МОСКОВСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. Ломоносова
МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ

Э. М. Галеев

Спецкурс годовой

Условия экстремума 2 порядка в теории экстремальных задач

(можно сдавать как два полугодовых:
1-семестр - Глава 1
2-семестр - Главы 3,5)

Москва - 2023

Оглавление

1 Конечномерные задачи **7**

1 Конечномерные гладкие задачи без ограничений 7

1.1 Постановка задачи 7

1.2 Необходимые и достаточные условия экстремума 8

1.2.1 Функции одной переменной . . . 8

1.2.2 Функции нескольких переменных 10

1.2.3 Теорема Вейерштрасса 12

1.2.4 Критерий Сильвестра 12

1.3 Примеры 14

2 Конечномерные гладкие задачи с равенствами . 18

2.1 Постановка задачи 18

2.2 Необходимые и достаточные условия экстремума 19

2.2.1 Принцип Лагранжа 19

2.2.2 Теорема об обратной функции . 20

2.2.3 Необходимое условие 2-го порядка 21

2.2.4 Достаточное условие 2-го порядка 23

2.3 Примеры 25

3 Конечномерные гладкие задачи с равенствами и неравенствами 26

3.1 Постановка задачи 26

3.2 Необходимые и достаточные условия экстремума 27
3.2.1 Принцип Лагранжа 27
3.2.2 Необходимое условие 2-го порядка 32
3.2.3 Достаточное условие 2-го порядка 33
3.3 Примеры 34
4 Выпуклый анализ 40
4.1 Элементы выпуклого анализа 40
4.1.1 Выпуклые множества 40
4.1.2 Выпуклые функции 42
4.1.3 Субдифференциал функции . . . 51
4.1.4 Сублинейные функции 53
4.2 Отделимость выпуклых множеств 56
4.2.1 Теоремы отделимости 57
4.2.2 Теоремы о строгой отделимости 61
4.3 Выпуклые задачи 63
4.3.1 Задачи без ограничений 63
4.3.2 Задачи с ограничением 64
4.3.3 Задача выпуклого программирования 65
4.3.4 Пример 69
5 Элементы функционального анализа 71
5.1 Линейные, нормированные пространства 71
5.1.1 Сопряженное пространство . . . 71
5.2 Определения производных 72
5.2.1 Производная по направлению . . 72
5.2.2 Вариация по Лагранжу 73
5.2.3 Производная Гато 73
5.2.4 Производная Фреше 74
5.2.5 Строгая дифференцируемость . 76
5.2.6 Частные производные 76
5.2.7 Производные высших порядков . 77
5.2.8 Контрпримеры 78
5.3 Основные теоремы 82
5.3.1 Теорема о суперпозиции 82

5.3.2 Формула Тейлора 82
5.4 Дополнительные сведения 83
6 Гладкая задача без ограничений 91
6.1 Постановка задачи 91
6.2 Необходимые условия 1-го порядка 91
6.3 Необходимые и достаточные условия . . . 92
7 Гладкая задача с равенствами в банаховых пространствах 95
7.1 Постановка задачи 95
7.2 Необходимые условия 1-го порядка 95
7.3 Необходимые условия 2-го порядка 98
7.4 Достаточные условия 2-го порядка 100
7.5 Квадратично-линейная задача 103
8 Гладкая задача с равенствами и неравенствами . 105
8.1 Постановка задачи 105
8.2 Необходимые условия 1-го порядка 105
8.3 Необходимые условия 1-го и 2-го порядка 106
8.4 Достаточные условия 2-го порядка 107
8.5 Квадратично-линейная задача 108

3 Вариационное исчисление 110
1 Простейшая задача ВИ 110
1.1 Постановка задачи 110
1.2 Вывод уравнения Эйлера с помощью основной леммы вариационного исчисления 111
1.3 Вывод уравнения Эйлера с помощью леммы Дюбуа-Реймона 113
1.4 Интегралы уравнения Эйлера 114
1.5 Примеры 115
2 Задача Больца 118
2.1 Постановка задачи 118
2.2 Необходимое условие экстремума 118
2.3 Примеры 121
3 Задача с подвижными концами 123
3.1 Постановка задачи 123

3.2 Необходимые условия экстремума 124
3.3 Примеры 125
4 Изопериметрическая задача 128
4.1 Постановка задачи 128
4.2 Необходимое условие экстремума 129
4.3 Примеры 129
4.4 Задача Дидоны 132
5 Задача со старшими производными 134
5.1 Постановка задачи 134
5.2 Вывод уравнения Эйлера–Пуассона 134
5.3 Пример 136
6 Задача Лагранжа 138
6.1 Постановка задачи 138
6.2 Необходимые условия экстремума 139
6.3 Примеры 145

5 Условия 2-го порядка в КВИ 149
1 Простейшая задача вариационного исчисления . 149
1.1 Сильный и слабый экстремум 149
1.2 Пример слабого, но не сильного 151
1.2.1 Лемма о скруглении углов 153
1.3 Теорема Боголюбова 155
1.4 Условия Лежандра, Якоби, Вейерштрасса 157
1.5 Необходимые и достаточные условия слабого и сильного экстремума 165
1.5.1 Игольчатые вариации. Условие Вейерштрасса. 165
1.5.2 Необходимые условия сильного . 169
1.5.3 Необходимые условия слабого . . 172
1.5.4 Поле экстремалей 175
1.5.5 Достаточные условия сильного . 181
1.5.6 Представление квадратичного функционала 185
1.5.7 Достаточные условия слабого . . 190

1.5.8 Квадратично-линейный функционал 193
2 Задача Больца 210
2.1 Сильный и слабый экстремум 210
2.2 Условия экстремума 2-го порядка 211
2.2.1 Необходимые условия слабого экстремума. 211
2.2.2 Достаточные условия сильного экстремума. 214
2.2.3 Квадратичный функционал . . . 217
2.3 Пример. 220
3 Изопериметрическая задача 223
3.1 Постановка задачи 223
3.2 Условия Лежандра, Якоби и регулярности 224
3.3 Условия экстремума 2-го порядка 227
3.3.1 Необходимые условия слабого экстремума 227
3.3.2 Достаточные условия сильного экстремума 227
3.3.3 Квадратичный функционал . . . 228
3.4 Пример 235
4 Задача со старшими производными 238
4.1 Постановка задачи 238
4.2 Условия Лежандра и Якоби 239
4.3 Условия экстремума II порядка 241
4.3.1 Необходимые условия слабого экстремума 241
4.3.2 Достаточные условия сильного экстремума 242
4.3.3 Квадратичный функционал . . . 245
4.4 Пример 248

Глава 1

Конечномерные задачи

1 Конечномерные гладкие задачи без ограничений

В этом параграфе даются необходимые и достаточные условия экстремума функций одной и нескольких переменных.

1.1 Постановка задачи

Пусть $f\colon \mathbb{R}^n \to \mathbb{R}$ — функция n действительных переменных, обладающая некоторой гладкостью. Под гладкостью мы понимаем определенную дифференцируемость функции. Если функция f дифференцируема k раз в точке $\widehat{x}$, мы пишем $f \in D^k(\widehat{x})$. *Гладкой конечномерной экстремальной задачей без ограничений* называется следующая задача:

$$f(x) \to \mathrm{extr}.$$

При решении задачи надо найти не только абсолютные (глобальные) экстремумы (минимумы и максимумы) функции, но и локальные экстремумы.

Определение 1. Точка $\widehat{x}$ является *точкой локального минимума* функции f (пишем $\widehat{x} \in \operatorname{locmin} f$), если существу-

ет окрестность $U_\varepsilon = \{x \mid |x - \widehat{x}| < \varepsilon\}$[1] точки $\widehat{x}$ такая, что $f(x) \geq f(\widehat{x})$ для любой точки x из этой окрестности.

Определение 2. Точка $\widehat{x}$ является *точкой локального максимума* функции f (пишем $\widehat{x} \in \operatorname{locmax} f$), если существует окрестность $U_\varepsilon = \{x \mid |x - \widehat{x}| < \varepsilon\}$ точки $\widehat{x}$ такая, что $f(x) \leq f(\widehat{x})$ для любой точки x из этой окрестности.

Если точка $\widehat{x}$ является точкой локального минимума или максимума функции f, то пишем $\widehat{x} \in \operatorname{locextr} f$.

1.2 Необходимые и достаточные условия экстремума

1.2.1 Функции одной переменной

Теорема 1.1 (Ферма)**.** *Пусть $f\colon \mathbb{R} \to \mathbb{R}$ — функция одного действительного переменного. Если $\widehat{x} \in \operatorname{locextr} f$ — точка локального экстремума функции f и функция $f \in D(\widehat{x})$ дифференцируема в точке $\widehat{x}$, то*

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

Это соотношение мы называем *условием стационарности* или необходимым условием экстремума 1-го порядка. Точки, удовлетворяющие условию стационарности, называются *стационарными.*

Теорема 1.2. `Необходимые условия экстремума:` *если $\widehat{x} \in \operatorname{locmin} f$ — точка локального минимума функции f и f дважды дифференцируема в точке $\widehat{x}$ $\big(f \in D^2(\widehat{x})\big)$, то*

$$f'(\widehat{x}) = 0, \qquad f''(\widehat{x}) \geq 0.$$

`Достаточные условия экстремума:` *если*

$$f'(\widehat{x}) = 0, \qquad f''(\widehat{x}) > 0,$$

[1] Напомним, что $|x| := \sqrt{\sum\limits_{i=1}^{n} |x_i|^2}$.

то $\widehat{x} \in \operatorname{locmin} f$ *— точка локального минимума функции* f.

Для локального максимума неравенства имеют противоположный вид: $f''(\widehat{x}) \le 0$ и $f''(\widehat{x}) < 0$ соответственно.

Теоремы 1 и 2 допускают обобщения на конечномерный и бесконечномерный случаи (см. соответственно Глава I, п. 1.2.2 и § 6).

В одномерном случае можно дать почти исчерпывающий ответ на вопрос о том, является ли данная точка $\widehat{x}$ локальным экстремумом или нет.

Теорема 1.3. `Необходимые условия минимума:` *если* $\widehat{x} \in \operatorname{locmin} f$ *— точка локального минимума функции* f *и функция* f n *раз дифференцируема в точке* $\widehat{x}$, *то либо* $f'(\widehat{x}) = \ldots = f^{(n)}(\widehat{x}) = 0$, *либо*

$$f'(\widehat{x}) = \ldots = f^{(2m-1)}(\widehat{x}) = 0, \quad f^{(2m)}(\widehat{x}) > 0 \tag{1}$$

при некотором $m:\ 2 \le 2m \le n$.
`Достаточные условия минимума:` *если выполняется условие* (1), *то* $\widehat{x} \in \operatorname{locmin} f$ *— точка локального минимума функции* f.

Доказательство. По формуле Тейлора

$$f(\widehat{x}+h) = \sum_{k=0}^{n} \frac{f^{(k)}(\widehat{x})}{k!} h^k + r(h),\ r(h) = o(h^n).$$

Необходимость при $n = 1$ следует из теоремы Ферма. Пусть далее $n > 1$. Тогда либо $f'(\widehat{x}) = \ldots = f^{(n)}(\widehat{x}) = 0$ (в этом случае выполняется утверждение теоремы), либо $f'(\widehat{x}) = \ldots = f^{(l-1)}(\widehat{x}) = 0$, $f^{(l)}(\widehat{x}) \ne 0$, $l \le n$. Значит,

$$f(\widehat{x}+h) - f(\widehat{x}) = \sum_{k=l}^{n} \frac{f^{(k)}(\widehat{x})}{k!} h^k + r(h) = \frac{f^{(l)}(\widehat{x})}{l!} h^l + r_1(h),$$

$r_1(h) = o(h^l)$. Поскольку $\widehat{x} \in \operatorname{locmin} f$, то $f(\widehat{x}+h) - f(\widehat{x}) \ge 0$ при достаточно малых h. Поэтому

$$\frac{f^{(l)}(\widehat{x})}{l!} h^l + r_1(h) \ge 0.$$

Так как $f^{(l)}(\widehat{x}) \neq 0$, то отсюда следует, что l — четно и $f^{(l)}(\widehat{x}) > 0$. Действительно, если бы l было нечетно, то слагаемое $\frac{f^{(l)}(\widehat{x})}{l!}h^l$, а вместе с ним и сумма $\frac{f^{(l)}(\widehat{x})}{l!}h^l + r_1(h)$ принимали бы знаки как больше, так и меньше нуля при разных знаках h. А это противоречит условию $\widehat{x} \in \operatorname{locmin} f$.

1.2.2 Функции нескольких переменных

Сформулируем необходимое условие экстремума 1-го порядка в конечномерной задаче без ограничений, являющееся аналогом теоремы Ферма.

Теорема 1.4. *Если $\widehat{x} = (\widehat{x}_1, \ldots, \widehat{x}_n) \in \operatorname{locextr} f$ — точка локального экстремума функции n переменных $f(x_1, \ldots, x_n)$ и функция $f \in D(\widehat{x})$ дифференцируема в точке $\widehat{x}$, то*

$$f'(\widehat{x}) = 0 \quad \left(\Leftrightarrow \frac{\partial f(\widehat{x})}{\partial x_1} = \ldots = \frac{\partial f(\widehat{x})}{\partial x_n} = 0 \right).$$

Сформулируем необходимые и достаточные условия экстремума 2-го порядка в конечномерной задаче без ограничений. Предварительно напомним, что второй производной функции нескольких переменных является симметричная матрица вторых производных

$$A = f''(\widehat{x}) = \left(\frac{\partial^2 f(\widehat{x})}{\partial x_i \partial x_j} \right)_{i,j=1}^n = \left(a_{ij} \right)_{i,j=1}^n,$$

а также приведем определения знакоопределенностей симметричных матриц.

Матрица A называется *неотрицательно определенной* (пишем $A \geq 0$), если $\langle Ah, h \rangle \geq 0 \ \forall\ h \in \mathbb{R}^n$
$\left(\Leftrightarrow \sum\limits_{i,j=1}^n a_{ij} h_i h_j \geq 0 \ \forall\ h = (h_1, \ldots, h_n) \in \mathbb{R}^n \right)$.

Матрица A называется *положительно определенной* (пишем $A > 0$), если $\langle Ah, h \rangle > 0 \quad \forall\ h \in \mathbb{R}^n,\ h \neq 0$.

Матрица A называется *строго положительно определенной*, если существует положительное число α такое, что

$$\langle Ah, h\rangle \geq \alpha |h|^2 \quad \forall\, h \in \mathbb{R}^n.$$

Аналогично определяются неположительная, отрицательная и строго отрицательная матрицы.

Отметим, что в конечномерном пространстве условие положительной определенности симметричной матрицы A эквивалентно условию строгой положительности матрицы A.

$$\langle Ah, h\rangle > 0 \,\forall\, h \in \mathbb{R}^n,\ h \neq 0 \iff \exists\, \alpha > 0 \mid \langle Ah, h\rangle \geq \alpha |h|^2 \,\forall\, h \in \mathbb{R}^n.$$

В бесконечномерных пространствах это не так (см. Гл. 1, п.6.3, Пример 2), то есть условие положительной определенности не является условием строгой положительности.

Теорема 1.5. `Необходимые условия минимума:` *если* $\widehat{x} \in \operatorname{locmin} f$ *— точка локального минимума функции* f *и функция* f *дважды дифференцируема в точке* $\widehat{x} = (\widehat{x}_1, \ldots, \widehat{x}_n)$ $\big(f \in D^2(\widehat{x})\big)$, *то*

$$f'(\widehat{x}) = 0,\ f''(\widehat{x}) \geq 0.$$

`Достаточные условия минимума:` *если*

$$f'(\widehat{x}) = 0,\ f''(\widehat{x}) > 0,$$

то $\widehat{x} \in \operatorname{locmin} f$ *— точка локального минимума функции* f.

Для локального максимума неравенства имеют противоположный вид: $f''(\widehat{x}) \leq 0$ и $f''(\widehat{x}) < 0$ соответственно.

1.2.3 Теорема Вейерштрасса и следствие из нее

Пусть $f\colon \mathbb{R}^n \to \mathbb{R}$ — функция n переменных. При исследовании вопроса о достижении функцией экстремума часто используется следующая теорема.

Теорема (Вейерштрасс[2]). *Непрерывная функция на непустом ограниченном замкнутом подмножестве конечномерного пространства (компакте) достигает своих абсолютных максимума и минимума.*

Выделим простое следствие из этой теоремы, которое часто будем использовать.

Следствие. *Если функция f непрерывна на $\mathbb{R}^n$ и $\lim\limits_{|x|\to\infty} f(x) = +\infty$, то она достигает своего абсолютного минимума на любом замкнутом подмножестве из $\mathbb{R}^n$ (не обязательно неограниченном).*

Аналогичное утверждение для абсолютного максимума выполняется, если $\lim\limits_{|x|\to\infty} f(x) = -\infty$.

1.2.4 Критерий Сильвестра

Знакоопределенность матрицы устанавливается с помощью критерия Сильвестра[3].

Напомним, что *последовательными главными минорами* (угловыми) матрицы A называются определители, составленные из первых k строк и k столбцов с номерами $1, 2, \ldots, k$:

$$A_{1\ldots k} := \det \begin{pmatrix} a_{11} & \ldots & a_{1k} \\ . & \ldots & . \\ a_{k1} & \ldots & a_{kk} \end{pmatrix}, \; k = 1, \ldots, n.$$

Главным минором $A_{i_1 \ldots i_k}$ матрицы A называется определитель матрицы размера $k \times k$, составленной из строк и столб-

[2]К. Вейерштрасс (1815–1897) — немецкий математик.

[3]Сильвестр, Джеймс Джозеф (1814-1897) — английский математик.

цов с номерами $i_1, \ldots, i_k$, $1 \le i_1 \le \ldots \le i_k \le n$:

$$A_{i_1 \ldots i_k} := \det \begin{pmatrix} a_{i_1 i_1} & \ldots & a_{i_1 i_k} \\ . & \ldots & . \\ a_{i_k i_1} & \ldots & a_{i_k i_k} \end{pmatrix}, \ k = 1, \ldots, n \ .$$

Теорема. *Пусть A — симметричная матрица. Тогда*

1. *Матрица A положительно определена ($A > 0$) тогда и только тогда, когда все её последовательные главные миноры положительны, т. е. $A_{1 \ldots k} > 0$, $k = 1, \ldots, n$.*

2. *Матрица A отрицательно определена ($A < 0$) тогда и только тогда, когда все её последовательные главные миноры чередуют знак, начиная с отрицательного, т. е. $(-1)^k A_{1 \ldots k} > 0$, $k = 1, \ldots, n$.*

3. *Матрица A неотрицательно определена ($A \ge 0$) тогда и только тогда, когда все её главные миноры неотрицательны, т. е. $A_{i_1 \ldots i_k} \ge 0$, $1 \le i_1 \le \ldots \le i_k \le n$, $k = 1, \ldots, n$.*

4. *Матрица A неположительно определена ($A \le 0$) тогда и только тогда, когда все её главные миноры чередуют знак, начиная с неположительного, т. е. $(-1)^k A_{i_1 \ldots i_k} \ge 0$, $1 \le i_1 \le \ldots \le i_k \le n$, $k = 1, \ldots, n$.*

1.3 Примеры

Пример 1. $f(x) = f(x_1, x_2) = x_1^2 - x_1x_2 + x_2^2 - 2x_1 + x_2 \to \text{extr}$.

Решение. Необходимое условие 1-го порядка экстремума:

$$f'(x) = 0 \iff \frac{\partial f(x)}{\partial x_1} = \frac{\partial f(x)}{\partial x_2} = 0 \iff \begin{cases} 2x_1 - x_2 - 2 = 0, \\ -x_1 + 2x_2 + 1 = 0. \end{cases}$$

Решая эту систему уравнений, находим единственную стационарную точку $\widehat{x} = (\widehat{x}_1, \widehat{x}_2) = (1, 0)$. Для проверки условий 2-го порядка выпишем матрицу вторых производных и проверим её знакоопределенность:

$$A = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right)_{i,j=1}^{2} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix},$$

$$A_1 = 2 > 0, \ A_{12} = \det \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = 3 > 0.$$

Матрица A по критерию Сильвестра положительно определена. Значит, функция f является выпуклой.

Для выпуклой функции точка $\widehat{x} \in \operatorname{absmin} f$, $S_{\text{absmin}} = f(1, 0) = -1$.

Покажем, что $S_{\text{absmax}} = +\infty$. Действительно, для $x_n = (n, 0)$ $f(x_n) = n^2 - 2n \to +\infty$ при $n \to \infty$.

Ответ. $(1, 0) \in \text{absmin}$, $S_{\text{absmin}} = -1$; $S_{\text{absmax}} = +\infty$.

Пример 2. $f(x) = f(x_1, x_2) = x_1^4 + x_2^4 - x_1^2 - 2x_1x_2 - x_2^2 \to \text{extr}$.

Решение. Необходимое условие 1-го порядка экстремума:

$$f'(\widehat{x}) = 0 \iff \frac{\partial f(x)}{\partial x_1} = \frac{\partial f(x)}{\partial x_2} = 0 \iff \begin{cases} 4x_1^3 - 2x_1 - 2x_2 = 0, \\ 4x_2^3 - 2x_1 - 2x_2 = 0. \end{cases}$$

Вычитая из первого уравнения системы второе, получим $4x_1^3 - 4x_2^3 = 0 \Leftrightarrow x_1^3 = x_2^3 \Leftrightarrow x_1 = x_2$. Подставляя $x_2 = x_1$ в первое уравнение системы, находим x_1 и x_2: $4x_1^3 - 2x_1 - 2x_1 = 0$ $\Leftrightarrow x_1^3 = x_1 \Leftrightarrow x_1 = -1; 0; 1$, соответственно $x_2 = -1; 0; 1$. Получили три стационарные точки $\widehat{x}^1 = (1, 1)$, $\widehat{x}^2 = (-1, -1)$, $\widehat{x}^3 = (0, 0)$.

Для проверки условий 2-го порядка выписываем матрицу вторых производных в каждой точке стационарности:

$$A = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right)_{i,j=1}^{2} = \begin{pmatrix} 12x_1^2 - 2 & -2 \\ -2 & 12x_2^2 - 2 \end{pmatrix}.$$

Для точек $(1,1)$ и $(-1,-1)$ матрица $A\big|_{(1,1)} = A\big|_{(-1,-1)} = \begin{pmatrix} 10 & -2 \\ -2 & 10 \end{pmatrix}$ по критерию Сильвестра положительно определена: $A_1 = 10 > 0$, $A_{12} = \det\begin{pmatrix} 10 & -2 \\ -2 & 10 \end{pmatrix} = 100 - 4 = 96 > 0$. По достаточному условию локального экстремума функции нескольких переменных эти точки доставляют локальный минимум функции f. Поскольку

$$\lim_{|x|\to\infty} f(x_1, x_2) = \lim_{|x|\to\infty} \left(x_1^4 + x_2^4 - (x_1 + x_2)^2\right) = +\infty,$$

то, во первых, $S_{\text{absmax}} = +\infty$, во вторых, по следствию из теоремы Вейерштрасса абсолютный минимум в задаче должен достигаться. Значит, эти точки будут доставлять не только локальный, но и абсолютный минимум,

$$S_{\text{absmin}} = f(1,1) = 1 + 1 - 1 - 2 - 1 = -2 = f(-1,-1).$$

Для точки $(0,0)$ матрица $A\big|_{(0,0)} = \begin{pmatrix} -2 & -2 \\ -2 & -2 \end{pmatrix}$ по критерию Сильвестра не является ни положительно, ни отрицательно определенной: $A_1 = -2 < 0$, $A_{12} = \det\begin{pmatrix} -2 & -2 \\ -2 & -2 \end{pmatrix} = 0$. Она является неположительно определенной матрицей ($A \leq 0$) и не является неотрицательно определенной матрицей ($A \not\geq 0$). Следовательно, не выполняется необходимое условие локального минимума. Поэтому $\widehat{x}^3 = (0,0) \notin \operatorname{locmin} f$.

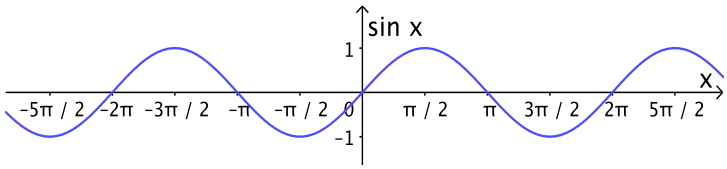
Проверим, доставляет ли точка локальный максимум. Поскольку $f(h,-h) = 2h^4 > 0 = f(\widehat{x}^3)$ при малых $h \neq 0$, то $\widehat{x}^3 \notin \operatorname{locmax} f$. Таким образом, точка $(0,0) \notin \operatorname{locextr} f$.

Ответ. $(0,0) \notin \operatorname{locextr}$; $(-1,-1), (1,1) \in \operatorname{absmin}$, $S_{\text{absmin}} = -2$; $S_{\text{absmax}} = +\infty$.

Приведем несколько примеров различных свойств экстремумов непрерывных функций.

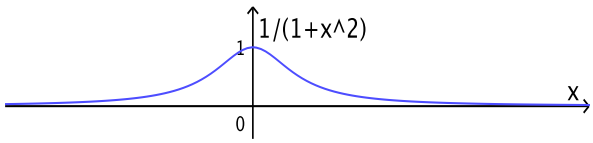
Пример 3. *Абсолютные минимумы и максимумы достигаются в бесконечном числе точек:*

$$f\colon \mathbb{R} \to \mathbb{R},\ \ f(x) = \sin x.$$



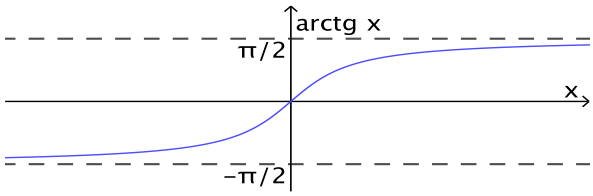
Пример 4. *Функция ограничена, абсолютный максимум достигается, минимум — нет:*

$$f\colon \mathbb{R} \to \mathbb{R},\ \ f(x) = \frac{1}{1+x^2}.$$



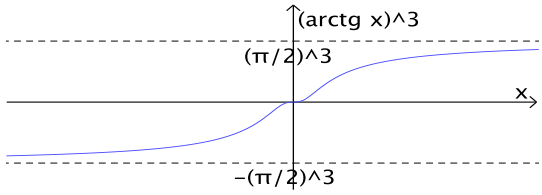
Пример 5. *Функция ограничена, но абсолютные максимум и минимум не достигаются:*

$$f\colon \mathbb{R} \to \mathbb{R},\ \ f(x) = \operatorname{arctg} x.$$



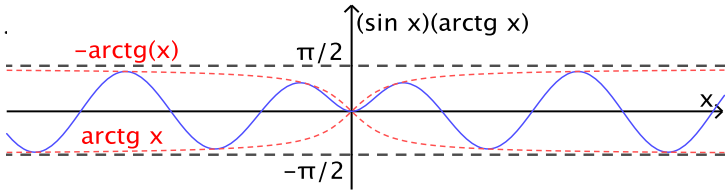
Пример 6. *Функция ограничена, имеет стационарные точки, но абсолютные максимум и минимум не достигаются:*

$$f\colon \mathbb{R} \to \mathbb{R},\ \ f(x) = (\operatorname{arctg} x)^3.$$



Пример 7. *Функция ограничена, имеет локальные максимумы и минимумы, но абсолютные максимум и минимум не достигаются:*

$$f\colon \mathbb{R} \to \mathbb{R},\ \ f(x) = \operatorname{arctg} x \cdot \sin x.$$



2 Конечномерные гладкие задачи с равенствами

2.1 Постановка задачи

Пусть $f_i\colon \mathbb{R}^n \to \mathbb{R}$, $i = 0, 1, \ldots, m$, — функции n переменных, отображающие пространство $\mathbb{R}^n$ в $\mathbb{R}$. Считаем, что все функции f_i обладают определенной гладкостью. *Гладкой конечномерной экстремальной задачей с ограничениями типа равенств* называется следующая задача:

$$f_0(x) \to \operatorname{extr}; \quad f_i(x) = 0,\ i = 1, \ldots, m. \tag{P}$$

Таким образом, в задаче (P) ищутся экстремумы функции на множестве, задаваемом конечным числом ограничений типа равенств. Точки, удовлетворяющие всем ограничениям типа равенств, называются *допустимыми точками* в задаче.

Определение 1. Говорим, что допустимая точка $\widehat{x}$ доставляет *локальный минимум* в задаче (P), и пишем $\widehat{x} \in \operatorname{locmin} P$, если существует $\delta > 0$ такое, что $f_0(x) \geq f_0(\widehat{x})$ для любой допустимой точки x, для которой[4] $|x - \widehat{x}| < \delta$.

Определение 2. Говорим, что допустимая точка $\widehat{x}$ доставляет *локальный максимум* в задаче (P), и пишем $\widehat{x} \in \operatorname{locmax} P$, если существует $\delta > 0$ такое, что $f_0(x) \leq f_0(\widehat{x})$ для любой допустимой точки x, для которой $|x - \widehat{x}| < \delta$.

Задачу на максимум всегда можно свести к задаче на минимум, взяв вместо функции f_0 функцию $\widetilde{f}_0 = -f_0$. Возможно и обратное сведение.

[4] Напомним, что $|x| := \sqrt{\sum\limits_{i=1}^{n} |x_i|^2}$.

2.2 Необходимые и достаточные условия экстремума

2.2.1 Принцип Лагранжа

Для решения задачи с ограничениями типа равенств

$$f_0(x) \to \text{extr}; \quad f_i(x) = 0,\ i = 1, \ldots, m. \qquad (P)$$

составляется функция Лагранжа[5]. Функция, экстремум которой ищется, умножается на множитель λ_0, функции ограничений $f_i(x) = 0$ умножаются на множители λ_i, складываются и выписывается *функция Лагранжа* $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$. Вектор λ называется *вектором множителей Лагранжа.* Необходимые условия экстремума в задаче с ограничениями выписываются через необходимые условия экстремума функции Лагранжа.

Теорема 2.1. *Пусть* $\widehat{x} \in \operatorname{locextr} P$ *— точка локального экстремума, а функции* f_i, $i = 0, 1, \ldots, m$, *непрерывно дифференцируемы в окрестности точки* $\widehat{x}$ (условие гладкости).

Тогда существует ненулевой вектор множителей Лагранжа $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$, $\lambda \neq 0$, *такой, что для функции Лагранжа* $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ *выполняется условие стационарности*[6]

$$\mathscr{L}'(\widehat{x}) = 0.$$

Точки, удовлетворяющие условию стационарности, называются *стационарными.* Это точки возможного локального экстремума.

---

[5] Ж. Л. Лагранж (1736–1813) — французский математик

[6] Условие стационарности можно расписать в эквивалентном виде: $\dfrac{\partial \mathscr{L}(\widehat{x})}{\partial x_j} = 0,\ j = 1, \ldots, n$ или $\sum\limits_{i=0}^{m} \lambda_i f_i'(\widehat{x}) = 0 \Leftrightarrow \sum\limits_{i=0}^{m} \lambda_i \langle f_i'(\widehat{x}), h\rangle = 0\, \forall\, h \in \mathbb{R}^n$.

Важную роль при исследовании полученных стационарных точек играет *пространство стационарных направлений*

$$\Pi(\widehat{x}) := \{h \in \mathbb{R}^n \mid \langle f_i'(\widehat{x}), h\rangle = 0,\ i = 1, \ldots, m\}.$$

Направления h из пространства стационарных направлений называются *стационарными*. Двигаясь не по стационарному направлению из точки $\widehat{x}$, мы сразу выходим из множества допустимых элементов.

Отметим, что в стационарной точке $\widehat{x}$ при $\lambda_0 \neq 0$ в силу условия стационарности

$$\lambda_0 \langle f_0'(\widehat{x}), h\rangle + \sum_{i=1}^{m} \lambda_i \langle f_i'(\widehat{x}), h\rangle = 0$$

выполняется также равенство $\langle f_0'(\widehat{x}), h\rangle = 0\ \forall\ h \in \Pi(\widehat{x})$.

2.2.2 Конечномерная теорема об обратной функции

Теорема (конечномерная теорема об обратной функции)**.** *Пусть $F\colon \mathbb{R}^n \to \mathbb{R}^n$ — непрерывно дифференцируемое отображение в некоторой окрестности $U \subset \mathbb{R}^n$ точки $\widehat{x}$, $F(\widehat{x}) = \widehat{y}$ и якобиан отображения F в точке $\widehat{x}$ отличен от нуля $\left(\det F'(\widehat{x}) = \det\left(\dfrac{dF_i(\widehat{x})}{dx_j}\right)_{i,j=1}^{n} \neq 0\right)$. Тогда существует обратное отображение F^{-1} некоторой окрестности V точки $\widehat{y}$ в окрестность U точки $\widehat{x}$ такое, что $F^{-1}(\widehat{y}) = \widehat{x}$ и*

$$F(F^{-1}(y)) = y, \quad |F^{-1}(y) - F^{-1}(\widehat{y})| \leq K|y - \widehat{y}| \quad \forall\ y \in V$$

с некоторой константой $K > 0$.

2.2.3 Необходимое условие экстремума 2-го порядка

Сформулируем необходимое условие 2-го порядка экстремума в гладкой конечномерной задаче с ограничениями типа равенств

$$f_0(x) \to \mathrm{extr}; \quad f_i(x) = 0,\ i = 1, \ldots, m. \qquad (P)$$

Здесь функционалы $f_i \colon \mathbb{R}^n \to \mathbb{R}$, $i = 0, 1, \ldots, m$.

Теорема 2.2. *Пусть* $\widehat{x} \in \operatorname{locextr} P$ *— точка локального экстремума, функции* $f_i \in D^2(\widehat{x})$, $i = 0, 1, \ldots, m$, *дважды дифференцируемы в точке* $\widehat{x}$ (условие гладкости), *векторы* $f_1'(\widehat{x}), \ldots, f_m'(\widehat{x})$ *линейно независимы* (условие регулярности).

Тогда существует вектор множителей Лагранжа $\lambda = (1, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$ *такой, что для функции Лагранжа* $\mathscr{L}(x) = f_0(x) + \sum\limits_{i=1}^{m} \lambda_i f_i(x)$ *выполняются условия стационарности* $\mathscr{L}'(\widehat{x}) = 0$ *и на пространстве стационарных направлений* $\Pi := \{h \in \mathbb{R}^n \mid \langle f_i'(\widehat{x}), h\rangle = 0,\ i = 1, \ldots, m\}$:

$$\langle \mathscr{L}''(\widehat{x})h, h\rangle \geq 0 \ \forall\ h \in \Pi \ \textit{для минимума},$$
$$\langle \mathscr{L}''(\widehat{x})h, h\rangle \leq 0 \ \forall\ h \in \Pi \ \textit{для максимума}.$$

Доказательство. Проведем доказательство для случая минимума. Случай максимума рассматривается аналогично.

Пусть $\widehat{x} \in \operatorname{locmin} P$. Тогда для гладкой конечномерной задачи с равенствами выполняется принцип Лагранжа. Согласно нему существует ненулевой вектор множителей Лагранжа $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$, такой, что для функции Лагранжа $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ выполняется условие стационарности $\sum\limits_{i=0}^{m} \lambda_i f_i'(\widehat{x}) = 0$. В силу условия регулярности векторы $f_1'(\widehat{x}), \ldots, f_m'(\widehat{x})$ линейно независимы. Поэтому $\lambda_0 \neq 0$ и, следовательно, можно положить $\lambda_0 = 1$. Тем самым условие стационарности с множителем Лагранжа $\lambda_0 = 1$ доказано.

Докажем выполнение неравенства. Возьмём произвольный вектор $h \in \Pi$. Для отображения $F(x) = (f_1(x), \ldots, f_m(x))$ и множества $M := \{x \mid F(x) = 0\}$ выполнены все условия теоремы о касательном пространстве (строгая дифференцируемость отображения F в точке $\widehat{x}$ и регулярность оператора $F'(\widehat{x})$). По теореме о касательном пространстве (см. § 5.4) $T_{\widehat{x}}M = \operatorname{Ker} F'(\widehat{x})$. Значит, вектор h будет касательным к множеству M в точке $\widehat{x}$.

По определению касательного вектора существуют число $\varepsilon > 0$ и отображение $r\colon [-\varepsilon; \varepsilon] \to \mathbb{R}^n$ такие, что точки $x_t := \widehat{x} + th + r(t) \in M \ \forall\ t \in [-\varepsilon; \varepsilon]$ и $|r(t)| = o(t)$ при $t \to 0$. Таким образом, точки x_t являются допустимыми в задаче и находятся в окрестности точки $\widehat{x}$.

Поскольку $f_i(x_t) = 0$ и $f_i(\widehat{x}) = 0$, $i = 1, \ldots, m$, то

$$f_0(x_t) = f_0(x_t) + \sum_{i=1}^{m} \lambda_i f_i(x_t) \overset{\text{def}\,\mathscr{L}}{=} \mathscr{L}(x_t),$$

$$f_0(\widehat{x}) = f_0(\widehat{x}) + \sum_{i=1}^{m} \lambda_i f_i(\widehat{x}) \overset{\text{def}\,\mathscr{L}}{=} \mathscr{L}(\widehat{x}).$$

Поэтому

$$f_0(x_t) - f_0(\widehat{x}) = \mathscr{L}(x_t) - \mathscr{L}(\widehat{x}) = \mathscr{L}(\widehat{x} + th + r(t)) - \mathscr{L}(\widehat{x}) =$$

(по формуле Тейлора)

$$= \langle \mathscr{L}'(\widehat{x}), th + r(t)\rangle + \frac{1}{2}\langle \mathscr{L}''(\widehat{x})(th + r(t)), th + r(t)\rangle + o(|th + r(t)|^2) =$$

$$\overset{\mathscr{L}'(\widehat{x})=0}{=} \frac{t^2}{2}\langle \mathscr{L}''(\widehat{x})h, h\rangle + o(t^2) \geq 0, \text{ так как } \widehat{x} \in \operatorname{locmin} P.$$

Деля это неравенство на t^2 и переходя к пределу при $t \to 0$, получаем $\langle \mathscr{L}''(\widehat{x})h, h\rangle \geq 0$. □

Для вектора $x = (x_1, x_2)$

$$\langle \mathscr{L}''h, h\rangle = \mathscr{L}_{x_1x_1}h_1^2 + 2\mathscr{L}_{x_1x_2}h_1h_2 + \mathscr{L}_{x_2x_2}h_2^2.$$

Для функции 3-х переменных

$$\langle \mathscr{L}''h, h\rangle = \mathscr{L}_{x_1x_1}h_1^2 + \mathscr{L}_{x_2x_2}h_2^2 + \mathscr{L}_{x_3x_3}h_3^2 + \\ +2(\mathscr{L}_{x_1x_2}h_1h_2 + \mathscr{L}_{x_1x_3}h_1h_3 + \mathscr{L}_{x_2x_3}h_2h_3).$$

2.2.4 Достаточное условие экстремума 2-го порядка

Сформулируем достаточное условие второго порядка экстремума в гладкой конечномерной задаче с ограничениями типа равенств

$$f_0(x) \to \text{extr}; \quad f_i(x) = 0,\ i = 1, \dots, m. \qquad (P)$$

Здесь функционалы $f_i \colon \mathbb{R}^n \to \mathbb{R}$, $i = 0, 1, \dots, m$.

Теорема 2.3. *Пусть в допустимой точке $\widehat{x}$ функции $f_i \in D^2(\widehat{x})$, $i = 0, 1, \dots, m$, дважды дифференцируемы, для функции Лагранжа $\mathscr{L}(x) = f_0(x) + \sum\limits_{i=1}^{m} \lambda_i f_i(x)$ с вектором множителей Лагранжа $\lambda = (1, \lambda_1, \dots, \lambda_m)$ выполняется условие стационарности $\mathscr{L}'(\widehat{x}) = 0$.*

Тогда, если матрица вторых производных положительно определена на пространстве стационарных направлений[7]:

$$\langle \mathscr{L}''(\widehat{x})h, h\rangle > 0 \quad \forall\ h \in \Pi,\ h \neq 0, \qquad (*)$$

то $\widehat{x} \in \operatorname{locmin} P$ — точка локального минимума.

Если матрица вторых производных отрицательно определена на пространстве стационарных направлений:

$$\langle \mathscr{L}''(\widehat{x})h, h\rangle < 0 \quad \forall\ h \in \Pi,\ h \neq 0,$$

то $\widehat{x} \in \operatorname{locmax} P$ — точка локального максимума.

Доказательство. Проведем доказательство для случая минимума. Случай максимума рассматривается аналогично.

Предположим противное, что точка $\widehat{x} \notin \operatorname{locmin} P$. Тогда существует последовательность допустимых[8] точек $x_n \to \widehat{x}$ таких, что $f_0(x_n) < f_0(\widehat{x})$. Обозначим $h_n := x_n - \widehat{x}$, $\overline{h}_n := \dfrac{h_n}{|h_n|}$.

---

[7] Пространство стационарных направлений
$\Pi = \Pi(\widehat{x}) := \{h \in \mathbb{R}^n \mid \langle f_i'(\widehat{x}), h\rangle = 0,\ i = 1, \dots, m\}$.

[8] $f_i(x_n) = f_i(\widehat{x}) = 0$, $i = 1, \dots, m$.

Последовательность точек $\overline{h}_n$ принадлежит единичной сфере в $\mathbb{R}^n$ (компакту). Поэтому из нее можно выделить подпоследовательность, сходящуюся к некоторой точке $\overline{h}$, принадлежащей этой единичной сфере. Поэтому $|\overline{h}| = 1$. Для простоты записи будем считать, что сама последовательность $\overline{h}_n$ сходится к $\overline{h}$.

Покажем, что $\overline{h} \in \Pi$. Имеем для $i = 1, \ldots, m$

$$f_i(x_n) = f_i(\widehat{x} + h_n) = f_i(\widehat{x}) + \langle f_i'(\widehat{x}), h_n\rangle + o(h_n)$$
$$\overset{f_i(x_n)=f_i(\widehat{x})=0}{\Longleftrightarrow} \langle f_i'(\widehat{x}), h_n\rangle + o(h_n) = 0.$$

Деля это равенство на $|h_n|$ и переходя к пределу, получаем:

$$\left\langle f_i'(\widehat{x}), \frac{h_n}{|h_n|}\right\rangle + \frac{o(h_n)}{|h_n|} = 0 \Longrightarrow \langle f_i'(\widehat{x}), \overline{h}\rangle = 0.$$

Значит, $\overline{h} \in \Pi$, $\overline{h} \neq 0$ (так как $|\overline{h}| = 1$).

Поскольку $f_i(x_n) = f_i(\widehat{x}) = 0$, $i = 1, \ldots, m$, то

$$f_0(x_n) = f_0(x_n) + \sum_{i=1}^{m} \lambda_i f_i(x_n) = \mathscr{L}(x_n),$$
$$f_0(\widehat{x}) = f_0(\widehat{x}) + \sum_{i=1}^{m} \lambda_i f_i(\widehat{x}) = \mathscr{L}(\widehat{x}).$$

Поэтому

$$f_0(x_n) - f_0(\widehat{x}) = \mathscr{L}(x_n) - \mathscr{L}(\widehat{x}) = \mathscr{L}(\widehat{x} + h_n) - \mathscr{L}(\widehat{x}) =$$

(по формуле Тейлора)

$$= \langle \mathscr{L}'(\widehat{x}), \overline{h}\rangle + \frac{1}{2}\langle \mathscr{L}''(\widehat{x})h_n, h_n\rangle + o(|h_n|^2) \overset{\mathscr{L}'(\widehat{x})=0}{=}$$
$$= \frac{1}{2}\langle \mathscr{L}''(\widehat{x})h_n, h_n\rangle + o(|h_n|^2) < 0,$$

так как по предположению $f_0(x_n) - f_0(\widehat{x}) < 0$. Разделим полученное неравенство на $|h_n|^2$: $\left\langle \mathscr{L}''(\widehat{x})\dfrac{h_n}{|h_n|}, \dfrac{h_n}{|h_n|}\right\rangle + \dfrac{o(|h_n|^2)}{|h_n|^2} < 0$. Переходя к пределу при $n \to +\infty$, получаем: $\langle \mathscr{L}''(\widehat{x})\overline{h}, \overline{h}\rangle \leq 0$. Пришли к противоречию с условием $(*)$ теоремы. Значит, наше предположение противного неверно и $\widehat{x} \in \operatorname{locmin} P$. $\square$

2.3 Примеры

Пример 1. $x_1 + 2x_2 \to \text{extr}; \quad x_1^2 + x_2^2 = 5.$

Решение. Функция Лагранжа $\mathscr{L} = \lambda_0(x_1+2x_2)+\lambda_1(x_1^2+x_2^2-5)$.

Необходимое условие локального экстремума:

$$\mathscr{L}_x = 0 \iff \begin{cases} \mathscr{L}_{x_1} = 0, \\ \mathscr{L}_{x_2} = 0, \end{cases} \iff \begin{cases} \lambda_0 + 2\lambda_1 x_1 = 0, \\ 2\lambda_0 + 2\lambda_1 x_2 = 0. \end{cases}$$

Если $\lambda_0 = 0$, то $\lambda_1 \neq 0$, так как не все множители Лагранжа равны нулю. Тогда из предыдущих уравнений вытекает, что $x_1 = x_2 = 0$, но эта точка не является допустимой.
Значит, в случае $\lambda_0 = 0$ стационарных точек нет.

Полагаем $\lambda_0 = 1$. Тогда $x_1 = -\dfrac{1}{2\lambda_1}$, $x_2 = -\dfrac{1}{\lambda_1}$.
Подставляя x_1, x_2 в ограничение $x_1^2 + x_2^2 = 5$, получаем:

$$\frac{1}{4\lambda_1^2} + \frac{1}{\lambda_1^2} = 5 \iff 5 = 20\lambda_1^2 \iff 1 = 4\lambda_1^2 \iff \lambda_1 = \pm\frac{1}{2}.$$

Поэтому имеем две стационарные точки:

$\lambda_1 = \dfrac{1}{2} \Rightarrow x_1 = -\dfrac{1}{2\lambda_1} = -1$, $x_2 = -\dfrac{1}{\lambda_1} = -2$, $\widehat{x}^1 = (-1, -2)$;

$\lambda_1 = -\dfrac{1}{2} \Rightarrow x_1 = -\dfrac{1}{2\lambda_1} = 1$, $x_2 = -\dfrac{1}{\lambda_1} = 2$, $\widehat{x}^2 = (1, 2)$.

По теореме Вейерштрасса (в силу компактности окружности) существуют решения задач на максимум и минимум. Значит, одна из найденных точек доставляет абсолютный минимум, другая — абсолютный максимум.
Вычислим значения функционала в стационарных точках:

$f_0(-1, -2) = -1 + 2(-2) = -5,$

$f_0(1, 2) = 1 + 2 \cdot 2 = 5.$

Значит, $S_{\text{absmin}} = -5$, $S_{\text{absmax}} = 5$.

Ответ. $(-1, -2) \in \text{absmin}$, $S_{\text{absmin}} = -5$; $(1, 2) \in \text{absmax}$, $S_{\text{absmax}} = 5$.

3 Конечномерные гладкие задачи с равенствами и неравенствами

В этом параграфе даются необходимые условия экстремума первого порядка в гладкой конечномерной задаче с ограничениями типа равенств и неравенств, необходимые и достаточные условия экстремума второго порядка, а также условия экстремума в квадратично-линейной задаче с линейными ограничениями.

3.1 Постановка задачи

Пусть $f_i\colon \mathbb{R}^n \to \mathbb{R}$, $i = 0, 1, \ldots, m$, — функции n переменных, отображающие пространство $\mathbb{R}^n$ в $\mathbb{R}$, обладающие некоторой гладкостью. *Гладкой конечномерной экстремальной задачей с ограничениями типа равенств и неравенств* называется следующая задача в $\mathbb{R}^n$:

$$f_0(x) \to \mathrm{extr}; \qquad \begin{array}{l} f_i(x) \le 0,\ i = 1, \ldots, m', \\ f_i(x) = 0,\ i = m'+1, \ldots, m. \end{array} \qquad (P)$$

Точки, удовлетворяющие всем ограничениям типа равенств и неравенств, называются *допустимыми точками* в задаче.

Определение 1. Говорим, что допустимая точка $\widehat{x}$ доставляет *локальный минимум* в задаче (P), и пишем $\widehat{x} \in \operatorname{locmin} P$, если существует $\delta > 0$ такое, что $f_0(x) \ge f_0(\widehat{x})$ для любой допустимой точки x, для которой $|x - \widehat{x}| < \delta$.

Определение 2. Говорим, что допустимая точка $\widehat{x}$ доставляет *локальный максимум* в задаче (P), и пишем $\widehat{x} \in \operatorname{locmax} P$, если существует $\delta > 0$ такое, что $f_0(x) \le f_0(\widehat{x})$ для любой допустимой точки x, для которой $|x - \widehat{x}| < \delta$.

Отметим, что в задаче (P) ограничения типа равенств можно было бы не писать, заменив любое из равенств $f(x) = 0$ ($\Leftrightarrow 0 \le f(x) \le 0$) двумя неравенствами $f(x) \le 0$, $-f(x) \le 0$.

3.2 Необходимые и достаточные условия экстремума

3.2.1 Принцип Лагранжа

Сформулируем необходимое условие 1-го порядка экстремума в гладкой конечномерной задаче с ограничениями типа равенств и неравенств — принцип Лагранжа. Имеем задачу

$$f_0(x) \to \text{extr}; \qquad \begin{array}{l} f_i(x) \le 0,\ i = 1, \dots, m', \\ f_i(x) = 0,\ i = m'+1, \dots, m. \end{array} \qquad (P)$$

Теорема 3.1. *Пусть* $\widehat{x} \in \operatorname{locextr} P$ *— точка локального экстремума, а функции* f_i, $i = 0, 1, \dots, m$, *непрерывно дифференцируемы в окрестности точки* $\widehat{x}$ (условие гладкости). *Тогда существует ненулевой вектор множителей Лагранжа* $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m) \in \mathbb{R}^{m+1}$, *такой, что для функции Лагранжа* $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ *выполняются условия*:

a) *стационарности*[9] : $\mathscr{L}'(\widehat{x}) = 0$;

b) *дополняющей нежёсткости*: $\lambda_i f_i(\widehat{x}) = 0,\ i = 1, \dots, m'$;

c) *знакоопределенности*:

$$\lambda_0 \ge 0, \quad \begin{array}{l} \lambda_i \ge 0,\ i = 1, \dots, m',\ \textit{для локального минимума}, \\ \lambda_i \le 0,\ i = 1, \dots, m',\ \textit{для локального максимума}. \end{array}$$

Доказательство. Проведем доказательство для случая минимума. Случай максимума рассматривается аналогично.

Будем считать также, что $f_0(\widehat{x}) = 0$, иначе рассмотрим функцию $\widetilde{f}_0(x) = f_0(x) - f_0(\widehat{x})$.

Если $f_i(\widehat{x}) < 0$ при некотором $1 \le i \le m$, то для гладкого отображения $f_i(x) < 0$ в некоторой окрестности точки $\widehat{x}$. Поэтому это неравенство $f_i(x) \le 0$ на локальный экстремум

[9] Условие стационарности можно расписать в эквивалентном виде: $\dfrac{\partial \mathscr{L}(\widehat{x})}{\partial x_j} = 0,\ j = 1, \dots, n$ или $\sum\limits_{i=0}^{m} \lambda_i f_i'(\widehat{x}) = 0 \Leftrightarrow \sum\limits_{i=0}^{m} \lambda_i \langle f_i'(\widehat{x}), h \rangle = 0 \,\forall\, h \in \mathbb{R}^n$.

не влияет и полагаем $\lambda_i = 0$. Таким образом, можно считать, что условия дополняющей нежёсткости уже выполнены и $f_i(\widehat{x}) = 0$, $i = 0, 1, \ldots, m$.

Для удобства записи введем отображение $F\colon \mathbb{R}^n \to \mathbb{R}^{m-m'}$, действующее по формуле $F(x) = (f_{m'+1}(x), \ldots, f_m(x))$.

A) *Вырожденный случай.* Пусть $\operatorname{Im} F'(\widehat{x}) \neq \mathbb{R}^{m-m'}$. Тогда $\operatorname{Im} F'(\widehat{x})$ есть замкнутое (в конечномерном пространстве все подпространства замкнутые) собственное подпространство в $\mathbb{R}^{m-m'}$. По лемме о нетривиальности аннулятора замкнутого собственного подпространства существует ненулевой функционал $y^* = (\lambda_{m'+1}, \ldots, \lambda_m) \in (\operatorname{Im} F'(\widehat{x}))^{\perp}$. Это означает, что $\langle y^*, y\rangle = 0 \ \forall\ y \in \operatorname{Im} F'(\widehat{x}) \Leftrightarrow \sum\limits_{i=m'+1}^{m} \lambda_i f_i'(\widehat{x}) = 0$. Остается положить $\lambda_i = 0$, $i = 0, 1, \ldots, m'$, и приходим к утверждению теоремы.

B) *Невырожденный случай.* Пусть $\operatorname{Im} F'(\widehat{x}) = \mathbb{R}^{m-m'}$. Рассмотрим множество

$$B = \{b = (b_0, b_1, \ldots, b_m) \in \mathbb{R}^{m+1} \mid \exists\ x \in \mathbb{R}^n :$$
$$b_i > f_i'(\widehat{x})[x],\ i = 0, 1, \ldots, m';\ b_i = f_i'(\widehat{x})[x],\ i = m'+1, \ldots, m\}.$$

Докажем, что B — выпуклое множество.
Действительно, возьмем две точки $b, b' \in B$. Тогда по определению множества B существуют точки x, x' такие, что

$$b_i > f_i'(\widehat{x})[x],\ i = 0, 1, \ldots, m',\ \ b_i = f_i'(\widehat{x})[x],\ i = m'+1, \ldots, m;$$

$$b_i' > f_i'(\widehat{x})[x'],\ i = 0, 1, \ldots, m',\ \ b_i' = f_i'(\widehat{x})[x'],\ i = m'+1, \ldots, m.$$

Покажем, что выпуклая комбинация $b_\theta = \theta b + (1-\theta) b' \in B$ $\forall\ 0 < \theta < 1$. Положим $x_\theta = \theta x + (1-\theta) x'$. Умножая первые два соотношения на θ, а вторые два на $1-\theta$ и складывая, получим соответственно:

$$\theta b_i + (1-\theta) b_i' > \langle f_i'(\widehat{x}), \theta x + (1-\theta) x'\rangle,\ i = 0, 1, \ldots, m';$$

$$\theta b_i + (1-\theta) b_i' = \langle f_i'(\widehat{x}), \theta x + (1-\theta) x'\rangle,\ i = m'+1, \ldots, m.$$

Поэтому $b_\theta \in B$.

Очевидно, что множество $B \neq \emptyset$. Действительно, возьмём точку $x = 0$. Тогда вектор $b \in B$ при $b_i > \langle f_i'(\widehat{x}), 0\rangle = 0$, $i = 0, 1, \ldots, m'$, $b_i = \langle f_i'(\widehat{x}), 0\rangle = 0$, $i = m'+1, \ldots, m$.

Покажем, что $0 \notin B$. Предположим противное и придем к противоречию с тем, что $\widehat{x}$ локальный минимум.

Если $0 \in B$, то существует элемент $h \in \mathbb{R}^n$ такой, что $f_i'(\widehat{x})[h] < 0$, $i = 0, 1, \ldots, m'$; $f_i'(\widehat{x})[h] = 0$, $i = m'+1, \ldots, m$ ($\Leftrightarrow F'(\widehat{x})[h] = 0$). В рассматриваемом нами невырожденном случае ($F'(\widehat{x})(\mathbb{R}^n) = \mathbb{R}^{m-m'}$) выполняются все условия теоремы о касательном пространстве для отображения F. Предположения о функциях f_i, гарантируют строгую дифференцируемость компонент F в точке $\widehat{x}$, а значит, и строгую дифференцируемость самого отображения F в точке $\widehat{x}$. По этой теореме $\operatorname{Ker} F'(\widehat{x}) = T_{\widehat{x}}N$, где $N := \{x \in \mathbb{R}^n \mid F(x) = F(\widehat{x}) = 0\}$. Поэтому $h \in T_{\widehat{x}}N$ и, следовательно, существует отображение $r\colon [-\varepsilon; \varepsilon] \to \mathbb{R}^n$ ($\varepsilon > 0$) такое, что $|r(t)| = o(t)$,

$$\widehat{x} + th + r(t) \in N \iff F\big(\widehat{x} + th + r(t)\big) = 0 \ \forall\, t \in [-\varepsilon, \varepsilon]. \quad (1)$$

При малых $t > 0$ для $i = 0, 1, \ldots, m'$ имеем неравенства

$$f_i\big(\widehat{x} + th + r(t)\big) = f_i(\widehat{x}) + t\langle f_i'(\widehat{x}), h\rangle + o(t) < 0. \qquad (1_i)$$

Из (1), (1_i), $i = 1, \ldots, m'$, следует, что $\widehat{x} + th + r(t) \in D(P)$. Но при этом неравенство (1_0) $\Leftrightarrow f_0\big(\widehat{x} + th + r(t)\big) < 0 = f_0(\widehat{x})$ противоречит тому, что $\widehat{x} \in \operatorname{locmin}$. Значит, $0 \notin B$.

По теореме отделимости выпуклое множество B и точку 0 можно отделить. Значит, существует ненулевой функционал $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$ такой, что

$$\inf_{b \in B} \langle \lambda, b\rangle \geq \langle \lambda, 0\rangle \iff \sum_{i=0}^{m} \lambda_i b_i \geq 0 \ \forall\, b \in B. \qquad (*)$$

Условие неотрицательности. Для любого $\varepsilon > 0$ вектор $(\varepsilon, \ldots, \varepsilon, \underset{i}{1}, \varepsilon, \ldots, \varepsilon, \underbrace{0, \ldots, 0}_{m-m'}) \in B$ (полагаем $x = 0$), $i = 0, \ldots, m'$.

Подставляя такие векторы в неравенство $(*)$, получаем, что

$$\varepsilon \sum_{j=0,\, j\neq i}^{m'} \lambda_j + \lambda_i \geq 0.$$

Устремляя ε к нулю, получаем, что $\lambda_i \geq 0$, $i = 0, 1, \ldots, m'$.

Условие стационарности. Для любых $x \in \mathbb{R}^n$ и $\varepsilon > 0$ вектор $\big(f_0'(\widehat{x})[x] + \varepsilon, \ldots, f_{m'}'(\widehat{x})[x] + \varepsilon, f_{m'+1}'(\widehat{x})[x], \ldots, f_m'(\widehat{x})[x]\big) \in B$. Подставляя такие векторы в $(*)$, приходим к неравенству

$$\varepsilon \sum_{i=0}^{m'} \lambda_i + \sum_{i=0}^{m} \lambda_i f_i'(\widehat{x})[x] \geq 0 \ \forall\ x \in \mathbb{R}^n.$$

В силу произвольности x отсюда следует, что $\sum\limits_{i=0}^{m} \lambda_i f_i'(\widehat{x}) = 0$.

$\square$

Поскольку условия стационарности и дополняющей нежёсткости определены с точностью до умножения на константу, то в задаче на минимум условие знакоопределенности можно записать в виде: $\lambda_i \geq 0$, $i = 0, 1, \ldots, m'$ (это условие называют еще *условием неотрицательности*); в задаче на максимум неравенства можно записать в виде: $\lambda_0 \leq 0$, $\lambda_i \geq 0$, $i = 1, \ldots, m'$.

Точки, удовлетворяющие необходимым условиям локального экстремума a)–c), называются *критическими*. Это точки возможного локального экстремума.

Если для критической точки $\lambda_0 > 0$, $\lambda_i \geq 0$, $i = 1, \ldots, m'$, то точка является *критической в задаче на минимум*.

Если для критической точки $\lambda_0 > 0$, $\lambda_i \leq 0$, $i = 1, \ldots, m'$, то точка является *критической в задаче на максимум*.

Если для критической точки $\lambda_0 > 0$, а $\lambda_i = 0$, $i = 1, \ldots, m'$, то точка является *критической в задаче и на минимум, и на максимум*.

Если для точки множители Лагранжа λ_i, $i = 1, \ldots, m'$, разного знака, то точка не является критической.

Поскольку множители Лагранжа определены с точностью до умножения на положительную константу, то обычно при решении рассматривают два случая: $\lambda_0 = 0$ или $\lambda_0 > 0$ (во втором случае часто полагают $\lambda_0 = 1$).

Если для ограничения типа неравенства в критической точке $f_i(\widehat{x}) < 0$, то $f_i(x) < 0$ в некоторой окрестности точки $\widehat{x}$ и при исследовании локального экстремума это неравенство можно не учитывать. В силу условия дополняющей нежёсткости соответствующий коэффициент $\lambda_i = 0$.

Если для ограничения типа неравенства в критической точке $f_i(\widehat{x}) = 0$, то это ограничение называется *активным*.

3.2.2 Необходимое условие 1-го и 2-го порядка

Рассмотрим задачу с ограничениями типа равенств и неравенств:

$$f_0(x) \to \text{extr}; \qquad \begin{array}{l} f_i(x) \le 0,\ i = 1, \dots, m', \\ f_i(x) = 0,\ i = m'+1, \dots, m, \end{array} \qquad (P)$$

где $f_i \colon \mathbb{R}^n \to \mathbb{R}$, $i = 0, 1, \dots, m$.

