma)\in B_X(\widehat{x},\delta)\subset U$. Из условия 3) теоремы следует, что для каждого $\sigma\in V_0$ отображение $F(x,\sigma)$ непрерывно на U_0, и тогда переходя к пределу в (19) при $n\to\infty$, получаем, что $F(\varphi(x,\sigma),\sigma)=0$.

В (21) доказано, что

$$\|x_n-x\|_X \le \frac{\gamma}{1-\theta}\|F(x,\sigma)\|_Y.$$

Переходя здесь к пределу при $n\to\infty$, приходим к неравенству $\|\varphi(x,\sigma)-x\|_X\le K\|F(x,\sigma)\|_Y$, где $K=\gamma/(1-\theta)$. □

Рассмотрим случай, когда Σ состоит из одного элемента (зависимость от него отмечать не будем). Тогда из теоремы 9, рассматривая вместо отображения $F(x)$ отображение $F(x)-F(\widehat{x})$, получаем

Следствие 4. *Пусть X и Y — банаховы пространства, U — окрестность точки $\widehat{x}\in X$ и $F\colon U\to Y$. Если F строго дифференцируемо в точке $\widehat{x}$ и $\operatorname{Im}F'(\widehat{x})=Y$, то найдется окрестность $U_0\subset U$ точки $\widehat{x}$, отображение $\varphi\colon U_0\to U$ и константа $K>0$ такие, что $F(\varphi(x))=F(\widehat{x})$ и $\|\varphi(x)-x\|_X\le K\|F(x)-F(\widehat{x})\|_Y$ для всех $x\in U_0$.*

Пусть M — непустое подмножество нормированного пространства X. Элемент $h\in X$ называется *касательным вектором к M в* *точке* $\widehat{x} \in M$, если существуют $\varepsilon > 0$ и отображение $r\colon (-\varepsilon, \varepsilon) \to X$ такие, что $\widehat{x}+th+r(t) \in M$ для всех $t \in (-\varepsilon, \varepsilon)$ и $\|r(t)\|_X/t \to 0$, при $t \to 0$. Множество всех касательных векторов к M в точке $\widehat{x} \in M$ обозначается через $T_{\widehat{x}}M$.

Теорема 10 (Люстерника)**.** *Пусть X, Y — банаховы пространства, U — окрестность точки $\widehat{x} \in X$, отображение $F\colon U \to Y$ — строго дифференцируемо в $\widehat{x}$, $\operatorname{Im} F'(\widehat{x}) = Y$ и $M = \{\, x \in U : F(x) = F(\widehat{x}) \,\}$. Тогда $T_{\widehat{x}}M = \operatorname{Ker} F'(\widehat{x})$.*

Доказательство. Пусть $h \in T_{\widehat{x}}M$ и r из определения касательного вектора. Тогда вследствие дифференцируемости F в точке $\widehat{x}$ имеем

$$0 = F(\widehat{x} + th + r(t)) - F(\widehat{x}) = tF'(\widehat{x})h + o(t),$$

откуда (деля на t и переходя к пределу при $t \to 0$) следует, что $h \in \operatorname{Ker} F'(\widehat{x})$.

Обратно, пусть $h \in \operatorname{Ker} F'(\widehat{x})$. Отображение $F(x)$ удовлетворяет условиям следствия 4. Следовательно, найдется такая окрестность $U_0 \subset U$ точки $\widehat{x}$, отображение $\varphi\colon U_0 \to U$ и константа $K > 0$, такие, что $F(\varphi(x)) = F(\widehat{x})$ и $\|\varphi(x) - x\|_X \le K\|F(x) - F(\widehat{x})\|_Y$ для всех $x \in U_0$. Пусть $\varepsilon > 0$ таково, что $\widehat{x} + th \in U_0$ при $t \in (-\varepsilon, \varepsilon)$. Положим $r(t) = \varphi(\widehat{x}+th) - \widehat{x} - th$. Имеем $F(\widehat{x}+th+r(t)) = F(\widehat{x})$ и

$$\|r(t)\|_X \le K\|F(\widehat{x}+th) - F(\widehat{x})\|_Y = K\|tF'(\widehat{x})h + o(t)\|_Y = K\|o(t)\|_Y,$$

т. е. h — касательный вектор. $\square$

10. Теоремы отделимости. Леммы об аннуляторах

Пусть A и B — непустые подмножества нормированного пространства X. Говорят, что ненулевой функционал $x^* \in X^*$ *отделяет множества A и B*, если

$$\sup_{x\in A}\langle x^*, x\rangle \le \inf_{x\in B}\langle x^*, x\rangle.$$

Если неравенство строгое, то говорят, что x^* *строго отделяет A и B*.

Пусть число $\gamma \in \mathbb{R}$ таково, что

$$\sup_{x\in A}\langle x^*, x\rangle \le \gamma \le \inf_{x\in B}\langle x^*, x\rangle.$$

Тогда, геометрически, отделимость множеств A и B означает, что они расположены по разные стороны от гиперплоскости

$$\{\, x \in X : \langle x^*, x\rangle = \gamma \,\}.$$

Напомним формулировку первой теоремы отделимости (см. [2, стр. 243]).

Теорема 11 (Первая теорема отделимости)**.** *Пусть A и B — непустые выпуклые подмножества нормированного пространства X, причем $\operatorname{int} A \neq \emptyset$ и $B \cap \operatorname{int} A = \emptyset$. Тогда множества A и B отделимы.*

Отсюда следует

Теорема 12 (Вторая теорема отделимости)**.** *Пусть A — непустое замкнутое выпуклое подмножество нормированного пространства X и $\widehat{x} \notin A$. Тогда множества A и $\widehat{x}$ строго отделимы.*

Доказательство. Так как A замкнуто, то дополнение к A открыто и поэтому существует такое $r > 0$, что открытый шар $B_X(\widehat{x}, r)$ не пересекается с A. Тогда по первой теореме отделимости существует ненулевой функционал $x^* \in X^*$ такой, что

$$\sup_{x \in A} \langle x^*, x \rangle \leq \inf_{x \in B_X(\widehat{x}, r)} \langle x^*, x \rangle.$$

Но

$$\inf_{x \in B_X(\widehat{x}, r)} \langle x^*, x \rangle < \langle x^*, \widehat{x} \rangle,$$

так как ненулевой линейный непрерывный функционал не может достигать точной нижней грани во внутренней точке. Следовательно, множества A и $\widehat{x}$ строго отделимы. □

Пусть L — подпространство нормированного пространства X. Множество

$$L^{\perp} = \{\, x^* \in X^* : \langle x^*, x \rangle = 0, \ \forall\, x \in L \,\}$$

называется *аннулятором* L. Легко видеть, что $L^{\perp}$ — замкнутое подпространство в X^*.

Лемма 3 (о нетривиальности аннулятора)**.** *Пусть L — замкнутое подпространство нормированного пространства X, не совпадающее с X. Тогда $L^{\perp}$ содержит ненулевой элемент.*

Доказательство. Так как $L \neq X$, то существует $\widehat{x} \notin L$. Множество L, очевидно, выпукло и по условию замкнуто, поэтому по второй теореме отделимости найдется ненулевой функционал $x^* \in X^*$ такой, что

$$\sup_{x \in L} \langle x^*, x \rangle < \langle x^*, \widehat{x} \rangle. \tag{22}$$

Тогда $x^* \in L^{\perp}$. Действительно, если $\langle x^*, x_0 \rangle \neq 0$ для некоторого $x_0 \in L$, то так как $\alpha x_0 \in L$ для любого $\alpha \in \mathbb{R}$, мы имеем

$$\sup_{x \in L} \langle x^*, x \rangle \geq \sup_{\alpha \in \mathbb{R}} \langle x^*, \alpha x_0 \rangle = \sup_{\alpha \in \mathbb{R}} \alpha \langle x^*, x_0 \rangle = +\infty,$$

что противоречит (22). □

Лемма 4 (об аннуляторе ядра)**.** *Пусть X и Y — банаховы пространства, $A \in \mathcal{L}(X, Y)$ и $\operatorname{Im} A = Y$. Тогда $(\operatorname{Ker} A)^{\perp} = \operatorname{Im} A^*$.*

Доказательство. Пусть $x^* \in \operatorname{Im} A^*$. Тогда $x^* = A^* y*$, где $y* \in Y^*$. Для любого $x \in \operatorname{Ker} A$ имеем

$$\langle x^*, x\rangle = \langle A^* y^*, x\rangle = \langle y^*, Ax\rangle = 0.$$

Тем самым $x^* \in (\operatorname{Ker} A)^{\perp}$.

Пусть $x^* \in (\operatorname{Ker} A)^{\perp}$. Образ оператора $M\colon X \to \mathbb{R} \times Y$,

$$Mx = (\langle x^*, x\rangle, Ax),$$

замкнут по лемме о замкнутости образа (лемма 2) и не совпадает с $\mathbb{R} \times Y$, так как $(1, 0) \notin \operatorname{Im} M$. Следовательно, по лемме о нетривиальности аннулятора (лемма 3) существует ненулевой функционал $(\alpha, y^*) \in \mathbb{R} \times Y^*$ такой, что

$$\alpha\langle x^*, x\rangle + \langle y^*, Ax\rangle = 0$$

для всех $x \in X$. При этом $\alpha \neq 0$, ибо в противном случае функционал y^* был бы нулевым в силу того, что $\operatorname{Im} A = Y$. Таким образом,

$$\langle x^* + \alpha^{-1} A^* y^*, x\rangle = 0$$

для всех $x \in X$. Следовательно,

$$x^* = A^*(-\alpha^{-1} y^*) \in \operatorname{Im} A^*.$$

□

11. Правило множителей Лагранжа для гладких задач с ограничениями типа равенств

Пусть U — открытое подмножество банахова пространства X, $f : U \to \mathbb{R}$ и $F\colon U \to Y$, где Y — банахово пространство. Задачу

$$f(x) \to \text{extr}, \quad F(x) = 0, \tag{23}$$

называют *задачей с ограничениями типа равенств.* Если функция f и отображение F обладают некоторой гладкостью, то говорят о *гладкой задаче с ограничениями типа равенств.*

Сопоставим задаче (23) *функцию Лагранжа*

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

где $\lambda_0 \in \mathbb{R}$ и $y^* \in Y^*$ называются *множителями Лагранжа.*

Теорема 13 (Правило множителей Лагранжа)**.** *Если $\widehat{x} \in U$ — локальный экстремум в задаче* (23)*, функция f — дифференцируема в $\widehat{x}$, отображение F — строго дифференцируемо в $\widehat{x}$ и $\operatorname{Im} F'(\widehat{x})$ — замкнутое подпространство в Y, то найдутся, не равные одновременно нулю, множители Лагранжа λ_0 и y^* такие, что*

$$\mathcal{L}_x(\widehat{x}, \lambda_0, y^*) = 0 \quad \Leftrightarrow \quad \lambda_0 f'(\widehat{x}) + (F'(\widehat{x}))^* y^* = 0. \tag{24}$$

Если $\operatorname{Im} F'(\widehat{x}) = Y$, то $\lambda_0 \neq 0$.

Доказательство. Рассмотрим сначала случай, когда $\operatorname{Im} F'(\widehat{x}) = Y$. Пусть $h \in \operatorname{Ker} F'(\widehat{x})$. Отображение F удовлетворяет условиям теоремы Люстерника (теорема 10) и поэтому $h \in T_{\widehat{x}}M$, т. е. существуют $\varepsilon > 0$ и отображение $r\colon (-\varepsilon, \varepsilon) \to X$ такие, что $F(\widehat{x}+th+r(t)) = 0$ для $t \in (-\varepsilon, \varepsilon)$ и $r(t) = o(t)$ при $t \to 0$. Таким образом, элементы $\widehat{x}+th+r(t)$, $t \in (-\varepsilon, \varepsilon)$, допустимы в (23) и так как $\widehat{x}$ — локальный экстремум в этой задаче, то

$$f(\widehat{x} + th + r(t)) - f(\widehat{x}) = t\langle f'(\widehat{x}), h\rangle + o(t)$$

сохраняет знак для достаточно малых t. Деля последнее соотношение на $t > 0$ и устремляя t к нулю, получаем, что $\langle f'(\widehat{x}), h\rangle$ сохраняет знак. Но h — произвольный элемент из $\operatorname{Ker} F'(\widehat{x})$ и поэтому $\langle f'(\widehat{x}), h\rangle = 0$ для любого $h \in \operatorname{Ker} F'(\widehat{x})$, т. е. $f'(\widehat{x}) \in (\operatorname{Ker} F'(\widehat{x}))^{\perp}$. Согласно лемме об аннуляторе ядра (лемма 4) $f'(\widehat{x}) \in \operatorname{Im}(F'(\widehat{x}))^*$ и, следовательно, существует функционал $y^* \in Y^*$ такой, что $f'(\widehat{x}) = -(F'(\widehat{x}))^* y^*$, или $f'(\widehat{x}) + (F'(\widehat{x}))^* y^* = 0$. Тем самым равенство (24) с $\lambda_0 = 1$ доказано.

Пусть теперь $\operatorname{Im} F'(\widehat{x}) \neq Y$. Так как по условию подпространство $\operatorname{Im} F'(\widehat{x})$ замкнуто, то по лемме о нетривиальности аннулятора (лемма 3) существует ненулевой функционал $y^* \in Y^*$ такой, что $\langle y^*, F'(\widehat{x})x\rangle = 0$ для любого $x \in X$, т. е. $(F'(\widehat{x}))^* y^* = 0$. Это доказывает утверждение теоремы в рассматриваемом случае с $\lambda_0 = 0$. $\square$

Рассмотрим частный случай задачи (23), когда $X = \mathbb{R}^{d_1}$, $Y = \mathbb{R}^{d_2}$, а отображение F задается функциями $f_j\colon \mathbb{R}^{d_1} \to \mathbb{R}$, $j = 1, \dots, d_2$, т. е. $F(x) = (f_1(x), \dots, f_{d_2}(x))^T$, $x \in \mathbb{R}^{d_1}$. Таким образом, рассматривается задача

$$f_0(x) \to \operatorname{extr}, \quad f_j(x) = 0, \quad j = 1, \dots, d_2. \tag{25}$$

Поскольку линейные функционалы на $\mathbb{R}^{d_2}$ являются вектор-строками $(\lambda_1, \dots, \lambda_{d_2})$, то функция Лагранжа задачи (25) записывается в виде

$$\mathcal{L}(x, \overline{\lambda}) = \sum_{j=0}^{d_2} \lambda_j f_j(x),$$

где $\overline{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_{d_2})$.

Классическим правилом множителей Лагранжа для гладких конечномерных задач является следующее утверждение.

