Вариационное исчисление и оптимальное управление (мехмат МГУ, лекции 2020 г.)

Е. Р. Аваков, Г. Г. Магарил-Ильяев

Содержание

Введение

Какие причины побуждают решать задачи на максимум и минимум?

Интерес к экстремальным задачам, т. е. задачам на максимум и минимум 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[1] 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".

Важная причина, побуждающая интерес к исследованию экстремальных задач, связана с тем, что многие законы природы, как оказалось, подчинены *экстремальным принципам*, т. е. многие природные процессы (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".

Первый экстремальный принцип в естествознании выдвинул Пьер Ферма (1662). Согласно этому принципу *свет "избирает" такую траекторию, вдоль которой время, затрачиваемое им на преодоление пути от одной точки до другой, минимально.* Законы преломления света, установленные экспериментально, находятся, исходя из этого принципа, как решения экстремальных задач. Отправляясь от принципа Ферма, И. Бернулли в 1696 году решил так называемую *задачу о брахистохроне* — о кривой наибыстрейшего ската, т. е. задачу 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 такой задачи возможно была навеяна более ранними размышлениями Галилея на эту тему). Историю развития вариационного исчисления принято отсчитывать именно с этого времени — года решения задачи о брахистохроне.

Нельзя не назвать также и прагматические причины, по которым приходится решать экстремальные задачи. Людям свойственно наилучшим образом использовать ресурсы, находящиеся в их распоряжении, и потому экстремальные задач естественно возникают при управлении различными процессами, в экономике, инженерии.

Первой задачей такого рода была *аэродинамическая задача Ньютона* о форме тела вращения, испытывающем наименьшее сопротивление в некоей "редкой среде", поставленная и решенная Нью-

[1] Ответ в ней приводил в своих сочинениях еще Аристотель (IV в. до н. э.)

тоном в его "Математических началах натуральной философии" (1687 г.). К числу экстремальных задач, возникающих в экономике и теории управления, относятся, например, *транспортная задача* (сформулированная в сороковые годы прошлого века) о наилучшем способе транспортировки продуктов со складов в магазины и *простейшая задача о быстродействии*, формализующая, в частности, задачу о наименьшем времени движения лифта в угольной шахте. Эта задача, также сформулированная в сороковые годы, послужила стимулом для создания *теории оптимального управления*. Считалось, что эта теория родилась в пятидесятые годы прошлого века, но когда она, в основном, сложилась, выяснилось, что первой задачей оптимального управления была именно аэродинамическая задача Ньютона.

Экстремальные задачи и их формализация

Мы видим, что зачастую экстремальные задачи ставятся 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*. Точно поставленная экстремальная задача включает в себя функционал f (со значениями, вообще говоря, в расширенной прямой $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$) вместе со своей областью определения X и множество ограничений $C \subset X$. Формализованная экстремальная задача записывается так

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

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 (P) или ее *решениями*. Если нас интересуют и точки минимума и точки максимума, то вместо $\min(\max)$ пишем extr и говорим о задаче на экстремум функционала f.

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

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

П. Ферма в 1638 году иллюстрировал свой метод нахождения экстремумов на решении следующей задачи: *найти прямоугольный треугольник наибольшей площади при данной сумме длин его катетов*. Если a — сумма длин катетов, а x — длина одного из них, то формализованная постановка выглядит так $x(a-x)/2 \to \max,\ 0 \le x \le a$.

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

Одна из основных целей курса — изложить единый взгляд на необходимые условия экстремума первого порядка для различных классов экстремальных задач (нелинейного и выпуклого программирования, вариационного исчисления и оптимального управления). С этой целью мы рассматриваем экстремальную задачу общего вида на выпуклом множестве с ограничениями типа равенств и выводим для нее необходимые условия минимума. Оказывается, что необходимые условия экстремума во всех перечисленных классах задач являются непосредственными следствиями необходимых условий, полученных в общей задаче.

Глава 1. Аппарат теории экстремальных задач

В этой главе собраны те факты, на которые опираются доказательства утверждений, представленные в этих лекциях.

§1.1. Элементы функционального анализа

1.1.1. Нормированные пространства и линейные операторы

Векторное пространство X называется (вещественным) *нормированным пространством*, если каждому элементу $x \in X$ сопоставлено вещественное число $\|x\|_X$, называемое нормой x, удовлетворяющее условиям: (a) $\|x\|_X \geq 0$ для любого $x \in X$ и $\|x\|_X = 0$ тогда и только тогда, когда $x = 0$; (b) $\|\alpha x\|_X = |\alpha|\|x\|_X$ для любых $\alpha \in \mathbb{R}$ и $x \in X$; (c) $\|x + y\|_X \leq \|x\|_X + \|y\|_X$ для любых $x, y \in X$ (неравенство треугольника).

Если $x, y \in X$, то величина $\rho(x, y) = \|x - y\|_X$ называется *расстоянием между элементами x и y*.

Расстоянием от точки $\widehat{x} \in X$ до множества $A \subset X$ называется величина $\operatorname{dist}(\widehat{x}, A) = \inf_{x \in A} \|\widehat{x} - x\|_X$.

Пусть $\widehat{x} \in X$ и $r > 0$. Множества $U_X(\widehat{x}, r) = \{\, x \in: \|x - \widehat{x}\|_X < r \,\}$ и $B_X(\widehat{x}, r) = \{\, x \in X : \|x - \widehat{x}\|_X \leq r \,\}$ называется соответственно *открытым* и *замкнутым* шарами в X с центром в $\widehat{x}$ радиуса r.

Пусть $A \subset X$. Точка $x \in A$ называется *внутренней* точкой A, если она входит в A с некоторым шаром с центром в x. Множество внутренних точек множества A называется *внутренностью* A и обозначается $\operatorname{int} A$.

Множество A называется *открытым*, если $\operatorname{int} A = A$ (проверьте, что открытый шар — открытое множество).

Любое открытое множество, содержащее данную точку, называется *окрестностью* этой точки.

Точка $x \in X$ называется *граничной* точкой множества A, если в любой ее окрестности содержатся точки из A и из $X \backslash A$. Множество всех граничных точек множества A обозначаем ∂A (проверьте, что если $x \in A$, то либо $x \in \operatorname{int} A$, либо $x \in \partial A$).

Множество $B \subset X$ называется *замкнутым*, если его дополнение открыто (проверьте, что замкнутый шар — замкнутое множество).

Пусть $A \subset X$. Точка $x \in X$ называется *предельной точкой* множества A, если каждая ее окрестность имеет непустое пересечение с A. Совокупность всех предельных точек множества A называется *замыканием* A и обозначается $\mathrm{cl}\,\mathrm{A}$ (проверьте, что $\mathrm{cl}\,\mathrm{A}$ — замкнутое множество и что множество A замкнуто тогда и только тогда, когда оно совпадает со своим замыканием).

Множество $A \subset X$ называется *компактным* (или *компактом*), если из любого его покрытия открытыми множествами можно выделить конечное подпокрытие.

Нормированное пространство X называется *банаховым* пространством, если каждая фундаментальная последовательность его элементов сходится.

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

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

Пусть $\Lambda \in \mathcal{L}(X,Y)$. Множества $\operatorname{Im}\Lambda = \{\, y \in Y \mid \exists\, x \in X : \Lambda x = y \,\}$ и $\operatorname{Ker}\Lambda = \{x \in X : \Lambda x = 0\,\}$ называются соответственно *образом* и *ядром* оператора Λ.

Если $\operatorname{Im}\Lambda = Y$, то говорят, что оператор Λ *сюръективен* (или Λ — *эпиморфизм* или Λ — *отображение на*).

Если $\operatorname{Ker}\Lambda = \{0\}$, то говорят, что оператор Λ *инъективен* (или Λ — *мономорфизм* или Λ — *взаимно однозначное отображение*).

Если оператор $\Lambda \in \mathcal{L}(X,Y)$ и сюръективен и инъективен, то говорят, что он *биективен.*

Пусть X и Y — нормированные пространства. Говорят, что пространства *изометрически изоморфны*, если существует такое линейное сюръективное отображение $A\colon X \to Y$, что $\|Ax\|_Y = \|x\|_X$. Отсюда следует, что отображение A непрерывно и взаимно однозначно. Изометрически изоморфные пространства, как нормированные пространства, не различимы.

Множество всех линейных непрерывных функционалов на нормированном пространстве X называется *сопряженным пространством* к X и обозначается X^*. Значение линейного функционала $x^* \in X^*$ на элементе $x \in X$ обозначаем так: $\langle x^*, x\rangle$.

Пространство X^* является банаховым пространством с нормой

$$\|x^*\|_{X^*} = \sup_{\|x\|_X \le 1} \langle x^*, x\rangle.$$

Пусть X^* и Y^* — сопряженные пространства соответственно к X и Y и $\Lambda \in \mathcal{L}(X,Y)$. Отображение $\Lambda^*\colon Y^* \to X^*$, определенное по правилу $\langle \Lambda^* y^*, x\rangle = \langle y^*, \Lambda x\rangle$ (т. е. функционалу y^* сопоставляется функционал $x \mapsto \langle y^*, \Lambda x\rangle$), называется *сопряженным оператором к* Λ. При этом, $\Lambda^* \in \mathcal{L}(Y^*, X^*)$.

Заметим, что если оператор $\Lambda \in \mathcal{L}(X,Y)$ сюръективен, то сопряженный оператор инъективен. Действительно, пусть $y^* \in \operatorname{Ker}\Lambda^*$. Для любого $y \in Y$ существует $x \in X$ такое, что $y = \Lambda x$ и мы имеем $\langle y^*, y\rangle = \langle y^*, \Lambda x\rangle = \langle \Lambda^* y^*, x\rangle = 0$, т. е. $y^* = 0$ и значит, $\operatorname{Ker}\Lambda^* = \{0\}$.

Произведение $X \times Y$ нормированных пространств X и Y (которое, как векторное пространство, есть совокупность пар (x,y), где $x \in X$ и $y \in Y$, с операциями покоординатного сложения и умножения на числа) есть нормированное пространство с нормой $\|(x,y)\|_{X\times Y} = \|x\|_X + \|y\|_Y$. Легко проверить, что если X и Y — банаховы пространства, то $X \times Y$ — банахово пространство.

В пространстве $X \times Y$ иногда рассматривают другие, но эквивалентные[1] введенной нормы, например, $\sqrt{\|x\|_X^2 + \|y\|_Y^2}$ или $\max(\|x\|_X, \|y\|_Y)$ (проверьте эквивалентность).

---

[1]Нормы $\|\cdot\|_1$ и $\|\cdot\|_2$ в нормированном пространстве Z эквивалентны, если существуют такие константы $c_i > 0$, $i = 1,2$, что $c_1\|z\|_1 \le \|z\|_2 \le c_2\|z\|_1$ для всех $z \in Z$.

Лемма 1 (о сопряженном к произведению пространств)**.** *Пусть X и Y — нормированные пространства. Тогда для любого $z^* \in (X \times Y)^*$ найдутся элементы $x^* \in X^*$ и $y^* \in Y^*$ такие, что для всех $(x, y) \in X \times Y$ справедлива формула*

$$\langle z^*, (x, y)\rangle = \langle x^*, x\rangle + \langle y^*, y\rangle.$$

Доказательство. Определим функционалы $x^* \in X^*$ и $y^* \in Y^*$ соответственно по формулам $\langle x^*, x\rangle = \langle z^*, (x, 0)\rangle$ для всех $x \in X$ и $\langle y^*, y\rangle = \langle z^*, (0, y)\rangle$ для всех $y \in Y$. Тогда в силу линейности z^*

$$\langle z^*, (x, y)\rangle = \langle z^*, (x, 0)\rangle + \langle z^*, (0, y)\rangle = \langle x^*, x\rangle + \langle y^*, y\rangle.$$

□

1.1.2. Примеры нормированных пространств

Приведем здесь примеры банаховых пространств, с которыми, в основном, будем иметь дело в дальнейшем.

1. Пространство $\mathbb{R}^n$. Это совокупность всех упорядоченных наборов $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ из n действительных чисел (если $n = 1$, то это просто множество действительных чисел, и мы пишем $\mathbb{R}$ вместо $\mathbb{R}^1$), называемых *векторами* или *вектор-столбцами*, а числа x_i, $1 \le i \le n$ — *координатами вектора* x. Ради экономии места, элементы $\mathbb{R}^n$ иногда будем записывать так $x = (x_1, \ldots, x_n)^T$, где ^{T} обозначает *транспонирование.* В $\mathbb{R}^n$ естественным образом вводится операция (покоординатного) сложения векторов и операция (покоординатного) умножения вектора на число, превращающие это множество в вещественное векторное пространство.

Пусть $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$. Величина $|x| = \sqrt{x_1^2 + \ldots + x_n^2}$ называется *длиной* или *модулем* или *евклидовой нормой* вектора x. Элементарная проверка показывает, что это, действительно, норма в $\mathbb{R}^n$. Из полноты множества действительных чисел следует, что $\mathbb{R}^n$ — банахово пространство.

В пространстве $\mathbb{R}^n$ равносильное определение компактности следующее: множество $A \subset \mathbb{R}^n$ *компактно*, если оно ограничено (т. е. содержится в некотором шаре) и замкнуто (докажите это утверждение).

Пусть $a = (a_1, \ldots, a_n)$ — вектор-строка из n действительных чисел. Отображение $x \mapsto \langle a, x\rangle = \sum_{i=1}^n a_i x_i$, где $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, есть, очевидно, линейный функционал на $\mathbb{R}^n$. Легко видеть, что и любой линейный функционал l на $\mathbb{R}^n$ задается подобным образом с $a = (l(e_1), \ldots, l(e_n))$, где $e_1 = (1, 0, \ldots, 0)^T, \ldots, e_n = (0, \ldots, 0, 1)^T$ — *стандартный базис* в $\mathbb{R}^n$. Таким образом, пространство $(\mathbb{R}^n)^*$ можно отождествить с множеством наборов из n действительных чисел, но расположенных в строку (с аналогичными операциями сложения и умножения на числа и с той же евклидовой нормой). Из неравенства Коши–Буняковского следует, что элементы $(\mathbb{R}^n)^*$ — непрерывные функционалы.

Пусть $\Lambda \colon \mathbb{R}^n \to \mathbb{R}^m$ — линейный оператор. Мы будем его отождествлять с его матрицей (размера $m \times n$) в стандартных базисах $\mathbb{R}^n$ и $\mathbb{R}^m$ соответственно и обозначать той же буквой. В этом случае Λx — произведение матрицы Λ на вектор $x = (x_1, \ldots, x_n)^T$. Проверьте, что линейные операторы непрерывны.

2. Пространство $C([t_0,t_1],\mathbb{R}^n)$. Это совокупность непрерывных вектор-функций на $[t_0,t_1]$ со значениями в $\mathbb{R}^n$ с обычными операциями сложения и умножения на числа и нормой $\|x(\cdot)\|_{C([t_0,t_1],\mathbb{R}^n)} = \max_{t\in[t_0,t_1]}|x(t)|$ (если $n=1$, то вместо $C([t_0,t_1],\mathbb{R})$ пишем $C([t_0,t_1])$). Несложная проверка показывает, что $C([t_0,t_1],\mathbb{R}^n)$ — банахово пространство.

3. Пространство $C^1([t_0,t_1],\mathbb{R}^n)$. Это совокупность всех непрерывно дифференцируемых вектор-функций на отрезке $[t_0,t_1]$ с обычными операциями сложения и умножения на числа и нормой $\|x(\cdot)\|_{C^1([t_0,t_1],\mathbb{R}^n)} = \max(\|x(\cdot)\|_{C([t_0,t_1],\mathbb{R}^n)}, \|\dot{x}(\cdot)\|_{C([t_0,t_1],\mathbb{R}^n)})$ (если $n=1$, то вместо $C^1([t_0,t_1],\mathbb{R})$ пишем $C^1([t_0,t_1])$). Снова, простая проверка показывает, что $C^1([t_0,t_1],\mathbb{R}^n)$ — банахово пространство.

1.1.3. Теорема отделимости и критерий граничной точки выпуклого множества

Пусть A и B — непустые подмножества $\mathbb{R}^n$. Говорят, что ненулевой функционал $x^*\in(\mathbb{R}^n)^*$ *отделяет множества* A *и* B, если

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

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

Пусть число $\gamma\in\mathbb{R}$ таково, что $\sup_{x\in A}\langle x^*,x\rangle \le \gamma \le \inf_{x\in B}\langle x^*,x\rangle$. Тогда, геометрически, отделимость множеств A и B означает, что они расположены по разные стороны от гиперплоскости $\{\, x\in\mathbb{R}^n : \langle x^*,x\rangle = \gamma \,\}$.

Пусть X — линейное пространство. Непустое множество $A\subset X$ называется *выпуклым*, если с любыми двумя своими точками оно содержит и отрезок $[x,y] = \{\, z\in\mathbb{R}^n : z=(1-\alpha)x+\alpha y,\ 0\le\alpha\le 1 \,\}$, соединяющий эти точки.

Пустое множество выпукло по определению.

Пусть A — непустое подмножество $\mathbb{R}^n$. Наименьшее выпуклое множество, содержащее A, называется *выпуклой оболочкой множества* A и обозначается $\operatorname{co}\mathrm{A}$. Проверьте, что $\operatorname{co}\mathrm{A}$ состоит из выпуклых комбинаций элементов из A, т. е. векторов вида $\alpha_1x_1+\ldots+\alpha_mx_m$, где $x_i\in A$, $\alpha_i\ge 0$, $1\le i\le m$, и $\sum_{i=1}^m\alpha_i=1$.

Теорема 1 (Теорема отделимости)**.** *Пусть A и B — непустые выпуклые подмножества $\mathbb{R}^n$ и $A\cap B=\emptyset$. Тогда множества A и B отделимы.*

Доказательство этой теоремы можно найти в [**?**].

Предложение 1. *Пусть A — непустое выпуклое подмножество $\mathbb{R}^n$.*

(*a*) *Если $\operatorname{int}A=\emptyset$, то A содержится в некоторой гиперплоскости.*

(*b*) *Если $x\in\operatorname{int}A$ и $y\in A$, то $(1-\alpha)x+\alpha y\in\operatorname{int}A,\ \ \forall\,\alpha\in(0,1)$.*

Доказательство. (*a*) Докажем эквивалентное утверждение: если A не содержится ни в какой гиперплоскости, то $\operatorname{int}A\ne\emptyset$. Без ограничения общности считаем, что $0\in A$. Покажем сначала, что если A не содержится ни в какой собственном подпространстве (а тем самым ни в какой гиперплоскости, проходящей через ноль), то A содержит n линейно независимых векторов. Действительно, пусть максимальное число линейно независимых векторов из A равно $k < n$. Тогда, если $x_1, \dots, x_k$ — такие векторы и L — подпространство, натянутое на них, то по предположению A не принадлежит L. Следовательно, существует вектор $x_{k+1} \in A$ такой, что $x_{k+1} \notin L$ и значит, векторы $x_1, \dots, x_k, x_{k+1}$ линейно независимы, что противоречит максимальности k.

Пусть теперь $T\colon \mathbb{R}^n \to \mathbb{R}^n$ — линейный оператор, переводящий линейно независимые векторы $x_1, \dots, x_n$ из A в векторы стандартного базиса $e_1, \dots, e_n$. Если обозначить $S = \operatorname{co}\{0, x_1, \dots, x_n\}$ и $S_0 = \operatorname{co}\{0, e_1, \dots, e_n\}$, то, как нетрудно проверить, $T(S) = S_0$. Множество $\{\alpha = (\alpha_1, \dots, \alpha_n)^T \in \mathbb{R}^n : \alpha_i > 0,\ i = 1, \dots, n,\ \sum_{i=1}^n \alpha_i < 1\}$, очевидно, открыто и содержится в S_0. Следовательно, прообраз этого множества при непрерывном отображении T открыт и содержится в S. Поскольку A — выпуклое множество и элементы $0, x_1, \dots, x_n$ содержатся в A, то $S \subset A$ по определению выпуклой оболочки и поэтому $\operatorname{int} \mathrm{A} \neq \emptyset$.

(b) Пусть $z = (1-\alpha)x + \alpha y$ для некоторого $0 < \alpha < 1$. Так как $x \in \operatorname{int} \mathrm{A}$, то существует такое $\delta > 0$, что $U_{\mathbb{R}^n}(x, \delta) \subset A$. Множество $(1-\alpha)U_{\mathbb{R}^n}(x,\delta) + \alpha y$ открыто (почему?), содержит z и принадлежит A, т. е. $z \in \operatorname{int} \mathrm{A}$. □

Пусть A — непустое подмножество $\mathbb{R}^n$ и $\widehat{x} \in A$. Множество

$$N_A(\widehat{x}) = \{\, x^* \in (\mathbb{R}^n)^* : \langle x^*, x - \widehat{x} \rangle \le 0,\ \forall\, x \in A \,\}$$

называется *нормальным конусом* ко множеству A в точке $\widehat{x}$.

Предложение 2 (Критерий граничной точки)**.** *Пусть K — выпуклое подмножество $\mathbb{R}^n$ и $\widehat{x} \in K$. Тогда $\widehat{x} \in \partial K$ в том и только в том случае, если $N_K(\widehat{x}) \neq \{0\}$.*

Доказательство. Необходимость. Пусть $\widehat{x} \in \partial K$ и $\operatorname{int} K \neq \emptyset$. Тогда $\widehat{x} \notin \operatorname{int} K$, и так как $\operatorname{int} K$ — выпуклое множество (что сразу следует из утверждения (b) предложения 1), то по теореме отделимости найдется ненулевой функционал $x^* \in X^*$ такой, что $\langle x^*, x \rangle \le \langle x^*, \widehat{x} \rangle$ для всех $x \in \operatorname{int} K$. Отсюда, снова в силу утверждения (b) и непрерывности x^*, следует, что $\langle x^*, x - \widehat{x} \rangle \le 0$ для всех $x \in K$, т. е. $x^* \in N_K(\widehat{x})$.

Если $\operatorname{int} K = \emptyset$, то согласно утверждению (a) предложения 1 множество K содержится в некоторой гиперплоскости $\{x \in \mathbb{R}^n : \langle x^*, x \rangle = \gamma\}$. Ясно, что $\gamma = \langle x^*, \widehat{x} \rangle$, и мы получаем, что $\langle x^*, x - \widehat{x} \rangle = 0$ для любого $x \in K$ и значит, $x^* \in N_K(\widehat{x})$.

Достаточность. Пусть $x^* \in N_K(\widehat{x})$ и $x^* \neq 0$. Существует такой вектор $x \in \mathbb{R}^n$, что $\langle x^*, x \rangle > 0$ и значит, $\langle x^*, \alpha x \rangle > 0$ для любого $\alpha > 0$. Следовательно, $\widehat{x} + \alpha x \notin K$ для любого $\alpha > 0$ и тем самым $\widehat{x} \in \partial K$. □

§1.2. Дифференциальное исчисление в нормированных пространствах

1.2.1. Основные понятия и теоремы

Пусть X, Y — нормированные пространства, U — открытое подмножество X и задано отображение $F\colon U \to Y$.

Определение 1. *Отображение F называется дифференцируемым в точке $\widehat{x} \in U$, если найдется такой оператор $\Lambda \in \mathcal{L}(X, Y)$, что для всех $h \in X$, для которых*

$x + h \in U$ справедливо представление

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

где $r(h) = o(\|h\|_X)$ ($\|r(h)\|_Y / \|h\|_X \to 0$ при $h \to 0$). Оператор Λ называется производной отображения F в точке $\widehat{x}$ и обозначается $F'(\widehat{x})$.