Теорема 3.2. *Пусть точка* $\widehat{x} \in \operatorname{locextr} P$, *функции* $f_i \in C^2(\mathcal{O}(\widehat{x}))$, $i = 0, 1, \dots, m$, *векторы* $f'_{m'+1}(\widehat{x}), \dots, f'_m(\widehat{x})$ — *линейно независимы. Тогда существует вектор множителей Лагранжа* $\lambda \in \mathbb{R}^{m+1}$, $\lambda \ne 0$, *такой, что для функции Лагранжа* $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ *выполняются условия*:

a) стационарности: $\mathscr{L}'(\widehat{x}) = 0$;

b) дополняющей нежёсткости: $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \dots, m'$;

c) знакоопределенности множителей Лагранжа:

$$\lambda_0 \ge 0, \quad \begin{array}{l} \lambda_i \ge 0,\ i = 1, \dots, m',\ \textit{для минимума}, \\ \lambda_i \le 0,\ i = 1, \dots, m',\ \textit{для максимума}; \end{array}$$

d) для любого h *из конуса допустимых направлений*[10]

$$K = K(\widehat{x}) := \left\{ h \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \langle f'_i(\widehat{x}), h \rangle \le 0,\ i = 0, 1, \dots, m', \\ \langle f'_i(\widehat{x}), h \rangle = 0,\ i = m'+1, \dots, m. \end{array} \right\}$$

найдётся вектор множителей Лагранжа $\lambda \ne 0$, *удовлетворяющий условиям a)–c) и такой, что*

$$\begin{array}{l} \mathscr{L}''(\widehat{x})[h, h] \ge 0 \quad \textit{для минимума}, \\ \mathscr{L}''(\widehat{x})[h, h] \le 0 \quad \textit{для максимума}. \end{array} \qquad (\mathscr{L}'')$$

Замечание. Отметим, что если для точки локального экстремума $\widehat{x}$ вектор множителей Лагранжа λ единственен (с точностью до умножения на положительную константу), то неравенства $(\mathscr{L}'')$ выполняются для любого $h \in K$.

---

[10] Напомним, что в формировании конуса допустимых направлений из ограничений типа неравенств берутся только активные ограничения, т. е. только те, для которых $f_i(\widehat{x}) = 0,\ i = 1, \dots, m'$.

3.2.3 Достаточное условие экстремума 2-го порядка

Сформулируем достаточное условие второго порядка локального экстремума в гладкой конечномерной задаче с ограничениями типа равенств и неравенств:

$$f_0(x) \to \text{extr}; \qquad \begin{array}{l} f_i(x) \le 0,\ i = 1, \dots, m', \\ f_i(x) = 0,\ i = m'+1, \dots, m, \end{array} \qquad (P)$$

$f_i \colon \mathbb{R}^n \to \mathbb{R}$, $i = 0, 1, \dots, m$.

Теорема 3.3. *Пусть функции f_i, $i = 0, \dots, m$, дважды непрерывно дифференцируемы в некоторой окрестности допустимой точки $\widehat{x}$* (условие гладкости)*, для функции Лагранжа $\mathscr{L}(x) = f_0(x) + \sum\limits_{i=1}^{m} \lambda_i f_i(x)$ и вектора множителей Лагранжа $\lambda = (1, \lambda_1, \dots, \lambda_m)$ в точке $\widehat{x}$ выполняются условия:*

a) стационарности: $\mathscr{L}'(\widehat{x}) = 0$;

b) дополняющей нежёсткости: $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \dots, m'$;

c) неотрицательности: $\lambda_i \ge 0$, $i = 1, \dots, m'$;

d) положительной определенности матрицы вторых производных функции Лагранжа на пространстве стационарных направлений[11]

$$\Pi(\widehat{x}) := \left\{ h \in \mathbb{R}^n \;\middle|\; \begin{array}{rl} \lambda_i \langle f_i'(\widehat{x}), h \rangle = 0, & i = 1, \dots, m', \\ \langle f_i'(\widehat{x}), h \rangle = 0, & i = m'+1, \dots, m, \end{array} \right\} :$$

$$\langle \mathscr{L}''(\widehat{x}) h, h \rangle > 0 \quad \forall\ h \in \Pi,\ h \ne 0. \qquad (*)$$

Тогда $\widehat{x} \in \operatorname{locmin} P$ — точка локального минимума.

Если условия c) и d) выполняются в виде:

c′) $\lambda_i \le 0$, $i = 1, \dots, m'$;

d′) $\langle \mathscr{L}''(\widehat{x}) h, h \rangle < 0 \quad \forall\ h \in \Pi,\ h \ne 0$,

то $\widehat{x} \in \operatorname{locmax} P$ — точка локального максимума.

---

[11] Напомним, что в формировании пространства стационарных направлений из ограничений типа неравенств берутся только активные ограничения, т. е. только те, для которых $f_i(\widehat{x}) = 0$.

3.3 Примеры

Пример 1. $x_1^2+x_2^2+x_1x_2-3x_1-3x_2 \to \text{extr}; \quad 3x_1+2x_2 \le 8.$

Решение. Функция Лагранжа

$$\mathscr{L} = \lambda_0(x_1^2+x_2^2+x_1x_2-3x_1-3x_2)+\lambda_1(3x_1+2x_2-8).$$

Необходимые условия локального экстремума:

a) стационарности:

$$\mathscr{L}_x = 0 \iff \begin{cases} \mathscr{L}_{x_1} = 0, \\ \mathscr{L}_{x_2} = 0, \end{cases} \iff \begin{cases} \lambda_0(2x_1+x_2-3)+3\lambda_1 = 0, \\ \lambda_0(2x_2+x_1-3)+2\lambda_1 = 0; \end{cases}$$

b) дополняющей нежёсткости: $\lambda_1(3x_1+2x_2-8)=0$;

c) знакоопределенности:

$$\lambda_0 \ge 0, \quad \begin{array}{l} \lambda_1 \ge 0, \text{ в задаче на минимум,} \\ \lambda_1 \le 0, \text{ в задаче на максимум.} \end{array}$$

Если $\lambda_0 = 0$, то из уравнений пункта a) выводим, что $\lambda_1 = 0$ — все множители Лагранжа — нули, а этого быть не может. Значит, при $\lambda_0 = 0$ допустимых точек нет.

Пусть $\lambda_0 \ne 0$. Полагаем для удобства $\lambda_0 = 1$. Получим

$$\begin{cases} 2x_1+x_2-3+3\lambda_1 = 0, \\ x_1+2x_2-3+2\lambda_1 = 0, \end{cases} \iff \begin{cases} x_1 = 1-\frac{4}{3}\lambda_1, \\ x_2 = 1-\frac{1}{3}\lambda_1. \end{cases} \qquad (a')$$

Предположим $\lambda_1 = 0$. Тогда $x_1 = x_2 = 1$. Получили критическую точку на минимум и на максимум $\widehat{x}^1 = (1,1)$.

Пусть $\lambda_1 \ne 0$, тогда в силу условия b) $3x_1+2x_2-8=0$. Подставим x_1, x_2 из (a') в уравнение: $3\big(1-\frac{4}{3}\lambda_1\big)+2\big(1-\frac{1}{3}\lambda_1\big)=8$ $\Leftrightarrow -4\lambda_1-\frac{2}{3}\lambda_1 = 3 \Leftrightarrow -\frac{14}{3}\lambda_1 = 3 \Leftrightarrow \lambda_1 = -\frac{9}{14} < 0$ (критическая точка будет в задаче на максимум). Соответственно,

$$\begin{cases} x_1 = 1-\frac{4}{3}\big(-\frac{9}{14}\big) = 1+\frac{6}{7} = \frac{13}{7}, \\ x_2 = 1-\frac{1}{3}\big(-\frac{9}{14}\big) = 1+\frac{3}{14} = \frac{17}{14}. \end{cases}$$

Значит, в случае $\lambda_1 \neq 0$ имеется критическая точка в задаче на максимум $\widehat{x}^2 = \left(\frac{13}{7}, \frac{17}{14}\right)$.

Исследование. Вторая производная функционала

$$f_0'' = A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} > 0 \quad (A_1 = 2,\ A_{12} = 4 - 1 > 0)$$

по критерию Сильвестра является положительно определенной матрицей. Значит, функция $f_0(x) = x_1^2 + x_2^2 + x_1x_2 - 3x_1 - 3x_2 \to +\infty$ при $|x| \to \infty$. Тогда по следствию из теоремы Вейерштрасса решение задачи на минимум существует, а в силу единственности критической точки на минимум решением может быть только точка $(1, 1)$. При этом $S_{\text{absmax}} = +\infty$, так как множество допустимых точек неограничено.

Осталось проверить, доставляет ли критическая точка $\widehat{x}^2 = \left(\frac{13}{7}, \frac{17}{14}\right)$ с $\lambda_1 = -\frac{9}{14}$ локальный максимум. Исследуем её с помощью необходимых условий 2-го порядка. В ней неравенство $f_1(x) = 3x_1 + 2x_2 - 8 \leq 0$ обращается в равенство (это можно не проверять, поскольку $\lambda_1 \neq 0$, а в силу условия дополняющей нежёсткости $\lambda_1(3x_1 + 2x_2 - 8) = 0$), значит, ограничение является активным.

Пространство стационарных направлений

$$\Pi := \{h \in \mathbb{R}^2 \mid \langle f_1'(\widehat{x}), h\rangle = 0 \Leftrightarrow 3h_1 + 2h_2 = 0\} \neq 0.$$

Для функции Лагранжа $\mathscr{L} = x_1^2 + x_2^2 + x_1x_2 - 3x_1 - 3x_2 + \lambda_1(3x_1 + 2x_2 - 8)$ вторая производная

$$\langle \mathscr{L}''h, h\rangle = 2(h_1^2 + h_2^2 + h_1h_2) > 0 \quad \forall\ h \neq 0.$$

Матрица A положительно определена, поэтому необходимое условие максимума в задаче с неравенствами ($\langle Ah, h\rangle \leq 0$ $\forall\ h \in \Pi$) не выполняется. Значит, $\widehat{x}^2 \notin \text{locmax}$.

Ответ. $(1, 1) \in \text{absmin}$, $S_{\text{absmin}} = -\frac{3}{2}$, критическая точка в задаче на максимум $\left(\frac{13}{7}, \frac{17}{14}\right) \notin \text{locmax}$, $S_{\text{absmax}} = +\infty$.

Пример 2. $\frac{x_1^2+x_2^2+x_3^2}{2} \to \text{extr}$; $2x_1 - x_2 + x_3 \le 5$, $x_1 + x_2 + x_3 = 3$.

Решение. Функция Лагранжа

$$\mathscr{L} = \lambda_0 \frac{x_1^2+x_2^2+x_3^2}{2} + \lambda_1(2x_1 - x_2 + x_3 - 5) + \lambda_2(x_1 + x_2 + x_3 - 3).$$

Необходимые условия локального экстремума:

a) стационарности:

$$\mathscr{L}_x = 0 \iff \begin{cases} \mathscr{L}_{x_1} = 0, \\ \mathscr{L}_{x_2} = 0, \\ \mathscr{L}_{x_3} = 0, \end{cases} \iff \begin{cases} \lambda_0 x_1 + 2\lambda_1 + \lambda_2 = 0, \\ \lambda_0 x_2 - \lambda_1 + \lambda_2 = 0, \\ \lambda_0 x_3 + \lambda_1 + \lambda_2 = 0; \end{cases}$$

b) дополняющей нежёсткости: $\lambda_1(2x_1 - x_2 + x_3 - 5) = 0$;

c) знакоопределенности: $\lambda_0 \ge 0$, $\begin{matrix} \lambda_1 \ge 0, \text{ на минимум}, \\ \lambda_1 \le 0, \text{ на максимум}. \end{matrix}$

Если $\lambda_0 = 0$, то из уравнений пункта a) выводим, что $\lambda_1 = \lambda_2 = 0$ — все множители Лагранжа — нули, а этого быть не может. Значит, при $\lambda_0 = 0$ допустимых точек нет.

Пусть $\lambda_0 \ne 0$. Полагаем для удобства $\lambda_0 = 1$. Получим

$$\begin{cases} x_1 + 2\lambda_1 + \lambda_2 = 0, \\ x_2 - \lambda_1 + \lambda_2 = 0, \\ x_3 + \lambda_1 + \lambda_2 = 0, \end{cases} \iff \begin{cases} x_1 = -2\lambda_1 - \lambda_2, \\ x_2 = \lambda_1 - \lambda_2, \\ x_3 = -\lambda_1 - \lambda_2. \end{cases} \qquad (a')$$

Предположим $\lambda_1 = 0$. Тогда из (a') $x_1 = x_2 = x_3 = -\lambda_2$. Подставляя $x_i = -\lambda_2$ в равенство $x_1 + x_2 + x_3 = 3$, получаем, что $-3\lambda_2 = 3 \Leftrightarrow \lambda_2 = -1 \Rightarrow x_1 = x_2 = x_3 = 1$. Получили точку $\widehat{x}^1 = (1, 1, 1)$ критическую в задаче и на минимум, и на максимум.

Пусть $\lambda_1 \ne 0$, тогда в силу условия b) $2x_1 - x_2 + x_3 - 5 = 0$. Подставим x_1, x_2, x_3 в уравнения $x_1 + x_2 + x_3 = 3$, $2x_1 - x_2 + x_3 = 5$:

$$\begin{cases} -2\lambda_1 - \lambda_2 + \lambda_1 - \lambda_2 - \lambda_1 - \lambda_2 = 3, \\ -4\lambda_1 - 2\lambda_2 - \lambda_1 + \lambda_2 - \lambda_1 - \lambda_2 = 5, \end{cases} \iff \begin{cases} -2\lambda_1 - 3\lambda_2 = 3, \\ -6\lambda_1 - 2\lambda_2 = 5, \end{cases}$$

откуда $\lambda_1 = -\frac{9}{14} < 0$ (в этом случае критические точки будут в задаче на максимум), $\lambda_2 = -\frac{4}{7}$. Соответственно,

$$\begin{cases} x_1 = -2\lambda_1 - \lambda_2 = \frac{13}{7}, \\ x_2 = \lambda_1 - \lambda_2 = -\frac{1}{14}, \\ x_3 = -\lambda_1 - \lambda_2 = \frac{17}{14}. \end{cases}$$

Значит, в случае $\lambda_1 \neq 0$ имеется критическая точка в задаче на максимум $\widehat{x}^2 = \left(\frac{13}{7}, -\frac{1}{14}, \frac{17}{14}\right)$.

Исследование. Функция $f_0(x) = x_1^2 + x_2^2 + x_3^2 \to +\infty$ при $|x| \to \infty$, значит, по следствию из теоремы Вейерштрасса решение задачи на минимум существует. В силу единственности критической точки в задаче на минимум решением может быть только точка $(1, 1, 1)$. При этом $S_{\text{absmax}} = +\infty$, так как множество допустимых точек неограничено.

Осталось проверить, доставляет ли критическая точка $\widehat{x}^2 = \left(\frac{13}{7}, -\frac{1}{14}, \frac{17}{14}\right)$ с $\lambda_1 = -\frac{9}{14}$ локальный максимум. Исследуем её с помощью необходимых условий 2-го порядка. В ней неравенство $f_1(x) = 2x_1 - x_2 + x_3 - 5 \leq 0$ обращается в равенство (это можно не проверять, поскольку $\lambda_1 \neq 0$, а в силу условия дополняющей нежёсткости $\lambda_1(2x_1 - x_2 + x_3 - 5) = 0$), значит, ограничение является активным.

Пространство стационарных направлений

$$\Pi := \left\{ h \in \mathbb{R}^2 \;\middle|\; \begin{array}{l} \langle f_1'(\widehat{x}), h\rangle = 0, \\ \langle f_2'(\widehat{x}), h\rangle = 0, \end{array} \Leftrightarrow \begin{array}{r} 2h_1 - h_2 + h_3 = 0, \\ h_1 + h_2 + h_3 = 0, \end{array} \right\} \neq 0.$$

Для функции Лагранжа $\mathscr{L} = \frac{x_1^2 + x_2^2 + x_3^2}{2} + \lambda_1(2x_1 - x_2 + x_3 - 5) + \lambda_2(x_1 + x_2 + x_3 - 3)$ вторая производная

$$\langle \mathscr{L}''(\widehat{x})h, h\rangle = h_1^2 + h_2^2 + h_3^2 > 0 \quad \forall\, h \neq 0.$$

Матрица A положительно определена, поэтому необходимое условие максимума в с неравенствами ($\langle Ah, h\rangle \leq 0 \;\forall\; h \in \Pi$) не выполняется. Значит, $\widehat{x}^2 \notin \text{locmax}$.

Ответ. $(1, 1, 1) \in \text{absmin}$, $S_{\text{absmin}} = 3$, критическая точка в задаче на максимум $\left(\frac{13}{7}, -\frac{1}{14}, \frac{17}{14}\right) \notin \text{locmax}$, $S_{\text{absmax}} = +\infty$.

Необходимое условие экстремума 2-го порядка

Теорема. *Пусть* $\widehat{x} \in \operatorname{locextr} P$, $f_i \in C^2(\mathcal{O}(\widehat{x}))$, $i = 0, 1, \ldots, m$, *векторы* $f'_{m'+1}(\widehat{x}), \ldots, f'_m(\widehat{x})$ *— линейно независимы. Тогда существует ненулевой вектор множителей Лагранжа* $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$, *такой, что для функции Лагранжа* $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ *выполняются условия:*

a) стационарности[12]*:* $\mathscr{L}'(\widehat{x}) = 0$*;*

b) дополняющей нежёсткости: $\lambda_i f_i(\widehat{x}) = 0,\ i = 1, \ldots, m'$;

c) знакоопределенности:

$$\lambda_0 \geq 0, \quad \begin{array}{ll} \lambda_i \geq 0,\ i=1,\ldots,m', & \text{для локального минимума,} \\ \lambda_i \leq 0,\ i=1,\ldots,m', & \text{для локального максимума.} \end{array}$$

для функции Лагранжа $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ $(\lambda \neq 0)$ *выполняются условия:*

стационарности: $\mathscr{L}'(\widehat{x}) = 0$;

дополняющей нежёсткости: $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \ldots, m'$;

знакоопределенности множителей Лагранжа:

$$\lambda_0 \geq 0, \quad \begin{array}{ll} \lambda_i \geq 0,\ i = 1,\ldots,m', & \text{для минимума,} \\ \lambda_i \leq 0,\ i = 1,\ldots,m', & \text{для максимума;} \end{array}$$

непусто;

2) *для любого* h *из конуса допустимых направлений*

$$K = K(\widehat{x}) := \left\{ h \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \langle f'_i(\widehat{x}), h \rangle \leq 0,\ i = 0, 1, \ldots, m', \\ \langle f'_i(\widehat{x}), h \rangle = 0,\ i = m'+1, \ldots, m. \end{array} \right\}^{13}$$

найдётся, зависящий от h, *множитель Лагранжа* $\lambda \in \Lambda$ *такой, что*

$$\begin{array}{ll} \mathscr{L}''(\widehat{x})[h,h] \geq 0 & \text{для минимума,} \\ \mathscr{L}''(\widehat{x})[h,h] \leq 0 & \text{для максимума.} \end{array} \qquad (\mathscr{L}'')$$

---

[12] Условие стационарности можно расписать в эквивалентном виде: $\dfrac{\partial \mathscr{L}(\widehat{x})}{\partial x_j} = 0,\ j = 1, \ldots, n$ или $\sum\limits_{i=0}^{m} \lambda_i f'_i(\widehat{x}) = 0 \Leftrightarrow \sum\limits_{i=0}^{m} \lambda_i \langle f'_i(\widehat{x}), h \rangle = 0 \,\forall\, h \in \mathbb{R}^n$.

[13] Напомним, что в формировании конуса допустимых направлений из ограничений типа неравенств берутся только активные ограничения, т. е. только те, для которых $f_i(\widehat{x}) = 0$.

Достаточное условие экстремума 2-го порядка

Теорема. $f_i \in C^2(\mathcal{O}(\widehat{x}))$, $i = 0, \ldots, m$, *для функции Лагранжа* $\mathscr{L}(x) = f_0(x) + \sum\limits_{i=1}^{m} \lambda_i f_i(x)$ *выполняются условия*:

a) *стационарности*: $\mathscr{L}'(\widehat{x}) = 0$;

b) *дополняющей нежёсткости*: $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \ldots, m'$;

c) *неотрицательности*: $\lambda_i \geq 0$, $i = 1, \ldots, m'$;

d) $\langle \mathscr{L}''(\widehat{x})h, h\rangle > 0 \ \forall\ h \in \Pi,\ h \neq 0 \Longrightarrow \widehat{x} \in \operatorname{locmin} P$.

Если условия c) *и d*) *выполняются в виде*:

c$'$) $\lambda_i \leq 0,\ i = 1, \ldots, m'$;

d$'$) $\langle \mathscr{L}''(\widehat{x})h, h\rangle < 0 \ \forall\ h \in \Pi,\ h \neq 0 \Longrightarrow \widehat{x} \in \operatorname{locmax} P$,

где $\Pi(\widehat{x}) := \left\{ h \in \mathbb{R}^n \;\middle|\; \begin{array}{rl} \lambda_i \langle f_i'(\widehat{x}), h\rangle = 0, & i = 1, \ldots, m', \\ \langle f_i'(\widehat{x}), h\rangle = 0, & i = m'+1, \ldots, m. \end{array} \right\}$

4 Выпуклый анализ

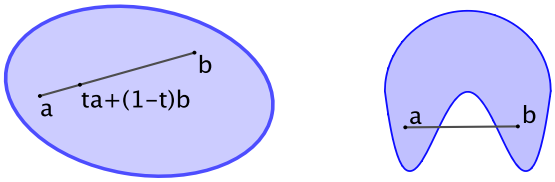
Пусть в этом пункте X — линейное нормированное пространство (для простоты понимания можно считать, что $X = \mathbb{R}^n$ — конечномерное пространство).

4.1 Элементы выпуклого анализа

4.1.1 Выпуклые множества

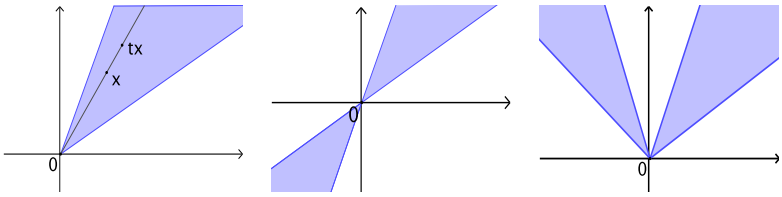
Введем некоторые понятия, которые используются в выпуклом анализе:

- *отрезок* $[a;\, b] = \{x \in X \mid x = ta + (1-t)b,\ 0 \le t \le 1\}$ или в виде $(1-t)a + tb$, $a + (1-t)b$;
- *интервал* $(a;\, b) = \{x \in X \mid x = ta + (1-t)b,\ 0 < t < 1\}$;
- *выпуклое множество* A, если $\forall\, a, b \in A$ отрезок $[a; b] \subset A$, т. е. $\forall\ a, b \in A$ точка $ta + (1-t)b \in A\ \forall\ 0 \le t \le 1$;



Слева — выпуклое множество; справа — невыпуклое множество.

- *конус* K $(K \ne \emptyset)$, если $\forall\ x \in K$ точка $tx \in K\ \forall\ t \ge 0$;



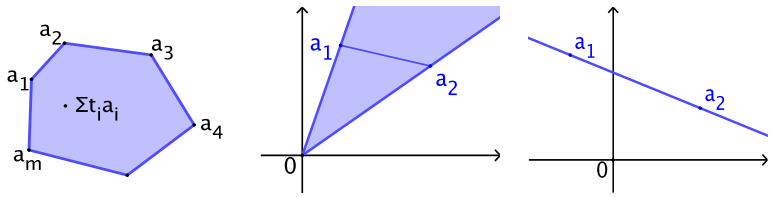
Слева — выпуклый конус; справа — два невыпуклых конуса.

- *аффинное множество* A, если $\forall\ a, b \in A$ точка $ta + (1-t)b \in A\ \forall\ t \in \mathbb{R}$.

Очевидно, что аффинное множество выпукло.

Возьмем фиксированные точки $a_1, \ldots, a_m \in X$. Пусть $\sum\limits_{i=1}^{m} t_i a_i$ — комбинация этих точек, $t_i \in \mathbb{R}$. Дадим различные определения таких комбинаций и оболочек этих комбинаций:

- *выпуклая комбинация*, если $t_i \geq 0$, $\sum\limits_{i=1}^{m} t_i = 1$;
- *коническая комбинация*, если $t_i \geq 0$;
- *аффинная комбинация*, если $\sum\limits_{i=1}^{m} t_i = 1$;
- *выпуклая оболочка* $\operatorname{conv}\{a_1, \ldots, a_m\} := \left\{ a = \sum\limits_{i=1}^{m} t_i a_i \,\middle|\, t_i \geq 0,\ \sum\limits_{i=1}^{m} t_i = 1 \right\}$ — совокупность всех выпуклых комбинаций;
- *коническая выпуклая оболочка* $\operatorname{conev}\{a_1, \ldots, a_m\} := \left\{ a = \sum\limits_{i=1}^{m} t_i a_i \,\middle|\, t_i \geq 0 \right\}$;
- *аффинная оболочка* $\operatorname{aff}\{a_1, \ldots, a_m\} := \left\{ a = \sum\limits_{i=1}^{m} t_i a_i \,\middle|\, \sum\limits_{i=1}^{m} t_i = 1 \right\}$.



Слева — выпуклая оболочка точек $a_1, \ldots, a_m$; посередине — коническая выпуклая оболочка двух точек a_1, a_2; справа — аффинная оболочка двух точек a_1, a_2.

4.1.2 Выпуклые функции

Пусть задана функция (функционал), отображающая линейное нормированное пространство в “расширенную” прямую:

$$f\colon X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}.$$

С каждой такой функцией f связываются два множества:
$\operatorname{dom} f = \{x \in X \mid f(x) < +\infty\}$ — *эффективное множество*
и $\operatorname{epi} f = \{(a, x) \in \mathbb{R} \times X \mid a \geq f(x)\}$ — *надграфик* функции f.

- Функция f называется *выпуклой*,
  если надграфик f — выпуклое множество.

- Функция f называется *замкнутой*,
  если надграфик f — замкнутое множество.

- Функция f называется *собственной*,
  если $f(x) > -\infty \ \forall\, x$ и $f \not\equiv +\infty$.

Случай $f(x) \equiv +\infty$ не интересен.

Если же для выпуклой функции существует точка z_0 такая, что $f(z_0) = -\infty$, а $|f(z_1)| < +\infty$, то в силу выпуклости надграфика $f(z) = -\infty \ \forall\, z \in [z_0, z_1)$. Если при этом функция f является к тому же замкнутой, то $f(z) = -\infty \ \forall\, z \in \operatorname{dom} f$.

Мы будем изучать выпуклые собственные функции. Для краткости будем называть их просто “выпуклые” функции. Иногда их называют ещё *выпуклые вниз* функции. Можно рассматривать также *выпуклые вверх* (*вогнутые*) функции. Они аналогично определяются с помощью подграфика функции.

Можно выпуклость определять не на всём пространстве, а на некотором интервале или на выпуклом подмножестве некоторого векторного пространства. Если не оговорено иначе, то будем выпуклость определять на всём пространстве.

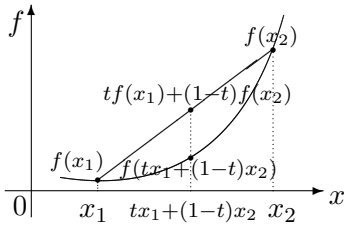
Из определения выпуклого множества сразу следует, что функция выпукла тогда и только тогда, когда выполнено *неравенство Йенсена*[14]:

$$f(tx_1+(1-t)x_2) \le tf(x_1)+(1-t)f(x_2)\ \forall\, x_1, x_2 \in X,\ \forall\, t \in (0;\,1).$$

Действительно, если надграфик выпуклое множество, то выпуклая оболочка точек $(x_1, f(x_1))$ и $(x_2, f(x_2))$ точка $t(x_1,f(x_1))+(1-t)(x_2,f(x_2)) = (tx_1+(1-t)x_2, tf(x_1)+(1-t)f(x_2))$ также принадлежит надграфику. Тогда по определению надграфика

$$f(x) \le a \iff f(tx_1 + (1-t)x_2) \le tf(x_1) + (1-t)f(x_2).$$

Геометрически это означает, что точка на хорде, соединяющая две точки на графике, лежит выше соответствующей точки графика.



Ясно, что сумма $\lambda_1 f_1 + \lambda_2 f_2$ выпуклых функций f_1 и f_2 с неотрицательными коэффициентами λ_1 и λ_2 является функцией выпуклой.

Суперпозиция двух выпуклых функций не всегда является выпуклой функцией. Например, функция e^{-x^2}, являющаяся суперпозицией выпуклых функций $f(x) = x^2$ и $g(t) = e^{-t}$, не является выпуклой функцией.

Непрерывно дифференцируемая функция одной переменной выпукла на интервале тогда и только тогда, когда её гра-

[14]Иоган Людвиг Виллиам Вальдемар Йенсен — датский математик и инженер (1859-1925).

фик лежит не ниже касательной, проведённой к этому графику в любой точке промежутка выпуклости.

Определение 3. *Точка d выпуклого множества D называется крайней (угловой), если не существует точек $d_1, d_2 \in D$, $d_1 \neq d_2$, и числа $t \in (0; 1)$ таких, что $d = td_1 + (1-t)d_2$.*

Определение 4. *Замкнутое выпуклое множество называется строго выпуклым, если все его граничные точки крайние.*

У отрезка $[a; b]$ крайними точками являются концы отрезка — точки a и b.

У треугольника крайними точками являются его вершины. Точки сторон треугольника, не являющиеся вершинами, не являются крайними.

У круга крайними точками являются точки порождающей его окружности. У круга все граничные точки являются крайними.

Определение 5. *Функция f называется строго выпуклой, если её надграфик — строго выпуклое множество.*

Из определения строго выпуклого множества сразу следует, что функция f строго выпукла тогда и только тогда, когда для любых точек $x_1, x_2 \in X$, $x_1 \neq x_2$, и для любого $t \in (0; 1)$ выполнено *строгое неравенство Йенсена*:

$$f(tx_1 + (1-t)x_2) < tf(x_1) + (1-t)f(x_2).$$

Теорема 4.1. *Пусть $f\colon \mathbb{R} \to \mathbb{R}$ — дифференцируемая функция (на всей числовой прямой). Тогда функция f выпукла тогда и только тогда, когда её производная f' не убывает (на всей числовой прямой).*

Доказательство. Необходимость. Пусть дифференцируемая функция f выпукла. Докажем, что ее производная не убывает. Возьмем произвольную точку $x = tx_1 + (1-t)x_2$, $0 < t < 1$, из интервала $(x_1; x_2)$. Из равенства для x находится $t = \frac{x_2 - x}{x_2 - x_1}$, $1 - t = \frac{x - x_1}{x_2 - x_1}$. Тогда неравенство Йенсена для выпуклой функции f можно записать в виде:

$$f(x) \le tf(x_1) + (1-t)f(x_2) \iff$$
$$f(x) \le \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2) \iff$$

(умножим обе части неравенства на положительное число $x_2 - x_1$ равное $x_2 - x + x - x_1$)

$$(x_2 - x + x - x_1)f(x) \le (x_2 - x)f(x_1) + (x - x_1)f(x_2)$$
$$\iff (x_2 - x)(f(x) - f(x_1) \le (x - x_1)(f(x_2) - f(x)) \iff$$

(разделим обе части неравенства на положительное число $(x_2 - x)(x - x_1)$)

$$\frac{f(x) - f(x_1)}{x - x_1} \le \frac{f(x_2) - f(x)}{x_2 - x}. \qquad (*)$$

Устремляя в неравенстве $(*)$ x к x_1, получим

$$f'(x_1) \le \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Устремляя в неравенстве $(*)$ x к x_2, получим

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \le f'(x_2).$$

Из двух последних неравенств следует, что $f'(x_1) \le f'(x_2)$.

Достаточность. Пусть функция f' не убывает. Докажем, выпуклость функции f. Возьмем точки $x_1 < x_2$.
В обозначениях предыдущего пункта по теореме Лагранжа

$$f(x) - f(x_1) = f'(\xi_1)(x - x_1) \iff \frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1);$$
$$f(x_2) - f(x) = f'(\xi_2)(x_2 - x) \iff \frac{f(x_2) - f(x)}{x_2 - x} = f'(\xi_2),$$

где $x_1 < \xi_1 < x < \xi_2 < x_2$. Поскольку функция f' не убывает, то $f'(\xi_1) \le f'(\xi_2)$ и, следовательно,

$$\frac{f(x) - f(x_1)}{x - x_1} \le \frac{f(x_2) - f(x)}{x_2 - x}.$$

Таким образом, для функции f выполнено соотношение $(*)$, которое эквивалентно выполнению неравенства Йенсена.
Значит, f — выпуклая функция. □

Выпуклость многих классических функций одной переменной сразу вытекает из следующей теоремы.

Теорема 4.2. *Пусть функция $f\colon \mathbb{R} \to \mathbb{R}$ дважды непрерывно дифференцируема ($f \in D^2(\mathbb{R})$). Тогда она выпукла тогда и только тогда, когда её вторая производная неотрицательна ($f''(x) \ge 0\ \forall\ x \in \mathbb{R}$).*

Если $f''(x) > 0\ \forall\ x \in \mathbb{R}$, то функция является строго выпуклой.

Доказательство. Достаточность. Пусть $f''(x) \ge 0\ \forall\ x \in \mathbb{R}$. Докажем, что тогда f — выпуклая функция. Поскольку выпуклость функции, как было сказано выше, эквивалентна выполнению неравенства Йенсена, то достаточно доказать, что выполняется неравенство Йенсена.

Возьмём две произвольные точки x_1, x_2 и их выпуклую комбинацию — точку $x_0 = tx_1 + (1-t)x_2$, $0 < t < 1$. Разлагая функцию f в ряд Тейлора в точке x_0 с остаточным членом

в форме Лагранжа, выпишем значения функции f в точках x_1, x_2:

$$f(x_1) = f(x_0) + f'(x_0)(x_1 - x_0) + \tfrac{1}{2} f''(\xi_1)(x_1 - x_0)^2 \geq \\ \geq f(x_0) + f'(x_0)(x_1 - x_0),$$

$$f(x_2) = f(x_0) + f'(x_0)(x_2 - x_0) + \tfrac{1}{2} f''(\xi_2)(x_2 - x_0)^2 \geq \\ \geq f(x_0) + f'(x_0)(x_2 - x_0),$$

где $\xi_1 \in [x_1; x_0]$, $\xi_2 \in [x_0; x_2]$ (неравенства выполняются в силу неотрицательности второй производной). Умножая первое неравенство на t, второе на $1 - t$, и складывая, получаем неравенство Йенсена:

$$tf(x_1) + (1-t)f(x_2) \geq tf(x_0) + (1-t)f(x_0) + \\ + f'(x_0)\big(t(x_1 - x_0) + (1-t)(x_2 - x_0)\big) = \\ = f(x_0) + f'(x_0)\big(\underbrace{tx_1 + (1-t)x_2}_{x_0} - x_0\big) = f(tx_1 + (1-t)x_2).$$

Если $f''(x) > 0$, то неравенства выполняются строгие.

Необходимость. Пусть выпуклая функция $f\colon \mathbb{R} \to \mathbb{R}$ дважды дифференцируема. Докажем, что $f''(x) \geq 0 \ \forall\ x \in \mathbb{R}$. Действительно, поскольку f выпуклая функция, то по Теореме 4.1 ее производная f' не убывает и значит, вторая производная $f''(x) \geq 0$. $\square$

Отметим, что достаточное условие строгой выпуклости функции ($f''(x) > 0$) не является необходимым. Например, функция $f(x) = x^4$ является строго выпуклой. Однако строгое неравенство $f''(x) > 0$ в этом случае не выполняется. так как в точке $x = 0$ её вторая производная равна нулю.

Приведем несколько примеров выпуклых функций.

1. $f(x) = e^{ax}$, $a \in \mathbb{R}$.

Выпуклость функции следует из Теоремы 4.2

2. $f(x) = |x|^p$, $p \geq 1$.

Выпуклость следует из выпуклости надграфика функции.

3. *Аффинная функция* (в многомерном случае $f\colon \mathbb{R}^n \to \mathbb{R}$, $f(x) = \langle a, x\rangle + b$, $a \in \mathbb{R}^n$, $b \in \mathbb{R}$).

Аффинная функция является выпуклой по неравенству Йенсена. Модуль аффинной функции $|\langle a, x\rangle + b|$ также является выпуклой функцией, поскольку её надграфик — пересечение полупространств — выпуклое множество.

Выпуклость функций нескольких переменных можно определять также из следующего многомерного обобщения Теоремы 4.2.

Теорема 4.3. *Пусть функция $f\colon \mathbb{R}^n \to \mathbb{R}$ дважды непрерывно дифференцируема ($f \in D^2(\mathbb{R}^n)$). Тогда она выпукла тогда и только тогда, когда её матрица вторых производных всюду неотрицательно определена*

$$\left(f''(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)_{i,j=1}^{n} \geq 0 \ \forall\ x \in \mathbb{R}^n \right).$$

4. *Квадратичная функция* $Q(x) = \langle Ax, x\rangle$ (A — симметричная матрица), является выпуклой тогда и только тогда, когда матрица A неотрицательно определена.

Это сразу вытекает из Теоремы 4.3.

Выпуклыми функциями многих переменных являются также следующие функции:

5. *Функция нормы*

$$f(x) = \|x\|_p = \|x\|_{l_p} := \begin{cases} \left(\sum\limits_{j=1}^{n} |x_j|^p \right)^{\frac{1}{p}}, & 1 \leq p < \infty, \\ \max\{|x_1|, \ldots, |x_n|\}, & p = +\infty. \end{cases}$$

6. *Индикаторная функция* выпуклого множества $A \subset X$

$$\delta A(x) = \begin{cases} 0, & x \in A, \\ +\infty, & x \notin A. \end{cases}$$

7. *Функция Минковского* выпуклого множества $A \subset X$

$$\mu A(x) = \begin{cases} 0, & \alpha^{-1}x \in A\ \forall\ \alpha > 0, \\ +\infty, & \alpha^{-1}x \notin A\ \forall\ \alpha > 0, \\ \inf\{\alpha \mid \alpha > 0,\ \alpha^{-1}x \in A\}, & \text{иначе.} \end{cases}$$

Функция Минковского означает наименьшее число, в которое надо уменьшить вектор x, чтобы он попал в множество A. Функцию Минковского более кратко можно задать формулой

$$\mu A(x) = \inf\{\alpha > 0 \mid x \in \alpha A\}. \qquad (*)$$

Свойства функции Минковского:

1°. *Функция Минковского является положительно однородной функцией*:

$$\mu A(\lambda x) = \lambda \mu A(x) \quad \forall\ \lambda > 0,\ \forall\ x \in X.$$

2°. *Функция Минковского является субаддитивной функцией*:

3°. *Функция Минковского является выпуклой функцией.*

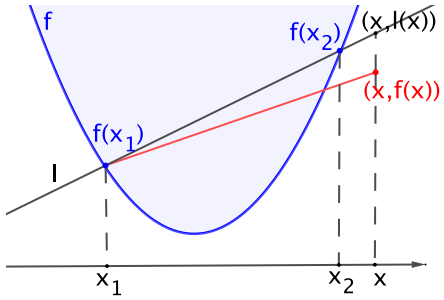
8. *Опорная функция* непустого множества $A \subset X$

$$sA(x^*) = \max_{x \in A} \langle x^*, x \rangle \qquad (sA\colon X^* \to \mathbb{R}).$$

Опорная функция является выпуклой как верхняя грань выпуклых функций (надграфик является пересечением полупространств — выпуклое множество).

Теорема 4.4. $f\colon \mathbb{R} \to \mathbb{R}$ — *выпуклая ограниченная функция. Тогда* $f(x) = \mathrm{const}$.

Доказательство. Предположим противное, что $f(x) \neq \mathrm{const}$. Тогда существуют точки x_1, x_2 такие, что $x_1 < x_2$ и $f(x_1) \neq f(x_2)$. Не ограничивая общности, считаем, что $f(x_1) < f(x_2)$. Поскольку f — выпуклая ограниченная функция, то надграфик функции $\operatorname{epi} f$ — выпуклое множество, следовательно, хорда $\big[(x_1, f(x_1)), (x_2, f(x_2))\big] \subset \operatorname{epi} f$. Значит, на отрезке $[x_1, x_2]$ график функции f лежит ниже этой хорды (или совпадает с ней).



Пусть $l(x)$ — прямая, проходящая через точки $(x_1, f(x_1))$ и $(x_2, f(x_2))$. Покажем, что

$$f(x) \geq l(x) \quad \forall\, x > x_2. \tag{$*$}$$

Если это не так, то существует точка $x > x_2$ такая, что $f(x) < l(x)$, т. е. точка $(x, f(x))$ лежит ниже прямой l. Так как функция f выпукла, то отрезок $\big[(x_1, f(x_1)), (x, f(x))\big] \subset \operatorname{epi} f$ и лежит ниже прямой l, но это не так в точке x_2 (ниже точки $(x_2, f(x_2))$ не может быть точек из $\operatorname{epi} f$). Поэтому неравенство $(*)$ действительно выполняется.

Так как $l(x) \to +\infty$ при $x \to +\infty$, то $f(x) \to +\infty$ при $x \to +\infty$. Получаем противоречие с ограниченностью функции f. Значит, предположение противного неверно и $f(x) = \mathrm{const}$. $\square$

4.1.3 Субдифференциал выпуклой функции

Дадим определение важного понятия выпуклого анализа — субдифференциала функции, обобщающего для выпуклых функций понятие производной в гладком анализе.

Субдифференциалом выпуклой функции f в точке $\widehat{x}$ называется следующее множество в сопряженном пространстве X^*:

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

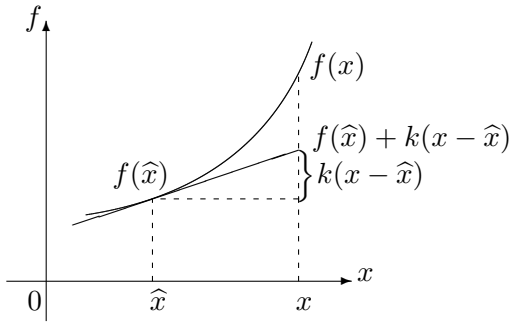
Напомним, что *сопряженным пространством* X^* называется пространство линейных непрерывных функционалов на X. В случае $X = \mathbb{R}^n$ сопряженное пространство $(\mathbb{R}^n)^* = \mathbb{R}^n$.

Из определения сразу вытекает, что субдифференциал — выпуклое множество в X^*. Легко доказать, что оно замкнуто. Субдифференциал дифференцируемой функции совпадает с её производной.

Неравенство в определении субдифференциала можно переписать в виде

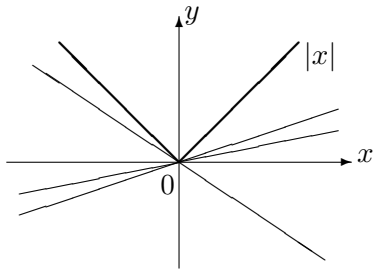
$$f(\widehat{x}) + \langle x^*, x - \widehat{x}\rangle \le f(x).$$

Для функции одной переменной субдифференциал $\partial f(\widehat{x})$ — это совокупность угловых коэффициентов $k = x^*$, при которых прямые $y = kx + b$, проходящие через точку $(\widehat{x}, f(\widehat{x}))$, лежат под графиком функции $y = f(x)$.



$$\partial f(\widehat{x}) = \{k \in \mathbb{R} \mid f(\widehat{x}) + k(x - \widehat{x}) \le f(x)\ \forall\, x \in \mathbb{R}\}.$$

Пример. $f(x) = |x|$.



$$\partial|x| = \begin{cases} \operatorname{sign} x, & x \neq 0, \\ [-1;\, 1], & x = 0. \end{cases}$$

Для субдифференциала суммы функций имеет место теорема аналогичная теореме о производной суммы функций.

Теорема (Моро–Рокафеллар)**.** *Пусть f_1 и f_2 — выпуклые функции на X. Существует точка x_0, в которой функция f_1 конечна ($|f_1(x_0)| < \infty$), а функция f_2 непрерывна ($f_2 \in C(x_0)$). Тогда*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x) \quad \forall\ x.$$

При доказательстве необходимых условий экстремума в гладкой задаче с равенствами и неравенствами нам понадобится следующая теорема о субдифференциале максимума.

Теорема (Дубовицкий–Милютин)**.** *Пусть f_1 и f_2 — выпуклые функции, непрерывные в точке $\widehat{x}$, $f_1(\widehat{x}) = f_2(\widehat{x})$. Тогда*

$$\partial \max\{f_1, f_2\}(\widehat{x}) = \operatorname{conv}\big(\partial f_1(\widehat{x}) \cup \partial f_2(\widehat{x})\big).$$

4.1.4 Сублинейные функции

Важным примером выпуклой функции являются сублинейные функции.

- Функция называется *сублинейной*, если её надграфик является выпуклым конусом с вершиной в нуле.

Из неравенства Йенсена следует, что собственная функция p является сублинейной тогда и только тогда, когда она положительно однородна и субаддитивна:

a) $p(\lambda x) = \lambda p(x)$ для любого $\lambda > 0$, для любого $x \in X$;

b) $p(x_1 + x_2) \leq p(x_1) + p(x_2)$ для любых $x_1, x_2 \in X$.

Если X — нормированное пространство, то $x \to \|x\|$ — сублинейная функция.

Для непустого множества $A \subset X$ функция Минковского $\mu A(x)$ и опорная функции $sA(x^*)$ также являются сублинейными функциями.

Субдифференциал модуля функции

Пусть $f(x)$ — дифференцируемая функция n переменных. Рассмотрим функцию $|f(x)|$. Найдём её субдифференциал.

Если $f(x) < 0$, то $|f(x)| = -f(x)$ и $\partial|f(x)| = -f'(x)$.

Если $f(x) > 0$, то $|f(x)| = f(x)$ и $\partial|f(x)| = f'(x)$.

Если $f(x) = 0$, то запишем функцию $|f(x)|$ в виде $|f(x)| = \max\{-f(x), f(x)\}$. Тогда по теореме Дубовицкого–Милютина

$$\partial|f(x)| = \partial \max\{-f(x), f(x)\} = \operatorname{conv}\big\{\partial(-f)(x) \cup \partial f(x)\big\} =$$

$$= \operatorname{conv}\{-f'(x), f'(x)\} = [-f'(x);\, f'(x)].$$

Субдифференциал является отрезком в пространстве $\mathbb{R}^n$. Таким образом,

$$\partial|f(x)| = \begin{cases} -f'(x), & f(x) < 0, \\ \{\alpha f'(x),\ \alpha \in [-1;\, 1]\}, & f(x) = 0, \\ f'(x), & f(x) > 0. \end{cases}$$

Напомним, что для функции n переменных производная есть вектор частных производных $f'(x) = \left(\dfrac{\partial f(x)}{\partial x_1}, \ldots, \dfrac{\partial f(x)}{\partial x_n}\right)$. Поэтому

$$\partial|f(x)| = \begin{cases} -\left(\dfrac{\partial f(x)}{\partial x_1}, \ldots, \dfrac{\partial f(x)}{\partial x_n}\right), & f(x) < 0, \\ \left\{\alpha\left(\dfrac{\partial f(x)}{\partial x_1}, \ldots, \dfrac{\partial f(x)}{\partial x_n}\right),\ \alpha \in [-1;\, 1]\right\}, & f(x) = 0, \\ \left(\dfrac{\partial f(x)}{\partial x_1}, \ldots, \dfrac{\partial f(x)}{\partial x_n}\right), & f(x) > 0. \end{cases}$$

Субдифференциал нормы в нуле

Пусть X — линейное нормированное пространство, функционал $f\colon X \to \mathbb{R}$, $f(x) = \|x\|$. Найдём субдифференциал нормы в нуле.

По определению субдифференциала

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

У нас $\widehat{x} = 0$, $f(\widehat{x}) = \|0\| = 0$. Следовательно,

$$x^* \in \partial f(0) \iff \langle x^*, x\rangle \le \|x\| \quad \forall\, x \in X. \qquad (*)$$

Пусть $x^* \in \partial f(0)$. По лемме Банаха (см. п. 5.4) существует вектор $x \in X$ такой, что $\|x\| = 1$, $\langle x^*, x\rangle = \|x^*\|$. Подставляя этот вектор x в неравенство $(*)$, получим

$$\|x^*\| = \langle x^*, x\rangle \overset{(*)}{\le} \|x\| = 1 \Longrightarrow \|x^*\| \le 1.$$

Значит, $\partial f(0) \subset B^* := \left\{x^* \in X^* \mid \|x^*\|_{X^*} \le 1\right\}$ — единичный шар сопряженного пространства. Докажем, что $\partial f(0) = B^*$.

Для этого надо доказать обратное включение $\partial f(0) \supset B^*$. Возьмем $x^* \in B^*$, тогда $\|x^*\| \le 1$. По неравенству Коши–Буняковского

$$\langle x^*, x\rangle \le \|x^*\| \cdot \|x\| \overset{\|x^*\| \le 1}{\le} \|x\| \quad \forall\, x \in X.$$

Неравенство $(*)$ выполняется, значит, $x^* \in \partial f(0)$.
Таким образом, $\partial f(0) = \{x^* \in X^* \mid \|x^*\|_{X^*} \le 1\} = B^*$.

4.2 Отделимость выпуклых множеств

При выводе необходимых условий экстремума (принципа Лагранжа) в выпуклых задачах и в задачах с равенствами и неравенствами мы будем использовать свойство отделимости непересекающихся выпуклых множеств.

Определение 6. *Множества A и B из пространства X называются отделимыми, если существует линейный непрерывный функционал $\lambda \in X^*$, $\lambda \neq 0$, для которого*

$$\inf_{a \in A} \langle \lambda, a \rangle \geq \sup_{b \in B} \langle \lambda, b \rangle.$$

Из определения следует, что множества являются отделимыми, если можно провести гиперплоскость $H := \{x \in X \mid \langle \lambda, x \rangle = c\}$, $c \in \mathbb{R}$, так, что одно из множеств лежит в одном замкнутом полупространстве $H_+ := \{x \in X \mid \langle \lambda, x \rangle \geq c\}$, а другое — в другом замкнутом полупространстве $H_- := \{x \in X \mid \langle \lambda, x \rangle \leq c\}$.

Замечание. Поскольку функционал λ в определении отделимости является непрерывным, то для отделимости (не строгой!) вместо множеств A и B можно рассматривать их внутренности или замыкания.

Определение 7. *Множества A и B называются строго отделимыми, если существует линейный непрерывный функционал $\lambda \in X^*$, для которого*

$$\inf_{a \in A} \langle \lambda, a \rangle > \sup_{b \in B} \langle \lambda, b \rangle.$$

4.2.1 Теоремы отделимости

Приведем результаты об отделимости в конечномерном случае и в случае линейных нормированных пространств.

Теорема 4.5 (отделимости множеств в конечномерном пространстве)**.** *Пусть A и B — непустые выпуклые множества в $\mathbb{R}^n$, $A \cap B = \emptyset$. Тогда множества A и B отделимы.*

Замечание. Поскольку в определении отделимости (не строгой) стоит нестрогое неравенство, то, если внутренность одного из множеств непуста, например, $\operatorname{int} A \neq \emptyset$, то вместо условия $A \cap B = \emptyset$ в теореме отделимости можно требовать, что бы $\operatorname{int} A \cap B = \emptyset$.

Теорема 4.6 (отделимости множеств в линейном топологическом пространстве)**.** *Пусть A и B — непустые выпуклые множества в X,* $\operatorname{int} A \neq \emptyset$, $(\operatorname{int} A) \cap B = \emptyset$. *Тогда множества A и B отделимы.*

Примеры двух непересекающихся выпуклых подмножеств, которые нельзя отделить.

В теореме отделимости в конечномерном пространстве $\operatorname{int} A$ может быть пуста. В теореме отделимости в бесконечномерном пространстве это не так. Условие $\operatorname{int} A \neq \emptyset$ является существенным. Покажем это на примере в бесконечномерном гильбертовом пространстве $l_2 = \left\{x = (x_1, x_2, \ldots) \mid \sum\limits_{k\in\mathbb{N}} x_k^2 < \infty\right\}$.

Напомним, что из определения выпуклого множества следует, что пустое множество является выпуклым.

Пример 3. *В пространстве l_2 выпуклое множество $A = \left\{a \in l_2 \mid \sum\limits_{k\in\mathbb{N}} a_k = 1\right\}$ и точку $b = 0$ нельзя отделить.*

Решение. Множество A имеет пустую внутренность. Действительно, если предположить противное, что внутренность A непуста, то существует шар $B(\widehat{a}, \delta) = \{x \mid \|\widehat{a} - x\|_{l_2} < \delta\}$ с центром в точке $\widehat{a}$ радиуса δ, содержащийся во множестве A. Точка $x = \widehat{a} + (\varepsilon, 0, \ldots) \in B(\widehat{a}, \delta)$ при $0 < \varepsilon < \delta$. С другой стороны, $\sum\limits_{k\in\mathbb{N}} x_k = \varepsilon + \sum\limits_{k\in\mathbb{N}} \widehat{a}_k \neq 1$. Это противоречит тому, что $x \in A$.

Покажем, что эти множества нельзя отделить. Предположим противное, что существует ненулевой линейный непрерывный функционал $\lambda = (\lambda_1, \lambda_2, \ldots) \in l_2^*$ такой, что

$$\inf_{a\in A} \langle \lambda, a\rangle \geq \langle \lambda, b\rangle = 0. \qquad (*)$$

Поскольку $\lambda \neq 0$, то $\exists\ \lambda_i \neq \lambda_j$ (для определенности, $\lambda_i < \lambda_j$). Вектор со всеми равными не нулю координатами не принадлежит l_2^*: $(\lambda, \lambda, \ldots) \notin l_2^*$ при $\lambda \neq 0$. Возьмем точку $a = (a_1, a_2, \ldots)$, со следующими компонентами: $a_i = n$, $a_j = 1 - n$, $a_k = 0$ при $k \neq i, j$. Тогда точка $a \in A$ и

$\langle \lambda, a\rangle = \lambda_i a_i + \lambda_j a_j = \lambda_i n + \lambda_j(1 - n) = n(\lambda_i - \lambda_j) + \lambda_j \to -\infty$

при $n \to +\infty$. Приходим к противоречию с неравенством $(*)$. Значит, выпуклое множество A и точка b не отделимы.

Пример 4. *Компактный эллипсоид* $A = \left\{ a \in l_2 \,\middle|\, \sum\limits_{k\in\mathbb{N}} k^2 a_k^2 \le 1 \right\}$ *и луч* $B = \left\{ b = \beta\left(1, \frac{1}{2}, \ldots, \frac{1}{k}, \ldots\right) \mid \beta > 0 \right\}$ *нельзя отделить.*

Решение. Поскольку функция $f(a) = \sum k^2 a_k^2$ является выпуклой, то в силу неравенства Йенсена множество $A = \{a \mid f(a) \le 1\}$ является выпуклым. Действительно, если две точки $a, a' \in A$, то

$$f(ta + (1-t)a') \le tf(a) + (1-t)f(a') \le t + (1-t) = 1.$$

Значит, выпуклая комбинация $ta + (1-t)a' \in A$.

Покажем, что множество A имеет пустую внутренность. Предположим противное, что внутренность A непуста. Тогда существует шар $B(\widehat{a}, \delta) = \{x \mid \|\widehat{a} - x\|_{l_2} < \delta\}$ с центром в точке $\widehat{a}$ радиуса δ, содержащийся во множестве A. Точка $x = \widehat{a} + \varepsilon e_n \in B(\widehat{a}, \delta)$ $\forall\ |\varepsilon| < \delta$ $(e_n = (0, \ldots, 0, 1, 0, \ldots))$. С другой стороны, для точки x

$$\sum_{k\in\mathbb{N}} k^2 x_k^2 = \sum_{\substack{k\in\mathbb{N}\\ k\ne n}} k^2 \widehat{a}_k^2 + n^2(\widehat{a}_n + \varepsilon)^2 \to +\infty \text{ при } n \to +\infty.$$

А это противоречит тому, что $x \in A$.

Очевидно, что луч B является выпуклым множеством в l_2 и имеет пустую внутренность.

Множества A и B не пересекаются, так как точки луча B $\beta\left(1, \frac{1}{2}, \ldots, \frac{1}{k}, \ldots\right) \notin A$ при $\beta > 0$.

Покажем, что эти множества нельзя отделить. Возьмем произвольный функционал $\lambda \in l_2^*$, $\lambda \ne 0$. Тогда

$$\sup_{b\in B} \langle \lambda, b \rangle = \sup_{\beta>0} \beta \sum_{k\in\mathbb{N}} \frac{\lambda_k}{k} \ge 0.$$

В то время как $\inf\limits_{a\in A} \langle \lambda, a \rangle = \inf\limits_{a\in A} \sum\limits_{k\in\mathbb{N}} \lambda_k a_k < 0$, так как компоненты a_k можем брать с любыми знаками. Неравенство $\inf\limits_{a\in A} \langle \lambda, a \rangle \ge \sup\limits_{b\in B} \langle \lambda, b \rangle$ не может выполняться, значит, множества нельзя отделить.

4.2.2 Теоремы о строгой отделимости

Теорема 4.7 (о строгой отделимости точки от множества в нормированном пространстве)**.** *Пусть A — непустое выпуклое замкнутое множество в X, $b \notin A$. Тогда точку b можно строго отделить от множества A.*

Доказательство. Поскольку точка $b \notin A$ — замкнутому множеству, то существует $\varepsilon > 0$, для которого $A \cap (b + \varepsilon U) = \emptyset$, где $U := \{x \in X \mid \|x\| \leq 1\}$ — единичный шар пространства X. Следовательно, по теореме отделимости в нормированном пространстве выпуклые множества A и $B := b + \varepsilon U$ ($\operatorname{int} B \neq \emptyset$, $\operatorname{int} B \cap A = \emptyset$) можно отделить, т. е., существует линейный непрерывный функционал $\lambda \in X^*$, $\lambda \neq 0$ такой, что

$$\inf_{a \in A} \langle \lambda, a \rangle \geq \sup_{x \in b + \varepsilon U} \langle \lambda, x \rangle = \sup_{u \in U} \langle \lambda, b + \varepsilon u \rangle =$$

$$= \langle \lambda, b \rangle + \varepsilon \sup_{\|u\| \leq 1} \langle \lambda, u \rangle = \langle \lambda, b \rangle + \varepsilon \|\lambda\|_{X^*} > \langle \lambda, b \rangle.$$

Отсюда $\inf\limits_{a \in A} \langle \lambda, a \rangle > \langle \lambda, b \rangle$. Это и означает, что точка b строго отделена от множества A. □

Рассмотрим следующее обобщение доказанной теоремы.

Теорема 4.8 (о строгой отделимости компакта)**.** *A и B — непустые выпуклые замкнутые подмножества нормированного пространства X, причём B — компакт и $A \cap B = \emptyset$. Тогда множества A и B строго отделимы.*

Приведем еще одно доказательство теоремы о строгой отделимости компакта, использующее лемму о замкнутости суммы (разности) замкнутого множества и компакта, которая также имеет самостоятельный интерес.

Лемма (о замкнутости суммы (разности) замкнутого множества и компакта)**.** *Пусть X — нормированное пространство, множества $A, B \subset X$, A — замкнуто, B — компакт. Тогда алгебраическая сумма $C = A + B$ (разность $A - B$) замкнута.*

Второе доказательство теоремы о строгой отделимости компакта

Доказательство. Обозначим $C := A - B$. Тогда множество C замкнуто по лемме о замкнутости суммы (разности) замкнутого множества и компакта и $0 \notin C$. Следовательно, по теореме о строгой отделимости точки от выпуклого замкнутого множества точку 0 можно строго отделить от множества C, т. е., существует линейный непрерывный функционал $\lambda \in X^*$, $\lambda \neq 0$, такой, что

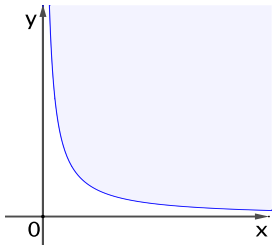
$$\inf_{c \in C} \langle \lambda, c \rangle > 0 \iff \inf_{a \in A, b \in B} \langle \lambda, a - b \rangle > 0$$

$$\inf_{a \in A} \langle \lambda, a \rangle - \sup_{b \in B} \langle \lambda, b \rangle > 0 \iff \inf_{a \in A} \langle \lambda, a \rangle > \sup_{b \in B} \langle \lambda, b \rangle.$$

Это и означает, что множества A и B строго отделимы. □

Замечание. *Требование компактности одного из множеств в теореме о строгой отделимости существенно.*

Действительно, непересекающиеся замкнутые множества ось x и множество точек над гиперболой $y = \dfrac{1}{x}$, $x > 0$, не пересекаются, но не являются строго отделимыми.



4.3 Выпуклые задачи

4.3.1 Задачи без ограничений

Пусть $f\colon X \to \overline{\mathbb{R}}$ — выпуклая функция, отображающая нормированное пространство X в расширенную прямую. *Выпуклой задачей без ограничений* называется следующая экстремальная задача:

$$f(x) \to \min\,. \tag{P}$$

Теорема (аналог теоремы Ферма)**.** *Для того, чтобы точка $\widehat{x}$ доставляла в выпуклой задаче без ограничений (P) абсолютный минимум ($\widehat{x} \in \operatorname{absmin} P$), необходимо и достаточно, чтобы выполнялось соотношение*

$$0 \in \partial f(\widehat{x}).$$

Доказательство. Имеем эквивалентные утверждения:

$$\widehat{x} \in \operatorname{absmin} P \iff f(x) \geq f(\widehat{x}) \iff f(x) - f(\widehat{x}) \geq 0 \iff$$
$$f(x) - f(\widehat{x}) \geq \langle 0, x - \widehat{x}\rangle \ \forall\, x \in X \iff 0 \in \partial f(\widehat{x}). \qquad \square$$

Поскольку из выпуклости функции f не следует, вообще говоря, выпуклость функции $-f$, то существенно, что рассматривается задача на минимум, а не на максимум.