Теорема 14 (Правило множителей Лагранжа в конечномерном случае). *Если $\widehat{x}$ — локальный экстремум в задаче* (25), *функция f_0 дифференцируема в $\widehat{x}$, а функции f_j, $j = 1, \dots, d_2$, строго дифференцируемы в $\widehat{x}$, то найдутся, не равные одновременно нулю, множители Лагранжа $\lambda_0, \lambda_1, \dots, \lambda_{d_2}$ такие, что*

$$\mathcal{L}_x(\widehat{x}, \overline{\lambda}) = 0 \quad \Leftrightarrow \quad \sum_{j=0}^{d_2} \lambda_j f_j'(\widehat{x}) = 0.$$

Если векторы $f_1'(\widehat{x}), \dots, f_{d_2}'(\widehat{x})$ линейно независимы, то $\lambda_0 \neq 0$.

Доказательство сразу следует из предыдущей теоремы, если учесть, что подпространство $\operatorname{Im} F'(\widehat{x})$ конечномерно и поэтому замкнуто, а линейная независимость векторов $f_1'(\widehat{x}), \dots, f_{d_2}'(\widehat{x})$ эквивалентна условию $\operatorname{Im} F'(\widehat{x}) = Y$.

Приведем пример, показывающий, что с $\lambda_0 \neq 0$ правило множителей Лагранжа может не выполняться. Рассмотрим задачу

$$x_1 \to \min, \quad x_1^3 - x_2^2 = 0.$$

Здесь $d_1 = 2$, $d_2 = 1$, $f_0(x) = x_1$, $f_1(x) = x_1^3 - x_2^2$,

$$\mathcal{L}(x, \lambda_0, \lambda_1) = \lambda_0 x_1 + \lambda_1(x_1^3 - x_2^2).$$

Нетрудно убедиться, что точка $\widehat{x} = (0, 0)$ является точкой минимума в рассматриваемой задаче. В силу того, что

$$\mathcal{L}_x(0, \lambda_0, \lambda_1) = (\lambda_0, 0) = 0,$$

λ_0 не может быть отличным от нуля.

12. Условия экстремума второго порядка для гладких задач с ограничениями типа равенств

Теорема 15 (Необходимые условия экстремума второго порядка). *Если $\widehat{x}$ — локальный минимум (максимум) в задаче* (23), *функция f и отображение F дважды дифференцируемы в $\widehat{x}$ и $\operatorname{Im} F'(\widehat{x}) = Y$, то найдется множитель Лагранжа $y^* \in Y^*$ такой, что*

$$\mathcal{L}_x(\widehat{x}, 1, y^*) = 0 \quad \Leftrightarrow \quad f'(\widehat{x}) + (F'(\widehat{x}))^* y^* = 0 \tag{26}$$

и для всех $h \in \operatorname{Ker} F'(\widehat{x})$

$$\mathcal{L}_{xx}(\widehat{x}, 1, y^*)[h, h] \geq 0 \ (\leq 0) \quad \Leftrightarrow \quad f''(\widehat{x})[h, h] + \langle y^*, F''(\widehat{x})[h, h] \rangle \geq 0 \ (\leq 0) \tag{27}$$

Доказательство. Из дважды дифференцируемости отображения F в точке $\widehat{x}$ следует непрерывность F' в точке $\widehat{x}$. В силу предложения 1 отображение F является строго дифференцируемым. Поэтому соотношение (26) сразу следует из (24).

Докажем (27). Пусть $h \in \operatorname{Ker} F'(\widehat{x})$. Отображение F удовлетворяет условиям теоремы Люстерника (теорема 10) и поэтому $h \in T_{\widehat{x}} M$, т. е. существуют $\varepsilon > 0$ и отображение $r\colon (-\varepsilon, \varepsilon) \to X$ такие, что $F(\widehat{x} + th + r(t)) = 0$ для $t \in (-\varepsilon, \varepsilon)$ и $r(t) = o(t)$ при $t \to 0$. Если $\widehat{x}$ — локальный минимум, то $f(\widehat{x} + th + r(t)) \geq f(\widehat{x})$ для достаточно малых t. Теперь по формуле Тейлора имеем (учитывая (26)

$$0 \le f(\widehat{x}+th+r(t)) - f(\widehat{x}) = \mathcal{L}(\widehat{x}+th+r(t),1,y^*) - \mathcal{L}(\widehat{x},1,y^*)$$
$$= \frac{1}{2}\mathcal{L}_{xx}(\widehat{x},1,y^*)[th+r(t),th+r(t)] + o(t^2)$$
$$= \frac{t^2}{2}\mathcal{L}_{xx}(\widehat{x},1,y^*)[h,h] + o(t^2),$$

откуда следует (27). □

Теорема 16 (Достаточные условия экстремума второго порядка в задаче с ограничениями типа равенств)**.** *Пусть в задаче* (23) *функция f и отображение F дважды дифференцируемы в допустимой точке $\widehat{x}$ и $\operatorname{Im} F'(\widehat{x})$ — замкнутое подпространство в Y. Тогда если найдутся множитель Лагранжа $y^* \in Y^*$ и число $\alpha > 0$ такие, что*

$$\mathcal{L}_x(\widehat{x},1,y^*) = 0 \tag{28}$$

и при всех $h \in \operatorname{Ker} F'(\widehat{x})$

$$\mathcal{L}_{xx}(\widehat{x},1,y^*)[h,h] \ge \alpha\|h\|_X^2 \quad (\le \alpha\|h\|_X^2), \tag{29}$$

то $\widehat{x}$ — локальный минимум (максимум) в задаче (23).

Доказательство. Рассмотрим отображение $G\colon X \to \operatorname{Im} F'(\widehat{x})$, определенное по формуле $G(x) = F'(\widehat{x})x$. Для него, очевидно, выполнены условия следствия 4, согласно которому существуют окрестность U_0 точки ноль, отображение $\varphi\colon U_0 \to X$ и константа $K > 0$ такие, что

$$F'(\widehat{x})(\varphi(x)) = 0 \tag{30}$$

и

$$\|\varphi(x) - x\|_X \le K\|F'(\widehat{x})x\|_Y \tag{31}$$

для всех $x \in U_0$.

Можно считать, что в любой окрестности $\widehat{x}$ есть допустимые в (23) точки (т.к. изолированная точка автоматически является и локальным минимумом, и локальным максимумом). Пусть $x \in U_0$ и $\widehat{x}+x$ — допустимый элемент в задаче (23). Тогда по формуле Тейлора

$$0 = F(\widehat{x}+x) = F'(\widehat{x})x + \frac{1}{2}F''(\widehat{x})[x,x] + o(\|x\|_X^2).$$

отсюда следует, что для достаточно малых x справедливо неравенство

$$\|F'(\widehat{x})x\|_Y \le \left(\frac{1}{2}\|F''(\widehat{x})\| + 1\right)\|x\|_X^2,$$

а тогда из (31) получаем, что

$$\|\varphi(x) - x\|_X \le \gamma\|x\|_X^2,$$

где $\gamma = K(\|F''(\widehat{x})\|/2 + 1)$. Следовательно,

$$\|\varphi(x)\|_X \le \|x\| + \gamma\|x\|_X^2 = (1 + \gamma\|x\|_X)\|x\|_X.$$

Считая, что $\|x\|_X < 1/\gamma$, имеем также оценку

$$\|\varphi(x)\|_X \ge \|x\|_X - \|\varphi(x) - x\|_X \ge (1 - \gamma\|x\|_X)\|x\|_X.$$

Обозначая, для краткости, $L(x) = \mathcal{L}(x, 1, y^*)$, снова по формуле Тейлора получаем (учитывая (28) и то, что $\widehat{x} + x$ — допустимая точка)

$$f(\widehat{x} + x) = f(\widehat{x}) + \frac{1}{2}L''(\widehat{x})[x, x] + o(\|x\|_X^2).$$

Отсюда, полагая $B = \|L''(\widehat{x})\|$, учитывая, что $\varphi(x) \in \operatorname{Ker} F'(\widehat{x})$ согласно (30), полученные выше оценки, и считая, что выполнено первое из неравенств (29), будем иметь

$$\begin{aligned}
f(\widehat{x}+x)-f(\widehat{x}) &= \frac{1}{2}L''(\widehat{x})[\varphi(x)-\varphi(x)+x, \varphi(x)-\varphi(x)+x]+o(\|x\|_X^2)\\
&= \frac{1}{2}(L''(\widehat{x})[\varphi(x), \varphi(x)] - 2L''(\widehat{x})[\varphi(x) - x, \varphi(x)]\\
&+ L''(\widehat{x})[\varphi(x) - x, \varphi(x) - x]) + o(\|x\|_X^2) \ge \frac{1}{2}(\alpha\|\varphi(x)\|_X^2\\
&- 2B\|\varphi(x)\|_X\|\varphi(x) - x\|_X - B\|\varphi(x) - x\|_X^2) + o(\|x\|_X^2)\\
&\ge \frac{1}{2}\|x\|_X^2(\alpha(1 - \gamma\|x\|_X)^2 - 2B\gamma\|x\|_X(1 + \gamma\|x\|_X)\\
&\qquad - B\gamma^2\|x\|_X^2) + o(\|x\|_X^2).
\end{aligned}$$

Нетрудно убедиться, что выражение справа неотрицательно для достаточно малых x и поэтому $\widehat{x}$ — локальный минимум. Случай, когда выполнено второе из неравенств (29) исследуется аналогично. □

13. Гладкие задачи с ограничениями типа равенств и неравенств

Пусть U — открытое подмножество банахова пространства X, $f_j : U \to \mathbb{R}$, $j = 0, 1, \dots, m$ и $F\colon U \to Y$, где Y — банахово пространство. Задачу

$$(32) \qquad f_0(x) \to \min, \quad f_j(x) \le 0, \quad 1 \le j \le m, \quad F(x) = 0,$$

называют *задачей с ограничениями типа равенств и неравенств*. Если функции f_j, $j = 0, 1, \dots, m$, и отображение F обладают некоторой гладкостью, то говорят о *гладкой задаче с ограничениями типа равенств и неравенств*.

Сопоставим задаче (32) *функцию Лагранжа*

$$\mathcal{L}(x, \overline{\lambda}, y^*) = \sum_{j=0}^{m} \lambda_j f_j(x) + \langle y^*, F(x)\rangle,$$

где $\overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in (\mathbb{R}^{m+1})^*$ и $y^* \in Y^*$. Числа λ_j, $j = 0, 1, \ldots, m$, и функционал y^* называются *множителями Лагранжа.*

Теорема 17 (Правило множителей Лагранжа в задаче с ограничениями типа равенств и неравенств)**.** *Если $\widehat{x}$ — локальный минимум в задаче* (32)*, функции f_j, $j = 0, 1, \ldots, m$, дифференцируемы в $\widehat{x}$, отображение F строго дифференцируемо в $\widehat{x}$ и $\operatorname{Im} F'(\widehat{x})$ — замкнутое подпространство в Y, то найдутся такие множители Лагранжа $\lambda_0, \lambda_1, \ldots, \lambda_m$ и y^*, не равные нулю одновременно, для которых выполнены условия*

- (a) $\mathcal{L}_x(\widehat{x}, \overline{\lambda}, y^*) = 0 \Leftrightarrow \sum_{j=0}^{m} \lambda_j f_j'(\widehat{x}) + (F'(\widehat{x}))^* y^* = 0$ *(условие стационарности);*
- (b) $\lambda_j \geq 0$, $j = 0, 1, \ldots, m$ *(условие неотрицательности);*
- (c) $\lambda_j f_j(\widehat{x}) = 0$, $j = 1, \ldots, m$ *(условие дополняющей нежесткости).*

Если $\operatorname{Im} F'(\widehat{x}) = Y$ и существует вектор $h \in \operatorname{Ker} F'(\widehat{x})$ такой, что $\langle f_j'(\widehat{x}), h \rangle < 0$ для всех $j \in J_0 = \{j : f_j(\widehat{x}) = 0, \ 1 \leq j \leq m\}$, то $\lambda_0 \neq 0$.

Доказательство. Заметим сначала, что утверждение (c) можно считать выполненным всегда. В самом деле, отбросим те ограничения среди неравенств, для которых $f_j(\widehat{x}) < 0$. Тогда $\widehat{x}$ будет локальным экстремумом и в новой задаче. Если для этой задачи доказаны утверждения (a) и (b), то (c) выполняется автоматически. Дополнив найденный набор множителей Лагранжа нулевыми компонентами, соответствующими тем номерам, где $f_j(\widehat{x}) < 0$, получим утверждения (a), (b) и (c) для исходной задачи.

Как и в доказательстве правила множителей Лагранжа для задачи с ограничениями типа равенств рассмотрим отдельно два случая.

A) Вырожденный случай: $\operatorname{Im} F'(\widehat{x}) \neq Y$. Здесь, фактически, повторяется доказательство правила множителей Лагранжа для вырожденного случая в гладкой задаче с равенствами. В силу того, что подпространство $\operatorname{Im} F'(\widehat{x})$ замкнуто, по лемме о нетривиальности аннулятора (лемма 3) существует ненулевой функционал $y^* \in Y^*$ такой, что $\langle y^*, F'(\widehat{x})x \rangle = 0$ для любого $x \in X$, т. е. $(F'(\widehat{x}))^* y^* = 0$. Остается положить $\lambda_j = 0$, $j = 0, 1, \ldots, m$.

B) Невырожденный случай: $\operatorname{Im} F'(\widehat{x}) = Y$. Рассмотрим множество

$$C = \{ (\mu_0, \mu_1, \ldots, \mu_m, y) \in \mathbb{R}^{m+1} \times Y : \exists x \in X : \mu_j > \langle f_j'(\widehat{x}), x \rangle, \\ j = 0, 1, \ldots, m, \ y = F'(\widehat{x})x \}.$$

Очевидно, что C — выпуклое множество. Докажем, что $0 \notin C$. Предположим, что $0 \in C$. Тогда существует такое $x_0 \in X$, что

$\langle f_j'(\widehat{x}), x_0\rangle < 0$, $j = 0, 1, \dots, m$, и $F'(\widehat{x})x_0 = 0$. По теореме Люстерника $x_0 \in T_{\widehat{x}}M$, где $M = \{\, x \in X : F(x) = F(\widehat{x}) = 0 \,\}$, т. е. существуют $\varepsilon > 0$ и отображение $r\colon (-\varepsilon, \varepsilon) \to X$ такие, что $F(\widehat{x}+tx_0+r(t)) = 0$ для всех $t \in (-\varepsilon, \varepsilon)$ и $\|r(t)\|_X = o(t)$ при $t \to 0$. В силу дифференцируемости функций f_j, $j = 0, 1, \dots, m$, в точке $\widehat{x}$ имеем

$$f_j(\widehat{x} + tx_0 + r(t)) = f_j(\widehat{x}) + \langle f_j'(\widehat{x}), x_0\rangle t + o(t) < f_j(\widehat{x})$$

для достаточно малых $t > 0$. Это значит, что для таких t точки $\widehat{x} + tx_0 + r(t)$ допустимы в задаче (32), а значение функционала f_0 на них меньше, чем $f_0(\widehat{x})$, в противоречие с тем, что $\widehat{x}$ — локальный минимум. Итак, $0 \notin C$.

