

**Вариационное исчисление и оптимальное
управление**

К. Ю. Осипенко

Оглавление

Предисловие	5
Введение	5
Глава 1. Экстремальные задачи без ограничений	7
1. Экстремальные задачи, их формализация	7
2. Производная отображения	7
3. Теорема о среднем	10
4. Теорема Ферма для гладких задач без ограничений	10
5. Вторая производная отображения	11
6. Теорема о смешанных производных	12
7. Формула Тейлора	13
8. Необходимые и достаточные условия экстремума второго порядка	14
Глава 2. Вспомогательные результаты	16
1. Строгая дифференцируемость	16
2. Теорема о суперпозиции	16
3. Теорема о полном дифференциале. Оператор Немыцкого	18
4. Производная оператора Немыцкого	19
5. Обобщенный оператор Немыцкого	19
6. Лемма о правом обратном и лемма о замкнутости образа	20
7. Теорема о неявной функции	20
8. Теорема Люстерника	22
9. Теоремы отделимости	22
10. Леммы об аннуляторах	23
Глава 3. Экстремальные задачи с ограничениями	25
1. Правило множителей Лагранжа для гладких задач с ограничениями типа равенств	25
2. Условия экстремума второго порядка для гладких задач с ограничениями типа равенств	26
3. Гладкие задачи с ограничениями типа равенств и неравенств	28
Глава 4. Выпуклые экстремальные задачи	31
1. Выпуклые задачи без ограничений	31
2. Субдифференциал. Теорема Ферма	32
3. Выпуклые задачи с ограничениями. Теорема Каруша–Куна–Таккера	33
Глава 5. Вариационное исчисление	35
1. Простейшая задача вариационного исчисления. Уравнение Эйлера	35
2. Задача Больца	37
3. Интегралы уравнения Эйлера	38
4. Задача Лагранжа. Общая постановка	38
5. Задача со старшими производными. Уравнение Эйлера–Пуассона	42
6. Изопериметрическая задача	43

Оглавление	4
Глава 6. Оптимальное управление	45
1. Задачи оптимального управления	45
2. Принцип максимума	45
3. Доказательство принципа максимума	48
4. Необходимые условия сильного экстремума в простейшей задаче вариационного исчисления. Условие Вейрштрасса и Лежандра	48
5. Необходимые условия слабого экстремума в простейшей задаче вариационного исчисления. Условие Якоби	51
6. Теория поля и достаточные условия сильного экстремума в простейшей задаче вариационного исчисления	53
7. Избранные задачи	56
Литература	61

Предисловие

В основе предлагаемого пособия лежат лекции, который автор читал на механико-математическом факультете МГУ. В пособии подробно излагается метод множителей Лагранжа и основные задачи оптимального управления, поэтому оно может быть полезно студентам МФТИ, изучающим курсы “Методы оптимизации” и “Методы оптимального управления”. Автор существенно использовал в пособии книгу В. М. Алексеева, В. М. Тихомирова, С. В. Фомина [1] и конспекты лекций Г. Г. Магарил-Ильяева. Своим приятным долгом автор считает выражение глубокой благодарности за ценные обсуждения и советы своим коллегам профессорам Э. М. Галееву, Г. Г. Магарил-Ильяеву и В. М. Тихомирову.

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, по-своему. 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”.

Экстремальные задачи без ограничений

1. 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

$$(1) \quad f(x) \rightarrow \min(\max), \quad x \in C,$$

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 ext и говорим о задаче на экстремум функционала f .

Отметим еще, что если \hat{x} — решение задачи (1) на минимум (максимум), то ясно, что \hat{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 $\hat{x} \in C$ называется *локальным минимумом (максимумом)* в задаче (1), если существует такая ее окрестность U , что $f(x) \geq f(\hat{x})$ ($f(x) \leq f(\hat{x})$) для всех допустимых $x \in U$ (т. е. для всех $x \in C \cap U$).

2. Производная отображения

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 $\Lambda: X \rightarrow Y$ обозначим через $\mathcal{L}(X, Y)$. Это нормированное пространство с нормой

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

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 , $j = 1, \dots, d$, — *координатами вектора x* . Для экономии места, элементы \mathbb{R}^d будем записывать так $x = (x_1, \dots, x_d)^T$, 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 $x = (x_1, \dots, x_d)^T \in \mathbb{R}^d$. Величина

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

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 $a = (a_1, \dots, a_d)$ — вектор-строка из d действительных чисел. Для каждого $x = (x_1, \dots, x_d)^T \in \mathbb{R}^d$ 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 $a = (l(e_1), \dots, l(e_d))$, где

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

— *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 $F: U \rightarrow Y$ называется дифференцируемым в точке $\hat{x} \in U$, если найдется такой оператор $\Lambda \in \mathcal{L}(X, Y)$, что для всех $h \in X$, для которых $\hat{x} + h \in U$ справедливо представление

$$(2) \quad F(\hat{x} + h) = F(\hat{x}) + \Lambda h + r(h),$$

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

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

Рассмотрим некоторые примеры.

Пример 2.1. Пусть U — окрестность точки $\hat{x} \in \mathbb{R}^d$. Тогда функция $f: U \rightarrow \mathbb{R}$ дифференцируема в точке \hat{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 $\hat{x} + h \in U$ справедливо представление

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

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

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

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

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

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

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

Пример 2.2. Пусть теперь $U \subset \mathbb{R}^{d_1}$ и задано отображение $F: U \rightarrow \mathbb{R}^{d_2}$. 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) $\Lambda = (a_{ij})_{1 \leq i \leq d_2, 1 \leq j \leq d_1}$ 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 $\hat{x} \in U$, 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 $\hat{x} + h \in U$, справедливо представление

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

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

Отображение $F: U \rightarrow \mathbb{R}^{d_2}$ можно записать в виде

$$F(x) = (f_1(x), \dots, f_{d_2}(x))^T,$$

где $f_j: U \rightarrow \mathbb{R}$, $j = 1, \dots, d_2$ ($f_j(x)$ — это j -ая координата вектора $F(x)$ в стандартном базисе в \mathbb{R}^{d_2}). Легко проверить, что F дифференцируемо в точке \hat{x} тогда и только тогда, когда функции f_j , $j = 1, \dots, d_2$, дифференцируемы в \hat{x} . При этом строки матрицы $F'(\hat{x})$ являются векторами $f'_1(\hat{x}), \dots, f'_{d_2}(\hat{x})$. Производную $F'(\hat{x})$ называют *матрицей Якоби* отображения F в точке \hat{x} . Тем самым

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

3. Теорема о среднем

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

ТЕОРЕМА 1 (о среднем). Пусть X, Y — нормированные пространства, U — открытое множество в X , отображение $F: U \rightarrow Y$ 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. [3, стр. 195]) найдется элемент $y^* \in Y^*$ такой, что $\|y^*\|_{Y^*} = 1$ и

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

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),

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

□

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

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

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 $f: U \rightarrow \mathbb{R}$. Рассмотрим задачу

$$(5) \quad f(x) \rightarrow \text{extr}, \quad x \in U.$$

ТЕОРЕМА 2 (Ферма). Если \hat{x} — локальный экстремум в задаче (5) и функция f дифференцируема в \hat{x} , то

$$(6) \quad f'(\hat{x}) = 0.$$

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

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

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

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

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

5. Вторая производная отображения

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

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

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

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

Пространство $\mathcal{L}(X, \mathcal{L}(X, Y))$ изометрически изоморфно пространству $\mathcal{L}^2(X, Y)$ всех непрерывных билинейных отображений $B: X \times X \rightarrow Y$ 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 $B: X \times X \rightarrow Y$, 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 $h, \xi, \eta \in X$

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

Пример 5.2. Пусть U — окрестность точки $\hat{x} \in \mathbb{R}^d$ и функция $f: U \rightarrow \mathbb{R}$ дважды дифференцируема во всех точках из U . Тогда, как было показано в примере 2.1,

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

аналогично примеру 2.2 (здесь $f': U \rightarrow (\mathbb{R}^d)^*$) получаем, что если функция f' дифференцируема в точке \hat{x} , то функции $\frac{\partial f}{\partial x_j}$, $j = 1, \dots, d$, дифференцируемы в точке \hat{x} и

$$(8) \quad f''(\hat{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\hat{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\hat{x}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_d}(\hat{x}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\hat{x}) & \frac{\partial^2 f}{\partial x_2^2}(\hat{x}) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_d}(\hat{x}) \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_d \partial x_1}(\hat{x}) & \frac{\partial^2 f}{\partial x_d \partial x_2}(\hat{x}) & \dots & \frac{\partial^2 f}{\partial x_d^2}(\hat{x}) \end{pmatrix}.$$

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

Известно (см., например, [2, 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.

6. Теорема о смешанных производных

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

ТЕОРЕМА 3 (о смешанных производных). *Если отображение $F: U \rightarrow Y$ дважды дифференцируемо в точке $\hat{x} \in U$, то для всех $\xi, \eta \in X$*

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

ДОКАЗАТЕЛЬСТВО. Из определения второй производной для x достаточно близких к \hat{x}

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

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

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

Имеем

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

При $x = \hat{x}$ получаем

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

Таким образом,

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

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

$$(10) \quad \begin{aligned} \|\varphi'(x) - \varphi'(\hat{x})\| &\leq \varepsilon(\|x - \hat{x}\|_X + \|\eta\|_X) + \varepsilon\|x - \hat{x}\|_X + \varepsilon\|\eta\|_X \\ &= 2\varepsilon(\|x - \hat{x}\|_X + \|\eta\|_X). \end{aligned}$$

Для достаточно малых η и ξ положим

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

Из (9) имеем

$$\varphi'(\widehat{x})\xi = F''(\widehat{x})[\eta, \xi] + \alpha(\widehat{x} + \eta)\xi\|\eta\|_X.$$

Отсюда

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

Используя (4), получаем

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

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

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

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

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

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

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

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

7. Формула Тейлора

ТЕОРЕМА 4 (Формула Тейлора). Пусть X и Y — нормированные пространства, U — окрестность точки $\widehat{x} \in X$. Если отображение $F: U \rightarrow Y$ 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. $\|r(h)\|_Y / \|h\|_X^2 \rightarrow 0$ при $h \rightarrow 0$.

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

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

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

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

\square

8. Необходимые и достаточные условия экстремума второго порядка

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

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

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

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

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

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 $f''(\hat{x})$ — матрица Гесса (8). Тем самым из доказанной теоремы вытекает

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

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

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 \hat{x} , $f'(\hat{x}) = 0$ и существует $\alpha > 0$ такое, что для всех $h \in X$*

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

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

ДОКАЗАТЕЛЬСТВО. Предположим для определенности, что

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

Тогда для достаточно малых h по формуле Тейлора

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

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

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

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

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

8. НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ ЭКСТРЕМУМА ВТОРОГО ПОРЯДКА⁵

ДОКАЗАТЕЛЬСТВО. Пусть $h^T f''(\hat{x})h > 0$ для любого ненулевого $h \in \mathbb{R}^d$. Функция $h^T f''(\hat{x})h$ 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

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

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

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

Теперь утверждение следствия непосредственно вытекает из теоремы 6. \square

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).