4.3.2 Задачи с ограничением

Пусть $f\colon X \to \overline{\mathbb{R}}$ — выпуклая функция, отображающая нормированное пространство X в расширенную прямую, $A \subset X$ — выпуклое множество. *Выпуклой задачей с ограничением* (или просто *выпуклой задачей*) называется следующая экстремальная задача:

$$f(x) \to \min; \quad x \in A. \tag{P}$$

Теорема. *Пусть $\widehat{x} \in \operatorname{locmin} P$ — доставляет локальный минимум в выпуклой задаче (P). Тогда $\widehat{x} \in \operatorname{absmin} P$ — доставляет абсолютный минимум в этой задаче.*

Доказательство. Пусть $\widehat{x} \in \operatorname{locmin} P$. Это означает, что существует окрестность U точки $\widehat{x}$ такая, что

$$f(\widehat{x}) \le f(x) \quad \forall\, x \in U \cap A. \tag{$*$}$$

Возьмем произвольную допустимую точку $x \in A$. Поскольку x и $\widehat{x}$ из A, то их выпуклая комбинация $x_\alpha = \alpha x + (1-\alpha)\widehat{x} \in U \cap A$ при достаточно малом $\alpha > 0$. Следовательно,

$$f(\widehat{x}) \overset{(*)}{\le} f(x_\alpha) \overset{\operatorname{def} x_\alpha}{=} f(\alpha x + (1-\alpha)\widehat{x}) \le \text{(по неравенству Иенсена)}$$
$$\le \alpha f(\widehat{x}) + (1-\alpha) f(x) = f(\widehat{x}) + \alpha\big(f(x) - f(\widehat{x})\big),$$

откуда

$$\alpha\big(f(x) - f(\widehat{x})\big) \ge 0 \overset{\alpha>0}{\Leftrightarrow} f(x) - f(\widehat{x}) \ge 0 \Leftrightarrow f(\widehat{x}) \le f(x)\ \forall\, x \in A.$$

Значит, $\widehat{x}$ доставляет абсолютный минимум в задаче (P). $\square$

Из теоремы следует, что в выпуклой задаче локальный минимум является и абсолютным (глобальным). Поэтому в дальнейшем в выпуклых задачах, говоря “минимум”, имеем в виду абсолютный минимум.

4.3.3 Задача выпуклого программирования

Пусть $f_i\colon X \to \overline{\mathbb{R}}$, $i = 0, 1, \ldots, m$, — выпуклые функции, отображающие линейное нормированное пространство X в расширенную прямую, $A \subset X$ — выпуклое множество. *Задачей выпуклого программирования* называется следующая экстремальная задача:

$$f_0(x) \to \min; \quad f_i(x) \le 0,\ i = 1, \ldots, m,\ x \in A. \qquad (P)$$

Точка x называется *допустимой* в задаче, если $x \in A$ и выполняются все заданные ограничения типа неравенств.

При проверке достаточных условий минимума в задаче выпуклого программирования будет использоваться некоторое условие регулярности множества допустимых элементов — условие Слейтера[15].

Множество допустимых элементов в задаче (P) удовлетворяет *условию Слейтера*, если существует точка $\overline{x} \in A$, для которой $f_i(\overline{x}) < 0$, $i = 1, \ldots, m$.

Теорема (Каруш[16]–Кун[17]–Таккер[18]). **1.** *Если* $\widehat{x} \in \operatorname{absmin} P$ — *решение задачи выпуклого программирования, то найдется вектор множителей Лагранжа* $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$, $\lambda \ne 0$, *такой, что для функции Лагранжа* $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ *выполняются условия:*

a) принцип минимума функции Лагранжа:

$$\min_{x \in A} \mathscr{L}(x) = \mathscr{L}(\widehat{x});$$

b) дополняющей нежёсткости: $\lambda_i f_i(\widehat{x}) = 0,\ i = 1, \ldots, m$;

c) неотрицательности: $\lambda_i \ge 0,\ i = 0, 1, \ldots, m$.

---

[15] Слейтер, Джон Кларк (1900–1976) — американский физик, химик и математик.

[16] Каруш, Вильям (1917–1997) — американский математик.

[17] Кун, Гарольд Уильям (1925–2014) — американский математик.

[18] Таккер, Альберт (1905–1995) — канадский математик.

2. *Если для допустимой точки* $\widehat{x}$ *выполнены условия* $a) - c)$ *и* $\lambda_0 \neq 0$, *то* $\widehat{x} \in \operatorname{absmin} P$.

3. *Если для допустимой точки* $\widehat{x}$ *выполнены условия* $a) - c)$ *и множество допустимых элементов удовлетворяет условию Слейтера, то* $\widehat{x} \in \operatorname{absmin} P$.

Доказательство. Пусть $\widehat{x} \in \operatorname{absmin} P$. Не ограничивая общности, считаем, что $f_0(\widehat{x}) = 0$ — иначе введем новую функцию $\widetilde{f}_0(x) = f_0(x) - f_0(\widehat{x})$. Положим

$$B = \{b = (b_0, \ldots, b_m) \in \mathbb{R}^{m+1} \mid \exists\, x \in A : f_i(x) \leq b_i,\ i = 0, \ldots, m\}.$$

Покажем, что B — *непустое выпуклое множество.* Действительно, неотрицательный октант $\mathbb{R}^{m+1}_+ \subset B$, т. е. любой вектор с неотрицательными компонентами принадлежит множеству B, ибо в определении B можно положить $x = \widehat{x}$. Докажем выпуклость множества B. Пусть точки b и b' принадлежат множеству B. Надо доказать, что выпуклая комбинация точек b и b' точка $tb + (1-t)b' \in B$ для любого $t \in (0; 1)$. Поскольку точки $b, b' \in B$, то по определению множества B существуют точки $x, x' \in A$ такие, что $f_i(x) \leq b_i$, $f_i(x') \leq b'_i$, $i = 0, 1, \ldots, m$. Положим $x_t = tx + (1-t)x'$. Тогда $x_t \in A$, поскольку множество A — выпукло, а ввиду выпуклости функций f_i, $i = 0, 1, \ldots, m$,

$$\begin{aligned} f_i(x_t) = f_i\big(tx + (1-t)x'\big) &\leq (\text{по неравенству Иенсена}) \\ &\leq t f_i(x) + (1-t) f_i(x') \leq t b_i + (1-t) b'_i, \end{aligned}$$

т. е. точка $tb + (1-t)b' \in B$.

Докажем, что точка $0 \notin \operatorname{int} B$. Если бы $0 \in \operatorname{int} B$, то при достаточно малом $\varepsilon < 0$ точка $(\varepsilon, 0, \ldots, 0) \in B$. Тогда по определению множества B существует точка $\widetilde{x} \in A$ такая, что $f_0(\widetilde{x}) \leq \varepsilon < 0 = f_0(\widehat{x})$, $f_i(\widetilde{x}) \leq 0$, $i = 1, \ldots, m$. Значит, $\widetilde{x}$ является допустимой точкой ($\widetilde{x} \in D(P)$) и $f_0(\widetilde{x}) < f_0(\widehat{x})$, т. е. $\widehat{x} \notin \operatorname{absmin} P$ — противоречие. Следовательно, $0 \notin \operatorname{int} B$.

По теореме отделимости в конечномерном пространстве непустое выпуклое множество B и точку 0 можно отделить, т. е. существует вектор $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \neq 0$, для которого

$$\inf_{b \in B} \langle \lambda, b \rangle \geq \langle \lambda, 0 \rangle \iff \sum_{i=0}^{m} \lambda_i b_i \geq 0 \quad \forall\, b \in B. \qquad (*)$$

Из неравенства $(*)$ будут выведены условия неотрицательности, дополняющей нежёсткости и принцип минимума.

1. *Условие неотрицательности.* Поскольку, как мы уже говорили, любой вектор с неотрицательными компонентами принадлежит B, то вектор $(0, \ldots, 0, 1, 0, \ldots, 0) \in B$, где единица стоит на i-ом месте (счет начинаем с нуля).
Подставив эту точку в неравенство $(*)$, получим, что $\lambda_i \geq 0$.

Условие дополняющей нежёсткости. Нетрудно видеть, что точка $(0, \ldots, 0, f_i(\widehat{x}), 0, \ldots, 0) \in B$. Действительно, в определении множества B возьмем $x = \widehat{x}$, тогда $x \in A$ и нужные неравенства выполняются: $f_j(\widehat{x}) \leq 0$, $j \neq i$, $f_i(\widehat{x}) \leq f_i(\widehat{x})$.
Подставив эту точку в неравенство $(*)$, имеем $\lambda_i f_i(\widehat{x}) \geq 0$. Если $\lambda_i f_i(\widehat{x}) > 0$, то (так как по уже доказанному условию неотрицательности $\lambda_i \geq 0$) $\lambda_i > 0 \Rightarrow f_i(\widehat{x}) > 0$ — это противоречит допустимости точки $\widehat{x}$. Значит, $\lambda_i f_i(\widehat{x}) = 0$.

Принцип минимума. Возьмем точку $x \in A$, тогда точка $(f_0(x),\, f_1(x),\, \ldots, f_m(x)) \in B$. Поэтому по неравенству $(*)$

$$\mathscr{L}(x) = \sum_{i=0}^{m} \lambda_i f_i(x) \geq 0 = \sum_{i=0}^{m} \lambda_i f_i(\widehat{x}) = \mathscr{L}(\widehat{x}),$$

т. к., не ограничивая общности, положили $f_0(\widehat{x}) = 0$ и по уже доказанному условию дополняющей нежёсткости $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \ldots, m$. Таким образом, принцип минимума для функции Лагранжа доказан.

2. Пусть для допустимой точки $\widehat{x}$ выполнены условия $a) - c)$ и $\lambda_0 \neq 0$. Положим $\lambda_0 = 1$. Тогда для любой допустимой точки x будет выполняться неравенство

$$f_0(\widehat{x}) \overset{b)}{=} f_0(\widehat{x}) + \sum_{i=1}^{m} \lambda_i f_i(\widehat{x}) \overset{\text{def}}{=}$$
$$= \mathscr{L}(\widehat{x}) \overset{a)}{\leq} \mathscr{L}(x) \overset{\text{def}}{=} f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \overset{c)}{\leq} f_0(x)$$

(в последней сумме все слагаемые $\lambda_i f_i(x)$ неположительны: $\lambda_i \geq 0$, $f_i(x) \leq 0$). Неравенство $f_0(\widehat{x}) \leq f_0(x)$ для любой допустимой точки x означает, что $\widehat{x} \in \operatorname{absmin} P$.

3. Пусть для допустимой точки $\widehat{x}$ выполнены условия $a) - c)$ и множество допустимых элементов удовлетворяет условию Слейтера в точке $\overline{x}$ ($\exists\ \overline{x} \in A : f_i(\overline{x}) < 0$, $i = 1, \ldots, m$). Предположим, что при этом $\lambda_0 = 0$. Так как вектор $\lambda \neq 0$, то в силу условия неотрицательности существует хотя бы одно $\lambda_j > 0$, $j \in \{1, \ldots, m\}$. Следовательно, имеем

$$\mathscr{L}(\overline{x}) = \sum_{i=1}^{m} \lambda_i f_i(\overline{x}) \overset{c)}{\leq} \lambda_j f_j(\overline{x}) < 0 \overset{b)}{=} \sum_{i=1}^{m} \lambda_i f_i(\widehat{x}) = \mathscr{L}(\widehat{x})$$

отметим, что в первой сумме все слагаемые $\lambda_i f_i(\overline{x})$ в силу условия c) неположительны: $\lambda_i \geq 0$, а $f_i(\overline{x}) < 0$). Но полученное неравенство $\mathscr{L}(\overline{x}) < \mathscr{L}(\widehat{x})$ противоречит выполненному условию a). Значит, наше предположение, что $\lambda_0 = 0$ неверно. Поэтому $\lambda_0 \neq 0$ и по доказанному п. 2 $\widehat{x} \in \operatorname{absmin} P$.

Теорема Каруша–Куна–Таккера полностью доказана. $\square$

4.3.4 Пример

$$f(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2 + 3|x_1 + x_2 - 2| \to \min.$$

Решение. Функция f является выпуклой функцией как сумма двух выпуклых функций. Действительно, обозначим $g(x) = x_1^2+x_1x_2+x_2^2$ и $h(x) = x_1+x_2-2$. Тогда функция g выпукла по Теореме 4.3 п. 4.1.2, так как по критерию Сильвестра матрица вторых производных

$$g''(x) = \left(\frac{\partial^2 g(x)}{\partial x_i \partial x_j}\right)_{i,j=1}^{2} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

положительно определена для любого x. Функция $|h|$, являющаяся модулем аффинной функции, также является выпуклой (см. п. 4.1.2).

Необходимое и достаточное условие минимума в выпуклой задаче без ограничений:

$$\widehat{x} \in \operatorname{absmin} \iff 0 \in \partial f(\widehat{x}). \tag{1}$$

Найдём субдифференциалы функций g и $|h|$. Для дифференцируемой функции её субдифференциал совпадает с производной, а производная функции двух переменных есть вектор частных производных. Поэтому $\partial g(x) = (g_{x_1}, g_{x_2}) = (2x_1 + x_2, x_1 + 2x_2)$.

По формуле субдифференциала модуля (см. п. 4.1.4)

$$\partial|h(x)| = \begin{cases} -h'(x), & h(x) < 0, \\ \{\alpha h'(x),\ \alpha \in [-1;\ 1]\}, & h(x) = 0, \\ h'(x), & h(x) > 0, \end{cases}$$

для $h(x) = x_1 + x_2 - 2$ имеем:

$$\partial|h(x)| = \begin{cases} (-1, -1), & x_1 + x_2 - 2 < 0, \\ (\alpha, \alpha),\ |\alpha| \le 1, & x_1 + x_2 - 2 = 0, \\ (1, 1), & x_1 + x_2 - 2 > 0. \end{cases}$$

Значит, по теореме Моро–Рокафеллара о субдифференциале суммы

$$\partial f(x) = \partial g(x) + 3\partial|h(x)| =$$

$$= (2x_1 + x_2, x_1 + 2x_2) + 3\begin{cases} (-1,-1), & x_1 + x_2 - 2 < 0, \\ (\alpha,\alpha),\ |\alpha| \le 1, & x_1 + x_2 - 2 = 0, \\ (1,1), & x_1 + x_2 - 2 > 0, \end{cases} =$$

$$= \begin{cases} (2x_1 + x_2 - 3, x_1 + 2x_2 - 3), & x_1 + x_2 - 2 < 0, \\ (2x_1 + x_2 + 3\alpha, x_1 + 2x_2 + 3\alpha),\ |\alpha| \le 1, & x_1 + x_2 - 2 = 0, \\ (2x_1 + x_2 + 3, x_1 + 2x_2 + 3), & x_1 + x_2 - 2 > 0. \end{cases}$$

Следовательно, соотношение (1) запишется в виде

$$(0,0) \in \partial f(\widehat{x}) \iff \left[\begin{array}{ll} \begin{cases} 2x_1 + x_2 - 3 = 0, \\ x_1 + 2x_2 - 3 = 0, \\ x_1 + x_2 - 2 < 0, \end{cases} & (i) \\ \begin{cases} 2x_1 + x_2 + 3\alpha = 0, \\ x_1 + 2x_2 + 3\alpha = 0, \\ x_1 + x_2 - 2 = 0 \quad (|\alpha| \le 1), \end{cases} & (ii) \\ \begin{cases} 2x_1 + x_2 + 3 = 0, \\ x_1 + 2x_2 + 3 = 0, \\ x_1 + x_2 - 2 > 0. \end{cases} & (iii) \end{array}\right.$$

В случае (i) нет критических точек, так как, складывая два первых уравнения системы и деля на 3, получаем, что $x_1 + x_2 - 2 = 0$ — противоречие с условием $x_1 + x_2 - 2 < 0$.

В случае (iii) также нет критических точек, так как из уравнений системы следует, что $x_1+x_2+2 = 0$ — противоречие с условием $x_1 + x_2 - 2 > 0$.

В случае (ii) система из трёх уравнений с тремя неизвестными имеет единственное решение $\widehat{x}_1 = \widehat{x}_2 = 1$, $\alpha = -1$.

Ответ. $\widehat{x} = (\widehat{x}_1, \widehat{x}_2) = (1,1) \in \mathrm{absmin}$, $S_{\mathrm{absmin}} = 3$.

5 Элементы функционального анализа

5.1 Линейные, нормированные, банаховы пространства

Определение 1. Последовательность точек $\{x_n\}_{n=1}^{\infty}$ метрического пространства называется *фундаментальной*, если она удовлетворяет *критерию Коши*, т. е. если для любого $\varepsilon > 0$ существует N_ε такое, что $d(x_{n_1}, x_{n_2}) < \varepsilon$ для всех $n_1, n_2 > N_\varepsilon$.

Определение 2. Метрическое пространство называется *полным*, если любая фундаментальная последовательность сходится.

Определение 3. Полное нормированное пространство называется *банаховым* пространством.

5.1.1 Сопряженное пространство, сопряженный оператор

Совокупность X^* всех линейных непрерывных функционалов на нормированном пространстве X образует *сопряженное* к X пространство. Оно является банаховым пространством относительно нормы

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

где $\langle x^*, x \rangle$ означает действие функционала x^* на элемент x.

Пусть X и Y — нормированные пространства. Обозначим через $L(X, Y)$ пространство линейных непрерывных операторов из X в Y.

Для оператора $\Lambda \in L(X, Y)$ можно определить *сопряженный оператор* $\Lambda^* \colon Y^* \to X^*$, действующий по формуле

$$\langle \Lambda^* y^*, x \rangle = \langle y^*, \Lambda x \rangle \quad \forall\, x \in X.$$

5.2 Определения производных

Для вещественнозначных функций одного вещественного переменного два определения — существование конечного предела

$$\lim_{h\to 0}\frac{f(\widehat{x}+h)-f(\widehat{x})}{h} \tag{1}$$

и возможность асимптотического разложения при $h\to 0$

$$f(\widehat{x}+h)=f(\widehat{x})+f'(\widehat{x})h+o(h) \tag{2}$$

— приводят к одному и тому же понятию дифференцируемости. Но уже для функций двух и большего числа переменных существует несколько различных подходов к понятию дифференцируемости (гладкости). Определение (1), ведущее начало от Ньютона, ведет к понятиям производной по направлению, вариации по Лагранжу и производной Гато. Определение (2), ведущее начало от Лейбница, ведет к понятиям производной Гато, Фреше и строгой дифференцируемости.

Пусть далее в этом пункте X, Y, Z — линейные нормированные пространства. Как правило (если это не оговорено иначе), $f\colon X\to Y$ — отображение пространства X или некоторой окрестности точки $\widehat{x}\in X$ в пространство Y.[19]

5.2.1 Производная по направлению

Будем говорить, что отображение f имеет в точке $\widehat{x}$ *производную по направлению* h, если существует предел справа

$$\lim_{\lambda\to +0}\frac{f(\widehat{x}+\lambda h)-f(\widehat{x})}{\lambda}=:\delta_+ f(\widehat{x},h),$$

который обозначается $\delta_+ f(\widehat{x},h)$. Плюс в индексе здесь указывает на то, что берется именно предел справа.

[19] Но вполне содержателен пример, когда $X=\mathbb{R}^n$, $Y=\mathbb{R}^m$. Элемент из $L(X,Y)$ определяется в этом случае матрицей размера $n\times m$.

5.2.2 Вариация по Лагранжу

Если для отображения f в точке $\widehat{x}$ существует предел

$$\delta f(\widehat{x}, h) := \lim_{\lambda \to 0} \frac{f(\widehat{x} + \lambda h) - f(\widehat{x})}{\lambda} \quad \forall\, h \in X,$$

то говорят, что отображение f *имеет в точке* $\widehat{x}$ *вариацию по Лагранжу по направлению* h. При этом отображение $h \to \delta f(\widehat{x}, h)$ называют *вариацией по Лагранжу*, $\delta f(\widehat{x}, \cdot)\colon X \to Y$.

Таким образом, для существования вариации по Лагранжу необходимо существование в точке $\widehat{x}$ производной по всем направлениям h и выполнения равенств

$\delta_+ f(\widehat{x}, h) = -\delta_+ f(\widehat{x}, -h)\ \forall\, h \in X$.

5.2.3 Производная Гато

Если оператор вариации по Лагранжу $\delta f(\widehat{x}, \cdot)\colon X \to Y$ линеен и непрерывен по h ($\delta f(\widehat{x}, \cdot) \in L(X, Y)$), то говорят, что отображение f *дифференцируемо по Гато в точке* $\widehat{x}$, а оператор $\delta f(\widehat{x}, \cdot)$ называется *производной Гато отображения* f *в точке* $\widehat{x}$ и обозначается $f'_G(\widehat{x})$.

Отсюда следует, что если отображение f дифференцируемо по Гато в точке $\widehat{x}$, то для любого фиксированного вектора h имеет место разложение

$$f(\widehat{x} + \lambda h) = f(\widehat{x}) + \lambda f'_G(\widehat{x})[h] + r(h, \lambda),$$

где $\|r(h, \lambda)\|_Y = o(|\lambda|)$ при $\lambda \to 0$. Через $f'_G(\widehat{x})[h]$ обозначен образ действия оператора $f'_G(\widehat{x})$ на элемент h.

Отметим, что отображение, дифференцируемое по Гато в точке $\widehat{x}$, не обязано быть непрерывным в этой точке (см. Пример 4, п. 5.2.8).

5.2.4 Производная Фреше

Отображение f называют *дифференцируемым по Фреше в точке* $\widehat{x}$ и пишут $f \in D(\widehat{x})$, если существуют линейный непрерывный оператор $f'(\widehat{x})\colon X \to Y$ и отображение $r\colon X \to Y$ такие, что

$$f(\widehat{x}+h) = f(\widehat{x}) + f'(\widehat{x})[h] + r(h), \qquad (*)$$

где $\|r(h)\|_Y = o(\|h\|_X)$ при $\|h\|_X \to 0$. Оператор $f'(\widehat{x})$ называется *производной Фреше*. Разложение $(*)$ можно кратко записать так:

$$f(\widehat{x}+h) = f(\widehat{x}) + f'(\widehat{x})[h] + o(h),$$

понимая $o(h)$ как элемент пространства Y, для которого $\|o(h)\|_Y = o(\|h\|)$ при $\|h\| \to 0$.

Из разложения $(*)$ следует, что функция, дифференцируемая по Фреше, непрерывна в точке дифференцируемости, а также дифференцируема в этой точке по Гато. Уже в двумерном случае эти два понятия различаются: из дифференцируемости функции по Гато не следует её дифференцируемость по Фреше (см. Пример 4, п. 5.2.8). Аналогично из дифференцируемости по Гато по определению вытекает существование вариации по Лагранжу. И снова (уже в двумерном случае) эти понятия различны. Отметим, что производная Фреше линейного оператора совпадает с самим оператором.

Если в каждой точке x открытого множества U отображение $f \in D(x)$, а отображение $x \to f'(x)$ непрерывно, то пишем $f \in C^1(U)$.

На языке ε–δ определение дифференцируемости по Фреше отображения f в точке $\widehat{x}$ формулируется так: существует оператор $f'(\widehat{x}) \in L(X,Y)$ такой, что для любого $\varepsilon > 0$ найдется $\delta > 0$, при котором для любого $\|h\| < \delta$ выполняется неравенство

$$\|f(\widehat{x}+h) - f(\widehat{x}) - f'(\widehat{x})[h]\|_Y \le \varepsilon\|h\|_X.$$

Из разложения $(*)$ следует, что производная Фреше определена однозначно, ибо

$$\begin{cases} f(\widehat{x}+h) = f(\widehat{x}) + \Lambda_1[h] + r_1(h), \\ f(\widehat{x}+h) = f(\widehat{x}) + \Lambda_2[h] + r_2(h), \end{cases} \Longrightarrow \Lambda_1 h - \Lambda_2 h = o(h),$$

а последнее равенство для линейных непрерывных операторов Λ_1 и Λ_2 возможно лишь при $\Lambda_1 = \Lambda_2$. Чтобы убедиться в этом достаточно разделить обе части равенства на $\|h\|$ и устремить h к нулю.

Для функций одной переменной производная по Фреше совпадает с обычной производной, которая изучается в курсе математического анализа.

Пример 1. Для функции нескольких переменных $f\colon \mathbb{R}^n \to \mathbb{R}$ производная в точке $\widehat{x} = (\widehat{x}_1, \ldots, \widehat{x}_n)$ является вектором $f'(\widehat{x}) = \left(\frac{\partial f(\widehat{x})}{\partial x_1}, \ldots, \frac{\partial f(\widehat{x})}{\partial x_n}\right)$. Оператор производной Фреше определяется следующим образом:

$$f'(\widehat{x})[h] = \langle f'(\widehat{x}), h\rangle = \left\langle \left(\frac{\partial f(\widehat{x})}{\partial x_1}, \ldots, \frac{\partial f(\widehat{x})}{\partial x_n}\right), (h_1, \ldots, h_n) \right\rangle = \sum_{i=1}^{n} \frac{\partial f(\widehat{x})}{\partial x_i} h_i.$$

Пример 2. Для отображений $f\colon \mathbb{R}^n \to \mathbb{R}^m$, где $f(x) = \big(f_1(x), \ldots, f_m(x)\big)$, производная в точке $\widehat{x} = (\widehat{x}_1, \ldots, \widehat{x}_n)$ является матрицей $f'(\widehat{x}) = \begin{pmatrix} \frac{\partial f_1(\widehat{x})}{\partial x_1} & \ldots & \frac{\partial f_1(\widehat{x})}{\partial x_n} \\ \ldots & \ldots & \ldots \\ \frac{\partial f_m(\widehat{x})}{\partial x_1} & \ldots & \frac{\partial f_m(\widehat{x})}{\partial x_n} \end{pmatrix}$. Оператор производной Фреше определяется следующим образом:

$$f'(\widehat{x})[h] = f'(\widehat{x})h = \begin{pmatrix} \frac{\partial f_1(\widehat{x})}{\partial x_1} & \ldots & \frac{\partial f_1(\widehat{x})}{\partial x_n} \\ \ldots & \ldots & \ldots \\ \frac{\partial f_m(\widehat{x})}{\partial x_1} & \ldots & \frac{\partial f_m(\widehat{x})}{\partial x_n} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

5.2.5 Строгая дифференцируемость

Во многих задачах конечномерного и бесконечномерного анализа дифференцируемости по Фреше в точке недостаточно для получения содержательного результата. Это побуждает к следующему усилению дифференцируемости в точке.

Пусть отображение f дифференцируемо по Фреше в точке $\widehat{x}$. Оно называется *строго дифференцируемым в точке* $\widehat{x}$ (при этом пишут $f \in SD(\widehat{x})$), если для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что для всех x_1 и x_2, удовлетворяющих неравенствам $\|x_1 - \widehat{x}\| < \delta$, $\|x_2 - \widehat{x}\| < \delta$, выполнено неравенство

$$\|f(x_1) - f(x_2) - f'(\widehat{x})[x_1 - x_2]\|_Y \leq \varepsilon \|x_1 - x_2\|_X.$$

Из определения следует, что функция, строго дифференцируемая по Фреше в точке, непрерывна в некоторой окрестности этой точки.

Ниже будет показано, что если производная Фреше (даже производная Гато) отображения непрерывна в точке, то отображение будет строго дифференцируемо в этой точке.

5.2.6 Частные производные

Пусть X, Y, Z — линейные нормированные пространства. Рассмотрим отображение $f\colon X \times Y \to Z$, $(\widehat{x}, \widehat{y}) \in X \times Y$.

Если отображение $x \to f(x, \widehat{y})$ дифференцируемо в точке $\widehat{x}$ по Фреше, то его производная называется *частной производной по x отображения f в точке* $(\widehat{x}, \widehat{y})$ и обозначается $f'_x(\widehat{x}, \widehat{y})$ или $\dfrac{\partial f(\widehat{x}, \widehat{y})}{\partial x}$.

Аналогично определяется частная производная по y

$f'_y(\widehat{x}, \widehat{y}) = \dfrac{\partial f(\widehat{x}, \widehat{y})}{\partial y}$.

5.2.7 Производные высших порядков

Дадим теперь определение второй производной Фреше. Если отображение $f\colon X \to Y$ дифференцируемо в каждой точке $x \in X$, то определено отображение $x \to f'(x)$ пространства X в пространство $L(X,Y)$. Поскольку $L(X,Y)$ также является линейным нормированным пространством, то можно определять производную Фреше этого отображения $f'\colon X \to L(X,Y)$, которая и называется *второй производной* отображения f:

$$f''(\widehat{x}) = (f')'(\widehat{x}) \in L\big(X, L(X,Y)\big).$$

Для вектора $h_1 \in X$ определен оператор $f''(\widehat{x})[h_1] \in L(X,Y)$. Для $h_2 \in X$ определен элемент $f''(\widehat{x})[h_1][h_2] \in Y$. Тем самым определено отображение $f''(\widehat{x})[h_1,h_2] := f''(\widehat{x})[h_1][h_2]$. Значит, определено билинейное (линейное по каждому аргументу) отображение $f''(\widehat{x})\colon X \times X \to Y$.

Аналогично определяются производные высших порядков.

Теорема (о смешанных производных). *Если отображение f дважды дифференцируемо в точке $\widehat{x}$ $\big(f \in D^2(\widehat{x})\big)$, то*

$$f''(\widehat{x})[h_1,h_2] = f''(\widehat{x})[h_2,h_1] \quad \forall\ h_1,h_2 \in X.$$

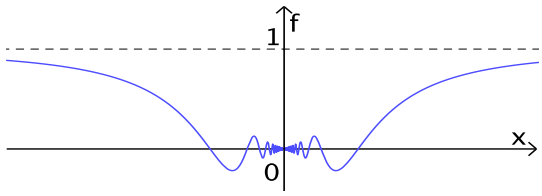
Замечание. Можно считать, что отображения определены (дифференцируемы) не на всем пространстве, а в окрестности рассматриваемых точек.

5.2.8 Контрпримеры на дифференцируемость

Приведем несколько контрпримеров, показывающих, что введенные выше понятия дифференцируемости действительно различны.

Пример 1. *Непрерывная функция не имеет в фиксированной точке производной ни по какому направлению:*

$$f\colon \mathbb{R} \to \mathbb{R},\ f(x) = \begin{cases} x \sin \dfrac{1}{x}, & x \neq 0, \\ 0, & x = 0, \end{cases} \qquad \widehat{x} = 0.$$



На прямой $\mathbb{R}$ имеются с точностью до умножения на положительную константу два направления: $h_1 = 1$ и $h_2 = -1$. Однако производных ни по какому направлению не существует. Действительно, предел

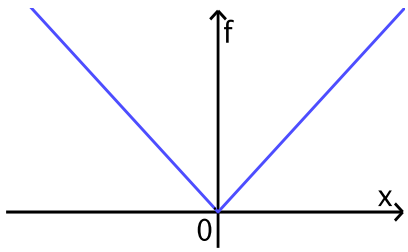
$$\delta_+ f(\widehat{x}, h)\Big|_{\substack{\widehat{x}=0 \\ h=1}} = \lim_{\lambda \to +0} \frac{f(\widehat{x} + \lambda h) - f(\widehat{x})}{\lambda}\Bigg|_{\substack{\widehat{x}=0 \\ h=1}} =$$

$$= \lim_{\lambda \to +0} \frac{\lambda \sin \frac{1}{\lambda}}{\lambda} = \lim_{\lambda \to +0} \sin \frac{1}{\lambda}$$

не существует. В силу чётности функции f не существует также предел по направлению $h = -1$.

Пример 2. *Непрерывная функция имеет в фиксированной точке производную по всем направлениям, но не имеет в этой точке вариации по Лагранжу:*

$$f\colon \mathbb{R} \to \mathbb{R},\ f(x) = |x|,\ \widehat{x} = 0.$$



Как и в примере 1 на прямой $\mathbb{R}$ имеются с точностью до умножения на положительную константу два направления: $h_1 = 1$ и $h_2 = -1$. Пределы по обоим направлениям существуют:

$$\delta_+ f(\widehat{x}, h)\Big|_{\substack{\widehat{x}=0\\ h=h_i}} = \lim_{\lambda\to+0} \frac{f(\widehat{x}+\lambda h) - f(\widehat{x})}{\lambda}\Bigg|_{\substack{\widehat{x}=0\\ h=h_i}} =$$

$$= \lim_{\lambda\to+0} \frac{|\lambda h_i|}{\lambda} = \lim_{\lambda\to+0} \frac{|\lambda|}{\lambda} = 1,\ i = 1, 2.$$

Но предел

$$\delta f(\widehat{x}, h)\Big|_{\substack{\widehat{x}=0\\ h=1}} = \lim_{\lambda\to 0} \frac{f(\widehat{x}+\lambda h) - f(\widehat{x})}{\lambda}\Bigg|_{\substack{\widehat{x}=0\\ h=1}} = \lim_{\lambda\to 0} \frac{|\lambda|}{\lambda}$$

не существует. Следовательно, отображение f не имеет в точке $\widehat{x} = 0$ вариации по Лагранжу.

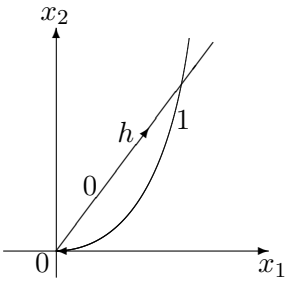
Отметим, что если $\widehat{x} \neq 0$, то $\delta f(\widehat{x}, h) = f'(\widehat{x})[h] = \operatorname{sign}\widehat{x} \cdot h$.

Пример 3. *Отображение имеет в фиксированной точке вариацию по Лагранжу, непрерывно в этой точке, но не имеет в этой точке производной Гато:*

Возьмём отображение $f\colon \mathbb{R}^2 \to \mathbb{R}$, в полярных координатах $x = (x_1, x_2) = (r\cos\varphi,\ r\sin\varphi)$ задающееся по формуле $f(x) = r\cos 3\varphi$.

Пример 4. *Функция f имеет в фиксированной точке производную Гато, но не имеет в этой точке производной Фреше:*

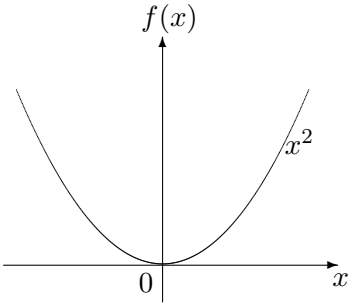
$$f\colon \mathbb{R}^2 \to \mathbb{R},\ f(x) = \begin{cases} 1, & x_2 = x_1^2,\ x_1 > 0, \\ 0, & \textit{в остальных случаях}, \end{cases} \qquad \widehat{x} = (0,0).$$



Так как $\delta f(\widehat{x}, h) = \lim\limits_{\lambda\to 0} \dfrac{f(\widehat{x}+\lambda h) - f(\widehat{x})}{\lambda} = \lim\limits_{\lambda\to 0} \dfrac{f(\lambda h)}{\lambda} = 0$ для любого h, то производная Гато существует и $f'_G(\widehat{x}) = 0$. С другой стороны, функция f разрывна в точке $\widehat{x} = (0,0)$, а функция, дифференцируемая по Фреше, должна быть непрерывна в точке дифференцируемости.

Пример 5. *Функция имеет в фиксированной точке производную Фреше, но не строго дифференцируема в этой точке:*

$$f\colon \mathbb{R} \to \mathbb{R},\ f(x) = \begin{cases} x^2, & x \textit{ рационально}, \\ 0, & x \textit{ иррационально}, \end{cases} \qquad \widehat{x} = 0.$$



Выписанная функция одной (!) переменной дифференцируема в точке $\widehat{x} = 0$. Значит она дифференцируема по Фреше в этой точке.

С другой стороны, функция имеет разрывы в любой окрестности нуля, а строго дифференцируемая функция должна быть непрерывна в некоторой окрестности $\widehat{x}$.

5.3 Некоторые теоремы дифференциального исчисления в нормированных пространствах.

Приведем несколько теорем, наиболее часто используемых в теории экстремальных задач.

5.3.1 Теорема о суперпозиции

Теорема. *Пусть X, Y, Z — линейные нормированные пространства, $\varphi\colon X \to Y$, $\psi\colon Y \to Z$, $f = \psi \circ \varphi\colon X \to Z$ — суперпозиция отображений φ и ψ, $\varphi(\widehat{x}) = \widehat{y}$. Тогда, если отображение ψ дифференцируемо по Фреше в точке $\widehat{y}$, а отображение φ в точке $\widehat{x}$ имеет вариацию по Лагранжу (дифференцируемо по Гато, дифференцируемо по Фреше), то отображение f обладает в точке $\widehat{x}$ тем же свойством, что и отображение φ, и при этом соответственно*

A) $\delta f(\widehat{x}, h) = \psi'(\widehat{y})[\delta\varphi(\widehat{x}, h)] \quad \forall\ h \in X$,

B) $f'_G(\widehat{x}) = \psi'(\widehat{y}) \circ \varphi'_G(\widehat{x})$,

C) $f'(\widehat{x}) = \psi'(\widehat{y}) \circ \varphi'(\widehat{x}) \quad \big(\Leftrightarrow f'(\widehat{x})[h] = \psi'(\widehat{y})[\varphi'(\widehat{x})[h]]\big)$,

D) Если $\psi \in SD(\widehat{y})$, *а* $\varphi \in SD(\widehat{x})$, *то* $f \in SD(\widehat{x})$.

5.3.2 Формула Тейлора

Теорема. *Пусть X, Y — линейные нормированные пространства, отображение $f : X \to Y$, причём $f \in D^n(\widehat{x})$. Тогда имеет место разложение в ряд Тейлора*

$$f(\widehat{x} + h) =$$
$$= f(\widehat{x}) + f'(\widehat{x})[h] + \frac{1}{2} f''(\widehat{x})[h, h] + \ldots + \frac{1}{n!} f^{(n)}(\widehat{x})[h, \ldots, h] + r(h),$$

где $\|r(h)\| = o(\|h\|^n)$ *при* $h \to 0$.

5.4 Дополнительные сведения из алгебры и функционального анализа

В этом пункте приводятся дополнительные сведения из алгебры и функционального анализа, которые понадобятся для доказательства теорем об условиях экстремума в гладких экстремальных задачах в нормированных пространствах.

Определение. *Аннулятором $A^\perp$ множества A линейного нормированного пространства X называется множество линейных непрерывных функционалов x^*, для которых $\langle x^*, x\rangle = 0$ для любого $x \in A$:*

$$A^\perp := \{x^* \in X^* \mid \langle x^*, x\rangle = 0 \quad \forall\, x \in A\}.$$

Отметим, что аннулятор $A^\perp$ всегда содержит нулевой элемент ($x^* = 0$) сопряженного пространства X^*.

Лемма (о нетривиальности аннулятора)**.** *Пусть L — замкнутое собственное* (т. е. $L \neq X$) *подпространство линейного нормированного пространства X. Тогда аннулятор $L^\perp$ содержит ненулевой элемент $x^* \in X^*$.*

Далее нам понадобятся следующие теоремы из функционального анализа и некоторые леммы.

Теорема (Банаха об открытости)**.** *Пусть X, Y — банаховы пространства, Λ — линейный непрерывный оператор из X в Y, являющийся эпиморфизмом*[20]*. Тогда образ каждого открытого множества в X открыт в Y.*

Теорема (Банаха об обратном операторе)**.** *Пусть X, Y — банаховы пространства, Λ — линейный непрерывный оператор из X в Y, являющийся эпиморфизмом и $\operatorname{Ker}\Lambda = 0$. Тогда существуют обратный оператор $\Lambda^{-1}\colon Y \to X$ так же линейный и непрерывный.*

---

[20] $\Lambda\colon X \overset{\text{на}}{\to} Y$

Лемма (Банаха)**.** *Пусть X — линейное нормированное пространство, точка $x_0 \in X$. Тогда существует линейный непрерывный функционал $x^* \in X^*$ такой, что $\|x^*\| = 1$, $\langle x^*, x_0 \rangle = \|x_0\|$.*

Лемма Банаха также является следствием теоремы Хана–Банаха о продолжении линейного функционала.

Лемма (о правом обратном отображении)**.** *Пусть X, Y — банаховы пространства, Λ — непрерывный линейный оператор из X в Y, являющийся эпиморфизмом. Тогда существуют обратное отображение $M\colon Y \to X$* (необязательно непрерывное и необязательно линейное) *и константа $K > 0$ такие, что*

$$\Lambda \circ M = I_Y{}^{21}, \quad \|My\| \le C\|y\| \quad \forall\, y \in Y.$$

Доказательство. Обозначим $B_X := \{x \in X \mid \|x\| < 1\}$ — открытый шар радиуса 1 в пространстве X. По теореме Банаха об открытости образ открытого множества ΛB_X открыт. Поскольку при линейном отображении ноль переходит в ноль, то множество ΛB_X содержит внутренней точкой ноль, и, следовательно, содержит некоторый открытый шар с центром в нуле $\delta B_Y := \{y \in Y \mid \|y\| < \delta\}$.
Значит, для любого $y \in \delta B_Y$ найдется $x(y)$ такой, что

$$\Lambda x(y) = y, \tag{i}$$

$$\|x(y)\| < 1. \tag{ii}$$

Для любого $y \in Y$, $y \neq 0$, обозначим $My := \dfrac{2\|y\|}{\delta} x\Big(\dfrac{\delta y}{2\|y\|}\Big)$.
Тогда $\Big\|\dfrac{\delta y}{2\|y\|}\Big\| = \dfrac{\delta}{2} < \delta$, т. е. $\dfrac{\delta y}{2\|y\|} \in \delta B_Y$ и из определения отображения M имеем:

$$\Lambda \circ My = \Lambda\Big(\frac{2\|y\|}{\delta} x\Big(\frac{\delta y}{2\|y\|}\Big)\Big) = \frac{2\|y\|}{\delta}\Lambda\Big(x\Big(\frac{\delta y}{2\|y\|}\Big)\Big) \overset{(i)}{=}$$
$$= \frac{2\|y\|}{\delta} \cdot \frac{\delta y}{2\|y\|} = y,$$

$$\|My\| = \Big\|\frac{2\|y\|}{\delta} x\Big(\frac{\delta y}{2\|y\|}\Big)\Big\| = \frac{2\|y\|}{\delta} \cdot \Big\|x\Big(\frac{\delta y}{2\|y\|}\Big)\Big\| \overset{(ii)}{\le} \frac{2}{\delta}\|y\|. \qquad \square$$

[21] Через I_Y обозначается тождественный оператор на пространстве Y.

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

Лемма (об аннуляторе ядра регулярного[22] оператора)**.** *Пусть* X, Y *— банаховы пространства,* $A\colon X \to Y$ *— линейный непрерывный эпиморфизм. Тогда* $(\operatorname{Ker} A)^{\perp} = \operatorname{Im} A^*$.

Касательные векторы, конусы, пространства

При выводе условий экстремума в задачах с ограничениями понадобятся понятия касательных векторов, конусов, пространств и теоремы для их нахождения. В этом пункте приводятся теорема об обратном отображении, теорема Люстерника и теорема о касательном пространстве.

Пусть X — линейное нормированное пространство, M — некоторое его подмножество. Вектор $h \in X$ называется *односторонним касательным (полукасательным) вектором к множеству* M *в точке* $\widehat{x} \in X$, если существуют $\varepsilon > 0$ и отображение $r\colon (0; \varepsilon] \to X$, такие, что

a) $\widehat{x} + th + r(t) \in M \ \forall\, t \in (0; \varepsilon]$;

b) $\|r(t)\| = o(t)$ при $t \to +0$.

Вектор h называется *касательным к множеству* M *в точке* $\widehat{x} \in X$, если векторы h и $-h$ являются односторонними касательными векторами к M в точке $\widehat{x}$, т. е. если существуют $\varepsilon > 0$ и отображение $r\colon [-\varepsilon; \varepsilon] \to X$ такие, что

a) $\widehat{x} + th + r(t) \in M \ \forall\, t \in [-\varepsilon; \varepsilon],\ t \neq 0$;

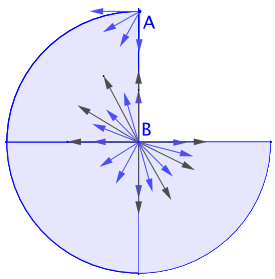
b) $\|r(t)\| = o(t)$ при $t \to 0$.

[22] Оператор называется *регулярным*, если он является линейным непрерывным эпиморфизмом.

Множество всех касательных векторов к M в точке $\widehat{x}$ обозначается $T_{\widehat{x}}M$, множество односторонних касательных векторов обозначается $T^+_{\widehat{x}}M$. Очевидно, что $T_{\widehat{x}}M$ и $T^+_{\widehat{x}}M$ — конусы. Если множество $T_{\widehat{x}}M$ является подпространством в X, то оно называется *касательным пространством к множеству M в точке $\widehat{x}$*.

Если точка $\widehat{x} \in M$, то в обоих определениях t может равняться нулю.

Пример. На круге с вырезанной первой четвертью в точках A и B проведены односторонние касательные (короткие) и касательные векторы (более длинные).



Теорема (об обратном отображении)**.** *Пусть X, Z — банаховы пространства, отображение $F\colon X \to Z$, $F(\widehat{x}) = \widehat{z}$, $F \in SD(\widehat{x})$, оператор $F'(\widehat{x})$ является эпиморфизмом. Тогда существуют обратное отображение $F^{-1}\colon U \subset Z \to X$ некоторой окрестности U точки $\widehat{z}$ и константа $K > 0$ такие, что $F^{-1}(\widehat{z}) = \widehat{x}$ и*

$$F(F^{-1}(z)) = z, \quad \|F^{-1}(z) - F^{-1}(\widehat{z})\| \le K\|z - \widehat{z}\| \quad \forall\, z \in U.$$

Теорема (Люстерник)**.** *Пусть X, Z — банаховы пространства, отображение $F\colon X \to Z$, $F \in SD(\widehat{x})$, оператор $F'(\widehat{x})$ является эпиморфизмом. Тогда существуют отображение $\varphi\colon U \subset X \to X$ некоторой окрестности U точки $\widehat{x}$ и число $C > 0$ такие, что*

$$F\big(x+\varphi(x)\big) = F(\widehat{x}), \;\; \|\varphi(x)\| \le C\|F(x)-F(\widehat{x})\| \;\; \forall\, x \in U. \quad (1.1)$$

Если обозначить $x + \varphi(x) = \psi(x)$, то соотношения (1.1) можно переформулировать в следующем виде:

$$F\big(\psi(x)\big) = F(\widehat{x}), \;\; \|\psi(x) - x\| \le C\|F(x) - F(\widehat{x})\| \;\; \forall\, x \in U. \quad (1.2)$$

Геометрический смысл теоремы Люстерника заключается в том, что если обозначить $M(\widehat{x}) = \{x \in X \mid F(x) = F(\widehat{x})\}$, то справедливо неравенство

$$d(x, M) \le C\|F(x) - F(\widehat{x})\| \quad \forall\, x \in U, \qquad (1.3)$$

где $d(x, M)$ — расстояние от точки x до множества M.

Лемма (Хоффман)**.** *Пусть X, Y — банаховы пространства, оператор $\Lambda \in L(X,Y)$, $\Lambda X = Y$, функционалы $x_i^* \in X^*$, $i=1,\dots,n$, конус $K := \{h \in X \mid \langle x_i^*, h\rangle \le 0,\ i=1,\dots,n,\ \Lambda h = 0\}$. Тогда существует константа $C > 0$ такая, что*[23]

$$d(x,K) \le C\left(\sum_{i=1}^{n} \langle x_i^*, x\rangle_+ + \|\Lambda x\|\right).$$

[23]Напомним, что для числа $a \in \mathbb{R}$ $a_+ \stackrel{\text{def}}{=} \begin{cases} a, & a \ge 0, \\ 0, & a < 0. \end{cases}$

Лемма 1. *Пусть X, Y — банаховы пространства, оператор $\Lambda \in L(X, Y)$, подпространство $\operatorname{Im} A$ замкнуто в Y. Тогда любой вектор $h \in X$ можно представить в виде суммы $h = h_1 + h_2$, где*

$$\Lambda h_1 = 0, \qquad \|h_2\| \le C\|\Lambda h\|$$

с некоторой константой $C > 0$, не зависящей от h.

Доказательство. Замкнутое подпространство банахового пространства само банахово и, значит, линейный непрерывный оператор $\Lambda\colon X \to \widetilde{Y} := \operatorname{Im}\Lambda$ является эпиморфизмом. Тогда по лемме о правом обратном операторе существуют оператор $M\colon \widetilde{Y} \to X$ и константа $C > 0$ такие, что

$$\Lambda M = I_{\widetilde{Y}}, \quad \|My\| \le C\|y\| \quad \forall\, y \in \widetilde{Y}.$$

Для вектора $h \in X$ вектор $\Lambda h \in \widetilde{Y}$. Поэтому определён вектор $h_2 := M(\Lambda h) \in X$. Обозначим $h_1 := h - h_2$. Тогда $h = h_1 + h_2$,

$$\Lambda h_1 = \Lambda(h - h_2) = \Lambda h - \Lambda h_2 = \Lambda h - \Lambda M(\Lambda h) \overset{\Lambda M = I_{\widetilde{Y}}}{=} \Lambda h - \Lambda h = 0,$$

а для вектора h_2 по лемме о правом обратном операторе выполняется оценка: $\|h_2\| = \|M(\Lambda h)\| \le C\|\Lambda h\|$. $\square$

6 Гладкая задача без ограничений

В этом пункте даются необходимые и достаточные условия экстремума функционалов в нормированных пространствах.

6.1 Постановка задачи

Пусть $f\colon X \to \mathbb{R}$ — отображение линейного нормированного пространства X во множество действительных чисел (в этом случае обычно говорим *функционал* на пространстве X), обладающее некоторой гладкостью, т. е. определенными свойствами дифференцируемости.

Гладкой задачей без ограничений называется задача об отыскании экстремумов этого функционала:

$$f(x) \to \mathrm{extr}.$$

6.2 Необходимые условия 1-го порядка

Теорема 6.1 (аналог теоремы Ферма в нормированных пространствах)**.** *Пусть $\widehat{x} \in \operatorname{locextr} f$ — точка локального экстремума функционала f и функционал f дифференцируем по Фреше (имеет вариацию по Лагранжу) в точке $\widehat{x}$. Тогда*

$$f'(\widehat{x}) = 0 \qquad (\delta f(\widehat{x}, h) = 0 \quad \forall\ h \in X).$$

Доказательство. Возьмем произвольный элемент $h \in X$, зафиксируем его. Рассмотрим функцию $\varphi(\lambda) = f(\widehat{x} + \lambda h)$. Поскольку точка $\widehat{x} \in \operatorname{locextr} f$, то $0 \in \operatorname{locextr} \varphi$ — локальный экстремум функции φ. По теореме Ферма для функций одной переменной $\varphi'_\lambda(0) = 0$. В терминах производной по направлению это эквивалентно тому, что $\delta f(\widehat{x}, h) = 0$. В силу произвольности h вариация $\delta f(\widehat{x}, h) = 0\ \forall\ h \in X$. Значит, вариация по Лагранжу $\delta f(\widehat{x}, \cdot) = 0$.

Если функционал f дифференцируем по Фреше в точке $\widehat{x}$, то в этой точке он имеет вариацию по Лагранжу и $f'(\widehat{x})[h] = \delta f(\widehat{x}, h)$. Поскольку из уже доказанного следует, что эта вариация $\delta f(\widehat{x}, \cdot) = 0$, то и $f'(\widehat{x}) = 0$ в силу определения дифференцируемости по Фреше. $\square$

6.3 Необходимые и достаточные условия 2-го порядка

Теорема 6.2. `Необходимые условия минимума:` *если $\widehat{x}$ — точка локального минимума функционала f $\left(\widehat{x} \in \operatorname{locmin} f\right)$ и функционал f дважды дифференцируем в точке $\widehat{x}$ $\left(f \in D^2(\widehat{x})\right)$, то*

$$f'(\widehat{x}) = 0, \;\; f''(\widehat{x})[h,h] \geq 0 \quad \forall\, h \in X.$$

`Достаточные условия минимума:` *если $f'(\widehat{x}) = 0$ и*

$$f''(\widehat{x})[h,h] \geq \alpha \|h\|^2 \quad \forall\, h \in X \tag{$*$}$$

при некотором $\alpha > 0$, то $\widehat{x} \in \operatorname{locmin} f$.

Доказательство. Возьмём произвольный вектор $h \in X$.
По формуле Тейлора

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

Необходимость. Поскольку $\widehat{x} \in \operatorname{locmin} f$, то по аналогу теоремы Ферма $f'(\widehat{x}) = 0$. Поэтому в силу формулы Тейлора при фиксированном h и достаточно малых λ

$$f(\widehat{x} + \lambda h) - f(\widehat{x}) = \frac{\lambda^2}{2} f''(\widehat{x})[h,h] + r(\lambda h), \quad r(\lambda h) = o(|\lambda|^2).$$

Поскольку $\widehat{x} \in \operatorname{locmin} f$, то $f(\widehat{x} + \lambda h) - f(\widehat{x}) \geq 0$ при малых λ. Поэтому

$$\tfrac{\lambda^2}{2} f''(\widehat{x})[h,h] + r(\lambda h) \geq 0.$$

Разделим обе части последнего неравенства на λ^2 и устремим λ к нулю. Поскольку $\frac{r(\lambda h)}{\lambda^2} \to 0$, то отсюда

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

Достаточность. Так как $f'(\widehat{x}) = 0$, то по формуле Тейлора в силу заданного условия $f''(\widehat{x})[h,h] \geq \alpha ||h||^2$ имеем:

$$f(\widehat{x} + h) - f(\widehat{x}) = \frac{1}{2} f''(\widehat{x})[h,h] + r(h) \geq \frac{\alpha}{2} ||h||^2 + r(h) \geq 0$$

при достаточно малых h, так как $r(h) = o(\|h\|^2)$. Следовательно, $\widehat{x} \in \operatorname{locmin} f$. $\square$

Для локального максимума неравенства имеют противоположный вид: $f''(\widehat{x})[h,h] \le 0$ и $f''(\widehat{x})[h,h] \le -\alpha\|h\|^2$ $(\alpha > 0)$ соответственно.

Условие $(*)$ называется условием *строгой положительности* второй производной Фреше функционала f.

В конечномерных пространствах условие положительной определенности симметричной матрицы A, т. е. условие

$$\langle Ah, h\rangle > 0 \quad \forall\, h \in \mathbb{R}^n,\ h \ne 0,$$

гарантирует её строгую положительность (и, значит, является достаточным условием минимума в стационарной точке).
В бесконечномерных пространствах это не так.

Пример 1. *В бесконечномерных пространствах условие положительной определенности отображения не гарантирует строгой положительности отображения.*

Рассмотрим отображение[24] $f\colon l_2 \to \mathbb{R}$, задающееся формулой

$$f(x) = \sum_{n=1}^{\infty} \frac{x_n^2}{n}, \quad x = (x_1, \ldots, x_n, \ldots).$$

Производная Фреше $f'(x)[h] = \sum\limits_{n=1}^{\infty} \dfrac{2x_n h_n}{n}$. Вторая производная $f''(x)[h,h] = 2\sum\limits_{n=1}^{\infty} \dfrac{h_n^2}{n} > 0\ \forall\, h \ne 0$ и, следовательно, положительно определена.

Но вместе с тем вторая производная отображения f не является строго положительной. Действительно, неравенство $f''(x)[h,h] \ge \alpha\|h\|^2\ \forall\, h$ не может выполняться ни для какой константы $\alpha > 0$, поскольку на последовательности векторов $\{e_n\}$ (e_n — n-й базисный вектор пространства l_2),

$$f''(x)[e_n, e_n] = \frac{2}{n} = \frac{2}{n}\|e_n\|^2 \ngeq \alpha\|e_n\|^2.$$

[24] Напомним, что l_2 — гильбертово пространство с нормой $\|x\|_{l_2} = \left(\sum\limits_{n=1}^{+\infty} x_n^2\right)^{\frac{1}{2}}$.

Пример 2. *В бесконечномерных пространствах условие равенства нулю первой производной и положительной определенности второй производной не гарантируют локального экстремума отображения.*

Рассмотрим задачу без ограничений

$$f(x) = \sum_{n=1}^{\infty} \left(\frac{x_n^2}{n^3} - x_n^4 \right) \to \min$$

в гильбертовом пространстве l_2, $x = (x_1, \ldots, x_n, \ldots)$.
Производная Фреше $f'(x)[h] = \sum\limits_{n=1}^{\infty} \left(\dfrac{2x_n h_n}{n^3} - 4x_n^3 h_n \right)$.
В точке $\widehat{x} = 0$ производная $f'(\widehat{x}) = 0$.
Следовательно, точка $\widehat{x} = 0$ является стационарной.
Вторая производная $f''(x)[h,h] = \sum\limits_{n=1}^{\infty} \left(\dfrac{2h_n^2}{n^3} - 12x_n^2 h_n^2 \right)$.
Отсюда $f''(0)[h,h] = 2 \sum\limits_{n=1}^{\infty} \dfrac{h_n^2}{n^3} > 0 \ \forall\ h \neq 0$ и, значит, вторая производная является положительно определенным оператором.

Но вместе с тем, как и в предыдущем примере, вторая производная не является строго положительным оператором:

$$f''(0)[e_n, e_n] = \frac{2}{n^3} = \frac{2}{n^3} \|e_n\|^2 \ngeq \alpha \|e_n\|^2$$

ни для какой константы $\alpha > 0$ (e_n — n-й базисный вектор пространства l_2).

Точка $\widehat{x} = 0 \notin \mathrm{locmin} f$, поскольку на последовательности векторов $\{x^n\}$, $x^n = \dfrac{e_n}{n}$, $f(x^n) = \dfrac{1}{n^5} - \dfrac{1}{n^4} < 0 = f(0)$ при $n > 1$, а сама последовательность $\{x^n\} \to \widehat{x} = 0$ в пространстве l_2 при $n \to +\infty$.

7 Гладкая задача с равенствами в банаховых пространствах

7.1 Постановка задачи

Пусть X, Y — банаховы пространства. Функционал $f\colon X \to \mathbb{R}$ и отображение $F\colon X \to Y$ обладают определенной гладкостью. *Гладкой экстремальной задачей с ограничениями типа равенств* называется следующая задача:

$$f(x) \to \operatorname{extr}; \quad F(x) = 0. \tag{P}$$

Поскольку пространство Y может быть бесконечномерным, то количество уравнений в ограничении $F(x) = 0$ может быть бесконечным.

7.2 Необходимые условия 1-го порядка

Теорема 7.1 (принцип Лагранжа)**.** *Пусть $\widehat{x} \in \operatorname{locextr} P$ — точка локального экстремума, X, Y — банаховы пространства* (условие банаховости)*, функционал f и отображение $F \in SD(\widehat{x})$ — строго дифференцируемы в точке $\widehat{x}$* (условие гладкости)*, $\operatorname{Im} F'(\widehat{x})$ — замкнутое подпространство в Y* (ослабленное условие регулярности)*. Тогда существуют множители Лагранжа $\lambda_0 \in \mathbb{R}$ и функционал $y^* \in Y^*$ не равные одновременно нулю, и такие, что для функции Лагранжа*

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

выполняется условие стационарности:

$$\mathscr{L}'(\widehat{x}) = 0 \quad \Big(\Longleftrightarrow \sum_{i=0}^{m} \lambda_i f_i'(\widehat{x})[h] + \langle y^*, F'(\widehat{x})[h] \rangle = 0 \ \forall\ h \in X \Big).$$

Точки, удовлетворяющие условию стационарности, называются *стационарными*. Это точки возможного локального экстремума.

Доказательство. Не ограничивая общности, считаем, что $f(\widehat{x}) = 0$. Определим отображение Φ, действующее по формуле $\Phi(x) := (f(x), F(x))$. Тогда $\Phi(\widehat{x}) = (f(\widehat{x}), F(\widehat{x})) = (0, 0)$, отображение Φ строго дифференцируемо в точке $\widehat{x}$, производная Фреше $\Phi'(\widehat{x}) = \big(f'(\widehat{x}),\ F'(\widehat{x})\big)$, оба отображения Φ и $\Phi'(\widehat{x})$ отображают пространство X в пространство $\mathbb{R} \times Y$.

Перепишем условие стационарности в виде

$$\langle(\lambda_0, y^*), (\langle f'(\widehat{x}), h\rangle, F'(\widehat{x})[h])\rangle = 0 \iff$$

$$\langle(\lambda_0, y^*), \Phi'(\widehat{x})[h]\rangle = 0 \quad \forall\, h \in X \iff \langle(\lambda_0, y^*), \operatorname{Im}\Phi'(\widehat{x})\rangle = 0$$

$$\iff (\lambda_0, y^*) \in (\operatorname{Im}\Phi'(\widehat{x}))^{\perp} \iff (\operatorname{Im}\Phi'(\widehat{x}))^{\perp} \neq 0.$$

Таким образом, условие стационарности означает, что аннулятор $(\operatorname{Im}\Phi'(\widehat{x}))^{\perp}$ содержит ненулевые элементы.

Поскольку подпространство $\operatorname{Im} F'(\widehat{x})$ замкнуто по условию, образ $f'(\operatorname{Ker} F'(\widehat{x}))$ тоже замкнут (в $\mathbb{R}$ имеются всего два подпространства: 0 и $\mathbb{R}$, и оба они замкнутые), то к отображению $\Phi'(\widehat{x})$ можно применить лемму о замкнутости образа. По этой лемме образ $\operatorname{Im}\Phi'(\widehat{x})$ замкнут в $\mathbb{R} \times Y$.

Если подпространство $\operatorname{Im}\Phi'(\widehat{x})$ является собственным (вырожденный случай), то по лемме о нетривиальности аннулятора замкнутого собственного подпространства аннулятор $(\operatorname{Im}\Phi'(\widehat{x}))^{\perp}$ содержит ненулевые элементы и, следовательно, условие стационарности выполняется.