Покажем теперь, что $\operatorname{int} C \neq \emptyset$. Рассмотрим множество

$$\begin{aligned} C_0 = \{\, (\mu_0, \mu_1, \dots, \mu_m, y) \in \mathbb{R}^{m+1} \times Y : \mu_j > d,\ j = 0, 1, \dots, m,& \\ y \in F'(\widehat{x})(B_X(0,1)) \,\}, \quad d = \max_{0\le j\le m} \|f_j'(\widehat{x})\|.& \end{aligned}$$

По теореме Банаха об открытом отображении [2, стр. 243] множество $F'(\widehat{x})(B_X(0,1))$ открыто. Тем самым C_0 — открытое множество. Покажем, что $C_0 \subset C$. Действительно, пусть $(\mu_0, \mu_1, \dots, \mu_m, y) \in C_0$ и $x \in B_X(0,1)$ такое, что $y = F'(\widehat{x})x$. Тогда $\mu_j > d \ge \langle f_j'(\widehat{x}), x\rangle$, $j = 0, \dots, m$, и значит, $(\mu_0, \mu_1, \dots, \mu_m, y) \in C$, т. е. $\operatorname{int} C \neq \emptyset$.

В силу первой теоремы отделимости (теорема 11) множество C можно отделить от нуля, т.е. найдется ненулевой функционал $(\lambda_0, \lambda_1, \dots, \lambda_m, y^*) \in (\mathbb{R}^{m+1})^* \times Y^*$ такой, что

$$\sum_{j=0}^{m} \lambda_j \mu_j + \langle y^*, y\rangle \ge 0 \tag{33}$$

для всех $(\mu_0, \mu_1, \dots, \mu_m, y) \in C$. Наборы $(\mu_0, \mu_1, \dots, \mu_m, 0)$, где $\mu_j > 0$, $j = 0, 1 \dots, m$, принадлежат C (надо взять $x = 0$). Подставляя их в (33), получаем, что

$$\sum_{j=0}^{m} \lambda_j \mu_j \ge 0$$

для всех $\mu_j > 0$, $j = 0, 1 \dots, m$. Отсюда вытекают неравенства $\lambda_j \ge 0$, $j = 0, 1 \dots, m$, и утверждение (b) теоремы доказано.

Для любого $x \in X$ и любого $\varepsilon > 0$

$$(\langle f_0'(\widehat{x}), x\rangle + \varepsilon, \dots, \langle f_m'(\widehat{x}), x\rangle + \varepsilon, F'(\widehat{x})x) \in C.$$

Из (33), получаем, что

$$\sum_{j=0}^{m} \lambda_j \langle f_j(\widehat{x}), x\rangle + \langle y^*, F'(\widehat{x})x\rangle \ge -\varepsilon \sum_{j=0}^{m} \lambda_j.$$

В силу произвольности ε левая часть этого неравенства (которая есть линейный функционал) неотрицательна на X и значит, она равна нулю, а это равносильно утверждению (a) теоремы.

Докажем последнее утверждение теоремы. Пусть выполнены его предположения и $\lambda_0 = 0$. Если при некотором $1 \le j \le m$, $f_j(\widehat{x}) \neq 0$, то из c) следует, что $\lambda_j = 0$. Тем самым из (a) вытекает, что

$$\sum_{j \in J_0} \lambda_j \langle f_j'(\widehat{x}), h \rangle = 0.$$

В силу b) и того, что $\langle f_j'(\widehat{x}), h \rangle < 0$, $j \in J_0$, получаем, что $\lambda_j = 0$, $j = 0, 1, \ldots, m$. Следовательно, из a) вытекает, что $\langle y^*, F'(\widehat{x})x \rangle = 0$ для всех $x \in X$. Так как $\operatorname{Im} F'(\widehat{x}) = Y$, то $\langle y^*, y \rangle = 0$ при всех $y \in Y$, т.е. $y^* = 0$. Это противоречит тому, что не все множители Лагранжа равны нулю. $\square$

14. Выпуклые задачи без ограничений. Субдифференциал. Теорема Ферма

Пусть X — вещественное линейное пространство и $f \colon X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$. Множества

$$\operatorname{dom} f = \{\, x \in X : f(x) < +\infty \,\},$$
$$\operatorname{epi} f = \{\, (x, \alpha) \in X \times \mathbb{R} : \alpha \ge f(x),\ x \in \operatorname{dom} f \,\}$$

называются соответственно *эффективным множеством* и *надграфиком* (или *эпиграфом*) *функции* f. Функцию f называют *собственной*, если $\operatorname{dom} f \neq \emptyset$.

Для элементов расширенной прямой считается, что $a + (+\infty) = +\infty$ для всех $a \in \mathbb{R}$, $a \cdot (+\infty) = +\infty$, если $a > 0$, $0 \cdot (+\infty) = 0$ и $+\infty + (+\infty) = +\infty$.

Функция $f \colon X \to \overline{\mathbb{R}}$ называется *выпуклой*, если ее надграфик выпуклое множество в $X \times \mathbb{R}$. Нетрудно проверить, что функция f выпукла тогда и только тогда, когда для любых $x_1, x_2 \in X$ и любого $0 \le \alpha \le 1$ выполняется неравенство

$$f((1 - \alpha)x_1 + \alpha x_2) \le (1 - \alpha)f(x_1) + \alpha f(x_2),$$

которое называется *неравенством Йенссена.*

Теорема 18. *Пусть X — линейное нормированное пространство и $f \colon X \to \mathbb{R}$ — дважды дифференцируема на X. Тогда f — выпуклая функция тогда и только тогда, когда $f''(x)[h, h] \ge 0$ для всех $x \in X$ и всех $h \in X$.*

Доказательство. Пусть f — выпуклая функция. Предположим, что существует $x \in X$ и $h \in X$ такие, что $f''(x)[h, h] < 0$. По формуле Тейлора для $t \in \mathbb{R}$ имеем

$$f(x+th) = f(x) + f'(x)ht + f''(x)[h,h]\frac{t^2}{2} + o(t^2),$$
$$f(x-th) = f(x) - f'(x)ht + f''(x)[h,h]\frac{t^2}{2} + o(t^2).$$

Отсюда, складывая эти равенства, получаем

$$f(x+th) - 2f(x) + f(x-th) = f''(x)[h,h]t^2 + o(t^2).$$

Следовательно, при достаточно малых t

$$f(x+th) - 2f(x) + f(x-th) < 0,$$

что противоречит выпуклости f.

Пусть теперь $f''(x)[h,h] \geq 0$ для всех $x \in X$ и всех $h \in X$. Для произвольных $x_1, x_2 \in X$ рассмотрим функцию

$$F(t) = f(x_1 + t(x_2 - x_1)) - f(x_1) - t(f(x_2) - f(x_1)).$$

Имеем $F(0) = F(1) = 0$,

$$F''(t) = f''(x_1 + t(x_2 - x_1))[x_2 - x_1, x_2 - x_1] \geq 0.$$

Предположим, что при некотором $t \in (0,1)$ $F(t) > 0$. Тогда найдется точка $t_0 \in (0,1)$, в которой функция F будет достигать максимального значения и, значит, $F'(t_0) = 0$. Поскольку $F''(t) \geq 0$ при всех $t \in [t_0, 1]$, то $F'(t) \geq 0$ при $t \in [t_0, 1]$. Тем самым функция F не убывает на отрезке $[t_0, 1]$, а значит, $F(1) > 0$. Полученное противоречие доказывает, что $F(t) \leq 0$ для всех $t \in [0,1]$. Таким образом, для всех $t \in [0,1]$ справедливо неравенство

$$f((1-t)x_1 + tx_2) \leq (1-t)f(x_1) + tf(x_2),$$

что и означает выпуклость функции f. □

Пусть $X = \mathbb{R}^d$ и $f\colon \mathbb{R}^d \to \mathbb{R}$. Если f — дважды дифференцируема, то $f''(x)$ — гессиан f в точке x (см. (8)).

Следствие 5. *Если функция $f\colon \mathbb{R}^d \to \mathbb{R}$ дважды дифференцируема на $\mathbb{R}^d$, то она является выпуклой в том, и только в том случае, если ее гессиан в любой точке $x \in X$ удовлетворяет условию*

$$h^T f''(x) h \geq 0$$

для всех $h \in X$ (матрица гессиана в любой точке является неотрицательно определенной).

Пусть X — линейное нормированное пространство, $f\colon X \to \overline{\mathbb{R}}$, $\widehat{x} \in X$ и функция f конечна в точке $\widehat{x}$. *Субдифференциалом функции f в точке $\widehat{x}$* называется множество (возможно пустое)

$$\partial f(\widehat{x}) = \{\, x^* \in X^* : f(x) - f(\widehat{x}) \geq \langle x^*, x - \widehat{x}\rangle,\ \forall\, x \in X \,\}.$$

Следующее предложение показывает, что субдифференциал достаточно естественное обобщение понятия производной на выпуклые функции.

Предложение 3. *Пусть X — линейное нормированное пространство и $f\colon X \to \overline{\mathbb{R}}$ — выпуклая функция, дифференцируемая в точке $\widehat{x}$. Тогда $\partial f(\widehat{x}) = \{f'(\widehat{x})\}$.*

Доказательство. Пусть $x \in X$. Для любого $0 < \alpha < 1$ имеем по неравенству Йенссена

$$f((1-\alpha)\widehat{x} + \alpha x) \leq (1-\alpha)f(\widehat{x}) + \alpha f(x),$$

откуда

$$f(\widehat{x} + \alpha(x - \widehat{x})) - f(\widehat{x}) \leq \alpha(f(x) - f(\widehat{x})).$$

В силу дифференцируемости функции f в точке $\widehat{x}$ имеем

$$\alpha\langle f'(\widehat{x}), x - \widehat{x}\rangle + o(\alpha) \leq \alpha(f(x) - f(\widehat{x})).$$

Сокращая на α и переходя к пределу при $\alpha \to 0$, получаем, что $f'(\widehat{x}) \in \partial f(\widehat{x})$.

Обратно, если $x^* \in \partial f(\widehat{x})$, то для любого $x \in X$ и любого $t > 0$ имеем $f(\widehat{x} + tx) - f(\widehat{x}) \geq t\langle x^*, x\rangle$. Следовательно,

$$t\langle f'(\widehat{x}), x\rangle + o(t) \geq t\langle x^*, x\rangle,$$

т. е. $\langle f'(\widehat{x}), x\rangle \geq \langle x^*, x\rangle$ для любого x и значит, $x^* = f'(\widehat{x})$. □

Пусть X — линейное нормированное пространство и $f\colon X \to \overline{\mathbb{R}}$ — собственная функция. Рассмотрим задачу

$$f(x) \to \min, \quad x \in X. \tag{34}$$

Теорема 19 (Ферма в субдифференциальной форме)**.** *Точка $\widehat{x}$ является глобальным минимумом в задаче* (34) *тогда и только тогда, когда $0 \in \partial f(\widehat{x})$.*

Доказательство. Если $\widehat{x}$ — глобальный минимум, то $f(x) - f(\widehat{x}) \geq 0 = \langle 0, x - \widehat{x}\rangle$ для любого $x \in X$, т. е. $0 \in \partial f(\widehat{x})$. Если $0 \in \partial f(\widehat{x})$, то $f(x) - f(\widehat{x}) \geq \langle 0, x - \widehat{x}\rangle = 0$, т. е. $f(x) \geq f(\widehat{x})$ для любого $x \in X$. □

Если в задаче (34) функция f — выпуклая, то она называется *выпуклой задачей без ограничений.* Отметим, что в этом случае нет смысла говорить о локальных минимумах, поскольку любой локальный минимум является и глобальным. Действительно, пусть $\widehat{x}$ — локальный минимум, т. е. существует такая окрестность U точки $\widehat{x}$, что $f(\widehat{x}) \leq f(x)$ для всех $x \in U$. Пусть теперь x — произвольная точка из X. Для достаточно малых $0 < \alpha \leq 1$ точки $(1-\alpha)\widehat{x} + \alpha x$ принадлежат U и поэтому (по неравенству Йенссена) $f(\widehat{x}) \leq f((1-\alpha)\widehat{x} + \alpha x) \leq (1-\alpha)f(\widehat{x}) + \alpha f(x)$, откуда следует, что $f(\widehat{x}) \leq f(x)$.

Из предложения 3 и теоремы 19 вытекает

Следствие 6. *Если в задаче* (34) f — *выпуклая функция, дифференцируемая в точке* $\widehat{x}$, *то* $\widehat{x}$ — *глобальный минимум в том и только в том случае, если* $f'(\widehat{x}) = 0$.

15. Выпуклые задачи с ограничениями. Теорема Каруша–Куна–Таккера

Пусть X — вещественное линейное пространство, $f_j\colon X \to \mathbb{R}$, $j = 0, 1, \dots, m$, — выпуклые функции и A — непустое выпуклое подмножество X. Задачу

$$(35) \qquad f_0(x) \to \min, \quad f_j(x) \le 0, \quad j = 1, \dots, m, \quad x \in A$$

называют *выпуклой задачей* или *задачей выпуклого программирования.*

Свяжем с задачей (35) следующую функцию Лагранжа

$$\mathcal{L}(x, \overline{\lambda}) = \sum_{j=0}^{m} \lambda_j f_j(x),$$

где $\overline{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_m)$ — набор множителей Лагранжа.

Теорема 20 (Каруша–Куна–Таккера)**.** *Если* $\widehat{x}$ — *минимум в задаче* (35), *то найдется такой ненулевой набор множителей Лагранжа* $\overline{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_m)$, *что выполнены следующие условия*

- (a) $\min_{x\in A} \mathcal{L}(x, \overline{\lambda}) = \mathcal{L}(\widehat{x}, \overline{\lambda})$ (*условие минимума*)*;*
- (b) $\lambda_j \ge 0$, $j = 0, 1, \dots, m$ (*условие неотрицательности*)*;*
- (c) $\lambda_j f_j(\widehat{x}) = 0$, $j = 1, \dots, m$ (*условие дополняющей нежесткости*)*.*

Если существует допустимая в (35) *точка* $\widehat{x}$ *и набор множителей Лагранжа* $\overline{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_m)$, *удовлетворяющие условиям* (a), (b) *и* (c) *и при этом* $\lambda_0 > 0$, *то* $\widehat{x}$ — *решение задачи* (35).

Если найдется точка $\overline{x} \in A$ *такая, что* $f_j(\overline{x}) < 0$, $1 \le j \le m$, *то* $\lambda_0 \ne 0$ *(условие Слейтера).*

Доказательство. Пусть $\widehat{x}$ — решение задачи (35). Рассмотрим множество

$$M = \{\, \mu = (\mu_0, \mu_1, \dots, \mu_m)^T \in \mathbb{R}^{m+1} : \exists\, x \in A : f_0(x) - f_0(\widehat{x}) < \mu_0, \\ f_j(x) \le \mu_j,\ j = 1, \dots, m \,\}.$$

Непосредственная проверка показывает, что это множество выпукло. Кроме того, легко видеть, что оно содержит все векторы с положительными компонентами (надо взять $x = \widehat{x}$) и тем самым его внутренность не пуста. Наконец, $0 \notin M$, так как в противном случае нашелся бы элемент $\overline{x} \in A$ такой, что $f_j(\overline{x}) \le 0$, $j = 1, \dots, m$, и $f_0(\overline{x}) - f_0(\widehat{x}) < 0$, в противоречие с тем, что $\widehat{x}$ — минимум.