Легко проверить, что производная определена однозначно.

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

Определение 2. *Отображение F называется строго дифференцируемым в точке $\widehat{x} \in U$, если найдется такой оператор $\Lambda \in \mathcal{L}(X, Y)$, что для любого $\varepsilon > 0$ существует $\delta = \delta(\varepsilon) > 0$, обладающее тем свойством, что если $\|x_i - \widehat{x}\|_X < \delta$, $i = 1, 2$, то справедливо неравенство*

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

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

Обратное не верно. Действительно, из определения строгой дифференцируемости в $\widehat{x}$ легко следует, что отображение непрерывно в некоторой окрестности $\widehat{x}$. Рассмотрим функцию $f \colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^2 D(x)$, где $D(\cdot)$ — функция Дирихле (равная единице, если x рационально и нулю, если иррационально). Легко проверить, что f дифференцируема в нуле и, очевидно, разрывна во всех остальных точках и значит, не может быть строго дифференцируемой в нуле.

Доказательство этой теоремы см. в [2], п. 2.2.2.

Покажем, что если отображение непрерывно дифференцируемо в $\widehat{x}$, то оно строго дифференцируемо в этой точке. Сначала сформулируем следующую теорему.

Теорема 1 (о среднем)**.** *Пусть отображение F дифференцируемо в каждой точке отрезка $[x_1, x_2] \subset U$. Тогда*

$$\|F(x_1) - F(x_2)\|_Y \le \sup_{x \in [x_1, x_2]} \|F'(x)\| \|x_1 - x_2\|_X.$$

Доказательство см. в [2], п. 2.2.3.

Следствие 1. *Пусть отображение F непрерывно дифференцируемо в точке $\widehat{x} \in U$. Тогда F строго дифференцируемо в $\widehat{x}$.*

Доказательство. Пусть $\varepsilon > 0$ и $\delta = \delta(\varepsilon) > 0$ такое, что $\|F'(x) - F'(\widehat{x})\| < \varepsilon$ для $x \in U_X(\widehat{x}, \delta)$. Если $x_i \in U_X(\widehat{x}, \delta)$, $i = 1, 2$, то легко видеть, что $[x_1, x_2] \subset U_X(\widehat{x}, \delta)$ и тогда применяя теорему о среднем к отображению $x \mapsto F(x) - F'(\widehat{x})x$, получаем, что

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

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

Теорема 2 (о суперпозиции дифференцируемых отображений)**.** *Пусть X, Y и Z — нормированные пространства, U — окрестность точки $\widehat{x} \in X$, $F\colon U \to Y$, V — окрестность точки $F(\widehat{x})$, $G\colon V \to Z$ и $H = G \circ F\colon U \to Z$. Тогда, если F дифференцируемо (строго дифференцируемо, непрерывно диффеенцируемо) в точке $\widehat{x}$, G дифференцируемо (строго дифференцируемо, непрерывно дифференцируемо) в точке $F(\widehat{x})$, то отображение H дифференцируемо (строго дифференцируемо, непрерывно дифференцируемо) в точке $\widehat{x}$ и $H'(\widehat{x}) = G'(F(\widehat{x})) \circ F'(\widehat{x})$.*

Пусть X, Y_1 и Y_2 — нормированные пространства, U — окрестность точки $\widehat{x}$ и $F_i\colon U \to Y_i$, $i = 1, 2$. Рассмотрим отображение $F = (F_1, F_2)\colon U \to Y_1 \times Y_2$, действующее по правилу: $F(x) = (F_1(x), F_2(x))$ для всех $x \in U$.

Предложение 1. *Отображение F дифференцируемо (строго дифференцируемо, непрерывно дифференцируемо) в точке $\widehat{x}$ тогда и только тогда, когда отображения F_1 и F_2 дифференцируемы (строго дифференцируемы, непрерывно дифференцируемы) в точке $\widehat{x}$ и при этом $F'(\widehat{x}) = (F_1'(\widehat{x}), F_2'(\widehat{x}))$.*

Доказательство. Пусть отображения F_1 и F_2 дифференцируемы в точке $\widehat{x}$. Тогда по определению

$$F(\widehat{x}+h) = (F_1(\widehat{x}+h), F_2(\widehat{x}+h)) = (F_1(\widehat{x}), F_2(\widehat{x})) + (F_1'(\widehat{x}), F_2'(\widehat{x}))h + (r_1(h), r_2(h)),$$

где $\|r_i(h)\|_{Y_i}/\|h\|_X \to 0$ при $h \to 0$, $i = 1, 2$. Следовательно, отображение F дифференцируемо в точке $\widehat{x}$ и $F'(\widehat{x}) = (F_1'(\widehat{x}), F_2'(\widehat{x}))$.

Для проверки того, что из дифференцируемости F следует дифференцируемость F_1 и F_2, надо заметить, что $\Lambda \in \mathcal{L}(X, Y_1 \times Y_2)$ тогда и только тогда, когда $\Lambda = (\Lambda_1, \Lambda_2)$, где $\Lambda_i \in \mathcal{L}(X, Y_i)$, $i = 1, 2$.

Рассуждения, связанные со строгой дифференцируемостью и непрерывной дифференцируемостью, вполне аналогичны. $\square$

Пусть X, Y и Z — нормированные пространства, W — окрестность точки $(\widehat{x}, \widehat{y}) \in X \times Y$ и $F\colon W \to Z$. Если отображение $x \to F(x, \widehat{y})$, определенное на открытом множестве $\{x \in X : (x, y) \in W\}$, дифференцируемо в точке $\widehat{x}$, то соответствующую производную называют *частной производной отображения F по x в точке $(\widehat{x}, \widehat{y})$* и обозначают $F_x(\widehat{x}, \widehat{y})$. Аналогично определяется частная производная F по y в точке $(\widehat{x}, \widehat{y})$, которую обозначаем $F_y(\widehat{x}, \widehat{y})$.

Предложение 2. *Если отображение $F\colon W \to Z$ дифференцируемо в точке $(\widehat{x}, \widehat{y})$, то оно имеет в этой точке частные производные по x и y, и для всех $\xi \in X$ и $\eta \in Y$ справедливо представление*

$$F'(\widehat{x}, \widehat{y})[\xi, \eta] = F_x(\widehat{x}, \widehat{y})[\xi] + F_y(\widehat{x}, \widehat{y})[\eta]^1$$

Доказательство. Так как F дифференцируемо в точке $(\widehat{x}, \widehat{y})$, то, в частности, справедливо представление $F(\widehat{x}+h, \widehat{y}) = F(\widehat{x}, \widehat{y}) + F'(\widehat{x}, \widehat{y})(h, 0) + r(h)$, где $\|r(h)\|_Z/\|h\|_X \to 0$ при $h \to 0$. Отсюда, в силу единственности, следует, что линейный оператор из

[1] $F'(\widehat{x}, \widehat{y})[\xi, \eta]$ обозначает действие (значение) оператора $F'(\widehat{x}, \widehat{y})$ на элементе (ξ, η). Аналогично для частных производных.

X в Z, действующий по правилу: $\xi \mapsto F'(\widehat{x}, \widehat{y})[\xi, 0]$ совпадает с $F_x(\widehat{x}, \widehat{y})$. Аналогично, линейный оператор $\eta \mapsto F'(\widehat{x}, \widehat{y})[0, \eta]$ из Y в Z совпадает с $F_y(\widehat{x}, \widehat{y})$. Но тогда $F'(\widehat{x}, \widehat{y})[\xi, \eta] = F'(\widehat{x}, \widehat{y})([\xi, 0] + [0, \eta]) = F_x(\widehat{x}, \widehat{y})[\xi] + F_y(\widehat{x}, \widehat{y})[\eta]$. $\square$

Простые примеры показывают, что существование частных производных у отображения в данной точке не только не гарантирует его дифференцируемость в этой точке, но даже его непрерывность. Однако справедлива следующая

Теорема 3 (о полном дифференциале)**.** *Отображение $F\colon W \to Z$ непрерывно дифференцируемо на W тогда и только тогда, когда его частные производные F_x и F_y непрерывно дифференцируемы на W и справедливо равенство $F'(\widehat{x}, \widehat{y})[\xi, \eta] = F_x(\widehat{x}, \widehat{y})[\xi] + F_y(\widehat{x}, \widehat{y})[\eta]$.*

Доказательство см. в [2], п. 2.2.4.

Дадим определение второй производной. Пусть отображение $F'\colon U \to \mathcal{L}(X, Y)$ дифференцируемо в точке $\widehat{x}$. Тогда говорят, что F дважды дифференцируема в $\widehat{x}$ и соответствующую (вторую) производную обозначают $F''(\widehat{x})$. Ясно, что $F''(\widehat{x}) \in \mathcal{L}(X, \mathcal{L}(X, Y))$.

Пространство $\mathcal{L}(X, \mathcal{L}(X, Y))$ изометрически изоморфно пространству $\mathcal{L}^2(X, Y)$ всех непрерывных билинейных (т. е. линейных по каждому аргументу) отображений $B\colon X \times X \to Y$ с нормой $\|B\| = \sup_{\|x_1\|_X \le 1, \|x_2\|_X \le 1} \|B[x_1, x_2]\|_Y$. Действительно, пусть отображение A сопоставляет $\Lambda \in \mathcal{L}(X, \mathcal{L}(X, Y))$ отображение $B\colon X \times X \to Y$, действующее 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. Далее, $\|\Lambda\| = \sup_{\|x_1\|_X \le 1} \|\Lambda x_1\|_{\mathcal{L}(X,Y)} = \sup_{\|x_1\|_X \le 1, \, \|x_2\|_X \le 1} \|\Lambda x_1[x_2]\|_Y = \|B\|$, т. е. A изометрично отображает $\mathcal{L}(X, \mathcal{L}(X, Y))$ в $\mathcal{L}^2(X, Y)$. Линейность и сюръективность A проверяются без труда.

Теорема 4 (Формула Тейлора)**.** *Если отображение $F\colon U \to Y$ дважды дифференцируемо в точке $\widehat{x}$, то $F''(\widehat{x})$ — симметричное билинейное отображение и справедлива формула Тейлора*

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

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

1.2.2. Дифференцируемость некоторых отображений

Рассмотрим здесь три примера дифференцируемых отображений.

Пример 1. *Отображение конечномерных пространств.* Пусть U — открытое подмножество $\mathbb{R}^n$ и заданы функции $f_i\colon U \to \mathbb{R}$, $i = 1, \ldots, m$. Тогда определено отображение $F\colon U \to \mathbb{R}^m$ по формуле $F(x) = (f_1(x), \ldots, f_m(x))^T$, $x \in U$. Пусть это отображение дифференцируемо в точке $\widehat{x} \in U$. Рассматривая его как отображение в произведение пространств $\mathbb{R} \times \ldots \times \mathbb{R}$ (m раз), получаем в силу предложения 1, что функции f_i дифференцируемы в точке $\widehat{x}$ и $F'(\widehat{x}) = (f_1'(\widehat{x}), \ldots, f_m'(\widehat{x}))$. Далее, так как U принадлежит произведению пространств $\mathbb{R} \times \ldots \times \mathbb{R}$ (n раз), то согласно предложению 2 функции f_i имеют в точке $\widehat{x}$ частные производные $\partial f_i(\widehat{x})/\partial x_j$, $1 \le i \le m$, $1 \le j \le n$ и для любого $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ справедливо представление $\langle f_i'(\widehat{x}), x\rangle = (\partial f_i(\widehat{x})/\partial x_1)x_1 + \ldots + (\partial f_i(\widehat{x})/\partial x_n)x_n$. Отсюда следует, что

$f_i'(\widehat{x}) = (\partial f_i(\widehat{x})/\partial x_1, \ldots, \partial f_i(\widehat{x})/\partial x_n)$, $i = 1, \ldots, m$, и значит, $F'(\widehat{x})$ — матрица вида $(\partial f_i(\widehat{x})/\partial x_j)_{1 \le i \le m,\, 1 \le j \le n}$, которую называют *матрицей Якоби* отображения F в точке $\widehat{x}$.

Если функция $f\colon U \to \mathbb{R}$ дважды дифференцируемо в $\widehat{x}$, то вторая производная — симметричная билинейная форма с матрицей $(\partial^2 f(x)/\partial x_j \partial x_i)$, $i, j = 1, \ldots, n$, которую называют *матрицей Гесса* или *гессианом функции f в точке $\widehat{x}$*.

Пример 2. *Квадратично линейное отображение.* Пусть X и Y – нормированные пространства, отображение $B \in \mathcal{L}^2(X, Y)$ симметрично, $\Lambda \in \mathcal{L}(X, Y)$ и $a \in Y$. Отображение $F\colon X \to Y$, действующее по правилу

$$F(x) = \frac{1}{2}B[x, x] + \Lambda x + a,$$

называется *квадратично линейным* отображением.

Пусть $x \in X$. Для любого $h \in X$ имеем

$$F(x + h) = \frac{1}{2}B[x + h, x + h] + \Lambda(x + h) + a = \frac{1}{2}B[x, x] + \Lambda x + a \\ + B[x, h] + \Lambda h + \frac{1}{2}B[h, h] = F(x) + B[x, h] + \Lambda h + \frac{1}{2}B[h, h].$$

Так как $\|B[h, h]\| \le \|B\|\|h\|_X^2$, то оператор F дифференцируем в точке x и его производная в этой точке действует по правилу: $F'(x)h = B[x, h] + \Lambda h$.

Вторая производная действует по правилу: $F''(x)[h_1, h_2] = B[h_1, h_2]$ для любых $h_1, h_2 \in X$.

Пример 3. *Обобщенный оператор Немыцкого.* Пусть $[t_0, t_1]$ — конечный отрезок, G — открытое подмножество $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r$ и

$$\mathcal{U} = \{\, (x(\cdot), u(\cdot)) \in C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r) : (t, x(t), u(t) \in G, \quad t \in [t_0, t_1] \,\}.$$

Нетрудно проверить, что это открытое множество в $C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$.

Пусть $f\colon G \to \mathbb{R}^m$ — непрерывная функция переменных $t \in \mathbb{R}$, $x \in \mathbb{R}^n$ и $u \in \mathbb{R}^r$. Определим отображение $F\colon \mathcal{U} \to C([t_0, t_1], \mathbb{R}^m)$ по правилу

$$F(x(\cdot), u(\cdot))(t) = f(t, x(t), u(t)), \quad \forall\, t \in [t_0, t_1],$$

которое называют *обобщенным оператором Немыцкого.*

Если фиксирована пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$, то для сокращения записи пишем $\widehat{f}(t) = f(t, \widehat{x}(t), \widehat{u}(t))$ и аналогично для первых и вторых производных функции f по переменным x и u.

Предложение 3. *Пусть функция f и ее частные производные по x и u непрерывны на G. Тогда обобщенный оператор Немыцкого непрерывно дифференцируем на $\mathcal{U}$ и для любой пары $(\widehat{x}(\cdot), \widehat{u}(\cdot)) \in \mathcal{U}$ его производная действует по правилу*

$$F'(\widehat{x}(\cdot), \widehat{u}(\cdot))[h(\cdot), v(\cdot)](t) = \widehat{f}_x(t)h(t) + \widehat{f}_u(t)v(t), \quad t \in [t_0, t_1], \tag{1}$$

где $(h(\cdot), v(\cdot)) \in C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$.

Доказательство. Пусть $(\widehat{x}(\cdot),\widehat{u}(\cdot))\in\mathcal{U}$. Вычислим частные производные по $x(\cdot)$ и $u(\cdot)$ оператора f в точке $(\widehat{x}(\cdot),\widehat{u}(\cdot))$.

Нетрудно проверить, что существует $\delta_0>0$ такое, что компакт $K=\{\,(t,x,\widehat{u}(t)) : |x-\widehat{x}(t)|\le\delta_0,\ \ t\in[t_0,t_1]\,\}$ принадлежит G.

Пусть $\varepsilon>0$. Функция f_x равномерно непрерывна на K и поэтому найдется $0<\delta\le\delta_0$ такое, что если $|x_1-x_2|<\delta$, то $\|f_x(t,x_1,\widehat{u}(t))-f_x(t,x_2,\widehat{u}(t))\|<\varepsilon$ для всех $(t,x_i,\widehat{u}(t))\in K$, $i=1,2$.

Для любого $t\in[t_0,t_1]$ отображение $g\colon U_{\mathbb{R}^n}(\widehat{x}(t),\delta)\to\mathbb{R}^m$, $g(x)=f(t,x,\widehat{u}(t))-\widehat{f}_x(t)x$, дифференцируемо на $U_{\mathbb{R}^n}(\widehat{x}(t),\delta)$ и его производная в точке x имеет вид: $g'(x)=f_x(t,x,\widehat{u}(t))-\widehat{f}_x(t)$.

Пусть $x(\cdot)\in U_{C^1([t_0,t_1],\mathbb{R}^n)}(\widehat{x}(\cdot),\delta)$. Тогда $x(t)\in U_{\mathbb{R}^n}(\widehat{x}(t),\delta)$, и мы имеем по теореме о среднем, примененной к отображению g (учитывая, что если $x\in[x(t),\widehat{x}(t)]$, то $x\in U_{\mathbb{R}^n}(\widehat{x}(t),\delta)$)

$$
\begin{aligned}
&|f(t,x(t),\widehat{u}(t))-\widehat{f}(t)-\widehat{f}_x(t)(x(t)-\widehat{x}(t))|\\
&\qquad\le\sup_{x\in[x(t),\widehat{x}(t)]}\|f_x(t,x,\widehat{u}(t))-\widehat{f}_x(t)\||x(t)-\widehat{x}(t)|\le\varepsilon|x(t)-\widehat{x}(t)|\\
&\qquad\qquad\le\varepsilon\max(|x(t)-\widehat{x}(t)|,|\dot{x}(t)-\dot{\widehat{x}}(t)|).
\end{aligned}
$$

Так как это верно для любого $t\in[t_0,t_1]$, то отсюда следует, что у отображения F в каждой точке $(\widehat{x}(\cdot),\widehat{u}(\cdot))\in\mathcal{U}$ существует частная производная по $x(\cdot)$, и она действует по правилу: $F_{x(\cdot)}(\widehat{x}(\cdot),\widehat{u}(\cdot))[h(\cdot)](t)=\widehat{f}_x(t)h(t)$, $h(\cdot)\in C^1([t_0,t_1],\mathbb{R}^n)$, $t\in[t_0,t_1]$.

Покажем, что эта частная производная непрерывна на $\mathcal{U}$. Пусть $(\widehat{x}(\cdot),\widehat{u}(\cdot))\in\mathcal{U}$. Существует такое $\delta_1>0$, что $K_1=\{\,(t,x,u) : |x-\widehat{x}(t)|+|u-\widehat{u}(t)|\le\delta_1,\ \ t\in[t_0,t_1]\,\}\subset G$.

Пусть $\varepsilon>0$. В силу равномерной непрерывности f_x на компакте K_1 найдется $\delta_2\le\delta_1/2$ такое, что если $|x_1-x_2|<\delta_2$ и $|u_1-u_2|<\delta_2$ то $\|f_x(t,x_1,u_1)-f_x(t,x_2,u_2)\|<\varepsilon$ для всех $(t,x_i,u_i)\in K_1$, $i=1,2$.

Пусть $(x(\cdot),u(\cdot))\in U_{C^1([t_0,t_1],\mathbb{R}^n)}(\widehat{x}(\cdot),\delta_2)\times U_{C([t_0,t_1],\mathbb{R}^r)}(\widehat{x}(\cdot),\delta_2)$. Для любого $h(\cdot)\in C^1([t_0,t_1],\mathbb{R}^n)$ и $t\in[t_0,t_1]$ имеем

$$
\begin{aligned}
|(F_{x(\cdot)}(\widehat{x}(\cdot),\widehat{u}(\cdot))-F_{x(\cdot)}(x(\cdot),u(\cdot)))[h(\cdot)](t)|&=|(f_x(t,\widehat{x}(t),\widehat{u}(t))-\\
&-f_x(t,x(t),u(t)))h(t)|\le\varepsilon\|h(\cdot)\|_{C^1([t_0,t_1],\mathbb{R}^n)}.
\end{aligned}
$$

Отсюда следует непрерывность $F_{x(\cdot)}$ в произвольной точке $(\widehat{x}(\cdot),\widehat{u}(\cdot))$ и значит, на $\mathcal{U}$.

Совершенно аналогично доказывается, что у отображения F на $\mathcal{U}$ существует непрерывная частная производная по $u(\cdot)$, которая в каждой точке $(\widehat{x}(\cdot),\widehat{u}(\cdot))\in\mathcal{U}$ действует по правилу: $F_{u(\cdot)}(\widehat{x}(\cdot),\widehat{u}(\cdot))[v(\cdot)](t)=\widehat{f}_u(t)v(t)$, $v(\cdot)\in C([t_0,t_1],\mathbb{R}^r)$, $t\in[t_0,t_1]$.

Теперь из теоремы 3 о полном дифференциале следует, что отображение F непрерывно дифференцируемо на $\mathcal{U}$ и справедлива формула (1).

□

§1.3. Метрическая регулярность. Теорема о возмущении. Метрическая регулярность линейного отображения

Всюду далее, чтобы избежать лишних обозначений, считаем, что вводимые отображения определены на всем пространстве. Это никак не ограничивает общности, поскольку мы будем иметь дело с локальными понятиями (дифференцируемость в точке, локальный экстремум и т. д.), так что если отображение априори определено в некоторой окрестности данной точки, то продолжая его как угодно за пределы этой окрестности, ситуация, с локальной точки зрения, не изменится.

Определение 1. *Пусть X и Y — нормированные пространства, $K \subset X$ и $\widehat{x} \in K$. Будем говорить, что отображение $F\colon X \to Y$ метрически регулярно в окрестности точки $\widehat{x}$ относительно множества K, если найдутся окрестности V_1 и V_2 соответственно точек $\widehat{x}$ и $F(\widehat{x})$ и константа $c > 0$ такие, что*

$$\operatorname{dist}(x, F^{-1}(y) \cap K) \le c\|y - F(x)\|_Y \tag{1}$$

для всех $(x, y) \in (V_1 \cap K) \times V_2$ (где $F^{-1}(y) = \{x \in X : F(x) = y\}$).

В случае, если важна величина константы c, то будем говорить о метрической регулярности с константой c.

Если $K = X$, то говорим просто о метрической регулярности в окрестности точки $\widehat{x}$ (опуская слова “относительно множества X”).

Если $V_1 = X$, $V_2 = Y$, то будем говорить, что отображение $F\colon X \to Y$ *глобально метрически регулярно в точке $\widehat{x}$ относительно множества K.*

Из определения сразу следует, что для любой пары $(x, y) \in (V_1 \cap K) \times V_2$ и $\varepsilon > 0$ найдется элемент $x_\varepsilon(x, y) \in K$ такой, что $F(x_\varepsilon(x, y)) = y$ и при этом

$$\|x_\varepsilon(x, y) - x\|_X \le (c + \varepsilon)\|y - F(x)\|_Y.$$

В частности, если $x = \widehat{x}$ и $K = X$, то это есть утверждение о существовании обратной функции у отображения F.

Отметим здесь одно предложение, которое понадобится впоследствии.