Предположим, что подпространство $\operatorname{Im}\Phi'(\widehat{x})$ не является собственным, то есть $\operatorname{Im}\Phi'(\widehat{x}) = \mathbb{R} \times Y$ (невырожденный случай). В этом случае для отображения $\Phi\colon X \to \mathbb{R} \times Y$, $\Phi(\widehat{x}) = (f(\widehat{x}), F(\widehat{x})) = (0, 0) = (\widehat{\alpha}, \widehat{y})$, по теореме об обратном отображении существуют отображение $\Phi^{-1}\colon W \subset \mathbb{R} \times Y \to X$ некоторой окрестности W точки $(\widehat{\alpha}, \widehat{y})$ и константа $K > 0$ такие, что $\Phi^{-1}(\widehat{\alpha}, \widehat{y}) = \widehat{x}$, $\Phi(\Phi^{-1}(\alpha, y)) = (\alpha, y)$,

$$\|\Phi^{-1}(\alpha, y) - \Phi^{-1}(\widehat{\alpha}, \widehat{y})\| \leq K\|(\alpha, y) - (\widehat{\alpha}, \widehat{y})\|$$

$$\iff \|\Phi^{-1}(\alpha, y) - \widehat{x}\| \leq K\|(\alpha, y)\| \quad \forall\, (\alpha, y) \in W.$$

Положим $x(\varepsilon) = \Phi^{-1}(\varepsilon, 0)$ для малых по модулю ε. Тогда

$$\Phi(x(\varepsilon)) = (\varepsilon, 0) \iff f(x(\varepsilon)) = \varepsilon,\ F(x(\varepsilon)) = 0,$$

$$\|\Phi^{-1}(\varepsilon, 0) - \widehat{x}\| = \|x(\varepsilon) - \widehat{x}\| \le K\|(\varepsilon, 0)\| = K|\varepsilon|.$$

Из этих соотношений следует, что вектор $\widehat{x}$ не доставляет в задаче экстремума, ибо вблизи него существуют допустимые векторы $x(\varepsilon)$, на которых функционал f принимает значения как большие, так и меньшие чем $f(\widehat{x})$. Получили противоречие с тем, что $\widehat{x} \in \operatorname{locextr} P$. Таким образом, невырожденный случай невозможен, и тем самым теорема доказана. $\square$

Замечание. *Если в условиях теоремы выполнено условие регулярности отображения F в точке $\widehat{x}$, т. е. $\operatorname{Im} F'(\widehat{x}) = Y$, то множитель $\lambda_0 \neq 0$, и, следовательно, можем считать его равным единице* ($\lambda_0 = 1$).

Доказательство. Действительно, если $\lambda_0 = 0$, то $y^* \neq 0$ в силу того, что множители Лагранжа одновременно в ноль не обращаются. Поэтому условие стационарности приобретает вид: $\langle y^*, F'(\widehat{x})[h]\rangle = 0\ \forall\ h \in X \Leftrightarrow \langle y^*, Y\rangle = 0 \Leftrightarrow y^* = 0$. Получили противоречие. $\square$

7.3 Необходимые условия 2-го порядка

Выведем необходимые условия 2-го порядка экстремума в задаче с равенствами:

$$f(x) \to \operatorname{extr}; \quad F(x) = 0. \tag{P}$$

Здесь X, Y — банаховы пространства, функционал $f\colon X \to \mathbb{R}$, отображение $F\colon X \to Y$.

Теорема 7.2. *Пусть $\widehat{x} \in \operatorname{locextr} P$ — точка локального экстремума, X, Y — банаховы пространства* (условие банаховости)*, функционал f и отображение F имеют в этой точке вторые производные Фреше* ($f, F \in D^2(\widehat{x})$) (условие гладкости), $\operatorname{Im} F'(\widehat{x}) = Y$ (условие регулярности).

Тогда существуют множители Лагранжа $\lambda_0 = 1$, $y^* \in Y^*$, такие, что для функции Лагранжа $\mathscr{L}(x) = f(x) + \langle y^*, F(x)\rangle$ выполняется условие стационарности

$$\mathscr{L}'(\widehat{x}) = 0$$

и вторая производная функции Лагранжа знакоопределена на пространстве стационарных направлений $\Pi := \{h \in X \mid F'(\widehat{x})[h] = 0\}$:

$$\mathscr{L}''(\widehat{x})[h,h] \geq 0 \ \forall\ h \in \Pi \ \textit{для минимума},$$
$$\mathscr{L}''(\widehat{x})[h,h] \leq 0 \ \forall\ h \in \Pi \ \textit{для максимума}.$$

Доказательство. По Замечанию к Следствию 2 п. 5.3.3 существование второй производной Фреше в точке гарантирует строгую дифференцируемость отображения в этой точке. Тогда в силу выполнения условия банаховости пространств, условия гладкости отображений и выполнения условия ослабленной регулярности по теореме о необходимых условиях 1-го порядка для функции Лагранжа выполняется условие стационарности. Причём, в силу выполнения условия регулярности множитель Лагранжа λ_0 можем положить равным единице согласно Замечанию к этой теореме.

Проведем доказательство неотрицательности второй производной функции Лагранжа на пространстве стационарных направлений для минимума. Случай максимума рассматривается аналогично.

Возьмем произвольный вектор $h \in \Pi = \operatorname{Ker} F'(\widehat{x})$. Для отображения F и множества $M := \{x \mid F(x) = 0\}$ — множества допустимых точек в задаче (P) — выполнены все условия теоремы о касательном пространстве (строгая дифференцируемость отображения F в точке $\widehat{x}$ и регулярность оператора $F'(\widehat{x})$). По теореме о касательном пространстве (см. § 5.4) $T_{\widehat{x}}M = \operatorname{Ker} F'(\widehat{x})$. Значит, вектор h будет касательным к множеству M в точке $\widehat{x}$.

По определению касательного вектора существуют число $\varepsilon > 0$ и отображение $r\colon [-\varepsilon; \varepsilon] \to X$ такие, что точка $x_t := \widehat{x} + th + r(t) \in M$ $(\Leftrightarrow F(x_t) = 0)$ для любого $t \in [-\varepsilon; \varepsilon]$ и норма $\|r(t)\| = o(t)$ при $t \to 0$. Таким образом, x_t — допустимые точки в задаче и находятся в окрестности точки $\widehat{x}$.

Поскольку $F(x_t) = 0$ и $F(\widehat{x}) = 0$, то

$$f_0(x_t) = f_0(x_t) + \langle y^*, F(x_t)\rangle \overset{\text{def}}{=} \mathscr{L}(x_t),$$
$$f_0(\widehat{x}) = f_0(\widehat{x}) + \langle y^*, F(\widehat{x})\rangle \overset{\text{def}}{=} \mathscr{L}(\widehat{x}).$$

Поэтому

$$f_0(x_t) - f_0(\widehat{x}) = \mathscr{L}(x_t) - \mathscr{L}(\widehat{x}) = \mathscr{L}(\widehat{x} + th + r(t)) - \mathscr{L}(\widehat{x}) =$$

(по формуле Тейлора)

$$= \mathscr{L}'(\widehat{x})[th + r(t)] + \frac{1}{2}\mathscr{L}''(\widehat{x})[th + r(t), th + r(t)] + o(\|th + r(t)\|^2) =$$

$$\overset{\mathscr{L}'(\widehat{x})=0}{=} \frac{t^2}{2}\mathscr{L}''(\widehat{x})[h, h] + o(t^2) \geq 0, \text{ так как } \widehat{x} \in \operatorname{locmin} P.$$

Деля это неравенство на t^2 и переходя к пределу при $t \to 0$, получаем $\mathscr{L}''(\widehat{x})[h, h] \geq 0 \ \forall\ h \in \Pi$. $\square$

7.4 Достаточные условия 2-го порядка

Выведем достаточные условия 2-го порядка локального экстремума в гладкой задаче с ограничениями типа равенств в банаховых пространствах:

$$f(x) \to \mathrm{extr}; \quad F(x) = 0. \tag{P}$$

Теорема 7.3. *Пусть X, Y — банаховы пространства* (условие банаховости)*, функционалы f_i, $i = 0, 1, \ldots, m$, и отображение $F \in C^2(\mathcal{O}(\widehat{x}))$* (условие гладкости)*, $\operatorname{Im} F'(\widehat{x})$ — замкнутое подпространство в Y* (ослабленное условие регулярности)*. В допустимой точке $\widehat{x}$ для функции Лагранжа $\mathscr{L}(x) = f(x) + \langle y^*, F(x)\rangle$ с множителем Лагранжа $\lambda_0 = 1$ выполняется условие стационарности ($\mathscr{L}'(\widehat{x}) = 0$).*

Тогда, если вторая производная функции Лагранжа $\mathscr{L}''(\widehat{x})$ строго положительно определена на пространстве стационарных направлений[25]*:*

$$\mathscr{L}''(\widehat{x})[h, h] \geq \alpha \|h\|^2 \quad \forall\ h \in \Pi$$

с некоторой константой $\alpha > 0$, то $\widehat{x} \in \operatorname{locmin} P$ — точка локального минимума.

Если $\mathscr{L}''(\widehat{x})[h, h] \leq -\alpha\|h\|^2$ $\forall\ h \in \Pi$, то $\widehat{x} \in \operatorname{locmax} P$ — точка локального максимума.

Доказательство. Проведем доказательство для случая минимума. Случай максимума рассматривается аналогично.

Возьмем $\widehat{x}+h$ — допустимый элемент в задаче ($F(\widehat{x}+h)=0$). По Лемме 1 п. 5.4, применённой к линейному непрерывному оператору $F'(\widehat{x})$ вектор $h \in X$ можно представить в виде суммы $h = h_1 + h_2$, где

$$F'(\widehat{x})[h_1] = 0, \qquad \|h_2\| \leq C\|F'(\widehat{x})h\| \tag{L}$$

[25] Пространство стационарных направлений $\Pi = \Pi(\widehat{x}) := \{h \in X \mid F'(\widehat{x})[h] = 0\}$.

с некоторой константой $C > 0$, не зависящей от h.
Значит, вектор $h_1 \in \Pi$.

По формуле Тейлора

$$\underbrace{F(\widehat{x}+h)}_{=0} = \underbrace{F(\widehat{x})}_{=0} + F'(\widehat{x})[h] + \frac{1}{2}F''(\widehat{x})[h,h] + r(h),$$

где $r(h) = o(\|h\|^2)$. Отсюда $F'(\widehat{x})[h] = -\dfrac{1}{2}F''(\widehat{x})[h,h] - r(h)$.

Следовательно, при достаточно малых h

$$\|F'(\widehat{x})[h]\| \le \frac{1}{2}\|F''(\widehat{x})[h,h]\| + \|r(h)\| \le C_1\|h\|^2$$

с некоторой константой $C_1 > 0$. Поэтому при таких h

$$\|h_2\| \overset{(L)}{\le} C\|F'(\widehat{x})[h]\| \le CC_1\|h\|^2 = \gamma\|h\|^2 \qquad (h_2)$$

для $\gamma := CC_1$, а по неравенству треугольника для норм

$$\|h\| - \|h_2\| \le \|h_1\| \overset{\text{def}}{=} \|h - h_2\| \le \|h\| + \|h_2\| \overset{\|h_2\|<\gamma\|h\|^2}{\Longrightarrow}$$

$$\|h\| - \gamma\|h\|^2 \le \|h_1\| \le \|h\| + \gamma\|h\|^2 \iff$$

$$(1-\gamma\|h\|)\|h\| \le \|h_1\| \le (1+\gamma\|h\|)\|h\|. \qquad (h_1)$$

Поскольку для допустимых точек $F(\widehat{x}+h) = 0$ и $F(\widehat{x}) = 0$, то

$$f(\widehat{x}+h) = f(\widehat{x}+h) + \langle y^*, F(\widehat{x}+h)\rangle \overset{\text{def}\,\mathscr{L}}{=} \mathscr{L}(\widehat{x}+h),$$

$$f(\widehat{x}) = f(\widehat{x}) + \langle y^*, F(\widehat{x})\rangle \overset{\text{def}\,\mathscr{L}}{=} \mathscr{L}(\widehat{x}).$$

Поэтому

$$f(\widehat{x}+h) - f(\widehat{x}) = \mathscr{L}(\widehat{x}+h) - \mathscr{L}(\widehat{x}) =$$

(по формуле Тейлора)

$$= \mathscr{L}'(\widehat{x})[h] + \tfrac{1}{2}\mathscr{L}''(\widehat{x})[h,h] + o(\|h\|^2) =$$

$$\overset{\mathscr{L}'(\widehat{x})=0}{=} \tfrac{1}{2}\mathscr{L}''(\widehat{x})[h,h] + o(\|h\|^2) =$$

(подставим $h = h_1 + h_2$)

$$= \frac{1}{2}\mathscr{L}''(\widehat{x})[h_1 + h_2, h_1 + h_2] + o(\|h\|^2) =$$

$$= \frac{1}{2}\Big(\mathscr{L}''(\widehat{x})[h_1, h_1] + 2\mathscr{L}''(\widehat{x})[h_1, h_2] + \mathscr{L}''(\widehat{x})[h_2, h_2]\Big) + o(\|h\|^2) \geq$$

(поскольку по условию $\mathscr{L}''(\widehat{x})[h_1, h_1] \geq \alpha\|h_1\|^2$ для вектора $h_1 \in \Pi$, а для непрерывного билинейного оператора имеет место оценка $\|\mathscr{L}''(\widehat{x})[\xi, \eta]\| \leq B\|\xi\|\cdot\|\eta\|$)

$$\geq \frac{1}{2}\Big(\alpha\|h_1\|^2 - 2B\|h_1\|\cdot\|h_2\| - B\|h_2\|^2\Big) + o(\|h\|^2) \overset{(h_1)}{\geq}$$

$$\geq \frac{1}{2}\Big(\alpha(1-\gamma\|h\|)^2\|h\|^2 - 2B(1+\gamma\|h\|)\|h\|\cdot\|h_2\| - B\|h_2\|^2\Big) + o(\|h\|^2) \overset{(h_2)}{\geq}$$

$$\geq \frac{1}{2}\Big(\alpha(1-\gamma\|h\|)^2\|h\|^2 - 2B(1+\gamma\|h\|)\|h\|\gamma\|h\|^2 - B\gamma^2\|h\|^4\Big) + o(\|h\|^2) =$$

$$= \frac{1}{2}\Big(\alpha(1-\gamma\|h\|)^2 - 2B(1+\gamma\|h\|)\gamma\|h\| - B\gamma^2\|h\|^2\Big)\|h\|^2 + o(\|h\|^2) \geq$$

(возьмём $\|h\| < \varepsilon$)

$$\geq \frac{1}{2}\Big(\alpha(1-\gamma\varepsilon)^2 - 2B(1+\gamma\varepsilon)\varepsilon\gamma - B\gamma^2\varepsilon^2\Big)\|h\|^2 + o(\|h\|^2) \geq 0,$$

так как при достаточно малых $\varepsilon > 0$ (при $\varepsilon = 0$ множитель в круглых скобках равен $\alpha > 0$). Из полученного неравенства $f(\widehat{x} + h) - f(\widehat{x}) \geq 0$ следует, что $\widehat{x} \in \operatorname{locmin} P$. $\square$

7.5 Квадратично-линейная задача с линейным ограничением типа равенства

Постановка задачи

Пусть X, Y — линейные нормированные пространства. Рассмотрим квадратично-линейную задачу с линейным ограничением типа равенства

$$f_0(x) = A(x,x) + \langle x_0^*, x\rangle + b_0 \to \mathrm{extr}; \quad \Lambda x = b, \qquad (P)$$

где $A\colon X \times X \to \mathbb{R}$ — квадратичный симметрический билинейный функционал, $x_0^*\colon X \to \mathbb{R}$ — линейный непрерывный функционал, $b_0 \in \mathbb{R}$, $\Lambda\colon X \to Y$ — линейный непрерывный оператор, элемент $b \in Y$.

Теорема 7.4. *Пусть $\widehat{x}$ — допустимый элемент в задаче (P), удовлетворяющий условию стационарности $(\mathscr{L}'(\widehat{x}) = 0)$ функции Лагранжа $\mathscr{L}(x) = f_0(x) + \langle y^*, \Lambda x\rangle$ с множителем Лагранжа $\lambda_0 = 1$.*

Если $A(h,h) \geq 0 \ \forall\ h \in \Pi := \operatorname{Ker}\Lambda = \{h \in X \mid \Lambda h = 0\}$, то $\widehat{x} \in \operatorname{absmin} P$. Если для A не выполнено условие $A(h,h) \geq 0$ на пространстве Π, то $\widehat{x} \notin \operatorname{locmin} P$, а $S_{\mathrm{absmin}} = -\infty$.

Если $A(h,h) \leq 0 \ \forall\ h \in \Pi$, то $\widehat{x} \in \operatorname{absmax} P$. Если для A не выполнено условие $A(h,h) \leq 0$ на пространстве Π, то $\widehat{x} \notin \operatorname{locmax} P$, а $S_{\mathrm{absmax}} = +\infty$.

Доказательство. Проведем доказательство для случая минимума. Случай максимума рассматривается аналогично.

Пусть вектор h такой, что вектор $\widehat{x} + h$ является допустимым. Для линейного ограничения это эквивалентно тому, что $\Lambda h = 0$. Действительно,

$$\Lambda(\widehat{x} + h) = b \iff \Lambda\widehat{x} + \Lambda h = b \iff \Lambda h = 0,$$

поскольку $\widehat{x}$ — допустимый элемент, то $\Lambda\widehat{x} = b$. Значит, вектор $\widehat{x} + h$ является допустимым тогда и только тогда, когда $h \in \Pi$ — пространству стационарных направлений.

Поскольку в точке $\widehat{x}$ выполняется условие стационарности функции Лагранжа, то

$$\mathscr{L}'(\widehat{x}) = 0 \iff f_0'(\widehat{x})[h] + \langle y^*, \Lambda h\rangle = 0 \ \forall\ h \in X.$$

Отсюда

$$f_0'(\widehat{x})[h] = 0 \ \forall\ h \in \Pi. \qquad (*)$$

Для функционала $f_0(x) = A(x,x) + \langle x_0^*, x\rangle + b_0$ вторая производная $f_0''(\widehat{x}) = 2A$, следовательно, по формуле Тейлора имеет место равенство

$$f_0(\widehat{x}+h) = f_0(\widehat{x}) + f_0'(\widehat{x})[h] + \tfrac{1}{2} f_0''(\widehat{x})[h,h] \iff$$
$$f_0(\widehat{x}+h) - f_0(\widehat{x}) = f_0'(\widehat{x})[h] + A(h,h) \ \forall\ h \in X.$$

Из последнего равенства в силу равенства $(*)$ и неотрицательной определенности квадратичного функционала A на пространстве стационарных направлений Π вытекает, что

$$f_0(\widehat{x}+h) - f_0(\widehat{x}) \geq 0 \ \forall\ h \in \Pi.$$

Значит, $\widehat{x} \in \operatorname{absmin} P$.

Предположим, что не выполнено условие неотрицательности формы A на пространстве стационарных направлений, то есть существует элемент $h \in \Pi$, для которого $A(h,h) < 0$. Тогда $th \in \Pi \ \forall\ t \in \mathbb{R}$, а

$$f_0(\widehat{x}+th) - f_0(\widehat{x}) = A(th,th) = t^2 A(h,h) < 0.$$

Отсюда, с одной стороны, $f_0(\widehat{x}+th) < f_0(\widehat{x})$, а $\widehat{x}+th \to \widehat{x}$ при $t \to 0$, поэтому, $\widehat{x} \notin \operatorname{locmin} P$.

С другой стороны, $f_0(\widehat{x}+th) = f_0(\widehat{x}) + t^2 A(h,h) \to -\infty$ при $t \to +\infty$, и, значит, $S_{\text{absmin}} = -\infty$. $\square$

8 Гладкая задача с равенствами и неравенствами

8.1 Постановка задачи

Пусть X, Y — банаховы пространства. Отображение $F\colon X \to Y$, функционалы $f_i\colon X \to \mathbb{R}$, $i = 0, 1, \ldots, m$, обладают некоторой гладкостью. *Гладкой экстремальной задачей с ограничениями типа равенств и неравенств* в нормированных пространствах называется следующая задача:

$$f_0(x) \to \operatorname{extr}; \quad f_i(x) \le 0,\ i = 1, \ldots, m,\ F(x) = 0. \qquad (P)$$

8.2 Необходимые условия 1-го порядка

Теорема 8.1 (правило множителей Лагранжа)**.** *Пусть точка $\widehat{x} \in \operatorname{locextr} P$, X, Y — банаховы пространства, отображения $F, f_i \in SD(\widehat{x})$, $i = 0, 1, \ldots, m$, $\operatorname{Im} F'(\widehat{x})$ — замкнутое подпространство в Y. Тогда существуют множители Лагранжа — вектор $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$ и функционал $y^* \in Y^*$, $(\lambda, y^*) \ne 0$, такие, что для функции Лагранжа $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x) + \langle y^*, F(x)\rangle$ выполняются условия:*

a) стационарности:

$$\mathscr{L}'(\widehat{x}) = 0 \ \Big(\Longleftrightarrow \sum_{i=0}^{m} \lambda_i f_i'(\widehat{x})[h] + \langle y^*, F'(\widehat{x})[h]\rangle = 0\ \forall\, h \in X\Big);$$

b) дополняющей нежёсткости: $\lambda_i f_i(\widehat{x}) = 0,\ i = 1, \ldots, m$;

c) знакоопределенности:

$$\lambda_0 \ge 0,\ \begin{array}{l} \lambda_i \ge 0,\ i = 1, \ldots, m,\ \text{для локального минимума,} \\ \lambda_i \le 0,\ i = 1, \ldots, m,\ \text{для локального максимума.} \end{array}$$

8.3 Необходимые условия 1-го и 2-го порядка

Рассмотрим гладкую задачу с ограничениями типа равенств и неравенств в банаховых пространствах:

$$f_0(x) \to \mathrm{extr}; \quad f_i(x) \le 0,\ i = 1, \dots, m,\ F(x) = 0, \qquad (P)$$

функционалы $f_i : X \to \mathbb{R}$, $i = 0, 1, \dots, m$, отображение $F : X \to Y$.

Теорема 8.2. *Пусть точка $\widehat{x} \in \mathrm{locextr}\, P$, X, Y — банаховы пространства, функционалы f_i, $i = 0, 1, \dots, m$, и отображение $F \in C^2(\mathcal{O}(\widehat{x}))$, $\mathrm{Im}\, F'(\widehat{x}) = Y$. Тогда существуют множители Лагранжа — вектор $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m) \in \mathbb{R}^{m+1}$, $\lambda \ne 0$, и функционал $y^* \in Y^*$, такие, что для функции Лагранжа $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x) + \langle y^*, F(x) \rangle$ выполняются условия:*

a) стационарности: $\mathscr{L}'(\widehat{x}) = 0$;

b) дополняющей нежёсткости: $\lambda_i f_i(\widehat{x}) = 0,\ i = 1, \dots, m$;

c) знакоопределенности множителей Лагранжа:

$$\lambda_0 \ge 0, \quad \begin{array}{l} \lambda_i \ge 0,\ i = 1, \dots, m,\ \textit{для минимума}, \\ \lambda_i \le 0,\ i = 1, \dots, m,\ \textit{для максимума}; \end{array}$$

d) для любого h из конуса допустимых направлений[26]

$$K = K(\widehat{x}) := \{h \in X \mid f_i'(\widehat{x})[h] \le 0,\ i = 0, 1, \dots, m,\ F'(\widehat{x})[h] = 0\}$$

найдутся множители Лагранжа $(\lambda, y^*) \ne 0$, удовлетворяющие условиям a)–c) и такие, что

$$\begin{array}{ll} \mathscr{L}''(\widehat{x})[h, h] \ge 0 & \textit{для минимума}, \\ \mathscr{L}''(\widehat{x})[h, h] \le 0 & \textit{для максимума}. \end{array} \qquad (\mathscr{L}'')$$

Замечание. Отметим, что если для точки локального экстремума $\widehat{x}$ вектор множителей Лагранжа (λ, y^*) единственен (с точностью до умножения на положительную константу), то неравенства $(\mathscr{L}'')$ выполняются для любого $h \in K$.

---

[26] Напомним, что в формировании конуса допустимых направлений из ограничений типа неравенств берутся только активные ограничения, т. е. только те, для которых $f_i(\widehat{x}) = 0,\ i = 1, \dots, m$.

8.4 Достаточные условия 2-го порядка

Выведем достаточные условия 2-го порядка локального экстремума в гладкой задаче с ограничениями типа равенств и неравенств в банаховых пространствах:

$$f_0(x) \to \text{extr}; \quad f_i(x) \le 0,\ i = 1, \dots, m,\ F(x) = 0, \qquad (P)$$

функционалы $f_i : X \to \mathbb{R}$, $i = 0, 1, \dots, m$, отображение $F : X \to Y$.

Теорема 8.3. *Пусть X, Y — банаховы пространства* (условие банаховости)*, функционалы f_i, $i = 0, 1, \dots, m$, и отображение $F \in C^2(\mathcal{O}(\widehat{x}))$* (условие гладкости)*, $\operatorname{Im} F'(\widehat{x}) = Y$* (условие регулярности)*. В допустимой точке $\widehat{x}$ для функции Лагранжа $\mathscr{L}(x) = f_0(x) + \sum\limits_{i=1}^{m} \lambda_i f_i(x) + \langle y^*, F(x) \rangle$ выполняются условия*:

a) стационарности: $\mathscr{L}'(\widehat{x}) = 0$;

b) дополняющей нежёсткости: $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \dots, m$;

c) неотрицательности: $\lambda_i \ge 0,\ i = 1, \dots, m$;

d) строгой положительности $\mathscr{L}''(\widehat{x})$ на конусе допустимых направлений[27]:

$$\mathscr{L}''(\widehat{x})[h, h] \ge \alpha \|h\|^2 \quad \forall\ h \in K$$

с некоторой константой $\alpha > 0$.

Тогда $\widehat{x} \in \operatorname{locmin} P$ — точка локального минимума.

Если условия c) и d) выполняются в виде:

c) $\lambda_i \le 0,\ i = 1, \dots, m$;

d) $\mathscr{L}''(\widehat{x})[h, h] \le -\alpha \|h\|^2 \quad \forall\ h \in K(\widehat{x})$,

то $\widehat{x} \in \operatorname{locmax} P$ — точка локального максимума.

---

[27] Конус допустимых направлений
$K = K(\widehat{x}) := \{h \in X \mid f_i'(\widehat{x})[h] \le 0,\ i = 1, \dots, m,\ F'(\widehat{x})[h] = 0\}$.
Напомним, что в формировании конуса допустимых направлений из ограничений типа неравенств берутся только активные ограничения, т. е. только те, для которых $f_i(\widehat{x}) = 0$.

8.5 Квадратично-линейная задача с линейными ограничениями типа неравенств и равенства

Постановка задачи

Пусть X, Y — линейные нормированные пространства. Рассмотрим квадратично-линейную задачу с линейными ограничениями типа неравенств и равенства

$$f_0(x) := A(x,x) + \langle x_0^*, x\rangle + b_0 \to \text{extr};$$
$$\langle x_i^*, x\rangle + b_i \le 0,\ i = 1,\dots,m,\ \Lambda x = b, \quad (P)$$

где $A\colon X\times X \to \mathbb{R}$ — квадратичный симметрический билинейный функционал, $f_0\colon X \to \mathbb{R}$, $x_i^* \in X^*$, $b_i \in \mathbb{R}$, $i = 0,1,\dots,m$, $\Lambda\colon X \to Y$ — линейный непрерывный оператор, элемент $b \in Y$.

Теорема 8.4. *Пусть $\widehat{x}$ — допустимый элемент в задаче (P), удовлетворяющий условию стационарности функции Лагранжа $\mathscr{L}(x) = f_0(x) + \sum\limits_{i=1}^{m} \lambda_i(\langle x_i^*, x\rangle + b_i) + \langle y^*, \Lambda x\rangle$ и условию знакоопределенности множителей Лагранжа, а все ограничения типа неравенств активны.*

Тогда, если $A(h,h) \ge 0$ на конусе допустимых направлений $K := \{h \in X \mid \langle x_i^*, h\rangle \le 0,\ i = 1,\dots,m,\ \Lambda h = 0\}$ в критической точке $\widehat{x}$ на минимум, то $\widehat{x} \in \operatorname{absmin} P$. Если для A не выполнено условие $A(h,h) \ge 0$ на конусе K, то $S_{\text{absmin}} = -\infty$.

Если $A(h,h) \le 0$ на конусе K в критической точке $\widehat{x}$ на максимум, то $\widehat{x} \in \operatorname{absmax} P$. Если для A не выполнено условие $A(h,h) \le 0$ на конусе K, то $S_{\text{absmax}} = +\infty$.

Доказательство. Проведем доказательство для случая критической точки в задаче на минимум, т. е. $\lambda_i \ge 0$, $i = 1,\dots,m$. Случай максимума рассматривается аналогично.

Пусть вектор h такой, что вектор $\widehat{x}+h$ является допустимым. Тогда в силу активности всех ограничений $(\langle x_i^*, \widehat{x}\rangle + b_i = 0)$

$$\langle x_i^*, \widehat{x} + h\rangle + b_i \le 0 \iff \langle x_i^*, h\rangle \le 0,\ i = 1,\dots,m; \qquad (i)$$

$$\Lambda(\widehat{x} + h) = b \iff \Lambda\widehat{x} + \Lambda h = b \iff \Lambda h = 0. \qquad (ii)$$

Следовательно, вектор $\widehat{x}+h$ является допустимым в задаче тогда и только тогда, когда $h \in K$ — конусу допустимых направлений.

Поскольку в точке $\widehat{x}$ выполняется условие стационарности функции Лагранжа, то

$$\mathscr{L}'(\widehat{x}) = 0 \iff f_0'(\widehat{x})[h] + \sum_{i=1}^{m} \lambda_i \langle x_i^*, h\rangle + \langle y^*, \Lambda h\rangle = 0 \ \forall\ h \in X.$$

Отсюда

$$f_0'(\widehat{x})[h] = -\sum_{i=1}^{m} \lambda_i \langle x_i^*, h\rangle \overset{(i)}{\geq} 0 \ \forall\ h \in K. \qquad (*)$$

Для функционала $f_0(x) = A(x,x) + \langle x_0^*, x\rangle + b_0$ вторая производная $f_0''(\widehat{x})[h,h] = 2A(h,h)$, следовательно, по формуле Тейлора имеет место равенство

$$f_0(\widehat{x}+h) = f_0(\widehat{x}) + f_0'(\widehat{x})[h] + \tfrac{1}{2} f_0''(\widehat{x})[h,h] \iff$$
$$f_0(\widehat{x}+h) - f_0(\widehat{x}) = f_0'(\widehat{x})[h] + A(h,h) \ \forall\ h \in X.$$

Из последнего равенства в силу неравенства $(*)$ и неотрицательной определенности квадратичного функционала A на конусе допустимых направлений K вытекает, что

$$f_0(\widehat{x}+h) - f_0(\widehat{x}) \geq 0 \ \forall\ h \in K.$$

Значит, $\widehat{x} \in \operatorname{absmin} P$.

Предположим, что не выполнено условие неотрицательной определенности матрицы A на конусе допустимых направлений, то есть существует элемент $h \in K$, для которого $\langle Ah, h\rangle < 0$. Тогда $th \in K \ \forall\ t > 0$, а

$$f_0(\widehat{x}+th) = f_0(\widehat{x}) + t f_0'(\widehat{x})[h] + t^2 (Ah,h) \to -\infty$$

при $t \to \infty$ и, значит, $S_{\text{absmin}} = -\infty$. $\square$

Глава 3

Вариационное исчисление

1 Простейшая задача вариационного исчисления

1.1 Постановка задачи

Простейшей задачей вариационного исчисления называется экстремальная задача в пространстве $C^1([t_0;\, t_1], \mathbb{R})$[1]:

$$J(x(\cdot)) = \int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt \to \text{extr};\ x(t_0) = x_0,\ x(t_1) = x_1. \quad (P)$$

Здесь $L = L(t, x, \dot{x})$ — заданная функция трех переменных, называемая *интегрантом*. Отрезок $[t_0;\, t_1]$ предполагается фиксированным и конечным, $t_0 < t_1$. Экстремум в задаче рассматривается среди непрерывно дифференцируемых

[1] $C^1([t_0;\, t_1], \mathbb{R}) = C^1([t_0;\, t_1])$ — пространство отображений отрезка $[t_0;\, t_1]$ в пространство $\mathbb{R}$, имеющих на отрезке $[t_0;\, t_1]$ непрерывную производную.

функций $x(\cdot) \in C^1([t_0;\, t_1], \mathbb{R})$, удовлетворяющих *условиям на концах*, или *краевым условиям*: $x(t_0) = x_0$, $x(t_1) = x_1$. Такие функции называются *допустимыми*.

Определение. Говорим, что допустимая функция $\widehat{x}$ доставляет *слабый локальный минимум* в задаче (P), и пишем $\widehat{x}(\cdot) \in \operatorname{wlocmin}$ [2] P, если существует $\delta > 0$ такое, что $J(x(\cdot)) \geq J(\widehat{x}(\cdot))$ для любой допустимой функции $x(\cdot)$ $(x(\cdot) \in D(P))$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^1([t_0;\, t_1])}$ [3] $< \delta$.

Аналогично определяется слабый локальный максимум.

1.2 Вывод уравнения Эйлера с помощью основной леммы вариационного исчисления

Теорема 1.1. *Пусть функция $\widehat{x}(\cdot)$ доставляет слабый локальный экстремум в задаче (P), функции $L, L_x, L_{\dot{x}}$ — непрерывны как функции трех переменных, $\widehat{L}_{\dot{x}}(\cdot)$* [4] *$\in C^1([t_0;\, t_1])$. Тогда функция $\widehat{x}(\cdot)$ удовлетворяет уравнению Эйлера*

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \quad \forall\, t \in [t_0;\, t_1].$$

Функции, удовлетворяющие уравнению Эйлера, называются *экстремалями*. Множество экстремалей будем обозначать $E(P)$. Допустимые функции (класса C^1 с заданными граничными условиями), удовлетворяющие уравнению Эйлера, называются *допустимыми экстремалями*. Множество допустимых экстремалей в задаче (P) обозначаем $DE(P)$.

---

[2] weak — слабый

[3] напомним, что $\|y(\cdot)\|_{C^1([t_0;\, t_1])} := \max\{\|y(\cdot)\|_{C([t_0;\, t_1])}, \|\dot{y}(\cdot)\|_{C([t_0;\, t_1])}\}$, где $\|y(\cdot)\|_{C([t_0;\, t_1])} := \max\{|y(t)| \mid t \in [t_0;\, t_1]\}$.

[4] $\widehat{L}_{\dot{x}}(t) := \dfrac{\partial}{\partial \dot{x}} L(t, x, \dot{x})\Big|_{\substack{x=\widehat{x}(t)\\ \dot{x}=\dot{\widehat{x}}(t)}}$, $\widehat{L}_x(t) := \dfrac{\partial}{\partial x} L(t, x, \dot{x})\Big|_{\substack{x=\widehat{x}(t)\\ \dot{x}=\dot{\widehat{x}}(t)}}$.

Лемма (Лагранж). [5] *Пусть функция* $a(\cdot) \in C([t_0;\, t_1])$ *и*
$\int\limits_{t_0}^{t_1} a(t)h(t)\,dt = 0 \;\forall\; h(\cdot) \in C_0^1([t_0;\, t_1])$. *Тогда* $a(t) \equiv 0$.

[5] Лемму Лагранжа иногда называют еще *основной леммой вариационного исчисления.*

1.3 Вывод уравнения Эйлера с помощью леммы Дюбуа-Реймона

В этом пункте мы выведем уравнение Эйлера для меньших условий гладкостей, наложенных на интегрант L. При выводе уравнения Эйлера вместо леммы Лагранжа будем использовать лемму Дюбуа-Реймона[6].

Теорема 1.2. *Пусть функция $\widehat{x}$ доставляет слабый локальный экстремум в задаче (P), функции $L, L_x, L_{\dot{x}}$ — непрерывны в некоторой окрестности расширенного графика $\Gamma_{\widehat{x}\dot{\widehat{x}}}$*[7]. *Тогда $\widehat{L}_{\dot{x}}$ непрерывно дифференцируемая функция и выполнено уравнение Эйлера*

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \quad \forall\ t \in [t_0;\, t_1].$$

Лемма (Дюбуа-Реймон)**.** *Пусть функции $a_0, a_1 \in C([t_0;\, t_1])$ и $\int\limits_{t_0}^{t_1} \big(a_1(t)\dot{h}(t) + a_0(t)h(t)\big)dt = 0\ \forall\ h(\cdot) \in C_0^1([t_0;\, t_1])$.*

Тогда функция $a_1(\cdot) \in C^1([t_0;\, t_1])$ и выполняется дифференциальное уравнение $-\dfrac{d}{dt}a_1(t) + a_0(t) = 0\ \forall\ t \in [t_0;\, t_1]$ (аналог уравнения Эйлера).

---

[6] Дюбуа-Реймон, Поль Давид Густав (1831–1889) — немецкий математик.

[7] $\Gamma_{\widehat{x}\dot{\widehat{x}}} := \{(t, \widehat{x}(t), \dot{\widehat{x}}(t)) \mid t \in [t_0;\, t_1]\}$ — расширенный график.

1.4 Интегралы уравнения Эйлера

Если интегрант $L = L(t, x, \dot{x})$ не зависит явно от одной из переменных, то уравнение Эйлера сводится к более простым уравнениям.

1. Если интегрант $L = L(t, \dot{x})$ не зависит явно от x, то имеет место *интеграл импульса*

$$\widehat{L}_{\dot{x}}(t) = \text{const}.$$

Интеграл импульса выводится из уравнения Эйлера:

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

2. Если интегрант $L = L(x, \dot{x})$ не зависит явно от t, то имеет место обобщенный *интеграл энергии* (оба названия интегралов взяты из классической механики)

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

Для доказательства интеграла энергии достаточно продифференцировать последнее равенство по t и воспользоваться уравнением Эйлера:

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

Замечание. Отметим, что при выводе интеграла энергии мы использовали дополнительное предположение о существовании второй производной $\ddot{\widehat{x}}$. Интеграл энергии имеет также всегда экстремаль $\widehat{x}(t) = \text{const}$, которая быть может не удовлетворяет уравнению Эйлера.

1.5 Примеры

Пример 1. $J(x(\cdot)) = \int\limits_0^1 \dot{x}^2\,dt \to \text{extr}$; $x(0) = 0$, $x(1) = 1$.

Решение. Необходимое условие экстремума — уравнение Эйлера для интегранта $L = \dot{x}^2$:

$$-\frac{d}{dt}L_{\dot{x}} + L_x = 0 \iff -\frac{d}{dt}2\dot{x} = 0 \iff -2\ddot{x} = 0 \iff \ddot{x} = 0.$$

Общее решение уравнения Эйлера: $x = C_1 t + C_2$.
Неизвестные константы C_1, C_2 находим из условий на концах:

$x(0) = 0 \Rightarrow C_2 = 0;$

$x(1) = 1 \Rightarrow C_1 = 1.$

Значит, имеется единственная допустимая экстремаль $\widehat{x} = t$.

Проверим, доставляет ли найденная допустимая экстремаль минимум или максимум в задаче. Для этого надо сравнить значение функционала $J(\widehat{x})$ со значением функционала $J(x)$ на произвольной допустимой функции x. Обозначим $h = x - \widehat{x}$. Тогда $x = \widehat{x} + h$, $h \in C^1([0;\,1])$, $h(0) = h(1) = 0$. Функционал J является квадратичным. Тогда для задачи с квадратичным функционалом и линейными ограничениями

$$J(\widehat{x} + h) - J(\widehat{x}) = J(h) = \int\limits_0^1 \dot{h}^2\,dt \geq 0.$$

Разность $J(\widehat{x} + h) - J(\widehat{x})$ неотрицательна, значит, $\widehat{x} \in \text{absmin}$,

$$S_{\text{absmin}} = \int\limits_0^1 \dot{\widehat{x}}^2\,dt = \int\limits_0^1 1\,dt = 1.$$

Покажем, что $S_{\text{absmax}} = +\infty$. Действительно, возьмем последовательность допустимых функций $x_n(t) = \widehat{x}(t) + nh(t)$, например, $h = t(t-1)$. Тогда

$$J(x_n) = \int\limits_0^1 \dot{x}_n^2\,dt = \int\limits_0^1 (1+n\dot{h})^2\,dt = n^2 \int\limits_0^1 \dot{h}^2\,dt + 1 \to +\infty \text{ при } n \to \infty.$$

Ответ. Допустимая экстремаль $\widehat{x} = t \in \text{absmin}$, $S_{\text{absmin}} = 1$; $S_{\text{absmax}} = +\infty$.

Замечание. В Примере 1 и во многих других задачах КВИ (где в отличии от задач оптимального управления нет ограничения на величину производной) показать, что $S_{\text{absmax}} = +\infty$ можно рассматривая значение функционала на допустимых функциях (типа "пилы") с большой по модулю производной. В Главе 5 (см. Лемма о скруглении углов) будет показано, что такие негладкие функции можно "сгладить".

Пример 2. $J(x(\cdot)) = \int\limits_0^{\frac{3\pi}{2}} (\dot{x}^2 - x^2)\,dt \to \text{extr}$; $x(0) = x\Big(\dfrac{3\pi}{2}\Big) = 0$.

Решение. Необходимое условие экстремума — уравнение Эйлера для интегранта $L = \dot{x}^2 - x^2$:

$$-\frac{d}{dt}L_{\dot{x}} + L_x = 0 \iff -\frac{d}{dt}2\dot{x} - 2x = 0 \iff$$
$$-2\ddot{x} - 2x = 0 \iff \ddot{x} + x = 0.$$

Общее решение уравнения Эйлера: $x = C_1 \sin t + C_2 \cos t$.
Неизвестные константы C_1, C_2 находим из условий на концах:

$x(0) = 0 \Rightarrow C_2 = 0;$

$x\Big(\dfrac{3\pi}{2}\Big) = 0 \Rightarrow C_1 = 0.$

Значит, имеется единственная допустимая экстремаль $\widehat{x} \equiv 0$, $J(\widehat{x}) = 0$.

Проверим, доставляет ли найденная допустимая экстремаль минимум или максимум в задаче. Рассмотрим допустимые функции вида $x_\lambda(t) = \lambda \sin \dfrac{2t}{3}$. Очевидно, что $x_\lambda(\cdot) \to \widehat{x}(\cdot)$ в метрике пространства $C^1\big([0;\, \frac{3\pi}{2}]\big)$ при $\lambda \to 0$, но при этом

$$J(x_\lambda) = \lambda^2 \int\limits_0^{\frac{3\pi}{2}} \Big(\frac{4}{9}\cos^2\frac{2t}{3} - \sin^2\frac{2t}{3}\Big)dt =$$

$$= \lambda^2 \int\limits_0^{\frac{3\pi}{2}} \Big(\frac{4}{9}\cdot\frac{1+\cos\frac{4t}{3}}{2} - \frac{1-\cos\frac{4t}{3}}{2}\Big)dt = \lambda^2\frac{3\pi}{4}\Big(\frac{4}{9}-1\Big) < 0 = J(\widehat{x}).$$

Значение функционала на x_λ меньше чем на $\widehat{x}$, значит, $\widehat{x}$ не доставляет слабого локального минимума.

Поскольку $J(x_\lambda(\cdot)) \to -\infty$ при $\lambda \to \infty$, то $S_{\text{absmin}} = -\infty$.

Если взять допустимые функции $x_\lambda(t) = \lambda \sin \dfrac{4t}{3}$, то аналогично получим

$$J(x_\lambda) = \lambda^2 \int\limits_0^{\frac{3\pi}{2}} \Big(\frac{16}{9}\cos^2\frac{4t}{3} - \sin^2\frac{4t}{3}\Big)dt =$$

$$= \lambda^2 \int\limits_0^{\frac{3\pi}{2}} \Big(\frac{16}{9}\cdot\frac{1+\cos\frac{8t}{3}}{2} - \frac{1-\cos\frac{8t}{3}}{2}\Big)dt = \lambda^2\frac{3\pi}{4}\Big(\frac{16}{9}-1\Big) > 0 = J(\widehat{x}).$$

Значение функционала на x_λ больше чем на $\widehat{x}$, значит, $\widehat{x}$ не доставляет слабого локального максимума.

Поскольку $J(x_\lambda(\cdot)) \to +\infty$ при $\lambda \to \infty$, то $S_{\text{absmax}} = +\infty$.

Из этого примера видно, что уравнение Эйлера — необходимое, но не достаточное условие экстремума.

Ответ. $S_{\text{absmin}} = -\infty$; $S_{\text{absmax}} = +\infty$; допустимая экстремаль $\widehat{x} \equiv 0 \notin \text{wlocextr}$.

2 Задача Больца

2.1 Постановка задачи

Задачей Больца[8] называется следующая экстремальная задача без ограничений в пространстве $C^1([t_0;\, t_1])$:

$$B(x(\cdot)) = \int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt + l(x(t_0), x(t_1)) \to \text{extr}. \qquad (P)$$

Здесь $L = L(t, x, \dot{x})$ — заданная функция трех, а $l = l(x(t_0), x(t_1))$ — заданная функция двух переменных. Задача Больца — элементарная задача вариационного исчисления. Функционал B называется *функционалом Больца*, функция l — *терминантом*. Любые функции класса $C^1([t_0;\, t_1])$ являются *допустимыми* в задаче.

Определение. Говорим, что допустимая функция $\widehat{x}(\cdot)$ доставляет *слабый локальный минимум* в задаче (P), и пишем $\widehat{x}(\cdot) \in \operatorname{wlocmin} P$, если существует $\delta > 0$ такое, что $B(x(\cdot)) \geq B(\widehat{x}(\cdot))$ для любой допустимой функции $x(\cdot)$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^1([t_0;\, t_1])} < \delta$.

2.2 Необходимое условие экстремума

Теорема. *Пусть функция $\widehat{x}(\cdot)$ доставляет слабый локальный экстремум в задаче (P) ($\widehat{x}(\cdot) \in \operatorname{wlocextr} P$), функции $L, L_x, L_{\dot{x}}$ — непрерывны в некоторой окрестности расширенного графика $\Gamma_{\widehat{x}\dot{\widehat{x}}}$ ($L, L_x, L_{\dot{x}} \in C(\mathcal{O}(\Gamma_{\widehat{x}\dot{\widehat{x}}}))$), функция l — непрерывно дифференцируема в окрестности точки $(\widehat{x}(t_0), \widehat{x}(t_1))$ ($l \in C^1(\mathcal{O}(\widehat{x}(t_0), \widehat{x}(t_1)))$. Тогда $\widehat{L}_{\dot{x}}(\cdot)$ — непрерывно дифференцируемая функция ($\widehat{L}_{\dot{x}}(\cdot) \in C^1([t_0;\, t_1])$) и выполнены*

a) уравнение Эйлера

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \quad \forall\, t \in [t_0;\, t_1];$$

b) условия трансверсальности

$$\widehat{L}_{\dot{x}}(t_0) = \widehat{l}_{x(t_0)},\ \widehat{L}_{\dot{x}}(t_1) = -\widehat{l}_{x(t_1)}.$$

[8] Больц, Оскар (1857–1942) — немецкий математик.

Доказательство. Возьмём произвольную функцию $h(\cdot) \in C^1([t_0;\, t_1])$ и зафиксируем её. Поскольку $\widehat{x} \in \operatorname{locextr} P$, то функция одного переменного

$$\varphi(\lambda) := B(\widehat{x}(\cdot) + \lambda h(\cdot)) = \int\limits_{t_0}^{t_1} L\big(t, \widehat{x}(t) + \lambda h(t), \dot{\widehat{x}}(t) + \lambda \dot{h}(t)\big) dt + \\ + l\big(\widehat{x}(t_0) + \lambda h(t_0), \widehat{x}(t_1) + \lambda h(t_1)\big)$$

имеет экстремум при $\lambda = 0$. Тогда по теореме Ферма $\varphi'(0) = 0$. Дифференцируя функцию φ и полагая $\lambda = 0$, получаем:

$$\varphi'(0) = \int\limits_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}}(t)\dot{h}(t) + \widehat{L}_x(t)h(t)\big) dt + \widehat{l}_{x(t_0)} h(t_0) + \widehat{l}_{x(t_1)} h(t_1) = 0 \\ \forall\ h(\cdot) \in C^1([t_0;\, t_1]). \quad (1)$$

Равенство (1) выполняется для любой функции $h(\cdot) \in C^1([t_0;\, t_1])$, а, значит, и для функций $h(\cdot) \in C_0^1([t_0;\, t_1])$. Следовательно, из соотношения (1) вытекает, что

$$\int\limits_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}}(t)\dot{h}(t) + \widehat{L}_x(t)h(t)\big) dt = 0 \quad \forall\ h(\cdot) \in C_0^1([t_0;\, t_1]).$$

Отсюда по лемме Дюбуа-Реймона функция $\widehat{L}_{\dot{x}}(\cdot) \in C^1([t_0;\, t_1])$ и выполняется уравнение Эйлера

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \quad \forall\ t \in [t_0;\, t_1].$$

Для завершения доказательства теоремы осталось вывести условия трансверсальности. Проинтегрируем по частям первый интеграл в соотношении (1) (интегрирование по частям возможно в силу уже доказанной непрерывной дифференцируемости $\widehat{L}_{\dot{x}}(\cdot)$):

$$\int\limits_{t_0}^{t_1} \widehat{L}_{\dot{x}}(t)\dot{h}(t)\, dt = \int\limits_{t_0}^{t_1} \widehat{L}_{\dot{x}}(t) dh(t) = \widehat{L}_{\dot{x}}(t)h(t)\Big|_{t_0}^{t_1} - \int\limits_{t_0}^{t_1} h(t) \frac{d}{dt}\widehat{L}_{\dot{x}}(t)\, dt.$$

Подставляя полученное выражение в соотношение (1) и учитывая уже доказанное уравнение Эйлера, имеем

$$\int\limits_{t_0}^{t_1}\Big(-\frac{d}{dt}\widehat{L}_{\dot{x}}(t)+\widehat{L}_x(t)\Big)h(t)\,dt+$$
$$+\widehat{L}_{\dot{x}}(t_1)h(t_1)-\widehat{L}_{\dot{x}}(t_0)h(t_0)+\widehat{l}_{x(t_0)}h(t_0)+\widehat{l}_{x(t_1)}h(t_1)=$$
$$=\big(\widehat{L}_{\dot{x}}(t_1)+\widehat{l}_{x(t_1)}\big)h(t_1)+\big(-\widehat{L}_{\dot{x}}(t_0)+\widehat{l}_{x(t_0)}\big)h(t_0)=0$$
$$\forall\ h(\cdot)\in C^1([t_0;\,t_1]). \quad (3.1)$$

Подставляя в (2) последовательно $h(t)=t-t_1$ и $h(t)=t-t_0$, придем к условиям трансверсальности $\widehat{L}_{\dot{x}}(t_0)=\widehat{l}_{x(t_0)}$ и $\widehat{L}_{\dot{x}}(t_1)=-\widehat{l}_{x(t_1)}$. $\square$

2.3 Примеры

Пример. $B(x(\cdot)) = \int\limits_0^1 (\dot{x}^2 - x)\,dt + x^2(1) \to \text{extr}.$

Решение. Необходимые условия экстремума:

a) уравнение Эйлера для интегранта $L = \dot{x}^2 - x$

$$-\frac{d}{dt}L_{\dot{x}} + L_x = 0 \iff -\frac{d}{dt}2\dot{x} - 1 = 0 \iff$$
$$-2\ddot{x} - 1 = 0 \iff \ddot{x} = -\frac{1}{2};$$

b) условия трансверсальности для терминанта $l = x^2(1)$

$$\begin{cases} L_{\dot{x}}(0) = l_{x(0)}, \\ L_{\dot{x}}(1) = -l_{x(1)}, \end{cases} \iff \begin{cases} 2\dot{x}(0) = 0, \\ 2\dot{x}(1) = -2x(1), \end{cases} \iff \begin{cases} \dot{x}(0) = 0, \\ \dot{x}(1) = -x(1). \end{cases}$$

Решение уравнения Эйлера: $\dot{x} = -\dfrac{t}{2} + C_1$, $x = -\dfrac{t^2}{4} + C_1 t + C_2$. Из условий трансверсальности находим константы C_1 и C_2:

$$\begin{cases} \dot{x}(0) = 0, \\ \dot{x}(1) = -x(1), \end{cases} \iff \begin{cases} C_1 = 0, \\ -\dfrac{1}{2} = \dfrac{1}{4} - C_2 \Leftrightarrow C_2 = \dfrac{3}{4}. \end{cases}$$

Единственная допустимая экстремаль $\widehat{x}(t) = \dfrac{-t^2 + 3}{4}$.

Проверим, доставляет ли найденная допустимая экстремаль минимум или максимум в задаче. Сравним значение функционала $B(\widehat{x})$ со значением функционала $B(x)$ на произвольной допустимой функции x. Обозначим $h = x - \widehat{x}$. Тогда $x = \widehat{x} + h$, где $h \in C^1([0;\,1])$. Функционал B с квадратично-линейным интегрантом $L = \dot{x}^2 - x$ и квадратичной терминальной частью $l = x^2(1)$ является квадратично-линейным. Тогда для задачи без ограничений с квадратично-линейным

функционалом

$$B(\widehat{x}+h)-B(\widehat{x})=\frac{1}{2}B''(h)=\int\limits_0^1 \dot{h}^2\,dt+h^2(1)\geq 0.$$

Разность $B(\widehat{x}+h)-B(\widehat{x})$ неотрицательна, значит, $\widehat{x}\in$ absmin,

$$S_{\text{absmin}}=B(\widehat{x})=\int\limits_0^1 (\dot{\widehat{x}}^2-\widehat{x})\,dt+\widehat{x}^2(1)=$$

$$=\int\limits_0^1 \Big(\frac{t^2}{4}+\frac{t^2}{4}-\frac{3}{4}\Big)dt+\frac{1}{4}=\Big(\frac{t^3}{6}-\frac{3t}{4}\Big)\Big|_0^1+\frac{1}{4}=\frac{2-9+3}{12}=-\frac{1}{3}.$$

Покажем, что $S_{\text{absmax}}=+\infty$. Действительно, возьмем последовательность допустимых функций $x_n(t)=n$. Тогда

$$B(x_n(\cdot))=\int\limits_0^1 (\dot{x}_n^2-x_n)\,dt+x_n^2(1)=-n+n^2\to+\infty \text{ при } n\to\infty.$$

Ответ. Допустимая экстремаль $\dfrac{-t^2+3}{4}\in$ absmin, $S_{\text{absmin}}=-\dfrac{1}{3}$; $S_{\text{absmax}}=+\infty$.

3 Задача с подвижными концами

3.1 Постановка задачи

Задачей с подвижными концами называется следующая экстремальная задача в пространстве $C^1(\Delta) \times \mathbb{R}^2$:

$$J(\xi) = \int\limits_{t_0}^{t_1} f(t, x, \dot{x})\, dt + l_0(t_0, x(t_0), t_1, x(t_1)) \to \text{extr}; \qquad (P)$$

$$l_i(t_0, x(t_0), t_1, x(t_1)) = 0, \ i = 1, \ldots, m, \qquad (1)$$

где $\xi = (x(\cdot), t_0, t_1)$, Δ — заданный конечный отрезок, $t_0, t_1 \in \Delta$, $t_0 < t_1$. Здесь t_0, t_1 — концы отрезка интегрирования являются подвижными и поэтому значение как интегральной так и терминальной частей функционала зависят от величин t_0, t_1.

Частным случаем является задача, в которой один из концов или даже оба закреплены.

Элемент $\xi = (x(\cdot), t_0, t_1)$ называется *допустимым*, если $x \in C^1(\Delta)$, $t_0, t_1 \in \Delta$, $t_0 < t_1$, и выполняются условия (1) на концах.

Определение. Говорим, что допустимый элемент $\widehat{\xi} = (\widehat{x}(\cdot), \widehat{t}_0, \widehat{t}_1))$ доставляет *слабый локальный минимум* в задаче (P), и пишем $\widehat{\xi} \in \text{wlocmin}\, P$, если существует $\delta > 0$ такое, что $J(\xi) \geq J(\widehat{\xi})$ для любого допустимого элемента $\xi = (x(\cdot), t_0, t_1)$, для которого $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^1(\Delta)} < \delta$, $|t_0 - \widehat{t}_0| < \delta$, $|t_1 - \widehat{t}_1| < \delta$.

Аналогично определяется слабый локальный максимум.

3.2 Необходимые условия экстремума

Теорема. *Пусть элемент* $\widehat{\xi} = (\widehat{x}(\cdot), \widehat{t}_0, \widehat{t}_1)$ *доставляет слабый локальный экстремум в задаче* (P) $(\widehat{\xi} \in \operatorname{wlocextr} P)$, *функции* $L, L_x, L_{\dot{x}}$ *— непрерывны в некоторой окрестности расширенного графика* $\Gamma_{\widehat{x}\dot{\widehat{x}}}$[9] $(L, L_x, L_{\dot{x}} \in C(\mathcal{O}(\Gamma_{\widehat{x}\dot{\widehat{x}}})))$, *функции* l_i *— непрерывно дифференцируемы в окрестности точки* $(\widehat{t}_0, \widehat{x}(\widehat{t}_0), \widehat{t}_1, \widehat{x}(\widehat{t}_1))$ $(l_i \in C^1(\widehat{t}_0, \widehat{x}(\widehat{t}_0), \widehat{t}_1, \widehat{x}(\widehat{t}_1)))$, $i = 0, 1, \ldots, m$. *Тогда существует ненулевой вектор множителей Лагранжа* $\lambda = (\lambda_0, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$, $\lambda \neq 0$, *такой, что для функции Лагранжа*

$$\mathscr{L} = \mathscr{L}(x(\cdot), t_0, t_1) = \int\limits_{t_0}^{t_1} \lambda_0 f(t, x, \dot{x})\, dt + \sum_{i=0}^{m} \lambda_i l_i(t_0, x(t_0), t_1, x(t_1))$$

выполнены условия:

a) стационарности по x *— уравнение Эйлера для интегранта* $L = \lambda_0 f(t, x, \dot{x})$

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \ \forall\ t \in \Delta \iff -\frac{d}{dt}\lambda_0 \widehat{f}_{\dot{x}}(t) + \lambda_0 \widehat{f}_x(t) = 0;$$

b) трансверсальности по x *для терминанта* $l = \sum\limits_{i=0}^{m} \lambda_i l_i(t_0, x(t_0), t_1, x(t_1))$

$$\begin{cases} \widehat{L}_{\dot{x}}(\widehat{t}_0) = \widehat{l}_{x(t_0)}, \\ \widehat{L}_{\dot{x}}(\widehat{t}_1) = -\widehat{l}_{x(t_1)} \end{cases} \iff \begin{cases} \lambda_0 \widehat{f}_{\dot{x}}(\widehat{t}_0) = \widehat{l}_{x(t_0)}, \\ \lambda_0 \widehat{f}_{\dot{x}}(\widehat{t}_1) = -\widehat{l}_{x(t_1)}; \end{cases}$$

c) стационарности по подвижным концам (выписывается только для подвижных концов отрезка интегрирования):

$$\widehat{\mathscr{L}_{t_0}} = 0 \iff -\lambda_0 \widehat{f}(\widehat{t}_0) + \widehat{l}_{t_0} + \widehat{l}_{x(t_0)} \dot{\widehat{x}}(\widehat{t}_0) = 0,$$

$$\widehat{\mathscr{L}_{t_1}} = 0 \iff \lambda_0 \widehat{f}(\widehat{t}_1) + \widehat{l}_{t_1} + \widehat{l}_{x(t_1)} \dot{\widehat{x}}(\widehat{t}_1) = 0.$$

[9] $\Gamma_{\widehat{x}\dot{\widehat{x}}} := \{(t, \widehat{x}(t), \dot{\widehat{x}}(t)) \mid t \in \Delta\}$ — расширенный график.

Необходимые условия экстремума в задаче со подвижными концами непосредственно будут вытекать из необходимых условий экстремума в задаче Лагранжа п. 6.2.

3.3 Примеры

Пример. $J(x(\cdot),T)=\int\limits_0^T(\dot x^2+4x+4)\,dt\to\text{extr};\ x(0)=0.$

Решение. Функция Лагранжа

$$\mathscr{L}=\mathscr{L}(x(\cdot),T)=\int\limits_0^T\lambda_0(\dot x^2+4x+4)\,dt+\lambda_1x(0).$$

Необходимые условия экстремума:

a) уравнение Эйлера для интегранта $L=\lambda_0(\dot x^2+4x+4)$

$$-\frac{d}{dt}L_{\dot x}+L_x=0\iff-\frac{d}{dt}2\lambda_0\dot x+4\lambda_0=0\iff-2\lambda_0(\ddot x-2)=0;$$

b) трансверсальность по x для терминанта $l=\lambda_1x(0)$

$$\begin{cases}L_{\dot x}(0)=l_{x(0)},\\ L_{\dot x}(T)=-l_{x(T)},\end{cases}\iff\begin{cases}2\lambda_0\dot x(0)=\lambda_1,\\ 2\lambda_0\dot x(T)=0;\end{cases}$$

c) стационарность по T (выписываем только для подвижного конца отрезка интегрирования)

$$\mathscr{L}_T=0\iff\lambda_0\big(\dot x^2(T)+4x(T)+4\big)=0.$$

Если $\lambda_0=0$, то из b) следует, что $\lambda_1=0$ — все множители Лагранжа оказались нулями. Этого не может быть. Значит, $\lambda_0\neq0$.

Положим $\lambda_0=1$. Тогда условия a)–c) преобразуются к виду

$$\ddot x=2,\ \dot x(T)=0,\ x(T)=-1.$$

Найдем допустимую экстремаль. Из уравнения Эйлера и начального условия в нуле

$\ddot{x} = 2 \Leftrightarrow \dot{x} = 2t + C_1 \Leftrightarrow x = t^2 + C_1 t + C_2;$

$x(0) = 0 \Leftrightarrow C_2 = 0 \Rightarrow x = t^2 + C_1 t;$

Из полученных условий на конце T находим C_1, T:

$$\begin{cases} \dot{x}(T) = 0, \\ x(T) = -1, \end{cases} \iff \begin{cases} 2T + C_1 = 0, \\ T^2 + C_1 T = -1, \end{cases} \iff \begin{cases} \widehat{T} = 1, \\ C_1 = -2. \end{cases}$$

Таким образом, в задаче имеется единственный допустимый экстремальный элемент $\widehat{\xi} = (\widehat{x}(\cdot), \widehat{T}) = (t^2 - 2t, 1)$.