Согласно первой теореме отделимости найдется такой ненулевой функционал, т. е. вектор $\overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in (\mathbb{R}^{m+1})^*$, что

$$\sum_{j=0}^{m} \lambda_j \mu_j \geq 0 \tag{36}$$

для всех $\mu = (\mu_0, \mu_1, \ldots, \mu_m)^T \in M$. Пусть $\delta > 0$. Подставляя в (36) векторы $(1, \delta, \ldots, \delta)^T, \ldots, (\delta, \ldots, \delta, 1)^T$, а затем устремляя δ к нулю, получаем, что $\lambda_j \geq 0$, $j = 0, 1, \ldots, m$, т. е. доказано утверждение (b) теоремы.

Теперь подставим в (36) векторы $(\delta, \ldots, \delta, f_j(\widehat{x}), \delta, \ldots, \delta)^T$, $j = 1, \ldots, m$ (они принадлежат M, надо взять $x = \widehat{x}$) и снова, устремляя δ к нулю, получим, что $\lambda_j f_j(\widehat{x}) \geq 0$. Но $\lambda_j f_j(\widehat{x}) \leq 0$, так как $\lambda_j \geq 0$, а $f_j(\widehat{x}) \leq 0$ и поэтому $\lambda_j f_j(\widehat{x}) = 0$, $j = 1, \ldots, m$, что доказывает утверждение (c).

Пусть $x \in A$. Ясно, что $(f_0(x) - f_0(\widehat{x}) + \delta, f_1(x), \ldots, f_m(x))^T \in M$. Подставляя этот вектор в (36), приходим (в пределе при $\delta \to 0$) к неравенству $\sum_{j=0}^{m} \lambda_j f_j(x) \geq \lambda_0 f_0(\widehat{x})$. Добавляя справа нулевые слагаемые $\lambda_j f_j(\widehat{x})$, $j = 1, \ldots, m$, получаем, что $\mathcal{L}(x, \overline{\lambda}) \geq \mathcal{L}(\widehat{x}, \overline{\lambda})$ и (a) доказано.

Докажем второе утверждение теоремы. Пусть x — допустимый элемент в задаче (36). Тогда, используя это обстоятельство вместе с (b), (a) и (c), будем иметь

$$\lambda_0 f_0(x) \geq \lambda_0 f_0(x) + \sum_{j=1}^{m} \lambda_j f_j(x) = \mathcal{L}(x, \overline{\lambda}) \geq \mathcal{L}(\widehat{x}, \overline{\lambda})$$

$$= \lambda_0 f_0(x) + \sum_{j=1}^{m} \lambda_j f_j(\widehat{x}) = \lambda_0 f_0(\widehat{x}).$$

Деля на λ_0, получаем требуемое.

Докажем последнее утверждение теоремы. Если $\lambda_0 = 0$, то ненулевой множитель Лагранжа находится среди остальных и поэтому (с учетом (c)) $\mathcal{L}(\overline{x}, \overline{\lambda}) = \sum_{i=1}^{m} \lambda_i f_i(\overline{x}) < 0 = \sum_{i=1}^{m} \lambda_i f_i(\widehat{x}) = \mathcal{L}(\widehat{x}, \overline{\lambda})$, что противоречит (a). $\square$

16. Простейшая задача вариационного исчисления. Уравнение Эйлера

Принято считать, что вариационное исчисление родилось с задачи о брахистохроне, предложенной в 1696 г. И. Бернулли для решения своим современникам. Задача была решена самим Бернулли, его братом Яковом, Ньютоном, Лейбницем и Лопиталем. Решения были разные, и вскоре еще было решено несколько сходных задач.

В начале 18 века И. Бернулли предложил Л. Эйлеру (тогда молодому человеку, которого он консультировал по научным вопросам) найти общие методы решения подобных задач. Начиная с 1732 г.

Л. Эйлер начал активно этим заниматься и через 12 лет завершил свой фундаментальный труд “Modus invineniondi lineas curvas maximive proprietate gemdenties sive soluto problematis isoperimetrice latissimo sensu accepti” (“Метод нахождения кривых линий, обладающих свойствами максимума или минимума, или решение изопериметрической задачи, взятой в самом широком смысле”), Лозанна, 1744 г. Там, в частности, была рассмотрена задача, которая ныне называется простейшей задачей (классического) вариационного исчисления.

Пусть $[t_0, t_1]$ — отрезок числовой прямой, G — открытое подмножество $\mathbb{R}^3$, $L\colon G \to \mathbb{R}$ — функция переменных t, x, $\dot{x}$ и $x_0, x_1 \in \mathbb{R}$. Задача

$$(37)\quad J(x) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt \to \text{extr}, \quad x(t_0) = x_0, \quad x(t_1) = x_1,$$

называется *простейшей задачей* (*классического*) *вариационного исчисления.* Функцию L называют *интегрантом* или *лагранжианом* задачи.

Уточним постановку. Обозначим через $C([t_0, t_1])$ и $C^1([t_0, t_1])$ множества всех непрерывных и непрерывно дифференцируемых функций x на $[t_0, t_1]$. Это нормированные пространства соответственно с нормами

$$\|x\|_{C([t_0,t_1])} = \max_{t \in [t_0,t_1]} |x(t)|,$$
$$\|x\|_{C^1([t_0,t_1])} = \max(\|x\|_{C([t_0,t_1])}, \|\dot{x}\|_{C([t_0,t_1])}).$$

Функция $x \in C^1([t_0, t_1])$ называется *допустимой в задаче* (37), если $\Gamma(x) = \{(t, x(t), \dot{x}(t)) : t \in [t_0, t_1]\} \subset G$ и $x(t_j) = x_j$, $j = 0, 1$.

Допустимая функция $\widehat{x}$ называется *слабым локальным минимумом* (*максимумом*) в задаче (37), если существует такое $\varepsilon > 0$, что для любой допустимой функции x, для которой $\|x - \widehat{x}\|_{C^1([t_0,t_1])} < \varepsilon$ выполняется неравенство $J(x) \geq J(\widehat{x})$ ($J(x) \leq J(\widehat{x})$). *Слабый локальный экстремум* — это либо слабый локальный минимум, либо слабый локальный максимум.

Далее, если фиксирована функция $\widehat{x}$, то для сокращения записи используем обозначения: $\widehat{L}_x(t) = L_x(t, \widehat{x}(t), \dot{\widehat{x}}(t))$ и аналогично для частной производной L по $\dot{x}$.

Теорема 21 (Необходимые условия экстремума в задаче (37))**.** *Пусть $\widehat{x}$ доставляет слабый локальный экстремум в задаче* (37). *Тогда, если функция L непрерывна вместе со своими частными производными по x и $\dot{x}$ в окрестности $\Gamma(\widehat{x})$, то $\widehat{L}_{\dot{x}} \in C^1([t_0, t_1])$ и для всех $t \in [t_0, t_1]$ выполнено уравнение Эйлера*

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

Доказательство. Пусть $x \in C^1([t_0, t_1])$ и $x(t_0) = x(t_1) = 0$. Положим $x_\alpha = \widehat{x} + \alpha x$, где $\alpha \in \mathbb{R}$. Для достаточно малых α функция x_α принадлежит окрестности $\Gamma(\widehat{x})$. Кроме того, очевидно, что $x_\alpha \in C^1([t_0, t_1])$ и $x_\alpha(t_j) = x_j$, $j = 0, 1$. Тем самым для достаточно малых α функции x_α допустимы в задаче (37). Функция $\widehat{J}(\alpha) = J(x_\alpha)$ имеет в нуле локальный экстремум. В силу теоремы о производной суперпозиции функций (теорема 7) и дифференцировании оператора Немыцкого (следствие 3) эта функция дифференцируема в нуле и тогда по теореме Ферма ее производная в нуле равна нулю. Следовательно,

$$\frac{d\widehat{J}(\alpha)}{d\alpha}\bigg|_{\alpha=0} = \int_{t_0}^{t_1} \left(\widehat{L}_x(t)x(t) + \widehat{L}_{\dot{x}}(t)\dot{x}(t)\right) dt = 0. \tag{38}$$

Пусть p такая функция, что $\dot{p} = \widehat{L}_x$. Тогда, интегрирую по частям, получаем

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

Поэтому, из (38) следует, что для всех функций $x \in C^1([t_0, t_1])$, для которых $x(t_0) = x(t_1) = 0$ справедливо равенство

$$\int_{t_0}^{t_1} \left(-p(t) + \widehat{L}_{\dot{x}}(t)\right) \dot{x}(t)\, dt = 0.$$

Но тогда для любой константы $c \in \mathbb{R}$ справедливо и такое равенство

$$\int_{t_0}^{t_1} \left(-p(t) + \widehat{L}_{\dot{x}}(t) + c\right) \dot{x}(t)\, dt = 0. \tag{39}$$

Выберем c так, чтобы

$$\int_{t_0}^{t_1} (-p(t) + \widehat{L}_{\dot{x}}(t) + c)\, dt = 0.$$

Рассмотрим функцию

$$x(t) = \int_{t_0}^{t} (-p(\tau) + \widehat{L}_{\dot{x}}(\tau) + c)\, d\tau.$$

Ясно, что $x \in C^1([t_0, t_1])$ и $x(t_0) = x(t_1) = 0$. Подставим эту функцию в (39). Тогда получим

$$\int_{t_0}^{t_1} \left(-p(t) + \widehat{L}_{\dot{x}}(t) + c\right)^2 dt = 0.$$

Отсюда следует, что $-p + \widehat{L}_{\dot{x}} + c = 0$. Тем самым $\widehat{L}_{\dot{x}} \in C^1([t_0, t_1])$. Дифференцируя равенство $p = \widehat{L}_{\dot{x}} + c$ и учитывая, что $\dot{p} = \widehat{L}_x$, получаем уравнение Эйлера. $\square$

17. Задача Больца. Интегралы уравнения Эйлера

Пусть, как и в предыдущем случае, $[t_0, t_1]$ — отрезок числовой прямой, G — открытое подмножество $\mathbb{R}^3$, $L\colon G \to \mathbb{R}$ — непрерывная функция переменных t, x и $\dot{x}$. Пусть, кроме того, задана функция $l\colon W \to \mathbb{R}$, где W — открытое подмножество $\mathbb{R}^2$. Задача

$$(40)\qquad \mathcal{B}(x) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\,dt + l(x(t_0), x(t_1)) \to \mathrm{extr}$$

называется *задачей Больца.*

Функция $x \in C^1([t_0, t_1])$ называется *допустимой в задаче* (40), если $\Gamma(x) = \{(t, x(t), \dot{x}(t)) : t \in [t_0, t_1]\} \subset G$ и $(x(t_0), x(t_1)) \in W$.

Слабый локальный экстремум определяется аналогично предыдущему случаю.

Функции $\widehat{L}_x$ и $\widehat{L}_{\dot{x}}$ определяются как и раньше, и кроме того, для функции $l(\xi_0, \xi_1)$ полагаем $\widehat{l}_{\xi_j} = l_{\xi_j}(\widehat{x}(t_0), \widehat{x}(t_1))$, $j = 0, 1$.

Теорема 22 (Необходимые условия экстремума в задаче (40))**.** *Пусть $\widehat{x}$ доставляет слабый локальный экстремум в задаче* (40). *Тогда, если функция L непрерывна вместе со своими частными производными по x и $\dot{x}$ в окрестности $\Gamma(\widehat{x})$, а функция l непрерывна вместе со своими частными производными по ξ_0 и ξ_1 в окрестности точки $(\widehat{x}(t_0), \widehat{x}(t_1))$, то $\widehat{L}_{\dot{x}} \in C^1([t_0, t_1])$, для всех $t \in [t_0, t_1]$ выполнено уравнение Эйлера*

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

а, кроме того, выполняется условие трансверсальности

$$\widehat{L}_{\dot{x}}(t_j) = (-1)^j \widehat{l}_{\xi_j}, \quad j = 0, 1.$$

Доказательство. Пусть $x \in C^1([t_0, t_1])$ и $x_\alpha = \widehat{x} + \alpha x$. Рассуждая точно так же, как и в предыдущей теореме, приходим к соотношению

$$(41)\quad \left.\frac{d\mathcal{B}(x_\alpha)}{d\alpha}\right|_{\alpha=0} = \int_{t_0}^{t_1} \left(\widehat{L}_x(t)x(t) + \widehat{L}_{\dot{x}}(t)\dot{x}(t)\right)\,dt + \widehat{l}_{\xi_0}x(t_0) + \widehat{l}_{\xi_1}x(t_1) = 0.$$

Пусть p — решение задачи Коши $\dot{p} = \widehat{L}_x(t)$, $p(t_1) = -\widehat{l}_{\xi_1}$. Тем самым

$$p(t) = -\widehat{l}_{\xi_1} - \int_t^{t_1} \widehat{L}_x(\tau)\,d\tau, \quad t \in [t_0, t_1].$$

Подставляя $\dot{p}$ в (41) вместо $\widehat{L}_x$ и интегрируя по частям, получаем, что

$$(42)\qquad \int_{t_0}^{t_1} \left(-p(t) + \widehat{L}_{\dot{x}}(t)\right)\dot{x}(t)\,dt + (\widehat{l}_{\xi_0} - p(t_0))x(t_0) = 0.$$

Пусть теперь x — решение задачи Коши $\dot{x} = -p(t) + \widehat{L}_{\dot{x}}(t)$, $x(t_0) = \widehat{l}_{\xi_0} - p(t_0)$, т. е.

$$x(t) = \widehat{l}_{\xi_0} - p(t_0) + \int_{t_0}^{t} (-p(\tau) + \widehat{L}_{\dot{x}}(\tau))\, d\tau, \quad t \in [t_0, t_1].$$

Ясно, что $x \in C^1([t_0, t_1])$. Подставляя x в (42), приходим к равенству

$$\int_{t_0}^{t_1} \left(-p(t) + \widehat{L}_{\dot{x}}(t)\right)^2 dt + (\widehat{l}_{\xi_0} - p(t_0))^2 = 0,$$

откуда следует равенство $p = \widehat{L}_{\dot{x}}$, равносильное, в силу определения p, уравнению Эйлера, а также соотношение $p(t_0) = \widehat{l}_{\xi_0}$, или $\widehat{L}_{\dot{x}}(t_0) = \widehat{l}_{\xi_0}$. Условие $\widehat{L}_{\dot{x}}(t_1) = -\widehat{l}_{\xi_1}$ входит в определение p. $\square$

Мы рассмотрели "одномерные" варианты простейшей задачи и задачи Больца. Совершенно аналогично рассматриваются их векторные аналоги, когда $x = (x_1, \ldots, x_d)^T$. В этом случае роль пространств $C([t_0, t_1])$ и $C^1([t_0, t_1])$ играют пространства $C([t_0, t_1], \mathbb{R}^d)$ и $C^1([t_0, t_1], \mathbb{R}^d)$ — соответственно непрерывных и непрерывно дифференцируемых вектор-функций со значениями в $\mathbb{R}^d$. Они определяются аналогично одномерным вариантам, где $|x(t)| = \sqrt{x_1^2(t) + \ldots + x_d^2(t)}$. Необходимые условия экстремума здесь имеют тот же вид и их доказательства остаются прежними. Но формулы, разумеется, надо понимать векторно. Например,

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \ \Leftrightarrow\ -\frac{d}{dt}\widehat{L}_{\dot{x}_j}(t) + \widehat{L}_{x_j}(t) = 0, \quad j = 1, \ldots, d.$$

Напомним, что первым интегралом дифференциального уравнения называется функция, которая постоянна на решениях данного уравнения.