Вспомогательные результаты

1. Строгая дифференцируемость

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

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

Отображение $F: U \rightarrow Y$, где U — открытое подмножество X , называется *строго дифференцируемым в точке* $\hat{x} \in U$, 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 $x_1, x_2 \in B_X(\hat{x}, \delta)$ справедливо неравенство

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

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

ПРЕДЛОЖЕНИЕ 1. Пусть X, Y — нормированные пространства, U — открытое подмножество X и отображение $F: U \rightarrow Y$ непрерывно дифференцируемо в точке $\hat{x} \in U$. Тогда F строго дифференцируемо в \hat{x} .

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

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

т. е. F строго дифференцируемо в \hat{x} . \square

2. Теорема о суперпозиции

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

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

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

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

$$(12) \quad \|\varphi(x_1) - \varphi(x_2) - L(x_1 - x_2)\|_Y \leq \varepsilon_1 \|x_1 - x_2\|_X,$$

$$(13) \quad \|\psi(y_1) - \psi(y_2) - M(y_1 - y_2)\|_Z \leq \varepsilon_1 \|y_1 - y_2\|_Y.$$

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 (12) и (13). Положим

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

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

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

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

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

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

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

Докажем теперь утверждение теоремы для случая, когда ψ дифференцируемо в точке \hat{y} и φ дифференцируемо в точке \hat{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

$$(15) \quad \|\varphi(\hat{x} + \Delta x) - \varphi(\hat{x}) - L\Delta x\|_Y \leq \varepsilon_1 \|\Delta x\|_X,$$

$$(16) \quad \|\psi(\hat{y} + \Delta y) - \psi(\hat{y}) - M\Delta y\|_Z \leq \varepsilon_1 \|\Delta y\|_Y.$$

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

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

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

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

□

3. Теорема о полном дифференциале. Оператор Немыцкого

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

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

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

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

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

Имеем

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

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

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

□

Пусть G открытое подмножество $\mathbb{R} \times \mathbb{R}^{d_1}$ и $f: G \rightarrow \mathbb{R}^{d_2}$ — 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 $F: U \rightarrow C([t_0, t_1], \mathbb{R}^{d_2})$, определенное по правилу

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

называется *оператором Немыцкого*.

4. Производная оператора 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 $\hat{x} \in U$. Существует $\delta_0 > 0$ такое, что компакт

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

принадлежит G . Пусть $\varepsilon > 0$. Функция f_x равномерно непрерывна на K и поэтому найдется $0 < \delta \leq \delta_0$ 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 $g: B_{\mathbb{R}^{d_1}}(\hat{x}(t), \delta) \rightarrow \mathbb{R}^{d_2}$, $g(x) = f(t, x) - f_x(t, \hat{x}(t))x$, дифференцируемо на $B_{\mathbb{R}^{d_1}}(\hat{x}(t), \delta)$ и его производная в точке x имеет вид $g'(x) = f_x(t, x) - f_x(t, \hat{x}(t))$. Пусть $x_j \in B_{C([t_0, t_1], \mathbb{R}^{d_1})}(\hat{x}, \delta)$, $j = 1, 2$. Тогда $x_j(t) \in B_{\mathbb{R}^{d_1}}(\hat{x}(t), \delta)$, $j = 1, 2$, и мы имеем по теореме о среднем, примененной к отображению g (учитывая, что если $x \in [x_1(t), x_2(t)]$, то $x \in B_{\mathbb{R}^{d_1}}(\hat{x}(t), \delta)$)

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

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

Докажем непрерывную дифференцируемость отображения F на U . В силу равномерной непрерывности f_x на компакте K для любого $\varepsilon > 0$ найдется $0 < \delta \leq \delta_0$ такое, что если $|x_1 - x_2| < \delta$, то $\|f_x(t, x_1) - f_x(t, x_2)\| < \varepsilon$ для всех $(t, x_j) \in K$, $j = 1, 2$. Пусть $h \in C([t_0, t_1], \mathbb{R}^{d_1})$ и $x \in B_{C([t_0, t_1], \mathbb{R}^{d_1})}(\hat{x}, \delta)$. Тогда $x(t) \in B_{\mathbb{R}^{d_1}}(\hat{x}(t), \delta)$ для любого $t \in [t_0, t_1]$. Следовательно, для любого $t \in [t_0, t_1]$

$$\begin{aligned} |F'(x)h(t) - F'(\hat{x})h(t)| &= |f_x(t, x(t))h(t) - f_x(t, \hat{x}(t))h(t)| \\ &\leq \|f_x(t, x(t)) - f_x(t, \hat{x}(t))\| |h(t)| \leq \varepsilon |h(t)|. \end{aligned}$$

Отсюда вытекает непрерывность F' для любой функции $\hat{x} \in U$. \square

Замечание. Предложение 2 остается справедливым, если U заменить на

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

5. Обобщенный оператор Немыцкого

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 $f: G \rightarrow \mathbb{R}^{d_2}$ — 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 U открыто в $C^1([t_0, t_1], \mathbb{R}^{d_1}) \times C([t_0, t_1], \mathbb{R}^{d_3})$. Определим отображение $F: U \rightarrow C([t_0, t_1], \mathbb{R}^{d_2})$ 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 , учитывая замечание, сделанное после доказательства предложения 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

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

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

6. Лемма о правом обратном и лемма о замкнутости образа

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

ДОКАЗАТЕЛЬСТВО. По теореме Банаха об открытом отображении [3, 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 $y \neq 0$, то

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

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

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

ДОКАЗАТЕЛЬСТВО. Пусть $(y, z) \in \text{cl Im } C$ и пусть $\{x_n\}$ — последовательность в X такая, что $y = \lim_{n \rightarrow \infty} Ax_n$ и $z = \lim_{n \rightarrow \infty} Bx_n$. Положим $h_n = R(Bx_n - z)$, где R — правый обратный к B , тогда $B(x_n - h_n) = z$. Так как $\|h_n\|_X = \|R(Bx_n - z)\|_X \leq \gamma \|Bx_n - z\|_Y$, то $h_n \rightarrow 0$ при $n \rightarrow \infty$ и поэтому $\lim_{n \rightarrow \infty} A(x_n - h_n) = y$. 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 $\bar{x} \in X$, что $B\bar{x} = z$ и $y = A\bar{x}$, т. е. $(y, z) \in \text{Im } C$. \square

7. Теорема о неявной функции

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

- 1) $F(\hat{x}, \hat{\sigma}) = 0$;
- 2) F непрерывно в точке $(\hat{x}, \hat{\sigma})$;

- 3) F дифференцируемо по x в точке $(\hat{x}, \hat{\sigma})$ и для каждого $\varepsilon > 0$ существуют окрестности $U(\varepsilon) \subset U$ и $V(\varepsilon)$ точек \hat{x} и $\hat{\sigma}$ такие, что для всех $x, x' \in U(\varepsilon)$ и $\sigma \in V(\varepsilon)$ выполняется соотношение

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

- 4) $\text{Im } F_x(\hat{x}, \hat{\sigma}) = Y$,

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

ДОКАЗАТЕЛЬСТВО. Обозначим для краткости $\Lambda = F_x(\hat{x}, \hat{\sigma})$. Так как $\text{Im } F_x(\hat{x}, \hat{\sigma}) = Y$, то по лемме о правом обратном (лемма 1) существует отображение $R: Y \rightarrow X$ и константа $\gamma > 0$ такие, что $\Lambda R(y) = y$ и $\|R(y)\|_X \leq \gamma \|y\|_Y$ 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 \hat{x} и $\hat{\sigma}$, соответствующие ε_0 (из формулировки теоремы). Пусть $\delta > 0$ такое, что $B_X(\hat{x}, \delta) \subset U(\varepsilon_0)$. Выберем окрестности U_0 и V_0 так, что $U_0 \subset B_X(\hat{x}, \delta/2)$, $V_0 \subset V(\varepsilon_0)$ 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

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

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

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

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

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

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

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

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

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

$$\begin{aligned} \|x_{n+m} - x_n\|_X &\leq \|x_{n+m} - x_{n+m-1}\|_X + \dots + \|x_{n+1} - x_n\|_X \leq \\ &\leq (\theta^{n+m-1} + \dots + \theta^n) \|x_1 - x\|_X \leq \frac{\theta^n}{1 - \theta} \|x_1 - x\|_X \leq \frac{\delta}{2} \theta^n. \end{aligned}$$

Положим $\varphi(x, \sigma) = \lim_{n \rightarrow \infty} x_n$. Из (21) следует, что $\varphi(x, \sigma) \in B_X(\hat{x}, \delta) \subset U$. 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 $n \rightarrow \infty$, получаем, что $F(\varphi(x, \sigma), \sigma) = 0$.

В (21) доказано, что

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

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

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

СЛЕДСТВИЕ 4. Пусть X и Y — банаховы пространства, U — окрестность точки $\hat{x} \in X$ и $F: U \rightarrow Y$. Если F строго дифференцируемо в точке \hat{x} и $\text{Im } F'(\hat{x}) = Y$, то найдется окрестность $U_0 \subset U$ точки \hat{x} , отображение $\varphi: U_0 \rightarrow U$ и константа $K > 0$ такие, что $F(\varphi(x)) = F(\hat{x})$ и $\|\varphi(x) - x\|_X \leq K \|F(x) - F(\hat{x})\|_Y$ для всех $x \in U_0$.

8. Теорема Люстерника

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

ТЕОРЕМА 10 (Люстерника). Пусть X, Y — банаховы пространства, U — окрестность точки $\hat{x} \in X$, отображение $F: U \rightarrow Y$ — строго дифференцируемо в \hat{x} , $\text{Im } F'(\hat{x}) = Y$ и $M = \{x \in U : F(x) = F(\hat{x})\}$. Тогда $T_{\hat{x}}M = \text{Ker } F'(\hat{x})$.

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

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

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

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

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

т. е. h — касательный вектор. \square

9. Теоремы отделимости

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

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

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

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

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

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. [3, стр. 243]).