Покажем, что $\widehat{\xi}$ не доставляет локального экстремума, т. е. что в любой его окрестности существует другой допустимый элемент, на котором значение функционала J в точке ξ как больше, так и меньше значения функционала J в точке $\widehat{\xi}$. Действительно, возьмем допустимый элемент $\xi = (x(\cdot), T) = (t^2 - 2t,\, T)$. Тогда $\xi \to \widehat{\xi}$ при $T \to \widehat{T}$, а

$$J(\xi) = \int\limits_0^T (\dot{x}^2 + 4x + 4)\, dt = \int\limits_0^T \big((2t-2)^2 + 4(t^2 - 2t) + 4\big) dt =$$

$$= \int\limits_0^T (8t^2 - 16t + 8)\, dt = 8 \int\limits_0^T (t-1)^2\, dt > J(\widehat{\xi}) = 8 \int\limits_0^1 (t-1)^2\, dt$$

при $T > 1$ и $J(\xi) < J(\widehat{\xi})$ при $T < 1$, поскольку под знаком интеграла стоит неотрицательная функция. Значит, $\widehat{\xi} \notin \mathrm{locextr}$.

Покажем, что $S_{\mathrm{absmin}} = -\infty$. Действительно, возьмем последовательность допустимых элементов $\xi_n = (x_n(\cdot), T_n) = (-t, n)$. Тогда

$$J(x_n, T_n) = \int\limits_0^{T_n} (\dot{x}_n^2 + 4x_n + 4)\, dt = \int\limits_0^n (5 - 4t)\ dt \to -\infty \text{ при } n \to +\infty.$$

Покажем, что $S_{\mathrm{absmax}} = +\infty$. Действительно, возьмем последовательность допустимых элементов $\xi_n = (x_n(\cdot), T_n) = (nt, 1)$. Тогда

$$J(x_n(\cdot), T_n) = \int\limits_0^{T_n} (\dot{x}_n^2 + 4x_n + 4)\, dt = \int\limits_0^1 (n^2 + 4nt + 4)\, dt \to +\infty$$

при $n \to +\infty$.

Ответ. Допустимая экстремальная пара $(\widehat{x}(\cdot), \widehat{T}) = (t^2 - 2t, 1) \notin \mathrm{wlocextr}$, $S_{\mathrm{absmin}} = -\infty$; $S_{\mathrm{absmax}} = +\infty$.

4 Изопериметрическая задача

4.1 Постановка задачи

Изопериметрической задачей в вариационном исчислении называется следующая экстремальная задача в пространстве $C^1([t_0;\, t_1])$:

$$J_0(x(\cdot)) = \int\limits_{t_0}^{t_1} f_0(t, x(t), \dot{x}(t))\, dt \to \text{extr}; \qquad (P)$$

$$J_i(x(\cdot)) = \int\limits_{t_0}^{t_1} f_i(t, x(t), \dot{x}(t))\, dt = \alpha_i,\ i = 1, \ldots, m, \qquad (1)$$

$$x(t_0) = x_0,\ x(t_1) = x_1, \qquad (2)$$

где $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ — заданные числа.

Отрезок $[t_0;\, t_1]$ является фиксированным и конечным, $t_0 < t_1$. Ограничения вида (1) называются *изопериметрическими*. Экстремум в задаче рассматривается среди функций $x(\cdot) \in C^1([t_0;\, t_1])$, удовлетворяющих изопериметрическим условиям (1) и условиям (2) на концах. Такие функции называются *допустимыми*.

Определение. Говорим, что допустимая функция $\widehat{x}(\cdot)$ доставляет *слабый локальный минимум* в задаче (P), и пишем $\widehat{x}(\cdot) \in \operatorname{wlocmin} P$, если существует $\delta > 0$ такое, что $J_0(x(\cdot)) \geq J_0(\widehat{x}(\cdot))$ для любой допустимой функции $x(\cdot)$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^1([t_0;\, t_1])} < \delta$.

4.2 Необходимое условие экстремума

Теорема. *Пусть функция* $\widehat{x}(\cdot)$ *доставляет слабый локальный экстремум в задаче* (P) $(\widehat{x}(\cdot) \in \operatorname{wlocextr} P)$, *функции* $f_i, f_{ix}, f_{i\dot{x}}$, $i = 0, 1, \ldots, m$, *— непрерывны в некоторой окрестности расширенного графика* $\Gamma_{\widehat{x}\dot{\widehat{x}}}$ $(f_i, f_{ix}, f_{i\dot{x}} \in C(\mathcal{O}(\Gamma_{\widehat{x}\dot{\widehat{x}}})))$.

Тогда существует вектор множителей Лагранжа $\lambda = (\lambda_0, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$, $\lambda \neq 0$, *такой, что для лагранжиана* $L = \sum\limits_{i=0}^{m} \lambda_i f_i(t, x, \dot{x})$ *выполняется условие гладкости* $\widehat{L}_{\dot{x}}(\cdot) \in C^1([t_0;\, t_1])$ *и выполнено уравнение Эйлера*

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \quad \forall\ t \in [t_0;\, t_1].$$

Замечание. Если функционал в изопериметрической задаче является квадратично-линейным, а ограничения являются линейными, то имеет место формула для задачи с квадратично-линейным функционалом и линейными ограничениями типа равенств (см. Глава 1 п. 7.5), согласно которой, если $\widehat{x}$ — допустимая экстремаль, а $\widehat{x} + h$ — допустимая функция, то для квадратично-линейного функционала

$$J_0(\widehat{x} + h) = J_0(\widehat{x}) + \frac{1}{2} J_0''[h, h],$$

а для квадратичного функционала

$$J_0(\widehat{x} + h) = J_0(\widehat{x}) + J_0(h).$$

4.3 Примеры

Пример. $J_0(x) = \int\limits_0^1 \dot{x}^2\, dt \to \operatorname{extr}$; $\int\limits_0^1 x\, dt = 0$, $x(0) = 0$, $x(1) = 1$.

Решение. Лагранжиан $L = \lambda_0 \dot{x}^2 + \lambda_1 x$.

Необходимое условие экстремума — уравнение Эйлера для лагранжиана L

$$-\frac{d}{dt}L_{\dot{x}} + L_x = 0 \iff -\frac{d}{dt}2\lambda_0\dot{x} + \lambda_1 = 0 \iff -2\lambda_0\ddot{x} + \lambda_1 = 0.$$

Если $\lambda_0 = 0$, то $\lambda_1 = 0$ — все множители Лагранжа — нули. Этого не может быть.

Положим $\lambda_0 = 1$. Тогда $2\ddot{x} = \lambda_1$.
Общее решение $x = C_2t^2 + C_1t + C_0$.

Неизвестные константы C_0, C_1, C_2 находим из условий на концах и изопериметрического условия:

$x(0) = 0 \Rightarrow C_0 = 0,$

$x(1) = 1 \Rightarrow C_2 + C_1 = 1,$

$\int\limits_0^1 x\,dt = 0 \Rightarrow \int\limits_0^1 (C_2t^2 + C_1t)\,dt = 0 \Rightarrow \dfrac{C_2}{3} + \dfrac{C_1}{2} = 0.$

Отсюда $C_2 = 3$, $C_1 = -2$. Таким образом, в задаче имеется единственная допустимая экстремаль $\widehat{x} = 3t^2 - 2t$.

Исследуем найденную допустимую экстремаль. Возьмём произвольную допустимую в задаче функцию x. Обозначим $h = x - \widehat{x}$. Тогда $x = \widehat{x} + h$, где $h \in C^1([0;\,1])$, $\int\limits_0^1 h\,dt = 0$ и $h(0) = h(1) = 0$. В задаче с квадратично-линейным функционалом и линейными ограничениями типа равенств на экстремали $\widehat{x}$

$$J_0(\widehat{x} + h) - J_0(\widehat{x}) = J_0(h) = \int\limits_0^1 \dot{h}^2\,dt \geq 0.$$

Разность $J_0(\widehat{x}+h) - J_0(\widehat{x})$ неотрицательна, значит, $\widehat{x} \in \mathrm{absmin}$,

$$S_{\mathrm{absmin}} = \int\limits_0^1 \dot{\widehat{x}}^2\,dt = \int\limits_0^1 (6t-2)^2\,dt = \frac{36t^3}{3} - \frac{24t^2}{2} + 4t\Big|_0^1 = 4.$$

Покажем, что $S_{\mathrm{absmax}} = +\infty$. Действительно, возьмем последовательность допустимых функций $x_n = \widehat{x} + nh$, $h(\cdot) \not\equiv 0$.

Для того, чтобы последовательность функций была допустима необходимо, чтобы $h(\cdot) \in C_0^1([0;\,1])$, $\int\limits_0^1 h\,dt = 0$, например, возьмем $h(t) = \sin 2\pi t$. Тогда

$$J_0(x_n) = \int\limits_0^1 \dot{x}_n^2\,dt = \int\limits_0^1 \left(\dot{\widehat{x}} + n\dot{h}\right)^2 dt \to +\infty \text{ при } n \to +\infty.$$

Ответ. Допустимая экстремаль $3t^2 - 2t \in \mathrm{absmin}$, $S_{\mathrm{absmin}} = 4$; $S_{\mathrm{absmax}} = +\infty$.

4.4 Задача Дидоны

Одними из первых задач на отыскание наибольших и наименьших величин являлись изопериметрические задачи известные еще с древних времен о нахождении замкнутой кривой, имеющей заданную длину и охватывающую наибольшую площадь, и о нахождении пространственной замкнутой поверхности, имеющей заданную площадь и охватывающей наибольший объем. Еще до Аристотеля (IV век до н. э.) было известно, что среди изопериметрических (имеющих равную длину) кривых наиболее вместимой является окружность, а среди изопифанных (имеющих равную площадь) поверхностей — сфера.

Изопериметрическая задача содержится также в легенде о царице Дидоне. Описываемые события легенда относит к 825 году до н. э.

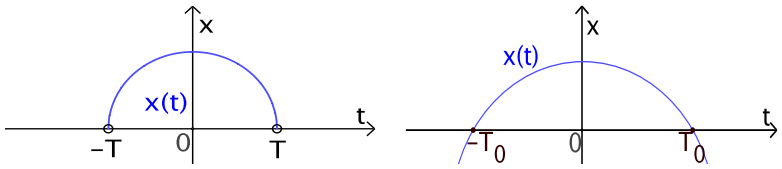
Финикийская царица Дидона и с ней небольшая часть жителей города Тира, спасаясь от преследований, покинули родной город и в поисках счастья отправились на кораблях на запад вдоль берегов Средиземного моря. Выбрав на африканском побережье удобное место (нынешний Тунисский залив), Дидона и её спутники решили основать здесь город. Эта идея не понравилась местным жителям, но все же финикийской царице удалось уговорить их предводителя Ярба, и он простодушно и неосторожно согласился уступить Дидоне клочок земли, “который можно окружить бычьей шкурой”. Хитрая финикиянка, разрезав шкуру на тонкие ремни, связала их в один длинный ремень и, окружив им значительную территорию, заложила на ней город Карфаген. В память об этой истории карфагенская цитадель получила название Бирса (шкура).

Из истории город Карфаген помнится нам еще войнами с Римом за обладание господством на Средиземном море (Пунические войны III–II век до н. э.), которые завершились взятием римлянами Карфагена и его разрушением.

Мы видим, что Дидона “решала” классическую изопериметрическую задачу о наибольшей вместимости. Естественно считать, что Дидона хотела сохранить выход к морю. Тогда мы получаем *первую задачу Дидоны*. Среди всех кривых длины l с концами на фиксированной прямой (прямолинейный берег), найти ту, которая ограничивает фигуру наибольшей площади. Формализованная задача имеет вид:

$$\int\limits_{-T}^{T} x\,dt \to \max; \quad \int\limits_{-T}^{T} \sqrt{1+\dot{x}^2}\,dt = l, \;\; x(-T) = x(T) = 0$$

(здесь T — подвижный конец). Эта задача относится к типу задач Лагранжа (см. § 6).



Давайте решим *вторую задачу Дидоны*, в которой оба конца кривой закреплены на прямой. Формализованная вторая задача Дидоны имеет вид:

$$\int\limits_{-T_0}^{T_0} x\,dt \to \max; \quad \int\limits_{-T_0}^{T_0} \sqrt{1+\dot{x}^2}\,dt = l, \;\; x(-T_0) = x(T_0) = 0$$

(здесь T_0 — фиксировано). Это — задача, укладывающаяся в схему изопериметрических задач п. 4.1.

5 Задача со старшими производными

5.1 Постановка задачи

Задачей со старшими производными в вариационном исчислении называется следующая экстремальная задача в пространстве $C^n([t_0;\, t_1])$:

$$J(x(\cdot)) = \int\limits_{t_0}^{t_1} L\big(t, x(t), \dot{x}(t), \ddot{x}(t), \ldots, x^{(n)}(t)\big)dt \to \operatorname{extr}; \qquad (P)$$

$$x^{(k)}(t_j) = x_{kj},\ k = 0, 1, \ldots, n-1,\ j = 0, 1. \qquad (1)$$

Здесь интегрант $L = L(t, x, \dot{x}, \ldots, x^{(n)})$ — функция $n+2$ переменных. Отрезок $[t_0;\, t_1]$ является фиксированным и конечным, $t_0 < t_1$. Экстремум в задаче рассматривается среди функций $x(\cdot) \in C^n([t_0;\, t_1])$, удовлетворяющих условиям на концах (1); такие функции называются *допустимыми*.

Определение. Говорим, что допустимая функция $\widehat{x}(\cdot)$ доставляет *слабый локальный минимум* в задаче (P), и пишем $\widehat{x}(\cdot) \in \operatorname{wlocmin} P$, если существует $\delta > 0$ такое, что $J(x(\cdot)) \geq J(\widehat{x}(\cdot))$ для любой допустимой функции $x(\cdot)$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^n([t_0;\, t_1])}$[10] $< \delta$.

5.2 Вывод уравнения Эйлера–Пуассона с помощью леммы Лагранжа

Теорема. *Пусть функция $\widehat{x}(\cdot)$ доставляет слабый локальный экстремум в задаче (P) ($\widehat{x}(\cdot) \in \operatorname{wlocextr} P$), функции $L, L_x,\ L_{\dot{x}}, \ldots,\ L_{x^{(n)}}$ — непрерывны в окрестности расширенного графика $\Gamma_{\widehat{x}\dot{\widehat{x}}\ldots\widehat{x}^{(n)}}$[11], $(L, L_x, \ldots, L_{x^{(n)}} \in C(\mathcal{O}(\Gamma_{\widehat{x}\dot{\widehat{x}}\ldots\widehat{x}^{(n)}})))$, $\widehat{L}_{x^{(k)}}(\cdot) \in C^k([t_0;\, t_1])$, $k = 1, \ldots, n$.*

---

[10] $\|y\|_{C^n([t_0;t_1])} := \max\big\{\|y\|_{C([t_0;t_1])}, \|\dot{y}\|_{C([t_0;t_1])}, \ldots, \|y^{(n)}\|_{C([t_0;t_1])}\big\}$

[11] $\Gamma_{\widehat{x}\dot{\widehat{x}}\ldots\widehat{x}^{(n)}} := \big\{(t, \widehat{x}(t), \dot{\widehat{x}}(t), \ldots, \widehat{x}^{(n)}(t)) \in \mathbb{R}^{n+2} \mid t \in [t_0;\, t_1]\big\}$

Тогда выполнено уравнение Эйлера–Пуассона

$$\sum_{k=0}^{n}(-1)^k \frac{d^k}{dt^k}\widehat{L}_{x^{(k)}}(t) = 0 \quad \forall\ t \in [t_0;\, t_1].$$

При $n = 1$ уравнение Эйлера–Пуассона совпадает с уравнением Эйлера. При $n = 2$ уравнение Эйлера–Пуассона выглядит следующим образом:

$$\frac{d^2}{dt^2}\widehat{L}_{\ddot{x}}(t) - \frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0.$$

5.3 Пример

$$J(x(\cdot)) = \int\limits_0^1 \ddot{x}^2\,dt \to \text{extr};\ x(0) = \dot{x}(0) = x(1) = 0,\ \dot{x}(1) = 1.$$

Решение. Необходимое условие — уравнение Эйлера–Пуассона

$$\frac{d^2}{dt^2}L_{\ddot{x}} - \frac{d}{dt}L_{\dot{x}} + L_x = 0 \iff \frac{d^2}{dt^2}2\ddot{x} = 0 \iff x^{(4)} = 0.$$

Общее решение уравнения Эйлера–Пуассона:

$$x = C_3t^3 + C_2t^2 + C_1t + C_0.$$

Неизвестные константы C_0, C_1, C_2, C_3 находим из условий на концах:

$$x(0) = 0 \Rightarrow C_0 = 0,$$
$$\dot{x}(0) = 0 \Rightarrow C_1 = 0,$$
$$\begin{cases} x(1) = 0, \\ \dot{x}(1) = 1, \end{cases} \iff \begin{cases} C_3 + C_2 = 0, \\ 3C_3 + 2C_2 = 1, \end{cases} \iff \begin{cases} C_2 = -1, \\ C_3 = 1. \end{cases}$$

Имеется единственная допустимая экстремаль $\widehat{x} = t^3 - t^2$.

Проверим, доставляет ли найденная допустимая экстремаль минимум или максимум в задаче. Сравним значение функционала $J(\widehat{x})$ со значением функционала $J(x)$ на произвольной допустимой функции x. Представим функцию x в виде суммы $x = \widehat{x} + h$, где функция h такова, что их сумма $\widehat{x} + h$ допустима. Для этого надо взять $h \in C^2([0;\,1])$, $h(0) = \dot{h}(0) = h(1) = \dot{h}(1) = 0$. Тогда в силу равенства

$$J(\widehat{x} + h) - J(\widehat{x}) = J(h) \quad \forall\ h \in C_0^2([0;\,1])$$

для квадратичного функционала J на экстремали $\widehat{x}$ имеем

$$J(\widehat{x} + h) - J(\widehat{x}) = \int\limits_0^1 \ddot{h}^2\,dt \geq 0. \qquad (*)$$

Разность $J(\widehat{x} + h) - J(\widehat{x})$ неотрицательна, значит, $\widehat{x} \in \text{absmin}$,

$$S_{\text{absmin}} = \int\limits_0^1 \ddot{\widehat{x}}^2\,dt = \int\limits_0^1 (6t - 2)^2\,dt = \int\limits_0^1 (36t^2 - 24t + 4)\,dt =$$

$$= 12t^3 - 12t^2 + 4t \Big|_0^1 = 4.$$

Покажем, что $S_{\text{absmax}} = +\infty$. Действительно, возьмем последовательность допустимых функций $x_n(\cdot) = \widehat{x}(\cdot) + nh(\cdot)$, $h \not\equiv 0$. Для того, чтобы последовательность функций была допустима необходимо, чтобы $h \in C_0^2([0;\,1])$, например, возьмем $h(t) = t^2(t-1)^2$. Тогда из равенства $(*)$

$$J(x_n) = J(\widehat{x} + nh) = J(\widehat{x}) + n^2 \int\limits_0^1 \ddot{h}^2\, dt \to +\infty \text{ при } n \to +\infty.$$

Ответ. Допустимая экстремаль $t^3 - t^2 \in \text{absmin}$, $S_{\text{absmin}} = 4$; $S_{\text{absmax}} = +\infty$.

6 Задача Лагранжа

Все задачи, изученные нами в предыдущих пунктах, являются частными случаями или могут быть сведены к задаче (мы сформулируем её ниже), поставленной Лагранжем в сочинении “Аналитическая механика” в 1788 году. Для её решения Лагранж использовал метод неопределенных множителей, который впоследствии стали называть методом множителей Лагранжа. Впрочем этот метод не был им аккуратно обоснован, и понадобилось более ста лет для того, чтобы придать рассуждениям Лагранжа вид строго доказанной теоремы.

6.1 Постановка задачи

Задачей Лагранжа называется экстремальная задача

$$B_0(\xi) \to \text{extr}; \qquad \begin{aligned} &B_i(\xi) \le 0,\ i = 1, \dots, m', \\ &B_i(\xi) = 0,\ i = m'+1, \dots, m, \end{aligned} \qquad (P)$$

$$\dot{x}_\alpha(t) - \varphi(t, x(t)) = 0 \quad \forall\, t \in \Delta, \qquad (*)$$

где $\xi = (x(\cdot), t_0, t_1)$, $x(\cdot) \in C^1(\Delta, \mathbb{R}^n)$[12], $t_0, t_1 \in \Delta$, $t_0 < t_1$, Δ — заданный конечный отрезок,

$$B_i(\xi) = \int\limits_{t_0}^{t_1} f_i(t, x(t), \dot{x}(t))\, dt + l_i(t_0, x(t_0), t_1, x(t_1)),\ i = 0, 1, \dots, m.$$

Условие $(*)$, называемое *дифференциальной связью*, может быть наложено не на все координаты вектор-функции $x(\cdot) = (x_1(\cdot), \dots, x_n(\cdot))$, а только на некоторые, для определенности на первые k координат: $\dot{x}_i(t) - \varphi_i(t, x(t)) = 0$, $i = 1, \dots, k$. Обозначим $x(\cdot) = (x_\alpha(\cdot), x_\beta(\cdot))$, где $x_\alpha(\cdot) = (x_1(\cdot), \dots, x_k(\cdot))$ —

[12] $C^1(\Delta, \mathbb{R}^n)$ — пространство вектор-функций $x(\cdot) = (x_1(\cdot), \dots, x_n(\cdot))$, где $x_i(\cdot) \in C^1(\Delta, \mathbb{R})$.

вектор с координатами, на которые наложена дифференциальная связь, $x_\beta(\cdot) = (x_{k+1}(\cdot), \ldots, x_n(\cdot))$ — вектор с координатами, на которые не наложена дифференциальная связь. Если дифференциальная связь отсутствует, то $k = 0$ и $x(\cdot) = x_\beta(\cdot)$.

Поскольку вместо $\dot{x}_\alpha(\cdot)$ в функции $f_i(t, x, \dot{x})$ можно подставить из $(*)$ равное ему выражение $\varphi(t, x)$, то в дальнейшем считаем, что $f_i = f_i(t, x, \dot{x}_\beta)$.

Частным случаем задачи (P) является задача, в которой один из концов t_0 или t_1 — подвижный, а другой закреплен, или оба конца отрезка интегрирования $[t_0; t_1]$ фиксированы.

Элемент ξ, для которого выполнены все указанные условия и ограничения задачи, называется *допустимым*.

Определение. Говорим, что допустимый элемент $\widehat{\xi} = (\widehat{x}(\cdot), \widehat{t}_0, \widehat{t}_1))$ доставляет *слабый локальный минимум* в задаче Лагранжа (P), и пишем $\widehat{\xi} \in \operatorname{wlocmin} P$, если существует $\delta > 0$ такое, что $B_0(\xi) \geq B_0(\widehat{\xi})$ для любого допустимого элемента $\xi = (x(\cdot), t_0, t_1)$, для которого $\|\xi - \widehat{\xi}\|_{C^1(\Delta)\times\mathbb{R}^2} < \delta$ $(\Longleftrightarrow \|x(\cdot) - \widehat{x}(\cdot)\|_{C^1(\Delta)} < \delta,\ |t_0 - \widehat{t}_0| < \delta,\ |t_1 - \widehat{t}_1| < \delta)$.

Аналогично формулируется определение слабого локального максимума.

6.2 Необходимые условия экстремума

Теорема (Эйлер–Лагранж)**.** *Пусть элемент* $\widehat{\xi} = (\widehat{x}(\cdot), \widehat{t}_0, \widehat{t}_1)$ *доставляет слабый локальный экстремум в задаче Лагранжа* (P) $(\widehat{\xi} \in \operatorname{wlocextr} P)$*, функции* $f_i, f_{ix}, f_{i\dot{x}}$ *непрерывны в некоторой окрестности расширенного графика* $\Gamma_{\widehat{x}\dot{\widehat{x}}}$[13] $(f_i, f_{ix}, f_{i\dot{x}} \in C(\mathcal{O}(\Gamma_{\widehat{x}\dot{\widehat{x}}})))$*,* $i = 0, 1, \ldots, m$*, функции* φ, φ_x *непрерывны в некоторой окрестности графика* $\Gamma_{\widehat{x}}$[14] $(\varphi, \varphi_x \in C(\mathcal{O}(\Gamma_{\widehat{x}}))$*, функции* l_i *непрерывно дифференцируемы в некоторой окрестности точки* $(\widehat{t}_0, \widehat{x}(\widehat{t}_0), \widehat{t}_1, \widehat{x}(\widehat{t}_1))$ $(l_i \in C^1(\mathcal{O}(\widehat{t}_0, \widehat{x}(\widehat{t}_0), \widehat{t}_1, \widehat{x}(\widehat{t}_1))))$*,* $i = 0, 1, \ldots, m$ (условие гладкости).

---

[13] $\Gamma_{\widehat{x}\dot{\widehat{x}}} := \{(t, \widehat{x}(t), \dot{\widehat{x}}(t)) \mid t \in \Delta\}$ — расширенный график.

[14] $\Gamma_{\widehat{x}} := \{(t, \widehat{x}(t)) \mid t \in \Delta\}$ — график.

Тогда найдутся множители Лагранжа $(\lambda, p(\cdot)) \in \mathbb{R}^{m+1} \times C^1(\Delta, \mathbb{R}^k)$, $\lambda \neq 0$, *такие, что для функции Лагранжа*

$$\mathscr{L}(x(\cdot), t_0, t_1) = \int\limits_{t_0}^{t_1} \Bigg(\overbrace{\sum_{i=0}^{m} \lambda_i f_i(t, x, \dot{x}_\beta)}^{f} + p(t)\big(\dot{x}_\alpha - \varphi(t, x)\big) \Bigg) dt + \\ + \underbrace{\sum_{i=0}^{m} \lambda_i l_i\big(t_0, x(t_0), t_1, x(t_1)\big)}_{l},$$

выполнены условия:

a) стационарности по x *— уравнение Эйлера*[15] *для лагранжиана* $L(t, x, \dot{x}) = f(t, x, \dot{x}_\beta) + p(\dot{x}_\alpha - \varphi(t, x))$:

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

b) трансверсальности[16] *по* x: $\begin{cases} \widehat{L}_{\dot{x}}(\widehat{t_0}) = \widehat{l}_{x(t_0)}, \\ \widehat{L}_{\dot{x}}(\widehat{t_1}) = -\widehat{l}_{x(t_1)}; \end{cases}$

c) стационарности по подвижным концам (выписывается только для подвижных концов отрезка интегрирования) [17]:

$\widehat{\mathscr{L}_{t_0}} = 0, \quad \widehat{\mathscr{L}_{t_1}} = 0;$

d) дополняющей нежесткости: $\lambda_i B_i(\widehat{\xi}) = 0,\ i = 1, \ldots, m'$;

---

[15] Уравнения Эйлера для переменных x_α и x_β распишутся следующим образом: $\begin{cases} -\dot{p}(t) - p(t)\widehat{\varphi}_{x_\alpha}(t) + \widehat{f}_{x_\alpha}(t) = 0, \\ -\dfrac{d}{dt}\widehat{f}_{\dot{x}_\beta}(t) - p(t)\widehat{\varphi}_{x_\beta}(t) + \widehat{f}_{x_\beta}(t) = 0. \end{cases}$

[16] Условия трансверсальности можно более подробно расписать в следующем виде: $\begin{cases} \widehat{L}_{\dot{x}_\alpha}(\widehat{t_0}) = \widehat{l}_{x_\alpha(t_0)}, \\ \widehat{L}_{\dot{x}_\alpha}(\widehat{t_1}) = -\widehat{l}_{x_\alpha(t_1)}, \\ \widehat{L}_{\dot{x}_\beta}(\widehat{t_0}) = \widehat{l}_{x_\beta(t_0)}, \\ \widehat{L}_{\dot{x}_\beta}(\widehat{t_1}) = -\widehat{l}_{x_\beta(t_1)}, \end{cases} \iff \begin{cases} p(\widehat{t_0}) = \widehat{l}_{x_\alpha(t_0)}, \\ p(\widehat{t_1}) = -\widehat{l}_{x_\alpha(t_1)}, \\ \widehat{f}_{\dot{x}_\beta}(\widehat{t_0}) = \widehat{l}_{x_\beta(t_0)}, \\ \widehat{f}_{\dot{x}_\beta}(\widehat{t_1}) = -\widehat{l}_{x_\beta(t_1)}. \end{cases}$

[17] Условия стационарности расписываются следующим образом:
$-\widehat{f}(\widehat{t_0}) + \widehat{l}_{t_0} + \widehat{l}_{x(t_0)}\dot{\widehat{x}}(\widehat{t_0}) = 0, \quad \widehat{f}(\widehat{t_1}) + \widehat{l}_{t_1} + \widehat{l}_{x(t_1)}\dot{\widehat{x}}(\widehat{t_1}) = 0.$

e) знакоопределенности:

$$\lambda_0 \geq 0, \quad \begin{array}{ll} \lambda_i \geq 0,\ i = 1, \ldots, m', & \textit{для локального минимума}, \\ \lambda_i \leq 0,\ i = 1, \ldots, m', & \textit{для локального максимума}. \end{array}$$

Доказательство. Проведем доказательство для случая минимума. Случай максимума рассматривается аналогично.

Поскольку равенство $B_i = 0$ можно заменить двумя неравенствами $B_i \leq 0$, $-B_i \leq 0$, то в дальнейшем для простоты записи считаем, что имеются только ограничения типа неравенств и $m' = m$.

Доказательство основано на правиле множителей Лагранжа для гладких задач с ограничениями типа равенств и неравенств в банаховых пространствах (Глава 1, п. 8.2). Напомним эту теорему для регулярного случая.

Пусть X, Y — банаховы пространства, отображение $F\colon X \to Y$, функционалы $f_i\colon X \to \mathbb{R}$, $i = 0, 1, \ldots, m$. Рассмотрим задачу

$$f_0(x) \to \min; \quad f_i(x) \leq 0,\ i = 1, \ldots, m,\ F(x) = 0. \qquad (P)$$

Теорема. *Пусть точка* $\widehat{x} \in \operatorname{locextr} P$, *отображения* $F, f_i \in SD(\widehat{x})$, $i = 0, 1, \ldots, m$, $\operatorname{Im} F'(\widehat{x}) = Y$.

Тогда существуют множители Лагранжа — вектор $\lambda \in \mathbb{R}^{m+1}$, $\lambda \neq 0$, *и функционал* $y^* \in Y^*$ *такие, что для функции Лагранжа* $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x) + \langle y^*, F(x) \rangle$ *выполняются условия:*

a) стационарности: $\mathscr{L}'(\widehat{x}) = 0$;

b) дополняющей нежёсткости: $\lambda_i f_i(\widehat{x}) = 0,\ i = 1, \ldots, m$;

c) неотрицательности: $\lambda_i \geq 0,\ i = 0, 1, \ldots, m$.

Покажем, что все условия теоремы о правиле множителей Лагранжа в полученной задаче

$$B_0(\xi) \to \min;\ B_i(\xi) \leq 0,\ i = 1, \ldots, m,\ F(\xi) = 0. \qquad (\widetilde{P})$$

выполняются. Здесь $X = C^1(\Delta, \mathbb{R}^n) \times \mathbb{R}^2$, $Y = C(\Delta, \mathbb{R}^k)$ — банаховы пространства. Отображение $F\colon X \to Y$ действует по формуле

$$F(\xi) = F(x(\cdot), t_0, t_1) = \dot{x}_\alpha(t) - \varphi(t, x(t)).$$

Из непрерывной дифференцируемости функций следует, что функционалы B_i и отображение F строго дифференцируемы в точке $\widehat{\xi}$ — условие гладкости выполняется.

Покажем, что $\operatorname{Im} F'(\widehat{\xi}) = C(\Delta, \mathbb{R}^k)$. Производная

$$F'(\widehat{\xi})[h(\cdot), \tau_0, \tau_1] = \dot{h}(t) - \widehat{\varphi}_{x_\alpha}(t)h(t).$$

Уравнение $F'(\widehat{\xi})[\xi] = y$ является системой линейных дифференциальных уравнений с непрерывными коэффициентами

$$\dot{h}(t) - \widehat{\varphi}_{x_\alpha}(t)h(t) = y(t). \tag{1}$$

Такая система для любой функции $y \in C(\Delta, \mathbb{R}^k)$ с любым граничным условием в форме Коши имеет решение. Возьмем граничное условие $h(\widehat{t}_0) = \gamma$. Таким образом, $\operatorname{Im} F'(\widehat{\xi}) = Y$.

Согласно этой теореме существуют множители Лагранжа $(\lambda, y^*) \in \mathbb{R}^{m+1} \times Y^*$, причём $\lambda \neq 0$ при $\operatorname{Im} F'(\widehat{\xi}) = Y$, такие, что для функции Лагранжа

$$\widetilde{\mathscr{L}}(\xi) = \sum_{i=0}^{m} \lambda_i B_i(\xi) + \langle y^*, F(\xi)\rangle = \widetilde{\mathscr{L}}(x(\cdot), t_0, t_1) =$$

$$= \int\limits_{t_0}^{t_1} f(t, x, \dot{x}_\beta)\, dt + l(t_0, x(t_0), t_1, x(t_1)) + \big\langle y^*, \dot{x}_\alpha(t) - \varphi(t, x(t))\big\rangle$$

выполняются условия:

a′) стационарности

$$\widehat{\widetilde{\mathscr{L}}}_\xi = 0 \iff \begin{array}{l} \widehat{\widetilde{\mathscr{L}}}_x = 0 \Leftrightarrow \widehat{\widetilde{\mathscr{L}}}_{x_\alpha} = 0,\ \widehat{\widetilde{\mathscr{L}}}_{x_\beta} = 0, \\ \widehat{\widetilde{\mathscr{L}}}_{t_0} = 0, \\ \widehat{\widetilde{\mathscr{L}}}_{t_1} = 0; \end{array}$$

b′) дополняющей нежесткости: $\lambda_i B_i(\widehat{\xi}) = 0,\ i = 1, \ldots, m$;

c′) неотрицательности: $\lambda_i \geq 0,\ i = 0, 1, \ldots, m$.

Покажем, что из уравнения $\widehat{\widetilde{\mathscr{L}}}_{x_\alpha} = 0$ следует существование функции $p \in C^1(\Delta, \mathbb{R}^k)$ такой, что

$$\langle y^*, y\rangle = \int\limits_{\widehat{t}_0}^{\widehat{t}_1} py\,dt \ \forall\ y \in C(\Delta, \mathbb{R}^k)$$

и для которой выполняются условия a),b) для x_α теоремы Эйлера–Лагранжа. Тогда $\widetilde{\mathscr{L}} = \mathscr{L}$ и теорема будет доказана. Уравнения Эйлера и условия трансверсальности по x_β будут вытекать из условия стационарности функции Лагранжа $\mathscr{L}$ по x_β. Они выводятся как и для задачи Больца.

Распишем условие стационарности $\widetilde{\mathscr{L}}$ по x_α:

$$\widehat{\widetilde{\mathscr{L}}}_{x_\alpha} = 0 \iff \widehat{\widetilde{\mathscr{L}}}_{x_\alpha}[h] = 0 \quad \forall\ h \in C^1(\Delta, \mathbb{R}^k) \iff$$

$$\int\limits_{\widehat{t}_0}^{\widehat{t}_1} \widehat{f}_{x_\alpha} h\,dt + \widehat{l}_{x_\alpha(t_0)} h(\widehat{t}_0) + \widehat{l}_{x_\alpha(t_1)} h(\widehat{t}_1) + \langle y^*, \dot{h} - \widehat{\varphi}_{x_\alpha} h\rangle = 0.$$

Отсюда в силу соотношения (1) ($\dot{h} - \widehat{\varphi}_{x_\alpha} h = y$)

$$\langle y^*, y\rangle = -\int\limits_{\widehat{t}_0}^{\widehat{t}_1} \widehat{f}_{x_\alpha} h\,dt - \widehat{l}_{x_\alpha(t_0)}\gamma - \widehat{l}_{x_\alpha(t_1)} h(\widehat{t}_1)$$
$$\forall\ y \in C(\Delta, \mathbb{R}^k),\ \forall\ \gamma \in \mathbb{R}^k. \qquad (2)$$

Определим функцию p из условий:

$$-\dot{p} - p\widehat{\varphi}_{x_\alpha} + \widehat{f}_{x_\alpha} = 0, \quad p(\widehat{t}_1) = -\widehat{l}_{x_\alpha(t_1)}. \qquad (3)$$

По теореме существования и единственности решения задачи Коши для линейной неоднородной системы функция $p \in C^1(\Delta, \mathbb{R}^k)$ определяется нашими условиями однозначно. Таким образом, уравнение Эйлера по x_α и условие стационарности по x_α в точке $\widehat{t}_1$ выполняются.

В силу соотношений (3) и (1) имеем:

$$\int\limits_{\widehat{t}_0}^{\widehat{t}_1} \frac{d}{dt}(ph)\,dt = p(\widehat{t}_1)h(\widehat{t}_1) - p(\widehat{t}_0)h(\widehat{t}_0) = \int\limits_{\widehat{t}_0}^{\widehat{t}_1} (\dot{p}h + p\dot{h})\,dt =$$

$$= \int\limits_{\widehat{t}_0}^{\widehat{t}_1} \big((\widehat{f}_{x_\alpha} - p\widehat{\varphi}_{x_\alpha})h + p(y + \widehat{\varphi}_{x_\alpha}h)\big)\,dt = \int\limits_{\widehat{t}_0}^{\widehat{t}_1} \big(\widehat{f}_{x_\alpha}h + py\big)\,dt.$$

Находя из последнего соотношения $\int\limits_{\widehat{t}_0}^{\widehat{t}_1} \widehat{f}_{x_\alpha}h\,dt$ и подставляя полученное выражение в (2), получим

$$\langle y^*, y\rangle = \int\limits_{\widehat{t}_0}^{\widehat{t}_1} py\,dt - p(\widehat{t}_1)h(\widehat{t}_1) + p(\widehat{t}_0)h(\widehat{t}_0) - \widehat{l}_{x_\alpha(t_0)}\gamma - \widehat{l}_{x_\alpha(t_1)}h(\widehat{t}_1) \overset{(3)}{=}$$

$$= \int\limits_{\widehat{t}_0}^{\widehat{t}_1} py\,dt + \gamma\big(p(\widehat{t}_0) - \widehat{l}_{x_\alpha(t_0)}\big) \quad \forall\, y \in C(\Delta, \mathbb{R}^k),\ \forall\, \gamma \in \mathbb{R}^k.$$

Поскольку функция y и вектор γ определяются независимо, то значение функционала $\langle y^*, y\rangle$ не зависит от γ, и, значит, коэффициент при γ должен обращаться в ноль. Откуда следует, что

$$\langle y^*, y(\cdot)\rangle = \int\limits_{\widehat{t}_0}^{\widehat{t}_1} p(t)y(t)\,dt,\ p(\widehat{t}_0) = \widehat{l}_{x_\alpha(t_0)}.$$

Таким образом, $\widetilde{\mathscr{L}} = \mathscr{L}$. Теорема полностью доказана. $\square$

6.3 Примеры

Пример 1. $J(x(\cdot)) = \int\limits_0^1 \dot{x}^2\,dt \to \mathrm{extr}; \int\limits_0^1 x\,dt = 0,\ x(1) = 1.$

Решение. Выписанная задача не является изопериметрической задачей, поскольку не задано граничное условие функции $x(\cdot)$ в нуле. Задачу надо решать как задачу Лагранжа.

Функция Лагранжа $\mathscr{L} = \int\limits_0^1 (\lambda_0\dot{x}^2 + \lambda_1 x)\,dt + \lambda_2(x(1) - 1)$.

Необходимые условия:

a) уравнение Эйлера для лагранжиана $L = \lambda_0\dot{x}^2 + \lambda_1 x$

$$-\frac{d}{dt}L_{\dot{x}} + L_x = 0 \iff -\frac{d}{dt}2\lambda_0\dot{x} + \lambda_1 = 0 \iff -2\lambda_0\ddot{x} + \lambda_1 = 0;$$

b) трансверсальность по x для терминанта $l = \lambda_2(x(1) - 1)$

$$\begin{cases} L_{\dot{x}}(0) = l_{x(0)}, \\ L_{\dot{x}}(1) = -l_{x(1)}, \end{cases} \iff \begin{cases} 2\lambda_0\dot{x}(0) = 0, \\ 2\lambda_0\dot{x}(1) = -\lambda_2. \end{cases}$$

Если $\lambda_0 = 0$, то из a) $\lambda_1 = 0$, а из b) $\lambda_2 = 0$ — все множители Лагранжа — нули. Этого не может быть.

Положим $\lambda_0 = 1$. Тогда $\ddot{x} = \frac{\lambda_1}{2}$. Общее решение $x = C_2t^2 + C_1t + C_0$. Неизвестные константы C_0, C_1, C_2 находим из условия трансверсальности $\dot{x}(0) = 0$, условия на конце в единице и изопериметрического условия:

$\dot{x}(0) = 0 \Rightarrow C_1 = 0,$

$$\begin{cases} x(1) = 1, \\ \int\limits_0^1 x\,dt = 0, \end{cases} \iff \begin{cases} C_2 + C_0 = 1, \\ \dfrac{C_2}{3} + C_0 = 0, \end{cases} \iff \begin{cases} C_0 = -\dfrac{1}{2}, \\ C_2 = \dfrac{3}{2}. \end{cases}$$

Единственная допустимая экстремаль $\widehat{x} = \dfrac{3t^2 - 1}{2}$.

Покажем с помощью непосредственной проверки, что функция $\widehat{x}$ доставляет абсолютный минимум в задаче. Возьмем функцию $h \in C^1([0;\,1])$ такую, чтобы $\widehat{x}+h$ была допустимой функцией. Для этого надо взять функцию h, для которой $\int\limits_0^1 h\,dt = 0$ и $h(1)=0$. Тогда в силу равенства для квадратичного функционала J в задаче с линейными ограничениями экстремали $\widehat{x}$ имеем

$$J(\widehat{x}+h) - J(\widehat{x}) = J(h) = \int\limits_0^1 \dot{h}^2\,dt \geq 0.$$

Разность $J(\widehat{x}+h) - J(\widehat{x})$ неотрицательна, значит, $\widehat{x} \in \mathrm{absmin}$,

$$S_{\mathrm{absmin}} = \int\limits_0^1 \dot{\widehat{x}}^2\,dt = \int\limits_0^1 (3t)^2\,dt = \int\limits_0^1 9t^2\,dt = \frac{9t^3}{3}\Big|_0^1 = 3.$$

Покажем, что $S_{\mathrm{absmax}} = +\infty$. Действительно, возьмем последовательность допустимых функций $x_n(\cdot) = \widehat{x}(\cdot) + nh(\cdot)$, $h \not\equiv 0$. Для того, чтобы последовательность функций была допустима необходимо, чтобы $h \in C^1([0;\,1])$, $h(1)=0$ и $\int\limits_0^1 h\,dt = 0$, например, возьмем $h(t) = \sin 2\pi t$. Тогда

$$J(x_n) = J(\widehat{x}+nh) = J(\widehat{x}) + J(nh) = J(\widehat{x}) + n^2\int\limits_0^1 \dot{h}^2 dt \to +\infty$$

при $n \to +\infty$.

Ответ. Допустимая экстремаль $\dfrac{3t^2-1}{2} \in \mathrm{absmin}$, $S_{\mathrm{absmin}} = 3$, $S_{\mathrm{absmax}} = +\infty$.

Пример 2. $J(x(\cdot)) = \int\limits_0^1 \ddot{x}^2\,dt \to \mathrm{extr}$; $x(0) = \dot{x}(0) = 0$, $x(1) = 1$.

Решение. Выписанная задача не является задачей со старшими производными, поскольку не задано значение $\dot{x}(1)$. Задачу надо свести к задаче Лагранжа, вводя вместо функции x вектор-функцию (x_1, x_2), и обозначения: $x_1 = x$, $x_2 = \dot{x}$. Тогда исходная задача сведется к задаче Лагранжа:

$$\int\limits_0^1 \dot{x}_2^2\,dt \to \mathrm{extr};\ \dot{x}_1 = x_2,\ x_1(0) = 0,\ x_2(0) = 0,\ x_1(1) = 1.$$

Функция Лагранжа

$$\mathscr{L} = \int\limits_0^1 \big(\lambda_0 \dot{x}_2^2 + p(t)(\dot{x}_1 - x_2)\big)dt + \lambda_1 x_1(0) + \lambda_2 x_2(0) + \lambda_3(x_1(1) - 1).$$

Необходимые условия:

a) система уравнений Эйлера для лагранжиана
$L = \lambda_0 \dot{x}_2^2 + p(\dot{x}_1 - x_2)$

$$\begin{cases} -\dfrac{d}{dt}L_{\dot{x}_1} + L_{x_1} = 0, \\ -\dfrac{d}{dt}L_{\dot{x}_2} + L_{x_2} = 0, \end{cases} \iff \begin{cases} -\dfrac{d}{dt}p = 0, \\ -\dfrac{d}{dt}2\lambda_0 \dot{x}_2 - p = 0, \end{cases} \iff \begin{cases} -\dot{p} = 0, \\ -2\lambda_0 \ddot{x}_2 - p = 0; \end{cases}$$

b) трансверсальность по x для терминанта
$l = \lambda_1 x_1(0) + \lambda_2 x_2(0) + \lambda_3(x_1(1) - 1)$

$$\begin{cases} L_{\dot{x}_1}(0) = l_{x_1(0)}, \\ L_{\dot{x}_1}(1) = -l_{x_1(1)}, \\ L_{\dot{x}_2}(0) = l_{x_2(0)}, \\ L_{\dot{x}_2}(1) = -l_{x_2(1)}, \end{cases} \iff \begin{cases} p(0) = \lambda_1, \\ p(1) = -\lambda_3, \\ 2\lambda_0 \dot{x}_2(0) = \lambda_2, \\ 2\lambda_0 \dot{x}_2(1) = 0. \end{cases}$$

Если $\lambda_0 = 0$, то из a) следует, что $p = 0$, а из b) $\lambda_1 = \lambda_2 = \lambda_3 = 0$ — все множители Лагранжа — нули.

Положим $\lambda_0 = 1$. Тогда из а)

$$\begin{cases} -\dot{p} = 0, \\ -2\ddot{x}_2 - p = 0, \end{cases} \Longleftrightarrow \begin{cases} p = C, \\ \ddot{x}_2 = -\frac{C}{2}, \end{cases} \Longrightarrow x_2^{(3)} = 0 \Longrightarrow x^{(4)} = 0.$$

Общее решение $x = C_3t^3 + C_2t^2 + C_1t + C_0$.
Неизвестные константы $C_0, \ldots, C_3$ находим из условия трансверсальности $\dot{x}_2(1) = 0 \Leftrightarrow \ddot{x}(1) = 0$ и условий на концах:

$x(0) = 0 \Rightarrow C_0 = 0,$

$\dot{x}(0) = 0 \Rightarrow C_1 = 0,$

$$\begin{cases} x(1) = 1, \\ \ddot{x}(1) = 0 \end{cases} \Longleftrightarrow \begin{cases} C_3 + C_2 = 1, \\ 6C_3 + 2C_2 = 0, \end{cases} \Longleftrightarrow \begin{cases} C_2 = \frac{3}{2}, \\ C_3 = -\frac{1}{2}. \end{cases}$$

Единственная допустимая экстремаль $\widehat{x} = -\dfrac{t^3}{2} + \dfrac{3t^2}{2}$.

Покажем, что функция $\widehat{x}$ доставляет абсолютный минимум в задаче. Возьмем функцию $h \in C^2([0;\,1])$ такую, чтобы $\widehat{x}+h$ была допустимой функцией. Для этого надо взять функцию h, для которой $h(0) = h(1) = \dot{h}(0) = 0$. Тогда для задачи с квадратичным функционалом и линейными ограничениями

$$J(\widehat{x} + h) - J(\widehat{x}) = J(h) = \int\limits_0^1 \ddot{h}^2\,dt \geq 0.$$

Разность $J(\widehat{x} + h) - J(\widehat{x})$ неотрицательна, значит, $\widehat{x} \in \mathrm{absmin}$,

$$S_{\mathrm{absmin}} = \int\limits_0^1 \ddot{\widehat{x}}^2\,dt = \int\limits_0^1 (-3t + 3)^2\,dt = 3.$$

Покажем, что $S_{\mathrm{absmax}} = +\infty$. Действительно, возьмем последовательность допустимых функций $x_n(\cdot) = \widehat{x}(\cdot) + nh(\cdot)$, $h \not\equiv 0$, например, $h(t) = t^2(t-1)$. Тогда

$J(x_n) = J(\widehat{x} + nh) = J(\widehat{x}) + J(nh) = J(\widehat{x}) + n^2 \int\limits_0^1 \ddot{h}^2\,dt \to +\infty$ при $n \to +\infty$.

Ответ. Допустимая экстремаль $-\dfrac{t^3}{2} + \dfrac{3t^2}{2} \in \mathrm{absmin}; S_{\mathrm{absmin}} = 3$, $S_{\mathrm{absmax}} = +\infty$.

Глава 5

Условия второго порядка в вариационном исчислении

1 Простейшая задача вариационного исчисления

Рассмотрим простейшую задачу вариационного исчисления (для определенности задачу на минимум)

$$\int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt \to \min;\ x(t_0) = x_0,\ x(t_1) = x_1. \qquad (P)$$

Функционал задачи будем обозначать $J(x(\cdot))$, если не сказано иначе.

1.1 Сильный и слабый экстремум

Задачу (P) мы рассматривали на слабый экстремум. Иногда, чтобы подчеркнуть, что задача рассматривается на слабый

экстремум, будем писать $P(W)$. Множество допустимых элементов в задаче на слабый экстремум $D(P(W))$ составляют непрерывно дифференцируемые функции класса $C^1([t_0;\, t_1])$ с заданными условиями на концах. Напомним определение слабого экстремума.

Определение 1. Говорим, что функция $\widehat{x}(\cdot) \in D(P(W))$ доставляет *слабый локальный минимум* в задаче (P) (и пишем $\widehat{x}(\cdot) \in \operatorname{wlocmin} P$), если она доставляет локальный минимум в пространстве $C^1([t_0;\, t_1])$, т. е. если существует $\delta > 0$ такое, что $J(x(\cdot)) \geq J(\widehat{x}(\cdot))$ для любой функции $x(\cdot) \in D(P(W))$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^1([t_0;\, t_1])}$[1] $< \delta$.

Наряду со слабым экстремумом простейшую задачу вариационного исчисления будем рассматривать на сильный экстремум $P(S)$ (буква S — начальная буква слова strong — сильный). Множество допустимых элементов в задаче на сильный экстремум $D(P(S))$ составляют кусочно-дифференцируемые функции $(PC^1([t_0;\, t_1]))$ с заданными условиями на концах.

Определение 2. Говорим, что функция $\widehat{x}(\cdot) \in D(P(S))$ доставляет *сильный локальный минимум* в задаче (P) (и пишем $\widehat{x}(\cdot) \in \operatorname{strlocmin} P$), если она доставляет локальный минимум в пространстве $C([t_0;\, t_1])$, т. е. если существует $\delta > 0$ такое, что $J(x(\cdot)) \geq J(\widehat{x}(\cdot))$ для любой функции $x(\cdot) \in D(P(S))$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C([t_0;\, t_1])} < \delta$.

Так как множество функций, среди которых доставляется сильный экстремум, шире, чем для слабого экстремума, то если функция $\widehat{x}(\cdot) \in C^1([t_0;\, t_1])$ доставляет сильный, то она доставляет и слабый экстремум. Поэтому для функций $\widehat{x}(\cdot) \in C^1([t_0;\, t_1])$ необходимое условие слабого экстремума является необходимым условием сильного, а достаточное условие сильного экстремума является достаточным условием сла-

[1] напомним, что $\|y(\cdot)\|_{C^1([t_0;\, t_1])} := \max\{\|y(\cdot)\|_{C([t_0;\, t_1])}, \|\dot{y}(\cdot)\|_{C([t_0;\, t_1])}\}$, где $\|y(\cdot)\|_{C([t_0;\, t_1])} := \max\{|y(t)| \mid t \in [t_0;\, t_1]\}$.

бого:

$$\text{ДУ} \implies \widehat{x} \in \operatorname{strlocmin} P \implies \widehat{x} \in \operatorname{wlocmin} P \implies \text{НУ}$$

1.2 Пример слабого, но не сильного экстремума

Приведем пример задачи, в которой допустимая экстремаль доставляет слабый локальный минимум, но не доставляет сильного локального минимума.

Пример 1. $\displaystyle\int_0^1 \dot{x}^3\,dt \to \min;\quad x(0)=0,\ x(1)=1.$

Решение. Необходимое условие слабого, а значит и сильного экстремума — уравнение Эйлера:

$$-\frac{d}{dt}L_{\dot{x}}+L_x=0 \iff \frac{d}{dt}3\dot{x}^2=0 \iff 3\dot{x}^2=C \iff \dot{x}=\text{const}.$$

Общее решение уравнения Эйлера: $x=C_1t+C_0$. Из условий на концах находим, что $C_1=1$, $C_0=0$. Таким образом, имеется единственная допустимая экстремаль $\widehat{x}=t$.

Покажем, что $\widehat{x}\in\operatorname{wlocmin}$. Возьмём произвольную допустимую в задаче функцию x. Обозначим $h=x-\widehat{x}$. Тогда $x=\widehat{x}+h$, где $h\in C_0^1([0;\,1])$[2]. Имеем

$$J(x)-J(\widehat{x})=J(\widehat{x}+h)-J(\widehat{x})=\int_0^1(\dot{\widehat{x}}+\dot{h})^3\,dt-\int_0^1\dot{\widehat{x}}^3\,dt=$$

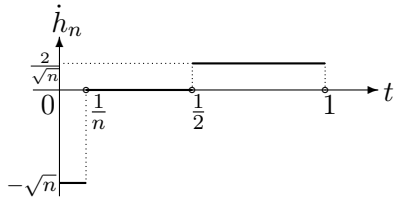
$$=\int_0^1(1+\dot{h})^3\,dt-\int_0^1 1^3\,dt=\int_0^1(3\dot{h}+3\dot{h}^2+\dot{h}^3)\,dt=\int_0^1\dot{h}^2(3+\dot{h})\,dt,\quad(*)$$

так как $\int_0^1\dot{h}\,dt=h(1)-h(0)=0$. Отсюда видно, что если $\|h\|_1<3$, то $3+\dot{h}(t)>0$ и, значит, $J(x)\geq J(\widehat{x})$, т. е. $\widehat{x}\in\operatorname{wlocmin}$.

[2] $C_0^1([t_0;\,t_1]):=\{h(\cdot)\in C^1([t_0;\,t_1])\mid h(t_0)=h(t_1)=0\}$.

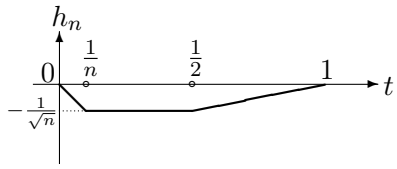
Покажем, что $\widehat{x}$ не доставляет сильного локального минимума ($\widehat{x} \notin \mathrm{strlocmin}$). Рассмотрим последовательность функций h_n ($n > 2$) такую, что

$$\dot{h}_n(t) = \begin{cases} -\sqrt{n}, & t \in \left[0;\, \frac{1}{n}\right], \\ 0, & t \in \left(\frac{1}{n};\, \frac{1}{2}\right), \\ \frac{2}{\sqrt{n}}, & t \in \left[\frac{1}{2};\, 1\right]. \end{cases}$$



Тогда $h_n(t) = \int\limits_0^t \dot{h}_n(\tau)d\tau =$

$$= \begin{cases} -\sqrt{n}\,t, & t \in \left[0;\, \frac{1}{n}\right], \\ -\frac{1}{\sqrt{n}}, & t \in \left(\frac{1}{n};\, \frac{1}{2}\right), \\ \frac{2t}{\sqrt{n}} - \frac{2}{\sqrt{n}}, & t \in \left[\frac{1}{2};\, 1\right]. \end{cases}$$



Функция $h_n \in PC_0^1([0;\,1])$ и $\|h_n\|_{C([0;\,1])} = \frac{1}{\sqrt{n}} \to 0$ при $n \to \infty$. Положим $x_n = \widehat{x} + h_n$. Получим последовательность допустимых (в задаче на сильный экстремум) функций x_n, $x_n(\cdot) \to \widehat{x}(\cdot)$ в метрике пространства $C([0;\,1])$, для которых в силу равенства $(*)$

$$J(x_n) - J(\widehat{x}) = \int\limits_0^1 \dot{h}_n^2(3 + \dot{h}_n)\,dt =$$

$$= \int\limits_0^{\frac{1}{n}} n(3-\sqrt{n})\,dt + \int\limits_{\frac{1}{2}}^1 \frac{4}{n}\left(3+\frac{2}{\sqrt{n}}\right)dt = 3-\sqrt{n}+\frac{6}{n}+\frac{4}{n\sqrt{n}} \to -\infty,$$

при $n \to \infty$, т. е. функция $\widehat{x}$ не доставляет сильного локального минимума, более того $S_{\mathrm{strabsmin}} = -\infty$.

Ниже в лемме о скруглении углов покажем, $S_{\mathrm{strabsmin}} = S_{\mathrm{wabsmin}}$. В нашей задаче можно было построить последовательность допустимых функций $\widetilde{x}_n$ класса C^1 такую, что $J(\widetilde{x}_n) \to -\infty$, т. е. сгладить функции x_n.

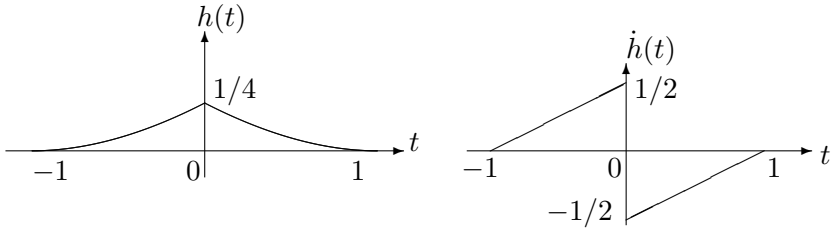
1.2.1 Лемма о скруглении углов

Для простоты записи сформулируем лемму и проведем доказательство для $n = 1$.

Лемма. *Пусть функция $\widehat{x}(\cdot) \in PC^1([t_0; t_1])$, интегрант L является непрерывной функцией. Тогда существует последовательность гладких функций $\{x_k(\cdot)\}_{k=1}^{\infty} \in C^1([t_0; t_1])$, $x_k(t_0) = \widehat{x}(t_0)$, $x_k(t_1) = \widehat{x}(t_1)$, такая, что $x_k(\cdot) \to \widehat{x}(\cdot)$ в метрике пространства $C([t_0; t_1])$ и $\lim\limits_{k\to\infty} J(x_k) = J(\widehat{x})$.*

Доказательство. Возьмём функцию

$$h(t) = \begin{cases} \frac{(t+1)^2}{4}, & -1 \le t \le 0, \\ \frac{(t-1)^2}{4}, & 0 \le t \le 1, \\ 0, & |t| > 1. \end{cases} \quad \text{Тогда } \dot{h}(t) = \begin{cases} \frac{t+1}{2}, & -1 \le t \le 0, \\ \frac{t-1}{2}, & 0 \le t \le 1, \\ 0, & |t| > 1. \end{cases}$$



Функция h непрерывна, а её производная в точке $t = 0$ имеет скачок $\dot{h}(-0) - \dot{h}(+0) = \lim\limits_{t\to-0} \dot{h}(t) - \lim\limits_{t\to+0} \dot{h}(t) = 1$, $|h(t)| \le \frac{1}{4}$, $|\dot{h}(t)| \le \frac{1}{2}$. Пусть $\tau_i \in (t_0; t_1)$, $i = 1, \ldots, m$, — точки разрыва производной $\dot{\widehat{x}}$ и $\Delta_i = \dot{\widehat{x}}(\tau_i - 0) - \dot{\widehat{x}}(\tau_i + 0)$ — её скачки в этих точках.

Функция $h_k(t, \tau_i) = \frac{1}{k} h\big(k(t - \tau_i)\big)$, график которой получается из графика функции h преобразованиями сдвига и подобия, отлична от нуля на интервале $\big(\tau_i - \frac{1}{k}, \tau_i + \frac{1}{k}\big)$ и непрерывна. Её производная непрерывна всюду кроме точки τ_i, где производная по-прежнему имеет скачок равный 1. Кроме того $|h_k(t, \tau_i)| \le \frac{1}{4k}$, $|\dot{h}_k(t, \tau_i)| \le \frac{1}{2}$.

Тогда функция $x_k(t) = \widehat{x}(t) - \sum\limits_{i=1}^{m} \Delta_i h_k(t, \tau_i)$ будет непрерывна вместе со своей производной на всем отрезке $[t_0;\, t_1]$. Непрерывность производной функции $x_k(t)$ в точках τ_i следует из того, что скачки производной функции $\widehat{x}(\cdot)$ равные Δ_i компенсируются скачками $-\Delta_i$ производной функции $h_k(\cdot,\tau_i)$. Отметим, что $x_k(t) = \widehat{x}(t)$ вне отрезков $\left[\tau_i - \frac{1}{k},\, \tau_i + \frac{1}{k}\right]$. В частности, для достаточно больших k (для определенности, $k > k_0$) эти отрезки не перекрываются и $x_k(t_0) = \widehat{x}(t_0)$, $x_k(t_1) = \widehat{x}(t_1)$. Поэтому

$$|x_k(t) - \widehat{x}(t)| = \left| \sum_{i=1}^{m} \Delta_i h_k(t, \tau_i) \right| \le \frac{1}{4k} \max_i |\Delta_i| = \frac{\Delta}{4k} \to 0$$

при $k \to +\infty$, $(\Delta := \max\limits_i |\Delta_i|)$, значит, $x_k(\cdot) \to \widehat{x}(\cdot)$ в метрике пространства $C([t_0;\, t_1])$;

$$|\dot{x}_k(t) - \dot{\widehat{x}}(t)| = \left| \sum_{i=1}^{m} \Delta_i \dot{h}_k(t, \tau_i) \right| \le \frac{1}{2} \max_i |\Delta_i| = \frac{\Delta}{2}.$$

На компакте $\left\{(t, x, \dot{x}) \mid t_0 \le t \le t_1,\, |x - \widehat{x}(t)| \le \frac{\Delta}{4},\, |\dot{x} - \dot{\widehat{x}}(t)| \le \frac{\Delta}{2}\right\}$ непрерывная функция L ограничена: $|L(t, x, \dot{x})| \le M$. Поэтому

$$|J(x_k) - J(\widehat{x})| = \left| \int\limits_{t_0}^{t_1} L(t, x_k, \dot{x}_k)\, dt - \int\limits_{t_0}^{t_1} L(t, \widehat{x}, \dot{\widehat{x}})\, dt \right| =$$

$$\overset{\text{при } k > k_0}{=} \left| \sum_{i=1}^{m} \int\limits_{\tau_i - \frac{1}{k}}^{\tau_i + \frac{1}{k}} \Big(L(t, x_k, \dot{x}_k) - L(t, \widehat{x}, \dot{\widehat{x}}) \Big) dt \right| \le \frac{4Mm}{k} \to 0$$

и, следовательно, $J(x_k) \to J(\widehat{x})$ при $k \to +\infty$. $\square$

Следствие. *Абсолютный экстремум в задаче* (P) *сильный и слабый совпадают:* $S_{\text{strabsmin}\, P} = S_{\text{wabsmin}\, P}$.