Если лагранжиан L не зависит от переменной x, то уравнение Эйлера имеет очевидный первый интеграл

$$p(t) = L_{\dot{x}}(t, \dot{\widehat{x}}(t)) = \text{const},$$

который называется *интегралом импульса.*

Если лагранжиан L не зависит от переменной t, то уравнение Эйлера имеет первый интеграл

$$H(t) = L_{\dot{x}}(\widehat{x}(t), \dot{\widehat{x}}(t))\dot{\widehat{x}}(t) - L(\widehat{x}(t), \dot{\widehat{x}}(t)) = \text{const}\,.$$

Он называется *интегралом энергии.*

Для доказательства вычислим производную функции H (учитывая, что $\widehat{x}(\cdot)$ удовлетворяет уравнению Эйлера)

$$\frac{dH}{dt} = L_{\dot{x}}(\widehat{x}(t), \dot{\widehat{x}}(t))\ddot{\widehat{x}}(t) + \frac{d}{dt}L_{\dot{x}}(\widehat{x}(t), \dot{\widehat{x}}(t))\dot{\widehat{x}}(t) - L_x(\widehat{x}(t), \dot{\widehat{x}}(t))\dot{\widehat{x}}(t)$$
$$- L_{\dot{x}}(\widehat{x}(t), \dot{\widehat{x}}(t))\ddot{\widehat{x}}(t) = \left(\frac{d}{dt}L_{\dot{x}}(\widehat{x}(t), \dot{\widehat{x}}(t)) - L_x(\widehat{x}(t), \dot{\widehat{x}}(t))\right)\dot{\widehat{x}}(t) = 0.$$

При доказательстве мы предположили существование $\dddot{\widehat{x}}$.

18. Задача Лагранжа. Общая постановка

Пусть $[t_0, t_1]$ — конечный отрезок, G — открытое подмножество $\mathbb{R} \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, W — открытое подмножество $\mathbb{R}^{d_1} \times \mathbb{R}^{d_1}$, функции $L_j\colon G \to \mathbb{R}$, $j = 0, 1, \dots, m$, и отображение $\varphi\colon G \to \mathbb{R}^{d_1}$ (переменных $t \in \mathbb{R}$, $x = (x_1, \dots, x_{d_1})^T \in \mathbb{R}^{d_1}$, $u = (u_1, \dots, u_{d_2})^T \in \mathbb{R}^{d_2}$) и функции $l_j\colon W \to \mathbb{R}$, $j = 0, 1, \dots, m$ (переменных ξ_0 и ξ_1) непрерывны на своей области определения. Задача

$$
\begin{gathered}
f_0(x,u) = \int_{t_0}^{t_1} L_0(t, x(t), u(t))\,dt + l_0(x(t_0), x(t_1)) \to \min, \qquad (43)\\
\dot{x} = \varphi(t, x, u),\\
f_j(x,u) = \int_{t_0}^{t_1} L_j(t, x(t), u(t))\,dt + l_j(x(t_0), x(t_1)) \le 0,\ 1 \le j \le m',\\
f_j(x,u) = \int_{t_0}^{t_1} L_j(t, x(t), u(t))\,dt + l_j(x(t_0), x(t_1)) = 0,\\
m'+1 \le j \le m,
\end{gathered}
$$

называется *задачей Лагранжа вариационного исчисления* (в понтрягинской форме).

Уточним постановку. Положим

$$Z = C^1([t_0, t_1], \mathbb{R}^{d_1}) \times C([t_0, t_1], \mathbb{R}^{d_2}).$$

Норму в Z введем следующим образом:

$$\|(x,u)\|_Z = \|x\|_{C^1([t_0,t_1],\mathbb{R}^{d_1})} + \|u\|_{C([t_0,t_1],\mathbb{R}^{d_2})}.$$

Пара $(x,u) \in Z$ называется *допустимой в задаче* (43), если

$$
\begin{gathered}
\Gamma(x,u) = \{\, (t, x(t), u(t)) : t \in [t_0, t_1] \,\} \subset G, \quad (x(t_0), x(t_1)) \in W,\\
\dot{x}(t) = \varphi(t, x(t), u(t)) \text{ для всех } t \in [t_0, t_1],\\
f_j(x,u) \le 0,\ 1 \le j \le m', \quad f_j(x,u) = 0,\ m'+1 \le j \le m.
\end{gathered}
$$

Допустимая пара $(\widehat{x}, \widehat{u})$ называется *слабым локальным минимумом в задаче* (43), если существует такое $\varepsilon > 0$, что для любой допустимой пары (x,u), для которой $\|(x,u) - (\widehat{x}, \widehat{u})\|_Z < \varepsilon$ выполнено неравенство $f_0(x,u) \ge f_0(\widehat{x}, \widehat{u})$.

Функцией Лагранжа задачи (43) назовем функцию

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

где

$$L(t,x,\dot{x},u)=\sum_{j=0}^{m}\lambda_j L_j(t,x,u)+p(t)(\dot{x}-\varphi(t,x,u)),$$

$$l(\xi_0,\xi_1)=\sum_{j=0}^{m}\lambda_j l_j(\xi_0,\xi_1)$$

и вектор $\overline{\lambda}=(\lambda_0,\lambda_1,\ldots,\lambda_m,p)\in(\mathbb{R}^{m+1})^*\times C^1([t_0,t_1],(\mathbb{R}^{d_1})^*)$ — набор множителей Лагранжа.

Если фиксирована пара $(\widehat{x},\widehat{u})$, то, как и раньше, для сокращения записи используем обозначения: $\widehat{L}_x(t)=L_x(t,\widehat{x}(t),\widehat{u}(t))$ и аналогично для частной производной по u, частных производных отображения φ, l и т. д.

Теорема 23 (Необходимые условия минимума в задаче (43))**.** *Пусть* $(\widehat{x}(\cdot),\widehat{u}(\cdot))$ *доставляет слабый локальный минимум в задаче* (43)*. Тогда, если функции* L_j*,* $0\le j\le m$*, и отображение* φ *непрерывны вместе со своими частными производными по* x *и* u *в окрестности множества* $\Gamma(\widehat{x},\widehat{u})$*, а функции* l_j*,* $0\le j\le m$*, непрерывно дифференцируемы в окрестности точки* $(\widehat{x}(t_0),\widehat{x}(t_1))$*, то найдется ненулевой набор множителей Лагранжа* $\overline{\lambda}=(\lambda_0,\lambda_1,\ldots,\lambda_m,p)\in(\mathbb{R}^{m+1})^*\times C^1([t_0,t_1],(\mathbb{R}^{d_1})^*)$ *такой, что выполняются*

(*a*) *условия стационарности* (*уравнения Эйлера–Лагранжа*)*:*

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t)+\widehat{L}_x(t)=0\ \Leftrightarrow\ -\dot{p}(t)=p(t)\widehat{\varphi}_x(t)-\sum_{j=0}^{m}\lambda_j\widehat{L}_{jx}(t),$$

$$\widehat{L}_u(t)=0\ \Leftrightarrow\ p(t)\widehat{\varphi}_u(t)=\sum_{j=0}^{m}\lambda_j\widehat{L}_{ju}(t);$$

(*b*) $\lambda_j\ge 0$*,* $j=0,1,\ldots m'$ (*условия неотрицательности*)*;*

(*c*) $\lambda_j f_j(\widehat{x},\widehat{u})=0$*,* $j=1,\ldots,m'$ (*условия дополняющей нежесткости*)*;*

(*d*) $\widehat{L}_{\dot{x}}(t_j)=(-1)^j\widehat{l}_{\xi_j}$*,* $j=0,1$ (*условия трансверсальности*)*.*

Доказательство. Положим $\xi=(x,u)$, $\widehat{\xi}=(\widehat{x},\widehat{u})$ и $\Phi(\xi)=(G(\xi),F(\xi))$, где $G(\xi)=(f_{m'+1}(\xi),\ldots,f_m(\xi))^T$, $F(\xi)=\dot{x}-\varphi(t,x,u)$. Тогда задача (43) может быть переписана в виде

$$(44)\qquad f_0(\xi)\to\min,\quad f_j(\xi)\le 0,\quad 1\le j\le m',\quad \Phi(\xi)=0.$$

Будем применять к этой задаче теорему 17. Из дифференциальных свойств функций L_j и l_j, $j=0,1,\ldots,m$, и φ, утверждений относительно дифференцируемости суперпозиций отображений и производной обобщенного оператора Немыцкого следует дифференцируемость отображений f_j, $j = 0, 1, \dots, m'$, и строгая дифференцируемость отображений F и G.

Покажем, что $\operatorname{Im} F'(\widehat{\xi}) = C([t_0, t_1], \mathbb{R}^{d_1})$. Пусть $y \in C([t_0, t_1], \mathbb{R}^{d_1})$. Имеем

$$F'(\widehat{\xi})(h, v)(t) = \dot{h}(t) - \widehat{\varphi}_x(t)h(t) - \widehat{\varphi}_u(t)v(t).$$

Положим $v = 0$. Уравнение

$$\dot{h} - \widehat{\varphi}_x(t)h = y(t) \tag{45}$$

является линейной системой дифференциальных уравнений с непрерывными коэффициентами. Поэтому для любого $y \in C([t_0, t_1], \mathbb{R}^{d_1})$ существует решение этой системы $h \in C^1([t_0, t_1], \mathbb{R}^{d_1})$. В силу леммы о замкнутости образа (лемма 2) $\operatorname{Im} \Phi'(\widehat{\xi})$ — замкнутое подпространство.

Таким образом, все условия теоремы 17 выполнены. Тогда существует ненулевой набор множителей Лагранжа $\overline{\lambda} = (\lambda_0, \dots, \lambda_{m'}, \eta^*)$ такой, что для функции Лагранжа

$$\mathcal{L}_1(\xi, \overline{\lambda}) = \sum_{j=0}^{m'} \lambda_j f_j(\xi) + \langle \eta^*, \Phi(\xi) \rangle$$

$$= \sum_{j=0}^{m'} \lambda_j f_j(\xi) + \langle \mu^*, G(\xi) \rangle + \langle y^*, F(\xi) \rangle = \sum_{j=0}^{m} \lambda_j f_j(\xi) + \langle y^*, F(\xi) \rangle,$$

выполняются условия стационарности $\mathcal{L}_1'(\widehat{\xi}, \overline{\lambda}) = 0$, а, кроме того, выполнены условия неотрицательности $\lambda_j \geq 0$, $j = 0, 1, \dots, m'$ и условия дополняющей нежесткости $\lambda_j f_j(\widehat{x}, \widehat{u}) = 0$, $j = 1, \dots, m'$.

Займемся исследованием условия стационарности, из которого получаем уравнения

$$\mathcal{L}_{1x}(\widehat{\xi}, \overline{\lambda})h(t) = \int_{t_0}^{t_1} \widehat{M}(t)h(t)\, dt + \widehat{l}_{\xi_0}h(t_0) + \widehat{l}_{\xi_1}h(t_1) + \langle y^*, \dot{h}(t) - \widehat{\varphi}_x(t)h(t) \rangle = 0, \tag{46}$$

где

$$\widehat{M} = \sum_{j=0}^{m} \lambda_j \widehat{L}_{jx},$$

а также

$$\mathcal{L}_{1u}(\widehat{\xi}, \overline{\lambda})v(t) = \int_{t_0}^{t_1} \left(\sum_{j=0}^{m} \lambda_j \widehat{L}_{ju}(t) \right) v(t)\, dt - \langle y^*, \widehat{\varphi}_u(t)v(t) \rangle = 0. \tag{47}$$

Пусть $y \in C([t_0,t_1],\mathbb{R}^{d_1})$, а $h \in C^1([t_0,t_1],\mathbb{R}^{d_1})$ является решением системы (45) с начальным условием $h(t_0)=h_0$. Тогда из (46) имеем

$$\langle y^*, y\rangle = -\int_{t_0}^{t_1} \widehat{M}(t)h(t)\,dt - \widehat{l}_{\xi_0}h_0 - \widehat{l}_{\xi_1}h(t_1).$$

Определим функцию p как решение системы

$$-\dot p - p\widehat{\varphi}_x(t) + \widehat{M}(t) = 0, \tag{48}$$

удовлетворяющее условию $p(t_1) = -\widehat{l}_{\xi_1}$. Из (45) получаем

$$p(t)\widehat{\varphi}_x(t)h(t) = p(t)\dot h(t) - p(t)y(t).$$

Тогда, выражая $\widehat{M}$ из (48), получаем

$$\begin{aligned}\langle y^*, y\rangle &= -\int_{t_0}^{t_1} (\dot p(t) + p(t)\widehat{\varphi}_x(t))h(t)\,dt - \widehat{l}_{\xi_0}h_0 - \widehat{l}_{\xi_1}h(t_1)\\ &= -\int_{t_0}^{t_1} (\dot p(t)h(t) + p(t)\dot h(t) - p(t)y(t))\,dt - \widehat{l}_{\xi_0}h_0 - \widehat{l}_{\xi_1}h(t_1)\\ &= \int_{t_0}^{t_1} p(t)y(t)\,dt - \int_{t_0}^{t_1} \frac{d}{dt}(p(t)h(t))\,dt - \widehat{l}_{\xi_0}h_0 - \widehat{l}_{\xi_1}h(t_1)\\ &= \int_{t_0}^{t_1} p(t)y(t)\,dt + (p(t_0) - \widehat{l}_{\xi_0})h_0.\end{aligned}$$

Полагая $h_0 = 0$, получаем, что

$$\langle y^*, y\rangle = \int_{t_0}^{t_1} p(t)y(t)\,dt,$$

а полагая $y = 0$, в силу произвольности h_0 получаем, что $p(t_0) = \widehat{l}_{\xi_1}$.

Учитывая вид функционала y^*, из (47) имеем

$$\int_{t_0}^{t_1} \left(\sum_{j=0}^{m} \lambda_j \widehat{L}_{ju}(t) - p(t)\widehat{\varphi}_u(t)\right) v(t)\,dt = 0.$$

В силу того, что это равенство справедливо для любой функции $v \in C([t_0,t_1],\mathbb{R}^{d_2})$, получаем, что

$$p(t)\widehat{\varphi}_u(t) = \sum_{j=0}^{m} \lambda_j \widehat{L}_{ju}(t).$$

□

19. Задача со старшими производными. Уравнение Эйлера–Пуассона. Изопериметрическая задача

В качестве следствия теоремы, доказанной в предыдущем разделе, получим необходимые условия экстремума в задаче со старшими производными и изопериметрической задаче.