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

Отсюда следует

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

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

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

Но

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

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

10. Леммы об аннуляторах

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 $\hat{x} \notin L$. Множество L , очевидно, выпукло и по условию замкнуто, поэтому по второй теореме отделимости найдется ненулевой функционал $x^* \in X^*$ такой, что

$$(22) \quad \sup_{x \in L} \langle x^*, x \rangle < \langle x^*, \hat{x} \rangle.$$

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). \square

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

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

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

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

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

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

замкнут по лемме о замкнутости образа (лемма 2) и не совпадает с $\mathbb{R} \times Y$, так как $(1, 0) \notin \text{Im } M$. 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 $\text{Im } A = Y$. Таким образом,

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

для всех $x \in X$. Следовательно,

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

□

Экстремальные задачи с ограничениями

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

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

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

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

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

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

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

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

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

Пусть теперь $\text{Im } F'(\hat{x}) \neq Y$. Так как по условию подпространство $\text{Im } F'(\hat{x})$ замкнуто, то по лемме о нетривиальности аннулятора (лемма 3) существует ненулевой функционал $y^* \in Y^*$ такой, что $\langle y^*, F'(\hat{x})x \rangle = 0$ для любого $x \in X$, т. е. $(F'(\hat{x}))^* y^* = 0$. 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 $f_j: \mathbb{R}^{d_1} \rightarrow \mathbb{R}$, $j = 1, \dots, d_2$, т. е. $F(x) = (f_1(x), \dots, f_{d_2}(x))^T$, $x \in \mathbb{R}^{d_1}$. Таким образом, рассматривается задача

$$(25) \quad f_0(x) \rightarrow \text{ext}, \quad f_j(x) = 0, \quad j = 1, \dots, d_2.$$

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

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

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

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

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

$$\mathcal{L}_x(\hat{x}, \bar{\lambda}) = 0 \quad \Leftrightarrow \quad \sum_{j=0}^{d_2} \lambda_j f'_j(\hat{x}) = 0.$$

Если векторы $f'_1(\hat{x}), \dots, f'_{d_2}(\hat{x})$ линейно независимы, то $\lambda_0 \neq 0$.

Доказательство сразу следует из предыдущей теоремы, если учесть, что подпространство $\text{Im } F'(\hat{x})$ конечномерно и поэтому замкнуто, а линейная независимость векторов $f'_1(\hat{x}), \dots, f'_{d_2}(\hat{x})$ эквивалентна условию $\text{Im } F'(\hat{x}) = Y$.

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

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

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 $\hat{x} = (0, 0)$ 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.

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

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

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

и для всех $h \in \text{Ker } F'(\hat{x})$

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

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

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

$$\begin{aligned} 0 \leq f(\hat{x} + th + r(t)) - f(\hat{x}) &= \mathcal{L}(\hat{x} + th + r(t), 1, y^*) - \mathcal{L}(\hat{x}, 1, y^*) \\ &= \frac{1}{2} \mathcal{L}_{xx}(\hat{x}, 1, y^*)[th + r(t), th + r(t)] + o(t^2) \\ &= \frac{t^2}{2} \mathcal{L}_{xx}(\hat{x}, 1, y^*)[h, h] + o(t^2), \end{aligned}$$

откуда следует (27). \square

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 \hat{x} и $\text{Im } F'(\hat{x})$ — замкнутое подпространство в Y . Тогда если найдутся множитель Лагранжа $y^* \in Y^*$ и число $\alpha > 0$ такие, что

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

и при всех $h \in \text{Ker } F'(\hat{x})$

$$(29) \quad \mathcal{L}_{xx}(\hat{x}, 1, y^*)[h, h] \geq \alpha \|h\|_X^2 \quad (\leq -\alpha \|h\|_X^2),$$

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

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

$$(30) \quad F'(\hat{x})(\varphi(x)) = 0$$

и

$$(31) \quad \|\varphi(x) - x\|_X \leq K \|F'(\hat{x})x\|_Y$$

для всех $x \in U_0$.

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

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

отсюда следует, что для достаточно малых x справедливо неравенство

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

а тогда из (31) получаем, что

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

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

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

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

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

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

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

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

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

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

3. Гладкие задачи с ограничениями типа равенств и неравенств

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

$$(32) \quad f_0(x) \rightarrow \min, \quad f_j(x) \leq 0, \quad 1 \leq j \leq m, \quad F(x) = 0,$$

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*

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

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

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

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

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

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

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

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

В) Невырожденный случай: $\text{Im } F'(\hat{x}) = Y$. Рассмотрим множество

$$C = \{ ((\mu_0, \mu_1, \dots, \mu_m)^T, y) \in \mathbb{R}^{m+1} \times Y : \exists x \in X : \mu_j > \langle f'_j(\hat{x}), x \rangle, \\ j = 0, 1, \dots, m, y = F'(\hat{x})x \}.$$

Очевидно, что C — выпуклое множество. Докажем, что $0 \notin C$. Предположим, что $0 \in C$. Тогда существует такое $x_0 \in X$, что $\langle f'_j(\hat{x}), x_0 \rangle < 0$, $j = 0, 1, \dots, m$, и $F'(\hat{x})x_0 = 0$. По теореме Люстерника $x_0 \in T_{\hat{x}}M$, где $M = \{x \in X : F(x) = F(\hat{x}) = 0\}$, т. е. существуют $\varepsilon > 0$ и отображение $r: (-\varepsilon, \varepsilon) \rightarrow X$ такие, что $F(\hat{x} + tx_0 + r(t)) = 0$ для всех $t \in (-\varepsilon, \varepsilon)$ и $\|r(t)\|_X = o(t)$ при $t \rightarrow 0$. В силу дифференцируемости функций f_j , $j = 0, 1, \dots, m$, в точке \hat{x} имеем

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

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

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

$$C_0 = \{ ((\mu_0, \mu_1, \dots, \mu_m)^T, y) \in \mathbb{R}^{m+1} \times Y : \mu_j > d, j = 0, 1, \dots, m, \\ y \in F'(\hat{x})(B_X(0, 1)) \}, \quad d = \max_{0 \leq j \leq m} \|f'_j(\hat{x})\|.$$

По теореме Банаха об открытом отображении [3, стр. 243] множество $F'(\hat{x})(B_X(0, 1))$ открыто. Тем самым C_0 — открытое множество. Покажем, что $C_0 \subset C$. Действительно, пусть $((\mu_0, \mu_1, \dots, \mu_m)^T, y) \in C_0$ и $x \in B_X(0, 1)$ такое, что $y = F'(\hat{x})x$. Тогда $\mu_j > d \geq \langle f'_j(\hat{x}), x \rangle$, $j = 0, \dots, m$, и значит, $((\mu_0, \mu_1, \dots, \mu_m)^T, y) \in C$, т. е. $\text{int } C \neq \emptyset$.

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

$$(33) \quad \sum_{j=0}^m \lambda_j \mu_j + \langle y^*, y \rangle \geq 0$$

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

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

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

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

$$((\langle f'_0(\hat{x}), x \rangle + \varepsilon, \dots, \langle f'_m(\hat{x}), x \rangle + \varepsilon)^T, F'(\hat{x})x) \in C.$$

Из (33), получаем, что

$$\sum_{j=0}^m \lambda_j \langle f'_j(\hat{x}), x \rangle + \langle y^*, F'(\hat{x})x \rangle \geq -\varepsilon \sum_{j=0}^m \lambda_j.$$

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 $1 \leq j \leq m$, $f_j(\hat{x}) \neq 0$, то из c) следует, что $\lambda_j = 0$. Тем самым из (a) вытекает, что

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

В силу b) и того, что $\langle f'_j(\hat{x}), h \rangle < 0$, $j \in J_0$, получаем, что $\lambda_j = 0$, $j = 0, 1, \dots, m$. Следовательно, из a) вытекает, что $\langle y^*, F'(\hat{x})x \rangle = 0$ для всех $x \in X$. Так как $\text{Im } F'(\hat{x}) = Y$, 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

Выпуклые экстремальные задачи

1. Выпуклые задачи без ограничений

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

$$\begin{aligned} \text{dom } f &= \{x \in X : f(x) < +\infty\}, \\ \text{epi } f &= \{(x, \alpha) \in X \times \mathbb{R} : \alpha \geq f(x), x \in \text{dom } f\} \end{aligned}$$

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

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 $f: X \rightarrow \overline{\mathbb{R}}$ 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 $0 \leq \alpha \leq 1$ выполняется неравенство

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

которое называется *неравенством Йенсена*.

ТЕОРЕМА 18. Пусть X — линейное нормированное пространство и $f: X \rightarrow \mathbb{R}$ — дважды дифференцируема на X . Тогда f — выпуклая функция тогда и только тогда, когда $f''(x)[h, h] \geq 0$ 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

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

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 . \square

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

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

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

для всех $h \in \mathbb{R}^d$ (матрица гессиана в любой точке является неотрицательно определенной).

2. Субдифференциал. Теорема Ферма

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

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

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

ПРЕДЛОЖЕНИЕ 3. Пусть X — линейное нормированное пространство и $f: X \rightarrow \overline{\mathbb{R}}$ — выпуклая функция, дифференцируемая в точке \hat{x} . Тогда $\partial f(\hat{x}) = \{f'(\hat{x})\}$.

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

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

откуда

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

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

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

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

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

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

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

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

$$(34) \quad f(x) \rightarrow \min, \quad x \in X.$$

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

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

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 \hat{x} — локальный минимум, т. е. существует такая окрестность U точки \hat{x} , что $f(\hat{x}) \leq f(x)$ для всех $x \in U$. Пусть теперь x — произвольная точка из X . Для достаточно малых $0 < \alpha \leq 1$ точки $(1 - \alpha)\hat{x} + \alpha x$ принадлежат U и поэтому (по неравенству Йенссена) $f(\hat{x}) \leq f((1 - \alpha)\hat{x} + \alpha x) \leq (1 - \alpha)f(\hat{x}) + \alpha f(x)$, откуда следует, что $f(\hat{x}) \leq f(x)$.

Из предложения 3 и теоремы 19 вытекает

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

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

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

$$(35) \quad f_0(x) \rightarrow \min, \quad f_j(x) \leq 0, \quad j = 1, \dots, m, \quad x \in A$$

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

Свяжем с задачей (35) следующую функцию Лагранжа

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

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

ТЕОРЕМА 20 (Каруша–Куна–Таккера). 1. *Если \hat{x} — минимум в задаче (35), то найдется такой ненулевой набор множителей Лагранжа $\bar{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_m)$, что выполнены следующие условия*

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

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

3. *Если найдется точка $\bar{x} \in A$ такая, что $f_j(\bar{x}) < 0$, $1 \leq j \leq m$, то в 1 $\lambda_0 \neq 0$ (условие Слейтера).*

ДОКАЗАТЕЛЬСТВО. Пусть \hat{x} — решение задачи (35). Рассмотрим множество

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

Непосредственная проверка показывает, что это множество выпукло. Кроме того, легко видеть, что оно содержит все векторы с положительными компонентами (надо взять $x = \hat{x}$) и тем самым его внутренность не пуста. Наконец, $0 \notin M$, так как в противном случае нашелся бы элемент $\bar{x} \in A$ такой, что

$f_j(\bar{x}) \leq 0$, $j = 1, \dots, m$, и $f_0(\bar{x}) - f_0(\hat{x}) < 0$, в противоречие с тем, что \hat{x} — минимум.