Для локальных экстремумов это может быть не так (см. п. 1.2.).

1.3 Теорема Боголюбова

Рассмотрим простейшую задачу вариационного исчисления

$$J(x) = \int\limits_{t_0}^{t_1} L(t, x, \dot{x})\, dt \to \min; \quad \begin{array}{l} x(t_0) = x_0, \\ x(t_1) = x_1. \end{array} \qquad (P)$$

Теорема (Боголюбов[3]). *Пусть $L\colon \mathbb{R}^{2n+1} \to \mathbb{R}$ — непрерывный интегрант, $\widetilde{L}(t, x, \cdot)$ — вторая сопряженная[4] (в смысле выпуклого анализа) функции $\dot{x} \to L(t, x, \dot{x})$. Рассмотрим задачу*

$$\widetilde{J}(x(\cdot) = \int\limits_{t_0}^{t_1} \widetilde{L}(t, x, \dot{x})\, dt \to \min; \quad \begin{array}{l} x(t_0) = x_0, \\ x(t_1) = x_1. \end{array} \qquad (\widetilde{P})$$

Тогда $S_{\operatorname{absmin} P} = S_{\operatorname{absmin} \widetilde{P}}$. Более того, для любой допустимой функции $x \in PC^1([t_0;\, t_1])$ существует последовательность гладких допустимых функций $\{x_k(\cdot) \in C^1([t_0;\, t_1])\}_{k \geq 1}$ такая, что $x_k(\cdot) \overset{C([t_0;\, t_1])}{\to} x(\cdot)$ и $\lim\limits_{k \to \infty} J(x_k) = \widetilde{J}(x)$.

Задачу на максимум можно свести к задаче на минимум, взяв вместо функционала J функционал $-J$.

$$-J(x) = \int\limits_{t_0}^{t_1} \big(- L(t, x, \dot{x}) \big)\, dt \to \min; \quad \begin{array}{l} x(t_0) = x_0, \\ x(t_1) = x_1. \end{array} \qquad (P')$$

При этом $S_{\operatorname{absmax} P} = -S_{\operatorname{absmin} P'}$. Далее рассмотрим задачу

$$\widetilde{(-J)}(x) = \int\limits_{t_0}^{t_1} \widetilde{(-L)}(t, x, \dot{x})\, dt \to \min; \quad \begin{array}{l} x(t_0) = x_0, \\ x(t_1) = x_1. \end{array} \qquad (\widetilde{P'})$$

---

[3] Н. Н. Боголюбов (1909–1992) — советский математик

[4] Напомним, что *сопряженной* функцией к функции $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ называется функция $f^*(y) := \sup\limits_x \{\langle y, x\rangle - f(x)\}$.

Функция $f^{**}(x) := \sup\limits_y \{\langle x, y\rangle - f^*(y)\}$, являющаяся сопряженной к f^*, называется *второй сопряженной* к f.

Вторую сопряженную надо брать к функции $-L$.
Таким образом, $S_{\operatorname{absmax} P} = -S_{\operatorname{absmin} \widetilde{(P')}}$.

1.4 Условия Лежандра, Якоби, Вейерштрасса

Рассмотрим простейшую задачу вариационного исчисления, для определенности, задачу на минимум

$$\int\limits_{t_0}^{t_1} L\left(t, x(t), \dot{x}(t)\right)\, dt \to \min;\ x(t_0) = x_0,\ x(t_1) = x_1. \qquad (P)$$

Пусть $\widehat{x}(\cdot) \in C^1([t_0;\, t_1])$ — некоторая фиксированная допустимая ($\widehat{x}(t_0) = x_0$, $\widehat{x}(t_1) = x_1$) экстремаль (т. е. функция $\widehat{x}$ удовлетворяет уравнению Эйлера). Далее предполагаем, что интегрант L по меньшей мере дважды непрерывно дифференцируем в некоторой окрестности расширенного графика $\Gamma_{\widehat{x}\dot{\widehat{x}}}$. Возьмём произвольную допустимую функцию $x \in C^1([t_0;\, t_1])$. Тогда функция $h = x - \widehat{x}$ принадлежит пространству $C_0^1([t_0;\, t_1])$ ($x = \widehat{x} + h$). По формуле Тейлора

$$J(x) = J(\widehat{x}+h) = J(\widehat{x}) + J'(\widehat{x})[h] + \frac{1}{2} J''(\widehat{x})[h,h] + o\big(\|h\|^2_{C^1([t_0;\, t_1])}\big).$$

Из этого разложения выводятся необходимые и достаточные условия экстремума.

Необходимое условие 1-го порядка локального экстремума

$$J'(\widehat{x})[h] = \int\limits_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}}(t)\dot{h}(t) + \widehat{L}_x(t)h(t)\big)dt = 0\ \forall\ h(\cdot) \in C_0^1([t_0;\, t_1])$$

равносильно, как было показано в Гл. 3 § 1, уравнению Эйлера.

Необходимое условие 2-го порядка локального минимума

$$J''(\widehat{x})[h,h] = \int\limits_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\widehat{L}_{x\dot{x}}(t)\dot{h}(t)h(t) + \widehat{L}_{xx}(t)h^2(t)\big)dt \geq 0$$
$$\forall\ h(\cdot) \in C_0^1([t_0;\, t_1]). \qquad (5.1)$$

Из последнего неравенства выводятся условия 2-го порядка минимума в простейшей задаче вариационного исчисления.

Важную роль для выполнения неравенства играет коэффициент $\widehat{L}_{\dot{x}\dot{x}}(t)$ при $\dot{h}^2$.

Определение 3. Говорим, что на экстремали $\widehat{x}$ выполнено *условие Лежандра*[5], если $\widehat{L}_{\dot{x}\dot{x}}(t) \geq 0 \ \forall\ t \in [t_0;\ t_1]$,
усиленное условие Лежандра, если $\widehat{L}_{\dot{x}\dot{x}}(t) > 0 \ \forall\ t \in [t_0;\ t_1]$.

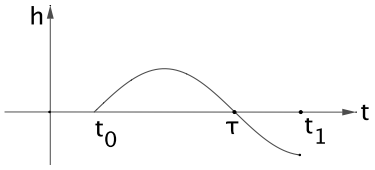
Пусть далее $\widehat{L}_{\dot{x}\dot{x}}, \widehat{L}_{\dot{x}x}, \widehat{L}_{xx} \in C^1([t_0;\ t_1])$ и выполнено усиленное условие Лежандра. Уравнение Эйлера по h для интегранта $\widetilde{L} = \widetilde{L}(t, h, \dot{h}) := \widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\widehat{L}_{x\dot{x}}(t)\dot{h}(t)h(t) + \widehat{L}_{xx}(t)h^2(t)$

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

называется *уравнением Якоби* для исходной задачи на экстремали $\widehat{x}$.

Уравнение Якоби — однородное линейное дифференциальное уравнение 2-го порядка, которое (из-за усиленного условия Лежандра) можно разрешить относительно второй производной.

Точка τ называется *сопряженной к точке* t_0, если для нетривиального решения уравнения Якоби $h(\cdot)$ с начальным условием $h(t_0) = 0$ функция h в точке τ обращается в ноль ($h(\tau) = 0$). Обычно второе граничное условие задается как $\dot{h}(t_0) = 1$ (или $\dot{h}(t_0) \neq 0$, поскольку все такие решения однородного уравнения пропорциональны и корни функции h одни и те же).

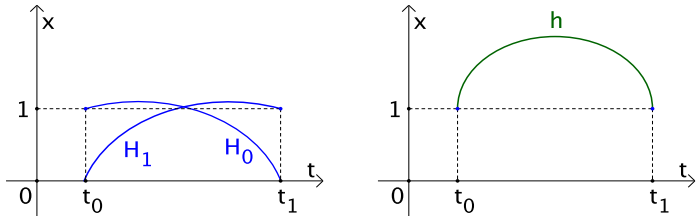


[5]Андриен Мари Лежандр — французский математик (1752-1833).

Определение 4. Говорят, что на $\widehat{x}$ 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.

Утверждение 1. *Если выполнено усиленное условие Якоби, то a) для любых $h_0, h_1 \in \mathbb{R}$ существует и единственно решение уравнения Якоби $h(\cdot)$, для которого $h(t_0) = h_0$, $h(t_1) = h_1$; b) существует решение уравнения Якоби $h(\cdot)$, для которого $h(t) > 0\ \forall\ t \in [t_0;\, t_1]$.*

Доказательство. Если выполнено усиленное условие Якоби, то существуют решения уравнения Якоби $H_0(\cdot)$, $H_1(\cdot)$ с граничными условиями $H_0(t_0) = 1$, $H_0(t_1) = 0$, $H_1(t_1) = 1$, $H_1(t_0) = 0$. Причём, из-за выполнения усиленного условия Якоби, эти функции не имеют нулей на полуинтервалах и, следовательно, сохраняют знак на отрезке $[t_0;\, t_1]$.



a) b)

a) Положим $h(t) = H_0(t)h_0 + H_1(t)h_1$. Тогда

$$h(t_0) = H_0(t_0)h_0 + H_1(t_0)h_1 = h_0,$$
$$h(t_1) = H_0(t_1)h_0 + H_1(t_1)h_1 = h_1.$$

Если существует другое решение уравнения Якоби $\overline{h} \neq h$, то функция $\widetilde{h} = \overline{h} - h \not\equiv 0$ тоже является решением уравнения Якоби с граничными условиями $\widetilde{h}(t_0) = \widetilde{h}(t_1) = 0$, что противоречит выполнению усиленного условия Якоби.

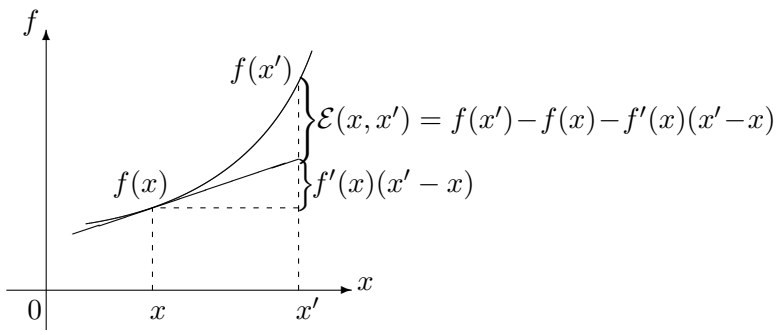
b) Очевидно, что для функции $h = H_0 + H_1$ выполняется условие $h(t) > 0\ \forall\ t \in [t_0;\, t_1]$. $\square$

Пусть f — дифференцируемая функция. Функцию

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

назовем *функцией Вейерштрасса* функции f.

Геометрический смысл функции Вейерштрасса. $\mathcal{E}(x,x')$ — разность между значением f в точке x' и значением в точке x' линейной функции, касательной к графику f в точке x.



Отсюда ясно, что если f выпукла, то $\mathcal{E}(x,x') \geq 0$. Можно показать, что верно и обратное. Если $\mathcal{E}(x,x') \geq 0$, то функция f выпукла.

Пусть $L(\dot{x}) = L(t,x,\dot{x})$ — интегрант простейшей задачи вариационного исчисления. Функция

$$\mathcal{E}(t,x,\dot{x},u) := L(t,x,u) - L(t,x,\dot{x}) - L_{\dot{x}}(t,x,\dot{x})(u-\dot{x})$$

называется *функцией Вейерштрасса интегранта* L. Таким образом, $\mathcal{E}(t,x,\dot{x},u)$ — функция Вейерштрасса функции $\dot{x} \to L(t,x,\dot{x})$, где t,x играют роль параметров.

Определение 5. Говорят, что на экстремали $\widehat{x}$ выполнено *условие Вейерштрасса*, если

$$\mathcal{E}(t,\widehat{x}(t),\dot{\widehat{x}}(t),u) = L(t,\widehat{x}(t),u) - L(t,\widehat{x}(t),\dot{\widehat{x}}(t)) - \widehat{L}_{\dot{x}}(t)(u-\dot{\widehat{x}}(t)) \geq 0$$
$$\forall\ u \in \mathbb{R},\ \forall\ t \in [t_0;\, t_1].$$

Геометрический смысл условия Вейерштрасса на экстремали $\widehat{x}$: для любого фиксированного $t \in [t_0;\, t_1]$ график функции $L = L(\dot{x}) = L(t, \widehat{x}(t), \dot{x})$ (как функции от $\dot{x}$) лежит выше касательной к кривой L в точке $\dot{\widehat{x}}(t)$.

Если функция $L(\dot{x})$ выпуклая по $\dot{x}$ для любых t, x, то условие Вейерштрасса выполняется на любой экстремали $\widehat{x}$.

Отметим, что для квадратичного по $\dot{x}$ интегранта $L(t, x, \dot{x}) = L(\dot{x}) = A\dot{x}^2 + C\dot{x} + B$ (A, B, C — функции от t, x) функция Вейерштрасса

$$\mathcal{E}(t, x, \dot{x}, u) = L(t, x, u) - L(t, x, \dot{x}) - L_{\dot{x}}(u - \dot{x}) =$$
$$= Au^2 + Cu + B - (A\dot{x}^2 + C\dot{x} + B) - (2A\dot{x} + C)(u - \dot{x}) =$$
$$= Au^2 + Cu + B - A\dot{x}^2 - C\dot{x} - B - 2Au\dot{x} - Cu + 2A\dot{x}^2 + C\dot{x} =$$
$$= Au^2 - 2Au\dot{x} + A\dot{x}^2 = A(u - \dot{x})^2 = \frac{1}{2} L_{\dot{x}\dot{x}}(u - \dot{x})^2.$$

Полученная формула для функции Вейерштрасса квадратичного по $\dot{x}$ интегранта

$$\mathcal{E}(t, x, \dot{x}, u) = \frac{1}{2} L_{\dot{x}\dot{x}}(u - \dot{x})^2 \tag{5.2}$$

выводится также из формулы Тейлора для квадратичной по $\dot{x}$ функции:

$$L(u) = L(\dot{x}) + L_{\dot{x}}(u - \dot{x}) + \frac{1}{2} L_{\dot{x}\dot{x}}(u - \dot{x})^2.$$

Условия Лежандра, Якоби, Вейерштрасса в векторном случае.

Необходимое условие 1-го порядка локального минимума для допустимой вектор-функции $\widehat{x} = (\widehat{x}_1, \ldots, \widehat{x}_n) \in C^1([t_0;\, t_1], \mathbb{R}^n)$

$$J'(\widehat{x})[h] = \int\limits_{t_0}^{t_1} \big(\langle \widehat{L}_{\dot{x}}, \dot{h}\rangle + \langle \widehat{L}_x, h\rangle\big) dt = 0 \ \forall\ h \in C_0^1([t_0;\, t_1], \mathbb{R}^n)$$

равносильно системе уравнений Эйлера.

Необходимое условие 2-го порядка локального минимума запишется в виде неравенства

$$\int\limits_{t_0}^{t_1} \Big(\big\langle \widehat{L}_{\dot{x}\dot{x}}\dot{h}, \dot{h}\big\rangle + 2\big\langle \widehat{L}_{x\dot{x}}\dot{h}, h\big\rangle + \big\langle \widehat{L}_{xx}h, h\big\rangle\Big) dt \geq 0 \ \forall\ h \in C_0^1([t_0;\, t_1], \mathbb{R}^n).$$

Для вектор-функций

$$L_{\dot{x}\dot{x}} = \begin{pmatrix} L_{\dot{x}_1\dot{x}_1} & \ldots & L_{\dot{x}_1\dot{x}_n} \\ \ldots & \ldots & \ldots \\ L_{\dot{x}_n\dot{x}_1} & \ldots & L_{\dot{x}_n\dot{x}_n} \end{pmatrix}, \quad L_{\dot{x}x} = \begin{pmatrix} L_{\dot{x}_1x_1} & \ldots & L_{\dot{x}_1x_n} \\ \ldots & \ldots & \ldots \\ L_{\dot{x}_nx_1} & \ldots & L_{\dot{x}_nx_n} \end{pmatrix},$$

$$L_{x\dot{x}} = \begin{pmatrix} L_{x_1\dot{x}_1} & \ldots & L_{x_1\dot{x}_n} \\ \ldots & \ldots & \ldots \\ L_{x_n\dot{x}_1} & \ldots & L_{x_n\dot{x}_n} \end{pmatrix}, \quad L_{xx} = \begin{pmatrix} L_{x_1x_1} & \ldots & L_{x_1x_n} \\ \ldots & \ldots & \ldots \\ L_{x_nx_1} & \ldots & L_{x_nx_n} \end{pmatrix}$$

— матрицы размера $n \times n$. Матрица $L_{\dot{x}x}$ является транспонированной к матрице $L_{x\dot{x}}$ и $\langle L_{x\dot{x}}\dot{h}, h\rangle = \langle \dot{h}, L^*_{x\dot{x}}h\rangle = \langle \dot{h}, L_{\dot{x}x}h\rangle$.

Условие Лежандра $\widehat{L}_{\dot{x}\dot{x}}(t) \geq 0$ означает неотрицательную определенность матрицы, *усиленное условие Лежандра* $\widehat{L}_{\dot{x}\dot{x}}(t) > 0$ — положительную определенность матрицы.

Крышка над функцией $L(t, x(t), \dot{x}(t))$ и её производными означает, что функция $L(t, x(t), \dot{x}(t))$ и её производные рассматриваются на экстремали, т. е. $\widehat{L}(t) = L(t, \widehat{x}(t), \dot{\widehat{x}}(t))$.

Пусть далее $\widehat{L}_{\dot{x}\dot{x}}, \widehat{L}_{\dot{x}x}, \widehat{L}_{xx} \in C^1([t_0;\, t_1], \mathbb{R}^{n^2})$ и выполнено усиленное условие Лежандра.

Система уравнений Эйлера по $h = (h_1, \ldots, h_n)$ для интегранта $\widetilde{L} = \widetilde{L}(t, h, \dot{h}) := \langle \widehat{L}_{\dot{x}\dot{x}}\dot{h}, \dot{h} \rangle + 2\langle \widehat{L}_{x\dot{x}}\dot{h}, h \rangle + \langle \widehat{L}_{xx}h, h \rangle$

$$-\frac{d}{dt}\widetilde{L}_{\dot{h}}(t) + \widetilde{L}_h(t) = 0 \iff$$

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

называется *системой уравнений Якоби* для исходной задачи на экстремали $\widehat{x}$.

Точка τ называется *сопряженной к точке* t_0, если существует нетривиальное решение системы уравнений Якоби вектор $h = (h_1, \ldots, h_n)$, для которого $h(t_0) = h(\tau) = 0$. Говорят, что на $\widehat{x}$ выполнено *условие Якоби*, если в интервале $(t_0;\, t_1)$ нет точек, сопряженных с t_0, и *усиленное условие Якоби*, если в полуинтервале $(t_0;\, t_1]$ нет точек, сопряженных с t_0.

Если выполнено усиленное условие Лежандра, то матрица $\widehat{L}_{\dot{x}\dot{x}}$ положительно определена и, значит, невырождена и обратима. Поэтому, продифференцировав по t в уравнении Якоби (J), получим

$$-\dot{\widehat{L}}_{\dot{x}\dot{x}}\dot{h} - \widehat{L}_{\dot{x}\dot{x}}\ddot{h} - \dot{\widehat{L}}_{\dot{x}x}h - \underline{\widehat{L}_{\dot{x}x}\dot{h}} + \underline{\widehat{L}_{x\dot{x}}\dot{h}} + \widehat{L}_{xx}h = 0 \iff$$

$$\widehat{L}_{\dot{x}\dot{x}}\ddot{h} + \dot{\widehat{L}}_{\dot{x}\dot{x}}\dot{h} + (\dot{\widehat{L}}_{\dot{x}x} - \widehat{L}_{xx})h = 0.$$

Эту систему из n дифференциальных уравнений 2-го порядка с непрерывными коэффициентами в силу обратимости матрицы $\widehat{L}_{\dot{x}\dot{x}}$ можно разрешить относительно старших производных.

Дадим аналитический способ нахождения сопряженных точек. Система уравнений Якоби — система n линейных дифференциальных уравнений 2-го порядка, которую (из-за усиленного условия Лежандра) можно разрешить относительно 2-й производной. Пусть $h^1(\cdot), \ldots, h^n(\cdot)$ — фундаментальная система решений системы уравнения Якоби. Выберем её с начальными условиями $H(t_0) = 0$ (нулевая матрица), $\dot{H}(t_0) = I$

(единичная матрица) или $\det \dot{H}(t_0) \neq 0$, где

$$H(t) = \big(h^1(t) \ldots h^n(t)\big) = \begin{pmatrix} h_1^1(t) & \ldots & h_1^n(t) \\ \ldots & \ldots & \ldots \\ h_n^1(t) & \ldots & h_n^n(t) \end{pmatrix}.$$

Столбцы матрицы H являются векторами решений системы уравнений Якоби. Покажем, что точка τ является сопряженной к t_0 тогда и только тогда, когда матрица $H(\tau)$ является вырожденной.

Действительно, условие $\det H(\tau) = 0$ равносильно тому, что столбцы матрицы $H(\tau)$ являются линейно зависимыми. Это означает, что существуют числа $\alpha_1, \ldots, \alpha_n$, не все равные нулю, такие что для функции $h = \sum\limits_{i=1}^{n} \alpha_i h^i$ выполняются граничные условия $h(\tau) = 0$. Граничное условие $H(t_0) = 0$ равносильно тому, что $h_j^i(t_0) = 0$, $i, j = 1, \ldots, n$. Следовательно, $h(\tau) = 0$, т. е. τ — сопряжённая точка.

Утверждение 2. *Если выполнены усиленные условия Лежандра и Якоби, то для любых $h_0, h_1 \in \mathbb{R}^n$ существует и единственно решение уравнения Якоби $h(\cdot)$, для которого $h(t_0) = h_0$, $h(t_1) = h_1$.*

Доказательство. По теореме существования и единственности решения системы дифференциальных уравнений 2-го порядка существуют фундаментальные матрицы решений системы уравнений Якоби $H(t, t_0)$ и $H(t, t_1)$ с граничными условиями $H(t_0, t_0) = 0$, $\dot{H}(t_0, t_0) = I$, и $H(t_1, t_1) = 0$, $\dot{H}(t_1, t_1) = I$.

Из усиленного условия Якоби (на полуинтервале $(t_0;\, t_1]$ нет сопряженных точек) вытекает, что матрицы $H(t, t_0)$ и $H(t, t_1)$ невырождены для $t \in (t_0;\, t_1]$ и $[t_0,\, t_1)$ соответственно. Положим $H_0(t) = H(t, t_1) H^{-1}(t_0, t_1)$, $H_1(t) = H(t, t_0) H^{-1}(t_1, t_0)$. Тогда $H_0(t_0) = I$, $H_0(t_1) = 0$, $H_1(t_0) = 0$, $H_1(t_1) = I$, и, значит, функция $h(t) = H_0(t) h_0 + H_1(t) h_1$ является решением уравнения Якоби с граничными условиями

$$h(t_0) = H_0(t_0)h_0 + H_1(t_0)h_1 = h_0,$$
$$h(t_1) = H_0(t_1)h_0 + H_1(t_1)h_1 = h_1.$$

Если существует другое решение системы уравнений Якоби $\overline{h} \neq h$, то вектор-функция $\widetilde{h} = \overline{h} - h \not\equiv 0$ тоже является решением уравнения Якоби с граничными условиями $\widetilde{h}(t_0) = \widetilde{h}(t_1) = 0$. Тогда для матрицы фундаментальных решений системы уравнений Якоби $\det \widetilde{H}(t_1) = 0$, что противоречит выполнению усиленного условия Якоби. $\square$

Условие Вейерштрасса для вектор-функций запишется в виде

$$\mathcal{E}(t, \widehat{x}(t), \dot{\widehat{x}}(t), u) =$$
$$= L(t, \widehat{x}(t), u) - L(t, \widehat{x}(t), \dot{\widehat{x}}(t)) - \left\langle L_{\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t)), u - \dot{\widehat{x}}(t) \right\rangle \geq 0$$
$$\forall\ u \in \mathbb{R}^n,\ \forall\ t \in [t_0;\, t_1].$$

1.5 Необходимые и достаточные условия слабого и сильного экстремума

Рассмотрим простейшую задачу вариационного исчисления для вектор-функций $x = (x_1, \ldots, x_n)$ (для определенности задачу на минимум)

$$\int\limits_{t_0}^{t_1} L\left(t, x(t), \dot{x}(t)\right)\, dt \to \min;\ x(t_0) = x_0,\ x(t_1) = x_1. \qquad (P)$$

1.5.1 Игольчатые вариации. Условие Вейерштрасса.

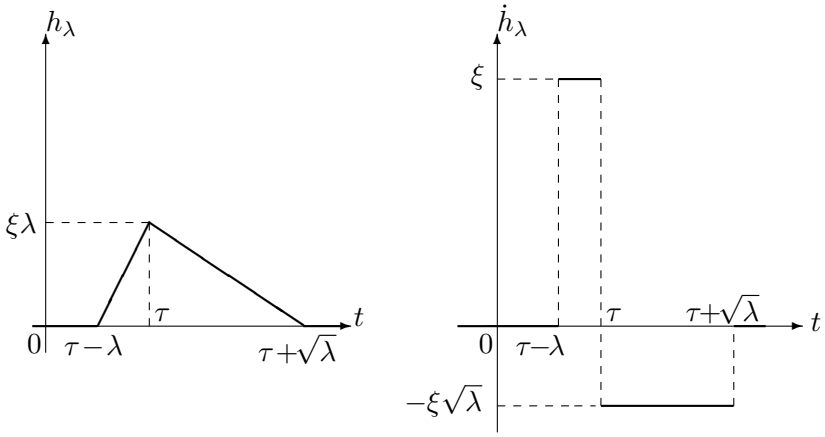
Понятие сильного экстремума ввел в вариационное исчисление Вейерштрасс. Для доказательства необходимого условия сильного минимума Вейерштрасс применил специальные игольчатые вариации функции $\widehat{x}$.

В дальнейшем игольчатые вариации и пакет иголок использовались для доказательства принципа максимума.

Пусть функция $\widehat{x}(t) \in PC^1[t_0;\, t_1]$. Зафиксируем точку $\tau \in (t_0;\, t_1) \cap T$ ($T \subset [t_0;\, t_1]$ — множество точек непрерывности

функции $\dot{\widehat{x}}$) и число $\xi \in \mathbb{R}$. Положим $x_\lambda(t) = \widehat{x}(t) + h_\lambda(t)$, где $\lambda \geq 0$ достаточно мало,

$$h_\lambda(t) = h_\lambda(t;\tau,\xi) = \begin{cases} \xi\lambda + (t-\tau)\xi, & t \in [\tau-\lambda;\, \tau], \\ \xi\lambda - (t-\tau)\xi\sqrt{\lambda}, & t \in [\tau;\, \tau+\sqrt{\lambda}], \\ 0, & t \notin [\tau-\lambda;\, \tau+\sqrt{\lambda}]. \end{cases}$$



Производная вариации $h_\lambda(\cdot)$ имеет вид, изображенный на рисунке (для удобства изображения взято $n = 1$, $\xi > 0$). При малом λ изменение функции $\widehat{x}$ производится на очень маленьком отрезке, в связи с чем подобные вариации называют “игольчатыми”. Очевидно, что $x_\lambda(\cdot) \to \widehat{x}(\cdot)$ в метрике пространства $C([t_0;\, t_1])$ при $\lambda \to 0$.

С помощью игольчатых вариаций докажем условие Вейерштрасса — необходимое условие сильного минимума в простейшей задаче вариационного исчисления на классе кусочно-гладких функций.

Теорема 1.1. *Пусть функция* $\widehat{x} \in \operatorname{strlocmin} P$, *интегрант* L *непрерывно дифференцируем в некоторой окрестности расширенного графика* $\Gamma_{\widehat{x}\dot{\widehat{x}}}$, $T \subset [t_0;\, t_1]$ — *множество точек непрерывности функции* $\dot{\widehat{x}}$. *Тогда на* $\widehat{x}$ *выполняется условие Вейерштрасса*

$$\mathcal{E}(t,\widehat{x}(t),\dot{\widehat{x}}(t),u)=L(t,\widehat{x}(t),u)-L(t,\widehat{x}(t),\dot{\widehat{x}}(t))-\widehat{L}_{\dot{x}}(t)(u-\dot{\widehat{x}}(t))\geq 0$$
$$\forall\ u\in\mathbb{R}^n,\ \forall\ t\in T.$$

Доказательство. Возьмём точку $\tau\in(t_0;\,t_1)\cap T$ (случаи $\tau=t_0,t_1$ доказываются предельными переходами $\tau\to t_0$, $\tau\to t_1$ в неравенстве). Рассмотрим выписанную выше игольчатую вариацию $x_\lambda=\widehat{x}+h_\lambda$ функции $\widehat{x}$.

При достаточно малых $\lambda\geq 0$ отрезок $[\tau-\lambda,\,\tau+\sqrt{\lambda}\,]\subset T$, функция x_λ допустима в задаче на сильный экстремум: $x_\lambda\in PC^1([t_0;\,t_1])$, $x_\lambda(t_i)=\widehat{x}(t_i)+h_\lambda(t_i)=x_i$, $i=0,1$. Очевидно, что функция $x_\lambda\to\widehat{x}$ в метрике пространства $C([t_0;\,t_1])$ при $\lambda\to 0$.

Поскольку $\widehat{x}\in\operatorname{strlocmin}P$, то при достаточно малых $\lambda\geq 0$

$$J(x_\lambda)-J(\widehat{x})\geq 0\iff\int\limits_{t_0}^{t_1}\big(L(t,x_\lambda,\dot{x}_\lambda)-L(t,\widehat{x},\dot{\widehat{x}})\big)dt\geq 0$$

Разобьем отрезок интегрирования $[t_0;\,t_1]$ на четыре отрезка $[t_0;\,\tau-\lambda]$, $[\tau-\lambda;\,\tau]$, $[\tau;\,\tau+\sqrt{\lambda}]$, $[\tau+\sqrt{\lambda};\,t_1]$. Поскольку функции $\widehat{x}(t)$ и $x_\lambda(t)$ совпадают при $t\in[t_0;\,\tau-\lambda]$ и $t\in[\tau+\sqrt{\lambda};\,t_1]$, то $L(t,x_\lambda,\dot{x}_\lambda)=L(t,\widehat{x},\dot{\widehat{x}})$ и, соответственно, интегралы на этих промежутках интегрирования обращаются в ноль. Значит,

$$\int\limits_{\tau-\lambda}^{\tau}\big(L(t,x_\lambda,\dot{x}_\lambda)-\widehat{L}\big)dt+\int\limits_{\tau}^{\tau+\sqrt{\lambda}}\big(L(t,x_\lambda,\dot{x}_\lambda)-\widehat{L}\big)dt\geq 0\iff$$

(по теореме о среднем для определенных интегралов[6])

$$\lambda\big(L(t,x_\lambda,\dot{x}_\lambda)-\widehat{L}\big)\big|_{t\in[\tau-\lambda;\,\tau]}+$$
$$+\sqrt{\lambda}\big(L(t,x_\lambda,\dot{x}_\lambda)-\widehat{L}\big)\big|_{t\in[\tau;\,\tau+\sqrt{\lambda}]}\geq 0\iff$$

[6] **Теорема** (о среднем для определённых интегралов).
Пусть $f(x)$ — непрерывная функция на отрезке $[a;\,b]$.
Тогда $\int\limits_a^b f(x)dx=(b-a)f(\xi)$, где ξ — некоторая точка из отрезка $[a;\,b]$.

(разделим обе части неравенства на положительный множитель λ и подставим x_λ и $\dot{x}_\lambda$ на соответствующих промежутках)

$$\big(L(t,x_\lambda,\dot{x}_\lambda)-\widehat{L}\big)\big|_{t\in[\tau-\lambda;\,\tau]}+\frac{\big(L(t,x_\lambda,\dot{x}_\lambda)-\widehat{L}\big)\big|_{t\in[\tau;\,\tau+\sqrt{\lambda}]}}{\sqrt{\lambda}}\geq 0. \quad (*)$$

При $t\in[\tau-\lambda;\,\tau]$ функции $x_\lambda=\widehat{x}+\xi\lambda+(t-\tau)\xi$, $\dot{x}_\lambda=\dot{\widehat{x}}+\xi$. Поэтому при $\lambda\to 0$ точка $\tau-\lambda\to\tau$ и, следовательно, первая разность $L(t,x_\lambda,\dot{x}_\lambda)-\widehat{L}=L(t,\widehat{x}+\xi\lambda+(t-\tau)\xi,\dot{\widehat{x}}+\xi)-\widehat{L}\to L(\tau,\widehat{x}(\tau),\dot{\widehat{x}}(\tau)+\xi)-\widehat{L}(\tau)$.

При $t\in[\tau;\,\tau+\sqrt{\lambda}]$ функции $h_\lambda=\xi\lambda-(t-\tau)\xi\sqrt{\lambda}$, $\dot{h}_\lambda=-\xi\sqrt{\lambda}$. В силу дифференцируемости

$$L(t,x_\lambda,\dot{x}_\lambda)-\widehat{L}=L(t,\widehat{x}+h_\lambda,\dot{\widehat{x}}+\dot{h}_\lambda)-\widehat{L}=\widehat{L}_x h_\lambda+\widehat{L}_{\dot{x}}\dot{h}_\lambda+o(\sqrt{\lambda})=$$
$$=\widehat{L}_x\cdot(\xi\lambda-(t-\tau)\xi\sqrt{\lambda})+\widehat{L}_{\dot{x}}\cdot(-\xi\sqrt{\lambda})+o(\sqrt{\lambda}).$$

Поэтому вторая дробь в неравенстве $(*)$ при $\lambda\to 0$

$$\frac{\big(\widehat{L}_x\xi\lambda-\widehat{L}_x(t-\tau)\xi\sqrt{\lambda}-\widehat{L}_{\dot{x}}\xi\sqrt{\lambda}+o(\sqrt{\lambda})\big)\big|_{t\in[\tau;\,\tau+\sqrt{\lambda}]}}{\sqrt{\lambda}}\to-\widehat{L}_{\dot{x}}(\tau)\xi.$$

Тогда неравенство $(*)$ перепишется в виде

$$L(\tau,\widehat{x}(\tau),\dot{\widehat{x}}(\tau)+\xi)-\widehat{L}(\tau)-\widehat{L}_{\dot{x}}(\tau)\xi\geq 0.$$

Нетрудно видеть, что выписанное неравенство является тем же самым, что и в формулировке теоремы, где $t=\tau$, $u=\dot{\widehat{x}}(\tau)+\xi$. Условие Вейерштрасса доказано. $\square$

В задаче на максимум условие Вейерштрасса меняет свой знак.

1.5.2 Необходимые условия сильного экстремума

Теорема 1.2. *Пусть функция* $\widehat{x} \in PC^1([t_0;\, t_1], \mathbb{R}^n)$ *доставляет сильный локальный минимум в задаче* (P), *интегрант* L *непрерывно дифференцируем в некоторой окрестности расширенного графика* $\Gamma_{\widehat{x}\dot{\widehat{x}}}$. *Тогда на* $\widehat{x}$ *выполняется уравнение Эйлера*

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \quad \forall\, t \in T^7$$

и удовлетворяется условие Вейерштрасса

$$\mathcal{E}(t, \widehat{x}(t), \dot{\widehat{x}}(t), u) = L(t, \widehat{x}(t), u) - L(t, \widehat{x}(t), \dot{\widehat{x}}(t)) - \widehat{L}_{\dot{x}}(t)(u - \dot{\widehat{x}}(t)) \geq 0$$
$$\forall\, u \in \mathbb{R}^n,\ \forall\, t \in T.$$

Если при этом существует $\widehat{L}_{\dot{x}\dot{x}}(t)\ \forall\, t \in T$, *то выполняется также условие Лежандра:* $\widehat{L}_{\dot{x}\dot{x}}(t) \geq 0\ \forall\, t \in T$.

В задаче на максимум условия Лежандра и Вейерштрасса меняют свой знак.

Доказательство. Формализуем задачу (P) как задачу оптимального управления

$$\int\limits_{t_0}^{t_1} L(t, x, u)\, dt \to \min;\ \dot{x} = u,\ x(t_0) = x_0,\ x(t_1) = x_1. \qquad (P')$$

Условие $\widehat{x} \in \operatorname{strlocmin} P$ равносильно тому, что пара $(\widehat{x}, \widehat{u})$, где $\widehat{u} = \dot{\widehat{x}}$, является оптимальным процессом в задаче оптимального управления (P'). Поэтому согласно принципу максимума Понтрягина найдутся множители Лагранжа $\lambda = (\lambda_0, \lambda_1, \lambda_2)$ и $p(\cdot) \in PC^1([t_0;\, t_1])$, $\lambda \neq 0$, такие, что для функции Лагранжа задачи (P')

$$\mathscr{L} = \int\limits_{t_0}^{t_1} \underbrace{\big(\lambda_0 L(t, x, u) + p(t)(\dot{x} - u)\big)}_{\widetilde{L}} dt + \underbrace{\lambda_1(x(t_0) - x_0) + \lambda_2(x(t_1) - x_1)}_{l}$$

---

[7] $T \subset [t_0;\, t_1]$ — множество точек непрерывности функции $\dot{\widehat{x}}(t)$.

выполняются условия:

a) уравнение Эйлера для интегранта $\widetilde{L}$:

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

$$-\dot{p}(t) + \lambda_0 \widehat{L}_x(t) = 0 \quad \forall\, t \in T;$$

b) трансверсальности по x:

$$\begin{cases} \widetilde{L}_{\dot{x}}(t_0) = l_{x(t_0)}, \\ \widetilde{L}_{\dot{x}}(t_1) = -l_{x(t_1)}, \end{cases} \iff \begin{cases} p(t_0) = \lambda_1, \\ p(t_1) = -\lambda_2; \end{cases}$$

c) оптимальности по u:

$$\min_{u\in\mathbb{R}^n}\{\lambda_0 L(t,\widehat{x}(t),u) - p(t)u\} = \lambda_0 L(t,\widehat{x}(t),\dot{\widehat{x}}(t)) - p(t)\dot{\widehat{x}}(t)\ \forall\, t \in T.$$

d) неотрицательности: $\lambda_0 \geq 0$.

Если $\lambda_0 = 0$, то условие оптимальности перепишется в виде $\min\limits_{u\in\mathbb{R}^n}\{-p(t)u\} = -p(t)\dot{\widehat{x}}(t)$. Поскольку минимум конечен и равен $-p(t)\dot{\widehat{x}}(t)$, то это может быть только при $p(t) \equiv 0$. Тогда из b) следует что все множители Лагранжа нули.

Значит, $\lambda_0 \neq 0$. Полагаем в задаче на минимум $\lambda_0 = 1$. Тогда из c)

$$\min_{u\in\mathbb{R}^n}\{\underbrace{L(t,\widehat{x}(t),u) - p(t)u}_{f(u)}\} = L(t,\widehat{x}(t),\dot{\widehat{x}}(t)) - p(t)\dot{\widehat{x}}(t).$$

По необходимому условию 1-го порядка минимума функции f следует, что

$$f'(\widehat{u}) = 0 \iff \widehat{L}_u(t) - p(t) = 0 \iff p(t) = \widehat{L}_{\dot{x}}(t). \qquad (5.3)$$

Подставляя $p = \widehat{L}_{\dot{x}}$ в условие a), получаем уравнение Эйлера для интегранта L.

А из необходимого условия 2-го порядка минимума функции f следует выполнение условия Лежандра:

$$f''(\widehat{u}) \geq 0 \iff \widehat{L}_{\dot{x}\dot{x}}(t) \geq 0\ \forall\, t \in T.$$

Условие оптимальности по u при $\lambda_0 = 1$ и $p = \widehat{L}_{\dot{x}}$

$$L(t, \widehat{x}(t), u) - \widehat{L}_{\dot{x}}(t)u \geq L(t, \widehat{x}(t), \dot{\widehat{x}}(t)) - \widehat{L}_{\dot{x}}(t)\dot{\widehat{x}}(t) \iff$$
$$L(t, \widehat{x}(t), u) - L(t, \widehat{x}(t), \dot{\widehat{x}}(t)) - \widehat{L}_{\dot{x}}(t)(u - \dot{\widehat{x}}(t)) \geq 0$$

есть не что иное, как условие Вейерштрасса. $\square$

Выполнение уравнения Эйлера для сильного экстремума при $\widehat{x} \in C^1$ также следует из соотношения

$$\widehat{x} \in \operatorname{strlocextr} P \Longrightarrow \widehat{x} \in \operatorname{wlocextr} P \Longrightarrow \text{уравнение Эйлера}$$

и уже доказанного выполнения необходимого условия слабого экстремума — уравнения Эйлера.

1.5.3 Необходимые условия слабого экстремума

Теорема 1.3. *Пусть функция* $\widehat{x} \in C^2([t_0;\, t_1], \mathbb{R}^n)$ *доставляет слабый локальный минимум в задаче* (P), *интегрант* L *трижды непрерывно дифференцируем в некоторой окрестности расширенного графика* $\Gamma_{\widehat{x}\dot{\widehat{x}}}$. *Тогда на* $\widehat{x}$ *выполняется уравнение Эйлера, условие Лежандра, а если на экстремали* $\widehat{x}$ *выполнено усиленное условие Лежандра, то выполняется и условие Якоби.*

В задаче на максимум условие Лежандра меняет свой знак.

Доказательство. Для простоты записи проведем доказательство для $n = 1$.

1) *Вывод уравнения Эйлера* слабого локального экстремума был проведен в Главе 3 § 1.

2) *Вывод условия Лежандра.* В Главе 5 п. 1.3 было показано, что необходимым условием 2-го порядка слабого локального минимума является выполнения неравенства (5.1):

$$K(h(\cdot)) := \int\limits_{t_0}^{t_1} \Big(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\widehat{L}_{x\dot{x}}(t)\dot{h}(t)h(t) + \widehat{L}_{xx}(t)h^2(t) \Big)\, dt \geq 0 \qquad \forall\ h(\cdot) \in C_0^1([t_0;\, t_1]).$$

Из неотрицательности и вида функционала K следует, что функция $\overline{h}(t) \equiv 0$ доставляет абсолютный минимум (слабый) в задаче

$$K(h(\cdot)) = \int\limits_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{x\dot{x}}\dot{h}h + \widehat{L}_{xx}h^2 \big) dt \to \min; \qquad h(t_0) = h(t_1) = 0. \quad (P'')$$

По следствию из леммы о скруглении углов функция $\overline{h} \equiv 0$ доставляет в задаче (P'') и сильный абсолютный минимум. Тогда в силу Теоремы 1.2 о необходимых условиях сильного минимума в задаче (P'') на $\overline{h}$ выполняется условие Лежандра

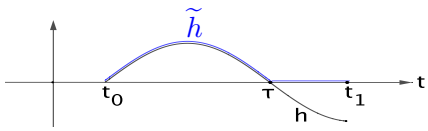
для интегранта

$$\widetilde{L}(t, h(t), \dot{h}(t)) := \widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\widehat{L}_{x\dot{x}}(t)\dot{h}(t)h(t) + \widehat{L}_{xx}(t)h^2(t),$$

т. е. $\widetilde{L}_{\dot{h}\dot{h}}(t) \geq 0 \Leftrightarrow \widehat{L}_{\dot{x}\dot{x}}(t) \geq 0 \ \forall\, t \in [t_0;\, t_1]$.

Таким образом, условие Лежандра в задаче (P) выполнено.

3) *Вывод условия Якоби.* Предположим противное, что условие Якоби не выполнено, т. е. существует точка $\tau \in (t_0;\, t_1)$ и нетривиальное $(h \not\equiv 0)$ решение $h \in C^1([t_0;\, t_1])$ уравнения Якоби, для которого $h(t_0) = h(\tau) = 0$. Отметим, что из нетривиальности решения h однородного линейного дифференциального уравнения второго порядка с условием $h(\tau) = 0$ вытекает, что $\dot{h}(\tau) \neq 0$. Положим $\widetilde{h}(t) = \begin{cases} h(t), & t_0 \leq t \leq \tau, \\ 0, & \tau \leq t \leq t_1. \end{cases}$



Тогда

$$K(\widetilde{h}(\cdot)) = \int\limits_{t_0}^{\tau} \left(\widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{x\dot{x}}\dot{h}h + \widehat{L}_{xx}h^2\right) dt =$$

$$= \int\limits_{t_0}^{\tau} \left(\widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + \widehat{L}_{\dot{x}x}h\dot{h} + \widehat{L}_{x\dot{x}}\dot{h}h + \widehat{L}_{xx}h^2\right) dt =$$

$$= \int\limits_{t_0}^{\tau} \left(\widehat{L}_{\dot{x}\dot{x}}\dot{h} + \widehat{L}_{\dot{x}x}h\right)\dot{h}\, dt + \int\limits_{t_0}^{\tau} \left(\widehat{L}_{x\dot{x}}\dot{h} + \widehat{L}_{xx}h\right)h\, dt =$$

$$= \left(\widehat{L}_{\dot{x}\dot{x}}\dot{h} + \widehat{L}_{\dot{x}x}h\right)h\Big|_{t_0}^{\tau} +$$

$$+ \int\limits_{t_0}^{\tau} \left(-\frac{d}{dt}\left(\widehat{L}_{\dot{x}\dot{x}}\dot{h} + \widehat{L}_{\dot{x}x}h\right) + \widehat{L}_{x\dot{x}}\dot{h} + \widehat{L}_{xx}h\right)h\, dt = 0,$$

(так как функция h удовлетворяет уравнению Якоби). Таким образом, $K(\widetilde{h}) = 0$, а это означает, что $\widetilde{h} \in \operatorname{strabsmin} P''$ (наряду с функцией $\overline{h} \equiv 0$). Следовательно, пара $(\widetilde{h}, \widetilde{u} = \dot{\widetilde{h}})$ является оптимальным управляемым процессом в задаче оптимального управления

$$\int\limits_{t_0}^{t_1} \widetilde{L}(t, h, u)\, dt \to \min;\ \dot{h} = u,\ h(t_0) = h(t_1) = 0.$$

Проводя рассуждения как и в Теореме 1.2, получим, что найдется функция $\widetilde{p} \in PC^1([t_0;\, t_1])$ такая, что для лагранжиана $\widetilde{L}(t, h, \dot{h}) = \widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{\dot{x}x}h\dot{h} + \widehat{L}_{xx}h^2$ на экстремали $\widetilde{h}$ выполняется равенство аналогичное (5.3) ($p = \widehat{L}_{\dot{x}}$)

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

Поскольку $\widetilde{h}(t) = \dot{\widetilde{h}}(t) = 0$ при $t > \tau$, то $\widetilde{p}(\tau + 0) = 0$ и в силу непрерывности функции $\widetilde{p}$

$$0 = \widetilde{p}(\tau - 0) = 2\widehat{L}_{\dot{x}\dot{x}}(\tau)\dot{\widetilde{h}}(\tau - 0) = 2\widehat{L}_{\dot{x}\dot{x}}(\tau)\dot{h}(\tau) = 0,$$

откуда $\dot{h}(\tau) = 0$ (ибо $\widehat{L}_{\dot{x}\dot{x}}(\tau) \neq 0$ из-за усиленного условия Лежандра). Мы пришли к противоречию с условием $\dot{h}(\tau) \neq 0$. Таким образом, предположение противного неверно, и условие Якоби выполнено. $\square$

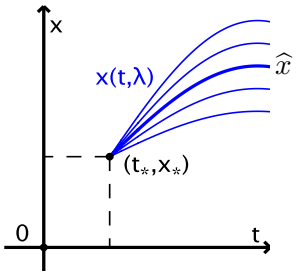
1.5.4 Поле экстремалей

Рассмотрим простейшую задачу вариационного исчисления

$$\int\limits_{t_0}^{t_1} L\left(t, x(t), \dot{x}(t)\right)\, dt \to \min;\ x(t_0) = x_0,\ x(t_1) = x_1. \qquad (P)$$

Полем экстремалей[8] в задаче (P) называется множество функций $\{x(t,\lambda)\}$ с параметром λ, удовлетворяющих уравнению Эйлера. Кроме того необходимо, чтобы функции $x(t,\lambda) \in C^1([t_0;\, t_1] \times \Lambda, \mathbb{R}^n)$, Λ — некоторое открытое множество в пространстве $\mathbb{R}^n$.

Если для поля экстремалей существует точка (t_*, x_*) такая, что $x(t_*, \lambda) = x_*$ для всех $\lambda \in \Lambda$, то это поле называется *центральным полем экстремалей*, а точка (t_*, x_*) называется *центром поля экстремалей*.



Пусть $\widehat{x}$ — допустимая экстремаль[9] в задаче (P).

Экстремаль $\widehat{x}$ включена в поле экстремалей $\{x(\cdot,\lambda)\}$, если $\widehat{x}(\cdot) = x(\cdot, \widehat{\lambda})$ при некотором $\widehat{\lambda} \in \Lambda$.

Экстремаль $\widehat{x}$ окружена полем экстремалей, если она включена в поле экстремалей и существует окрестность G графика $\Gamma_{\widehat{x}}$ такая, что для любой точки (τ, ξ) из этой окрестности имеется единственная экстремаль семейства, проходящая

[8] Экстремаль — это функция, удовлетворяющая уравнению Эйлера.

[9] Допустимая экстремаль — это функция с заданными в задаче граничными условиями, удовлетворяющая уравнению Эйлера.

через эту точку. Единственность экстремали, проходящей через точку (τ, ξ), означает, что по точке $(\tau, \xi) \in G$ значение $\lambda = \lambda(\tau, \xi)$ отыскивается единственным образом. Кроме того необходимо, чтобы функция $\lambda = \lambda(\tau, \xi) \in C^1(G)$.

Функция $u\colon G \to \mathbb{R}^n$, $u(\tau, \xi) = \dfrac{d}{dt} x(t, \lambda(\tau, \xi))\Big|_{t=\tau} =: \dot{x}(\tau, \lambda(\tau, \xi))$ называется *функцией наклона поля*. Отметим, что на экстремали $\widehat{x}$ функция наклона поля $u(t, \widehat{x}(t)) \equiv \dot{\widehat{x}}(t)$ совпадает с производной функции $\widehat{x}(t)$.

Построение центрального поля экстремалей

Теорема 1.4. *Пусть $\widehat{x} \in C^2([t_0;\, t_1], \mathbb{R}^n)$ — допустимая экстремаль в задаче (P), интегрант L трижды непрерывно дифференцируем в некоторой окрестности расширенного графика $\Gamma_{\widehat{x}\dot{\widehat{x}}}$, на $\widehat{x}$ выполнены усиленные условия Лежандра и Якоби. Тогда $\widehat{x}$ можно окружить центральным полем экстремалей.*

S-функция и её дифференциал

Пусть $x(\cdot, \lambda)$ — дважды непрерывно дифференцируемое центральное поле экстремалей с центром в точке (t_*, x_*), окружающее допустимую экстремаль $\widehat{x}(\cdot)$, и интегрант L — дважды непрерывно дифференцируем в некоторой окрестности расширенного графика функции $\widehat{x}(\cdot)$ $(L \in C^2(\mathcal{O}(\Gamma_{\widehat{x}\dot{\widehat{x}}})))$. Функция

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

называется *S-функцией центрального поля* $x(\cdot, \lambda)$.

Найдём дифференциал S-функции. Он понадобится нам для вывода основной формулы Вейерштрасса при доказательстве достаточных условий сильного экстремума. Для нахождения дифференциала найдём частные производные S-функции.

При этом нам понадобятся некоторые полезные соотношения, вытекающие из определения функции $\lambda(\tau, \xi)$ и функции наклона поля. Имеем тождество

$$x(\tau, \lambda(\tau, \xi)) = \xi.$$

Дифференцируя обе части этого тождества по τ, получим

$$\dot{x}(\tau, \lambda(\tau, \xi)) + x_\lambda(\tau, \lambda(\tau, \xi))\lambda_\tau(\tau, \xi) = 0 \Longrightarrow$$
$$x_\lambda \lambda_\tau = -\dot{x}(\tau, \lambda(\tau, \xi)) \stackrel{\text{def}}{=} -u(\tau, \xi) \qquad (\tau)$$

($u(\tau,\xi)$ — функция наклона поля). Дифференцируя обе части тождества $x(\tau,\lambda(\tau,\xi))=\xi$ по ξ, получим

$$x_\lambda(\tau,\lambda(\tau,\xi))\lambda_\xi(\tau,\xi)=I \qquad (\xi)$$

(I — единичная матрица). Поскольку (t_*,x_*) — центр поля, то $x(t_*,\lambda)=x_*$ для любого λ, и, значит, выполняется следующее соотношение

$$x_\lambda(t_*,\lambda(\tau,\xi))=0 \quad \forall\ (\tau,\xi)\in\mathcal{O}(\Gamma_{\widehat{x}}). \qquad (*)$$

Найдём $\frac{\partial S}{\partial\tau}$, дифференцируя по τ интеграл с переменным верхним пределом, и используя непрерывность $\dot{x}_\lambda$, вытекающую из того, что $x(\cdot,\lambda)\in C^2$. Имеем для $\lambda=\lambda(\tau,\xi)$, $u=u(\tau,\xi)$ и $S(\tau,\xi)=\int\limits_{t_*}^{\tau}L(t,x(t,\lambda(\tau,\xi)),\dot{x}(t,\lambda(\tau,\xi)))\,dt$

$$\frac{\partial S}{\partial\tau}=L(\tau,x(\tau,\lambda),\dot{x}(\tau,\lambda))+\int\limits_{t_*}^{\tau}\Big(L_x(t,x(t,\lambda),\dot{x}(t,\lambda))x_\lambda(t,\lambda)\lambda_\tau+$$

$$+L_{\dot{x}}(t,x(t,\lambda),\dot{x}(t,\lambda))\dot{x}_\lambda(t,\lambda)\lambda_\tau\Big)dt=$$

$$=L(\tau,\xi,u)+\int\limits_{t_*}^{\tau}(L_xx_\lambda\lambda_\tau+L_{\dot{x}}\dot{x}_\lambda\lambda_\tau)\,dt=$$

$$=L(\tau,\xi,u)+\int\limits_{t_*}^{\tau}L_xx_\lambda\lambda_\tau\,dt+\int\limits_{t_*}^{\tau}L_{\dot{x}}\,d(x_\lambda\lambda_\tau)=$$

$$=L(\tau,\xi,u)+L_{\dot{x}}x_\lambda\lambda_\tau\Big|_{t_*}^{\tau}+\int\limits_{t_*}^{\tau}\Big(-\frac{d}{dt}L_{\dot{x}}+L_x\Big)x_\lambda\lambda_\tau\,dt=$$

$$\overset{(\tau),(*)}{=}L(\tau,\xi,u(\tau,\xi))-L_{\dot{x}}(\tau,\xi,u(\tau,\xi))\,u(\tau,\xi).$$

При выводе мы воспользовались тем, что функции $x(\cdot,\lambda)$ — экстремали, т. е. удовлетворяют уравнению Эйлера.

Формула для $\frac{\partial S}{\partial \xi}$ выводится аналогично. Дифференцируя по ξ, получим

$$\frac{\partial S}{\partial \xi} = \int\limits_{t_*}^{\tau} (L_x x_\lambda \lambda_\xi + L_{\dot x} \dot x_\lambda \lambda_\xi)\, dt = \int\limits_{t_*}^{\tau} L_x x_\lambda \lambda_\xi\, dt + \int\limits_{t_*}^{\tau} L_{\dot x}\, d(x_\lambda \lambda_\xi) =$$

$$= L_{\dot x} x_\lambda \lambda_\xi \Big|_{t_*}^{\tau} + \int\limits_{t_*}^{\tau} \Big(-\frac{d}{dt} L_{\dot x} + L_x \Big) x_\lambda \lambda_\xi dt \overset{(\xi),(*)}{=} L_{\dot x}(\tau,\xi,u(\tau,\xi)).$$

Таким образом, имеет место следующая формула для дифференциала функции S:

$$dS(\tau,\xi) = \frac{\partial S}{\partial \tau} d\tau + \frac{\partial S}{\partial \xi} d\xi =$$

$$= \big(L(\tau,\xi,u(\tau,\xi)) - L_{\dot x}(\tau,\xi,u(\tau,\xi))u(\tau,\xi)\big)d\tau + L_{\dot x}(\tau,\xi,u(\tau,\xi))d\xi.$$

Основная формула Вейерштрасса

В частности, для произвольной функции $x \in PC^1([t_0;\, t_1],\mathbb{R}^n)$ формула для дифференциала функции S примет вид:

$$dS(t,x(t)) =$$

$$= \big(L(t,x(t),u(t,x(t))) - L_{\dot x}(t,x(t),u(t,x(t)))u(t,x(t))\big)dt +$$

$$+ L_{\dot x}(t,x(t),u(t,x(t)))\, dx(t) =$$

$$= \big(L(t,x(t),u(t,x(t))) + L_{\dot x}(t,x(t),u(t,x(t)))(\dot x(t) - u(t,x(t)))\big)dt. \tag{5.4}$$

В таком виде мы ей и будем пользоваться в дальнейшем. Отметим также, что поскольку $u(t,\widehat{x}(t)) = \dot{\widehat{x}}(t)$, то из соотношения (5.4)

$$dS(t,\widehat{x}(t)) = \widehat{L}(t)\, dt.$$

Поэтому для допустимой экстремали $\widehat{x}(\cdot)$ и для произвольной допустимой функции $x(\cdot)$ имеем равенство

$$J(\widehat{x})=\int\limits_{t_0}^{t_1}\widehat{L}(t)\,dt=\int\limits_{t_0}^{t_1}dS(t,\widehat{x}(t))=S(t_1,x_1)-S(t_0,x_0)=\int\limits_{t_0}^{t_1}dS(t,x(t)).$$

Следовательно,

$$J(x)-J(\widehat{x})=\int\limits_{t_0}^{t_1}L(t,x(t),\dot{x}(t))\,dt-\int\limits_{t_0}^{t_1}dS(t,x(t))\overset{(5.4)}{=}$$

$$=\int\limits_{t_0}^{t_1}\Big(L(t,x(t),\dot{x}(t))-L(t,x(t),u(t,x(t)))-L_{\dot{x}}(t,x(t),u(t,x(t)))\times$$

$$\times\big(\dot{x}(t)-u(t,x(t))\big)\Big)\,dt=\int\limits_{t_0}^{t_1}\mathcal{E}\big(t,x(t),u(t,x(t)),\dot{x}(t)\big)\,dt. \quad (5.5)$$

Эту формулу называют *основной формулой Вейерштрасса.*

Для квадратичного по $\dot{x}$ интегранта $L(\dot{x})$ по формуле (5.2) функция Вейерштрасса $\mathcal{E}(t,x,\dot{x},u)=\frac{1}{2}L_{\dot{x}\dot{x}}(u-\dot{x})^2$. Поэтому в этом случае основная формула Вейерштрасса запишется следующим образом:

$$J(x)-J(\widehat{x})=\int\limits_{t_0}^{t_1}\frac{1}{2}L_{\dot{x}\dot{x}}(\dot{x}-u)^2\,dt. \quad (5.6)$$

1.5.5 Достаточные условия сильного экстремума

Теорема 1.5. *Пусть функция $\widehat{x} \in C^2([t_0;\, t_1], \mathbb{R}^n)$ — допустимая экстремаль в задаче (P), интегрант $L \in C^3(V \times \mathbb{R}^n)$, где $V \subset \mathbb{R}^{n+1}$ — некоторая окрестность графика $\Gamma_{\widehat{x}}$, на $\widehat{x}$ выполнены усиленные условия Лежандра и Якоби, интегрант L является выпуклым по $\dot{x}$ на V. Тогда $\widehat{x}$ доставляет сильный локальный минимум ($\widehat{x} \in \operatorname{strlocmin} P$).*

Доказательство. По Теореме 1.4 окружим экстремаль $\widehat{x}$ центральным полем экстремалей $x(\cdot, \lambda)$, покрывающем некоторую окрестность $U \subset V$ графика $\Gamma_{\widehat{x}}$. Пусть $x \in PC^1([t_0;\, t_1])$ — произвольная допустимая функция, график которой расположен в этой окрестности. Тогда по основной формуле Вейерштрасса (5.5)

$$J(x) - J(\widehat{x}) = \int\limits_{t_0}^{t_1} \mathcal{E}\big(t, x(t), u(t, x(t)), \dot{x}(t)\big)\, dt \geq 0,$$

поскольку L выпукло по $\dot{x}$ для $(t, x) \in U$, следовательно, $\mathcal{E}(t, x, u, \dot{x}) \geq 0 \ \forall\ (u, \dot{x}) \in \mathbb{R}^n \times \mathbb{R}^n$. Значит, $J(x) \geq J(\widehat{x})$, т. е. функция $\widehat{x}$ доставляет сильный локальный минимум. $\square$

Рассмотрим простейшую задачу вариационного исчисления в случае, когда интегрант лишь линейно зависит от x, т. е. $L = L(t, x, \dot{x}) = f(t, \dot{x}) + a(t)x$, где $a(\cdot) \in C([t_0;\, t_1])$:

$$J(x) = \int\limits_{t_0}^{t_1} \big(f(t, \dot{x}) + a(t)x\big) dt \to \min; \qquad \begin{array}{l} x(t_0) = x_0, \\ x(t_1) = x_1. \end{array} \qquad (P')$$

Теорема 1.6. *Пусть $\widehat{x} \in DE(P'(S))$, $f, f_{\dot{x}} \in C(\mathbb{R}^2)$, на $\widehat{x}$ выполнено условие Вейерштрасса. Тогда $\widehat{x} \in \operatorname{absmin} P'$.*

Доказательство. Уравнение Эйлера:

$$-\frac{d}{dt}\widehat{L}_{\dot{x}} + \widehat{L}_x = 0 \iff -\frac{d}{dt}\widehat{f}_{\dot{x}} + a = 0.$$

На экстремали $\widehat{x}$ выполняется условие Вейерштрасса, значит,

$$\mathcal{E}(t,\widehat{x}(t),\dot{\widehat{x}}(t),u)=L(t,\widehat{x}(t),u)-L(t,\widehat{x}(t),\dot{\widehat{x}}(t))-\widehat{L}_{\dot{x}}(t)(u-\dot{\widehat{x}}(t))\geq 0$$
$$\forall\ u\in\mathbb{R},\ \forall\ t\in T^{10}.$$

Для интегранта $L=f(t,\dot{x})+ax$ это неравенство перепишется в следующем виде

$$f(t,u)+\underline{a\widehat{x}}-f(t,\dot{\widehat{x}})-\underline{a\widehat{x}}-\widehat{f}_{\dot{x}}(u-\dot{\widehat{x}})\geq 0\ \forall\ u\in\mathbb{R},\ \forall\ t\in T.$$

Возьмём $x(\cdot)$ — произвольную допустимую функцию в задаче на сильный экстремум. Тогда для $u=\dot{x}(t)$

$$f(t,\dot{x}(t))-f(t,\dot{\widehat{x}})-\widehat{f}_{\dot{x}}(\dot{x}(t)-\dot{\widehat{x}})\geq 0\ \forall\ t\in T.$$

Проинтегрируем обе части этого неравенства:

$$\int\limits_{t_0}^{t_1} f(t,\dot{x})\,dt-\int\limits_{t_0}^{t_1} f(t,\dot{\widehat{x}})\,dt-\int\limits_{t_0}^{t_1}\widehat{f}_{\dot{x}}(\dot{x}-\dot{\widehat{x}})\,dt\geq 0\iff$$

$$\int\limits_{t_0}^{t_1} f(t,\dot{x})\,dt-\int\limits_{t_0}^{t_1} f(t,\dot{\widehat{x}})\,dt-\int\limits_{t_0}^{t_1}\widehat{f}_{\dot{x}}\,d(x-\widehat{x})\geq 0\overset{\frac{d}{dt}\widehat{f}_{\dot{x}}=a}{\iff}$$

$$\int\limits_{t_0}^{t_1} f(t,\dot{x})\,dt-\int\limits_{t_0}^{t_1} f(t,\dot{\widehat{x}})\,dt-\widehat{f}_{\dot{x}}(x-\widehat{x})\Big|_{t_0}^{t_1}+\int\limits_{t_0}^{t_1} a(x-\widehat{x})\,dt\geq 0.$$

Свободные члены при интегрировании по частям обращаются в ноль, поскольку $x(t_0)=\widehat{x}(t_0)$, $x(t_1)=\widehat{x}(t_1)$. Поэтому

$$\int\limits_{t_0}^{t_1}\big(f(t,\dot{x})+ax\big)dt\geq\int\limits_{t_0}^{t_1}\big(f(t,\dot{\widehat{x}})+a\widehat{x}\big)dt\iff J(x)\geq J(\widehat{x}).$$

Следовательно, $\widehat{x}\in\operatorname{absmin}P'$. □

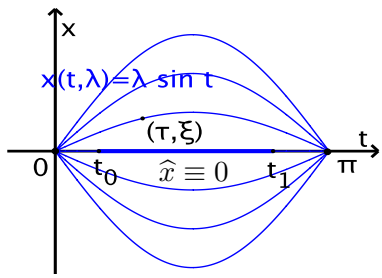
---

[10] $T\subset[t_0;\,t_1]$ — множество точек непрерывности функции $\dot{\widehat{x}}$.