Пусть $[t_0,t_1]$ — отрезок числовой прямой, G — открытое подмножество $\mathbb{R}^{d+2}$, $L\colon G\to\mathbb{R}$ — непрерывная функция (переменные которой обозначаем $t,x,\dot x,\dots,x^{(d)}$) и $x_j^{(k)}\in\mathbb{R}$, $k=0,1,\dots,d-1$, $j=0,1$. Задача

$$(49)\quad J(x)=\int_{t_0}^{t_1}L(t,x(t),\dot x(t),\dots,x^{(d)}(t))\,dt\to\min,$$
$$x^{(k)}(t_j)=x_j^{(k)},\quad k=0,1,\dots,d-1,\quad j=0,1,$$

называется *задачей со старшими производными.*

Обозначим через $C^d([t_0,t_1])$ пространство всех d раз непрерывно дифференцируемых функций x на $[t_0,t_1]$ с нормой

$$\|x\|_{C^d([t_0,t_1])}=\max(\|x\|_{C([t_0,t_1])},\|\dot x\|_{C([t_0,t_1])},\dots,\|x^{(d)}\|_{C([t_0,t_1])}).$$

Функция $x\in C^d([t_0,t_1])$ называется *допустимой в задаче* (49), если

$$\Gamma(x)=\{(t,x(t),\dot x(t),\dots,x^{(d)}(t))^T:t\in[t_0,t_1]\}\subset G$$

и $x^{(k)}(t_j)=x_j^{(k)}$, $k=0,1,\dots,d-1$, $j=0,1$.

Допустимая функция $\widehat x$ называется *локальным минимумом в задаче* (49), если существует такое $\varepsilon>0$, что для любой допустимой функции x, для которой $\|x-\widehat x\|_{C^d([t_0,t_1])}<\varepsilon$ выполняется неравенство $J(x)\ge J(\widehat x)$.

Теорема 24 (Необходимые условия минимума в задаче (49). Уравнение Эйлера–Пуассона)**.** *Пусть $\widehat x$ — локальный минимум в (49), функция L непрерывна вместе с частными производными $L_{x^{(k)}}$ в окрестности $\Gamma(\widehat x)$, $\widehat L_{x^{(k)}}\in C^k([t_0,t_1])$, $k=1,\dots,d-1$, и $\widehat L_{x^{(d)}}\in C^{d-1}([t_0,t_1])$. Тогда $\widehat L_{x^{(d)}}\in C^d([t_0,t_1])$ и для всех $t\in[t_0,t_1]$ выполнено уравнение Эйлера–Пуассона*

$$\sum_{k=0}^{d}(-1)^k\frac{d^k}{dt^k}\widehat L_{x^{(k)}}(t)=0.$$

Доказательство. Обозначая $x=x_1$, $\dot x_1=x_2$, …, $\dot x_{d-1}=x_d$, $\dot x_d=u$, задачу (49) можно записать как задачу Лагранжа

$$(50)\quad\int_{t_0}^{t_1}L(t,x_1(t),\dots,x_d(t),u(t))\,dt\to\min,$$
$$\dot x_1=x_2,\ \dots,\ \dot x_{d-1}=x_d,\ \dot x_d=u,$$
$$x_k(t_j)=x_j^{(k-1)},\ k=1,\dots,d,\ j=0,1.$$

Простая проверка показывает, что если $\widehat x$ — локальный минимум в задаче (49), то вектор-функция $(\widehat x_1,\dots,\widehat x_d,\widehat u)^T$ — локальный минимум в данной задаче. Согласно общей теореме 43 найдутся такие множители Лагранжа λ_0 и $p = (p_1, \ldots, p_d) \in C^1([t_0, t_1], (\mathbb{R}^d)^*)$, не равные одновременно нулю, что имеют место равенства

$$
\begin{gathered}
-\dot{p}(t) = p(t)\widehat{\varphi}_x(t) - \lambda_0 \widehat{L}_x(t), \\
p(t)\widehat{\varphi}_u(t) = \lambda_0 \widehat{L}_u(t).
\end{gathered}
$$

Поскольку $\varphi(t, x, u) = (x_2, \ldots, x_d, u)^T$, то эти равенства имеют вид

$$
\begin{aligned}
-\dot{p}_1(t) &= -\lambda_0 \widehat{L}_{x_1}(t), \\
-\dot{p}_2(t) &= p_1(t) - \lambda_0 \widehat{L}_{x_2}(t), \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
-\dot{p}_d(t) &= p_{d-1}(t) - \lambda_0 \widehat{L}_{x_d}(t), \\
p_d(t) &= \lambda_0 \widehat{L}_u(t).
\end{aligned} \tag{51}
$$

Если $\lambda_0 = 0$, то сразу видно, что и $p = 0$. Будем считать, что $\lambda_0 = 1$. Из второго равенства следует, что $p_2 \in C^2([t_0, t_1])$, из третьего — $p_3 \in C^3([t_0, t_1])$, наконец, из предпоследнего — $p_d \in C^d([t_0, t_1])$. Тогда из последнего равенства следует, что $\widehat{L}_{x^{(d)}} \in C^d([t_0, t_1])$.

Из соотношений (51) выводим (учитывая условия теоремы и переходя к прежним обозначениям: $x_1 = x$, $x_2 = \dot{x}$, $\ldots$, $x_d = x^{(d-1)}$, $u = x^{(d)}$), что

$$
\begin{aligned}
\widehat{L}_x(t) = \dot{p}_1(t) = -\ddot{p}_2(t) + \frac{d}{dt}\widehat{L}_{\dot{x}}(t) = \ldots = (-1)^{d-1} p_d^{(d)}(t) \\
+ \sum_{k=1}^{d-1} (-1)^{k-1} \frac{d^k}{dt^k} \widehat{L}_{x^{(k)}}(t) = (-1)^{d-1} \frac{d^d}{dt^d} \widehat{L}_{x^{(d)}}(t) \\
+ \sum_{k=1}^{d-1} (-1)^{k-1} \frac{d^k}{dt^k} \widehat{L}_{x^{(k)}}(t).
\end{aligned}
$$

Это, очевидно, равносильно уравнению Эйлера–Пуассона. □

Пусть $[t_0, t_1]$ — отрезок числовой прямой, G — открытое подмножество $\mathbb{R}^{2d+1}$, функции $f_j\colon G \to \mathbb{R}$, $j = 0, 1, \ldots, m$ (переменных $t \in \mathbb{R}$, $x \in \mathbb{R}^d$ и $\dot{x} \in \mathbb{R}^d$) непрерывны на G, $\alpha_j \in \mathbb{R}$, $j = 1, \ldots, m$, и $x_j \in \mathbb{R}^d$, $j = 0, 1$. Задача

$$
J(x) = \int_{t_0}^{t_1} f_0(t, x(t), \dot{x}(t))\, dt \to \min, \tag{52}
$$

$$
\int_{t_0}^{t_1} f_j(t, x(t), \dot{x}(t))\, dt = \alpha_j, \ 1 \le j \le m, \quad x(t_0) = x_0, \quad x(t_1) = x_1
$$

называется *изопериметрической задачей*. Ясно, что это частный случай задачи (43) (дифференциальная связь: $\dot{x} = u$).

Функция x называется *допустимой в задаче* (52), если $x \in C^1([t_0,t_1],\mathbb{R}^d)$,

$$\Gamma(x) = \{(t, x(t), \dot{x}(t))^T : t \in [t_0,t_1]\} \subset G,$$

$$\int_{t_0}^{t_1} f_j(t, x(t), \dot{x}(t))\,dt = \alpha_j,\ 1 \le j \le m, \quad x(t_j) = x_j,\ j = 0,1.$$

Допустимая функция $\widehat{x}$ называется *локальным минимумом* в задаче (52), если существует такое $\varepsilon > 0$, что для любой допустимой функции x, для которой $\|x - \widehat{x}\|_{C^1([t_0,t_1],\mathbb{R}^d)} < \varepsilon$ выполняется неравенство $J(x) \ge J(\widehat{x})$.

Положим

$$L(t, x, \dot{x}, \overline{\lambda}) = \sum_{j=0}^{m} \lambda_j f_j(t, x, \dot{x}), \quad \overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m).$$

Теорема 25 (Необходимые условия минимума в задаче (52))**.** *Пусть $\widehat{x}$ доставляет слабый локальный минимум в задаче* (52)*. Тогда, если функции f_j, $j = 0,1,\ldots,m$, непрерывны вместе со своими частными производными по x и $\dot{x}$ в окрестности $\Gamma(\widehat{x})$, то найдется такой ненулевой набор множителей Лагранжа $\overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, что $\widehat{L}_{\dot{x}} \in C^1([t_0,t_1],\mathbb{R}^d)$ и для всех $t \in [t_0,t_1]$ выполнено уравнение Эйлера*

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0. \tag{53}$$

Доказательство. Рассмотрим следующую задачу Лагранжа

$$\int_{t_0}^{t_1} f_0(t, x(t), u(t))\,dt \to \min, \quad \dot{x} = u, \tag{54}$$

$$\int_{t_0}^{t_1} f_j(t, x(t), u(t))\,dt = \alpha_j,\ 1 \le j \le m, \quad x(t_0) = x_0, \quad x(t_1) = x_1.$$

Несложная проверка показывает, что если $\widehat{x}$ — локальный экстремум в задаче (52), то $(\widehat{x}, \widehat{u})$, где $\widehat{u} = \dot{\widehat{x}}$, — локальный экстремум в данной задаче.

Согласно теореме о необходимых условиях минимума в задаче (43) найдутся такие множители Лагранжа $(\lambda_0, \lambda_1, \ldots, \lambda_m)$ и $p \in C^1([t_0,t_1],(\mathbb{R}^d)^*)$, не все равные нулю, что

$$-\dot{p} = p\widehat{\varphi}_x(t) - \sum_{j=0}^{m} \lambda_j f_{jx}(t, \widehat{x}, \widehat{u}),$$

$$p(t)\widehat{\varphi}_u(t) = \sum_{j=0}^{m} \lambda_j f_{ju}(t, \widehat{x}, \widehat{u}).$$

Поскольку $\varphi(t,x,u)=u$, то получаем равенства

$$-\dot p = -\sum_{j=0}^{m} \lambda_j f_{jx}(t,\widehat{x},\widehat{u}),$$

$$p(t) = \sum_{j=0}^{m} \lambda_j f_{ju}(t,\widehat{x},\widehat{u}).$$

Подставляя в эти равенства $\dot{\widehat{x}}$ вместо $\widehat{u}$ убеждаемся, что $\widehat{L}_{\dot x} \in C^1([t_0,t_1],\mathbb{R}^d)$ и имеет место равенство (53). $\square$

20. Задачи оптимального управления

Вариационное исчисление, как уже говорилось, интенсивно развивалось в 18 веке (в основном усилиями Эйлера, Лагранжа и Лежандра). В 19 веке в его развитии приняли участие такие математики как Пуассон, Вейерштрасс, Гильберт и Пуанкаре. К началу 20 века предмет, в существенном, оказался исчерпанным. Построение теории экстремума (как она представлялась в те времена), казалось, завершено. Но впоследствии появились выпуклые задачи, а затем, в начале 50-х годов прошлого века родилось оптимальное управление — новое направление в теории экстремума, охватывающее вариационное исчисление. Необходимые условия экстремума в задачах оптимального управления были получены в школе Л. С. Понтрягина. Основной результат называется принципом максимума Понтрягина. В этом разделе рассматривается задача оптимального управления и доказываются для нее необходимые условия минимума.

Пусть $[t_0,t_1]$ — конечный отрезок, G — открытое подмножество $\mathbb{R}\times\mathbb{R}^{d_1}$, U — непустое подмножество $\mathbb{R}^{d_2}$, функция $f\colon G\times U\to\mathbb{R}$ и отображение $\varphi\colon G\times U\to\mathbb{R}^{d_1}$ (переменных $t\in\mathbb{R}$, $x=(x_1,\dots,x_{d_1})^T\in\mathbb{R}^{d_1}$ и $u=(u_1,\dots,u_{d_2})^T\in\mathbb{R}^{d_2}$) непрерывны на $G\times U$ и $x_0,x_1\in\mathbb{R}^{d_1}$. Задача

$$(55)\quad J(x,u)=\int_{t_0}^{t_1} f(t,x(t),u(t))\,dt\to\min,\quad \dot x=\varphi(t,x,u),$$
$$u(t)\in U,\quad x(t_0)=x_0,\quad x(t_1)=x_1$$

называется *задачей оптимального управления.* Переменную x часто называют фазовой переменной, а u — управлением.

Уточним постановку. Пусть $PC^1([t_0,t_1],\mathbb{R}^{d_1})$ — совокупность всех кусочно-непрерывно дифференцируемых, а $PC([t_0,t_1],\mathbb{R}^{d_2})$ — кусочно-непрерывных функций на $[t_0,t_1]$ со значениями соответственно в $\mathbb{R}^{d_1}$ и $\mathbb{R}^{d_2}$. Пара $(x,u)\in PC^1([t_0,t_1],\mathbb{R}^{d_1})\times PC([t_0,t_1],\mathbb{R}^{d_2})$ называется *допустимой в задаче* (55), если

$$\Gamma(x)=\{\,(t,x(t)) : t\in[t_0,t_1]\,\}\subset G,$$

включение $u(t) \in U$ и равенство $\dot{x}(t) = \varphi(t, x(t), u(t))$ выполняются для всех $t \in [t_0, t_1]$, где функция u непрерывна и $x(t_0) = x_0$, $x(t_1) = x_1$.

Допустимая пара $(\widehat{x}, \widehat{u})$ называется *сильным локальным минимумом в задаче* (55), если существует такое $\varepsilon > 0$, что для любой допустимой пары (x, u), для которой $\|x - \widehat{x}\|_{C([t_0,t_1],\mathbb{R}^{d_1})} < \varepsilon$ выполнено неравенство $J(x, u) \geq J(\widehat{x}, \widehat{u})$.

Функцией Лагранжа для задачи (55) будем называть функцию

$$\mathcal{L}(x, u, \overline{\lambda}) = \int_{t_0}^{t_1} L(t, x, \dot{x}, u, \lambda_0, p)\, dt,$$

где

$$L(t, x, \dot{x}, u, \lambda_0, p) = \lambda_0 f(t, x(t), u(t)) + p(t)(\dot{x}(t) - \varphi(t, x(t), u(t)).$$

Функцию

$$H(t, x, u, \lambda_0, p) = p\varphi(t, x, u) - \lambda_0 f(t, x, u)$$

называют *функцией Понтрягина задачи* (55).