Предложение 1. *Если отображение $F\colon X \to Y$ метрически регулярно в окрестности точки $\widehat{x} \in K \subset X$ относительно множества K, то $F(\widehat{x}) \in \operatorname{int} F(K \cap V)$ для любой окрестности V этой точки.*

Доказательство. Пусть окрестности V_1, V_2 и константа c из определения метрической регулярности отображения F, V — окрестность точки $\widehat{x}$, $r_1 > 0$ такое, что $U_X(\widehat{x}, r_1) \subset V$, а $0 < r_2 \le r_1/2c$ таково, что $U_Y(F(\widehat{x}), r_2) \subset V_2$. Покажем, что $U_Y(F(\widehat{x}), r_2) \subset F(K \cap V)$ и тем самым докажем предложение. Пусть $y \in U_Y(F(\widehat{x}), r_2)$. Тогда из (1), при $x = \widehat{x}$, следует, что существует $x(y) \in F^{-1}(y) \cap K$, для которого $\|x(y) - \widehat{x}\|_X \le 2\operatorname{dist}(\widehat{x}, F^{-1}(y) \cap K) \le 2c\|y - F(\widehat{x})\|_Y < 2cr_2 \le r_1$, т. е. $x(y) \in F^{-1}(y) \cap K \cap U_X(\widehat{x}, r_1)$. Таким образом, $y \in F(K \cap U_X(\widehat{x}, r_1)) \subset F(K \cap V)$. $\square$

Теорема 1 (о возмущении). *Пусть X — банахово пространство, Y — нормированное пространство, K — замкнутое подмножество X, отображение $\Phi\colon X \to Y$*

имеет замкнутый график и метрически регулярно в окрестности точки $\widehat{x} \in K$ *относительно множества* K *с константой* c, *а отображение* $G\colon X \to Y$ *липшицево в некоторой окрестности точки* $\widehat{x}$ *с константой Липшица* $l < c^{-1}$. *Тогда отображение* $F = \Phi + G$ *метрически регулярно в окрестности точки* $\widehat{x}$ *относительно множества* K *с константой* $c' = c/(1 - lc)$.

Доказательство. Доказательство проведем для случая, когда отображение Φ глобально метрически регулярно в точке $\widehat{x}$ относительно множества K с константой c, а отображение G липшицево на все пространстве X с константой Липшица $l < c^{-1}$.

Пусть $(x, y) \in K \times Y$. Рассмотрим последовательность $x_k = x_k(\varepsilon)$, $k \in \mathbb{N}$, заданную рекуррентными соотношениями

$$x_k \in \Phi^{-1}(y - G(x_{k-1})) \cap K, \quad x_0 = x, \tag{2}$$

и

$$\|x_k - x_{k-1}\|_X \le (1+\varepsilon)\operatorname{dist}(x_{k-1}, \Phi^{-1}(y - G(x_{k-1})) \cap K), \tag{3}$$

где $\varepsilon > 0$ столь мало, что $(1+\varepsilon)lc < 1$.

Так как отображение Φ глобально метрически регулярно в точке $\widehat{x}$ относительно множества K, то такая последовательность, очевидно, существует.

Покажем, что последовательность $\{x_k\}$ фундаментальна ($k \ge 2$). Используя (3), метрическую регулярность Φ, равенство $\Phi(x_{k-1}) = y - G(x_{k-2})$ (см. (2)), липшицивость G, а затем итерируя процедуру, будем иметь (обозначая $\alpha(\varepsilon) = (1+\varepsilon)lc$)

$$\begin{aligned} \|x_k - x_{k-1}\|_X &\le (1+\varepsilon)\operatorname{dist}(x_{k-1}, \Phi^{-1}(y - G(x_{k-1})) \cap K) \\ &\le (1+\varepsilon)c\|y - G(x_{k-1}) - \Phi(x_{k-1})\|_Y = (1+\varepsilon)c\|G(x_{k-1}) - G(x_{k-2})\|_Y \\ &\le (1+\varepsilon)cl\|x_{k-1} - x_{k-2}\|_X = \alpha(\varepsilon)\|x_{k-1} - x_{k-2}\|_X \\ &\le \ldots \le \alpha^{k-1}(\varepsilon)\|x_1 - x\|_X. \end{aligned} \tag{4}$$

Последовательность $\{x_k\}$ фундаментальна. Действительно, для любых $k, m \in \mathbb{N}$, используя неравенство треугольника, (4), формулу для суммы геометрической прогрессии, неравенство (3) при $k = 1$ и метрическую регулярность отображения Φ, получим оценку

$$\begin{aligned} \|x_{k+m} - x_k\|_X &\le \|x_{k+m} - x_{k+m-1}\|_X + \ldots + \|x_{k+1} - x_k\|_X \\ &\le (\alpha^{k+m-1}(\varepsilon) + \ldots + \alpha^k(\varepsilon))\|x_1 - x\|_X < \frac{\alpha^k(\varepsilon)}{1 - \alpha(\varepsilon)}\|x_1 - x\|_X \\ &\le \frac{\alpha^k(\varepsilon)(1+\varepsilon)}{1 - \alpha(\varepsilon)}\operatorname{dist}(x, \Phi^{-1}(y - G(x)) \cap K) \le \frac{\alpha^k(\varepsilon)(1+\varepsilon)c}{1 - \alpha(\varepsilon)}\|y - F(x)\|_Y, \end{aligned} \tag{5}$$

откуда следует фундаментальность последовательности $\{x_k\}$.

Положим $\varphi_\varepsilon(x, y) = \lim_{k\to\infty} x_k$ ($x_k = x_k(\varepsilon)$). Включение (2) равносильно равенству $\Phi(x_k) = y - G(x_{k-1})$. Переходя здесь к пределу при $k \to \infty$ и обозначая $\widehat{y}_\varepsilon = \lim_{k\to\infty} \Phi(x_k)$, получаем, что $\widehat{y}_\varepsilon = y - G(\varphi_\varepsilon(x, y))$. Так как график Φ замкнут, то $\widehat{y}_\varepsilon = \Phi(\varphi_\varepsilon(x, y))$ и тем самым $F(\varphi_\varepsilon(x, y)) = y$. Так как $x_k \in K$, $k \in \mathbb{N}$, и K замкнуто, то $\varphi_\varepsilon(x, y) \in K$.

Полагая $k = 0$ в (5) и переходя к пределу при $m \to \infty$, приходим к соотношению

$$\|\varphi_\varepsilon(x, y) - x\|_X \le ((1 + \varepsilon)c/(1 - \alpha(\varepsilon)))\|y - F(x)\|_Y.$$

Отсюда следует, что

$$\operatorname{dist}(x, F^{-1}(y) \cap K) \le ((1 + \varepsilon)c/(1 - \alpha(\varepsilon)))\|y - F(x)\|_Y.$$

Это справедливо для всех достаточно малых $\varepsilon > 0$ и поэтому

$$\operatorname{dist}(x, F^{-1}(y) \cap K) \le c'\|y - F(x)\|_Y$$

для всех $(x, y) \in K \times Y$, где $c' = c/(1 - lc)$, т. е. отображение F глобально метрически регулярно в точке $\widehat{x}$ относительно множества K (с константой c'). □

Следующее утверждение касается важного частного случая — метрической регулярности линейного отображения, и будет вместе с предыдущей теоремой использовано при доказательстве основной теоремы из главы 2. Его доказательство носит технических характер и здесь не приводится.

Лемма 1 (о метрической регулярности линейного отображения)**.** *Пусть X и Y — банаховы пространства, $\Lambda \in \mathcal{L}(X, Y)$, K — выпуклое замкнутое подмножество X, $\widehat{x} \in K$ и $0 \in \operatorname{int} \Lambda(K - \widehat{x})$. Тогда отображение Λ метрически регулярно в окрестности точки $\widehat{x}$ относительно множества K.*

§1.4. Обобщенная теорема об обратной функции и ее следствия: теорема о поправке и теорема Люстерника

Теорема 1 (Обобщенная теорема об обратной функции)**.** *Пусть X и Y — банаховы пространства, K — выпуклое замкнутое подмножество X, отображение $F\colon X \to Y$ строго дифференцируемо в точке $\widehat{x} \in K$ и $0 \in \operatorname{int} F'(\widehat{x})(K - \widehat{x})$. Тогда F метрически регулярно в окрестности $\widehat{x}$ относительно множества K.*

Доказательство. Положим $\Phi = F'(\widehat{x})$ и $G = F - \Phi$. Отображение Φ, по лемме о метрической регулярности линейного отображения, метрически регулярно в окрестности точки $\widehat{x}$ относительно множества K с некоторой константой $c > 0$ и, очевидно, имеет замкнутый график. Из строгой дифференцируемости F в точке $\widehat{x}$ легко следует, что найдется окрестность точки $\widehat{x}$, в которой отображение G липшицево с константой Липшица, меньшей c^{-1}. Но тогда по теореме о возмущении (см. теорему 1 §1.3) отображение $F = \Phi + G$ будет метрически регулярно в окрестности точки $\widehat{x}$ относительно множества K. □

Теперь выведем из этой теоремы два важных следствия, которые будут использоваться в дальнейшем.

Если $F\colon X \to Y$ и $\widehat{x} \in X$, то для краткости записи введем обозначение

$$M(\widehat{x}) = F^{-1}(F(\widehat{x})) = \{\, x \in X : F(x) = F(\widehat{x}) \,\},$$

т. е. $M(\widehat{x})$ — "поверхность уровня" отображения F.

Теорема 2 (Теорема о поправке)**.** *Пусть X и Y — банаховы пространства, отображение $F\colon X \to Y$ строго дифференцируемо в точке $\widehat{x}$ и $\operatorname{Im} F'(\widehat{x}) = Y$. Тогда существуют окрестность V точки $\widehat{x}$, отображение $\varphi\colon V \to X$ и константа $c > 0$ такие, что для всех $x \in V$ выполняются соотношения*

$$F(x + \varphi(x)) = F(\widehat{x}), \quad \|\varphi(x)\|_X \le 2c\|F(x) - F(\widehat{x})\|_Y.$$

Доказательство. Если в теореме $K = X$, то условие $0 \in \operatorname{int} F'(\widehat{x})(X - \widehat{x}) = \operatorname{int} F'(\widehat{x})X$, очевидно, равносильно тому, что $\operatorname{Im} F'(\widehat{x}) = Y$. Следовательно, согласно теореме 1, отображение F метрически регулярно в окрестности точки $\widehat{x}$ и значит, существуют окрестность V точки $\widehat{x}$ и константа $c > 0$ такие, что $\operatorname{dist}(x, M(\widehat{x})) \le c\|F(x) - F(\widehat{x})\|_Y$ для всех $x \in V$. Положим $\varphi(x) = \psi(x) - x$. Тогда $F(x + \varphi(x)) = F(\widehat{x})$ и $\|\varphi(x)\|_X \le 2\operatorname{dist}(x, M(\widehat{x})) \le 2c\|F(x) - F(\widehat{x})\|_Y$ для всех $x \in V$. $\square$

Название теоремы может быть оправдано тем, что для каждого $x \in V$ существует малая "поправка" $\varphi(x)$, переводящая x на поверхности уровня отображения F.

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

Теорема 3 (Теорема Люстерника)**.** *Пусть выполнены условия теоремы о поправке. Тогда*

$$T_{\widehat{x}}M(\widehat{x}) = \operatorname{Ker} F'(\widehat{x}).$$

Доказательство. Докажем, что $T_{\widehat{x}}M(\widehat{x}) \subset \operatorname{Ker} F'(\widehat{x})$. Пусть $h \in T_{\widehat{x}}M$ и отображение r из определения касательного вектора. Тогда вследствие дифференцируемости F в точке $\widehat{x}$ имеем $0 = F(\widehat{x} + th + r(t)) - F(\widehat{x}) = tF'(\widehat{x})h + o(t)$, откуда, деля на t и переходя к пределу при $t \to 0$, получаем, что $F'(\widehat{x})h = 0$, т. е. $h \in \operatorname{Ker} F'(\widehat{x})$.

Докажем теперь, что $\operatorname{Ker} F'(\widehat{x}) \subset T_{\widehat{x}}M(\widehat{x})$. Пусть $h \in \operatorname{Ker} F'(\widehat{x})$, а V — окрестность $\widehat{x}$ и отображение φ из утверждения теоремы о поправке. Ясно, что существует $\varepsilon > 0$ такое, что $\widehat{x} + th \in V$ для всех $t \in (-\varepsilon, \varepsilon)$. Тогда, согласно доказанному, $F(\widehat{x} + th + \varphi(\widehat{x} + th)) = F(\widehat{x})$ и $\|\varphi(\widehat{x} + th)\|_X \le 2c\|F(\widehat{x} + th) - F(\widehat{x})\|_Y = \|F(\widehat{x}) + tF'(\widehat{x})h + o(t) - F(\widehat{x})\|_Y = o(t)Y$. Если обозначить $r(t) = \varphi(\widehat{x} + th)$, то получим, что $\widehat{x} + th + r(t) \in M(\widehat{x})$ для всех $t \in (-\varepsilon, \varepsilon)$ и $r(t) = o(t)$ при $t \to 0$, т. е. $h \in T_{\widehat{x}}M(\widehat{x})$. $\square$

Глава 2. Условия экстремума в задачах нелинейного и выпуклого программирования

В этой главе будут получены необходимые условия минимума для различных классов экстремальных задач. Все эти условия подчиняются так называемому *принципу Лагранжа*, заключающегося в том, что решение исходной задачи удовлетворяет необходимым условиям минимума функции Лагранжа на множестве тех ограничений, которые в нее не вошли.

Сам Ж. Л. Лагранж высказал этот принцип для конечномерных задач в своей работе "Теория аналитических функций", 1897 г. Вот его слова: *Можно высказать следующий общий принцип. Если ищется максимум или минимум некоторой функции многих переменных при условии, что между этими переменными имеется связь, задаваемая одним или несколькими уравнениями, нужно прибавить к минимизируемой функции функции, задающие уравнения связи, умноженные на неопределенные множители, и искать затем максимум или минимум построенной суммы, как если бы переменные были независимы. Полученные уравнения, присоединенные к уравнениям связи, послужат для определения всех неизвестных.*

Из дальнейшего будет видно, что этот принцип (если придать ему чуть более расширительное толкование) носит универсальный характер.

Как уже говорилось выше, не ограничивая общности, считаем, что все функции и отображения определены на всем пространстве.

§2.1. Общая задача нелинейного программирования

2.1.1. Постановка задачи

Пусть K — непустое подмножество нормированного пространства X, $f_0\colon X \to \mathbb{R}$ и $F\colon X \to \mathbb{R}^m$. Задачу

$$f_0(x) \to \min, \quad F(x) = 0, \quad x \in K \tag{P_0}$$

будем называть *общей задачей нелинейного программирования.*

Предложение 1. *Если $\widehat{x}$ — локальный минимум в задаче (P_0) и $\Phi = (f_0, F)\colon X \to \mathbb{R} \times \mathbb{R}^m$, то найдется такая окрестность V точки $\widehat{x}$, что $\Phi(\widehat{x}) \in \partial\Phi(K \cap V)$.*

Если $\widehat{x}$ — глобальный минимум, то $\Phi(\widehat{x}) \in \partial\Phi(K)$.

Доказательство. Пусть $\widehat{x}$ — локальный минимум в задаче (P_0). По определению найдется такая окрестность V точки $\widehat{x}$, что $f_0(x) \geq f_0(\widehat{x})$ для всех $x \in K \cap V$, для которых $F(x) = 0$. Ясно, что $\Phi(\widehat{x}) = (f_0(\widehat{x}), F(\widehat{x})) \in \Phi(K \cap V)$. Далее, $(f_0(\widehat{x}) - \varepsilon, F(\widehat{x})) \notin \Phi(K \cap V)$ ни для какого $\varepsilon > 0$, ибо иначе нашлась бы точка $x_\varepsilon \in K \cap V$ такая, что $f_0(x_\varepsilon) = f_0(\widehat{x}) - \varepsilon < f_0(\widehat{x})$ и $F(x_\varepsilon) = F(\widehat{x}) = 0$ в противоречии с предположением. Следовательно, $\Phi(\widehat{x}) \in \partial\Phi(K \cap V)$.

В случае глобального минимума рассуждения такие же ($V = X$). $\square$

Сопоставим задаче (P_0) *функцию Лагранжа* $\mathcal{L}\colon X\times\mathbb{R}\times(\mathbb{R}^m)^*\to\mathbb{R}$, определенную формулой

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

где число λ_0 и линейный функционал $y^*\in Y^*$ называются *множителями Лагранжа.*

2.1.2. Правило множителей Лагранжа

Теорема 1 (Правило множителей Лагранжа — необходимые условия минимума первого порядка в задаче (P_0))**.** *Пусть $\widehat{x}$ — локальный минимум в задаче (P_0). Если X — банахово пространство, K — выпуклое замкнутое подмножество X, функция f_0 и отображение F строго дифференцируемы в точке $\widehat{x}$, то найдутся, не равные одновременно нулю, множители Лагранжа $\lambda_0\geq 0$ и y^* такие, что*

$$-\mathcal{L}_x(\widehat{x},\lambda_0,y^*)\in N_K(\widehat{x})\quad\Leftrightarrow\quad\langle\lambda_0 f_0'(\widehat{x})+(F'(\widehat{x}))^*y^*,x-\widehat{x}\rangle\geq 0,\ \ \forall\, x\in K. \tag{1}$$

Если $0\in\operatorname{int}F'(\widehat{x})(K-\widehat{x})$, то $\lambda_0\neq 0$.

Доказательство. Сопоставим задаче (P_0) следующую задачу

$$f(z)\to\min,\quad F(z)=0,\quad z\in K_1 \tag{P_0'}$$

переменной $z=(x,y)\in Z=X\times\mathbb{R}$, где $f(z)=f_0(x)+y$, $F(z)=F(x)$ и $K_1=K\times\mathbb{R}_+$.

Элементарная проверка показывает, что если $\widehat{x}$ — локальный минимум в задаче (P_0), то $\widehat{z}=(\widehat{x},0)$ — локальный минимум в задаче (P_0').

Пусть $\Phi=(f,F)$. Покажем, что $0\in\partial\Phi'(\widehat{z})(K_1-\widehat{z})$. Действительно, если это не так, то $0\in\operatorname{int}\Phi'(\widehat{z})(K_1-\widehat{z})$. Тогда согласно обобщенной теореме об обратной функции (теорема 1 §1.4) отображение Φ метрически регулярно в окрестности $\widehat{z}$ относительно множества K_1. Следовательно, по предложению 1 §1.3, $\Phi(\widehat{z})\in\operatorname{int}\Phi(K_1\cap V)$ для любой окрестности V точки $\widehat{z}$. Но, если $\widehat{z}$ — локальный минимум, то это противоречит предложению 1.

Итак, $0\in\partial\Phi'(\widehat{z})(K_1-\widehat{z})$. Множество $A=\Phi'(\widehat{z})(K_1-\widehat{z})$ выпукло, как образ выпуклого множества при линейном отображении. Тогда в силу критерия граничной точки (предложения 2 §1.1) $N_A\neq\{0\}$, т. е. существует ненулевой функционал $z^*\in(\mathbb{R}\times\mathbb{R}^m)^*$ такой, что справедливо неравенство $\langle z^*,z\rangle\leq 0$ для любого $z\in A$, или $\langle z^*,\Phi'(\widehat{z})(z-\widehat{z})\rangle\leq 0$ для любого $z=(x,y)\in K\times\mathbb{R}_+$. Для ненулевого функционала $-z^*$, полагая $-z^*=(\lambda_0,y^*)$, будем иметь, используя предложение 1 из §1.2 и лемму 1 из §1.1

$$\begin{aligned}0\leq\langle -z^*,\Phi'(\widehat{z})(z-\widehat{z})\rangle=\langle -z^*,(f'(\widehat{z}),F'(\widehat{z}))(z-\widehat{z})\rangle=\langle\lambda_0 f_0'(\widehat{x}),x-\widehat{x}\rangle+\lambda_0 y\\ +\langle y^*,F'(\widehat{x})(x-\widehat{x})\rangle=\lambda_0 y+\langle\mathcal{L}_x(\widehat{x},\lambda_0,y^*),x-\widehat{x}\rangle\end{aligned}\tag{2}$$

для всех $(x,y)\in K\times\mathbb{R}_+$.

Если $x=\widehat{x}$, то из (2) следует, что $\lambda_0\geq 0$. Полагая $y=0$, получаем включение $-\mathcal{L}_x(\widehat{x},\lambda_0,y^*)\in N_K(\widehat{x})$, которое равносильно в силу того, что $\langle\lambda_0 f_0'(\widehat{x}),x-\widehat{x}\rangle+\langle y^*,F'(\widehat{x})(x-\widehat{x})\rangle=\langle\lambda_0 f_0'(\widehat{x})+(F'(\widehat{x}))^*y^*,x-\widehat{x}\rangle$ неравенству $\langle\lambda_0 f_0'(\widehat{x})+(F'(\widehat{x}))^*y^*,x-\widehat{x}\rangle\geq 0$ для всех $x\in K$.

Докажем второе утверждение теоремы. Пусть $0 \in \operatorname{int} F'(\widehat{x})(K-\widehat{x})$. Допустим, что $\lambda_0 = 0$. Тогда $y^* \neq 0$ и из (2) следует, что $\langle -y^*, F'(\widehat{x})(x-\widehat{x})\rangle \leq 0$ для любого $x \in K$, или равносильно, $\langle -y^*, y\rangle \leq 0$ для любого $y \in B = F'(\widehat{x})(K-\widehat{x})$, т. е. $-y^* \in N_B(0)$. Следовательно, по критерию граничной точки $0 \in \partial F'(\widehat{x})(K-\widehat{x})$ в противоречии с предположением. □

§2.2. Гладкие задачи с ограничениями типа равенств

2.2.1. Постановка задачи

Здесь мы рассмотрим задачу, являющуюся естественным обобщением классической задачи, о которой говорил Лагранж, и исследуем ее наиболее полно, приводя также необходимые и достаточные условия минимума второго порядка.

Пусть X — нормированное пространство, $f_0\colon X \to \mathbb{R}$ и $F\colon X \to \mathbb{R}^m$. Задачу

$$f_0(x) \to \min, \quad F(x) = 0 \tag{P_1}$$

называют *задачей с ограничениями типа равенств.*

Функция Лагранжа здесь имеет тот же вид, что и в задаче (P_0).

2.2.2. Правило множителей Лагранжа

Теорема 1 (Правило множителей Лагранжа — необходимые условия минимума первого порядка в задаче (P_1))**.** *Пусть $\widehat{x}$ — локальный минимум в задаче (P_1). Если X — банахово пространство, функция f_0 и отображение F строго дифференцируемы в точке $\widehat{x}$, то найдутся, не равные одновременно нулю, множители Лагранжа $\lambda_0 \geq 0$ и y^* такие, что*

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

Если $\operatorname{Im} F'(\widehat{x}) = \mathbb{R}^m$*, то* $\lambda_0 \neq 0$.

Доказательство. Соотношения (1) сразу следует из теоремы о необходимых условиях минимума в задаче (P_0) при $K = X$, так как, очевидно, $N_X(x) = \{0\}$ для любого $x \in X$. Меняя знак, можно считать, что $\lambda_0 \geq 0$.