Пример (гармонический осциллятор)**.**

$$\int\limits_{t_0}^{t_1} (\dot{x}^2 - x^2)\, dt \to \min;\ x(t_0) = x(t_1) = 0 \quad (0 < t_0 < t_1 < \pi).$$

Решение. Уравнение Эйлера $\ddot{x} + x = 0$. Экстремали этого функционала имеют вид $x(t) = C_1 \sin t + C_2 \cos t$. Допустимая экстремаль $\widehat{x}(t) \equiv 0$. Совокупность экстремалей $x(t, \lambda) = \lambda \sin t$ есть центральное поле экстремалей с центром в точке $(0, 0)$, включающее, в частности, экстремаль $\widehat{x}$ при $\lambda = 0$, покрывающее полосу $0 < t < \pi$.



Найдём экстремаль поля, проходящую через точку (τ, ξ) $(0 < \tau < \pi)$:

$$x(\tau, \lambda) = \xi \Leftrightarrow \lambda \sin \tau = \xi \Leftrightarrow \lambda = \frac{\xi}{\sin \tau} \Rightarrow x(t, \lambda(\tau, \xi)) = \frac{\xi}{\sin \tau} \sin t.$$

Функция наклона поля

$$u(\tau, \xi) = \frac{d}{dt} x(t, \lambda(\tau, \xi))\Big|_{t=\tau} = \frac{d}{dt} \frac{\xi}{\sin \tau} \sin t\Big|_{t=\tau} = \xi \operatorname{ctg} \tau.$$

По основной формуле Вейерштрасса для любой допустимой функции x и функции наклона поля $u(t, x(t)) = x(t) \operatorname{ctg} t$

$$J(x) - J(\widehat{x}) = \int\limits_{t_0}^{t_1} \mathcal{E}\big(t, x, u(t, x), \dot{x}\big)\, dt =$$

$$= \int\limits_{t_0}^{t_1} \Big(L(t,x,\dot{x}) - L(t,x,u(t,x)) - L_{\dot{x}}(t,x,u(t,x))\big(\dot{x} - u(t,x)\big)\Big) dt =$$

$$= \int\limits_{t_0}^{t_1} \big((\dot{x}^2 - x^2) - (u^2 - x^2) - 2u(\dot{x} - u)\big) dt =$$

$$= \int\limits_{t_0}^{t_1} \big(\dot{x}^2 - 2u\dot{x} + u^2\big) dt = \int\limits_{t_0}^{t_1} (\dot{x} - u)^2 \, dt = \int\limits_{t_0}^{t_1} (\dot{x} - x \operatorname{ctg} t)^2 \, dt \geq 0.$$

Поэтому $\widehat{x} \in \mathrm{absmin}$.

1.5.6 Представление квадратичного функционала

В Главе 5 п. 1.3 было показано, что необходимым условием 2-го порядка слабого локального минимума является выполнения неравенства (5.1):

$$J''(\widehat{x})[h,h] = \int\limits_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{x\dot{x}}\dot{h}h + \widehat{L}_{xx}h^2\big)dt \geq 0 \ \forall\, h \in C_0^1([t_0;\, t_1]).$$

Поэтому важную роль при выводе необходимых и достаточных условий экстремума играет задача об оценке квадратичного функционала

$$K(h) = \int\limits_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{x\dot{x}}\dot{h}h + \widehat{L}_{xx}h^2\big)dt = \int\limits_{t_0}^{t_1} (\underbrace{A\dot{h}^2 + 2B\dot{h}h + Ch^2}_{\widetilde{L}})\, dt,$$

считаем, что $A(\cdot), B(\cdot) \in C^1([t_0;\, t_1])$, $C(\cdot) \in C([t_0;\, t_1])$.

Поскольку интегрант $\widetilde{L}$ является квадратичным, то уравнение Якоби совпадает с уравнением Эйлера

$$-\frac{d}{dt}(A\dot{u} + Bu) + B\dot{u} + Cu = 0. \qquad \text{(Э-Я)}$$

Лемма 2. *Пусть для интегранта $\widetilde{L}$ функционала K выполняется усиленное условие Лежандра ($A(t) \neq 0$ для любого $t \in [t_0;\, t_1]$) и условие Якоби (на интервале $(t_0;\, t_1)$ нет точек сопряженных с t_0). Тогда функционал K представим в виде*

$$\int\limits_{t_0}^{t_1} (A\dot{h}^2 + 2B\dot{h}h + Ch^2)\, dt = \int\limits_{t_0}^{t_1} A\left(\dot{h} - \frac{\dot{u}}{u}h\right)^2 dt \ \forall\, h(\cdot) \in C_0^1([t_0;\, t_1]),$$

где $u(\cdot)$ — любое решение уравнения Якоби для интегранта $\widetilde{L}$ с начальным условием $u(t_0) = 0$, $\dot{u}(t_0) \neq 0$.

Доказательство. По условию существует решение $u(\cdot)$ уравнения Эйлера–Якоби с начальными условиями $u(t_0) = 0$, $\dot{u}(t_0) = 1$, не имеющее нулей на интервале $(t_0; t_1)$. Как решение однородного линейного дифференциального уравнения 2-го порядка с непрерывными коэффициентами функция $u(\cdot) \in C^2([t_0; t_1])$, а любое решение с условиями $u(t_0) = 0$, $\dot{u}(t_0) \neq 0$ ему пропорционально. Поэтому представление достаточно доказать для одной функции $u(\cdot)$.

Сведём выражение в правой части равенства к левой.

$$\int\limits_{t_0}^{t_1} A\left(\dot{h} - \frac{\dot{u}}{u}h\right)^2 dt = \int\limits_{t_0}^{t_1} \left(A\dot{h}^2 + \frac{A\dot{u}^2}{u^2}h^2\right)dt - \int\limits_{t_0}^{t_1} 2\frac{A\dot{u}}{u}\dot{h}h\,dt =$$

$$= \int\limits_{t_0}^{t_1} \left(A\dot{h}^2 + \frac{A\dot{u}^2}{u^2}h^2\right)dt - \int\limits_{t_0}^{t_1} \frac{A\dot{u}}{u}\,dh^2 =$$

(возьмем последний интеграл по частям)

$$= \int\limits_{t_0}^{t_1} \left(A\dot{h}^2 + \frac{A\dot{u}^2}{u^2}h^2\right)dt - \frac{A\dot{u}}{u}h^2\Big|_{t_0}^{t_1} + \int\limits_{t_0}^{t_1} h^2 d\left(\frac{A\dot{u}}{u}\right) =$$

(Свободные члены при интегрировании по частям обращаются в ноль, поскольку функция $h(\cdot) \in C_0^1([t_0; t_1])$, а $u(\cdot) \in C^2([t_0; t_1])$ имеет однократный нуль в точке t_0 и быть может еще в точке t_1. Далее берем дифференциал частного.)

$$= \int\limits_{t_0}^{t_1} \left(A\dot{h}^2 + \frac{A\dot{u}^2}{u^2}h^2\right)dt + \int\limits_{t_0}^{t_1} \frac{(A\dot{u})^{\cdot}u - A\dot{u}\dot{u}}{u^2}h^2\,dt =$$

$$= \int\limits_{t_0}^{t_1} \left(A\dot{h}^2 + \frac{A\dot{u}^2 + (A\dot{u})^{\cdot}u - A\dot{u}^2}{u^2}h^2\right)dt = \int\limits_{t_0}^{t_1} \left(A\dot{h}^2 + \frac{(A\dot{u})^{\cdot}}{u}h^2\right)dt =$$

(из уравнения Эйлера–Якоби $-(A\dot{u}+Bu)^{\cdot}+B\dot{u}+Cu=0$ выразим $(A\dot{u})^{\cdot}=-(Bu)^{\cdot}+B\dot{u}+Cu=-\dot{B}u-B\dot{u}+B\dot{u}+Cu=$ $=-\dot{B}u+Cu=(-\dot{B}+C)u$ и подставим в подынтегральное выражение)

$$=\int\limits_{t_0}^{t_1}\left(A\dot{h}^2-\dot{B}h^2+Ch^2\right)dt=\int\limits_{t_0}^{t_1}\left(A\dot{h}^2+2B\dot{h}h+Ch^2\right)dt=K(h),$$

так как

$$\int\limits_{t_0}^{t_1}(-\dot{B}h^2)dt=-\int\limits_{t_0}^{t_1}h^2dB=-h^2B\Big|_{t_0}^{t_1}+\int\limits_{t_0}^{t_1}B\,dh^2\overset{\substack{h(t_1)=0\\h(t_0)=0}}{=}\int\limits_{t_0}^{t_1}2B\dot{h}h\,dt.$$

$\square$

Лемма 3. *Пусть для интегранта $\widetilde{L}$ функционала K выполняется усиленное условие Лежандра ($A(t) > 0$ для любого $t \in [t_0;\, t_1]$) и усиленное условие Якоби (на полуинтервале $(t_0;\, t_1]$ нет точек сопряженных с t_0). Тогда существует положительная константа $\alpha > 0$ такая, что*

$$K(h) = \int\limits_{t_0}^{t_1} (A\dot{h}^2 + 2B\dot{h}h + Ch^2)\, dt \geq \alpha \int\limits_{t_0}^{t_1} \dot{h}^2 dt \;\forall\, h(\cdot) \in C_0^1([t_0;\, t_1]).$$

Доказательство. Пусть $u(\cdot)$ — решение уравнения Якоби с начальными условиями $u(t_0) = 0$, $\dot{u}(t_0) = 1$. Для любого $0 < \alpha < \min\limits_{t \in [t_0;\, t_1]} A(t)$ функция $A_\alpha(t) = A(t) - \alpha$ положительна на отрезке $[t_0;\, t_1]$. Поэтому выполняется усиленное условие Лежандра для интегранта $\widetilde{L}_\alpha = A_\alpha \dot{h}^2 + 2B\dot{h}h + Ch^2$. Для таких α рассмотрим решение уравнения Якоби для этого интегранта с начальным условием $u_\alpha(t_0) = 0$, $\dot{u}_\alpha(t_0) = 1$.

По теореме о непрерывной зависимости решения дифференциального уравнения от параметров получаем, что $u_\alpha(\cdot) \to u(\cdot)$ при $\alpha \to +0$ в пространстве $C([t_0;\, t_1])$. Тогда для достаточно малых $\alpha > 0$ функции $u_\alpha(t) > 0 \;\forall\, t \in (t_0;\, t_1]$, т. е. также выполняется усиленное условие Якоби для интегранта $\widetilde{L}_\alpha$. По Лемме 2 для любого $h(\cdot) \in C_0^1([t_0;\, t_1])$

$$K_\alpha(h) := \int\limits_{t_0}^{t_1} \big(A_\alpha \dot{h}^2 + 2B\dot{h}h + Ch^2\big) dt = \int\limits_{t_0}^{t_1} A_\alpha \left(\dot{h} - \frac{\dot{u}_\alpha}{u_\alpha} h\right)^2 dt \geq 0.$$

Поскольку $A(t) = A_\alpha(t) + \alpha$, то

$$K(h) = \int\limits_{t_0}^{t_1} \big(A\dot{h}^2 + 2B\dot{h}h + Ch^2\big) dt = \int\limits_{t_0}^{t_1} \big((A_\alpha + \alpha)\dot{h}^2 + 2B\dot{h}h + Ch^2\big) dt =$$

$$= K_\alpha(h) + \alpha \int\limits_{t_0}^{t_1} \dot{h}^2 dt \geq \alpha \int\limits_{t_0}^{t_1} \dot{h}^2 dt. \qquad \square$$

Гармонический осциллятор

Вернемся к задаче о гармоническом осцилляторе и посмотрим, какое получается неравенство из Леммы 1. Возьмём функцию $u(\cdot)$, для которой $u(t_0) = 0$, $\dot{u}(t_0) \neq 0$, не имеющую нулей на интервале $(t_0;\, t_1)$, например,
$u(t) = \sin \dfrac{\pi(t-t_0)}{t_1-t_0}$. Для нее $\dfrac{\dot{u}}{u} = \dfrac{\pi}{t_1-t_0} \operatorname{ctg} \dfrac{\pi(t-t_0)}{t_1-t_0}$.
Функция $u(\cdot)$ удовлетворяет дифференциальному уравнению

$$\ddot{u} + \left(\frac{\pi}{t_1-t_0}\right)^2 u = 0.$$

Это уравнение является уравнением Эйлера–Якоби интегранта $L = \dot{h}^2 - \left(\dfrac{\pi}{t_1-t_0}\right)^2 h^2$.
По Лемме 2 для любой функции $h(\cdot) \in C_0^1([t_0;\, t_1])$

$$\int\limits_{t_0}^{t_1} \left(\dot{h}^2 - \left(\frac{\pi}{t_1-t_0}\right)^2 h^2\right) dt = \int\limits_{t_0}^{t_1} \left(\dot{h} - \frac{\dot{u}}{u} h\right)^2 dt =$$
$$= \int\limits_{t_0}^{t_1} \left(\dot{h} - \frac{\pi}{t_1-t_0} \operatorname{ctg} \frac{\pi(t-t_0)}{t_1-t_0} h\right)^2 dt \geq 0.$$

Следовательно, выполняется неравенство для любой функции $h(\cdot) \in C_0^1([t_0;\, t_1])$

$$\int\limits_{t_0}^{t_1} \dot{h}^2 dt \geq \frac{\pi^2}{(t_1-t_0)^2} \int\limits_{t_0}^{t_1} h^2 dt \iff \int\limits_{t_0}^{t_1} h^2 dt \leq \frac{(t_1-t_0)^2}{\pi^2} \int\limits_{t_0}^{t_1} \dot{h}^2 dt. \tag{5.7}$$

В виде неравенства для норм

$$\|h\|_{L_2([t_0;\, t_1])} \leq \frac{t_1-t_0}{\pi} \|\dot{h}\|_{L_2([t_0;\, t_1])}.$$

1.5.7 Достаточные условия слабого экстремума

Рассмотрим простейшую задачу вариационного исчисления, для определенности, на минимум

$$J(x(\cdot)) = \int\limits_{t_0}^{t_1} L\left(t, x(t), \dot{x}(t)\right) dt \to \inf;\ x(t_0)=x_0,\ x(t_1)=x_1. \quad (P)$$

Пусть $\widehat{x}(\cdot) \in C^1([t_0;\, t_1])$ — некоторая фиксированная допустимая экстремаль, т. е. функция $\widehat{x}$ удовлетворяет уравнению Эйлера, которое равносильно выполнению уравнения $J'(\widehat{x})[h] = 0\ \forall\ h(\cdot) \in C_0^1([t_0;\, t_1])$.

Возьмём произвольную допустимую функцию $x(\cdot)$. Тогда функция $h = x - \widehat{x}$ принадлежит пространству $C_0^1([t_0;\, t_1])$ $(x = \widehat{x} + h)$. По формуле Тейлора

$$J(x) = J(\widehat{x}+h) = J(\widehat{x})+J'(\widehat{x})[h]+\frac{1}{2}J''(\widehat{x})[h,h]+o\big(\|h\|^2_{C^1([t_0;\, t_1])}\big).$$

Достаточным условием локального минимума на допустимой экстремали $\widehat{x}$ для задачи с ограничением типа равенства в банаховых пространствах согласно теореме является выполнение следующего неравенства на пространстве $C_0^1([t_0;\, t_1])$ при некотором $\alpha > 0$

$$J''(\widehat{x})[h,h] \geq \alpha\|h\|^2_{C^1([t_0;\, t_1])} \iff$$

$$\int\limits_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}\dot{x}}\dot{h}^2+2\widehat{L}_{x\dot{x}}\dot{h}h+\widehat{L}_{xx}h^2\big)dt \geq \alpha\big(\max_{t\in[t_0;\, t_1]}|\dot{h}(t)|+\max_{t\in[t_0;\, t_1]}|h(t)|\big)^2.$$

Теорема 1.7. *Пусть функция $\widehat{x} \in C^2([t_0;\, t_1], \mathbb{R}^n)$ — допустимая экстремаль в задаче (P), интегрант L трижды непрерывно дифференцируем в некоторой окрестности расширенного графика $\Gamma_{\widehat{x}\dot{\widehat{x}}}$, на $\widehat{x}$ выполнены усиленные условия Лежандра и Якоби. Тогда $\widehat{x} \in \operatorname{wlocmin} P$.*

Доказательство. Пусть $U \in \mathcal{O}(\Gamma_{\widehat{x}\dot{\widehat{x}}})$ — некоторая окрестность расширенного графика $\Gamma_{\widehat{x}\dot{\widehat{x}}}$. Тогда при достаточно малом δ_0 компакт

$$\mathcal{K} := \{(t, x, \dot{x}) \mid t \in [t_0;\, t_1],\ |x - \widehat{x}(t)| \le \delta_0,\ |\dot{x} - \dot{\widehat{x}}(t)| \le \delta_0\} \subset U.$$

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 $\mathcal{K}$. Следовательно, для любого $\varepsilon > 0$ найдется $0 < \delta = \delta(\varepsilon) < \frac{\delta_0}{2}$ такое, что для любых $(t, x, \dot{x}) \in \mathcal{K}$, для которых $|x - \widehat{x}(t)| < \delta$ и $|\dot{x} - \dot{\widehat{x}}(t)| < \delta$

$$\begin{cases} |L_{\dot{x}\dot{x}}(t, x, \dot{x}) - \widehat{L}_{\dot{x}\dot{x}}(t)| < \varepsilon, \\ |L_{\dot{x}x}(t, x, \dot{x}) - \widehat{L}_{\dot{x}x}(t)| < \varepsilon, \\ |L_{xx}(t, x, \dot{x}) - \widehat{L}_{xx}(t)| < \varepsilon, \end{cases} \Rightarrow \begin{cases} L_{\dot{x}\dot{x}}(t, x, \dot{x}) > \widehat{L}_{\dot{x}\dot{x}}(t) - \varepsilon, \\ L_{\dot{x}x}(t, x, \dot{x}) > \widehat{L}_{\dot{x}x}(t) - \varepsilon, \\ L_{xx}(t, x, \dot{x}) > \widehat{L}_{xx}(t) - \varepsilon. \end{cases} \quad (1)$$

Возьмём произвольную допустимую функцию x. Тогда функция $h = x - \widehat{x}$ принадлежит пространству $C_0^1([t_0;\, t_1])$ $(x = \widehat{x} + h)$ и пусть $\|x - \widehat{x}\|_{C^1([t_0;\, t_1])} = \|h\|_{C^1([t_0;\, t_1])} < \delta$.
По формуле Тейлора с остаточным членом в форме Лагранжа

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

$(\theta = \theta(t) \in (0;\, 1))$

$$\overset{(1)}{\ge} \widehat{L} + \widehat{L}_{\dot{x}}\dot{h} + \widehat{L}_x h + \frac{1}{2}\big(\widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{\dot{x}x}\dot{h}h + \widehat{L}_{xx}h^2\big) - \frac{1}{2}\varepsilon(\dot{h}^2 + 2\dot{h}h + h^2).$$

Интегрируя это неравенство с учетом равенства

$$J'(\widehat{x})[h] = \int_{t_0}^{t_1} (\widehat{L}_{\dot{x}}\dot{h} + \widehat{L}_x h)\, dt = 0 \ \forall\ h \in C_0^1([t_0;\, t_1]),$$

получаем

$$J(x) = \int_{t_0}^{t_1} L(t, \widehat{x}+h, \dot{\widehat{x}}+\dot{h})\, dt \ge J(\widehat{x}) + \frac{1}{2}K(h) - \frac{\varepsilon}{2}\int_{t_0}^{t_1} (\dot{h}^2 + 2\dot{h}h + h^2)dt. \quad (2)$$

По Лемме 3 существует такое $\alpha > 0$, что

$$K(h) = \int\limits_{t_0}^{t_1} \left(\widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{x\dot{x}}\dot{h}h + \widehat{L}_{xx}h^2\right)dt \geq \alpha \int\limits_{t_0}^{t_1} \dot{h}^2 dt \ \forall\ h \in C_0^1([t_0;\, t_1]).$$

Оценим вычитаемый интеграл в правой части неравенства (2). В силу очевидного неравенства $2ab \leq a^2 + b^2$ имеем $\dot{h}^2 + 2\dot{h}h + h^2 \leq 2(\dot{h}^2 + h^2)$. Поэтому, а так же в силу неравенства (5.7) $\int\limits_{t_0}^{t_1} h^2 dt \leq \frac{(t_1-t_0)^2}{\pi^2} \int\limits_{t_0}^{t_1} \dot{h}^2 dt \ \forall\ h \in C_0^1([t_0;\, t_1])$, получим

$$\int\limits_{t_0}^{t_1} (\dot{h}^2 + 2\dot{h}h + h^2)\, dt \leq 2 \int\limits_{t_0}^{t_1} (\dot{h}^2 + h^2) dt \leq$$

$$\leq 2 \int\limits_{t_0}^{t_1} \left(1 + \frac{(t_1 - t_0)^2}{\pi^2}\right) \dot{h}^2 dt = c \int\limits_{t_0}^{t_1} \dot{h}^2 dt.$$

Следовательно, из (2)

$$J(x) - J(\widehat{x}) \geq \frac{\alpha - \varepsilon c}{2} \int\limits_{t_0}^{t_1} \dot{h}^2 dt \geq 0$$

при $\varepsilon < \frac{\alpha}{c}$. Значит, $\widehat{x} \in \operatorname{wlocmin} P$. $\square$

1.5.8 Квадратично-линейный функционал

Рассмотрим простейшую задачу вариационного исчисления с квадратично-линейным функционалом

$$J(x)=\int\limits_{t_0}^{t_1}(\underbrace{A\dot{x}^2+2C\dot{x}x+Bx^2+2Dx}_{L})\,dt\to\text{extr};\quad \begin{matrix} x(t_0)=x_0,\\ x(t_1)=x_1,\end{matrix}\qquad (P')$$

где функции $A(\cdot), C(\cdot) \in C^1([t_0;\, t_1])$, $B(\cdot), D(\cdot) \in C([t_0;\, t_1])$, $A(t) \neq 0\ \forall\ t \in [t_0;\, t_1]$.

Для квадратично-линейного функционала

$$J'(x)[h]=\int\limits_{t_0}^{t_1}\big(L_{\dot{x}}\dot{h}+L_x h\big)dt=2\int\limits_{t_0}^{t_1}\big(A\dot{x}\dot{h}+Cx\dot{h}+C\dot{x}h+Bxh+Dh\big)dt,$$

$$J''(\widehat{x})[h,h]=J''[h,h]=2\int\limits_{t_0}^{t_1}(\underbrace{A\dot{h}^2+2C\dot{h}h+Bh^2}_{\widetilde{L}})\,dt.$$

Уравнение Эйлера в задаче (P')

$$-\frac{d}{dt}L_{\dot{x}}+L_x=0 \iff -\frac{d}{dt}(A\dot{x}+Cx)+C\dot{x}+Bx+D=0. \quad (\text{Э})$$

Продифференцировав в уравнении Эйлера по t, получим уравнение Эйлера в эквивалентном виде:

$$\begin{aligned} -\dot{A}\dot{x}-A\ddot{x}-\dot{C}x-\underline{C\dot{x}}+\underline{C\dot{x}}+Bx+D=0 &\iff\\ A\ddot{x}+\dot{A}\dot{x}+(\dot{C}-B)x-D=0 &\iff\\ A\ddot{x}+\dot{A}\dot{x}+Ex-D=0, &\quad (5.8)\end{aligned}$$

где $E=\dot{C}-B$.

Интегрант квадратичного функционала $\widetilde{L}$ совпадает с квадратичной частью интегранта L. Следовательно, уравнение Якоби — уравнение Эйлера для интегранта $\widetilde{L}$ — совпадает с однородной частью уравнения Эйлера исходной задачи

$$-\frac{d}{dt}\widetilde{L}_{\dot{h}}+\widetilde{L}_h=0 \iff -\frac{d}{dt}(A\dot{h}+Ch)+C\dot{h}+Bh=0. \quad (\text{Я})$$

Аналогично, уравнение Якоби переписывается в эквивалентном виде:

$$A\ddot{h} + \dot{A}\dot{h} + Eh = 0. \tag{5.9}$$

Для квадратично-линейных функционалов по формуле Тейлора имеет место равенство

$$J(x+h) = J(x) + J'(x)[h] + \frac{1}{2}J''[h,h].$$

Если $\widehat{x}$ — допустимая экстремаль в задаче (P'), то $J'(\widehat{x})[h] = 0$ $\forall\ h \in C_0^1([t_0;\, t_1])$ (необходимое условие 1-го порядка эквивалентное уравнению Эйлера см. Глава 3 § 1). Значит, на допустимой экстремали $\widehat{x}$

$$J(\widehat{x}+h) = J(\widehat{x}) + \frac{1}{2}J''[h,h] \quad \forall\ h \in C_0^1([t_0;\, t_1]). \tag{$*$}$$

Утверждение 3. *a) Если функция $x \in C_0^1([t_0;\, t_1])$ удовлетворяет уравнению Эйлера (Э), то $J(x) = \int\limits_{t_0}^{t_1} Dx\, dt$.*

b) Если функция $h \in C_0^1([t_0;\, t_1])$ удовлетворяет уравнению Якоби (Я), то $J(h) = 2\int\limits_{t_0}^{t_1} Dh\, dt$, $J''[h,h] = 0$.

Доказательство. a) Пусть функция $x \in C_0^1([t_0;\, t_1])$ удовлетворяет уравнению Эйлера. Тогда

$$J(x) = \int\limits_{t_0}^{t_1} (A\dot{x}^2 + 2C\dot{x}x + Bx^2 + 2Dx)\, dt =$$

$$= \int\limits_{t_0}^{t_1} (A\dot{x} + Cx)\dot{x}\, dt + \int\limits_{t_0}^{t_1} (C\dot{x} + Bx)x\, dt + 2\int\limits_{t_0}^{t_1} Dx\, dt =$$

$$= \int\limits_{t_0}^{t_1} (A\dot{x} + Cx)\, dx + \int\limits_{t_0}^{t_1} (C\dot{x} + Bx)x\, dt + 2\int\limits_{t_0}^{t_1} Dx\, dt =$$

$$=(A\dot{x}+Cx)x\Big|_{t_0}^{t_1}+\int\limits_{t_0}^{t_1}\Big(-\frac{d}{dt}(A\dot{x}+Cx)+C\dot{x}+Bx\Big)x\,dt+2\int\limits_{t_0}^{t_1}Dx\,dt\overset{\substack{x(t_0)=0\\x(t_1)=0}}{=}$$

$$=\int\limits_{t_0}^{t_1}\Big(-\frac{d}{dt}(A\dot{x}+Cx)+C\dot{x}+Bx+D\Big)x\,dt+\int\limits_{t_0}^{t_1}Dx\,dt=\int\limits_{t_0}^{t_1}Dx\,dt.$$

Выражение в скобках под знаком интеграла обращается в ноль, так как функция x удовлетворяет уравнению Эйлера.

b) Пусть функция $h\in C_0^1([t_0;\,t_1])$ удовлетворяет уравнению Якоби. Тогда из предыдущего равенства

$$J(h)=\int\limits_{t_0}^{t_1}\Big(-\frac{d}{dt}(A\dot{h}+Ch)+C\dot{h}+Bh\Big)h\,dt+2\int\limits_{t_0}^{t_1}Dh\,dt=2\int\limits_{t_0}^{t_1}Dh\,dt.$$

Выражение в скобках под знаком интеграла обращается в ноль, так как функция h удовлетворяет уравнению Якоби.

Вторая производная функционала:

$$\frac{1}{2}J''[h,h]=\int\limits_{t_0}^{t_1}(A\dot{h}^2+2C\dot{h}h+Bh^2)\,dt=$$

$$=\int\limits_{t_0}^{t_1}(A\dot{h}^2+2C\dot{h}h+Bh^2+2Dh)\,dt-2\int\limits_{t_0}^{t_1}Dh\,dt=$$

$$=J(h)-2\int\limits_{t_0}^{t_1}Dh\,dt=0\Longrightarrow J''[h,h]=0.\qquad\square$$

Утверждение 4. *Пусть функция h удовлетворяет уравнению Якоби (5.9), $h(t) \neq 0 \ \forall \ t \in (t_0, t_1)$, функция $x \in C^2([t_0; t_1])$. Тогда функция x удовлетворяет уравнению Эйлера (5.8) тогда и только тогда, когда определитель Вронского*

$$w(t) = \begin{vmatrix} h(t) & x(t) \\ \dot{h}(t) & \dot{x}(t) \end{vmatrix} = h(t)\dot{x}(t) - \dot{h}(t)x(t)$$

удовлетворяет дифференциальному уравнению

$$(Aw)^{\cdot} = Dh. \tag{5.10}$$

Доказательство. Преобразуем левую часть уравнения (5.10), продифференцировав определитель Вронского по t,

$$\begin{aligned}(Aw)^{\cdot} &= \big(A(h\dot{x} - \dot{h}x)\big)^{\cdot} = \big(Ah\dot{x} - A\dot{h}x\big)^{\cdot} = \\ &= \dot{A}h\dot{x} + \underline{A\dot{h}\dot{x}} + Ah\ddot{x} - \dot{A}\dot{h}x - A\ddot{h}x - \underline{A\dot{h}\dot{x}} = \\ &= \dot{A}h\dot{x} + Ah\ddot{x} - \dot{A}\dot{h}x - A\ddot{h}x. \end{aligned} \tag{5.11}$$

Необходимость. Пусть функция x удовлетворяет уравнению Эйлера (5.8). Покажем, что тогда выполняется уравнение $(Aw)^{\cdot} = Dh$. Имеем

$$(Aw)^{\cdot} \overset{(5.11)}{=} \dot{A}h\dot{x} + Ah\ddot{x} - \dot{A}\dot{h}x - A\ddot{h}x =$$

$$\left(\text{подставим из уравнений} \quad \begin{matrix} (5.8)\ A\ddot{x} = -\dot{A}\dot{x} - Ex + D, \\ (5.9)\ A\ddot{h} = -\dot{A}\dot{h} - Eh, \end{matrix} \right)$$

$$= \dot{A}h\dot{x} + h(-\dot{A}\dot{x} - Ex + D) - \dot{A}\dot{h}x - (-\dot{A}\dot{h} - Eh)x = Dh.$$

Достаточность. Пусть функция x удовлетворяет уравнению $(Aw)^{\cdot} = Dh$. Докажем, что тогда x удовлетворяет уравнению Эйлера (5.8). Имеем

$$(Aw)^{\cdot} = Dh \overset{(5.11)}{\Longleftrightarrow} \dot{A}h\dot{x} + Ah\ddot{x} - \dot{A}\dot{h}x - A\ddot{h}x = Dh$$

(подставим из уравнения (5.9) $A\ddot{h} = -\dot{A}\dot{h} - Eh$)

$$\dot{A}h\dot{x} + Ah\ddot{x} - \underline{\dot{A}\dot{h}x} + \underline{\dot{A}\dot{h}x} + Ehx = Dh \Longleftrightarrow (A\ddot{x} + \dot{A}\dot{x} + Ex - D)h = 0.$$

Поскольку $h(t) \neq 0 \ \forall \ t \in (t_0, t_1)$, то $A\ddot{x} + \dot{A}\dot{x} + Ex - D = 0$, т. е. функция x удовлетворяет уравнению (5.8). $\square$

Теорема 1.8. *Пусть уравнение Якоби (5.9) имеет решение h с граничными условиями $h(t_0) = h(t_1) = 0$, не имеющее других нулей на интервале (t_0, t_1).*

Тогда решение уравнения Эйлера (5.8) с граничными условиями $x(t_0) = x_0$, $x(t_1) = x_1$ существует[11] *тогда и только тогда, когда выполнено условие*

$$G := A(t_1)\dot{h}(t_1)x_1 - A(t_0)\dot{h}(t_0)x_0 + \widehat{D} = 0, \qquad (G2)$$

где $\widehat{D} := \int\limits_{t_0}^{t_1} Dh\,dt$.

Доказательство. Необходимость. Пусть существует решение уравнения Эйлера (5.8) x с заданными граничными условиями $x(t_0) = x_0$, $x(t_1) = x_1$. Покажем, что тогда выполняется условие $G = 0$. Выпишем определитель Вронского для функций h и x:

$$w = \begin{vmatrix} h & x \\ \dot{h} & \dot{x} \end{vmatrix} = h\dot{x} - \dot{h}x.$$

По Утверждению 4 если функция h удовлетворяет уравнению Якоби (5.9), а функция x удовлетворяет уравнению Эйлера (5.8), то для пары h и x определитель Вронского удовлетворяет уравнению (5.10):

$$(Aw)^{\cdot} = Dh.$$

Проинтегрируем обе части этого уравнения:

$$\int\limits_{t_0}^{t_1} (Aw)^{\cdot}\,dt = \int\limits_{t_0}^{t_1} Dh\,dt \iff A(t_1)w(t_1) - A(t_0)w(t_0) = \widehat{D} \iff$$

$$A(t_1)\big(h(t_1)\dot{x}(t_1) - \dot{h}(t_1)x(t_1)\big) - A(t_0)\big(h(t_0)\dot{x}(t_0) - \dot{h}(t_0)x(t_0)\big) = \widehat{D}$$

$$\overset{h(t_0)=h(t_1)=0}{\iff} A(t_1)\dot{h}(t_1)x_1 - A(t_0)\dot{h}(t_0)x_0 + \widehat{D} = 0.$$

[11] Поскольку к фиксированному решению $\widehat{x}(\cdot)$ можно прибавить функцию $Ch(\cdot)$, $C \in \mathbb{R}$, не изменяющую граничные условия функции $x(\cdot)$, то все функции вида $\widehat{x}(\cdot) + Ch(\cdot)$ также будут являться решениями уравнения Эйлера с заданными граничными условиями.

Достаточность. Пусть для чисел x_0, x_1 выполняется условие $G = 0$. Надо показать, что существует решение уравнения Эйлера (5.8) с граничными условиями $x(t_0) = x_0$, $x(t_1) = x_1$.

Возьмём функцию w, удовлетворяющую уравнению (5.10):

$$(Aw)^{\cdot} = Dh \iff Aw = \int\limits_{t_0}^{t} Dh\,ds + C \iff w = \frac{\int\limits_{t_0}^{t} Dh\,ds + C}{A}. \quad (w)$$

Функция w удовлетворяет уравнению (5.10) для любого $C \in \mathbb{R}$ (константу C выберем позже). Найдём функцию x, для которой определитель Вронского пары h, x совпадает с w. Тогда функция x в силу Утверждения 4 будет удовлетворять уравнению (5.8).

Имеем:

$$h\dot{x} - \dot{h}x = w \iff \frac{h\dot{x} - \dot{h}x}{h^2} = \frac{w}{h^2} \iff \left(\frac{x}{h}\right)^{\cdot} = \frac{w}{h^2}.$$

Проинтегрируем обе части последнего уравнения:

$$\frac{x}{h} = \int\limits_{t_0}^{t} \frac{w(\tau)}{h^2(\tau)}\,d\tau + C' \iff x = h\int\limits_{t_0}^{t} \frac{w(\tau)}{h^2(\tau)}\,d\tau + C'h.$$

Выберем константу C, чтобы у функции x были заданные граничные условия. Поскольку у функции h нулевые граничные условия и не влияют на граничные условия функции x, то для простоты записи возьмём $C' = 0$. Перейдем к пределу в последнем равенстве при $t \to t_0$ и при $t \to t_1$, используя правило Лопиталя и учитывая, что $h(t_0) = h(t_1) = 0$:

$$x(t_0) = \lim_{t \to t_0} h(t) \int\limits_{t_0}^{t} \frac{w(\tau)}{h^2(\tau)}\,d\tau = \lim_{t \to t_0} \frac{\int\limits_{t_0}^{t} \frac{w(\tau)}{h^2(\tau)}\,d\tau}{\frac{1}{h(t)}} =$$

$$= \lim_{t\to t_0} \frac{\dfrac{w(t)}{h^2(t)}}{-\dfrac{\dot h(t)}{h^2(t)}} = -\frac{w(t_0)}{\dot h(t_0)} = -\frac{C}{A(t_0)\dot h(t_0)} = x_0$$

$$\left(w(t_0) \overset{(w)}{=} \frac{\int\limits_{t_0}^{t_0} Dh\,ds + C}{A(t_0)} = \frac{C}{A(t_0)} \right) \text{ при } C = -A(t_0)\dot h(t_0)x_0;$$

$$x(t_1) = \lim_{t\to t_1} h(t) \int\limits_{t_0}^{t} \frac{w(\tau)}{h^2(\tau)}\,d\tau = \lim_{t\to t_1} \frac{\int\limits_{t_0}^{t} \dfrac{w(\tau)}{h^2(\tau)}\,d\tau}{\dfrac{1}{h(t)}} = \lim_{t\to t_1} \frac{\dfrac{w(t)}{h^2(t)}}{-\dfrac{\dot h(t)}{h^2(t)}} =$$

$$= -\frac{w(t_1)}{\dot h(t_1)} \overset{(w)}{=} -\frac{\int\limits_{t_0}^{t_1} Dh\,ds + C}{A(t_1)\dot h(t_1)} = -\frac{\widehat{D} - A(t_0)\dot h(t_0)x_0}{A(t_1)\dot h(t_1)} = x_1$$

в силу выполнения соотношения (G2):

$A(t_1)\dot h(t_1)x_1 - A(t_0)\dot h(t_0)x_0 + \widehat{D} = 0.$

Значит, построенная функция $x(\cdot)$ удовлетворяет уравнению (5.8) и имеет заданные граничные условия. $\square$

Теорема 1.9. *Пусть* $A, C \in C^1([t_0; t_1])$, $B, D \in C([t_0; t_1])$, *выполнено усиленное условие Лежандра на минимум. Тогда*

a) если выполнено усиленное условие Якоби, то допустимая экстремаль $\widehat{x}$ *существует, единственна и доставляет абсолютный минимум;*

b) если выполнено условие Якоби и допустимая экстремаль существует, то она доставляет абсолютный минимум в задаче;

c) если выполнено условие Якоби, но не выполнено усиленное условие Якоби и допустимая экстремаль не существует, то $S_{\text{absmin}} = -\infty$*;*

d) если не выполнено условие Якоби, то $S_{\text{absmin}} = -\infty$*. Если при этом существует допустимая экстремаль* $\widehat{x}$*, то* $\widehat{x} \notin \operatorname{wlocmin} P_1$.

Доказательство. a) Существование допустимой экстремали. Пусть выполнено усиленное условие Якоби. Тогда по Утверждению 1 п. 1.3 для любых $h_0, h_1 \in \mathbb{R}$ существует и единственно решение уравнения Якоби h, для которого $h(t_0) = h_0$, $h(t_1) = h_1$. Значения h_0 и h_1 подберём позже.

Возьмём произвольное решение $\widetilde{x}$ уравнения Эйлера, например с заданными граничными условиями в форме Коши $\widetilde{x}(t_0) = 0$, $\dot{\widetilde{x}}(t_0) = 1$.

Подставим функцию $x = \widetilde{x} + h$ в уравнение Эйлера:

$$-\frac{d}{dt}(A\dot{x} + Cx) + C\dot{x} + Bx + D = 0.$$

В силу линейности однородной части уравнение перепишется в виде:

$$-\frac{d}{dt}(A\dot{\widetilde{x}} + C\widetilde{x}) + C\dot{\widetilde{x}} + B\widetilde{x} + D - \frac{d}{dt}(A\dot{h} + Ch) + C\dot{h} + Bh = 0.$$

Уравнение Эйлера для функции $x = \widetilde{x} + h$ выполняется, так как $\widetilde{x}$ удовлетворяет уравнению Эйлера, а h удовлетворяет уравнению Якоби. Тогда $x(t_0) = \widetilde{x}(t_0) + h_0 = h_0 = x_0$ при $h_0 := x_0$, $x(t_1) = \widetilde{x}(t_1) + h_1 = x_1$ при $h_1 := x_1 - \widetilde{x}(t_1)$.

Единственность. Если существует другая допустимая экстремаль $\overline{x} \in DE(P')$, $\overline{x} \neq \widehat{x}$, то $h = \overline{x} - \widehat{x} \not\equiv 0$ тоже является решением уравнения Якоби с условиями $h(t_0) = h(t_1) = 0$, что противоречит выполнению усиленного условия Якоби.

Докажем, что $\widehat{x} \in \operatorname{absmin} P'$. Возьмём произвольную допустимую функцию x. Тогда функция $h = x - \widehat{x}$ ($\Leftrightarrow x = \widehat{x} + h$) принадлежит пространству $C_0^1([t_0;\, t_1])$.
Для квадратичного функционала в задаче на экстремали $\widehat{x}$

$$J(x) - J(\widehat{x}) = \frac{1}{2} J''[h, h].$$

По Лемме 2 при выполнении усиленного условия Лежандра и условия Якоби

$$J''[h, h] = K(h) = \int\limits_{t_0}^{t_1} A\left(\dot{h} - \frac{\dot{u}}{u} h\right)^2 dt \geq 0,$$

где $u(\cdot)$ — любое решение уравнения Якоби с начальным условием $u(t_0) = 0$, $\dot{u}(t_0) \neq 0$.

Следовательно, $J(x) - J(\widehat{x}) \geq 0$, и допустимая экстремаль $\widehat{x}$ доставляет в задаче абсолютный минимум.

b) В пункте a) для доказательства того, что $\widehat{x} \in \operatorname{absmin} P'$ мы пользовались только существованием допустимой экстремали и выполнением условия Якоби. Поэтому существование допустимой экстремали и выполнение условия Якоби достаточно для того, чтобы допустимая экстремаль $\widehat{x}$ доставляла в задаче абсолютный минимум.

c) Пусть выполнено условие Якоби, но не выполнено усиленное условие Якоби. Это означает, что существует решение уравнения Якоби h, для которого $h(t_0) = h(t_1) = 0$. Отметим, что функция h на интервале (t_0, t_1) больше в ноль не обращается. Кроме того, по Утверждению 3 $J''[h,h] = 0$.
Возьмём произвольную допустимую функцию $x \in C^1([t_0;\, t_1])$, $x(t_0) = x_0$, $x(t_1) = x_1$, например, линейную функцию. Тогда функция $x + h$ будет также допустимой.

Следовательно, для квадратично-линейного функционала $J(x) = \int\limits_{t_0}^{t_1} \big(A\dot{x}^2 + 2C\dot{x}x + Bx^2 + 2Dx\big)dt$ по формуле Тейлора

$$J(x+h) - J(x) = J'(x)[h] + \frac{1}{2}J''[h,h] \overset{J''[h,h]=0}{=}$$

$$= J'(x)[h] = 2\int\limits_{t_0}^{t_1} \big(A\dot{x}\dot{h} + C\dot{x}h + Cx\dot{h} + Bxh + Dh\big)dt =$$

$$= 2\int\limits_{t_0}^{t_1} \big((A\dot{h} + Ch)\dot{x} + (C\dot{h} + Bh)x\big)dt + 2\int\limits_{t_0}^{t_1} Dh\,dt =$$

$$= 2\int\limits_{t_0}^{t_1} \big(A\dot{h} + Ch\big)dx + 2\int\limits_{t_0}^{t_1} \big(C\dot{h} + Bh\big)x\,dt + 2\int\limits_{t_0}^{t_1} Dh\,dt =$$

$$= 2(A\dot{h}+Ch)x\Big|_{t_0}^{t_1} + 2\int\limits_{t_0}^{t_1}\Big(-\frac{d}{dt}\big(A\dot{h}+Ch\big)+C\dot{h}+Bh\Big)x\,dt + 2\int\limits_{t_0}^{t_1} Dh\,dt =$$

$$= 2\Big(A(t_1)\dot{h}(t_1)x_1 - A(t_0)\dot{h}(t_0)x_0 + \int\limits_{t_0}^{t_1} Dh\,dt\Big) =: 2G.$$

После интегрирования по частям интеграл равен нулю, поскольку выражение в скобках под знаком интеграла обращается в ноль, так как функция h удовлетворяет уравнению Эйлера. При переходе использовали также, что $h(t_1) = h(t_0) = 0$, $x(t_1) = x_1$, $x(t_0) = x_0$.

Следовательно, $J(x + \lambda h) = J(x) + 2\lambda G$. По условию допустимая экстремаль не существует. Тогда по Теореме 1.8

$G \neq 0$. Поэтому $J(x+\lambda h) \to -\infty$ при $\lambda \to +\infty$ или $-\infty$, т. е. $S_{\text{absmin}} = -\infty$.

d) Предположим, что не выполнено условие Якоби. Так как условие Якоби является необходимым условием слабого локального минимума, то $\widehat{x} \notin \operatorname{wlocmin} P'$.

Докажем, что при этом $S_{\text{absmin}} = -\infty$.
Функция $\overline{h} \equiv 0$ не доставляет абсолютный минимум в задаче

$$J''[h,h] = 2\int\limits_{t_0}^{t_1} \big(A\dot{h}^2 + 2C\dot{h}h + Bh^2\big)dt \to \min; \quad \begin{matrix} h(t_0) = 0, \\ h(t_1) = 0, \end{matrix} \qquad (P'')$$

(если $\overline{h} \equiv 0 \in \operatorname{absmin} P''$, то по теореме о необходимых условиях слабого минимума выполнено условие Якоби). Значит, $S_{\operatorname{absmin} P''} < 0$. Поэтому существует функция $h \in C_0^1([t_0;\, t_1])$ такая, что $J''[h,h] < 0$. Возьмём произвольную допустимую в задаче (P') функцию x, например, линейную функцию. Тогда $x + \lambda h$ также допустимая функция и по формуле Тейлора

$$J(x+\lambda h) = J(x) + \lambda J'(x)[h] + \frac{1}{2}\lambda^2 J''[h,h] \to -\infty$$

при $\lambda \to +\infty$, т. е. $S_{\operatorname{absmin} P'} = -\infty$. $\square$

Примеры

Пример 1. $J(x(\cdot)) = \int\limits_0^1 \dot{x}^3\,dt \to \min;\ x(0) = 0,\ x(1) = 1.$

Исследуем с помощью условий 2-го порядка задачу, рассмотренную нами в п. 1.2. Мы выяснили ранее, что имеется единственная допустимая экстремаль $\widehat{x} = t$, доставляющая слабый локальный минимум в задаче и не доставляющая сильного. При этом нами была построена последовательность допустимых (в задаче на сильный экстремум) функций $x_n(\cdot) \in PC^1([0;\,1])$, $x_n(\cdot) \to \widehat{x}(\cdot)$ в метрике пространства $C([0;\,1])$, для которой $J(x_n(\cdot)) \to -\infty$ при $n \to \infty$.

Поскольку $\widehat{L}_{\dot{x}\dot{x}}(t) = 6\dot{\widehat{x}}(t) = 6 > 0\ \forall\ t \in [0;\,1]$, то выполняется усиленное условие Лежандра на минимум.

Выпишем уравнение Якоби, которое является уравнением Эйлера по h для интегранта $\widetilde{L} = \widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{\dot{x}x}\dot{h}h + \widehat{L}_{xx}h^2 = 6\dot{h}^2$:

$$-\frac{d}{dt}\widetilde{L}_{\dot{h}}(t) + \widetilde{L}_h(t) = 0 \iff -\frac{d}{dt}12\dot{h} = 0 \iff \ddot{h} = 0.$$

Общее решение уравнения Якоби: $h = C_1 t + C_2$. Начальным условиям $h(0) = 0$, $\dot{h}(0) = 1$, удовлетворяет функция $\widehat{h}(t) = t$. Эта функция не имеет нулей в полуинтервале $(0;\,1]$. Значит, сопряженных точек нет, и стало быть выполнено усиленное условие Якоби. По Теореме 1.7 выполнено достаточное условие слабого локального минимума, значит, $\widehat{x} \in \mathrm{wlocmin}$.

Функция $L = \dot{x}^3$ не выпукла по $\dot{x}$, значит, достаточное условие сильного минимума не выполняется. Проверим необходимое условие — условие Вейерштрасса:

$$\begin{aligned} E(t, \widehat{x}, \dot{\widehat{x}}, u) &= L(t, \widehat{x}, u) - L(t, \widehat{x}, \dot{\widehat{x}}) - \widehat{L}_{\dot{x}}(t)(u - \dot{\widehat{x}}) = \\ &= u^3 - \dot{\widehat{x}}^3 - 3\dot{\widehat{x}}^2(u - \dot{\widehat{x}}) = \\ &= u^3 - 1 - 3(u-1) \geq 0\ (?) \quad \forall\ u \in \mathbb{R},\ \forall\ t \in [0;\,1]. \end{aligned}$$

Оно не выполняется. Так как не выполняется необходимое условие, то функция $\widehat{x} \notin \mathrm{strlocmin}$.

Ответ. $\widehat{x} = t \in \text{wlocmin}$, $\widehat{x} \notin \text{strlocmin}$, $S_{\text{absmin}} = -\infty$.

Пример 2. $J(x(\cdot)) = \int\limits_0^{\frac{3\pi}{2}} (\dot{x}^2 - x^2)\,dt \to \min;\ x(0) = x\Big(\dfrac{3\pi}{2}\Big) = 0.$

Исследуем с помощью условий 2-го порядка задачу, рассмотренную нами в Главе 3 п. 1.6 (пример 2). Уравнение Эйлера: $\ddot{x} + x = 0$.

Мы выяснили ранее, что в задаче имеется единственная допустимая экстремаль $\widehat{x} = 0$, не доставляющая даже слабого локального минимума в задаче. При этом нами была построена последовательность допустимых функций $x_\lambda(t) = \lambda \sin \frac{2t}{3}$, $x_\lambda(\cdot) \to \widehat{x}(\cdot)$ в метрике пространства $C^1([0;\,1])$ при $\lambda \to 0$, для которых $J(x_\lambda(\cdot)) < 0 = J(\widehat{x}(\cdot))$ и $S_{\text{absmin}} = -\infty$.

Поскольку $\widehat{L}_{\dot{x}\dot{x}}(t) = 2 > 0\ \forall\ t \in \big[0;\,\frac{3\pi}{2}\big]$, то выполняется усиленное условие Лежандра на минимум.

Выпишем уравнение Якоби, которое является уравнением Эйлера по h для интегранта
$\widetilde{L} = \widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{\dot{x}x}\dot{h}h + \widehat{L}_{xx}h^2 = 2\dot{h}^2 - 2h^2$:

$$-\frac{d}{dt}\widetilde{L}_{\dot{h}}(t) + \widetilde{L}_h(t) = 0 \iff \ddot{h} + h = 0 \iff h = C_1 \sin t + C_2 \cos t.$$

Начальным условиям $h(0) = 0$, $\dot{h}(0) = 1$, удовлетворяет функция $\widehat{h}(t) = \sin t$. Эта функция в интервале $\big(0;\,\frac{3\pi}{2}\big)$ обращается в ноль в точке $\tau = \pi$. Таким образом, в интервале $\big(0;\,\frac{3\pi}{2}\big)$ имеется сопряженная точка, и стало быть не выполнено необходимое условие Якоби слабого локального минимума, значит, допустимая экстремаль $\widehat{x}$ не доставляет в задаче слабый минимум, и тем более не доставляет сильный минимум.

По теореме о необходимых и достаточных условиях экстремума в простейшей задаче вариационного исчисления с квадратичным функционалом из того, что не выполнено условие Якоби, будет следовать, что абсолютный минимум в задаче равен $-\infty$.

Ответ. Единственная допустимая экстремаль $\widehat{x} = 0 \notin \text{wlocmin}$ и, тем более, $\widehat{x} \notin \text{strlocmin}$, $S_{\text{absmin}} = -\infty$.

Пример 3 (*простейшая векторная задача ВИ, в которой допустимая экстремаль единственна и доставляет сильный экстремум*).

$$J(x(\cdot)) = J(x_1(\cdot), x_2(\cdot)) = \int\limits_0^1 (\dot{x}_1^2 + \dot{x}_2^2 + 2x_1x_2)\,dt \to \min;$$

$$x_1(0) = 0,\ x_2(0) = 0,\ x_1(1) = \sin 1,\ x_2(1) = -\sin 1.$$

Необходимое условие 1-го порядка экстремума — система уравнений Эйлера:

$$\begin{cases} -\frac{d}{dt}L_{\dot{x}_1} + L_{x_1} = 0, \\ -\frac{d}{dt}L_{\dot{x}_2} + L_{x_2} = 0, \end{cases} \iff \begin{cases} \ddot{x}_1 = x_2, \\ \ddot{x}_2 = x_1, \end{cases} \Rightarrow x_1^{(4)} = x_1.$$

Последнему дифференциальному уравнению 4-ого порядка соответствует характеристическое уравнение $k^4 = 1$. Его корни $k_{1,2} = \pm 1$, $k_{3,4} = \pm i$. Поэтому общее решение дифференциального уравнения $x_1 = C_1 \operatorname{sh} t + C_2 \operatorname{ch} t + C_3 \sin t + C_4 \cos t$. Тогда $x_2 = \ddot{x}_1 = C_1 \operatorname{sh} t + C_2 \operatorname{ch} t - C_3 \sin t - C_4 \cos t$. Из начальных условий находим неизвестные константы C_1, C_2, C_3, C_4:

$$\begin{cases} x_1(0) = C_2 + C_4 = 0, \\ x_2(0) = C_2 - C_4 = 0, \end{cases} \Rightarrow \begin{cases} C_2 = 0, \\ C_4 = 0, \end{cases} \Rightarrow \begin{cases} x_1 = C_1 \operatorname{sh} t + C_3 \sin t, \\ x_2 = C_1 \operatorname{sh} t - C_3 \sin t, \end{cases}$$

$$\begin{cases} x_1(1) = C_1 \operatorname{sh} 1 + C_3 \sin 1 = \sin 1, \\ x_2(1) = C_1 \operatorname{sh} 1 - C_3 \sin 1 = -\sin 1, \end{cases} \Rightarrow \begin{cases} C_1 = 0, \\ C_3 = 1. \end{cases}$$

Единственная допустимая экстремаль $\widehat{x}_1 = \sin t$, $\widehat{x}_2 = -\sin t$.

Условия экстремума 2-го порядка.

Условие Лежандра. $\widehat{L}_{\dot{x}\dot{x}} = \begin{pmatrix} \widehat{L}_{\dot{x}_1\dot{x}_1} & \widehat{L}_{\dot{x}_1\dot{x}_2} \\ \widehat{L}_{\dot{x}_2\dot{x}_1} & \widehat{L}_{\dot{x}_2\dot{x}_2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$

Эта матрица положительно определена, следовательно, выполняется усиленное условие Лежандра на минимум.

Условие Якоби.

$$\widehat{L}_{\dot{x}x} = \begin{pmatrix} \widehat{L}_{\dot{x}_1x_1} & \widehat{L}_{\dot{x}_1x_2} \\ \widehat{L}_{\dot{x}_2x_1} & \widehat{L}_{\dot{x}_2x_2} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \widehat{L}_{x_1\dot{x}_1} & \widehat{L}_{x_1\dot{x}_2} \\ \widehat{L}_{x_2\dot{x}_1} & \widehat{L}_{x_2\dot{x}_2} \end{pmatrix} = \widehat{L}_{x\dot{x}},$$

$$\widehat{L}_{xx} = \begin{pmatrix} \widehat{L}_{x_1x_1} & \widehat{L}_{x_1x_2} \\ \widehat{L}_{x_2x_1} & \widehat{L}_{x_2x_2} \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}.$$

Квадратичный интегрант

$$\widetilde{L}(t,h,\dot{h}) = \langle \widehat{L}_{\dot{x}\dot{x}}\dot{h}, \dot{h}\rangle + 2\langle \widehat{L}_{x\dot{x}}\dot{h}, h\rangle + \langle \widehat{L}_{xx}h, h\rangle =$$

$$\left\langle \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} \dot{h}_1 \\ \dot{h}_2 \end{pmatrix}, \begin{pmatrix} \dot{h}_1 \\ \dot{h}_2 \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\rangle =$$

$$= 2\dot{h}_1^2 + 2\dot{h}_2^2 + 4h_1h_2 = 2L(t,h,\dot{h}).$$

Отметим, что для квадратичного функционала всегда $\widetilde{L} = 2L$.

Система уравнений Якоби (система уравнений Эйлера для квадратичного интегранта $\widetilde{L}$):

$$\begin{cases} \ddot{h}_1 = h_2, \\ \ddot{h}_2 = h_1, \end{cases} \Rightarrow h_1^{(4)} = h_1.$$

Найдём фундаментальную систему решений уравнения Якоби со столбцами — решениями уравнения Якоби — матрицу $H(t) = \left(h^1(t)\, h^2(t)\right) = \begin{pmatrix} h_1^1(t) & h_1^2(t) \\ h_2^1(t) & h_2^2(t) \end{pmatrix}$ такую, что $H(0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ — нулевая матрица, $\dot{H}(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ — единичная матрица. Для векторов

$$h^1(t) = \begin{pmatrix} h_1^1(t) \\ h_2^1(t) \end{pmatrix} = \begin{pmatrix} C_1 \operatorname{sh} t + C_2 \operatorname{ch} t + C_3 \sin t + C_4 \cos t \\ C_1 \operatorname{sh} t + C_2 \operatorname{ch} t - C_3 \sin t - C_4 \cos t \end{pmatrix},$$

$$\dot{h}^1(t) = \begin{pmatrix} C_1 \operatorname{ch} t + C_2 \operatorname{sh} t + C_3 \cos t - C_4 \sin t \\ C_1 \operatorname{ch} t + C_2 \operatorname{sh} t - C_3 \cos t + C_4 \sin t \end{pmatrix},$$

должны выполняться граничные условия

$$\begin{cases} h_1^1(0) = C_2 + C_4 = 0, \\ h_2^1(0) = C_2 - C_4 = 0, \\ \dot{h}_1^1(0) = C_1 + C_3 = 1, \\ \dot{h}_2^1(0) = C_1 - C_3 = 0, \end{cases} \Leftrightarrow \begin{cases} C_1 = \frac{1}{2}, \\ C_2 = 0, \\ C_3 = \frac{1}{2}, \\ C_4 = 0. \end{cases}$$

Отсюда $h^1(t) = \begin{pmatrix} \frac{\operatorname{sh} t + \sin t}{2} \\ \frac{\operatorname{sh} t - \sin t}{2} \end{pmatrix}$. Аналогично, $h^2(t) = \begin{pmatrix} \frac{\operatorname{sh} t - \sin t}{2} \\ \frac{\operatorname{sh} t + \sin t}{2} \end{pmatrix}$.