Теорема 26 (Необходимые условия минимума в задаче (55). Принцип максимума Понтрягина)**.** *Пусть* $(\widehat{x}, \widehat{u})$ *доставляет сильный минимум в задаче* (55). *Тогда, если функция* f *и отображение* φ *непрерывны вместе со своими частными производными по* x *в* $G \times U$, *то найдется ненулевой набор множителей Лагранжа* $\overline{\lambda} = (\lambda_0, p) \in \mathbb{R} \times PC^1([t_0, t_1], (\mathbb{R}^{d_1})^*)$ *такой, что выполнено условие стационарности по* x

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

и в точках непрерывности $\widehat{u}$ *условие минимума по* u

$$\min_{u \in U} L(t, \widehat{x}(t), \dot{\widehat{x}}(t), u, \lambda_0, p(t)) = L(t, \widehat{x}(t), \dot{\widehat{x}}(t), \widehat{u}(t), \lambda_0, p(t)). \tag{56}$$

Условие (56) может быть записано в виде условия максимума по u

$$\max_{u \in U} H(t, \widehat{x}(t), u, \lambda_0, p(t)) = H(t, \widehat{x}(t), \widehat{u}(t), \lambda_0, p(t)).$$

В силу этого соотношения необходимые условия в задаче оптимального управления и называют “Принципом максимума Понтрягина”.

Сформулированная задача оптимального управления не самая общая, но достаточно представительная. Мы получим необходимые условия для более простого варианта, когда правый конец свободен, т. е. получим необходимые условия минимума в задаче

$$J(x, u) = \int_{t_0}^{t_1} f(t, x(t), u(t))\, dt \to \min, \quad \dot{x} = \varphi(t, x, u), \tag{57}$$
$$u(t) \in U, \quad x(t_0) = x_0.$$

Эти условия минимума те же, что и в задаче (55), но добавляется еще условие трансверсальности $p(t_1) = 0$.

Перед доказательством сформулируем две леммы (считаем, что функции f и φ удовлетворяют условиям теоремы 26).

Лемма 5 (об игольчатой вариации)**.** *Пусть* $(\widehat{x}, \widehat{u})$ *— допустимая пара в задаче* (55) (*или* (57))*,* $\tau \in (t_0, t_1)$ *— точка непрерывности* $\widehat{u}$*,* $\alpha > 0$ *столь мало, что функция* $\widehat{u}$ *непрерывна на* $[\tau - \alpha, \tau]$ *и* $v \in U$*. Положим*

$$u_\alpha(t, \tau, v) = \begin{cases} \widehat{u}(t), & t \notin [\tau - \alpha, \tau), \\ v, & t \in [\tau - \alpha, \tau). \end{cases}$$

Тогда найдется такое $\alpha_0 > 0$*, что для любого* $0 \leq \alpha < \alpha_0$ *существует единственное решение* $x_\alpha(t, \tau, v)$ *задачи Коши*

$$\dot{x} = \varphi(t, x, u_\alpha(t, \tau, v)), \quad x(t_0) = x_0,$$

определенное на всем отрезке $[t_0, t_1]$*. При этом,* $x_\alpha(t, \tau, v) \to \widehat{x}(t)$ *при* $\alpha \to 0$ *равномерно на* $[t_0, t_1]$*, отображение* $F(\alpha) = x_\alpha(t, \tau, v)$ *при каждом* $t \in [\tau, t_1]$ *непрерывно дифференцируемо для достаточно малых* $\alpha \geq 0$ *и* $y_{\tau v}(t) = F'(0)$ *на* $[\tau, t_1]$ *удовлетворяет уравнению*

$$\dot{y}_{\tau v} = \widehat{\varphi}_x(t) y_{\tau v}, \quad y_{\tau v}(\tau) = \varphi(\tau, \widehat{x}(\tau), v) - \varphi(\tau, \widehat{x}(\tau), \widehat{u}(\tau)). \tag{58}$$

Функцию $u_\alpha(t, \tau, v)$ называют *игольчатой вариацией* $\widehat{u}$, а пару (τ, v) — *иголкой.*

Лемма 6 (о производной функционала)**.** *Пусть* $(\widehat{x}, \widehat{u})$ *— допустимая пара в задаче* (55) (*или* (57))*. Тогда для функции*

$$J(\alpha) = J(x_\alpha(t, \tau, v), u_\alpha(t, \tau, v))$$

имеет место равенство

$$J'(0) = \Delta_{\tau v} f + \int_\tau^{t_1} \widehat{f}_x(t) y_{\tau v}(t)\, dt, \tag{59}$$

где $\Delta_{\tau v} f = f(\tau, \widehat{x}(\tau), v) - f(\tau, \widehat{x}(\tau), \widehat{u}(\tau))$ *и* $y_{\tau v}$ *— решение задачи* (58)*.*

Доказательство. Имеем

$$\begin{aligned} J'(0+0) &= \lim_{\alpha \to 0+0} \frac{J(\alpha) - J(0)}{\alpha} \\ &= \lim_{\alpha \to 0+0} \frac{1}{\alpha} \int_{\tau-\alpha}^{\tau} \bigl(f(t, x_\alpha(t, \tau, v), v) - f(t, \widehat{x}(t), \widehat{u}(t))\bigr)\, dt \\ &\quad + \lim_{\alpha \to 0+0} \frac{1}{\alpha} \int_\tau^{t_1} \bigl(f(t, x_\alpha(t, \tau, v), \widehat{u}(t)) - f(t, \widehat{x}(t), \widehat{u}(t))\bigr)\, dt. \end{aligned}$$

Второе слагаемое справа — производная в нуле интеграла по параметру, к которому применима стандартная теорема о дифференцируемости под знаком интеграла, ибо отображение $F(\alpha) = x_\alpha(t, \tau, v)$

непрерывно дифференцируемо для каждого $t \in [\tau, t_1]$, а отрезок $[\tau, t_1]$ можно разбить на конечное число интервалов, на каждом из которых функция $\widehat{u}$ непрерывна. К первому интегралу применим теорему о среднем для интегралов и тогда в итоге получим

$$\begin{aligned} J'(0+0) = \lim_{\alpha \to 0+0} (f(\xi, x_\alpha(\xi, \tau, v), v) - f(\xi, \widehat{x}(\xi), \widehat{u}(\xi))) \\ + \int_\tau^{t_1} \widehat{f}_x(t) y_{\tau v}(t)\, dt, \end{aligned}$$

где $\xi \in [\tau - \alpha, \tau]$. Когда $\alpha \to 0+0$, то, очевидно, $\xi \to \tau$, $x_\alpha(\xi, \tau, v) \to \widehat{x}(\tau)$ согласно лемме об игольчатой вариации, а $\widehat{u}(\xi) \to \widehat{u}(\tau)$, так как $\widehat{u}$ непрерывна в точке τ. Формула (59) доказана. $\square$

Доказательство принципа максимума. Пусть $(\widehat{x}, \widehat{u})$ — сильный минимум в задаче (55). Обозначая через p — решение линейного уравнения

$$-\dot{p} = p\widehat{\varphi}_x(t) - \widehat{f}_x(t), \quad p(t_1) = 0, \tag{60}$$

получаем условие стационарности по x и условие трансверсальности.

Так как пара $(\widehat{x}, \widehat{u})$ доставляет минимум, то необходимо $J'(0+0) \geq 0$, или согласно (59)

$$\Delta_{\tau v} f + \int_\tau^{t_1} \widehat{f}_x(t) y_{\tau v}(t)\, dt \geq 0.$$

Подставим сюда вместо функции $\widehat{f}_x$ ее выражение из (60), а затем вместо функции $\widehat{\varphi}_x y_{\tau v}$ ее выражение из (58) и, учитывая, что $p(t_1) = 0$, а $y_{\tau v}(\tau) = \varphi(\tau, \widehat{x}(\tau), v) - \varphi(\tau, \widehat{x}(\tau), \widehat{u}(\tau))$, будем иметь

$$\begin{aligned} 0 \leq \Delta_{\tau v} f + \int_\tau^{t_1} (\dot{p}(t) y_{\tau v}(t) + p(t) \dot{y}_{\tau v}(t))\, dt = \Delta_{\tau v} f + p(t) y_{\tau v}(t)\big|_\tau^{t_1} \\ = f(\tau, \widehat{x}(\tau), v) - f(\tau, \widehat{x}(\tau), \widehat{u}(\tau)) \\ - p(\tau)(\varphi(\tau, \widehat{x}(\tau), v) - \varphi(\tau, \widehat{x}(\tau), \widehat{u}(\tau))), \end{aligned}$$

или

$$\begin{aligned} p(\tau)\varphi(\tau, \widehat{x}(\tau), v) - f(\tau, \widehat{x}(\tau), v) \\ \leq p(\tau)\varphi(\tau, \widehat{x}(\tau), \widehat{u}(\tau)) - f(\tau, \widehat{x}(\tau), \widehat{u}(\tau)) \end{aligned}$$

Таким образом, для любой точки τ, где функция $\widehat{u}$ непрерывна, максимум выражения слева по всем $v \in U$ достигается в точке $\widehat{u}(\tau)$. Это и есть условие максимума из теоремы. $\square$

21. Необходимые условия сильного экстремума в простейшей задаче вариационного исчисления. Условие Вейрштрасса и Лежандра

Чтобы не усложнять выкладки, всюду далее будем иметь дело с классическим (скалярным) вариантом простейшей задачи вариационного исчисления, хотя все доказываемые утверждения справедливы и для векторного варианта. Для определенности будем рассматривать задачу на минимум

$$J(x) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt \to \min, \quad x(t_j) = x_j,\ j = 0, 1. \tag{61}$$

Здесь $[t_0, t_1]$ — отрезок числовой прямой, непрерывная функция L переменных $t \in \mathbb{R}$, $x \in \mathbb{R}$ и $\dot{x} \in \mathbb{R}$ определена на открытом подмножестве $G \subset \mathbb{R}^3$ и $x_j \in \mathbb{R}$, $j = 0, 1$.

Напомним также, что функция $x \in C^1([t_0, t_1])$ называется *допустимой в задаче* (61), если

$$\Gamma(x) = \{\, (t, x(t), \dot{x}(t)) : t \in [t_0, t_1] \,\} \subset G$$

и $x(t_j) = x_j$, $j = 0, 1$, и допустимая функция $\widehat{x}$ называется *слабым локальным минимумом* в задаче (61), если существует такое $\varepsilon > 0$, что для любой допустимой функции x, для которой $\|x - \widehat{x}\|_{C^1([t_0,t_1]} < \varepsilon$ выполняется неравенство $J(x) \geq J(\widehat{x})$.

В вариационном исчислении, наряду со слабым экстремумом, рассматривают еще и сильный локальный экстремум, где близость функций измеряется в пространстве $C([t_0, t_1])$. Точнее говоря, обозначим через $PC^1([t_0, t_1])$ пространство кусочно-непрерывно-дифференцируемых функций на $[t_0, t_1]$. Функция $x \in PC^1([t_0, t_1])$ называется допустимой в задаче (61), если для всех точек непрерывности функции $\dot{x}$ выполнено условие $(t, x(t), \dot{x}(t)) \in G$ и, кроме того, $x(t_j) = x_j$, $j = 0, 1$. Скажем, что допустимая функция $\widehat{x}$ доставляет *сильный локальный минимум* в задаче (61), если существует такое $\varepsilon > 0$, что для любой допустимой функции x, для которой $\|x - \widehat{x}\|_{C([t_0,t_1])} < \varepsilon$ выполняется неравенство $J(x) \geq J(\widehat{x})$.

Заметим, что если функция $\widehat{x}$ доставляет сильный минимум в задаче (61) и при этом $\widehat{x} \in C^1([t_0, t_1])$, то $\widehat{x}$ является и слабым минимумом в этой задаче. Действительно, пусть $\varepsilon > 0$ такое, что как только $x \in PC^1([t_0, t_1])$ и $\|x - \widehat{x}\|_{C([t_0,t_1])} < \varepsilon$, то $J(x) \geq J(\widehat{x})$. Если теперь $x \in C^1([t_0, t_1])$ и $\|x - \widehat{x}\|_{C^1([t_0,t_1])} < \varepsilon$, то так как, в частности, $x \in PC^1([t_0, t_1])$ и $\|x - \widehat{x}\|_{C([t_0,t_1])} \leq \|x - \widehat{x}\|_{C^1([t_0,t_1])} < \varepsilon$, получаем, что $J(x) \geq J(\widehat{x})$.

Таким образом, необходимые условия слабого минимума для $x \in C^1([t_0, t_1])$ являются необходимыми условиями и сильного минимума.

Функция, для которой выполнено уравнение Эйлера называется *экстремалью* (или *стационарной точкой*) задачи.

Пусть $\widehat{x}$ — экстремаль в задаче (61) и существует
$$\widehat{L}_{\dot{x}\dot{x}}(t) = L_{\dot{x}\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t)).$$
Говорят, что на функции $\widehat{x}$ выполнено *условие Лежандра*, если $\widehat{L}_{\dot{x}\dot{x}}(t) \geq 0$ для всех $t \in [t_0, t_1]$ и *усиленное условие Лежандра*, если $\widehat{L}_{\dot{x}\dot{x}}(t) > 0$ для всех $t \in [t_0, t_1]$.

Пусть $f\colon \mathbb{R} \to \mathbb{R}$ — дифференцируемая функция. Функция
$$\mathcal{E}(x, x') = f(x') - f(x) - f'(x)(x' - x)$$
называется *функцией Вейерштрасса* (соответствующей функции f). Геометрически, $\mathcal{E}(x, x')$ — разность между значением функции f и функции $g(y) = f(x) + f'(x)(y - x)$ (график которой есть касательная к графику функции f в точке x) в точке x'.

Если f — выпуклая функция, то $\mathcal{E}(x, x') \geq 0$ для всех $x, x' \in \mathbb{R}$. Действительно, пусть $x, x' \in \mathbb{R}$ и $0 < \alpha < 1$. По неравенству Йенссена
$$f((1-\alpha)x + \alpha x') \leq (1-\alpha)f(x) + \alpha f(x'),$$
откуда
$$\alpha^{-1}(f(x + \alpha(x' - x)) - f(x)) \leq f(x') - f(x).$$
Переходя к пределу при $\alpha \to 0$, получаем, что $f(x') - f(x) \geq f'(x)(x' - x)$.

Пусть L — интегрант в задаче (61). Если L — дифференцируемая функция на некотором открытом множестве $G \times \mathbb{R}$, где G — открытое подмножество $\mathbb{R} \times \mathbb{R}$, то функция
$$\mathcal{E}(t, x, \dot{x}, u) = L(t, x, u) - L(t, x, \dot{x}) - L_{\dot{x}}(t, x, \dot{x})(u - \dot{x}),$$
определенная на $G \times \mathbb{R} \times \mathbb{R}$, называется *функцией Вейерштрасса функционала J*. Ясно, что при каждых t и x — это функция Вейерштрасса, соответствующая функции $G(\dot{x}) = L(t, x, \dot{x})$.

Говорят, что на экстремали $\widehat{x}$ выполнено *условие Вейрштрасса*, если $\mathcal{E}(t, \widehat{x}(t), \dot{\widehat{x}}(t), u) \geq 0$ для всех $u \in \mathbb{R}$ и $t \in [t_0, t_1]$.