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

$$(36) \quad \sum_{j=0}^m \lambda_j \mu_j \geq 0$$

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

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

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

$$\begin{aligned} \lambda_0 f_0(x) &\geq \lambda_0 f_0(x) + \sum_{j=1}^m \lambda_j f_j(x) = \mathcal{L}(x, \bar{\lambda}) \geq \mathcal{L}(\hat{x}, \bar{\lambda}) \\ &= \lambda_0 f_0(\hat{x}) + \sum_{j=1}^m \lambda_j f_j(\hat{x}) = \lambda_0 f_0(\hat{x}). \end{aligned}$$

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)) $\mathcal{L}(\bar{x}, \bar{\lambda}) = \sum_{i=1}^m \lambda_i f_i(\bar{x}) < 0 = \sum_{i=1}^m \lambda_i f_i(\hat{x}) = \mathcal{L}(\hat{x}, \bar{\lambda})$, что противоречит (a). \square

Вариационное исчисление

1. 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 inveniondi lineas curvas maximive proprietate gementies 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 , $L: G \rightarrow \mathbb{R}$ — непрерывная функция переменных t, x, \dot{x} и $x_0, x_1 \in \mathbb{R}$. Задача

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

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 $\Gamma(x) = \{(t, x(t), \dot{x}(t))^T : t \in [t_0, t_1]\} \subset G$ и $x(t_j) = x_j, j = 0, 1$.

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

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

ТЕОРЕМА 21 (Необходимые условия экстремума в задаче (37)). Пусть \hat{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 $\Gamma(\hat{x})$, то $\hat{L}_{\dot{x}} \in C^1([t_0, t_1])$ и для всех $t \in [t_0, t_1]$ выполнено уравнение Эйлера

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

ДОКАЗАТЕЛЬСТВО. Пусть $x \in C^1([t_0, t_1])$ и $x(t_0) = x(t_1) = 0$. Положим $x_\alpha = \hat{x} + \alpha x$, где $\alpha \in \mathbb{R}$. Для достаточно малых α функция x_α принадлежит окрестности $\Gamma(\hat{x})$. 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 $\hat{J}(\alpha) = J(x_\alpha)$ имеет в нуле локальный экстремум. Имеет место представление

$$\hat{J}(\alpha) = I \circ F \circ D \circ S(\alpha),$$

где $S(\alpha) = \hat{x} + \alpha x$, $D(x) = (x, \dot{x})$, $F(x, u)(t) = L(t, x(t), u(t))$ и

$$I(x) = \int_{t_0}^{t_1} x(t) dt.$$

В силу дифференцируемости обобщенного оператора Немыцкого (следствие 3), теоремы о производной суперпозиции функций (теорема 7) и того, что производная линейного оператора есть сам этот линейный оператор ($I' = I$, $D' = D$), получаем

$$\left. \frac{d\hat{J}(\alpha)}{d\alpha} \right|_{\alpha=0} = I \circ F' \circ D \circ S'(\alpha) \Big|_{\alpha=0} = \int_{t_0}^{t_1} (\hat{L}_x(t)x(t) + \hat{L}_{\dot{x}}(t)\dot{x}(t)) dt.$$

По теореме Ферма производная $\hat{J}(\alpha)$ в нуле равна нулю. Следовательно,

$$(38) \quad \int_{t_0}^{t_1} (\hat{L}_x(t)x(t) + \hat{L}_{\dot{x}}(t)\dot{x}(t)) dt = 0.$$

Пусть p такая функция, что $\dot{p} = \hat{L}_{\dot{x}}$. Тогда, интегрирую по частям, получаем

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

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

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

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

$$(39) \quad \int_{t_0}^{t_1} (-p(t) + \hat{L}_{\dot{x}}(t) + c) \dot{x}(t) dt = 0.$$

Выберем c так, чтобы

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

Рассмотрим функцию

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

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

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

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

2. Задача Больца

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

$$(40) \quad \mathcal{B}(x) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt + l(x(t_0), x(t_1)) \rightarrow \text{extr}$$

называется *задачей Больца*.

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

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 $(\widehat{x}(t_0), \widehat{x}(t_1))^T$, 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 \hat{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 \hat{p} в (41) вместо \widehat{L}_x и интегрируя по частям, получаем, что

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

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 $x = (x_1, \dots, x_d)^T$. 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 $|x(t)| = \sqrt{x_1^2(t) + \dots + x_d^2(t)}$. Необходимые условия экстремума здесь имеют тот же вид и их доказательства остаются прежними. Но формулы, разумеется, надо понимать векторно. Например,

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

3. Интегралы уравнения Эйлера

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

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

который называется *интегралом импульса*.

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

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

Он называется *интегралом энергии*.

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

$$\begin{aligned} \frac{dH}{dt} &= L_{\dot{x}}(\hat{x}(t), \dot{\hat{x}}(t))\ddot{\hat{x}}(t) + \frac{d}{dt} L_{\dot{x}}(\hat{x}(t), \dot{\hat{x}}(t))\dot{\hat{x}}(t) - L_x(\hat{x}(t), \dot{\hat{x}}(t))\dot{\hat{x}}(t) \\ &\quad - L_{\dot{x}}(\hat{x}(t), \dot{\hat{x}}(t))\ddot{\hat{x}}(t) = \left(\frac{d}{dt} L_{\dot{x}}(\hat{x}(t), \dot{\hat{x}}(t)) - L_x(\hat{x}(t), \dot{\hat{x}}(t)) \right) \dot{\hat{x}}(t) = 0. \end{aligned}$$

При доказательстве мы предположили существование $\ddot{\hat{x}}$

4. 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 $L_j: G \rightarrow \mathbb{R}$, $j = 0, 1, \dots, m$, отображение $\varphi: G \rightarrow \mathbb{R}^{d_1}$ (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 $l_j: W \rightarrow \mathbb{R}$, $j = 0, 1, \dots, m$ (переменных ξ_0 и ξ_1) непрерывны на своей области определения. Задача

$$(43) \quad f_0(x, u) = \int_{t_0}^{t_1} L_0(t, x(t), u(t)) dt + l_0(x(t_0), x(t_1)) \rightarrow \min,$$

$$\dot{x} = \varphi(t, x, u),$$

$$f_j(x, u) = \int_{t_0}^{t_1} L_j(t, x(t), u(t)) dt + l_j(x(t_0), x(t_1)) \leq 0, \quad 1 \leq j \leq m',$$

$$f_j(x, u) = \int_{t_0}^{t_1} L_j(t, x(t), u(t)) dt + l_j(x(t_0), x(t_1)) = 0,$$

$$m' + 1 \leq j \leq m,$$

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

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

$$\dot{x}(t) = \varphi(t, x(t), u(t)) \text{ для всех } t \in [t_0, t_1],$$

$$f_j(x, u) \leq 0, \quad 1 \leq j \leq m', \quad f_j(x, u) = 0, \quad m' + 1 \leq j \leq m.$$

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

Функцией Лагранжа задачи (43) назовем функцию

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

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

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

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

(а) 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) $\lambda_j \geq 0$, $j = 0, 1, \dots, m'$ (условия неотрицательности);

(c) $\lambda_j f_j(\widehat{x}, \widehat{u}) = 0$, $j = 1, \dots, m'$ (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 $G(\xi) = (f_{m'+1}(\xi), \dots, f_m(\xi))^T$, $F(\xi) = \dot{x} - \varphi(t, x, u)$. Тогда задача (43) может быть переписана в виде

$$(44) \quad f_0(\xi) \rightarrow \min, \quad f_j(\xi) \leq 0, \quad 1 \leq j \leq m', \quad \Phi(\xi) = 0.$$

Будем применять к этой задаче теорему 17. Из дифференциальных свойств функций L_j и l_j , $j = 0, 1, \dots, m$, 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 $\text{Im } F'(\widehat{\xi}) = C([t_0, t_1], \mathbb{R}^{d_1})$. 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

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

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) $\text{Im } \Phi'(\widehat{\xi})$ — замкнутое подпространство.

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

$$\begin{aligned} \mathcal{L}_1(\xi, \bar{\lambda}) &= \sum_{j=0}^{m'} \lambda_j f_j(\xi) + \langle \eta^*, \Phi(\xi) \rangle \\ &= \sum_{j=0}^{m'} \lambda_j f_j(\xi) + \langle \mu^*, G(\xi) \rangle + \langle y^*, F(\xi) \rangle = \sum_{j=0}^m \lambda_j f_j(\xi) + \langle y^*, F(\xi) \rangle, \end{aligned}$$

выполняются условия стационарности $\mathcal{L}'_1(\widehat{\xi}, \bar{\lambda}) = 0$, 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

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

где

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

а также

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

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

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

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 $p(t_0) = \widehat{l}_{\xi_0}$.

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

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

□

5. Задача со старшими производными. Уравнение Эйлера–Пуассона

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

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

$$(49) \quad J(x) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t), \dots, x^{(n)}(t)) dt \rightarrow \min,$$

$$x^{(k)}(t_j) = x_j^{(k)}, \quad k = 0, 1, \dots, n-1, \quad j = 0, 1,$$

называется *задачей со старшими производными*.

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

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

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

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

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

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

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

$$\sum_{k=0}^n (-1)^k \frac{d^k}{dt^k} \hat{L}_{x^{(k)}}(t) = 0.$$

ДОКАЗАТЕЛЬСТВО. Обозначая $x = x_1$, $\dot{x}_1 = x_2$, \dots , $\dot{x}_{n-1} = x_n$, $\dot{x}_n = u$, задачу (49) можно записать как задачу Лагранжа

$$(50) \quad \int_{t_0}^{t_1} L(t, x_1(t), \dots, x_n(t), u(t)) dt \rightarrow \min,$$

$$\dot{x}_1 = x_2, \dots, \dot{x}_{n-1} = x_n, \dot{x}_n = u,$$

$$x_k(t_j) = x_j^{(k-1)}, \quad k = 1, \dots, n, \quad j = 0, 1.$$

Простая проверка показывает, что если \hat{x} — слабый локальный минимум в задаче (49), то вектор-функция $(\hat{x}_1, \dots, \hat{x}_n, \hat{u})^T$ — слабый локальный минимум в данной задаче. Согласно общей теореме 43 найдутся такие множители Лагранжа λ_0 и $p = (p_1, \dots, p_n) \in C^1([t_0, t_1], (\mathbb{R}^n)^*)$, не равные одновременно нулю, что имеют место равенства

$$-\dot{p}(t) = p(t)\hat{\varphi}_x(t) - \lambda_0\hat{L}_x(t),$$

$$p(t)\hat{\varphi}_u(t) = \lambda_0\hat{L}_u(t).$$

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

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

ДОКАЗАТЕЛЬСТВО. Рассмотрим следующую задачу Лагранжа

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

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

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

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

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

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

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

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

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

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

Оптимальное управление

1. 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 $f: G \times U \rightarrow \mathbb{R}$ и отображение $\varphi: G \times U \rightarrow \mathbb{R}^{d_1}$ (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

$$(55) \quad J(x, u) = \int_{t_0}^{t_1} f(t, x(t), u(t)) dt \rightarrow \min, \quad \dot{x} = \varphi(t, x, u),$$

$$u(t) \in U, \quad x(t_0) = x_0, \quad x(t_1) = x_1$$

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 $x(t_0) = x_0, x(t_1) = x_1$.