Пусть $\operatorname{Im} F'(\widehat{x}) = \mathbb{R}^m$. Если $\lambda_0 = 0$, то по доказанному $(F'(\widehat{x}))^* y^* = 0$. Но оператор $F'(\widehat{x})$ сюръективен, следовательно, его сопряженный инъективен (см. п. 1.1.1) и значит, $y^* = 0$, что невозможно. □

Если отображение $F\colon X \to \mathbb{R}^m$ задано набором функций $f_i\colon X \to \mathbb{R}$, $i = 1, ..., m$, т. е. $F(x) = (f_1(x), \ldots, f_m(x))^T$ для любого $x \in X$, то задача (P_1) приобретает вид

$$f_0(x) \to \min, \quad f_i(x) = 0, \quad i = 1, \ldots, m. \tag{P_1'}$$

Поскольку линейные функционалы на $\mathbb{R}^m$ суть вектор-строки $(\lambda_1, \ldots, \lambda_m)$ (см. п. 1.2.2), то функция Лагранжа задачи (P_1') имеет вид $\mathcal{L}(x, \lambda_0, \lambda) = \sum_{i=0}^m \lambda_i f_i(x)$, где $\lambda = (\lambda_1, \ldots, \lambda_m)$.

Следствие 1 (Правило множителей Лагранжа — необходимые условия минимума первого порядка в задаче (P_1'))**.** *Пусть* $\widehat{x}$ *— локальный минимум в задаче* (P_1')*. Если* X *— банахово пространство, функции* f_i*,* $i = 0, 1, \ldots, m$*, строго дифференцируемы в точке* $\widehat{x}$*, то найдутся, не равные одновременно нулю, множители Лагранжа* $\lambda_0 \geq 0$ *и* $\lambda = (\lambda_1, \ldots, \lambda_m)$ *такие, что*

$$\mathcal{L}_x(\widehat{x}, \lambda_0, \lambda) = 0 \;\Leftrightarrow\; \sum_{i=0}^{m} \lambda_i f_i'(\widehat{x}) = 0.$$

Если векторы $f_1'(\widehat{x}), \ldots, f_m'(\widehat{x})$ *линейно независимы, то* $\lambda_0 \neq 0$*.*

Доказательство сразу следует из предыдущей теоремы, если учесть, что $y^* = (\lambda_1, \ldots, \lambda_m)$ и поэтому с учетом леммы 2 из §1.1 имеем для любого $x \in X$

$$0 = \lambda_0 \langle f_0'(\widehat{x}), x \rangle + \langle y^*, F'(\widehat{x})x \rangle = \lambda_0 \langle f_0'(\widehat{x}), x \rangle + \sum_{i=0}^{m} \lambda_i \langle f_i'(\widehat{x}), x \rangle = \langle \sum_{i=0}^{m} \lambda_i f_i'(\widehat{x}), x \rangle,$$

что равносильно первому утверждению следствия.

Последнее утверждение очевидно.

2.2.3. Условия минимума второго порядка

Теорема 2 (Необходимые условия минимума второго порядка в задаче с ограничениями типа равенств)**.** *Пусть* $\widehat{x}$ *— локальный минимум в задаче* (P_1)*. Если* X *— банахово пространство, функция* f_0 *и отображение* F *дважды дифференцируемы в точке* $\widehat{x}$ *и* $\operatorname{Im} F'(\widehat{x}) = \mathbb{R}^m$*, то найдется множитель Лагранжа* $y^* \in (\mathbb{R}^m)^*$ *такой, что*

$$\mathcal{L}_x(\widehat{x}, 1, y^*) = 0 \;\Leftrightarrow\; f_0'(\widehat{x}) + (F'(\widehat{x}))^* y^* = 0 \tag{2}$$

и

$$\mathcal{L}_{xx}(\widehat{x}, 1, y^*)[h, h] \geq 0, \;\; \forall\, h \in \operatorname{Ker} F'(\widehat{x}) \;\Leftrightarrow\; f_0''(\widehat{x})[h, h] + \langle y^*, F''(\widehat{x})[h, h] \rangle \geq 0, \;\; \forall\, h \in \operatorname{Ker} F'(\widehat{x}). \tag{3}$$

Доказательство. Из дважды дифференцируемости отображений f_0 и F в точке $\widehat{x}$ следует непрерывность их производных в $\widehat{x}$, а значит, их строгая дифференцируемость в данной точке (см. следствие 1 §1.2). Следовательно, равенство (2) сразу вытекает из предыдущей теоремы (где можно считать, что $\lambda_0 = 1$).

Докажем (3). Пусть $h \in \operatorname{Ker} F'(\widehat{x})$. Отображение F удовлетворяет условиям теоремы Люстерника (см. следствие 3 §2.2) и поэтому $h \in T_{\widehat{x}} M(\widehat{x})$ ($M(\widehat{x}) = \{x \in X : F(x) = F(\widehat{x}) = 0\}$), т. е. существуют $\varepsilon > 0$ и отображение $r\colon (-\varepsilon, \varepsilon) \to X$ такие, что $F(\widehat{x} + th + r(t)) = 0$ для $t \in (-\varepsilon, \varepsilon)$ и $r(t) = o(t)$ при $t \to 0$. Таким образом, элементы $\widehat{x} + th + r(t)$, $t \in (-\varepsilon, \varepsilon)$, допустимы в (P_1) и так как $\widehat{x}$ — локальный минимум в этой задаче, то $f_0(\widehat{x} + th + r(t)) \geq f_0(\widehat{x})$ для достаточно малых по модулю t. Теперь по формуле Тейлора имеем (учитывая (2))

$$\begin{aligned} 0 \leq f_0(\widehat{x} + th + r(t)) - f_0(\widehat{x}) &= \mathcal{L}(\widehat{x} + th + r(t), 1, y^*) - \mathcal{L}(\widehat{x}, 1, y^*) \\ &= \frac{1}{2} \mathcal{L}_{xx}(\widehat{x}, 1, y^*)[th + r(t), th + r(t)] + o(t^2) = \frac{t^2}{2} \mathcal{L}_{xx}(\widehat{x}, 1, y^*)[h, h] + o(t^2), \end{aligned}$$

откуда, деля на t^2 и переходя к пределу при $t \to 0$, получаем (3). □

Теорема 3 (Достаточные условия минимума второго порядка в задаче с ограничениями типа равенств)**.** *Пусть в задаче* (P_1) X *— банахово пространство, функция* f_0 *и отображение* F *дважды дифференцируемы в допустимой точке* $\widehat{x}$*. Тогда, если найдутся множитель Лагранжа* $y^* \in (\mathbb{R}^m)^*$ *и число* $\alpha > 0$ *такие, что*

$$\mathcal{L}_x(\widehat{x}, 1, y^*) = 0 \tag{4}$$

и

$$\mathcal{L}_{xx}(\widehat{x}, 1, y^*)[h, h] \geq \alpha \|h\|_X^2, \ \ \forall\, h \in \operatorname{Ker} F'(\widehat{x}), \tag{5}$$

то $\widehat{x}$ *— локальный минимум в задаче* (P_1).

Доказательство. Рассмотрим отображение $F'(\widehat{x})\colon X \to \operatorname{Im} F'(\widehat{x})$. Для него, очевидно, выполнены условия теоремы о поправке (см. теорему 2 §2.2) в любой точке (в частности, в нуле), $Y = \operatorname{Im} F'(\widehat{x})$. Согласно этой теореме существуют окрестность нуля V_0, отображение $\varphi\colon V_0 \to X$ и константа $K > 0$ такие, что

$$F'(\widehat{x})(x + \varphi(x)) = F'(\widehat{x})0 = 0 \tag{6}$$

и

$$\|\varphi(x)\|_X \leq K|F'(\widehat{x})x| \tag{7}$$

для всех $x \in V_0$.

Можно считать, что в любой окрестности $\widehat{x}$ есть допустимые в (P_1) точки (иначе $\widehat{x}$ была бы изолированной точкой и автоматически локальным минимумом). Пусть $x \in V_0$ и $\widehat{x} + x$ — допустимый элемент в задаче (P_1). Тогда по формуле Тейлора

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

Отсюда следует, что для достаточно малых x справедливо неравенство $|F'(\widehat{x})x| \leq ((1/2)\|F''(\widehat{x})\| + 1)\|x\|_X^2$, а тогда из (7) получаем, что

$$\|\varphi(x)\|_X \leq \gamma \|x\|_X^2, \tag{8}$$

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

Обозначая, для краткости, $\mathcal{L}(x) = \mathcal{L}(x, 1, y^*)$ и $B = \|\mathcal{L}_{xx}(\widehat{x})\|$ и учитывая (4), допустимость точки $\widehat{x} + x$, включение $x + \varphi(x) \in \operatorname{Ker} F'(\widehat{x})$ и оценку (5), будем иметь

$$f_0(\widehat{x}+x) - f_0(\widehat{x}) = \mathcal{L}(\widehat{x}+x) - \mathcal{L}(\widehat{x}) = \frac{1}{2}\mathcal{L}_{xx}(\widehat{x})[x+\varphi(x)-\varphi(x), x+\varphi(x)-\varphi(x)] + o(\|x\|_X^2)$$
$$= \frac{1}{2}(\mathcal{L}_{xx}(\widehat{x})[x+\varphi(x), x+\varphi(x)] - 2\mathcal{L}_{xx}(\widehat{x})[x+\varphi(x), \varphi(x)] + \mathcal{L}_{xx}(\widehat{x})[\varphi(x), \varphi(x)]) + o(\|x\|_X^2)$$
$$\geq \frac{1}{2}(\alpha\|x + \varphi(x)\|_X^2 - 2B\|x + \varphi(x)\|_X\|\varphi(x)\|_X - B\|\varphi(x)\|_X^2) + o(\|x\|_X^2).$$

Нетрудно проверить, что в силу оценки (8) выражение справа положительно для достаточно малых x и поэтому $\widehat{x}$ — локальный минимум. □

Заметим, что доказано несколько больше: существует такое $\kappa > 0$ и окрестность нуля V_1 в X, что $f_0(\widehat{x} + x) - f_0(\widehat{x}) \geq \kappa\|x\|_X^2$ для всех таких $x \in V_1$, для которых точка $\widehat{x} + x$ допустима в задаче (P_1).

§2.3. Гладкие задачи с ограничениями вида $G(x) \in Q$

2.3.1. Постановка задачи

Пусть X — нормированное пространство, $\widehat{x} \in X$, $f_0\colon X \to \mathbb{R}$, $G\colon X \to \mathbb{R}^m$ и Q — непустое подмножество $\mathbb{R}^m$. Рассмотрим экстремальную задачу

$$f_0(x) \to \min, \quad G(x) \in Q. \tag{P_2}$$

Сопоставим задаче (P_2) *функцию Лагранжа* $\mathcal{L}\colon X \times \mathbb{R} \times (\mathbb{R}^m)^* \to \mathbb{R}$, определенную формулой

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

где $(\lambda_0, y^*) \in \mathbb{R} \times (\mathbb{R}^m)^*$.

2.3.2. Правило множителей Лагранжа

Теорема 1 (Правило множителей Лагранжа — необходимые условия минимума первого порядка в задаче (P_2))**.** *Пусть $\widehat{x}$ — локальный минимум в задаче (P_2). Если X — банахово пространство, функция f_0 и отображение G строго дифференцируемы в точке $\widehat{x}$, множество Q выпукло и замкнуто, то найдутся множители Лагранжа $\lambda_0 \geq 0$ и $y^* \in (\mathbb{R}^m)^*$, не равные одновременно нулю, такие, что*

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

и

$$y^* \in N_Q(G(\widehat{x})). \tag{2}$$

Если $0 \in \operatorname{int}(\operatorname{Im} G'(\widehat{x}) + G(\widehat{x}) - Q)$, *то* $\lambda_0 \neq 0$.

Доказательство. Сведем задачу (P_2) к задаче вида (P_0). Введем новую переменную $z = (x, y) \in Z = X \times \mathbb{R}^m$ и положим $f(z) = f_0(x)$, $F(z) = G(x) - y$, $K = X \times Q$.

Рассмотрим задачу

$$f(z) \to \min, \quad F(z) = 0, \quad z \in K.$$

Легко проверить, что если $\widehat{x}$ — локальный минимум в задаче (P_2), то точка $\widehat{z} = (\widehat{x}, G(\widehat{x}))$ — локальный минимум в данной задаче. Тогда согласно теореме 1 §2.2 найдутся множители Лагранжа $\lambda_0 \geq 0$ и $y^* \in (\mathbb{R}^m)^*$, не равные одновременно нулю, такие, что

$$\begin{aligned} 0 \leq \langle \lambda_0 f'(\widehat{z}), z - \widehat{z}\rangle + \langle y^*, F'(\widehat{z})(z - \widehat{z})\rangle = \langle \lambda_0 f_0'(\widehat{x}), x - \widehat{x}\rangle + \langle y^*, G'(\widehat{x})(x - \widehat{x}) \\ - (y - G(\widehat{x}))\rangle = \langle \lambda_0 f_0'(\widehat{x}) + (G'(\widehat{x}))^* y^*, x - \widehat{x}\rangle - \langle y^*, y - G(\widehat{x})\rangle \end{aligned} \tag{3}$$

для всех $x \in X$ и $y \in Q$.

Полагая здесь $y = G(\widehat{x})$, получаем, что $\langle \lambda_0 f_0'(\widehat{x}) + (G'(\widehat{x}))^* y^*, x - \widehat{x}\rangle \geq 0$ для всех $x \in X$ и значит, $\lambda_0 f_0'(\widehat{x}) + (G'(\widehat{x}))^* y^* = 0$.

Если $x = \widehat{x}$ в (3), то получаем соотношение (2).

Докажем последнее утверждение теоремы. Пусть $0 \in \operatorname{int}(\operatorname{Im} G'(\widehat{x}) + G(\widehat{x}) - Q)$. Предположим, что $\lambda_0 = 0$. Тогда $y^* \neq 0$ и из (3) следует, что $\langle y^*, G'(\widehat{x})(x - \widehat{x}) + G(\widehat{x}) - y\rangle \geq 0$ для всех $x \in X$ и $y \in Q$, или $\langle y^*, y\rangle \geq 0$ для всех $y \in A = \operatorname{Im} G'(\widehat{x}) + G(\widehat{x}) - Q$, т. е. $-y^* \in N_A(0)$ и значит, $N_A(0) \neq \{0\}$. Согласно критерию граничной точки (см. предложение 2 §1.1) ноль — граничная точка множества $\operatorname{Im} G'(\widehat{x}) + G(\widehat{x}) - Q$, что противоречит предположению. $\square$

2.3.3. Гладкие задачи с ограничениями типа равенств и неравенств

Пусть X — нормированное пространство, $f_i\colon X \to \mathbb{R}$, $i = 0, 1, \ldots, m_1$, и $F\colon X \to \mathbb{R}^m$. Задача

$$f_0(x) \to \min, \quad f_i(x) \le 0, \quad 1 \le i \le m_1, \quad F(x) = 0 \tag{P_3}$$

называется *задачей с ограничениями, задаваемые равенствами и неравенствами* или *задачей с ограничениями типа равенств и неравенств.*

Сопоставим задаче (P_3) *функцию Лагранжа* $\mathcal{L}\colon X \times \mathbb{R} \times (\mathbb{R}^{m_1})^* \times (\mathbb{R}^m)^* \to \mathbb{R}$, определенную формулой

$$\mathcal{L}(x, \lambda_0, \lambda, y^*) = \sum_{i=0}^{m_1} \lambda_i f_i(x) + \langle y^*, F(x)\rangle,$$

где $\lambda = (\lambda_1, \ldots, \lambda_{m_1})$.

Теорема 2 (Правило множителей Лагранжа — необходимые условия минимума первого порядка в задаче (P_3))**.** *Пусть $\widehat{x}$ — локальный минимум в задаче (P_3). Если X — банахово пространство, функции f_i, $i = 0, 1, \ldots, m_1$, и отображение F строго дифференцируемы в $\widehat{x}$, то найдутся множители Лагранжа $\lambda_0 \in \mathbb{R}$, $\lambda = (\lambda_1, \ldots, \lambda_{m_1}) \in (\mathbb{R}^{m_1})^*$ и $y^* \in (\mathbb{R}^m)^*$, не равные одновременно нулю, такие, что*

(*a*) $\mathcal{L}_x(\widehat{x}, \lambda_0, \lambda, y^*) = 0 \ \Leftrightarrow\ \sum\limits_{i=0}^{m_1} \lambda_i f_i'(\widehat{x}) + (F'(\widehat{x}))^* y^* = 0$;

(*b*) $\lambda_i \ge 0, \quad i = 0, 1, \ldots, m_1$;

(*c*) $\lambda_i f_i(\widehat{x}) = 0, \quad i = 1, \ldots, m_1$.

Если $\operatorname{Im} F'(\widehat{x}) = \mathbb{R}^m$ и существует вектор $h \in \operatorname{Ker} F'(\widehat{x})$ такой, что $\langle f_i'(\widehat{x}), h\rangle < 0$ для тех индексов $1 \le i \le m_1$, для которых $f_i(\widehat{x}) = 0$, то $\lambda_0 \ne 0$.

Доказательство. Если определить отображение $G\colon X \to \mathbb{R}^{m_1} \times \mathbb{R}^m$ по формуле $G(x) = (\Phi(x), F(x))$, где $\Phi(x) = (f_1(x), \ldots, f_{m_1}(x))^T$ и положить $Q = \mathbb{R}^{m_1}_- \times \{0\}$, где $\mathbb{R}^{m_1}_- = \{x = (x_1, \ldots, x_{m_1})^T \in \mathbb{R}^{m_1} : x_i \le 0,\ i = 1, \ldots, m_1\}$, а 0 — нулевой элемент в $\mathbb{R}^m$, то задача (P_3) примет в точности вид задачи (P_2).

В новых обозначениях, предположения теоремы, очевидно, удовлетворяют предположениям теоремы о необходимых условиях минимума для задачи (P_2). Тогда согласно этой теореме найдутся множители Лагранжа $\lambda_0 \ge 0$ и $(\lambda, y^*) = (\lambda_1, \ldots, \lambda_{m_1}, y^*) \in (\mathbb{R}^{m_1})^* \times (\mathbb{R}^m)^*$ (вместо y^*), не равные одновременно нулю, такие, что для всех $x \in X$ справедливо равенство

$$\begin{aligned} 0 = \lambda_0 \langle f_0'(\widehat{x}), x\rangle + \langle (\lambda, y^*), (\Phi'(\widehat{x})x, F'(\widehat{x})x)\rangle = \lambda_0 \langle f_0'(\widehat{x}), x\rangle \\ + \sum_{i=1}^{m_1} \lambda_i \langle f_i'(\widehat{x}), x\rangle + \langle y^*, F'(\widehat{x})x\rangle = \left\langle \sum_{i=0}^{m_1} \lambda_i f_i'(\widehat{x}) + (F'(\widehat{x}))^* y^*, x \right\rangle, \end{aligned}$$

откуда следует (a).

Условие (2) в нашем случае означает, что

$$\langle(\lambda, y^*), (y, 0) - (\Phi(\widehat{x}), 0)\rangle = \sum_{i=1}^{m_1} \lambda_i (y_i - f_i(\widehat{x})) \le 0, \ \ \forall\, y = (y_1, \ldots, y_{m_1})^T \in \mathbb{R}^{m_1}_-. \quad (4)$$

Пусть $1 \le i_0 \le m_1$. Подставляя сюда $y_{i_0} = f_{i_0}(\widehat{x}) - 1$ и $y_i = f_i(\widehat{x})$, если $i \ne i_0$, получаем, что $\lambda_{i_0} \ge 0$. Этим доказано соотношение (b). Из него вытекает, что $\lambda_i f_i(\widehat{x}) \le 0$, $i = 1, \ldots, m_1$. Подставляя в (4) $y_1 = \ldots = y_{m_1} = 0$, получим, что $\sum_{i=1}^{m_1} \lambda_i f_i(\widehat{x}) \ge 0$. Следовательно, $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \ldots, m_1$, и утверждение (c) доказано.

Докажем последнее утверждение. Пусть $\lambda_0 = 0$. Тогда из (a) последует, что среди чисел $\lambda_1, \ldots, \lambda_{m_1}$ есть ненулевые, ибо в противном случае, получили бы, что $y^* = 0$ (оператор F сюръективен и поэтому его сопряженный инъективен).

Пусть h обладает указанными свойствами. Тогда из (a) и того, что $h \in \operatorname{Ker} F'(\widehat{x})$ вытекает равенство $0 = \sum_{i=1}^{m_1} \lambda_i \langle f_i'(\widehat{x}), h\rangle$. Если $\lambda_{i_0} > 0$, то $f_{i_0}(\widehat{x}) = 0$ согласно (c). Следовательно, $\langle f_{i_0}'(\widehat{x}), h\rangle < 0$, по свойству h, и мы приходим к противоречию: $0 = \sum_{i=1}^{m_1} \lambda_i \langle f_i'(\widehat{x}), h\rangle < 0$. □

§2.4. Выпуклые задачи

2.4.1. Постановка задачи

Пусть X — нормированное пространство. Функция $f\colon X \to \mathbb{R}$ называется *выпуклой*, если множество

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

называемое *надграфиком* f, выпукло.

Нетрудно проверить, что функция f выпукла тогда и только тогда, когда для любых $x_1, x_1 \in X$ и любого $0 \le \alpha \le 1$ выполняется неравенство

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

которое называется *неравенством Иенссена.*

Пусть $f_i\colon X \to \mathbb{R}$, $i = 0, 1, \ldots, m$ — выпуклые функции и Q — непустое выпуклое подмножество X. Задача

$$f_0(x) \to \min, \quad f_i(x) \le 0, \quad 1 \le i \le m, \quad x \in Q \qquad (P_4)$$

называется *выпуклой задачей* или *задачей выпуклого программирования.*

Заметим, что в этой задаче любой локальный минимум является и глобальным. Действительно, пусть $\widehat{x}$ — локальный минимум, т. е. существует такая окрестность U точки $\widehat{x}$, что $f_0(\widehat{x}) \le f_0(x)$ для всех допустимых $x \in U$. Пусть x — произвольная допустимая точка в задаче (P_4). Для достаточно малых $0 < \alpha \le 1$ точки $(1-\alpha)\widehat{x} + \alpha x = \widehat{x} + \alpha(x - \widehat{x})$ принадлежат U и допустимы, так как $(1-\alpha)\widehat{x} + \alpha x \in Q$, и по неравенству Иенссена $f_i(1-\alpha)\widehat{x} + \alpha x) \le (1-\alpha) f_i(\widehat{x}) + \alpha f_i(x) \le 0$, $i = 1, \ldots, m$. Снова по неравенству Иенссена, $f_0(\widehat{x}) \le f_0((1-\alpha)\widehat{x} + \alpha x) \le (1-\alpha) f_0(\widehat{x}) + \alpha f_0(x)$, т. е. $f_0(\widehat{x}) \le f_0(x)$.

Сопоставим задаче (P_4) *функцию Лагранжа* $\mathcal{L}\colon X\times\mathbb{R}\times(\mathbb{R}^m)^*\to\mathbb{R}$, определенную формулой

$$\mathcal{L}(x,\lambda_0,\lambda)=\sum_{i=0}^{m}\lambda_i f_i(x),$$

где $\lambda=(\lambda_1,\dots,\lambda_m)$.

2.4.2. Правило множителей Лагранжа

Теорема 1 (Каруша–Куна–Таккера)**.** 1) *Необходимость. Если $\widehat{x}$ — минимум в задаче (P_4), то найдутся множители Лагранжа λ_0 и $\lambda=(\lambda_1,\dots,\lambda_m)$, не равные одновременно нулю, такие, что*

(a) $\min\limits_{x\in Q}\mathcal{L}(x,\lambda_0,\lambda)=\mathcal{L}(\widehat{x},\lambda_0,\lambda)$;

(b) $\lambda_i\ge 0,\quad i=0,1,\dots,m$;

(c) $\lambda_i f_i(\widehat{x})=0,\quad i=1,\dots,m$.