Теорема 27 (Необходимые условия сильного минимума в задаче (61))**.** *Пусть функция $\widehat{x} \in C^1([t_0, t_1])$ доставляет сильный локальный минимум в задаче* (61). *Тогда, если интегрант L непрерывен вместе с частными производными по x и $\dot{x}$ в $G \times \mathbb{R}$, то*

(a) *для всех $t \in [t_0, t_1]$ выполнено уравнение Эйлера*
$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0;$$

(b) *выполнено условие Вейерштрасса, т. е. для всех $t \in [t_0, t_1]$ и $u \in \mathbb{R}$ справедливо неравенство*
$$\mathcal{E}(t, \widehat{x}(t), \dot{\widehat{x}}(t), u) \geq 0;$$

(c) *если существует $\widehat{L}_{\dot{x}\dot{x}}$, то выполнено условие Лежандра, т. е. $\widehat{L}_{\dot{x}\dot{x}}(t) \geq 0$ для всех $t \in [t_0, t_1]$.*

Доказательство. Запишем задачу (61) как задачу оптимального управления

$$(62)\quad \int_{t_0}^{t_1} L(t,x(t),u(t))\,dt \to \min, \quad \dot x = u, \quad x(t_0)=x_0,\ x(t_1)=x_1.$$

Легко видеть, что $\widehat{x}$ доставляет сильный минимум в (61) тогда и только тогда, когда пара $(\widehat{x},\widehat{u})$, где $\widehat{u}=\dot{\widehat{x}}$ является сильным минимумом в (62).

Согласно принципу максимума (теорема 26) найдутся такой ненулевой набор множителей Лагранжа $\overline{\lambda} = (\lambda_0,p) \in \mathbb{R}\times PC^1([t_0,t_1])$, что для всех $t\in[t_0,t_1]$ выполнено условие стационарности по x

$$(63)\qquad -\dot p(t) + \lambda_0 \widehat{L}_x(t) = 0$$

и условие минимума по u

$$(64)\qquad \min_{u\in\mathbb{R}} (\lambda_0 L(t,\widehat{x}(t),u) - p(t)u) = \lambda_0 L(t,\widehat{x}(t),\dot{\widehat{x}}(t)) - p(t)\dot{\widehat{x}}(t).$$

Если $\lambda_0=0$, то $p=\mathrm{const}$ вследствие (63). Тогда из (64) следует, что эта константа обязана быть нулевой и тем самым все множители Лагранжа нулевые. Итак, $\lambda_0\ne 0$ и можно считать, что $\lambda_0=1$.

Условие (64) означает, что для всех $t\in[t_0,t_1]$ функция $f(u) = L(t,\widehat{x}(t),u)-p(t)u$ на $\mathbb{R}$ достигает минимума в точке $\dot{\widehat{x}}(t)$ и, следовательно, по теореме Ферма производная этой функции в данной точке равна нулю, т. е. $p(t)=\widehat{L}_{\dot x}(t)$. Вместе с (63) это дает уравнение Эйлера.

Необходимое условия минимума второго порядка функции f заключаются в том, что $\widehat{L}_{\dot x\dot x}(t)\ge 0$, т. е. выполнено условие Лежандра.

Из соотношения (64) и доказанного равенства $p(t)=\widehat{L}_{\dot x}(t)$ следует, что

$$L(t,\widehat{x}(t),u) - L_{\dot x}(t,\widehat{x}(t),\dot{\widehat{x}}(t))u \ge L(t,\widehat{x}(t),\dot{\widehat{x}}(t)) - L_{\dot x}(t,\widehat{x}(t),\dot{\widehat{x}}(t))\dot{\widehat{x}}(t)$$

для всех $u\in\mathbb{R}$ и $t\in[t_0,t_1]$ или, что то же

$$L(t,\widehat{x}(t),u) - L(t,\widehat{x}(t),\dot{\widehat{x}}(t)) - L_{\dot x}(t,\widehat{x}(t),\dot{\widehat{x}}(t))(u-\dot{\widehat{x}}(t)) \ge 0,$$

т. е. выполнено условие Вейерштрасса. □

22. Необходимые условия слабого экстремума в простейшей задаче вариационного исчисления. Условие Якоби

Предположим, что $\widehat{x}$ — экстремаль задачи (61), а интегрант L дважды непрерывно дифференцируем в некоторой окрестности

$\Gamma(\widehat{x})$. Пусть $h \in C^1([t_0,t_1])$, $h(t_0)=h(t_1)=0$. Рассмотрим функцию (одного переменного)

$$\varphi(\lambda)=\int_{t_0}^{t_1} L(t,\widehat{x}(t)+\lambda h(t),\dot{\widehat{x}}(t)+\lambda \dot{h}(t))\,dt. \tag{65}$$

Тогда для достаточно малых по модулю λ

$$\varphi'(\lambda)=\int_{t_0}^{t_1}(L_x(t,\widehat{x}(t)+\lambda h(t),\dot{\widehat{x}}(t)+\lambda\dot{h}(t))h(t) \\ + L_{\dot{x}}(t,\widehat{x}(t)+\lambda h(t),\dot{\widehat{x}}(t)+\lambda\dot{h}(t))\dot{h}(t))\,dt$$

и дифференцируя еще раз, получаем, что

$$\varphi''(0)=\int_{t_0}^{t_1}\Big(\widehat{L}_{xx}(t)h^2(t)+\widehat{L}_{x\dot{x}}(t)h(t)\dot{h}(t)+\widehat{L}_{\dot{x}x}(t)\dot{h}(t)h(t) \\ +\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t)\Big)\,dt=\int_{t_0}^{t_1}\Big(\widehat{L}_{xx}(t)h^2(t)+2\widehat{L}_{x\dot{x}}(t)h(t)\dot{h}(t) \\ +\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t)\Big)\,dt.$$

Функционал $\varphi''(0)$ (как функцию от h) обозначим через $Q(h)$ и рассмотрим задачу

$$Q(h)\to\min,\quad h(t_0)=h(t_1)=0.$$

Уравнение Эйлера для данной задачи имеет вид

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

и называется *уравнением Якоби* исходной задачи (61).

Пусть на $\widehat{x}$ выполнено усиленное условие Лежандра. Точка $\tau\in(t_0,t_1]$ называется *сопряженной точкой к* t_0, если существует нетривиальное решение h уравнения Якоби, для которого $h(t_0)=h(\tau)=0$.

Говорят, что на $\widehat{x}$ выполнено *условие Якоби*, если в интервале (t_0,t_1) нет точек сопряженных к t_0 и *усиленное условие Якоби*, если полуинтервал $(t_0,t_1]$ не содержит точек сопряженных к t_0.

При доказательстве необходимых условий слабого минимума в задаче (61) нам понадобится один несложный технический результат, который приводим без доказательства (см. [3, стр. 69]).

Лемма 7 (о скруглении углов)**.** *Пусть в задаче* (61) *интегрант* L *непрерывен по совокупности переменных. Тогда*

$$\inf\{\,J(x): x\in PC^1([t_0,t_1])\ x(t_0)=x_0,\ x(t_1)=x_1\,\} \\ =\inf\{\,J(x): x\in C^1([t_0,t_1]),\ x(t_0)=x_0,\ x(t_1)=x_1\,\}.$$

Теорема 28 (Необходимые условия слабого минимума в задаче (61))**.** *Пусть* $\widehat{x} \in C^1([t_0, t_1])$ *доставляет слабый локальный минимум в задаче* (61)*. Тогда, если интегрант* L *дважды непрерывно дифференцируем в окрестности* $\Gamma(\widehat{x})$*, то для всех* $t \in [t_0, t_1]$ *выполнено уравнение Эйлера, условие Лежандра и условие Якоби.*

Доказательство. Уравнение Эйлера, как необходимое условие слабого экстремума, уже было доказано раньше. Доказательство, заключается в том, что если $\widehat{x}$ — локальный минимум, то ноль есть локальный минимум для функции φ, определенной соотношением (65) и тогда необходимо $\varphi'(0) = 0$. Расшифровка этого условия и приводит к уравнению Эйлера.

Докажем условие Лежандра, расшифровывая необходимое условие минимума второго порядка $\varphi''(0) \geq 0$. Согласно формуле для $\varphi''(0)$, выписанной выше, данное условие равносильно тому, что $Q(h) \geq 0$ для всех $h \in C^1([t_0, t_1])$ таких, что $h(t_0) = h(t_1) = 0$. Это означает, что функция $\widehat{h} = 0$ есть слабый абсолютный минимум в задаче

$$(66) \quad Q(h) = \int_{t_0}^{t_1} \Big(\widehat{L}_{xx}(t)h^2(t) + 2\widehat{L}_{x\dot{x}}(t)h(t)\dot{h}(t) + \widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t)\Big)\,dt \to \min, \quad h(t_0) = h(t_1) = 0.$$

По лемме о скруглении углов $\widehat{h} = 0$ доставляет и сильный абсолютный минимум в этой задаче. Тогда, по уже доказанному, на $\widehat{h}$ должно выполняться условие Лежандра, которое в данном случае имеет тот же вид $\widehat{L}_{\dot{x}\dot{x}}(t) \geq 0$.

Докажем условие Якоби. Предположим противное, что существует точка $\tau \in (t_0, t_1)$ и нетривиальное решение $\overline{h}$ уравнения Якоби такое, что $\overline{h}(t_0) = \overline{h}(\tau) = 0$. Пусть функция $\widetilde{h}$ такова, что $\widetilde{h}(t) = \overline{h}(t)$, если $t_0 \leq t \leq \tau$ и $\widetilde{h}(t) = 0$, если $\tau \leq t \leq t_1$. Заметим, что $\dot{\overline{h}}(\tau) \neq 0$, так как в противном случае, по теореме единственности, функция $\overline{h}$ была бы тождественным нулем. Далее, интегрируя по частям ($\overline{h}(t_0) = \overline{h}(\tau) = 0$), получим

$$\begin{aligned} Q(\widetilde{h}) &= \int_{t_0}^{\tau} \Big(\widehat{L}_{xx}(t)\overline{h}^2(t) + 2\widehat{L}_{\dot{x}x}(t)\overline{h}(t)\dot{\overline{h}}(t) + \widehat{L}_{\dot{x}\dot{x}}(t)\dot{\overline{h}}^2(t)\Big)\,dt \\ &= \int_{t_0}^{\tau} \Big(\widehat{L}_{xx}(t)\overline{h}(t) + \widehat{L}_{x\dot{x}}(t)\dot{\overline{h}}(t)\Big)\overline{h}(t)\,dt + \int_{t_0}^{\tau} \Big(\widehat{L}_{\dot{x}x}(t)\overline{h}(t) \\ &\quad + \widehat{L}_{\dot{x}\dot{x}}(t)\dot{\overline{h}}(t)\Big)\dot{\overline{h}}(t)\,dt = \int_{t_0}^{\tau} \Big(-\frac{d}{dt}\Big(\widehat{L}_{\dot{x}x}(t)\overline{h}(t) + \widehat{L}_{\dot{x}\dot{x}}(t)\dot{\overline{h}}(t)\Big) \\ &\quad + \widehat{L}_{xx}(t)\overline{h}(t) + \widehat{L}_{x\dot{x}}(t)\dot{\overline{h}}(t)\Big)\overline{h}(t)\,dt. \end{aligned}$$

Поскольку $\overline{h}$ удовлетворяет уравнению Якоби, то отсюда следует, что $Q(\widetilde{h}) = 0$. Это означает, что наряду с $\widehat{h} = 0$, функция $\widetilde{h}$ также доставляет сильный минимум в задаче (66). Запишем эту задачу как задачу оптимального управления

$$\int_{t_0}^{t_1} \left(\widehat{L}_{xx}(t)h^2(t) + 2\widehat{L}_{\dot{x}x}(t)h(t)u(t) + \widehat{L}_{\dot{x}\dot{x}}(t)u^2(t)\right) dt \to \min,$$

$$\dot{h} = u, \quad h(t_0) = h(t_1) = 0.$$

Согласно принципу максимума найдутся такие множители Лагранжа λ_0 и $p \in PC^1([t_0, t_1])$, не равные одновременно нулю, что выполнено условие стационарности по h

$$-\dot{p}(t) + 2\lambda_0\widehat{L}_{xx}(t)\widetilde{h}(t) + 2\lambda_0\widehat{L}_{\dot{x}x}(t)\dot{\widetilde{h}}(t) = 0$$

и условие минимума по u

$$\begin{aligned}(67)\quad &\min_{u\in\mathbb{R}}(2\lambda_0\widehat{L}_{\dot{x}x}(t)\widetilde{h}(t)u + \lambda_0\widehat{L}_{\dot{x}\dot{x}}(t)u^2 - p(t)u) \\ &= 2\lambda_0\widehat{L}_{\dot{x}x}(t)\widetilde{h}(t)\dot{\widetilde{h}}(t) + \lambda_0\widehat{L}_{\dot{x}\dot{x}}(t)\dot{\widetilde{h}}^2(t) - p(t)\dot{\widetilde{h}}(t).\end{aligned}$$

Как и раньше проверяется, что $\lambda_0 \neq 0$. Пусть $\lambda_0 = 1/2$.

Из (67) следует, что для каждого $t \in [t_0, t_1]$ дифференцируемая на $\mathbb{R}$ функция

$$f(u) = \widehat{L}_{\dot{x}x}(t)\widetilde{h}(t)u + (1/2)\widehat{L}_{\dot{x}\dot{x}}(t)u^2 - p(t)u$$

достигает минимума в точке $\dot{\widetilde{h}}$. Следовательно, по теореме Ферма, ее производная в этой точке равна нулю

$$p(t) = \widehat{L}_{\dot{x}x}(t)\widetilde{h}(t) + \widehat{L}_{\dot{x}\dot{x}}(t)\dot{\widetilde{h}}(t). \tag{68}$$

По определению $\widetilde{h}(t) = 0$, если $t \geq \tau$ и поэтому из (67) вытекает, что $p(\tau + 0) = 0$. Но функция p непрерывна и поэтому (снова из (67)) получаем $0 = p(\tau - 0) = \widehat{L}_{\dot{x}\dot{x}}(\tau)\dot{\widetilde{h}}(\tau - 0) = \widehat{L}_{\dot{x}\dot{x}}(\tau)\dot{\overline{h}}(\tau) \neq 0$, так как $\dot{\overline{h}}(\tau) \neq 0$ (как уже было отмечено) и $\widehat{L}_{\dot{x}\dot{x}}(\tau) > 0$ в силу того, что выполнено усиленное условие Лежандра. Пришли к противоречию и тем самым условие Якоби доказано. □

Список литературы

[1] Ильин В. А., Позняк Э. Г. Основы математического анализа: В 2-х ч. Часть I. М: Физматлит, 2005.

[2] Колмогоров А. Н., Фомин С. В. Элементы теории функций и фукционального анализа. М.: ФИЗМАТЛИТ, 2004.

[3] Алексеев В. М., Тихомиров В. М., Фомин С. В. Оптимальное управление. М.: Наука, 1979.