2. Принцип максимума

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

Функцией Лагранжа для задачи (55) будем называть функцию

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

где

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

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 (\hat{x}, \hat{u}) доставляет сильный локальный 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 $\bar{\lambda} = (\lambda_0, p) \in \mathbb{R} \times PC^1([t_0, t_1], (\mathbb{R}^{d_1})^*)$ такой, что выполнено условие стационарности по x

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

и в точках непрерывности \hat{u} условие минимума по u

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

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

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

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

$$(57) \quad J(x, u) = \int_{t_0}^{t_1} f(t, x(t), u(t)) dt \rightarrow \min, \quad \dot{x} = \varphi(t, x, u),$$

$$u(t) \in U, \quad x(t_0) = x_0.$$

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

Для доказательства теоремы нам потребуются две леммы. Далее считаем, что функции f и φ удовлетворяют условиям теоремы 26.

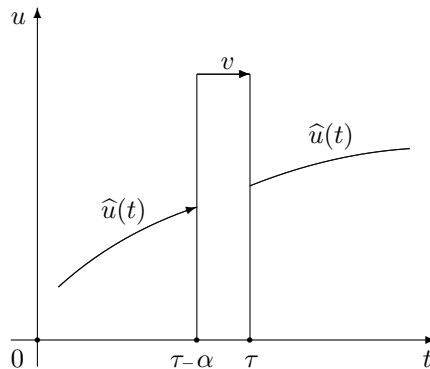


Рис. 1

Пусть (\hat{x}, \hat{u}) — допустимая пара в задаче (55) (или (57)), $\tau \in (t_0, t_1)$ — точка непрерывности \hat{u} , $\alpha > 0$ столь мало, что функция \hat{u} непрерывна на $[\tau - \alpha, \tau]$ и

$v \in U$. Положим

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

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

ЛЕММА 5 (об игольчатой вариации). *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 $[t_0, t_1]$, причем $x_\alpha(t, \tau, v) = \widehat{x}(t)$ при $t \in [t_0, \tau - \alpha]$ и, кроме того, при $\alpha \rightarrow 0$

- 1) *функция $x_\alpha(t, \tau, v) \rightarrow \widehat{x}(t)$ в метрике пространства $C([t_0, t_1], \mathbb{R}^{d_1})$;*
- 2) *функция*

$$\frac{x_\alpha(t, \tau, v) - \widehat{x}(t)}{\alpha} \rightarrow y_{\tau v}(t)$$

в метрике пространства $C([\tau, t_1], \mathbb{R}^{d_1})$, где функция $y_{\tau v}$ в точках непрерывности \widehat{u} , лежащих на отрезке $[\tau, t_1]$, удовлетворяет уравнению

$$(58) \quad \dot{y}_{\tau v} = \widehat{\varphi}_x(t) y_{\tau v}$$

с начальным условием

$$(59) \quad y_{\tau v}(\tau) = \varphi(\tau, \widehat{x}(\tau), v) - \varphi(\tau, \widehat{x}(\tau), \widehat{u}(\tau)).$$

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

$$(60) \quad J'(0+0) = \Delta_{\tau v} f + \int_{\tau}^{t_1} \widehat{f}_x(t) y_{\tau v}(t) dt,$$

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)–(59).

ДОКАЗАТЕЛЬСТВО. Имеем

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

Имеем в силу дифференцируемости f по x

$$\begin{aligned} f(t, x_\alpha(t, \tau, v), \widehat{u}(t)) - f(t, \widehat{x}(t), \widehat{u}(t)) \\ = \widehat{f}_x(t)[x_\alpha(t, \tau, v) - \widehat{x}(t)] + o(x_\alpha(t, \tau, v) - \widehat{x}(t)). \end{aligned}$$

Поэтому

$$\lim_{\alpha \rightarrow 0+0} \frac{1}{\alpha} \int_{\tau}^{t_1} (f(t, x_\alpha(t, \tau, v), \widehat{u}(t)) - f(t, \widehat{x}(t), \widehat{u}(t))) dt = \int_{\tau}^{t_1} \widehat{f}_x(t) y_{\tau v}(t) dt.$$

К первому интегралу применим теорему о среднем для интегралов и тогда в итоге получим

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

где $\xi \in [\tau - \alpha, \tau]$. Когда $\alpha \rightarrow 0+0$, то, очевидно, $\xi \rightarrow \tau$, $x_\alpha(\xi, \tau, v) \rightarrow \widehat{x}(\tau)$ согласно лемме об игольчатой вариации, а $\widehat{u}(\xi) \rightarrow \widehat{u}(\tau)$, так как \widehat{u} непрерывна в точке τ . Формула (60) доказана. \square

3. Доказательство принципа максимума

ДОКАЗАТЕЛЬСТВО ПРИНЦИПА МАКСИМУМА. Пусть $(\widehat{x}, \widehat{u})$ — сильный локальный минимум в задаче (57). Обозначая через p — решение линейной системы

$$(61) \quad -\dot{p} = p\widehat{\varphi}_x(t) - \widehat{f}_x(t), \quad p(t_1) = 0,$$

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 (60)

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

Подставим сюда вместо функции \widehat{f}_x ее выражение из (61), 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

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

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

4. 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

$$(62) \quad J(x) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt \rightarrow \min, \quad x(t_j) = x_j, \quad j = 0, 1.$$

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 множестве $G \times \mathbb{R}$, где $G \subset \mathbb{R}^2$ — открытое множество 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 (62)*, если

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

и $x(t_j) = x_j$, $j = 0, 1$, и допустимая функция \hat{x} называется *слабым локальным минимумом* в задаче (62), если существует такое $\varepsilon > 0$, что для любой допустимой функции x , для которой $\|x - \hat{x}\|_{C^1([t_0, t_1])} < \varepsilon$ выполняется неравенство $J(x) \geq J(\hat{x})$.

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 (62), 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 \hat{x} доставляет *сильный локальный минимум* в задаче (62), если существует такое $\varepsilon > 0$, что для любой допустимой функции x , для которой $\|x - \hat{x}\|_{C([t_0, t_1])} < \varepsilon$ выполняется неравенство $J(x) \geq J(\hat{x})$.

Заметим, что если функция \hat{x} доставляет сильный минимум в задаче (62) и при этом $\hat{x} \in C^1([t_0, t_1])$, то \hat{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 $\|x - \hat{x}\|_{C([t_0, t_1])} < \varepsilon$, то $J(x) \geq J(\hat{x})$. Если теперь $x \in C^1([t_0, t_1])$ и $\|x - \hat{x}\|_{C^1([t_0, t_1])} < \varepsilon$, то так как, в частности, $x \in PC^1([t_0, t_1])$ и $\|x - \hat{x}\|_{C([t_0, t_1])} \leq \|x - \hat{x}\|_{C^1([t_0, t_1])} < \varepsilon$, получаем, что $J(x) \geq J(\hat{x})$.

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

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

Пусть \hat{x} — экстремаль в задаче (62) и существует

$$\hat{L}_{\dot{x}\dot{x}}(t) = L_{\dot{x}\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)).$$

Говорят, что на функции \hat{x} выполнено *условие Лежандра*, если $\hat{L}_{\dot{x}\dot{x}}(t) \geq 0$ для всех $t \in [t_0, t_1]$ и *усиленное условие Лежандра*, если $\hat{L}_{\dot{x}\dot{x}}(t) > 0$ для всех $t \in [t_0, t_1]$.

Пусть $f: \mathbb{R} \rightarrow \mathbb{R}$ — дифференцируемая функция. Функция

$$\mathcal{E}(x, x') = f(x') - f(x) - f'(x)(x' - x)$$

называется *функцией Вейерштасса* (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 $\alpha \rightarrow 0$, получаем, что $f(x') - f(x) \geq f'(x)(x' - x)$.

Пусть L — интегрант в задаче (62). Если L — дифференцируемая функция по \dot{x} на множестве $G \times \mathbb{R}$, 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 \hat{x} выполнено *условие Вейерштрасса*, если $\mathcal{E}(t, \hat{x}(t), \dot{\hat{x}}(t), u) \geq 0$ для всех $u \in \mathbb{R}$ и $t \in [t_0, t_1]$.

ТЕОРЕМА 27 (Необходимые условия сильного минимума в задаче (62)). *Пусть функция $\hat{x} \in C^1([t_0, t_1])$ доставляет сильный локальный минимум в задаче (62). 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

$$\mathcal{E}(t, \hat{x}(t), \dot{\hat{x}}(t), u) \geq 0;$$

(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 (62) как задачу оптимального управления

$$(63) \quad \int_{t_0}^{t_1} L(t, \hat{x}(t), u(t)) dt \rightarrow \min, \quad \dot{x} = u, \quad x(t_0) = x_0, \quad x(t_1) = x_1.$$

Легко видеть, что \hat{x} доставляет сильный локальный минимум в (62) тогда и только тогда, когда пара (\hat{x}, \hat{u}) , где $\hat{u} = \dot{\hat{x}}$ является сильным локальным минимумом в (63).

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

$$(64) \quad -\dot{p}(t) + \lambda_0 \widehat{L}_x(t) = 0$$

и условие минимума по u

$$(65) \quad \min_{u \in \mathbb{R}} (\lambda_0 L(t, \hat{x}(t), u) - p(t)u) = \lambda_0 L(t, \hat{x}(t), \dot{\hat{x}}(t)) - p(t)\dot{\hat{x}}(t).$$

Если $\lambda_0 = 0$, то $p = \text{const}$ вследствие (64). Тогда из (65) следует, что эта константа обязана быть нулевой и тем самым все множители Лагранжа нулевые. Итак, $\lambda_0 \neq 0$ и можно считать, что $\lambda_0 = 1$.

Условие (65) означает, что для всех $t \in [t_0, t_1]$ функция $f(u) = L(t, \hat{x}(t), u) - p(t)u$ на \mathbb{R} достигает минимума в точке $\dot{\hat{x}}(t)$ и, следовательно, по теореме Ферма производная этой функции в данной точке равна нулю, т. е. $p(t) = \widehat{L}_{\dot{x}}(t)$. Вместе с (64) это дает уравнение Эйлера.