Если существует такая точка $\overline{x}\in Q$, что $f_i(\overline{x})<0$, $1\le i\le m$, то $\lambda_0\ne 0$ (условие Слейтера).

2) *Достаточность. Если существует допустимая в (P_4) точка $\widehat{x}$ и множители Лагранжа $\lambda_0>0$ и $\lambda=(\lambda_0,\lambda_1,\dots,\lambda_m)$, удовлетворяющие условиям (a), (b) и (c), то $\widehat{x}$ — решение задачи (P_4).*

Доказательство. 1) Введем новую переменную $z=(x,y)\in Z=X\times\mathbb{R}^{m+1}$, где $y=(y_0,y_1,\dots,y_m)^T$, и положим $f(z)=f_0(x)+y_0$, $F(z)=(f_1(x)+y_1,\dots,f_m(x)+y_m)^T$, $K=Q\times\mathbb{R}^{m+1}_+$.

Рассмотрим задачу

$$f(z)\to\min,\quad F(z)=0,\quad z\in K.$$

Легко проверить, что если $\widehat{x}$ — глобальный минимум в задаче (P_4), то точка $\widehat{z}=(\widehat{x},\widehat{y})$, где $\widehat{y}=(0,-f_1(\widehat{x}),\dots,-f_m(\widehat{x}))^T$, является глобальным минимумом в этой задаче. Тогда, если $\Phi=(f,F)\colon X\times\mathbb{R}^{m+1}\to\mathbb{R}^{m+1}$, то $\Phi(\widehat{z})=(f(\widehat{z}),F(\widehat{z}))\in\partial\Phi(K)$ согласно предложению 1 §2.1.

Покажем, что множество $\Phi(K)$ выпукло. Пусть $w_i\in\Phi(K)$, $i=1,2$. Надо доказать, что для любого $\alpha\in(0,1)$ найдется элемент $z(\alpha)\in K$ такой, что $\Phi(z(\alpha))=(1-\alpha)w_1+\alpha w_2$.

Пусть $z_i=(x_i,y^i)\in K$ такие, что $w_i=\Phi(z_i)$, $i=1,2$. Положим $G(x)=(f_0(x),f_1(x),\dots,f_m(x))$ и пусть $z(\alpha)=(x(\alpha),u(\alpha))$, где $x(\alpha)=(1-\alpha)x_1+\alpha x_2$ и $y(\alpha)=(1-\alpha)y^1+\alpha y^2+(1-\alpha)G(x_1)+\alpha G(x_2)-G(x(\alpha))$.

Так как множество Q выпукло, то $x(\alpha)\in Q$. Далее, поскольку функции $f_0,f_1,\dots,f_m$ выпуклы, то из неравенства Иенссена следует, что $(1-\alpha)G(x_1)+\alpha G(x_2)-G(x(\alpha))\ge 0$ (неравенство покоординатное), т. е. $y(\alpha)\in\mathbb{R}^{m+1}_+$ и тем самым $z(\alpha)\in K$.

Теперь, учитывая, что $\Phi(z) = (f(z), F(z)) = (f_0(x) + y_0, f_1(x) + y_1, \ldots, f_m(x) + y_m) = G(x) + y$ получаем соотношение

$$(1-\alpha)w_1+\alpha w_2 = (1-\alpha)\Phi(x_1, y^1)+\alpha\Phi(x_2, y^2) = (1-\alpha)G(x_1)+\alpha G(x_2)+(1-\alpha)y^1+\alpha y^2$$
$$= G(x(\alpha))+(1-\alpha)y^1+\alpha y^2+(1-\alpha)G(x_1)+\alpha G(x_2)-G(x(\alpha)) = G(x(\alpha))+y(\alpha) = \Phi(z(\alpha))$$

и значит, $\Phi(K)$ — выпуклое множество.

Так как $\Phi(\widehat{z}) \in \partial\Phi(K)$, то согласно критерию граничной точки (предложение 2 §1.1) найдется ненулевой вектор, принадлежащий нормальному конусу $N_K(\Phi(\widehat{z}))$. Обозначим через $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$ противоположный ему вектор. Тогда $\langle \lambda, \Phi(z) - \Phi(\widehat{z})\rangle \geq 0$ для всех $z \in K$, или

$$\lambda_0(f_0(x) + y_0 - f_0(\widehat{x})) + \sum_{i=1}^{m} \lambda_i(f_i(x) + y_i) \geq 0 \tag{1}$$

для всех $(x, y) \in Q \times \mathbb{R}^{m+1}_{+}$, где $y = (y_0, y_1, \ldots y_m)$.

Полагая здесь $x = \widehat{x}$, $y_0 = 1$ и $y_i = -f_i(\widehat{x})$, $i = 1, \ldots, m$, получаем, что $\lambda_0 \geq 0$. Если $1 \leq j \leq m$, то полагая $x = \widehat{x}$, $y_0 = 0$, $y_i = -f_i(\widehat{x})$ при $i \neq j$ и $y_j = -f_j(\widehat{x}) + 1$, получаем, что $\lambda_j \geq 0$. Этим доказано утверждение (b) теоремы.

Так как (b) справедливо, то $\lambda_i f_i(\widehat{x}) \leq 0$, $i = 1, \ldots, m$. Подставляя в (1) $x = \widehat{x}$ и $y = 0$, получаем, что $\sum_{i=1}^{m} \lambda_i f_i(\widehat{x}) \geq 0$ и поэтому $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \ldots, m$. Утверждение (c) доказано.

Пусть в (1) $y = 0$. Тогда, используя (c), будем иметь $\sum_{i=0}^{m} \lambda_i f_i(x) \geq \lambda_0 f_0(\widehat{x}) = \sum_{i=0}^{m} \lambda_i f_i(\widehat{x})$ для любого $x \in Q$. Это и есть утверждение (a) теоремы.

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

2) Пусть x — допустимый элемент в задаче (P_4). Тогда, используя это обстоятельство вместе с (b), затем (a) и (c), будем иметь

$$\lambda_0 f_0(x) \geq \lambda_0 f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) = \mathcal{L}(x, \lambda_0, \lambda) \geq \mathcal{L}(\widehat{x}, \lambda_0, \lambda) = \lambda_0 f_0(x)$$
$$+ \sum_{i=1}^{m} \lambda_i f_i(\widehat{x}) = \lambda_0 f_0(\widehat{x}).$$

Деля на λ_0, получаем требуемое. $\square$

Глава 3. Условия экстремума в задачах вариационного исчисления и оптимального управления

§3.1. Задача Майера вариационного исчисления

3.1.1. Постановка задачи

Пусть $[t_0, t_1]$ — конечный отрезок, задано отображение $\psi\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$ (переменных $t \in \mathbb{R}$, $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, $u = (u_1, \ldots, u_r)^T \in \mathbb{R}^r$) и функции $g_i\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $i = 0, 1, \ldots, m$ переменных $\xi_i \in \mathbb{R}^n$ и $i = 0, 1$.

В пространстве $Z = C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$[1] относительно переменной $(x(\cdot), u(\cdot)) \in Z$ рассмотрим задачу

$$g_0(x(t_0), x(t_1)) \to \min, \quad \dot{x} = \psi(t, x, u), \quad g_i(x(t_0), x(t_1)) \leq 0, \quad 1 \leq i \leq m',$$
$$g_i(x(t_0), x(t_1)) = 0, \quad m' + 1 \leq i \leq m, \quad (P_5)$$

которая называется *задачей Майера вариационного исчисления* (в понтрягинской форме).

Пара $(x(\cdot), u(\cdot))$ называется *допустимой* в задаче (P_5), если $\dot{x}(t) = \psi(t, x(t), u(t))$ для всех $t \in [t_0, t_1]$ и выполнены ограничения с равенствами и неравенствами.

Допустимая пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ называется *слабым локальным минимумом* в задаче (P_5), если существует такая окрестность $\mathcal{O}$ в Z пары $(\widehat{x}(\cdot), \widehat{u}(\cdot))$, что для любой допустимой пары $(x(\cdot), u(\cdot)) \in \mathcal{O}$ выполнено неравенство $g_0(x(t_0), x(t_1)) \geq g_0(\widehat{x}(t_0), \widehat{x}(t_1))$.

Для сокращения записи пишем $\widehat{\psi}(t) = \psi(t, \widehat{x}(t), \widehat{u}(t))$ и аналогично для производных по x и u этой функции, $\widehat{g}_{i\xi_j} = g_{i\xi_j}(\widehat{x}(t_0), \widehat{x}(t_1))$, $i = 0, 1, \ldots, m$, $j = 0, 1$.

В этом параграфе мы будем предполагать, что отображение ψ непрерывно вместе со своими частными производными по x и u на открытом множестве в $G \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r$, содержащим точки $\{\, (t, \widehat{x}(t), \widehat{u}(t)) : t \in [t_0, t_1] \,\}$, а функции g_i, $0 \leq i \leq m$, непрерывно дифференцируемы в окрестности точки $(\widehat{x}(t_0), \widehat{x}(t_1))$.

3.1.2. Уравнения Эйлера-Лагранжа

Положим $L(t, x, \dot{x}, u, p) = \langle p, \dot{x} - \psi(t, x, u) \rangle$, $p \in (\mathbb{R}^n)^*$. Если фиксированы функции $\widehat{x}(\cdot)$, $\widehat{u}(\cdot)$ и функция $p(\cdot)$ на $[t_0, t_1]$ со значениями в $(\mathbb{R}^n)^*$, для сокращения записи пишем $\widehat{L}(t) = \langle p(t), \dot{\widehat{x}}(t) - \widehat{\psi}(t) \rangle$ и аналогично для производных по x, $\dot{x}$ и u функции L.

Теорема 1 (Необходимые условия слабого локального минимума первого порядка в задаче (P_5))**.** *Если пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ доставляет слабый локальный минимум в задаче (P_5), то найдутся ненулевой вектор $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in (\mathbb{R}^{m+1})^*$ и функция $p(\cdot) \in C^1([t_0, t_1], (\mathbb{R}^n)^*)$ такие, что $\widehat{L}_{\dot{x}}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ и выполняются*

[1] Z — нормированное пространство с нормой $\|(x(\cdot), u(\cdot))\|_Z = \|x(\cdot)\|_{C^1([t_0,t_1],\mathbb{R}^n)} + \|u(\cdot)\|_{C([t_0,t_1],\mathbb{R}^r)}$.

(a) *условие стационарности по x:*

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

(b) *условия трансверсальности*

$$\widehat{L}_{\dot{x}}(t_j)=(-1)^j\sum_{i=0}^{m}\lambda_i\widehat{g}_{i\xi_j},\;\; j=0,1 \;\;\Leftrightarrow\;\; p(t_j)=(-1)^j\sum_{i=0}^{m}\lambda_1\widehat{g}_{i\xi_j},\;\; j=0,1,$$

(c) *условие стационарности по u*

$$\widehat{L}_u(t)=0 \;\;\Leftrightarrow\;\; p(t)\widehat{\psi}_u(t)=0.$$

(d) *условия дополняющей нежесткости*

$$\lambda_i g_i(\widehat{x}(t_0),\widehat{x}(t_1))=0, \quad i=1,\dots,m',$$

(e) *условия неотрицательности*

$$\lambda_i\geq 0, \quad i=0,1,\dots m'.$$

Чтобы не утяжелять изложение, рассмотрим задачу в постановке, которая не является столь общей как задача (P_5), но весьма содержательной. Доказательство общего результата потребует лишь некоторых технических усложнений приводимого ниже доказательства.

Итак, пусть $\psi\colon \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r\to\mathbb{R}^n$ и $g_i\colon\mathbb{R}^n\to\mathbb{R}$, $i=0,1,\dots,m$. Рассмотрим следующую задачу

$$g_0(x(t_1))\to\min, \quad \dot{x}=\psi(t,x,u), \quad x(t_0)=x_0, \quad g_i(x(t_1))=0, \quad i=1,\dots,m, \qquad (P_5')$$

на пространстве пар $(x(\cdot),u(\cdot))\in C^1([t_0,t_1],\mathbb{R}^n)\times C([t_0,t_1],\mathbb{R}^r)$.

Теорема 2 (Необходимые условия слабого локального минимума первого порядка в задаче (P_5'))**.** *Если пара $(\widehat{x}(\cdot),\widehat{u}(\cdot))$ доставляет слабый локальный минимум в задаче (P_5'), то найдутся ненулевой вектор $\lambda=(\lambda_0,\lambda_1,\dots,\lambda_m)\in(\mathbb{R}^{m+1})^*$ и функция $p(\cdot)\in C^1([t_0,t_1],(\mathbb{R}^n)^*)$ такие, что $\widehat{L}_{\dot{x}}(\cdot)\in C^1([t_0,t_1],\mathbb{R}^n)$ и выполняются*

(a) *условие стационарности по x:*

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

(b) *условия трансверсальности:*

$$\widehat{L}_{\dot{x}}(t_1)=-\sum_{i=0}^{m}\lambda_i g_i'(\widehat{x}(t_1)) \;\;\Leftrightarrow\;\; p(t_1)=-\sum_{i=0}^{m}\lambda_i g_i'(\widehat{x}(t_1)),$$

(*c*) *условие стационарности по* u:

$$\widehat{L}_u(t) = 0 \ \Leftrightarrow \ p(t)\widehat{\psi}_u(t) = 0.$$

Заметим, что в соответствии с общим результатом, следовало бы еще сказать, что найдется такой вектор $\mu = (\mu_1, \dots, \mu_n) \in (\mathbb{R}^n)^*$, что $\widehat{L}_{\dot{x}}(t_0) = p(t_0) = \mu$. Но это условие мы не пишем, так как оно не несет никакой информации, поскольку μ ни в какие другие соотношения не входит, и если $\lambda = 0$, то $p = 0$, а тем самым $\mu = 0$.

Доказательству теоремы предпошлем одну лемму о дифференцируемой зависимости решения дифференциального уравнения от параметра, доказывать которую здесь не будем, но отметим, что она есть непосредственное следствие классической теоремы о неявной функции (см. [2], п. 2.3.4).

Ниже для сокращения записи будем часто писать $\widehat{x}$ вместо $\widehat{x}(\cdot)$, $\widehat{u}$ вместо $\widehat{u}(\cdot)$ и т. д.

Лемма 1. *Пусть* $(\widehat{x}, \widehat{u}) \in C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$, *отображение* $\psi\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$ *удовлетворяет оговоренным выше условиям гладкости и функция* $\widehat{x}$ *является решением задачи Коши*

$$\dot{x} = \psi(t, x, \widehat{u}), \quad x(t_0) = x_0 \tag{1}$$

на отрезке $[t_0, t_1]$.

Тогда найдется такая окрестность $V \subset C([t_0, t_1], \mathbb{R}^r)$ *функции* $\widehat{u}$, *что для любого* $u \in V$ *существует единственное решение* $x(\cdot, x_0, u) = x(\cdot, u)$ *задачи Коши*

$$\dot{x} = \psi(t, x, u), \quad x(t_0) = x_0, \tag{2}$$

определенное на всем отрезке $[t_0, t_1]$. *При этом отображение* $u \mapsto x(\cdot, u)$, *как отображение из* V *в* $C^1([t_0, t_1], \mathbb{R}^n)$, *непрерывно дифференцируемо на* V *и для его производной по* u *в точке* $\widehat{u}$, *которую обозначим* $\widehat{x}'$, *и для любого* $v \in C([t_0, t_1], \mathbb{R}^r)$ *функция* $h = \widehat{x}'v$ *удовлетворяет на* $[t_0, t_1]$ *дифференциальному уравнению*

$$\dot{h} = \widehat{\psi}_x(t)h + \widehat{\psi}_u(t)v, \quad h(t_0) = 0. \tag{3}$$

Доказательство теоремы 2. Для пары $(\widehat{x}, \widehat{u})$ и отображения ψ выполнены условия леммы 1. Пусть V — окрестность функции $\widehat{u}$ из этой леммы. Рассмотрим следующую экстремальную задачу относительно переменной u:

$$g_0(x(t_1, u)) \to \min, \quad g_i(x(t_1, u)) = 0, \quad i = 1, \dots, m, \quad u \in V. \tag{4}$$

Точка $\widehat{u}$ доставляет локальный минимум в этой задаче. Действительно, если пара $(\widehat{x}, \widehat{u})$ является локальным минимумом в задаче (P_5'), то, по определению, существует такая окрестность $\mathcal{O} \subset C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$ этой пары, что $g_0(x(t_1)) \geq g_0(\widehat{x}(t_1))$ для всех допустимых пар $(x, u) \in \mathcal{O}$.

Так как отображение $u \mapsto x(\cdot, u)$, очевидно, непрерывно в точке $\widehat{u}$ и по условию $x(\cdot, \widehat{u}) = \widehat{x}$, то найдется окрестность $V_0 \subset V$ точки $\widehat{u}$ такая, что если $u \in V_0$, то $(x(\cdot, u), u) \in \mathcal{O}$.

Пусть точка $u \in V_0$ допустима в задаче (4). Тогда пара $(x(\cdot,u),u)$ допустима в задаче (P_5') и принадлежит $\mathcal{O}$. Следовательно, $g_0(x(t_1,u)) \geq g_0(\widehat{x}(t_1)) = g_0(x(t_1,\widehat{u}))$ и значит, $\widehat{u}$ — локальный минимум в задаче (4).

Если обозначить $f_i(u) = g_i(x(t_1,u))$ и $F = (f_1, f_2, ..., f_m)^T$ (считая, что f_0 и F продолжены произвольным образом за пределы V, то приходим к задаче

$$f_0(u) \to \min, \quad F(u) = 0 \tag{5}$$

на $C([t_0,t_1],\mathbb{R}^r)$, которая имеет вид задачи (P_1) и для которой $\widehat{u}$ — локальный минимум.

Функции g_i, $i = 0,1,\ldots,m$, непрерывно дифференцируемы как суперпозиции непрерывно дифференцируемых отображений: $u \mapsto x(\cdot,u) \mapsto x(t_1,u) \mapsto g_i(x(t_1,u))$ и поэтому строго дифференцируемы в точке $\widehat{u}(\cdot)$ (см. следствие 1 §1.2). Условия следствия 1 §3.1 выполнены и следовательно, найдутся множители Лагранжа $\lambda_0 \geq 0$ и $\lambda = (\lambda_1,\ldots,\lambda_m) \in (\mathbb{R}^m)^*$, не равные одновременно нулю, такие, что

$$\langle \sum_{i=0}^m \lambda_i f_i'(\widehat{u}), v\rangle = 0$$

для всех $v \in C([t_0,t_1],\mathbb{R}^r)$, или равносильно (согласно теореме о суперпозиции дифференцируемых отображений)

$$\langle \sum_{i=0}^m \lambda_i g_i'(\widehat{x}(t_1)), (\widehat{x}'v)(t_1)\rangle = \langle \sum_{i=0}^m \lambda_i g_i'(\widehat{x}(t_1)), h(t_1)\rangle = 0, \tag{6}$$

где $h = \widehat{x}'v$.

Пусть функция $p \in C^1([t_0,t_1],(\mathbb{R}^n)^*)$ является решением задачи Коши

$$\dot{p} = -p\,\widehat{\psi}_x(t), \quad p(t_1) = -\sum_{i=0}^m \lambda_i g_i'(\widehat{x}(t_1)). \tag{7}$$

Такое решение существует и единственно по теореме о существования решения линейного дифференциального уравнения. Отсюда, в частности, следует, что $\widehat{L}_{\dot{x}}(\cdot) \in C^1([t_0,t_1],\mathbb{R}^n)$.

Пусть $v \in C([t_0,t_1],\mathbb{R}^r)$. Используя равенство (6), начальные условия в (3) и (7), а затем подставляя вместо $\dot{h}(\cdot)$ и $\dot{p}(\cdot)$ их выражения в (3) и (7), будем иметь

$$0 = \langle \sum_{i=0}^m \lambda_i g_i'(\widehat{x}(t_1)), h(t_1)\rangle = -\langle p(t_1), h(t_1)\rangle = -\int_{t_0}^{t_1} (\langle \dot{p}(t), h(t)\rangle + \langle p(t), \dot{h}(t)\rangle)\,dt$$
$$-\langle p(t_0), h(t_0)\rangle = -\int_{t_0}^{t_1} \langle p(t)\widehat{\psi}_u(t), v(t)\rangle\,dt$$

для всех $v \in C([t_0,t_1],\mathbb{R}^r)$. Беря $v = (p\,\widehat{\psi}_u)^T$, получаем, что $\int_{t_0}^{t_1} |p(t)\widehat{\psi}_u(t)|^2\,dt = 0$ и значит, $p(t)\widehat{\psi}_u(t) = 0$ для всех $t \in [t_0,t_1]$. Вместе с (7) это доказывает теорему. □

§3.2. Задача Лагранжа вариационного исчисления

3.2.1. Постановка задачи

Пусть $[t_0, t_1]$ — конечный отрезок, заданы функции $f_i\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$, $i = 0, 1, \dots, m$, и отображение $\varphi\colon \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r \to \mathbb{R}^n$ (переменных $t \in \mathbb{R}$, $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, $u = (u_1, \dots, u_r)^T \in \mathbb{R}^r$) и функции $l_i\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $i = 0, 1, \dots, m$ переменных ξ_0 и ξ_1. Обозначим

$$\mathcal{B}_i(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} f_i(t, x(t), u(t))\,dt + l_i(x(t_0), x(t_1)), \quad i = 0, 1, \dots, m.$$

В пространстве $Z = C^1([t_0, t_1], \mathbb{R}^n)\times C([t_0, t_1], \mathbb{R}^r)$ относительно переменной $(x(\cdot), u(\cdot)) \in Z$ рассмотрим задачу

$$\mathcal{B}_0(x(\cdot), u(\cdot)) \to \min, \quad \dot x = \varphi(t, x, u),$$
$$\mathcal{B}_i(x(\cdot), u(\cdot)) \le 0, \quad 1 \le i \le m',$$
$$\mathcal{B}_i(x(\cdot), u(\cdot)) = 0, \quad m' + 1 \le i \le m, \quad (P_6)$$

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

Допустимая пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ называется *слабым локальным минимумом в задаче* (P_6), если существует такая окрестность $\mathcal{O}$ в Z пары $(\widehat{x}(\cdot), \widehat{u}(\cdot))$, что для любой допустимой пары $(x(\cdot), u(\cdot)) \in \mathcal{O}$ выполнено неравенство $\mathcal{B}_0(x(\cdot), u(\cdot)) \ge \mathcal{B}_0(\widehat{x}(\cdot), \widehat{u}(\cdot))$.

Соглашения относительно сокращенных обозначений аналогичны тем, которые были в предыдущем параграфе.

Также как и в предыдущем параграфе мы предполагаем, что функции f_i, $0 \le i \le m$, и отображение φ непрерывны вместе со своими частными производными по x и u на некотором открытом множестве $G \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r$, содержащим точки $\{\, (t, \widehat{x}(t), \widehat{u}(t)) : t \in [t_0, t_1] \,\}$, а функции l_i, $0 \le i \le m$, непрерывно дифференцируемы в окрестности точки $(\widehat{x}(t_0), \widehat{x}(t_1))$.

3.2.2. Уравнения Эйлера-Лагранжа

Положим $L(t, x, \dot x, u, \lambda, p) = \sum_{i=0}^m \lambda_i f_i(t, x, u) + \langle p, \dot x - \varphi(t, x, u)\rangle$, где $p \in (\mathbb{R}^n)^*$ и $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m) \in (\mathbb{R}^{m+1})^*$.