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

Из соотношения (65) и доказанного равенства $p(t) = \widehat{L}_{\dot{x}}(t)$ следует, что

$$L(t, \hat{x}(t), u) - L_x(t, \hat{x}(t), \dot{\hat{x}}(t))u \geq L(t, \hat{x}(t), \dot{\hat{x}}(t)) - L_x(t, \hat{x}(t), \dot{\hat{x}}(t))\dot{\hat{x}}(t)$$

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

$$L(t, \hat{x}(t), u) - L(t, \hat{x}(t), \dot{\hat{x}}(t)) - L_x(t, \hat{x}(t), \dot{\hat{x}}(t))(u - \dot{\hat{x}}(t)) \geq 0,$$

т. е. выполнено условие Вейерштрасса. \square

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

Предположим, что \hat{x} — экстремаль задачи (62) и выполнено условие:

$$(66) \quad L \text{ непрерывна вместе с } L_x, L_{\dot{x}}, L_{xx}, L_{x\dot{x}}, L_{\dot{x}x}, L_{\dot{x}\dot{x}} \\ \text{в некоторой окрестности } \Gamma(\hat{x}).$$

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

$$(67) \quad \varphi(\lambda) = \int_{t_0}^{t_1} L(t, \hat{x}(t) + \lambda h(t), \dot{\hat{x}}(t) + \lambda \dot{h}(t)) dt.$$

Тогда для достаточно малых по модулю λ

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

и дифференцируя еще раз, получаем, что

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

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

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

Уравнение Эйлера для данной задачи имеет вид

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

и называется *уравнением Якоби* исходной задачи (62).

Пусть на \hat{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 \hat{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 (62) нам понадобится один несложный технический результат, который приводим без доказательства (см. [1, стр. 69]).

ЛЕММА 7 (о скруглении углов). Пусть в задаче (62) 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 (62)). Пусть $\hat{x} \in C^1([t_0, t_1])$ доставляет слабый локальный минимум в задаче (62). Тогда, если выполнено условие (66), то для всех $t \in [t_0, t_1]$ выполнено уравнение Эйлера, условие Лежандра и в случае выполнения усиленного условия Лежандра — условие Якоби.

Доказательство. Уравнение Эйлера, как необходимое условие слабого экстремума, уже было доказано раньше. Доказательство, заключается в том, что если \hat{x} — локальный минимум, то ноль есть локальный минимум для функции φ , определенной соотношением (67) 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 $\hat{h} = 0$ есть слабый абсолютный минимум в задаче

$$(68) \quad Q(h) = \int_{t_0}^{t_1} \left(\hat{L}_{xx}(t)h^2(t) + 2\hat{L}_{x\dot{x}}(t)h(t)\dot{h}(t) + \hat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) \right) dt \rightarrow \min, \quad h(t_0) = h(t_1) = 0.$$

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

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

$$\begin{aligned} Q(\tilde{h}) &= \int_{t_0}^{\tau} \left(\hat{L}_{xx}(t)\bar{h}^2(t) + 2\hat{L}_{x\dot{x}}(t)\bar{h}(t)\dot{\bar{h}}(t) + \hat{L}_{\dot{x}\dot{x}}(t)\dot{\bar{h}}^2(t) \right) dt \\ &= \int_{t_0}^{\tau} \left(\hat{L}_{xx}(t)\bar{h}(t) + \hat{L}_{x\dot{x}}(t)\dot{\bar{h}}(t) \right) \bar{h}(t) dt + \int_{t_0}^{\tau} \left(\hat{L}_{\dot{x}\dot{x}}(t)\bar{h}(t) + \hat{L}_{x\dot{x}}(t)\dot{\bar{h}}(t) \right) \dot{\bar{h}}(t) dt \\ &= \int_{t_0}^{\tau} \left(-\frac{d}{dt} \left(\hat{L}_{\dot{x}\dot{x}}(t)\bar{h}(t) + \hat{L}_{x\dot{x}}(t)\dot{\bar{h}}(t) \right) + \hat{L}_{xx}(t)\bar{h}(t) + \hat{L}_{x\dot{x}}(t)\dot{\bar{h}}(t) \right) \bar{h}(t) dt. \end{aligned}$$

Поскольку \bar{h} удовлетворяет уравнению Якоби, то отсюда следует, что $Q(\tilde{h}) = 0$. Это означает, что наряду с $\hat{h} = 0$, функция \tilde{h} также доставляет сильный минимум в задаче (68). Запишем эту задачу как задачу оптимального управления

$$\int_{t_0}^{t_1} \left(\hat{L}_{xx}(t)h^2(t) + 2\hat{L}_{x\dot{x}}(t)h(t)u(t) + \hat{L}_{\dot{x}\dot{x}}(t)u^2(t) \right) dt \rightarrow \min, \\ \dot{h} = u, \quad h(t_0) = h(t_1) = 0.$$

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

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

и условие минимума по u

$$(69) \quad \min_{u \in \mathbb{R}} (2\lambda_0\hat{L}_{x\dot{x}}(t)\tilde{h}(t)u + \lambda_0\hat{L}_{\dot{x}\dot{x}}(t)u^2 - p(t)u) \\ = 2\lambda_0\hat{L}_{x\dot{x}}(t)\tilde{h}(t)\dot{\tilde{h}}(t) + \lambda_0\hat{L}_{\dot{x}\dot{x}}(t)\dot{\tilde{h}}^2(t) - p(t)\dot{\tilde{h}}(t).$$

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

Из (69) 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 \widetilde{h} . Следовательно, по теореме Ферма, ее производная в этой точке равна нулю

$$(70) \quad p(t) = \widehat{L}_{\dot{x}x}(t)\widetilde{h}(t) + \widehat{L}_{\dot{x}\dot{x}}(t)\dot{\widetilde{h}}(t).$$

По определению $\widetilde{h}(t) = 0$, если $t \geq \tau$ и поэтому из (70) вытекает, что $p(\tau+0) = 0$. Но функция p непрерывна и поэтому (снова из (70)) получаем $0 = p(\tau-0) = \widehat{L}_{\dot{x}x}(\tau)\dot{\widetilde{h}}(\tau-0) = \widehat{L}_{\dot{x}x}(\tau)\dot{\widetilde{h}}(\tau) \neq 0$, так как $\dot{\widetilde{h}}(\tau) \neq 0$ (как уже было отмечено) и $\widehat{L}_{\dot{x}x}(\tau) > 0$ в силу того, что выполнено усиленное условие Лежандра. Пришли к противоречию и тем самым условие Якоби доказано. \square

6. Теория поля и достаточные условия сильного экстремума в простейшей задаче вариационного исчисления

В этом параграфе рассматривается фрагмент теории поля и достаточных условий экстремума в вариационном исчислении на примере простейшей задачи.

Пусть $\{x(\cdot, \lambda)\}$ — семейство экстремалей простейшей задачи вариационного исчисления (62) (т. е. для функций $x(\cdot, \lambda) \in C^1([t_0, t_1], \mathbb{R})$ выполняется уравнение Эйлера) и параметр λ принадлежит некоторому открытому множеству в \mathbb{R} .

Пусть $\widehat{x}(\cdot)$ — фиксированный элемент данного семейства экстремалей. Будем говорить, что функция $\widehat{x}(\cdot)$ окружена полем экстремалей $\{x(\cdot, \lambda)\}$, если существует открытое множество G , содержащее множество (график $\widehat{x}(\cdot)$) $\{(t, \widehat{x}(t))^T \in \mathbb{R}^2 : t \in [t_0, t_1]\}$, что для любой точки $(\tau, \xi) \in G$ существует единственная экстремаль семейства, проходящая через эту точку, т. е. существует единственное $\lambda = \lambda(\tau, \xi)$ такое, что $x(\tau, \lambda(\tau, \xi)) = \xi$. При этом, для каждого $t \in [t_0, t_1]$ отображение $\lambda \mapsto x(t, \lambda)$ и отображение $(\tau, \xi) \mapsto \lambda(\tau, \xi)$ непрерывно дифференцируемы.

Функция $u: G \rightarrow \mathbb{R}$, $u(\tau, \xi) = \frac{d}{dt}x(t, \lambda(\tau, \xi))|_{t=\tau}$, называется *функцией наклона поля*.

Если существует такая точка (t_*, x_*) , что $x(t_*, \lambda) = x_*$ для всех λ , то семейство $\{x(\cdot, \lambda)\}$ называется *центральной полем экстремалей* (с центром в (t_*, x_*)).

ТЕОРЕМА 29 (Достаточные условия сильного минимума в простейшей задаче). Пусть $\widehat{x}(\cdot)$ — экстремаль в задаче (62), окруженная в области G центральной полем экстремалей с центром в точке $(t_*, \widehat{x}(t_*))$, $t_* < t_0$, такая, что $x(t_j) = x_j$, $j = 0, 1$. Если интегрант L непрерывно дифференцируем на $G \times \mathbb{R}$ и функция $\dot{x} \mapsto L(t, x, \dot{x})$ — выпуклая функция на \mathbb{R} , то $\widehat{x}(\cdot)$ — сильный локальный минимум в задаче (62).

ДОКАЗАТЕЛЬСТВО. Рассмотрим функцию $S: G \rightarrow \mathbb{R}$, определенную равенством

$$S(\tau, \xi) = \int_{t_*}^{\tau} L(t, x(t, \lambda(\tau, \xi)), \dot{x}(t, \lambda(\tau, \xi))) dt.$$

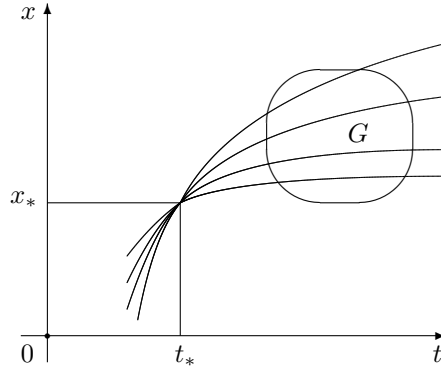


Рис. 2

Она называется *S-функцией центрального поля* $\{x(\cdot, \lambda)\}$. Эта функция непрерывно дифференцируема на G как суперпозиция непрерывно дифференцируемых функций. Подсчитаем ее частные производные. Имеем

$$\begin{aligned} \frac{\partial S(\tau, \xi)}{\partial \tau} &= L(\tau, x(\tau, \lambda(\tau, \xi)), \dot{x}(\tau, \lambda(\tau, \xi))) \\ &+ \int_{t_*}^{\tau} (L_x(t, x(t, \lambda(\tau, \xi)), \dot{x}(t, \lambda(\tau, \xi)))x_\lambda(t, \lambda(\tau, \xi))\lambda_\tau(\tau, \xi) \\ &+ L_{\dot{x}}(t, x(t, \lambda(\tau, \xi)), \dot{x}(t, \lambda(\tau, \xi)))\dot{x}_\lambda(t, \lambda(\tau, \xi))\lambda_\tau(\tau, \xi)) dt. \end{aligned}$$

Интегрируя по частям во втором слагаемом в интеграле и учитывая, что функции $x(\cdot, \lambda)$ — экстремали, получим

$$\begin{aligned} \frac{\partial S(\tau, \xi)}{\partial \tau} &= L(\tau, x(\tau, \lambda(\tau, \xi)), \dot{x}(\tau, \lambda(\tau, \xi))) \\ &+ L_{\dot{x}}(t, x(t, \lambda(\tau, \xi)), \dot{x}(t, \lambda(\tau, \xi)))x_\lambda(t, \lambda(\tau, \xi))\lambda_\tau(\tau, \xi)|_{t_*}^{\tau}. \end{aligned}$$

Дифференцируя тождество $x(\tau, \lambda(\tau, \xi)) = \xi$ по τ , получаем, что $\dot{x}(\tau, \lambda(\tau, \xi)) + x_\lambda(\tau, \lambda(\tau, \xi))\lambda_\tau(\tau, \xi) = 0$. Теперь учитывая, что по определению $\dot{x}(\tau, \lambda(\tau, \xi)) = u(\tau, \xi)$ и что $x_\lambda(t_*, \lambda(\tau, \xi)) = 0$ (так как поле центрально), приходим к равенству

$$(71) \quad \frac{\partial S(\tau, \xi)}{\partial \tau} = L(\tau, \xi, u(\tau, \xi)) - L_{\dot{x}}(\tau, \xi, u(\tau, \xi))u(\tau, \xi).$$

Аналогичные вычисления показывают, что

$$\begin{aligned} \frac{\partial S(\tau, \xi)}{\partial \xi} &= \int_{t_*}^{\tau} (L_x(t, x(t, \lambda(\tau, \xi)), \dot{x}(t, \lambda(\tau, \xi)))x_\lambda(t, \lambda(\tau, \xi))\lambda_\xi(\tau, \xi) \\ &+ L_{\dot{x}}(t, x(t, \lambda(\tau, \xi)), \dot{x}(t, \lambda(\tau, \xi)))\dot{x}_\lambda(t, \lambda(\tau, \xi))\lambda_\xi(\tau, \xi)) dt \\ &= L_{\dot{x}}(t, x(t, \lambda(\tau, \xi)), \dot{x}(t, \lambda(\tau, \xi)))x_\lambda(t, \lambda(\tau, \xi))\lambda_\xi(\tau, \xi)|_{t_*}^{\tau}. \end{aligned}$$

Дифференцируя тождество $x(\tau, \lambda(\tau, \xi)) = \xi$ по ξ , получаем

$$x_\lambda(\tau, \lambda(\tau, \xi))\lambda_\xi(\tau, \xi) = 1$$

и учитывая снова, что $\dot{x}(\tau, \lambda(\tau, \xi)) = u(\tau, \xi)$ и $x_\lambda(t_*, \lambda(\tau, \xi)) = 0$, приходим к соотношению

$$(72) \quad \frac{\partial S(\tau, \xi)}{\partial \xi} = L_{\dot{x}}(\tau, \xi, u(\tau, \xi)).$$

Пусть $x(\cdot) \in PC^1([t_0, t_1])$ — допустимая функция в задаче (62), график которой принадлежит G . Рассматривая функцию S на графике $x(\cdot)$, получаем функцию $t \mapsto S(t, x(t))$ на $[t_0, t_1]$. Ясно, что эта функция принадлежит $PC^1([t_0, t_1])$ и, за исключением конечного числа точек, будем иметь (учитывая выражения для частных производных функции S)

$$(73) \quad S'(t, x(t)) = L(t, x(t), u(t, x(t))) - L_{\dot{x}}(t, x(t), u(t, x(t))) u(t, x(t)) + L_{\dot{x}}(t, x(t), u(t, x(t))) \dot{x}(t).$$

Если $x(\cdot) = \hat{x}(\cdot)$, то $u(t, \hat{x}(t)) = \hat{x}(t)$ и тогда

$$(74) \quad S'(t, \hat{x}(t)) = L(t, \hat{x}(t), \hat{x}(t)).$$

Теперь, используя (74), а затем (73), получим

$$\begin{aligned} J(x(\cdot)) - J(\hat{x}(\cdot)) &= \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt - \int_{t_0}^{t_1} L(t, \hat{x}(t), \hat{x}(t)) dt \\ &= \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt - \int_{t_0}^{t_1} S'(t, \hat{x}(t)) dt = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt \\ &\quad - (S(t_1, x_1) - S(t_0, x_0)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) dt - \int_{t_0}^{t_1} S'(t, x(t)) dt \\ &= \int_{t_0}^{t_1} (L(t, x(t), \dot{x}(t)) - L(t, x(t), u(t, x(t)))) \\ &\quad - L_{\dot{x}}(t, x(t), u(t, x(t))) (\dot{x}(t) - u(t, x(t))) dt \\ &= \int_{t_0}^{t_1} \mathcal{E}(t, x(t), u(t, x(t)), \dot{x}(t)) dt. \end{aligned}$$

Формула

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

называется *основной формулой Вейрштасса*.

Поскольку функция $\dot{x} \mapsto L(t, x, \dot{x})$ выпукла на \mathbb{R} для любых $(t, x) \in G$, то $\mathcal{E}(t, x(t), u(t, x(t)), \dot{x}(t)) \geq 0$ для всех $t \in [t_0, t_1]$ и значит, $J(x(\cdot)) \geq J(\hat{x}(\cdot))$. \square

Замечание. Обозначим

$$\mathcal{H}(\tau, \xi, p) = p \cdot u(\tau, \xi) - L(\tau, \xi, u(\tau, \xi)).$$

Тогда на G функция S удовлетворяет следующему уравнению в частных производных

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

которое называется *уравнением Гамильтона–Якоби*.

Действительно, из равенств (72) и (71) получаем

$$\begin{aligned} \mathcal{H}\left(\tau, \xi, \frac{\partial S(\tau, \xi)}{\partial \xi}\right) &= \frac{\partial S(\tau, \xi)}{\partial \xi} u(\tau, \xi) - L(\tau, \xi, u(\tau, \xi)) \\ &= L_{\dot{x}}(\tau, \xi, u(\tau, \xi)) u(\tau, \xi) - L(\tau, \xi, u(\tau, \xi)) = -\frac{\partial S(\tau, \xi)}{\partial \tau}. \end{aligned}$$

7. Избранные задачи

1. Задача о гармоническом осцилляторе.

Согласно закону Гука, сила притяжения материальной точки пружиной пропорциональна отклонению точки. Второй закон Ньютона утверждает, что движение точки под воздействием силы пружины, описывается дифференциальным уравнением: $m\ddot{x} = F$ и тем самым $m\ddot{x} = -kx$. Соответствующий закон движения (как и вообще все законы классической механики) является следствием *принципа стационарного действия*, согласно которому траектория $x(t)$ материальной точки является стационарной точкой *функционала действия*

$$J(x(\cdot)) = \int \left(\frac{m\dot{x}^2(t)}{2} - \frac{kx^2(t)}{2} \right) dt.$$

Мы рассмотрим следующую задачу

$$J(x(\cdot)) = \int_0^T (\dot{x}^2(t) - x^2(t)) dt \rightarrow \min, \quad x(0) = x(T) = 0,$$

где $T > 0$.

Уравнение Эйлера имеет вид

$$\ddot{x} + x = 0.$$

Его общее решение

$$x(t) = C_1 \cos t + C_2 \sin t.$$

Если $T \neq k\pi$, $k \in \mathbb{N}$, то единственная экстремаль — тождественный нуль. Если $T = k\pi$, то имеется однопараметрическое семейство экстремалей $x(t, C) = C \sin t$.

Пусть $\hat{x}(\cdot)$ — допустимая экстремаль. На этой экстремали выполнено усиленное условие Лежандра, так как $L_{\dot{x}\dot{x}} = 2$ и условие Вейерштрасса, поскольку функция $\dot{x} \mapsto \dot{x}^2 - x^2$ выпукла. Уравнение Якоби имеет тот же вид, что и уравнение Эйлера

$$\ddot{h} + h = 0.$$

Если $T > \pi$, то $\hat{x}(\cdot)$ не может доставлять ни слабый, ни сильный минимум, поскольку на $(0, T)$ точка π является сопряженной к нулю и значит, не выполнено необходимое условие Якоби. Докажем, что в этом случае значение задачи $-\infty$. Положим

$$x_n(t) = n \sin \frac{\pi}{T} t, \quad n \in \mathbb{N}.$$

Имеем

$$\begin{aligned} \int_0^T (\dot{x}_n^2(t) - x_n^2(t)) dt &= n^2 \int_0^T \left(\frac{\pi^2}{T^2} \cos^2 \frac{\pi}{T} t - \sin^2 \frac{\pi}{T} t \right) dt \\ &= -n^2 T + n^2 \left(1 + \frac{\pi^2}{T^2} \right) \int_0^T \cos^2 \frac{\pi}{T} t dt = -n^2 T + n^2 \left(1 + \frac{\pi^2}{T^2} \right) \frac{T}{2} \\ &= n^2 \frac{\pi^2 - T^2}{2T}. \end{aligned}$$

Отсюда видно, что $J(x_n(\cdot)) \rightarrow -\infty$ при $n \rightarrow \infty$.

Если $T < \pi$, то $\hat{x}(\cdot) = 0$ можно окружить полем экстремалей $x(t, \lambda) = \lambda \sin(t + \varepsilon)$, где $0 < \varepsilon < \pi - T$. Действительно, для любых $\tau \in [0, T]$ и любых $\xi \in \mathbb{R}$ при $\lambda = \xi / \sin(\tau + \varepsilon)$ имеем $x(\tau, \lambda) = \lambda \sin(\tau + \varepsilon) = \xi$. Семейство $\{x(\cdot, \lambda)\}$ является центральным полем экстремалей с центром в точке $(-\varepsilon, 0)$. Таким образом, для экстремали $\hat{x}(\cdot) = 0$ выполнены условия теоремы о достаточных условиях сильного минимума и, следовательно, $\hat{x}(\cdot) = 0$ — сильный минимум.

Пусть $T = \pi$. Тогда для любой допустимой функции $x(\cdot) \in C^1([t_0, t_1])$

$$\begin{aligned} \int_0^\pi (\dot{x}(t) - x(t) \operatorname{ctg} t)^2 dt &= \int_0^\pi (\dot{x}^2(t) + x^2(t) \operatorname{ctg}^2 t dt - \int_0^\pi \operatorname{ctg} t dx^2(t) \\ &= \int_0^\pi (\dot{x}^2(t) + x^2(t) \operatorname{ctg}^2 t) dt - x^2(t) \operatorname{ctg} t \Big|_0^\pi - \int_0^\pi \frac{x^2(t)}{\sin^2 t} dt \\ &= \int_0^\pi (\dot{x}^2(t) - x^2(t)) dt. \end{aligned}$$

Отсюда вытекает, что любой допустимой функции $x(\cdot) \in C^1([t_0, t_1])$ выполняется неравенство $J(x(\cdot)) \geq 0$. Тем самым функция $\hat{x}(\cdot) = 0$ является глобальным минимумом.

Ответ. При $0 < T \leq \pi$ решением задачи о гармоническом осцилляторе является тождественный нуль. При $T > \pi$ значение задачи равно $-\infty$.