Если фиксированы пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$, вектор λ и функция $p = p(\cdot)$ на $[t_0, t_1]$ со значениями в $(\mathbb{R}^n)^*$, то пишем $\widehat{L}(t) = \sum_{i=0}^m \lambda_i \widehat{f}_i(t) + \langle p(t), \dot{\widehat{x}}(t) - \widehat{\varphi}(t)\rangle$ и аналогично для производных по x, $\dot x$ и u функции L, $\widehat{l}_{i\xi_j} = l_{i\xi_j}(\widehat{x}(t_0), \widehat{x}(t_1))$, $i = 0, 1, \dots, m$, $j = 0, 1$..

Теорема 1 (Уравнения Эйлера–Лагранжа — необходимые условия слабого минимума первого порядка в задаче (P_6))**.** *Если $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ доставляет слабый локальный минимум в задаче (P_6), то найдутся ненулевой вектор $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m) \in (\mathbb{R}^{m+1})^*$ и функция $p(\cdot) \in C^1([t_0, t_1], (\mathbb{R}^n)^*)$ такие, что $\widehat{L}_{\dot x}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ и выполняются*

(a) *условие стационарности по x:*

$$-\frac{d}{dt}\widehat{L}_{\dot x}(t) + \widehat{L}_x(t) = 0 \;\Leftrightarrow\; \dot p(t) = -p(t)\widehat{\varphi}_x(t) + \sum_{i=0}^m \lambda_i \widehat{f}_{ix}(t),$$

(b) *условия трансверсальности:*

$$\widehat{L}_{\dot{x}}(t_j) = (-1)^j \sum_{i=0}^{m} \lambda_i \widehat{l}_{i\xi_j}, \quad j = 0, 1 \quad \Leftrightarrow \quad p(t_j) = (-1)^j \sum_{i=0}^{m} \lambda_i \widehat{l}_{i\xi_j}, \quad j = 0, 1,$$

(c) *условие стационарности по u:*

$$\widehat{L}_u(t) = 0 \quad \Leftrightarrow \quad p(t)\widehat{\varphi}_u(t) - \sum_{i=0}^{m} \lambda_i \widehat{f}_{iu}(t) = 0,$$

(d) *условия дополняющей нежесткости:*

$$\lambda_i \mathcal{B}_i(\widehat{x}(\cdot), \widehat{u}(\cdot)) = 0, \quad i = 1, \dots, m',$$

(e) *условия неотрицательности:*

$$\lambda_i \geq 0, \quad i = 0, 1, \dots m'.$$

Уравнениями Эйлера–Лагранжа обычно называют условия стационарности по x и u.

Покажем, что необходимые условия минимума в задаче (P_6) являются следствием необходимых условий минимума в задаче (P_5). Действительно, сопоставим задаче (P_6) следующую постановку

$$\begin{aligned} &y_0(t_1) - y_0(t_0) + l_0(x(t_0), x(t_1)) \to \min, \\ &\dot{x} = \varphi(t, x, u), \quad \dot{y} = f(t, x, u)) = (f_0(t, x, u), \dots, f_m(t, x, u)), \\ &y_i(t_1) - y_i(t_0) + l_i(x(t_0), x(t_1)) \leq 0, \quad 1 \leq i \leq m', \\ &y_i(t_1) - y_i(t_0) + l_i(x(t_0), x(t_1)) = 0, \quad m' + 1 \leq i \leq m. \end{aligned} \qquad (P_6')$$

Пусть $z(\cdot) = (x(\cdot), y(\cdot)) \in C^1([t_0, t_1], \mathbb{R}^{n+m+1})$. Рассмотрим функции $g_i \colon \mathbb{R}^{n+m+1} \times \mathbb{R}^{n+m+1} \to \mathbb{R}$, $i = 0, 1, \dots, m$, переменных $\zeta_j = (\xi_j, \eta_j) \in \mathbb{R}^n \times \mathbb{R}^{m+1}$, $j = 0, 1$, и отображение $\psi \colon \mathbb{R} \times \mathbb{R}^{n+m+1} \times \mathbb{R}^r \to \mathbb{R}^{n+m+1}$ переменных t, z и u, определенные формулами: $g_i(\zeta_0, \zeta_1) = \eta_{1i} - \eta_{0i} + l_i(\xi_0, \xi_1)$, где $\eta_j = (\eta_{j0}, \eta_{j1}, \dots, \eta_{jm})^T$, $i = 0, 1, \dots, m$, $j = 0, 1$, и $\psi(t, z, u) = (\varphi(t, x, u), f(t, x, u))$. Тогда задача (P_6') приобретает вид задачи (P_5) относительно переменной $(z(\cdot), u(\cdot)) \in C^1([t_0, t_1], \mathbb{R}^{n+m+1}) \times C([t_0, t_1], \mathbb{R}^r)$.

Элементарная проверка показывает, что если $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ — локальный минимум в задаче (P_6), то пара $(\widehat{z}(\cdot), \widehat{u}(\cdot))$, где $\widehat{z}(\cdot) = (\widehat{x}(\cdot), \widehat{y}(\cdot))$ и $\dot{\widehat{y}}(t) = f(t, \widehat{x}(t), \widehat{u}(t))$, $t \in [t_0, t_1]$, является локальным минимумом в задаче (P_6').

Согласно теореме 1 §4.1, если пара $(\widehat{z}(\cdot), \widehat{u}(\cdot))$ доставляет слабый локальный минимум в задаче (P_6'), то найдутся ненулевой вектор $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m) \in (\mathbb{R}^{m+1})^*$ и функция $\overline{p}(\cdot) \in C^1([t_0, t_1], (\mathbb{R}^{n+m+1})^*)$ такие, что выполняются, в частности, условия стационарности по z:

$$\dot{\overline{p}}(t) = -\overline{p}(t)\widehat{\psi}_z(t)$$

и условия трансверсальности:

$$\overline{p}(t_j) = (-1)^j \sum_{i=0}^{m} \lambda_i \widehat{g}_{i\zeta_j}, \quad j = 0, 1.$$

Функцию $p(\cdot)$ можно представить в виде $\overline{p}(\cdot) = (p(\cdot), q(\cdot)) \in (\mathbb{R}^n)^* \times (\mathbb{R}^{m+1})^*$. Тогда условие стационарности по z равносильным образом записывается так:

$$\dot{p}(t) = -p(t)\widehat{\varphi}_x(t) - q(t)\widehat{f}_x(t) = -p(t)\widehat{\varphi}_x(t) - \sum_{i=0}^{m} q_i(t)\widehat{f}_{ix}(t),$$

где $q(t) = (q_0(t), q_1(t), \ldots, q_m(t))$ и $\dot{q}(t) = 0$, а условия трансверсальности так

$$p(t_j) = (-1)^j \sum_{i=0}^{m} \lambda_i \widehat{l}_{i\xi_j}, \;\; j = 0, 1, \qquad q(t_j) = -\lambda, \;\; j = 0, 1.$$

Таким образом, $q(t) \equiv -\lambda$ и мы сразу получаем условия стационарности по x и условия трансверсальности в теореме 1.

Условие стационарности по u в теореме 1 следует из соотношения $\overline{p}(t)\widehat{\psi}_u(t) = 0$. Условия дополняющей нежесткости и неотрицательности выполняются очевидным образом.

Необходимые условия минимума в задаче (P_6) доказаны.

§3.3. Классические задачи вариационного исчисления

В этом параграфе, в качестве следствия доказанных выше результатов, будут получены необходимые условия минимума в трех задачах классического вариационного исчисления: в задаче Больца, в простейшей задаче и в изопериметрической задаче. Требования относительно гладкости отображений, входящих в формулировки задач не оговариваем, так как они те же, что и в общей задаче Лагранжа.

3.3.1. Задача Больца

Пусть $[t_0, t_1]$ — отрезок числовой прямой, задана функция $L\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ переменных t, x и $\dot{x}$ и функция $l\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ переменных ξ_0 и ξ_1. Задача

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

относительно переменной $x(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ называется *задачей Больца.*

Стандартным образом определяется (слабый) локальный минимум в этой задаче.

Теорема 1 (Необходимые условия слабого минимума в задаче Больца)**.** *Если $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (B), то $\widehat{L}_{\dot{x}}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$, для всех $t \in [t_0, t_1]$ выполнено уравнение Эйлера:*

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

и условия трансверсальности:

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

Доказательство. Перепишем задачу (B) в форме задачи Лагранжа (P_6)

$$\mathcal{B}(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), u(t))\, dt + l(x(t_0), x(t_1)) \to \min, \quad \dot{x}(t) = u(t). \tag{1}$$

Легко видеть, что если $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (B), то пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$, где $\dot{\widehat{x}}(t) = \widehat{u}(t)$, доставляет слабый локальный минимум в задаче (1).

Согласно теореме 1 §4.2 найдутся положительное число λ_0 и функция $p(\cdot) \in C^1([t_0, t_1], (\mathbb{R}^n)^*)$ такие, что выполняются условие стационарности по x:

$$\dot{p}(t) = \lambda_0 \widehat{L}_x(t),$$

условия трансверсальности:

$$p(t_j) = (-1)^j \lambda_0 \widehat{l}_{\xi_j}, \quad j = 0, 1,$$

и условие стационарности по u:

$$p(t) = \lambda_0 \widehat{L}_u(t) = \lambda_0 \widehat{L}_{\dot{x}}(t).$$

Дифференцируя уравнение в условии стационарности по u и учитывая условие стационарности по x, получаем, деля на λ_0, уравнение Эйлера.

Условия трансверсальности следуют из выписанных условий и того, что $p(t_j) = \lambda_0 \widehat{L}_{\dot{x}}(t_j)$, $j = 0, 1$. □

3.3.2. Простейшая задача вариационного исчисления

Пусть $[t_0, t_1]$ — отрезок числовой прямой, задана функция $L\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ переменных t, x и $\dot{x}$ и векторы $x_i \in \mathbb{R}^n$, $i = 0, 1$. Задача

$$J(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt \to \min, \quad x(t_i) = x_i, \quad i = 0, 1, \tag{Pr}$$

называется *простейшей задачей* (*классического*) *вариационного исчисления.*

Эту задачу, как и задачу Больца, рассматриваем в пространстве $C^1([t_0, t_1], \mathbb{R}^n)$.

Теорема 2 (Необходимые условия минимума в задаче (Pr))**.** *Если $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (Pr), то $\widehat{L}_{\dot{x}}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ и для всех $t \in [t_0, t_1]$ выполнено уравнение Эйлера*

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

Доказательство. Перепишем задачу (Pr) в форме задачи Лагранжа (P_6)

$$J(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), u(t))\, dt \to \min, \quad \dot{x}(t) = u(t),$$

$$x(t_0) - x_0 = 0, \quad x(t_1) - x_1 = 0. \tag{2}$$

Легко видеть, что если $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (Pr), то пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$, где $\dot{\widehat{x}}(t) = \widehat{u}(t)$, доставляет слабый локальный минимум в задаче (2).

Согласно теореме 1 §4.2 найдутся, не равные одновременно нулю, число λ_0 и векторы $\lambda \in (\mathbb{R}^n)^*$ и $\mu \in (\mathbb{R}^n)^*$, а также вектор-функция $p(\cdot) \in C^1([t_0,t_1],(\mathbb{R}^n)^*)$ такие, что выполняются условие стационарности по x:

$$\dot p(t) = \lambda_0 \widehat{L}_x(t),$$

условия трансверсальности:

$$p(t_0) = \lambda, \qquad p(t_1) = -\mu,$$

условие стационарности по u:

$$p(t) = \lambda_0 \widehat{L}_u(t) = \lambda_0 \widehat{L}_{\dot x}(t)$$

и условие неотрицательности: $\lambda_0 \geq 0$.

Заметим, что $\lambda_0 \neq 0$, так как в противном случае из условий трансверсальности и стационарности по u последовало бы, что $\lambda = \mu = 0$. Теперь дифференцируя уравнение в условии стационарности по u и используя условие стационарности по x, получаем, деля на λ_0, уравнение Эйлера. Условия трансверсальности в качестве необходимых условий не пишем, так как они не несут никакой информации. $\square$

3.3.3. Изопериметрическая задача

Пусть $[t_0,t_1]$ — отрезок числовой прямой, заданы функция $f_i \colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $i = 0,1,\ldots,m$, переменных t, x и $\dot x$, векторы $x_i \in \mathbb{R}^n$, $i = 0,1$, и числа $\alpha_i \in \mathbb{R}$, $i = 1,\ldots,m$. Задача

$$J(x(\cdot)) = \int_{t_0}^{t_1} f_0(t,x(t),\dot x(t))\,dt \to \min, \quad \int_{t_0}^{t_1} f_i(t,x(t),\dot x(t))\,dt = \alpha_i,$$
$$i = 1,\ldots,m, \quad x(t_0) = x_0, \quad x(t_1) = x_1 \quad (Iz)$$

называется *изопериметрической задачей.*

Эту задачу также рассматриваем в пространстве $C^1([t_0,t_1],\mathbb{R}^n)$.

Положим $L(t,x,\dot x,\lambda) = \sum_{i=0}^m \lambda_i f_i(t,x,\dot x)$, где $\lambda = (\lambda_0,\lambda_1,\ldots,\lambda_m)$.

Теорема 3 (Необходимые условия минимума в задаче (Iz))**.** *Если $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (Iz), то найдется такой ненулевой набор множителей Лагранжа $\lambda = (\lambda_0,\lambda_1,\ldots,\lambda_m)$, что $\widehat{L}_{\dot x}(\cdot) \in C^1([t_0,t_1],\mathbb{R}^n)$ и для всех $t \in [t_0,t_1]$ выполнено уравнение Эйлера*

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

Доказательство. Перепишем задачу (Iz) в форме задачи (P_6)

$$J(x(\cdot),u(\cdot)) = \int_{t_0}^{t_1} f_0(t,x(t),u(t))\,dt \to \min, \quad \dot x = u,$$
$$\int_{t_0}^{t_1} f_i(t,x(t),u(t))\,dt - \alpha_i = 0, \quad i = 1,\ldots,m,$$
$$x(t_0) - x_0 = 0, \quad x(t_1) - x_1 = 0. \quad (3)$$

Ясно, что если $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (Iz), то пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$, где $\dot{\widehat{x}}(t) = \widehat{u}(t)$, доставляет слабый локальный минимум в задаче (3).

Согласно теореме 1 §4.2 найдутся, не равные одновременно нулю, число $\lambda_0 \geq 0$ и векторы $\lambda = (\lambda_1, \ldots, \lambda_m) \in (\mathbb{R}^m)^*$, $\mu \in (\mathbb{R}^n)^*$ и $\nu \in (\mathbb{R}^n)^*$, а также вектор-функция $p(\cdot) \in C^1([t_0, t_1], (\mathbb{R}^n)^*)$ такие, что выполняются условие стационарности по x:

$$\dot{p}(t) = -\sum_{i=0}^{m} \lambda_i \widehat{f}_{ix}(t),$$

условия трансверсальности:

$$p(t_0) = \mu, \qquad p(t_1) = -\nu,$$

и условие стационарности по u:

$$p(t) = \sum_{i=0}^{m} \lambda_i \widehat{f}_{iu}(t).$$

Дифференцируя уравнение в условии стационарности по u и используя условие стационарности по x, получаем (учитывая, что $\widehat{f}_{iu}(t) = \widehat{f}_{i\dot{x}}(t)$), уравнение Эйлера. Условия трансверсальности в качестве необходимых условий не пишем, так как они неинформативны. $\square$

§3.4. Задача оптимального управления

3.4.1. Постановка задачи

Пусть $[t_0, t_1]$ — конечный отрезок, U — непустое подмножество $\mathbb{R}^r$, заданы функции $f_i\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$, $i = 0, 1, \ldots, m$, и отображение $\varphi\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$ переменных $t \in \mathbb{R}$, $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, $u = (u_1, \ldots, u_r)^T \in \mathbb{R}^r$) и функции $l_i\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $i = 0, 1, \ldots, m$ переменных ξ_0 и ξ_1.

Обозначим через $PC([t_0, t_1], \mathbb{R}^r)$ совокупность кусочно-непрерывных функций, а через $PC^1([t_0, t_1], \mathbb{R}^n)$ совокупность всех кусочно-непрерывно дифференцируемых функций (т. е. непрерывных функций, у которых производная кусочно непрерывна) на $[t_0, t_1]$ со значениями соответственно в $\mathbb{R}^r$ и $\mathbb{R}^n$.

Пусть $\mathcal{B}_i(x(\cdot), u(\cdot))$, $i = 0, 1, \ldots, m$, обозначают тоже, что и в п. 4.2.1. В пространстве $Z_1 = PC^1([t_0, t_1], \mathbb{R}^n) \times PC([t_0, t_1], \mathbb{R}^r)$ относительно переменной $(x(\cdot), u(\cdot)) \in Z_1$ рассмотрим задачу

$$\begin{aligned} &\mathcal{B}_0(x(\cdot), u(\cdot)) \to \min, \quad \dot{x} = \varphi(t, x, u), \quad u(t) \in U, \\ &\qquad\qquad \mathcal{B}_i(x(\cdot), u(\cdot)) \leq 0, \quad 1 \leq i \leq m', \\ &\qquad\qquad\qquad \mathcal{B}_i(x(\cdot), u(\cdot)) = 0, \quad m' + 1 \leq i \leq m, \quad (P_7) \end{aligned}$$

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

Задача (P_7) называется *задачей оптимального управления*. Переменную $x(\cdot)$ часто называют фазовой переменной, $u(\cdot)$ — управлением, пару $(\widehat{x}(\cdot), u(\cdot))$ — процессом.

Допустимый процесс $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ называется *оптимальным процессом* или *сильным минимумом* в задаче (P_7), если существует такая окрестность $\mathcal{O}$ функции $\widehat{x}(\cdot)$ в $C([t_0, t_1], \mathbb{R}^n)$, что для любого допустимого процесса $(x(\cdot), u(\cdot))$, для которого $x(\cdot) \in \mathcal{O}$, выполнено неравенство $\mathcal{B}_0(x(\cdot), u(\cdot)) \geq \mathcal{B}_0(\widehat{x}(\cdot), \widehat{u}(\cdot))$.

3.4.2. Принцип максимума Понтрягина

Положим $H(t,x,u,\lambda,p) = \langle p, \varphi(t,x,u)\rangle - \sum_{i=0}^{m} \lambda_i f_i(t,x,u)$ (*функция Понтрягина* задачи (P_7)), где $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in (\mathbb{R}^{m+1})^*$ и $p \in (\mathbb{R}^n)^*$.

Если фиксирована пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$, то для функций и отображений на этой паре пользуемся теми же сокращенными обозначениями, что и в задаче Лагранжа. Если еще фиксирована функция $p = p(\cdot)$ на $[t_0, t_1]$ со значениями в $(\mathbb{R}^n)^*$, то пишем $\widehat{H}(t) = \langle p(t), \widehat{\varphi}(t)\rangle - \sum_{i=0}^{m} \lambda_i \widehat{f}_i(t)$ и аналогично для производной по x этой функции.

В этом параграфе мы предполагаем, что существует такое открытое множество $G \subset \mathbb{R} \times \mathbb{R}^n$, содержащее точки $\{\,(t, \widehat{x}(t)) : t \in [t_0, t_1]\,\}$, что функции f_i, $0 \le i \le m$, и отображение φ непрерывны вместе со своей частной производной по x на $V \times U$, а функции l_i, $0 \le i \le m$, непрерывно дифференцируемы в окрестности точки $(\widehat{x}(t_0), \widehat{x}(t_1))$.

Теорема 1 (Принцип максимума Понтрягина — необходимые условия сильного минимума первого порядка в задаче (P_7))**.** *Если $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ — оптимальный процесс в задаче (P_7), то найдутся ненулевой вектор $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in (\mathbb{R}^{m+1})^*$ и функция $p(\cdot) \in PC^1([t_0, t_1], (\mathbb{R}^n)^*)$ такие, что выполняются*

(a) *условие стационарности по x:*

$$\dot{p}(t) = -\widehat{H}_x(t) \;\Leftrightarrow\; \dot{p}(t) = -p(t)\widehat{\varphi}_x(t) + \sum_{i=0}^{m} \lambda_i \widehat{f}_{ix}(t),$$

(b) *условия трансверсальности:*

$$p(t_j) = (-1)^j \sum_{i=0}^{m} \lambda_i \widehat{l}_{i\xi_j}, \quad j = 0, 1,$$

(c) *условие максимума по u для всех $t \in [t_0, t_1]$, где функция $\widehat{u}(\cdot)$ непрерывна:*

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

(d) *условия дополняющей нежесткости:*

$$\lambda_i \mathcal{B}_i(\widehat{x}(\cdot), \widehat{u}(\cdot)) = 0, \quad i = 1, \ldots, m',$$

(e) *условия неотрицательности:*

$$\lambda_i \ge 0, \quad i = 0, 1, \ldots m'.$$

Из-за специфического условия (c) необходимые условия в задаче оптимального управления обычно называют *"Принципом максимума Понтрягина"*.

Доказательство принципа максимума проведем для задачи, аналогичной по структуре задаче (P_5').

Итак, пусть $\varphi\colon \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r\to\mathbb{R}^n$, $g_i\colon\mathbb{R}^n\to\mathbb{R}$, $i=0,1,\ldots,m$, и U — непустое подмножество $\mathbb{R}^r$. Рассмотрим следующую задачу

$$g_0(x(t_1))\to\min,\quad \dot x=\varphi(t,x,u),\quad u(t)\in U,\quad x(t_0)=x_0,$$
$$g_i(x(t_1))=0,\quad i=1,\ldots,m,\quad (P_7')$$

на пространстве пар $(x(\cdot),u(\cdot))\in PC^1([t_0,t_1],\mathbb{R}^n)\times PC([t_0,t_1],\mathbb{R}^r)$.

В данном случае $H(t,x,u,p)=\langle p,\varphi(t,x,u)\rangle$.

Теорема 2 (Необходимые условия сильного минимума в задаче (P_7'))**.** *Если $(\widehat{x}(\cdot),\widehat{u}(\cdot))$ — оптимальный процесс в задаче (P_7'), то найдутся ненулевой вектор $\lambda=(\lambda_0,\lambda_1,\ldots,\lambda_m)\in(\mathbb{R}^{m+1})^*$ и функция $p(\cdot)\in PC^1([t_0,t_1],(\mathbb{R}^n)^*)$ такие, что выполняются*

(a) *условие стационарности по x:*

$$\dot p(t)=-\widehat{H}_x(t)\ \Leftrightarrow\ \dot p(t)=-p(t)\widehat{\varphi}_x(t),$$

(b) *условия трансверсальности:*

$$p(t_1)=-\sum_{i=0}^m\lambda_i g_i'(\widehat{x}(t_1)),$$

(c) *условие максимума по u для всех $t\in[t_0,t_1]$, где функция $\widehat{u}(\cdot)$ непрерывна:*

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

Основания для того, чтобы не писать еще условие условие трансверсальности в точке t_0 те же, что и в соответствующей задаче Лагранжа.

Перед непосредственным доказательством теоремы приведем некоторые определения и сформулируем одно утверждение.

Пусть $\widehat{u}(\cdot)\in PC([t_0,t_1],\mathbb{R}^r)$, τ — точка непрерывности $\widehat{u}(\cdot)$ и $v\in U$. Для всех $\alpha\ge 0$ таких, что $\widehat{u}(\cdot)$ непрерывна на отрезке $[\tau-\alpha,\tau]$ определим функцию из $PC([t_0,t_1],\mathbb{R}^r)$ по формуле

$$u(t,\alpha;\tau,v)=\begin{cases}\widehat{u}(t), & \text{если}\quad t\notin[\tau-\alpha,\tau);\\ v, & \text{если}\quad t\in[\tau-\alpha,\tau),\end{cases}$$

которую называют *игольчатой вариацией* функции $\widehat{u}(\cdot)$, а пару (τ,v) — *иголкой.* При $\alpha=0$ игольчатая вариация совпадает с $\widehat{u}(\cdot)$.

Понятие игольчатой вариации естественным образом распространяется на конечный набор иголок. Пусть $k\in\mathbb{N}$ и $\mathcal{N}_k$ — совокупность пар (иголок) $(\tau_i,v_i)\in[t_0,t_1]\times U$, $1\le i\le k$, где $t_0<\tau_1<\ldots<\tau_k<t_1$ и точки τ_i суть точки непрерывности $\widehat{u}(\cdot)$. Для каждого $\overline{\alpha}=(\alpha_1,\ldots,\alpha_k)^T\in\mathbb{R}^k_+$ такого, что отрезки $[\tau_i-\alpha_i,\tau_i]$, $1\le i\le k$, не пересекаются и на них $\widehat{u}(\cdot)$ непрерывна, определим функцию из $PC([t_0,t_1],\mathbb{R}^r)$ по формуле

$$u(t,\overline{\alpha};\mathcal{N}_k)=\begin{cases}\widehat{u}(t), & \text{если}\quad t\notin\cup_{i=1}^k[\tau_i-\alpha_i,\tau_i);\\ v_i, & \text{если}\quad t\in[\tau_i-\alpha_i,\tau_i),\ 1\le i\le k,\end{cases}$$

которая также называется *игольчатой вариацией* $\widehat{u}(\cdot)$, а набор $\mathcal{N}_k$ называется *пакетом иголок.*

Лемма 1 (о пакете иголок). *Пусть* $(\widehat{x}(\cdot), \widehat{u}(\cdot)) \in PC^1([t_0,t_1],\mathbb{R}^n) \times PC([t_0,t_1],\mathbb{R}^r)$, U *— непустое подмножество* $\mathbb{R}^r$, *отображение* $\varphi\colon \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r \to \mathbb{R}^n$ *удовлетворяет условиям гладкости, оговоренным выше, а функция* $\widehat{x}(\cdot)$ *является решением задачи Коши*

$$\dot{x} = \varphi(t,x,\widehat{u}(t)), \quad x(t_0) = x_0 \tag{1}$$

на отрезке $[t_0,t_1]$.

Если $k \in \mathbb{N}$ *и* $\mathcal{N}_k = \{(\tau_i, v_i),\ 1 \le i \le k\}$ *— пакет иголок, то найдется такая окрестность нуля в* $\mathbb{R}^k$, *что для любого* $\overline{\alpha} = (\alpha_1,\ldots,\alpha_k)^T \in \mathbb{R}^k_+$, *принадлежащего этой окрестности, существует единственное решение* $x(\cdot,\overline{\alpha};\mathcal{N}_k)$ *задачи Коши*

$$\dot{x} = \varphi(t,x,u(t,\overline{\alpha};\mathcal{N}_k)), \quad x(t_0) = x_0, \tag{2}$$

определенное на всем отрезке $[t_0,t_1]$.

Для любого $t \in [t_0,t_1]$ *отображение* $\overline{\alpha} \mapsto x(t,\overline{\alpha};\mathcal{N}_k)$ *может быть продолжено до непрерывно дифференцируемого отображения в некоторой окрестности нуля в* $\mathbb{R}^k$. *Частная производная по* α_i, $1 \le i \le k$, *в нуле этого отображения, которую обозначим* $x_{\alpha_i}(t) = x_{\alpha_i}(t,0;\mathcal{N}_k)$, *удовлетворяет на* $[\tau_i,t_1]$ *уравнению*

$$\dot{x}_{\alpha_i}(t) = \widehat{\varphi}_x(t)x_{\alpha_i}(t),\ x_{\alpha_i}(\tau_i) = \varphi(\tau_i,\widehat{x}(\tau_i),v_i) - \varphi(\tau_i,\widehat{x}(\tau_i),\widehat{u}(\tau_i)). \tag{3}$$

Доказательство теоремы 2. Пусть $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ — оптимальный процесс в задаче (P_7'), $k \in \mathbb{N}$, $\mathcal{N}_k = \{(\tau_i,v_i),\ 1 \le i \le k\}$ — пакет иголок и $u(\cdot,\overline{\alpha};\mathcal{N}_k)$ — соответствующая игольчатая вариация управления $\widehat{u}(\cdot)$. Согласно лемме о пакете иголок найдется такая окрестность нуля V в $\mathbb{R}^k$, что для любого $\overline{\alpha} \in V \cap \mathbb{R}^k_+$ существует единственное решение $x(\cdot,\overline{\alpha};\mathcal{N}_k)$ задачи Коши (2).

Пусть $V_0 \subset V$ — окрестность нуля, на которой отображение $\overline{\alpha} \mapsto x(t,\overline{\alpha};\mathcal{N}_k)$ непрерывно дифференцируемо. Рассмотрим следующую экстремальную задачу относительно переменной $\overline{\alpha}$:

$$\begin{gathered} g_0(x(t_1,\overline{\alpha};\mathcal{N}_k)) \to \min, \quad -\alpha_i \le 0, \quad i = 1,\ldots,k, \\ g_i(x(t_1,\overline{\alpha};\mathcal{N}_k)) = 0, \quad i = 1,\ldots,m, \quad \overline{\alpha} = (\alpha_1,\ldots,\alpha_k)^T \in V_0. \end{gathered} \tag{4}$$

Точка $\overline{\alpha} = 0$ доставляет локальный минимум в этой задаче. Проверка этого факта, фактически, такая же, как и для задачи Лагранжа (см. доказательство того, что $\widehat{u}$ — локальный минимум в задаче (4) §4.1).

Задача (4) имеет вид задачи (P_3), где $X = \mathbb{R}^k$, $f_0(\overline{\alpha}) = g_0(x(t_1,\overline{\alpha};\mathcal{N}_k))$, $f_i(\overline{\alpha}) = -\alpha_i$, $i = 1,\ldots,k$, $g = (g_1,...,g_m)^T$ и $F(\overline{\alpha}) = g(x(t_1,\overline{\alpha};\mathcal{N}_k))$ (снова, для формального соответствия задаче (P_3) считаем, что все функции, как функции $\overline{\alpha}$, продолжены за пределы V_0, например, нулем).

Отображение F, непрерывно дифференцируемы как суперпозиции непрерывно дифференцируемых отображений: $\overline{\alpha} \mapsto x(t_1,\overline{\alpha};\mathcal{N}_k) \mapsto g(x(t_1,\overline{\alpha};\mathcal{N}_k))$ и поэтому строго дифференцируемы в нуле (см. следствие 1 §1.2). Таким образом, условия теоремы 2 §3.2 выполнены и следовательно, найдутся множители Лагранжа $\lambda_0 \ge 0$, $\mu = (\mu_1,\ldots,\mu_k) \in (\mathbb{R}^k)^*$, $\mu \ge 0$ и $\lambda = (\lambda_1,\ldots,\lambda_m) \in (\mathbb{R}^m)^*$, не равные одновременно нулю, такие, что

$$\langle \lambda_0 f_0'(0), \overline{\alpha}\rangle - \langle \mu, \overline{\alpha}\rangle + \langle \lambda, F'(0)\overline{\alpha}\rangle = 0 \tag{5}$$

для всех $\overline{\alpha} \in \mathbb{R}^k$.

Так как $\mu \geq 0$, то $\langle \mu, \overline{\alpha} \rangle \geq 0$ для $\overline{\alpha} \in \mathbb{R}^k_+$. Учитывая это, получаем из (5) по теореме о суперпозиции дифференцируемых отображений, что

$$\langle \sum_{j=0}^{m} \lambda_j g_j'(\widehat{x}(t_1)), \sum_{i=1}^{k} x_{\alpha_i}(t_1)\alpha_i \rangle \geq 0, \quad \forall\, \overline{\alpha} = (\alpha_1, \ldots, \alpha_k)^T \in \mathbb{R}^k_+, \tag{6}$$

где, напомним, $x_{\alpha_i}(t_1) = x_{\alpha_i}(t_1, 0; \mathcal{N}_k)$, $i = 1, \ldots, k$. Согласно (3) частная производная $x_{\alpha_i}(\cdot)$ зависит только от иголки (τ_i, v_i), так что будем также писать $x_{\alpha_i}(t) = x_{\alpha_i}(t, 0; \tau_i, v_i)$, $t \in [\tau_i, t_1]$.

Отметим, что соотношение (6) справедливо для любого $k \in \mathbb{N}$ и любого пакета иголок $\mathcal{N}_k$ с некоторым набором множителями Лагранжа $\lambda(\mathcal{N}_k) = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, причем $\lambda(\mathcal{N}_k) \neq 0$, поскольку, в противном случае, из (5) последовало бы, что и $\mu = 0$. Далее считаем, что $|\lambda(\mathcal{N}_k)| = 1$. Наша ближайшая цель — показать существование такого набора $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, $|\lambda| = 1$, что соотношение (6) будет выполняться с этим λ и любым пакетом иголок $\mathcal{N}_k$.

Для пакета иголок $\mathcal{N}_k$ обозначим через $\Lambda(\mathcal{N}_k)$ множество всех таких векторов $\lambda(\mathcal{N}_k)$, $|\lambda(\mathcal{N}_k)| = 1$, которые удовлетворяют (6). Ясно, что $\Lambda(\mathcal{N}_k)$ — замкнутое подмножество компакта — единичной сферы в $(\mathbb{R}^{m+1})^*$. Покажем, что семейство $\mathcal{A}$ всех таких подмножеств образует центрированную систему.

Пусть $\Lambda(\mathcal{N}_{k_j})$, $1 \leq j \leq s$, — произвольное конечное множество элементов $\mathcal{A}$. Проверим, что $\cap_{j=1}^{s} \Lambda(\mathcal{N}_{k_j}) \neq \emptyset$. Положим $\mathcal{N}_l = \cup_{j=1}^{s} \mathcal{N}_{k_j}$ (т. е. $\mathcal{N}_l$ — объединение иголок из пакетов $\mathcal{N}_{k_j}$, $1 \leq j \leq s$, и l — число элементов в этом объединении). Некоторые точки τ_i могут совпадать и поэтому $\mathcal{N}_l$, вообще говоря, не является пакетом иголок по нашему определению. Поступим следующим образом. Сопоставим каждой точке τ_i точку $\tau_{i\varepsilon} = \tau_i + i\varepsilon$, где $\varepsilon > 0$. Тогда для достаточно малых ε будем иметь $t_0 < \tau_{1\varepsilon} < \ldots < \tau_{l\varepsilon} < t_1$ и, очевидно, $\tau_{i\varepsilon} \to \tau_i$ при $\varepsilon \to 0$.

Для каждого такого ε набор иголок $\mathcal{N}_{l(\varepsilon)} = \{(\tau_{i\varepsilon}, v_i),\ 1 \leq i \leq l\}$ уже является пакетом иголок в нашем смысле и, следовательно, для него справедливо соотношение (6) с некоторым $\lambda(\mathcal{N}_{l(\varepsilon)})$, $|\lambda(\mathcal{N}_{l(\varepsilon)})| = 1$. В силу теоремы о непрерывной зависимости решения дифференциального уравнения от начальных данных $x_{\alpha_i}(t_1, 0; \tau_{i\varepsilon}, v_i) \to x_{\alpha_i}(t_1, 0; \tau_i, v_i)$ при $\varepsilon \to 0$ (см. (3)). Из соображений компактности можно считать, что $\lambda(\mathcal{N}_{l(\varepsilon)})$ также сходится к некоторому $\lambda(\mathcal{N}_l)$, $|\lambda(\mathcal{N}_l)| = 1$, при $\varepsilon \to 0$ и таким образом, соотношение (6) справедливо для набора $\mathcal{N}_l$ с $\lambda(\mathcal{N}_l)$.

Пусть $1 \leq j \leq s$. Положив в этом соотношении $\alpha_i = 0$ для тех индексов i, для которых (τ_i, v_i) не принадлежит набору $\mathcal{N}_{k_j}$, получим, что $\lambda(\mathcal{N}_l) \in \Lambda(\mathcal{N}_{k_j})$ и тем самым $\lambda(\mathcal{N}_l) \in \cap_{j=1}^{s} \Lambda(\mathcal{N}_{k_j})$.

Итак, совокупность множеств $\mathcal{A}$ образует центрированную систему и значит, согласно лемме о центрированной системе (см. лемму **??** §1.1) существует $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, для которого справедливо соотношение (6) при любом $k \in \mathbb{N}$ и любом наборе $\mathcal{N}_k$. В частности, для любого набора $\mathcal{N}_1 = \{(\tau, v)\}$, где $(\tau, v) \in (t_0, t_1) \times U$ и τ — точка непрерывности $\widehat{u}(\cdot)$, из (6) следует при $\alpha = 1$ неравенство

$$\langle \sum_{j=0}^{m} \lambda_j g_j'(\widehat{x}(t_1)), x_\alpha(t_1, 0; \tau, v) \rangle \geq 0. \tag{7}$$

Пусть $p(\cdot)$ — решение задачи Коши для линейного дифференциального уравнения

$$\dot{p} = -p\,\widehat{\varphi}_x(t), \quad p(t_1) = -\sum_{j=0}^{m} \lambda_j g_j'(\widehat{x}(t_1)). \tag{8}$$

Используя неравенство (7), начальные условия в (3) и (8), а затем подставляя вместо $\dot{h}(\cdot)$ и $\dot{p}(\cdot)$ их выражения из соотношений (3) и (8), будем иметь

$$0 \le \langle \sum_{j=0}^{m} \lambda_j g_j'(\widehat{x}(t_1)), x_\alpha(t_1)\rangle = -\langle p(t_1), x_\alpha(t_1)\rangle = -\int_\tau^{t_1} (\langle \dot{p}(t), x_\alpha(t)\rangle + \langle p(t), \dot{x}_\alpha(t)\rangle)\, dt$$
$$- \langle p(\tau), x_\alpha(\tau)\rangle = -\langle p(\tau), x_\alpha(\tau)\rangle = -\langle p(\tau), \varphi(\tau, \widehat{x}(\tau), v) - \varphi(\tau, \widehat{x}(\tau), \widehat{u}(\tau)\rangle. \tag{9}$$

Отсюда следует, что $H(\tau, \widehat{x}(\tau), v, p(\tau)) \le H(\tau, \widehat{x}(\tau), \widehat{u}(\tau), p(\tau))$ для всех точек непрерывности τ функции $\widehat{u}(\cdot)$ и $v \in U$. Вместе с (8) это доказывает все утверждения теоремы. □

§3.5. Необходимые и достаточные условия экстремума второго порядка для простейшей задачи вариационного исчисления

Здесь идет речь о простейшей задаче вариационного исчисления, которая сформулирована в §4.3 и которая, напомним, имеет вид

$$J(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt \to \min, \quad x(t_i) = x_i, \quad i = 0, 1. \tag{Pr}$$

Ради простоты записи будем предполагать, что $n = 1$, т. е. $x(\cdot)$ — обычная функция. В этом случае пространства непрерывных, непрерывно дифференцируемых, кусочно непрерывных и кусочно непрерывно дифференцируемых функций на отрезке $[t_0, t_1]$ обозначаем соответственно $C([t_0, t_1])$, $C^1([t_0, t_1])$, $PC([t_0, t_1])$ и $PC^1([t_0, t_1])$.

Кроме того, положим $C_0^1([t_0, t_1]) = \{\, h(\cdot) \in C^1([t_0, t_1]) : h(t_0) = h(t_1) = 0 \,\}$ и $PC_0^1([t_0, t_1]) = \{\, h(\cdot) \in PC^1([t_0, t_1]) : h(t_0) = h(t_1) = 0 \,\}$.

Пусть фиксирована функция $\widehat{x}(\cdot) \in C^1([t_0, t_1])$. В этом параграфе мы предполагаем, что функция L непрерывна вместе со своими вторыми частными производными по x и $\dot{x}$ на некотором открытом множестве $G \subset \mathbb{R}^3$, содержащим точки $\{\, (t, \widehat{x}(t), \dot{\widehat{x}}(t)) : t \in [t_0, t_1] \,\}$.

Для вторых частных производных (аналогично тому, как это было сделано для первых производных) используем обозначения: $\widehat{L}_{\dot{x}\dot{x}}(t) = L_{\dot{x}\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$, $\widehat{L}_{\dot{x}x}(t) = L_{\dot{x}x}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$ и $\widehat{L}_{xx}(t) = L_{xx}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$. В силу условий на L справедливо равенство $\widehat{L}_{\dot{x}x}(t) = \widehat{L}_{x\dot{x}}(t)$.

3.5.1. Необходимые условия экстремума второго порядка

Функции $\widehat{x}(\cdot) \in C^1([t_0, t_1])$ и каждой функции $h(\cdot) \in PC_0^1([t_0, t_1])$ сопоставим выражение

$$Q(\widehat{x}(\cdot))(h(\cdot)) = \int_{t_0}^{t_1} (\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\widehat{L}_{\dot{x}x}(t)\dot{h}(t)h(t) + \widehat{L}_{xx}(t)h^2(t))\, dt,$$

которая, очевидно, является квадратичной формой на $PC_0^1([t_0,t_1])$.

Нам понадобится один технический результат, который приведем без доказательства (см. [2], п. 1.4.3).

Лемма 1 (о скруглении углов)**.**

$$\inf\{\, Q(\widehat{x}(\cdot))(h(\cdot)) : h(\cdot)\in PC_0^1([t_0,t_1])\,\} = \inf\{\, Q(\widehat{x}(\cdot))(h(\cdot)) : h(\cdot)\in C_0^1([t_0,t_1])\,\}. \quad (1)$$

Теорема 1 (Квадратичные условия минимума для задачи (Pr))**.** *Если $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (Pr), то для всех $t\in[t_0,t_1]$ выполнено уравнение Эйлера*

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

и для всех $h(\cdot)\in PC_0^1([t_0,t_1])$ справедливо неравенство

$$Q(\widehat{x}(\cdot))(h(\cdot))\ge 0. \quad (2)$$

Доказательство. Условия этой теоремы, очевидно, гарантируют выполнение условий теоремы о необходимых условиях минимума первого порядка для задачи (Pr) (см. теорему 2 §4.3) и поэтому справедливо уравнение Эйлера.

Докажем второе утверждение теоремы. Пусть сначала $h=h(\cdot)\in C_0^1([t_0,t_1])$. Положим $x(\cdot,\alpha;h)=\widehat{x}(\cdot)+\alpha h(\cdot)$, где $\alpha\in\mathbb{R}$. Ясно, что $x(\cdot,\alpha;h)\in C^1([t_0,t_1])$, $x(t_i,\alpha;h)=x_i$, $i=0,1$. Для достаточно малых по модулю α множество $\{(t,x(t,\alpha;h),\dot{x}(t,\alpha;h)) : t\in[t_0,t_1]\}$, очевидно, принадлежит G и тем самым функция $x(\cdot,\alpha;h)$ допустима в задаче (Pr).

Функция $\alpha\mapsto J(\alpha;h(\cdot))=J(x(\cdot,\alpha;h(\cdot)))$ в нуле имеет локальный минимум. Ее производная в силу теоремы о суперпозиции дифференцируемых функций имеет вид

$$J'(\alpha;h(\cdot))=\int_{t_0}^{t_1}\big(L_{\dot x}(t,\widehat{x}(t)+\alpha h(t),\dot{\widehat{x}}(t)+\alpha\dot h(t))\dot h(t) \\ + L_x(t,\widehat{x}(t)+\alpha h(t),\dot{\widehat{x}}(t)+\alpha\dot h(t))h(t)\big)\,dt.$$

Как хорошо известно, необходимым условием минимума второго порядка для функции одного переменного $\alpha\mapsto J(\alpha;h(\cdot))$ является неотрицательность ее второй производной в нуле. Непосредственные вычисления показывают, что $J''(0;h(\cdot))=Q(\widehat{x}(\cdot))(h(\cdot))$ и значит, для всех функций $h(\cdot)\in C_0^1([t_0,t_1])$ справедливо неравенство (2). В силу леммы 1 это неравенство справедливо и для функций $h(\cdot)\in PC_0^1([t_0,t_1])$. □

Дальнейшее в этом пункте связано с проверкой соотношения (2), равносильного, очевидно, тому, что функция $\widehat{h}(\cdot)=0$ доставляет глобальный минимум в задаче

$$Q(\widehat{x}(\cdot))(h(\cdot))\to\min,\quad h(\cdot)\in PC_0^1([t_0,t_1]). \quad (3)$$

В силу однородности функционала $h(\cdot)\mapsto Q(\widehat{x}(\cdot))(h(\cdot))$ функция $\widehat{h}(\cdot)=0$ является глобальным минимумом (проверьте это). Знакоопределенность этого функционала напрямую связана с вопросом существования нетривиального глобального минимума в задаче (3).

Предложение 1 (Принцип максимума для задачи (3)). *Если функция* $\widehat{h}(\cdot)$ *является глобальным минимумом задачи* (3), *то найдется функция* $p(\cdot) \in PC^1([t_0, t_1])$ *такая, что*

$$\dot{p}(t) = 2\widehat{L}_{\dot{x}x}(t)\dot{\widehat{h}}(t) + 2\widehat{L}_{xx}(t)\widehat{h}(t), \tag{4}$$

и

$$\max_{u\in\mathbb{R}}(p(t)u - \widehat{L}_{\dot{x}\dot{x}}(t)u^2 - 2\widehat{L}_{\dot{x}x}(t)u\widehat{h}(t)) = p(t)\dot{\widehat{h}}(t) - \widehat{L}_{\dot{x}\dot{x}}(t)\dot{\widehat{h}}^2(t) - 2\widehat{L}_{\dot{x}x}(t)\dot{\widehat{h}}(t)\widehat{h}(t) \tag{5}$$

в точках t, *где функция* $\dot{\widehat{h}}(\cdot)$ *непрерывна.*

Доказательство. Рассмотрим задачу (3) как задачу оптимального управления

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

$$\dot{h} = u, \quad u \in \mathbb{R}, \quad h(t_0) = 0, \quad h(t_1) = 0. \tag{6}$$

Если $\widehat{h}(\cdot)$ — ее решение, то пара $(\widehat{h}(\cdot), \widehat{u}(\cdot))$, где $\widehat{u}(\cdot) = \dot{\widehat{h}}(\cdot)$, является сильным минимумом в задаче (6). Тогда согласно теореме 1 §4.3 найдутся числа $\lambda_0 \geq 0$, λ_1, λ_2, не равные одновременно нулю, и функция $p(\cdot) \in PC^1([t_0, t_1])$ такие, что справедливо соотношение, которое получается из уравнения (4), умножением его правой части на λ_0 и добавлением краевых условий: $p(t_0) = \lambda_1$, $p(t_1) = -\lambda_2$ и соотношение, которое получается из (5) умножением на λ_0 второго и третьего слагаемых справа и слева под знаком максимума.