2. Аэродинамическая задача Ньютона

В “Началах натуральной философии” (1687) Ньютон описал некую среду (типа разреженного воздуха), поставил и решил вопрос о том, когда “тело, образующееся при вращении кривой вокруг оси при движении в упомянутой среде [...] будет испытывать меньшее сопротивление, нежели всякое иное тело вращения при той же высоте и наибольшей ширине”.

Формализованная постановка этой задачи, как задачи оптимального управления, имеет вид

$$\int_0^{t_1} \frac{t dt}{1 + u^2(t)} \rightarrow \min, \quad \dot{x} = u, \quad x(0) = 0, \quad x(t_1) = x_1, \quad u \geq 0,$$

где $x_1 > 0$.

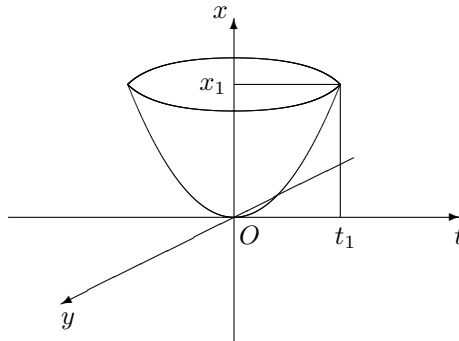


Рис. 3

Функция Лагранжа этой задачи такова

$$\mathcal{L}(x(\cdot), u(\cdot), \bar{\lambda}) = \int_0^{t_1} L(t, x(t), \dot{x}(t), u(t), p(t), \lambda_0) dt,$$

где $L(t, x, \dot{x}, u, p, \lambda_0) = \lambda_0 t / (1 + u^2) + p(\dot{x} - u)$ и $\bar{\lambda} = (\lambda_0, p(\cdot))$.

Согласно принципу максимума должно выполняться уравнение Эйлера по x :

$$(75) \quad -\frac{d}{dt} \hat{L}_{\dot{x}}(t) + \hat{L}_x(t) = 0 \Leftrightarrow -\dot{p} = 0 \Leftrightarrow p = \text{const}$$

и условие минимума по u :

$$(76) \quad \min_{u \geq 0} \left(\frac{\lambda_0 t}{1 + u^2} - pu \right) = \frac{\lambda_0 t}{1 + \hat{u}^2(t)} - p \hat{u}(t).$$

Проверим, что $\lambda_0 \neq 0$. Если $\lambda_0 = 0$, то необходимо $p \neq 0$, так как множители Лагранжа не могут быть все нулевыми. Но, если $p \neq 0$, то из (76) следует, что $\hat{u}(t) \equiv 0$ и тогда

$$\hat{x}(t_1) = \int_0^{t_1} \hat{u}(t) dt = 0$$

в противоречие с условием $x(t_1) = x_1 > 0$. Итак, $\lambda_0 \neq 0$ и мы можем считать, что $\lambda_0 = 1$.

Отметим еще, что $p < 0$. Действительно, если $p \geq 0$, то при любом $t \geq 0$ функция $u \rightarrow f(t, u) = t/(1+u^2) - pu$ монотонно убывает и поэтому соотношение (76) не может выполняться.

Определим, при каких t минимум функции $f(t, \cdot)$ достигается в точке $u = 0$. Это означает, что $f(t, 0) \leq f(t, u)$ при всех $u \geq 0$. Рассмотрим соответствующее неравенство

$$t \leq \frac{t}{1+u^2} - pu.$$

Это неравенство можно записать в виде

$$\frac{tu}{1+u^2} \leq -p.$$

Отсюда видно, что оно выполнено для $t \geq 0$ и таких, что

$$\frac{t}{-p} \leq \min_{u>0} \frac{1+u^2}{u} = 2.$$

Положим $\xi = -2p$. Тогда при $0 \leq t < \xi$ минимум функции $f(t, \cdot)$ достигается в единственной точке $u = 0$. Значит, $\hat{u}(t) = 0$, при $0 \leq t < \xi$. Следовательно, $\hat{x}(t) = 0$ при $0 \leq t < \xi$.

Пусть $t > \xi$. Тогда при $u \in [0, 1)$

$$\begin{aligned} f(t, u) - f(t, 1) &= \frac{t}{1+u^2} - pu - \frac{t}{2} + p = t \left(\frac{1}{1+u^2} - \frac{1}{2} \right) - p(u-1) \\ &> -2p \left(\frac{1}{1+u^2} - \frac{1}{2} \right) - p(u-1) = \frac{-2p}{1+u^2} - pu + 2p \\ &= -p \left(\frac{2}{1+u^2} - 2 + u \right) = -pu \frac{(1-u)^2}{1+u^2} \geq 0. \end{aligned}$$

Тем самым при $t > \xi$ минимум $f(t, \cdot)$ достигается в некоторой точке $u \geq 1$. В точке $t = \xi$ функция $\hat{u}(\cdot)$ имеет разрыв, а функция $\hat{x}(\cdot)$ — излом. Оптимальное управление $\hat{u}(\cdot)$ можно найти из условия $f_u(t, u) = 0$, из которого следует, что

$$t = -\frac{p(1+u^2)^2}{2u} = -\frac{p}{2} \left(\frac{1}{u} + 2u + u^3 \right).$$

Из того, что при $u \geq 1$

$$\frac{dt}{du} = -\frac{p}{2} \left(-\frac{1}{u^2} + 2 + 3u^2 \right) > 0.$$

вытекает, что каждому $t > \xi$ соответствует ровно одно $u \geq 1$, удовлетворяющее уравнению $f_u(t, u) = 0$. При этом u строго возрастает с ростом t .

В силу того, что $dx/dt = u$, имеем

$$\frac{dx}{du} = \frac{dx}{dt} \frac{dt}{du} = u \frac{dt}{du} = -\frac{p}{2} \left(-\frac{1}{u} + 2u + 3u^3 \right).$$

Отсюда

$$\hat{x} = -\frac{p}{2} \left(\ln \frac{1}{u} + u^2 + \frac{3}{4}u^4 \right) + c.$$

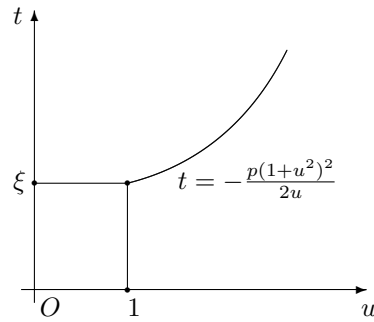


Рис. 4

Тем самым получаем параметрическое представление при $u \geq 1$ (т. е. при $t \geq \xi$)

$$\hat{x} = -\frac{p}{2} \left(\ln \frac{1}{u} + u^2 + \frac{3}{4}u^4 \right) + c,$$

$$t = -\frac{p}{2} \left(\frac{1}{u} + 2u + u^3 \right).$$

В силу непрерывности $\hat{x}(\cdot)$ и того, что $\hat{x}(t) = 0$ при $0 \leq t < \xi$, вытекает, что $\hat{x}(\xi) = 0$. Следовательно, $c = 7p/8$. Итак, если $0 \leq t \leq -2p$, то $\hat{x}(t) = 0$, а при $t \geq -2p$ кривая $\hat{x}(\cdot)$ задается параметрически

$$\hat{x}(t, p) = -\frac{p}{2} \left(\ln \frac{1}{u} + u^2 + \frac{3}{4}u^4 \right) + \frac{7}{8}p,$$

$$t = -\frac{p}{2} \left(\frac{1}{u} + 2u + u^3 \right),$$

где $p < 0$ и определяется из условия $\hat{x}(t_1, p) = x_1$. Эту кривую называют *кривой Ньютона*.

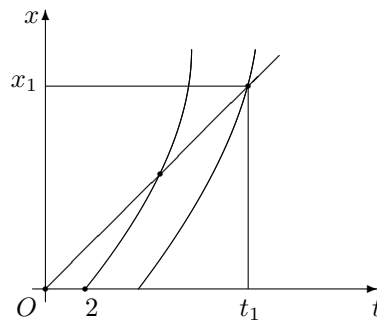


Рис. 5

Покажем, что всякая прямая $y = kt$ пересекает кривую Ньютона. Действительно, функция $\hat{x}(\cdot, p)$ непрерывна и выпукла, так как

$$\frac{d^2x}{dt^2} = \frac{du}{dt} = \frac{1}{dt/du} = \frac{1}{-\frac{p}{2} \left(-\frac{1}{u^2} + 2 + 3u^2 \right)} > 0.$$

Кроме того,

$$\lim_{t \rightarrow +\infty} \frac{x}{t} = \lim_{t \rightarrow +\infty} \frac{dx}{dt} = \lim_{t \rightarrow +\infty} u = +\infty.$$

Из параметрического представления кривой Ньютона видно, что кривая $\hat{x}(cd, p)$ получается из кривой $\hat{x}(cd, -1)$ гомотетией с центром O и коэффициентом $|p|$. Поэтому, чтобы провести кривую Ньютона через заданную точку (t_1, x_1) нужно найти точку пересечения прямой $x = x_1 t / t_1$ с кривой $\hat{x}(cd, -1)$, а затем сделать гомотетию кривой $\hat{x}(cd, -1)$, при которой точка пересечения переходит в точку (t_1, x_1) .

Убедимся, что кривая Ньютона доставляет абсолютный минимум в задаче. Пусть $x(\cdot)$ — допустимая функция в исходной задаче, т. е. она кусочно непрерывно дифференцируема, $x(0) = 0$ и $x(t_1) = x_1$. Тогда в силу (76) ($\lambda_0 = 1$)

$$\frac{t}{1 + \dot{x}^2(t)} - p\dot{x}(t) \geq \frac{t}{1 + \hat{u}^2(t)} - p\hat{u}(t).$$

Интегрируя это неравенство с учетом того, что

$$\int_0^{t_1} \dot{x}(t) dt = \int_0^{t_1} \hat{x}(t) dt = x_1,$$

получаем требуемое

$$\int_0^{t_1} \frac{t dt}{1 + \dot{x}^2(t)} \geq \int_0^{t_1} \frac{t dt}{1 + \hat{x}^2(t)}.$$

В результате поверхность, которая решает поставленную задачу, изображена на рис. 6.

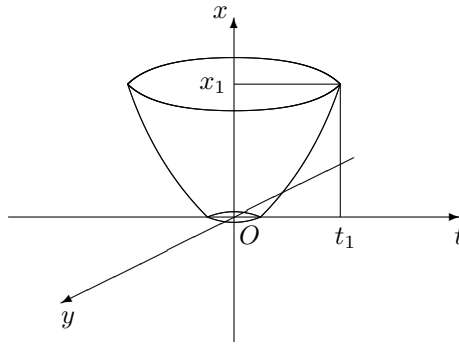


Рис. 6

Литература

- [1] Алексеев В. М., Тихомиров В. М., Фомин С. В. Оптимальное управление. М.: Наука, 1979.
- [2] Ильин В. А., Позняк Э. Г. Основы математического анализа: В 2-х ч. Часть I. М: Физматлит, 2005.
- [3] Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. М.: ФИЗМАТЛИТ, 2004.