Но $\lambda_0 \neq 0$, так как в противном случае, из (**??**) последовало бы, что $p(\cdot)$ — константа, а из (5) — что эта константа нулевая. Тогда из краевых условий получим, что $\lambda_1 = \lambda_2 = 0$, т. е. все множители Лагранжа нулевые. Следовательно, можно считать, что $\lambda_0 = 1$ и тем самым предложение 1 доказано. $\square$

Под знаком максимума в (5) стоит гладкая функция по u, которая в точке $\dot{\widehat{h}}(t)$, где t — точка непрерывности $\dot{\widehat{h}}(\cdot)$, достигает своего максимума и тем самым ее производная в этой точке равняется нулю:

$$p(t) = 2\widehat{L}_{\dot{x}\dot{x}}(t)\dot{\widehat{h}}(t) + 2\widehat{L}_{\dot{x}x}(t)\widehat{h}(t). \tag{7}$$

Отсюда и (4) следует, что если $\widehat{h}(\cdot)$ есть решение (глобальный минимум) задачи (3), то $\widehat{h}(\cdot)$ необходимо удовлетворяет (для всех $t \in [t_0, t_1]$, где функция $\dot{\widehat{h}}(\cdot)$ непрерывна) следующему уравнению

$$-\frac{d}{dt}\left[\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}(t) + \widehat{L}_{\dot{x}x}(t)h(t)\right] + \widehat{L}_{\dot{x}x}(t)\dot{h}(t) + \widehat{L}_{xx}(t)h(t) = 0, \tag{8}$$

которое называется *уравнением Якоби* задачи (Pr).

Введем теперь некоторые определения. Пусть $\widehat{x}(\cdot)$ — *экстремаль* задачи (Pr), т. е. $\widehat{x}(\cdot)$ удовлетворяет уравнению Эйлера для этой задачи.

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

(2) Пусть на $\widehat{x}(\cdot)$ выполнено усиленное условие Лежандра. Точка $\tau \in (t_0, t_1]$ называется *сопряженной точкой* к точке t_0, если существует нетривиальное решение $h(\cdot) \in C_0^1([t_0, t_1])$ уравнения Якоби такое, что $h(t_0) = h(\tau) = 0$.

(3) Говорят, что на функции $\widehat{x}(\cdot)$ выполнено *условие Якоби*, если в интервале (t_0, t_1) нет точек, сопряженных с t_0 и *усиленное условие Якоби*, если в полуинтервале $(t_0, t_1]$ нет точек, сопряженных с t_0.

Теорема 2 (Необходимое условие Лежандра)**.** *Если $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (Pr), то выполнено условие Лежандра.*

Доказательство. Из квадратичных условий минимума (см. теорему 1) следует, что $Q(\widehat{x}(\cdot))(h(\cdot)) \geq 0$ для любого $h(\cdot) \in PC_0^1([t_0, t_1])$ и значит, функция $\widehat{h}(\cdot) = 0$ является решением задачи (3). Тогда согласно предложению 1 для этой функции выполняется соотношение (5), которое в данном случае означает, что при каждом $t \in [t_0, t_1]$ функция $u \mapsto p(t)u - \widehat{L}_{\dot{x}\dot{x}}(t)u^2$ на $\mathbb{R}$ достигает максимума в нуле. Следовательно, ее первая производная равна нулю, т. е. $p(\cdot) = 0$, а вторая производная неположительна, или равносильно, $\widehat{L}_{\dot{x}\dot{x}}(t) \geq 0$ для всех $t \in [t_0, t_1]$. □

Для дальнейших рассмотрений понадобится следующая конструкция. Пусть функция $x(\cdot) \in C^1([t_0, t_1])$ является решением уравнения Якоби (8). Положим $y(\cdot) = \widehat{L}_{\dot{x}\dot{x}}(\cdot)\dot{x}(\cdot) + \widehat{L}_{\dot{x}x}(\cdot)x(\cdot)$. Согласно (8) это дифференцируемая функция и $\dot{y}(\cdot) = \widehat{L}_{\dot{x}x}(\cdot)\dot{x}(\cdot) + \widehat{L}_{xx}(\cdot)x(\cdot)$. Пусть выполнено усиленное условие Лежандра. Тогда выражая $\dot{x}(\cdot)$ из первого соотношения, а затем подставляя его во второе, получаем, что пара $(x(\cdot), y(\cdot))$ является решением следующей системы линейных уравнений

$$\begin{cases} \dot{x} = -\dfrac{\widehat{L}_{\dot{x}x}(t)}{\widehat{L}_{\dot{x}\dot{x}}(t)}\, x + \dfrac{1}{\widehat{L}_{\dot{x}\dot{x}}(t)}\, y \\[2ex] \dot{y} = \left(\widehat{L}_{xx}(t) - \dfrac{\widehat{L}^2_{\dot{x}x}(t)}{\widehat{L}_{\dot{x}\dot{x}}(t)}\right) x + \dfrac{\widehat{L}_{\dot{x}x}(t)}{\widehat{L}_{\dot{x}\dot{x}}(t)}\, y. \end{cases} \tag{9}$$

Верно, очевидно, и обратное, если пара $(x(\cdot), y(\cdot))$ — решение системы (9), то $x(\cdot)$ — решение уравнения Якоби (8).

Поскольку система (9) линейна, то для любых начальных условий она имеет единственное решение, определенное на всем отрезке $[t_0, t_1]$.

Теорема 3 (Необходимое условие Якоби)**.** *Если $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (Pr) и выполнено усиленное условие Лежандра, то выполнено условие Якоби.*

Доказательство. Доказываем от противного. Пусть $\tau \in (t_0, t_1)$ и $\overline{h}(\cdot) \in C_0^1([t_0, t_1])$ — нетривиальное решение уравнения Якоби такое, что $\overline{h}(\tau) = 0$. Положим

$$\widetilde{h}(t) = \begin{cases} \overline{h}(t), & t \in [t_0, \tau]; \\ 0, & t \in (\tau, t_1]. \end{cases}$$

Интегрируя по частям, будем иметь

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

Отсюда, так как $\overline{h}(\cdot)$ удовлетворяет уравнению Якоби, получаем, что $Q(\widehat{x}(\cdot))(\widetilde{h}(\cdot)) = 0$. Из квадратичных условий минимума (см. теорему 1) следует, что $Q(\widehat{x}(\cdot))(h(\cdot)) \geq 0$ для любого $h(\cdot) \in PC_0^1([t_0, t_1])$ и тем самым $\widetilde{h}(\cdot)$ является решением задачи (3). Тогда, как уже было отмечено выше, эта функция должна удовлетворять уравнению Якоби (8) (в тех точках t, где производная $\dot{\widetilde{h}}(\cdot)$ непрерывна). Покажем, что это невозможно.

Действительно, из уравнения Якоби (8) следует, что функция $q(\cdot) = \widehat{L}_{\dot{x}\dot{x}}(\cdot)\dot{\widetilde{h}}(\cdot) + \widehat{L}_{\dot{x}x}(\cdot)\widetilde{h}(\cdot)$ непрерывна. Далее, с одной стороны, $\widetilde{h}(t) = \dot{\widetilde{h}}(t) = 0$ для $t \in (\tau, t_1]$ и поэтому $q(\tau+0) = 0$. С другой, $q(\tau-0) = \widehat{L}_{\dot{x}\dot{x}}(\tau)\dot{\widetilde{h}}(\tau-0) = \widehat{L}_{\dot{x}\dot{x}}(\tau)\dot{\overline{h}}(\tau)$. Заметим теперь, что $\dot{\overline{h}}(\tau) \neq 0$, ибо в противном случае функция $\overline{h}(\cdot)$ удовлетворяла бы уравнению Якоби с начальными условиями $\overline{h}(\tau) = \dot{\overline{h}}(\tau) = 0$. Но тогда пара $(x(\cdot), y(\cdot))$, где $x(\cdot) = \overline{h}(\cdot)$ и $y(\cdot) = \widehat{L}_{\dot{x}\dot{x}}(\cdot)\dot{\overline{h}}(\cdot) + \widehat{L}_{\dot{x}x}(\cdot)\overline{h}(\cdot)$, была бы решением системы (9) с начальными условиями $x(\tau) = y(\tau) = 0$ и значит, в силу единственности $x(\cdot) = \overline{h}(\cdot) = 0$, что по предположению не так. Поскольку $\widehat{L}_{\dot{x}\dot{x}}(\tau) > 0$ вследствие того, что выполнено усиленное условие Лежандра, то $q(\tau - 0) \neq 0$ и значит, функция $q(\cdot)$ разрывна в точке τ. Полученное противоречие доказывает теорему. □

3.5.2. Достаточные условия экстремума второго порядка

Теорема 4 (Достаточные условия слабого минимума в задаче (Pr) в терминах квадратичной формы)**.** *Пусть $\widehat{x}(\cdot)$ — допустимая функция в задаче (Pr). Тогда, если $\widehat{x}(\cdot)$ для каждого $t \in [t_0, t_1]$ удовлетворяет уравнению Эйлера*

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

и существует такое $\alpha > 0$, что для всех $h(\cdot) \in C_0^1([t_0, t_1])$ справедливо неравенство

$$
\int_{t_0}^{t_1}\left(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\widehat{L}_{\dot{x}x}(t)\dot{h}(t)h(t) + \widehat{L}_{xx}(t)h^2(t)\right)dt \geq \alpha\int_{t_0}^{t_1}\dot{h}^2(t)\,dt, \tag{10}
$$

то $\widehat{x}(\cdot)$ доставляет слабый локальный минимум в задаче (Pr).

Доказательство. Покажем сначала, что существует такая константа $c > 0$, что для любой функции $h(\cdot) \in C^1([t_0, t_1])$, для которой $h(t_0) = 0$, справедливо неравенство

$$
\frac{1}{2}\int_{t_0}^{t_1}(\dot{h}^2(t) + 2|\dot{h}(t)h(t)| + h^2(t))\,dt \leq c\int_{t_0}^{t_1}\dot{h}^2(t)\,dt. \tag{11}
$$

Действительно, используя неравенство Коши–Буняковского, будем иметь для любого $t \in [t_0, t_1]$

$$|h(t)| = \left| \int_{t_0}^{t} \dot{h}(\tau)\, d\tau \right| \le \int_{t_0}^{t_1} |\dot{h}(t)|\, dt \le \sqrt{t_1 - t_0} \left(\int_{t_0}^{t_1} \dot{h}^2(t)\, dt \right)^{1/2}.$$

Следовательно,

$$\int_{t_0}^{t_1} h^2(t)\, dt \le (t_1 - t_0) \max_{t \in [t_0, t_1]} |h(t)|^2 \le (t_1 - t_0)^2 \int_{t_0}^{t_1} \dot{h}^2(t)\, dt.$$

Так как, очевидно, $|\dot{h}(t)h(t)| \le (\dot{h}^2(t) + h^2(t))/2$, то

$$\frac{1}{2} \int_{t_0}^{t_1} (\dot{h}^2(t) + 2|\dot{h}(t)h(t)| + h^2(t))\, dt \le \int_{t_0}^{t_1} (\dot{h}^2(t) + h^2(t))\, dt$$
$$\le (1 + (t_1 - t_0)^2) \int_{t_0}^{t_1} \dot{h}^2(t)\, dt = c \int_{t_0}^{t_1} \dot{h}^2(t)\, dt.$$

Пусть теперь $\varepsilon > 0$ таково, что $\varepsilon c \le \alpha$. Нетрудно проверить, что существует такое $\delta_0 > 0$, что компакт $K = \{\, (t, x, \dot{x}) : |x - \widehat{x}(t)| + |\dot{x} - \dot{\widehat{x}}(t)| \le \delta_0,\ t \in [t_0, t_1] \,\}$ принадлежит G. Функции $L_{\dot{x}\dot{x}}(t, x, \dot{x})$, $L_{\dot{x}x}(t, x, \dot{x})$ и $L_{xx}(t, x, \dot{x})$ равномерно непрерывны на K. Следовательно найдется $0 < \delta = \delta(\varepsilon) \le \delta_0/2$ такое, что

$$|L_{\dot{x}\dot{x}}(t, x_1, \dot{x}_1) - L_{\dot{x}\dot{x}}(t, x_2, \dot{x}_2)| < \varepsilon, \quad |L_{\dot{x}x}(t, x_1, \dot{x}_1) - L_{\dot{x}x}(t, x_2, \dot{x}_2)| < \varepsilon,$$
$$|L_{xx}(t, x_1, \dot{x}_1) - L_{xx}(t, x_2, \dot{x}_2)| < \varepsilon \qquad (12)$$

для $(t, x_i, \dot{x}_i) \in K$, $i = 1, 2$, для которых $|x_1 - x_2| < \delta$ и $|\dot{x}_1 - \dot{x}_2| < \delta$.

Пусть функция $x(\cdot)$ допустима в задаче (Pr) и $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^1([t_0,t_1])} < \delta$. Положим $h(\cdot) = x(\cdot) - \widehat{x}(\cdot)$ (тогда $h(\cdot) \in C_0^1([t_0, t_1])$). Для всех $t \in [t_0, t_1]$ точки $(t, \widehat{x}(t) + h(t), \dot{\widehat{x}}(t) + \dot{h}(t))$ принадлежат K и тем самым G. По формуле Тейлора

$$L(t, \widehat{x}(t) + h(t), \dot{\widehat{x}}(t) + \dot{h}(t)) = \widehat{L}(t) + \widehat{L}_x(t)h(t) + \widehat{L}_{\dot{x}}(t)\dot{h}(t) + \frac{1}{2}(L_{xx}(t, \widehat{x}(t)$$
$$+ \theta h(t), \dot{\widehat{x}}(t) + \theta \dot{h}(t))h^2(t) + 2L_{x\dot{x}}(t, \widehat{x}(t) + \theta h(t), \dot{\widehat{x}}(t) + \theta \dot{h}(t))h(t)\dot{h}(t)$$
$$+ L_{\dot{x}\dot{x}}(t, \widehat{x}(t) + \theta h(t), \dot{\widehat{x}}(t) + \theta \dot{h}(t))\dot{h}^2(t)), \qquad (13)$$

где $\theta = \theta(t) \in (0, 1)$.

Обозначим через $q(t)$ выражение под знаком интеграла слева в (10). Прибавим и вычтем величину $(1/2)q(t)$ из правой части равенства (13). Тогда в силу соотношений (12), будем иметь

$$L(t, \widehat{x}(t) + h(t), \dot{\widehat{x}}(t) + \dot{h}(t)) - \widehat{L}(t) \ge \widehat{L}_x(t)h(t) + \widehat{L}_{\dot{x}}(t)\dot{h}(t)$$
$$+ \frac{1}{2}q(t) - \frac{\varepsilon}{2}(h^2(t) + 2|h(t)\dot{h}(t)| + \dot{h}^2(t)).$$

Интегрируя это соотношение, затем интегрируя первое слагаемое справа по частям и используя оценки (10) и (11), получим

$$J(x(\cdot)) - J(\widehat{x}(\cdot)) \geq \int_{t_0}^{t_1} \left(-\int_{t_0}^{t} \widehat{L}_x(\tau)\, d\tau + \widehat{L}_{\dot{x}}(t) \right) \dot{h}(t)\, dt$$
$$+ \frac{\alpha}{2} \int_{t_0}^{t_1} \dot{h}^2(t)\, dt - \frac{\varepsilon c}{2} \int_{t_0}^{t_1} \dot{h}^2(t)\, dt$$

Первое слагаемое справа здесь равно нулю в силу того, что $\widehat{x}(\cdot)$ удовлетворяет уравнению Эйлера. Сумма остальных слагаемых неотрицательна в силу выбора ε. Таким образом, $J(x(\cdot)) \geq J(\widehat{x}(\cdot))$ для всех допустимых $x(\cdot)$ таких, что $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^1([t_0,t_1])} < \delta$ и значит, $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (Pr). □

Теорема 5 (Достаточные условия слабого минимума в простейшей задаче в терминах условий Лежандра и Якоби)**.** *Пусть интегрант L непрерывен вместе со своими частными производными до третьего порядка включительно на открытом множестве $G \subset \mathbb{R}^3$, допустимая в задаче (Pr) функция $\widehat{x}(\cdot)$ такова, что $\widehat{x}(\cdot) \in C^2([t_0, t_1])$, $\widehat{x}(\cdot)$ удовлетворяет уравнению Эйлера и на $\widehat{x}(\cdot)$ выполнены усиленные условия Лежандра и Якоби. Тогда $\widehat{x}(\cdot)$ доставляет слабый локальный минимум в задаче (Pr).*

Доказательство. Покажем, что в условиях теоремы справедливо неравенство (10). Тогда из теоремы о достаточных условиях в терминах квадратичной формы, очевидно, будет следовать утверждение данной теоремы.

Обозначим для краткости $A(\cdot) = \widehat{L}_{\dot{x}\dot{x}}(\cdot)$, $B(\cdot) = \widehat{L}_{\dot{x}x}(\cdot)$, $C(\cdot) = \widehat{L}_{xx}(\cdot)$ и докажем, что

$$Q(h(\cdot)) = \int_{t_0}^{t_1} \Big(A(t)\dot{h}^2(t) + 2B(t)\dot{h}(t)h(t) + C(t)h^2(t) \Big)\, dt \geq 0$$

для всех $h(\cdot) \in C_0^1([t_0, t_1])$.

Пусть $(x(\cdot), y(\cdot))$ — решение (9) с условиями: $x(t_0) = 0$, $y(t_0) = 1$. Ясно, что $x(\cdot) \neq 0$, $y(\cdot) \neq 0$ и из первого уравнения следует, что $\dot{x}(t_0) = \widehat{L}_{\dot{x}\dot{x}}^{-1}(t_0) > 0$. Как уже было отмечено ранее, $x(\cdot)$ является решением уравнения Якоби (см. (8)). Покажем, что справедлива формула

$$Q(h(\cdot)) = \int_{t_0}^{t_1} A(t) \left(\dot{h}(t) - \frac{\dot{x}(t)}{x(t)}\, h(t) \right)^2 dt. \tag{14}$$

Заметим, что величина справа имеет смысл, поскольку в силу усиленного условия Якоби функция $x(\cdot)$ не обращается в ноль на $(t_0, t_1]$, а отношение $h(t)/x(t)$ стремится к конечному пределу при $t \to t_0$ (согласно правилу Лопиталя).

Обозначим через $I(h(\cdot))$ интеграл справа в (14). Имеем

$$I(h(\cdot)) = \int_{t_0}^{t_1} \left(A(t)\dot{h}^2(t) - 2A(t)\frac{\dot{x}(t)}{x(t)}\dot{h}(t)h(t) + A(t)\frac{\dot{x}^2(t)}{x^2(t)}h^2(t) \right) dt.$$

Проинтегрируем по частям во втором слагаемом. Это возможно в силу того, что функция $A(\cdot)\dot{x}(\cdot)/x(\cdot)$ дифференцируема на $(t_0, t_1]$. Действительно, функции $x(\cdot)$ удовлетворяет уравнению Якоби, которое в новых обозначениях имеет вид

$$-\frac{d}{dt}\left[A(t)\dot{x}(t) + B(t)x(t)\right] + B(t)\dot{x}(t) + C(t)x(t) = 0. \tag{15}$$

Выражение в квадратных скобках дифференцируемо, а поскольку функция $B(\cdot)x(\cdot)$ также дифференцируема в силу условий теоремы, то и функция $A(\cdot)\dot{x}(\cdot)$ дифференцируема. Следовательно, функция $A(\cdot)\dot{x}(\cdot)/x(\cdot)$ дифференцируема на $(t_0, t_1]$. Итак, после интегрирования по частям получаем

$$I(h(\cdot)) = \int_{t_0}^{t_1} \left(A(t)\dot{h}^2(t) + \left(\frac{d}{dt} \left(A(t)\frac{\dot{x}(t)}{x(t)} \right) + A(t)\frac{\dot{x}^2(t)}{x^2(t)} \right) h^2(t) \right) dt.$$

Преобразуем выражение перед $h^2(t)$, дифференцируя первое слагаемое, как отношение двух функций, затем используя уравнение (15) и формулу дифференцирования произведения функций (штрих, как и точка, обозначает производную по t)

$$\begin{aligned}\frac{(A(t)\dot{x}(t))'x(t) - \dot{x}(t)A(t)\dot{x}(t)}{x^2(t)} + A(t)\frac{\dot{x}^2(t)}{x^2(t)} &= \frac{(A(t)\dot{x}(t))'}{x(t)} \\ &= \frac{-(B(t)x(t))' + B(t)\dot{x}(t) + C(t)x(t)}{x(t)} = -\dot{B}(t) + C(t).\end{aligned}$$

Подставляя полученное в предыдущее равенство и интегрируя по частям, приходим к соотношению

$$\begin{aligned}I(h(\cdot)) &= \int_{t_0}^{t_1} \left(A(t)\dot{h}^2(t) - \dot{B}(t)h^2(t) + C(t)h^2(t) \right) dt \\ &= \int_{t_0}^{t_1} \left(A(t)\dot{h}^2(t) + 2B(t)\dot{h}(t)h(t) + C(t)h^2(t) \right) dt,\end{aligned}$$

которое означает, что выполнено равенство (14) и тем самым $Q(h(\cdot)) \geq 0$ на $C_0^1([t_0, t_1])$.

Покажем теперь, что отсюда следует (10). Для любого $0 < \alpha < \min_{t\in[t_0,t_1]} A(t)$ функция $A_\alpha(\cdot) = A(\cdot) - \alpha$ положительна на $[t_0, t_1]$. Для таких α рассмотрим систему (9) с заменой $A(\cdot) = \widehat{L}_{\dot{x}\dot{x}}(\cdot)$ на $A_\alpha(\cdot)$. Пусть $(x_\alpha(\cdot), y_\alpha(\cdot))$ — ее решение с начальными условиями: $x_\alpha(t_0) = 0$, $y_\alpha(t_0) = 1$. Так как $A_\alpha(t_0) < A(t_0)$, то из первого уравнения этой системы получаем, что $\dot{x}_\alpha(t_0) = A_\alpha^{-1}(t_0) > A^{-1}(t_0) > 0$. Отсюда легко следует существование такого $\varepsilon > 0$, что $x_\alpha(t) > 0$ для $t \in (t_0, \varepsilon]$ и всех указанных α. По теореме о непрерывной зависимости решения дифференциального уравнения от параметров получаем, что $x_\alpha(\cdot) \to x(\cdot)$ при $\alpha \to 0$ в $C([t_0, t_1])$. Функция $x(\cdot)$ положительна на $(t_0, t_1]$, поскольку она не обращается в ноль на этом промежутке, а ее производная в нуле положительна. Тогда для достаточно малых $\alpha > 0$ функции $x_\alpha(\cdot)$ будут положительны на $[\varepsilon, t_1]$ и значит, на $(t_0, t_1]$. Фиксируем такое α и проводя те же рассуждения, что и выше с заменой $x(\cdot)$ на $x_\alpha(\cdot)$, получаем, что

$$Q_\alpha(h(\cdot)) = \int_{t_0}^{t_1} (A_\alpha(t)\dot{h}^2(t) + 2B(t)\dot{h}(t)h(t) + C(t)h^2(t))\, dt \geq 0$$

для всех $h(\cdot) \in C_0^1([t_0, t_1])$. Следовательно,

$$Q(h(\cdot)) = Q_\alpha(h(\cdot)) + \alpha \int_{t_0}^{t_1} \dot{h}^2(t)\, dt \geq \alpha \int_{t_0}^{t_1} \dot{h}^2(t)\, dt,$$

т. е. справедливо неравенство (10), что, как было отмечено выше, доказывает теорему. $\square$

Список литературы

[1] А. Д. Иоффе. Теория экстремальных задач. М.: Наука, 1974.

[2] В. М. Алексеев, В. М. Тихомиров, С. В. Фомин. Оптимальное управление. М.: Физматлит, 2005.

[3] Е. Р. Аваков, Г. Г. Магарил-Ильяев. Курс лекций по “Теории экстремума”. М.: Издательство Попечительского совета механико-математического факультета МГУ, 2018.