Сопряженные точки являются корнями уравнения

$$\det H(\tau) = 0 \iff \det \begin{pmatrix} \frac{\operatorname{sh}\tau + \sin\tau}{2} & \frac{\operatorname{sh}\tau - \sin\tau}{2} \\ \frac{\operatorname{sh}\tau - \sin\tau}{2} & \frac{\operatorname{sh}\tau + \sin\tau}{2} \end{pmatrix} = 0 \iff$$

$$(\operatorname{sh}\tau + \sin\tau)^2 - (\operatorname{sh}\tau - \sin\tau)^2 = 0 \iff \operatorname{sh}\tau \sin\tau = 0.$$

Отсюда $\tau = k\pi$, $k \in \mathbb{N}$. Первая точка сопряженная с нулём есть π. На полуинтервале $(0;\, 1]$ нет сопряженных точек, следовательно, выполняется усиленное условие Якоби. По Теореме 1.7 п. 1.5.7 вектор $\widehat{x} = (\widehat{x}_1, \widehat{x}_2) = (\sin t, -\sin t) \in \mathrm{wlocmin}$.

Условие Вейерштрасса — необходимое условие сильного минимума — выполняется:

$$\mathcal{E}(t, \widehat{x}, \dot{\widehat{x}}, u) \overset{\text{def}}{=} L(t, \widehat{x}, u) - L(t, \widehat{x}, \dot{\widehat{x}}) - \left\langle L_{\dot{x}}(t, \widehat{x}, \dot{\widehat{x}}), u - \dot{\widehat{x}} \right\rangle =$$

$$= u_1^2 + u_2^2 + 2\widehat{x}_1\widehat{x}_2 - \dot{\widehat{x}}_1^2 - \dot{\widehat{x}}_2^2 - 2\widehat{x}_1\widehat{x}_2 - \left\langle \begin{pmatrix} 2\dot{\widehat{x}}_1 \\ 2\dot{\widehat{x}}_2 \end{pmatrix}, \begin{pmatrix} u_1 - \dot{\widehat{x}}_1 \\ u_2 - \dot{\widehat{x}}_2 \end{pmatrix} \right\rangle =$$

$$= u_1^2 + u_2^2 - \dot{\widehat{x}}_1^2 - \dot{\widehat{x}}_2^2 - 2\dot{\widehat{x}}_1(u_1 - \dot{\widehat{x}}_1) - 2\dot{\widehat{x}}_2(u_2 - \dot{\widehat{x}}_2) =$$

$$= (u_1 - \dot{\widehat{x}}_1)^2 + (u_2 - \dot{\widehat{x}}_2)^2 \geq 0 \quad \forall\ (u_1, u_2) \in \mathbb{R}^2,\ \forall\ t \in [0;\, 1].$$

Выпуклость интегранта по $\dot{x}$. Функция $L(t, x, \dot{x}) = \dot{x}_1^2 + \dot{x}_2^2 + 2x_1x_2$ выпукла по $\dot{x}$, так как $\widehat{L}_{\dot{x}\dot{x}} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ — положительно определенная матрица для любых $(t, x) \in \mathbb{R}^3$. По Теореме 1.5 п. 1.5.5 вектор $\widehat{x} = (\widehat{x}_1, \widehat{x}_2) \in \mathrm{strlocmin}$. (Отметим, что из условия выпуклости интегранта L по $\dot{x}$ следует выполнение условия Вейерштрасса.)

Если воспользоваться Теоремой 1.9 п. а) для квадратичных функционалов, то можно было бы, проверив усиленные условия Лежандра и Якоби, сразу сказать, что $\widehat{x} = (\widehat{x}_1, \widehat{x}_2) \in \mathrm{absmin}$ (и слабый, и сильный).

$$S_{\mathrm{absmin}} = J(\widehat{x}) = \int\limits_0^1 \left(2\cos^2 t - 2\sin^2 t\right) dt = 2\int\limits_0^1 \cos 2t\, dt = \sin 2.$$

Ответ. Единственная допустимая экстремаль $(\sin t, -\sin t) \in \mathrm{absmin}$, $S_{\mathrm{absmin}} = \sin 2$.

2 Задача Больца

Рассмотрим задачу Больца (для определенности задачу на минимум) для вектор-функций $x(\cdot) = (x_1(\cdot), \ldots, x_n(\cdot))$

$$B(x(\cdot)) = \int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt + l(x(t_0), x(t_1)) \to \min . \qquad (P)$$

2.1 Сильный и слабый экстремум

Дадим определения слабого и сильного экстремума для задачи Больца.

Определение 6. Говорим, что допустимая функция $\widehat{x}(\cdot) \in C^1([t_0;\, t_1], \mathbb{R}^n)$ доставляет *слабый локальный минимум* в задаче (P), и пишем $\widehat{x}(\cdot) \in \operatorname{wlocmin} P$, если существует $\delta > 0$ такое, что $B(x(\cdot)) \geq B(\widehat{x}(\cdot))$ для любой допустимой функции $x(\cdot)$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^1([t_0;\, t_1], \mathbb{R}^n)} < \delta$.

Определение 7. Говорим, что допустимая функция $\widehat{x}(\cdot) \in PC^1([t_0;\, t_1], \mathbb{R}^n)$ доставляет *сильный локальный минимум* в задаче (P), и пишем $\widehat{x}(\cdot) \in \operatorname{strlocmin} P$, если существует $\delta > 0$ такое, что $B(x(\cdot)) \geq B(\widehat{x}(\cdot))$ для любой допустимой функции $x(\cdot)$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C([t_0;\, t_1], \mathbb{R}^n)} < \delta$.

Так как множество функций, среди которых доставляется сильный экстремум, шире, чем для слабого экстремума, то если функция $\widehat{x} \in C^1([t_0;\, t_1], \mathbb{R}^n)$ доставляет сильный, то она доставляет и слабый экстремум. Поэтому для функций $\widehat{x} \in C^1([t_0;\, t_1], \mathbb{R}^n)$ необходимое условие слабого экстремума является необходимым условием сильного, а достаточное условие сильного экстремума является достаточным условием слабого.

$$\text{ДУ} \implies \widehat{x} \in \operatorname{strlocmin} P \implies \widehat{x} \in \operatorname{wlocmin} P \implies \text{НУ}$$

В этом параграфе будут сформулированы и доказаны необходимые и достаточные условия слабого и сильного экстремума в задаче Больца, а также рассмотрена задача Больца с квадратичным функционалом. Доказательства условий 2-го порядка в задаче Больца во многом аналогичны доказательствам условий 2-го порядка в простейшей задаче классического вариационного исчисления. Поэтому они будут проведены более кратко.

2.2 Условия экстремума 2-го порядка

2.2.1 Необходимые условия слабого экстремума.

Теорема 2.1. *Пусть функция* $\widehat{x} \in C^2([t_0;\, t_1], \mathbb{R}^n)$ *доставляет слабый локальный минимум в задаче* (P) $(\widehat{x} \in \operatorname{wlocmin} P)$, *интегрант* L *трижды непрерывно дифференцируем в некоторой окрестности расширенного графика* $\Gamma_{\widehat{x}\dot{\widehat{x}}}$, *терминант* l *дважды непрерывно дифференцируем в точке* $(\widehat{x}(t_0), \widehat{x}(t_1))$. *Тогда на* $\widehat{x}$ *выполняются*

a) уравнение Эйлера:

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \quad \forall\ t \in [t_0;\, t_1]$$

и условия трансверсальности:

$$\widehat{L}_{\dot{x}}(t_0) = \widehat{l}_{x(t_0)},\ \widehat{L}_{\dot{x}}(t_1) = -\widehat{l}_{x(t_1)};$$

b) условие Лежандра: $\widehat{L}_{\dot{x}\dot{x}}(t) \geq 0\ \forall\ t \in [t_0;\, t_1]$;

c) если на $\widehat{x}$ *выполнено усиленное условие Лежандра, то выполняется и условие Якоби*;

d) если выполняются усиленное условие Лежандра и усиленное условие Якоби, то для любого вектора $(h_0, h_1) \in \mathbb{R}^{2n}$ *решение уравнения Якоби* h *с краевыми условиями* $h(t_0) = h_0$, $h(t_1) = h_1$, *существует, единственно и* $(P + \widehat{l}'')(h) \geq 0$, *где*

$$P(h) := \langle \widehat{L}_{\dot{x}\dot{x}}\dot{h} + \widehat{L}_{\dot{x}x}h, h\rangle\Big|_{t_0}^{t_1},$$

$$\widehat{l}''(h) := \widehat{l}''(h_0, h_1) := l''(\widehat{x}(t_0), \widehat{x}(t_1))[(h_0, h_1), (h_0, h_1)].$$

Доказательство. a) Уравнение Эйлера и условия трансверсальности в задаче Больца были доказаны нами ранее в Главе 3 п. 2.2.

b–c) Обозначим $\widehat{x}(t_0) = \widehat{x}_0$, $\widehat{x}(t_1) = \widehat{x}_1$. Поскольку функция $\widehat{x}$ доставляет локальный минимум функционалу

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

в пространстве $C^1([t_0;\, t_1])$, то она доставляет локальный минимум функционалу B и в более узком пространстве $C^1([t_0;\, t_1])$, $x(t_0) = \widehat{x}_0$, $x(t_1) = \widehat{x}_1$. Это означает, что функция $\widehat{x}$ доставляет слабый локальный минимум в простейшей задаче вариационного исчисления:

$$\int\limits_{t_0}^{t_1} L(t, x, \dot{x})\, dt \to \min; \qquad x(t_0) = \widehat{x}_0,\ x(t_1) = \widehat{x}_1,$$

и, следовательно, в соответствии с Теоремой 1.3 из п. 1.5.3 выполнены условия Лежандра и Якоби.

d) Если выполнены усиленные условия Лежандра и Якоби, то по Утверждению 2 п. 1.4 для любых $h_0, h_1 \in \mathbb{R}^n$ существует и единственно решение уравнения Якоби $h(\cdot)$, для которого $h(t_0) = h_0$, $h(t_1) = h_1$.

Поскольку $\widehat{x} \in \operatorname{wlocmin} P$, то по необходимому условию 2-го порядка локального минимума B

$$B''(\widehat{x})[h, h] = \int\limits_{t_0}^{t_1} \Big(\langle \widehat{L}_{\dot{x}\dot{x}} \dot{h}, \dot{h} \rangle + 2 \langle \widehat{L}_{x\dot{x}} \dot{h}, h \rangle + \langle \widehat{L}_{xx} h, h \rangle \Big)\, dt +$$

$$+ l''(\widehat{x}_0, \widehat{x}_1)[(h_0, h_1), (h_0, h_1)] = K(h) + \widehat{l}''(h) \geq 0$$

$$\forall\ h \in\ C^1([t_0;\, t_1], \mathbb{R}^n)\ (h_0 := h(t_0),\ h_1 := h(t_1)). \quad (*)$$

Пусть далее $h(\cdot)$ — решение уравнения Якоби с краевыми условиями $h(t_0) = h_0$, $h(t_1) = h_1$. Тогда

$$K(h) = \int\limits_{t_0}^{t_1} \Big(\langle \widehat{L}_{\dot{x}\dot{x}} \dot{h}, \dot{h} \rangle + 2 \langle \widehat{L}_{x\dot{x}} \dot{h}, h \rangle + \langle \widehat{L}_{xx} h, h \rangle \Big)\, dt =$$

$$= \int\limits_{t_0}^{t_1} \Big(\langle \widehat{L}_{\dot{x}\dot{x}} \dot{h}, \dot{h} \rangle + \langle \widehat{L}_{\dot{x}x} h, \dot{h} \rangle + \langle \widehat{L}_{x\dot{x}} \dot{h}, h \rangle + \langle \widehat{L}_{xx} h, h \rangle \Big) dt =$$

$$= \int\limits_{t_0}^{t_1} \langle \widehat{L}_{\dot{x}\dot{x}} \dot{h} + \widehat{L}_{\dot{x}x} h, dh \rangle + \int\limits_{t_0}^{t_1} \langle \widehat{L}_{x\dot{x}} \dot{h} + \widehat{L}_{xx} h, h \rangle dt =$$

$$= \int\limits_{t_0}^{t_1} \Big\langle -\frac{d}{dt} \Big(\widehat{L}_{\dot{x}\dot{x}} \dot{h} + \widehat{L}_{\dot{x}x} h \Big) + \widehat{L}_{x\dot{x}} \dot{h} + \widehat{L}_{xx} h, h \Big\rangle dt +$$

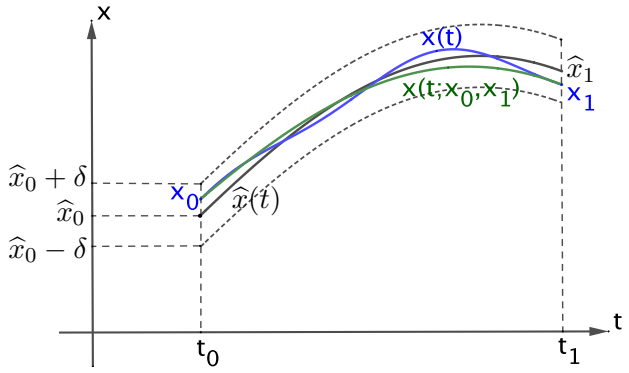
$$+ \langle \widehat{L}_{\dot{x}\dot{x}} \dot{h} + \widehat{L}_{\dot{x}x} h, h \rangle \Big|_{t_0}^{t_1} = P(h)$$

(в последнем интеграле подынтегральное выражение равно нулю, так как функция h удовлетворяет уравнению Якоби). Подставляя в неравенство $(*)$ вместо $K(h)$ равное ему $P(h)$, получим, что $(P + \widehat{l}'')(h) \geq 0$. $\square$

2.2.2 Достаточные условия сильного экстремума.

Теорема 2.2. *Пусть функция* $\widehat{x} \in C^2([t_0;\, t_1], \mathbb{R}^n)$ *— допустимая экстремаль в задаче* (P)*, интегрант* $L \in C^3(V \times \mathbb{R}^n)$*, где* $V \subset \mathbb{R}^{n+1}$ *— некоторая окрестность графика* $\Gamma_{\widehat{x}}$*, интегрант* L *является выпуклым по* $\dot{x}$ *на* V*, терминант* l *дважды непрерывно дифференцируем в точке* $(\widehat{x}(t_0), \widehat{x}(t_1))$*, на* $\widehat{x}$ *выполнены усиленные условия Лежандра и Якоби, квадратичная форма* $P + \widehat{l}''$ (см. п. 2.2.1) *положительно определена. Тогда* $\widehat{x} \in \operatorname{strlocmin} P$.

Доказательство. Для простоты изложения проведем доказательство для одномерных функций $x = x(t)$. Для вектор-функций доказательство аналогично. В некоторой δ-окрестности заданной экстремали $\widehat{x}(\cdot)$ возьмем произвольную допустимую функцию $x(\cdot)$ в задаче на сильный экстремум $(x(\cdot) \in PC^1([t_0;\, t_1]))$, т. е. $\|x(\cdot) - \widehat{x}(\cdot)\|_{C([t_0;\, t_1])} < \delta$. Величину δ выберем позже. Обозначим $x_0 := x(t_0)$, $x_1 := x(t_1)$, $\widehat{x}_0 := \widehat{x}(t_0)$, $\widehat{x}_1 := \widehat{x}(t_1)$. Тогда $|x_0 - \widehat{x}_0| < \delta$, $|x_1 - \widehat{x}_1| < \delta$.



Допустимая экстремаль $\widehat{x}(\cdot)$ удовлетворяет уравнению Эйлера с граничными условиями $\widehat{x}(t_0) = \widehat{x}_0$, $\widehat{x}(t_1) = \widehat{x}_1$. По теореме существования, единственности и непрерывной зависимости решения от начальных данных для любого $\varepsilon > 0$ су-

ществует $\delta = \delta(\varepsilon) > 0$ такое, что для любой точки (x_0, x_1), для которой $|x_0 - \widehat{x}_0| < \delta$, $|x_1 - \widehat{x}_1| < \delta$, существует единственное решение уравнения Эйлера $x(\cdot; x_0, x_1)$ с начальными условиями $x(t_0; x_0, x_1) = x_0$, $x(t_1; x_0, x_1) = x_1$ и $\|x(\cdot; x_0, x_1) - \widehat{x}(\cdot)\|_{C^1([t_0;\, t_1])} < \varepsilon$. При этом не ограничивая общности, будем считать, что $\delta \le \varepsilon$.

По Теореме 1.4 экстремаль $x(\cdot; x_0, x_1)$ можно окружить центральным полем экстремалей, так как в силу гладкости функций усиленные условия Лежандра и Якоби будут выполнены не только на экстремали $\widehat{x}$, но и в некоторой окрестности расширенного графика $\Gamma_{\widehat{x}\dot{\widehat{x}}}$. Имеем

$$B(x) = J(x) + l(x_0, x_1) =$$
$$= J(x) - J(x(\cdot; x_0, x_1)) + \underbrace{J(x(\cdot; x_0, x_1)) + l(x_0, x_1)}_{B(x(\cdot; x_0, x_1))}.$$

Из построения следует, что

$$\|x(\cdot) - x(\cdot; x_0, x_1)\|_{C([t_0;\, t_1])} = \|x(\cdot) - \widehat{x}(\cdot) + \widehat{x}(\cdot) - x(\cdot; x_0, x_1)\|_C \le$$
$$\le \|x(\cdot) - \widehat{x}(\cdot)\|_C + \|\widehat{x}(\cdot) - x(\cdot; x_0, x_1)\|_C < \delta + \varepsilon \le 2\varepsilon.$$

По основной формуле Вейерштрасса для экстремали $x(\cdot; x_0, x_1)$ и функции x близких в метрике пространства $C([t_0;\, t_1])$ в силу выпуклости интегранта L по $\dot{x}$

$$J(x) - J(x(\cdot; x_0, x_1)) = \int\limits_{t_0}^{t_1} \mathcal{E}\, dt \ge 0.$$

Покажем, что функция двух переменных $B(x(\cdot; x_0, x_1)) = B(x_0, x_1)$ имеет локальный минимум в точке $(\widehat{x}_0, \widehat{x}_1)$, т. е., что $B(x(\cdot; x_0, x_1)) \ge B(x(\cdot; \widehat{x}_0, \widehat{x}_1)) = B(\widehat{x})$. Тем самым

$$B(x) = J(x) - J(x(\cdot; x_0, x_1)) + B(x(\cdot; x_0, x_1)) \ge B(\widehat{x})$$

и теорема будет доказана.

Производная функции $B(x_0, x_1)$ согласно выводу необходимых условий экстремума в задаче Больца (см. Гл. 3 п. 2.2)

$$B'(\widehat{x}_0, \widehat{x}_1) = B'(x(\cdot; \widehat{x}_0, \widehat{x}_1)) = B'(\widehat{x}) = 0.$$

При выполнении усиленных условий Лежандра и Якоби по Утверждению 2 п. 1.3 для любых $h_0, h_1 \in \mathbb{R}$ существует и единственно решение уравнения Якоби $h(\cdot)$, для которого $h(t_0) = h_0$, $h(t_1) = h_1$. Тогда

$$B''(\widehat{x}_0, \widehat{x}_1)[h, h] = B''(x(\cdot; \widehat{x}_0, \widehat{x}_1))[h, h] = B''(\widehat{x}(\cdot))[h, h] =$$

$$= \int\limits_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + \widehat{L}_{\dot{x}x}h\dot{h} + \widehat{L}_{x\dot{x}}\dot{h}h + \widehat{L}_{xx}h^2\big)dt + \widehat{l}''[(h_0, h_1), (h_0, h_1)] =$$

$$= \int\limits_{t_0}^{t_1} (\widehat{L}_{\dot{x}\dot{x}}\dot{h} + \widehat{L}_{\dot{x}x}h)dh + \int\limits_{t_0}^{t_1} (\widehat{L}_{x\dot{x}}\dot{h} + \widehat{L}_{xx}h)h\,dt + \widehat{l}''(h_0, h_1) =$$

$$= \int\limits_{t_0}^{t_1} \Big(-\frac{d}{dt}\big(\widehat{L}_{\dot{x}\dot{x}}\dot{h} + \widehat{L}_{\dot{x}x}h\big) + \widehat{L}_{x\dot{x}}\dot{h} + \widehat{L}_{xx}h\Big)h\,dt +$$

$$+ (\widehat{L}_{\dot{x}\dot{x}}\dot{h} + \widehat{L}_{\dot{x}x}h)h\Big|_{t_0}^{t_1} + \widehat{l}''(h_0, h_1) = (P + \widehat{l}'')(h_0, h_1)$$

(функция $h(\cdot)$ удовлетворяет уравнению Якоби, поэтому выражение под знаком интеграла тождественно равно нулю). По условию теоремы квадратичная форма $P + \widehat{l}''$ положительно определена. Значит, и квадратичная форма $B''(\widehat{x}_0, \widehat{x}_1)$ положительно определена. Тогда по достаточному условию минимума функции нескольких переменных точка $(\widehat{x}_0, \widehat{x}_1) \in \operatorname{locmin} B(x_0, x_1)$. Следовательно, если точка (x_0, x_1) лежит в малой окрестности точки $(\widehat{x}_0, \widehat{x}_1)$, то

$$B(x_0, x_1) \geq B(\widehat{x}_0, \widehat{x}_1) = B(\widehat{x}(\cdot)).$$

Таким образом, если $\|x(\cdot) - \widehat{x}(\cdot)\|_{C([t_0;\, t_1])} < \delta$, то $B(x) \geq B(\widehat{x})$ и, значит, функция $\widehat{x} \in \operatorname{strlocmin} P$. $\square$

2.2.3 Квадратичный функционал

Рассмотрим задачу Больца с квадратичным функционалом и линейно-квадратичным терминантом

$$B(x(\cdot)) = \int\limits_{t_0}^{t_1} \big(\underbrace{A\dot{x}^2 + 2C\dot{x}x + Bx^2}_{L(t,x,\dot{x})}\big)dt +$$

$$+ \underbrace{\alpha x_1^2 + 2\gamma x_1 x_0 + \beta x_0^2 + a x_1 + b x_0}_{l(x_0,x_1)} \to \min \quad (P')$$

на слабый и сильный экстремум. Здесь $x_0 = x(t_0)$, $x_1 = x(t_1)$, $A = A(t)$, $B = B(t)$, $C = C(t)$, $\alpha, \beta, \gamma, a, b \in \mathbb{R}$.

Теорема 2.3. *Пусть в задаче (P') функции A и C непрерывно дифференцируемы, а B непрерывна, выполнено усиленное условие Лежандра на минимум. Тогда*

a) если существует допустимая экстремаль $\widehat{x}$, выполнено усиленное условие Якоби, матрица $P + l''$ неотрицательно определена, то $\widehat{x} \in \operatorname{absmin}$;

b) если не выполнено условие Якоби или выполнено усиленное условие Якоби, а матрица $P + l''$ не является неотрицательно определенной, то $S_{\operatorname{absmin}} = -\infty$, если же при этом существует допустимая экстремаль $\widehat{x}$, то $\widehat{x} \notin \operatorname{wlocmin} P'$.

Доказательство. a) Пусть существует допустимая экстремаль $\widehat{x}$, выполнено усиленное условие Якоби, матрица $P + l''$ неотрицательно определена. Докажем, что тогда $\widehat{x} \in \operatorname{absmin}$.

Возьмем произвольную функцию $x \in C^1([t_0;\, t_1])$. Обозначим $x_0 := x(t_0)$, $x_1 := x(t_1)$. Как уже отмечалось, для квадратичного интегранта уравнение Якоби совпадает с уравнением Эйлера. А при выполнении усиленного условия Якоби по Утверждению 1 п. 1.3 для любых $x_0, x_1 \in \mathbb{R}$ существует и единственно решение уравнения Эйлера–Якоби $x(\cdot; x_0, x_1)$, для которого $x(t_0; x_0, x_1) = x_0$, $x(t_1; x_0, x_1) = x_1$.

Представим функционал B в следующем виде:

$$B(x) = J(x) + l(x_0, x_1) =$$
$$= J(x) - J(x(\cdot; x_0, x_1)) + \underbrace{J(x(\cdot; x_0, x_1)) + l(x_0, x_1)}_{B(x(\cdot; x_0, x_1))}.$$

Функция $h(\cdot) := x(\cdot) - x(\cdot; x_0, x_1)$ ($\Leftrightarrow x(\cdot) = x(\cdot; x_0, x_1) + h(\cdot)$) принадлежит пространству $C_0^1([t_0; t_1])$. Для квадратичного функционала по формуле Тейлора на экстремали $x(\cdot; x_0, x_1)$

$$J(x) - J(x(\cdot; x_0, x_1)) = J'(x(\cdot; x_0, x_1))[h] + \frac{1}{2}J''[h, h] = \frac{1}{2}J''[h, h].$$

По Лемме 2 п. 1.5.6 при выполнении усиленного условия Лежандра в задаче на минимум и условия Якоби

$$\frac{1}{2}J''[h, h] = \int\limits_{t_0}^{t_1} (A\dot{h}^2 + 2B\dot{h}h + Ch^2)\, dt = \int\limits_{t_0}^{t_1} A\left(\dot{h} - \frac{\dot{u}}{u}h\right)^2 dt \geq 0,$$

где $u(\cdot)$ — любое решение уравнения Эйлера–Якоби с начальным условием $u(t_0) = 0$, $\dot{u}(t_0) \neq 0$. Следовательно, $J(x) - J(x(\cdot; x_0, x_1)) \geq 0$.

Покажем, что функция двух переменных $B(x(\cdot; x_0, x_1)) = B(x_0, x_1)$ имеет абсолютный минимум в точке $(\widehat{x}_0, \widehat{x}_1)$, т. е., что $B(x(\cdot; x_0, x_1)) \geq B(x(\cdot; \widehat{x}_0, \widehat{x}_1)) = B(\widehat{x})$. Тем самым

$$B(x) = J(x) - J(x(\cdot; x_0, x_1)) + B(x(\cdot; x_0, x_1)) \geq B(\widehat{x})$$

и, значит, $\widehat{x} \in \operatorname{absmin} P'$.

Действительно, производная функции $B(x_0, x_1)$ согласно выводу необходимых условий экстремума в задаче Больца (см. Гл. 3 п. 2.2)

$$B'(\widehat{x}_0, \widehat{x}_1) = B'(x(\cdot; \widehat{x}_0, \widehat{x}_1)) = B'(\widehat{x}) = 0.$$

Действие второй производной на решение уравнения Эйлера–Якоби $h(\cdot; h_0, h_1)$ с граничными условиями $h(t_0; h_0, h_1) = h_0$, $h(t_1; h_0, h_1) = h_1$

$$B''(\widehat{x}_0,\widehat{x}_1)[h,h] = B''(x(\cdot;\widehat{x}_0,\widehat{x}_1))[h,h] =$$

$$= (L_{\dot{x}\dot{x}}\dot{h}+L_{\dot{x}x}h)h\Big|_{t_0}^{t_1}+l''(h_0,h_1) = (P+l'')(h_0,h_1) \geq 0 \;\forall\, h_0,h_1\in\mathbb{R}$$

(по условию матрица $P + l''$ неотрицательно определена). Следовательно, по Теореме 2 Гл. 1 п. 4.1 функция $B(x_0,x_1)$ является выпуклой и её минимум в точке $(\widehat{x}_0,\widehat{x}_1)$ является абсолютным.

b) Предположим теперь, что не выполнено условие Якоби или выполнено усиленное условие Якоби, а матрица $P+l''$ не является неотрицательно определенной. Тогда функция $\overline{h}\equiv 0$ не доставляет абсолютный минимум в задаче

$$\widetilde{B}(h) = \frac{1}{2}B''(h) = \int\limits_{t_0}^{t_1}\big(A\dot{h}^2 + 2C\dot{h}h + Bh^2\big)dt+$$
$$+\alpha h_1^2 + 2\gamma h_1 h_0 + \beta h_0^2 \to \min \quad (P'')$$

(если $\overline{h}\equiv 0 \in \operatorname{absmin} P''$, то по теореме о необходимых условиях слабого минимума выполнено условие Якоби и, если выполнено усиленное условие Якоби, то матрица $P+l''$ является неотрицательно определенной). Значит, $S_{\operatorname{absmin} P''} < 0$. Поэтому существует функция $h\in C^1([t_0;\,t_1])$ такая, что $\widetilde{B}(h) < 0$. Но тогда

$$B(\lambda h) = \widetilde{B}(\lambda h) + \lambda(ah_1+bh_0) = \lambda^2\widetilde{B}(h) + \lambda(ah_1+bh_0) \to -\infty$$

при $\lambda\to+\infty$, т. е. $S_{\operatorname{absmin} P'} = -\infty$.

Если же при этом существует допустимая экстремаль $\widehat{x}$, то $\widehat{x}\notin \operatorname{wlocmin} P'$. Действительно, в этом случае

$$B(\widehat{x}+\lambda h) - B(\widehat{x}) = \frac{\lambda^2}{2}B''(h) = \lambda^2\widetilde{B}(h) < 0$$

при малых λ, т. е. $\widehat{x}\notin \operatorname{wlocmin} P'$. $\square$

2.3 Пример.

$$B(x(\cdot)) = \int\limits_0^{\frac{\pi}{2}} (\dot{x}^2 - x^2)\,dt + x^2(0) - x^2\Big(\frac{\pi}{2}\Big) + 4x\Big(\frac{\pi}{2}\Big) \to \mathrm{extr}\,.$$

Решение. Необходимые условия экстремума 1-го порядка:

a) *уравнение Эйлера*

$$-\frac{d}{dt}L_{\dot{x}} + L_x = 0 \iff -\frac{d}{dt}2\dot{x} - 2x = 0 \iff \ddot{x} + x = 0;$$

b) *условия трансверсальности*

$$\begin{cases} L_{\dot{x}}(0) = l_{x(0)}, \\ L_{\dot{x}}\big(\frac{\pi}{2}\big) = -l_{(\frac{\pi}{2})}, \end{cases} \iff \begin{cases} \dot{x}(0) = x(0), \\ \dot{x}\big(\frac{\pi}{2}\big) = x\big(\frac{\pi}{2}\big) - 2. \end{cases}$$

Общее решение уравнения Эйлера:

$$x = C_1 \sin t + C_2 \cos t \qquad \big(\dot{x} = C_1 \cos t - C_2 \sin t\big).$$

Из условий трансверсальности находим допустимую экстремаль:

$$\begin{cases} C_1 = C_2, \\ -C_2 = C_1 - 2, \end{cases} \iff C_1 = C_2 = 1.$$

Допустимая экстремаль $\widehat{x} = \sin t + \cos t$.

Исследование. Необходимые условия экстремума 2-го порядка.

<u>Условие Лежандра.</u> $\widehat{L}_{\dot{x}\dot{x}} = 2 > 0$ — выполняется усиленное условие Лежандра в задаче на минимум.

<u>Условие Якоби.</u> Уравнение Якоби, является уравнением Эйлера по h для интегранта

$$\widetilde{L} = \widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{\dot{x}x}\dot{h}h + \widehat{L}_{xx}h^2 = 2\dot{h}^2 - 2h^2$$

(в квадратичной задаче совпадает с уравнением Эйлера):

$$-\frac{d}{dt}\widetilde{L}_{\dot h} + \widetilde{L}_h = 0 \iff \ddot h + h = 0 \iff h = C_1 \sin t + C_2 \cos t.$$

Начальным условиям $h(0) = 0$, $\dot h(0) = 1$, удовлетворяет функция $\widehat h(t) = \sin t$. Нули функции h — точки $\tau = k\pi$, $k \in \mathbb{N}$, являются сопряженными точками. На полуинтервале $\left(0;\, \frac{\pi}{2}\right]$ нет сопряженных точек, следовательно, выполняется усиленное условие Якоби.

Проверим на неотрицательность квадратичную форму $P+l''$. Построим решение уравнения Якоби с краевыми условиями

$$\begin{cases} h(0) = h_0, \\ h\left(\frac{\pi}{2}\right) = h_1 \end{cases} \iff \begin{cases} C_2 = h_1, \\ C_1 = h_0, \end{cases} \implies \begin{array}{l} h(t) = h_0 \cos t + h_1 \sin t, \\ (\dot h(t) = -h_0 \sin t + h_1 \cos t). \end{array}$$

Квадратичная форма $P + l''$:

$$P+l'' = P(h_0, h_1)+l''(h_0, h_1) = \left(\widehat L_{\dot x\dot x}\dot h h+\widehat L_{\dot x x}h^2\right)\Big|_{t_0}^{t_1}+2h_0^2-2h_1^2 =$$

$$= 2(-h_0 \sin t + h_1 \cos t)(h_0 \cos t + h_1 \sin t)\Big|_0^{\frac{\pi}{2}} + 2h_0^2 - 2h_1^2 =$$

$$= 2(-h_0)h_1 - 2h_1 h_0 + 2h_0^2 - 2h_1^2 = 2h_0^2 - 4h_0 h_1 - 2h_1^2.$$

В матричном виде

$$P + l'' = \begin{pmatrix} 2 & -2 \\ -2 & -2 \end{pmatrix}$$

Поскольку $\det A_1 = 2 > 0$, $\det A_{12} = \det(P + l'') = -8 < 0$, то по критерию Сильвестра (см. Гл. 1 п. 1.2.5) квадратичная форма матрица $P+l''$ не является неотрицательно определенной.

Не выполняется необходимое условие слабого локального минимума — неотрицательность квадратичной формы $P + l''$. Следовательно, $\widehat x \notin \mathrm{wlocmin}$ и тем более $\widehat x \notin \mathrm{strlocmin}$.

По теореме об условиях экстремума для квадратичного функционала, если выполнено усиленное условие Якоби, но

матрица $P + l''$ не является неотрицательно определенной, то значение задачи равно $-\infty$.

Покажем, что $S_{\text{absmax}} = +\infty$. Возьмем последовательность допустимых функций $x_n(t) = \sin 2nt$. Отметим, что на них $l\big(x_n(0), x_n(\frac{\pi}{2})\big) = l(0,0) = 0$. Поэтому

$$B(x_n) = \int\limits_0^{\frac{\pi}{2}} (\dot{x}_n^2 - x_n^2)\, dt = \int\limits_0^{\frac{\pi}{2}} \big(4n^2 \cos^2 2nt - \sin^2 2nt\big)\, dt =$$

$$= 4n^2 \int\limits_0^{\frac{\pi}{2}} \frac{1+\cos 4nt}{2}\, dt - \int\limits_0^{\frac{\pi}{2}} \frac{1-\cos 4nt}{2}\, dt = \pi n^2 - \frac{\pi}{4} \to +\infty$$

при $n \to +\infty$.

Ответ. Допустимая экстремаль $\widehat{x} = \sin t + \cos t \notin \text{wlocextr}$, $S_{\text{absmin}} = -\infty$; $S_{\text{absmax}} = +\infty$.

3 Изопериметрическая задача

3.1 Постановка задачи

Рассмотрим изопериметрическую задачу (для определенности задачу на минимум) в пространствах функций $C^1([t_0;\, t_1])$ и $PC^1([t_0;\, t_1])$

$$J_0(x) \to \min;\ J_i(x) = \alpha_i,\ i = 1, \ldots, m;\ \begin{array}{l} x(t_0) = x_0, \\ x(t_1) = x_1, \end{array} \qquad (P)$$

где $\alpha_i \in \mathbb{R}$, $J_i(x) := \int\limits_{t_0}^{t_1} f_i(t, x, \dot{x})\, dt,\ i = 0, 1, \ldots, m.$

Дадим определения слабого и сильного экстремума для изопериметрической задачи.

Определение 8. Говорим, что допустимая функция $\widehat{x} \in C^1([t_0;\, t_1])$ доставляет *слабый локальный минимум* в задаче (P), и пишем $\widehat{x}(\cdot) \in \operatorname{wlocmin} P$, если существует $\delta > 0$ такое, что $J_0(x) \geq J_0(\widehat{x})$ для любой допустимой функции x, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^1([t_0;\, t_1])} < \delta$.

Определение 9. Говорим, что допустимая функция $\widehat{x}(\cdot) \in PC^1([t_0;\, t_1])$ доставляет *сильный локальный минимум* в задаче (P), и пишем $\widehat{x} \in \operatorname{strlocmin} P$, если существует $\delta > 0$ такое, что $J_0(x) \geq J_0(\widehat{x})$ для любой допустимой функции x, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C([t_0;\, t_1])} < \delta$.

Так как множество функций, среди которых доставляется сильный экстремум, шире, чем для слабого экстремума, то если функция $\widehat{x} \in C^1([t_0;\, t_1])$ доставляет сильный, то она доставляет и слабый экстремум. Поэтому для функций $\widehat{x} \in C^1([t_0;\, t_1])$ необходимое условие слабого экстремума является необходимым условием сильного, а достаточное условие сильного экстремума является достаточным условием слабого.

$$\text{ДУ} \implies \widehat{x} \in \operatorname{strlocmin} P \implies \widehat{x} \in \operatorname{wlocmin} P \implies \text{НУ}$$

3.2 Условия Лежандра, Якоби и регулярности

Будем далее предполагать, что все интегранты f_i по меньшей мере дважды непрерывно дифференцируемы в некоторой окрестности расширенного графика $\Gamma_{\widehat{x}\dot{\widehat{x}}}$.
Функция Лагранжа задачи (P):

$$\mathscr{L}(x) = \int\limits_{t_0}^{t_1} L(t, x, \dot{x})\, dt + \nu_0 x(t_0) + \nu_1 x(t_1),$$

где лагранжиан $L(t, x, \dot{x}) = \sum\limits_{i=0}^{m} \lambda_i f_i(t, x, \dot{x})$.

Пусть $\widehat{x}$ дважды непрерывно дифференцируемая допустимая экстремаль задачи (P) с множителем Лагранжа $\lambda_0 = 1$, т. е. на ней выполнено уравнение Эйлера для лагранжиана $L = f_0 + \sum\limits_{i=1}^{m} \lambda_i f_i$:

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

При наших допущениях относительно гладкости интегрантов f_i функция Лагранжа задачи (P) имеет вторую производную по Фреше в точке $\widehat{x}$:

$$\mathscr{L}''(\widehat{x})[h, h] = \int\limits_{t_0}^{t_1} \big(\widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{x\dot{x}}\dot{h}h + \widehat{L}_{xx}h^2\big)dt.$$

Говорим, что на экстремали $\widehat{x}$ выполнено *условие Лежандра*, если $\widehat{L}_{\dot{x}\dot{x}}(t) \geq 0\ \forall\, t \in [t_0;\, t_1]$ и *усиленное условие Лежандра*, если $\widehat{L}_{\dot{x}\dot{x}}(t) > 0\ \forall\ t \in [t_0;\, t_1]$.

Говорим, что на экстремали $\widehat{x}$ в задаче (P) выполнено *условие регулярности*, если на любом ненулевом отрезке $[t_0; \tau]$, $[\tau;\, t_1]$ при любом $\tau \in [t_0;\, t_1]$ функции $g_i := -\frac{d}{dt}\widehat{f}_{i\dot{x}} + \widehat{f}_{ix}$, $i = 1, \ldots, m$, линейно независимы.

Пусть далее на экстремали $\widehat{x}$ выполнено усиленное условие Лежандра и условие регулярности функций g_i, $i = 1, \ldots, m$.

Необходимое условие 2-го порядка в задачах с равенствами (Глава 1, Теорема 7.2) — 2-я производная функции Лагранжа неотрицательна на пространстве стационарных направлений[12]:

$$\mathscr{L}''(\widehat{x})[h,h] \geq 0 \ \forall \ h \in \Pi. \tag{5.12}$$

В задаче (P) пространство стационарных направлений

$$\Pi = \left\{ h \in C^1([t_0;\, t_1]) \ \Big| \ J_i'(\widehat{x})[h] = 0, \ i = 1, \ldots, m, \ \begin{array}{l} h(t_0) = 0, \\ h(t_1) = 0, \end{array} \right\}.$$

Неравенство (5.12) эквивалентно тому, что функция $\overline{h} \equiv 0$ доставляет абсолютный минимум в задаче

$$\mathscr{L}''(\widehat{x})[h,h] \to \min; \ J_i'(\widehat{x})[h] = 0, \ i = 1, \ldots, m; \ \begin{array}{l} h(t_0) = 0, \\ h(t_1) = 0, \end{array} \tag{$\widetilde{P}$}$$

$$J_i'(\widehat{x})[h] = \int\limits_{t_0}^{t_1} \big(\widehat{f}_{i\dot{x}}\dot{h} + \widehat{f}_{ix}h\big)dt = \int\limits_{t_0}^{t_1} \Big(-\frac{d}{dt}\widehat{f}_{i\dot{x}} + \widehat{f}_{ix}\Big)h\,dt = \int\limits_{t_0}^{t_1} g_i h\,dt.$$

Согласно правилу множителей Лагранжа для изопериметрической задачи существуют множители Лагранжа $\mu_0, \ldots, \mu_m$ не все равные нулю и такие, что для лагранжиана задачи $(\widetilde{P})$

$$\widetilde{L} = \widetilde{L}(t,h,\dot{h}) := \mu_0\big(\widehat{L}_{\dot{x}\dot{x}}\dot{h}^2 + 2\widehat{L}_{x\dot{x}}\dot{h}h + \widehat{L}_{xx}h^2\big) + \sum\limits_{i=1}^{m} \mu_i g_i h$$

выполняется уравнение Эйлера.

Уравнение Эйлера для лагранжиана $\widetilde{L}$ с $\mu_0 = \frac{1}{2}$ ($\mu_0 \neq 0$ в силу условия регулярности функций g_i)

$$-\frac{d}{dt}\widetilde{L}_{\dot{h}} + \widetilde{L}_h = 0 \iff$$

$$-\frac{d}{dt}\big(\widehat{L}_{\dot{x}\dot{x}}\dot{h} + \widehat{L}_{\dot{x}x}h\big) + \widehat{L}_{x\dot{x}}\dot{h} + \widehat{L}_{xx}h + \sum\limits_{i=1}^{m} \mu_i g_i = 0$$

[12]Напомним, что в задаче с равенствами $f_0(x) \to \mathrm{extr}$; $f_i(x) = 0$, $i = 1, \ldots, m$, пространство стационарных направлений $\Pi := \{h \in X \mid f_i'(\widehat{x})[h] = 0, \ i = 1, \ldots, m\}$.

называется *уравнением Якоби* для исходной задачи (P) на экстремали $\widehat{x}$. Это неоднородное уравнение в силу наличия свободного члена $\sum\limits_{i=1}^{m} \mu_i g_i$.

Точка τ называется *сопряженной к точке* t_0, если существует нетривиальное решение h неоднородного уравнения Якоби, для которого

$$\int\limits_{t_0}^{\tau} g_i(t)h(t)\,dt = 0,\ i = 1, \ldots, m,\ h(t_0) = h(\tau) = 0.$$

Говорят, что на экстремали $\widehat{x}$ 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.

Дадим аналитический способ нахождения сопряженных точек. Пусть $h_0(\cdot)$ — решение однородной части уравнения Якоби ($\mu_1 = \ldots = \mu_m = 0$) с краевыми условиями $h_0(t_0) = 0$, $\dot{h}_0(t_0) = 1$ (или $\dot{h}_0(t_0) \neq 0$).

Функции $h_i(\cdot)$ — решения неоднородного уравнения Якоби с $\mu_i = 1$ (или $\mu_i \neq 0$), $\mu_j = 0$, $j \neq i$, и краевыми условиями $h_i(t_0) = \dot{h}_i(t_0) = 0$, $i = 1, \ldots, m$.

Покажем, что точка τ является сопряженной к t_0 тогда и только тогда, когда матрица

$$H(\tau) = \begin{pmatrix} h_0(\tau) & \ldots & h_m(\tau) \\ \int\limits_{t_0}^{\tau} g_1 h_0 dt & \ldots & \int\limits_{t_0}^{\tau} g_1 h_m dt \\ \ldots & \ldots & \ldots \\ \int\limits_{t_0}^{\tau} g_m h_0 dt & \ldots & \int\limits_{t_0}^{\tau} g_m h_m dt \end{pmatrix}$$

является вырожденной, т. е. $\det H(\tau) = 0$.

Действительно, условие $\det H(\tau) = 0$ равносильно тому, что столбцы матрицы $H(\tau)$ являются линейно зависимыми.

Это означает, что существуют числа $\alpha_0, \alpha_1, \ldots, \alpha_m$, не все равные нулю, такие что для функции $h = \sum\limits_{k=0}^{m} \alpha_k h_k$ выполняются граничные условия $h(\tau) = \sum\limits_{k=0}^{m} \alpha_k h_k(\tau) = 0$; $\sum\limits_{k=0}^{m} \alpha_k \int\limits_{t_0}^{\tau} g_i h_k \, dt = 0$ $\Leftrightarrow \int\limits_{t_0}^{\tau} g_i \left(\sum\limits_{k=0}^{m} \alpha_k h_k \right) dt = 0 \Leftrightarrow \int\limits_{t_0}^{\tau} g_i h \, dt = 0$, $i = 1, \ldots, m$. При этом $h(t_0) = \sum\limits_{k=0}^{m} \alpha_k h_k(t_0) = 0$. Значит, для функции h в точке τ выполняются все условия определения сопряженной точки и τ — сопряжённая точка.

3.3 Условия экстремума 2-го порядка

3.3.1 Необходимые условия слабого экстремума

Теорема 3.1. *Пусть функция $\widehat{x} \in C^2([t_0; t_1], \mathbb{R})$ доставляет слабый локальный минимум в задаче (P) ($\widehat{x} \in \operatorname{wlocmin} P$), интегранты f_i, $i = 0, 1, \ldots, m$, трижды непрерывно дифференцируемы в некоторой окрестности расширенного графика $\Gamma_{\widehat{x}\dot{\widehat{x}}}$ ($f_i \in C^3(\mathcal{O}(\Gamma_{\widehat{x}\dot{\widehat{x}}}))$), а также на $\widehat{x}$ выполняется условие регулярности. Тогда на $\widehat{x}$ выполняется уравнение Эйлера, условие Лежандра и, если на экстремали выполнено усиленное условие Лежандра, то выполняется и условие Якоби.*

3.3.2 Достаточные условия сильного экстремума

Теорема 3.2. *Пусть функция $\widehat{x} \in C^3([t_0; t_1], \mathbb{R})$ — допустимая экстремаль в задаче (P), интегранты $f_i \in C^4(V \times \mathbb{R})$, где V — некоторая окрестность графика $\Gamma_{\widehat{x}}$, на $\widehat{x}$ выполнены усиленные условия Лежандра и Якоби и условие регулярности, интегрант $L = f_0 + \sum\limits_{i=1}^{m} \lambda_i f_i$ является выпуклым по $\dot{x}$ на V. Тогда $\widehat{x}$ доставляет сильный локальный минимум в задаче (P) ($\widehat{x} \in \operatorname{strlocmin} P$).*

3.3.3 Квадратичный функционал

Рассмотрим следующую изопериметрическую задачу с квадратичным функционалом и линейными ограничениями

$$\int\limits_{t_0}^{t_1} \big(A\dot{x}^2 + 2C\dot{x}x + Bx^2\big)dt \to \min;$$

$$\int\limits_{t_0}^{t_1} (a_i\dot{x} + b_i x)\, dt = \alpha_i,\ i = 1, \ldots, m, \quad \begin{matrix} x(t_0) = x_0, \\ x(t_1) = x_1. \end{matrix} \qquad (P')$$

где $A(\cdot), C(\cdot), a_i(\cdot) \in C^1([t_0;\, t_1])$, $B(\cdot), b_i(\cdot) \in C([t_0;\, t_1])$, $i = 1, \ldots, m$.

Функция Лагранжа задачи (P'):

$$\mathscr{L}(x) = \int\limits_{t_0}^{t_1} \left(\lambda_0\big(A\dot{x}^2 + 2C\dot{x}x + Bx^2\big) + \sum_{i=1}^{m} \lambda_i(a_i\dot{x} + b_i x)\right)dt + \\ + \nu_0 x(t_0) + \nu_1 x(t_1).$$

Первая производная функции Лагранжа

$$\mathscr{L}'(x)[h] = \int\limits_{t_0}^{t_1} \left(2\lambda_0\big(A\dot{x}\dot{h} + Cx\dot{h} + C\dot{x}h + Bxh\big) + \sum_{i=1}^{m} \lambda_i(a_i\dot{h} + b_i h)\right)dt + \\ + \nu_0 h(t_0) + \nu_1 h(t_1),$$

вторая производная функции Лагранжа

$$\mathscr{L}''(\widehat{x})[h, h] = \int\limits_{t_0}^{t_1} 2\lambda_0\big(A\dot{h}^2 + 2C\dot{h}h + Bh^2\big)dt = 2\lambda_0 J_0(h).$$

Пространство стационарных направлений

$$\Pi = \left\{ h \in C^1([t_0;\, t_1]) \;\middle|\; J_i'(\widehat{x})[h] = 0,\ i = 1, \ldots, m, \ \begin{matrix} h(t_0) = 0, \\ h(t_1) = 0. \end{matrix} \right\}$$

Условие $J_i'(\widehat{x})[h] = 0$ можно переписать в виде

$$\int\limits_{t_0}^{t_1} \big(a_i\dot{h} + b_i h\big)dt = 0 \iff \int\limits_{t_0}^{t_1} a_i\, dh + \int\limits_{t_0}^{t_1} b_i h\, dt = 0 \iff$$

$$\int\limits_{t_0}^{t_1} (-\dot{a}_i + b_i)h\, dt = 0.$$

Поскольку $\mathscr{L}''(\widehat{x})[h,h] = 2J_0(h)$, то необходимое условие минимума 2-го порядка — неотрицательность 2-й производной функции Лагранжа на пространстве стационарных направлений ($\mathscr{L}''(\widehat{x})[h,h] \geq 0\ \forall\ h \in \Pi$), запишется в виде

$$\mathscr{L}''(\widehat{x})[h,h] \geq 0\ \forall\, h : \int\limits_{t_0}^{t_1} (-\dot{a}_i + b_i)h\, dt = 0,\ i = 1,\ldots,m;\ \begin{array}{l} h(t_0) = 0, \\ h(t_1) = 0. \end{array}$$

Последнее неравенство эквивалентно тому, что функция $\overline{h} \equiv 0$ доставляет абсолютный минимум в задаче запишется в виде

$$\int\limits_{t_0}^{t_1} \lambda_0\big(A\dot{h}^2 + 2C\dot{h}h + Bh^2\big)dt \to \min;$$

$$\int\limits_{t_0}^{t_1} \big(a_i\dot{h} + b_i h\big)dt = 0,\ i = 1,\ldots,m,\ \begin{array}{l} h(t_0) = 0, \\ h(t_1) = 0. \end{array} \qquad (\widetilde{P}')$$

Функция $x = \widehat{x} + h \in C^1([t_0;\, t_1])$ в задаче (P') является допустимой, тогда и только тогда, когда для функции h выполняются условия

$$\int\limits_{t_0}^{t_1} \big(a_i\dot{h} + b_i h\big)dt = 0,\ i = 1,\ldots,m;\ h(t_0) = 0,\ h(t_1) = 0,$$

т. е. $h \in \Pi$ — пространству стационарных направлений.

Необходимое условие 1-го порядка минимума $\mathscr{L}'[h] = 0$ $\forall\ h \in \Pi$ в задаче (P') равносильно существованию вектора

множителей Лагранжа $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, $\lambda \neq 0$, такого, что для лагранжиана $L = \lambda_0(A\dot{x}^2 + 2C\dot{x}x + Bx^2) + \sum\limits_{i=1}^{m} \lambda_i(a_i\dot{x} + b_i x)$ на $\widehat{x}$ выполняется уравнение Эйлера.

Поскольку в функционале J_0 слагаемое

$$\int\limits_{t_0}^{t_1} 2C\dot{x}x\,dt = \int\limits_{t_0}^{t_1} C\,dx^2 = Cx^2\Big|_{t_0}^{t_1} - \int\limits_{t_0}^{t_1} \dot{C}x^2\,dt,$$

то его можно включить в коэффициент при x^2. Поэтому для упрощения записи считаем далее, что $C(t) \equiv 0$.

Теорема 3.3. *Пусть в задаче* (P') *функции* $A, a_1, \ldots, a_m$ *непрерывно дифференцируемы, функции* $B, b_1, \ldots, b_m$ *непрерывны и выполнены усиленное условие Лежандра и условие регулярности. Тогда*

a) если выполнено усиленное условие Якоби, то допустимая экстремаль существует, единственна и доставляет абсолютный минимум;

b) если не выполнено условие Якоби, то $S_{\text{absmin}} = -\infty$.

Доказательство. a) Пусть выполнено усиленное условие Якоби. Докажем, что допустимая экстремаль существует, единственна и доставляет абсолютный минимум.

Выпишем лагранжиан в задаче (P')
$L = \lambda_0(A\dot{x}^2 + Bx^2) + \sum\limits_{i=1}^{m} \lambda_i(a_i\dot{x} + b_i x)$. Уравнение Эйлера

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

$$-\frac{d}{dt}\Big(2\lambda_0 A\dot{x} + \sum_{i=1}^{m} \lambda_i a_i\Big) + 2\lambda_0 Bx + \sum_{i=1}^{m} \lambda_i b_i = 0.$$

Из условия регулярности (на любом ненулевом отрезке $[t_0; \tau]$, $[\tau;\, t_1]$ при любом $\tau \in [t_0;\, t_1]$ функции $g_i(t) := -\frac{d}{dt}\widehat{f}_{i\dot{x}} + \widehat{f}_{ix} = -\dot{a}_i + b_i$, $i = 1, \ldots, m$, линейно независимы) следует, что $\lambda_0 \neq 0$.

Положим $\lambda_0 = \frac{1}{2}$. Тогда уравнение Эйлера перепишется:

$$-\frac{d}{dt}(A\dot{x}) + Bx + \sum_{i=1}^{m} \lambda_i(-\dot{a}_i + b_i) = 0. \qquad (\text{Э})$$

Интегрант квадратичной задачи

$$\widetilde{L} = \mu_0(A\dot{h}^2 + Bh^2) + \sum_{i=1}^{m} \mu_i g_i h = \mu_0(A\dot{h}^2 + Bh^2) + \sum_{i=1}^{m} \mu_i(-\dot{a}_i + b_i)h.$$

Выпишем уравнение Эйлера для лагранжиана $\widetilde{L}$ с $\mu_0 = \frac{1}{2}$, являющееся уравнением Якоби исходной задачи:

$$-\frac{d}{dt}(A\dot{h}) + Bh + \sum_{i=1}^{m} \mu_i(-\dot{a}_i + b_i) = 0. \qquad (\text{Я})$$

Получили, что уравнения Эйлера и Якоби для задачи (P') с квадратичным функционалом и линейными ограничениями с точностью до переобозначения множителей Лагранжа совпадают.

Пусть $\widehat{x}(\cdot)$ — допустимая экстремаль в задаче. Возьмем произвольный допустимый элемент $x(\cdot)$. Тогда функция $h = x - \widehat{x}$ ($\Leftrightarrow x = \widehat{x} + h$) принадлежит пространству стационарных направлений

$$\Pi = \left\{ h \in C_0^1([t_0;\, t_1]) \;\Big|\; \int_{t_0}^{t_1} (a_i\dot{h} + b_i h)dt = 0,\ i = 1, \ldots, m \right\}.$$

Так как для квадратичного функционала J_0 имеет место равенство $J_0(x+h) = J_0(x) + J_0'(x)[h] + \frac{1}{2}J_0''(x)[h,h]$, то для экстремали $\widehat{x}$ для любого $h \in \Pi$

$$\frac{J_0'(\widehat{x})[h]}{2} = \int_{t_0}^{t_1} (A\dot{\widehat{x}}\dot{h} + B\widehat{x}h)dt = \int_{t_0}^{t_1} A\dot{\widehat{x}}\, dh + \int_{t_0}^{t_1} B\widehat{x}h\, dt =$$

$$= \int_{t_0}^{t_1} \left(-\frac{d}{dt}(A\dot{\widehat{x}}) + B\widehat{x} \right) h\, dt \overset{(\text{Э})}{=} -\int_{t_0}^{t_1} \left(\sum_{i=1}^{m} \lambda_i(-\dot{a}_i + b_i) \right) h\, dt =$$

$$= -\sum_{i=1}^{m} \lambda_i \int_{t_0}^{t_1} (-\dot{a}_i + b_i) h \, dt = -\sum_{i=1}^{m} \lambda_i \Big(-\int_{t_0}^{t_1} h \, da_i + \int_{t_0}^{t_1} b_i h \, dt \Big) =$$

$$= -\sum_{i=1}^{m} \lambda_i \int_{t_0}^{t_1} (a_i \dot{h} + b_i h) \, dt = 0.$$

Поскольку для квадратичного функционала $\frac{1}{2} J_0''(x)[h,h] = J_0(h)$, то на экстремали $\widehat{x}$ выполняется соотношение

$$J_0(\widehat{x} + h) = J_0(\widehat{x}) + J_0(h) \quad \forall \ h \in \Pi. \qquad (*)$$

Предположим, выполнено усиленное условие Якоби. Заметим, что существует и единственна экстремаль $\widehat{x}$, являющаяся решением уравнений Эйлера и Якоби с краевыми условиями $\widehat{x}(t_0) = x_0$, $\widehat{x}(t_1) = x_1$. Действительно, если бы $\overline{x}$ была бы другой допустимой экстремалью, то $h = \widehat{x} - \overline{x}$ было бы нетривиальным решением уравнения Якоби с условиями $h(t_0) = h(t_1) = 0$,

$$\int_{t_0}^{t_1} g_i h \, dt = \int_{t_0}^{t_1} (-\dot{a}_i + b_i) h \, dt = -\int_{t_0}^{t_1} h \, da_i + \int_{t_0}^{t_1} b_i h \, dt =$$

$$= \int_{t_0}^{t_1} (a_i \dot{h} + b_i h) dt = \int_{t_0}^{t_1} \big(a_i (\dot{\widehat{x}} - \dot{\overline{x}}) + b_i (\widehat{x} - \overline{x}) \big) dt =$$

$$= \int_{t_0}^{t_1} (a_i \dot{\widehat{x}} + b_i \widehat{x}) dt - \int_{t_0}^{t_1} (a_i \dot{\overline{x}} + b_i \overline{x}) dt = \alpha_i - \alpha_i = 0, \ i = 1, \ldots, m,$$

то точка $\tau = t_1$ является сопряженной точкой. Но это противоречит усиленному условию Якоби.

Функция $\overline{h} \equiv 0$ будет экстремалью с единственным тождественно нулевым набором множителей Лагранжа. Это следует из условия регулярности. Окружим её по Теореме 1.4

центральным полем экстремалей. Возьмем произвольную допустимую функцию $x \in C^1([t_0;\, t_1])$. Тогда $x - \widehat{x} = h \in \Pi$. По основной формуле Вейерштрасса для квадратичного функционала (5.6)

$$J_0(x) - J_0(\widehat{x}) = J_0(\widehat{x} + h) - J_0(\widehat{x}) \overset{(*)}{=} J_0(h) =$$

$$= J_0(h) - J_0(\overline{h} \equiv 0) \overset{(5.6)}{=} \frac{1}{2} \int\limits_{t_0}^{t_1} L_{\dot{x}\dot{x}} (\dot{h} - u)^2 \, dt \geq 0,$$

а в силу выполнения усиленного условия Лежандра $L_{\dot{x}\dot{x}}(t) = A(t) > 0 \ \forall \ t \in [t_0;\, t_1]$. Значит, $\widehat{x} \in \operatorname{absmin} P'$.

b) Пусть не выполнено условие Якоби. Докажем, что тогда $S_{\text{absmin}} = -\infty$.
Функция $\overline{h} \equiv 0$ не доставляет абсолютный минимум в задаче

$$J_0(h) = \frac{1}{2}J_0''(h) = \int\limits_{t_0}^{t_1}(A\dot{h}^2 + Bh^2)dt \to \min; \int\limits_{t_0}^{t_1}(a_i\dot{h} + b_i h)dt = 0,$$
$$i = 1, \ldots, m,\ h(t_0) = h(t_1) = 0 \quad (P'')$$

(если $\overline{h} \equiv 0 \in \operatorname{absmin} P''$, то по теореме о необходимых условиях слабого минимума выполнено условие Якоби). Значит, $S_{\operatorname{absmin} P''} < 0$. Поэтому существует функция $h \in \Pi$ такая, что $J_0''(h) < 0$. Возьмем произвольную допустимую в задаче (P') функцию x, например, линейную функцию. Тогда $x + \lambda h$ также допустимая функция и по формуле Тейлора

$$J_0(x + \lambda h) = J_0(x) + \lambda J_0'(x)[h] + \frac{1}{2}\lambda^2 J_0''(h) \to -\infty$$

при $\lambda \to +\infty$, т. е. $S_{\operatorname{absmin} P'} = -\infty$. $\square$

3.4 Пример

$$\int\limits_0^{T_0} (\dot{x}^2 - x^2)\, dt \to \text{extr}; \quad \int\limits_0^{T_0} x dt = 0,\ x(0) = x(T_0) = 0. \qquad (P)$$

Решение. Условия экстремума 1-го порядка.

Лагранжиан $L = \lambda_0(\dot{x}^2 - x^2) + \lambda_1 x$.

Необходимое условие экстремума 1-го порядка — *уравнение Эйлера*

$$-\frac{d}{dt} L_{\dot{x}} + L_x = 0 \iff -2\lambda_0(\ddot{x} + x) + \lambda_1 = 0.$$

Если $\lambda_0 = 0$, то $\lambda_1 = 0$ — все множители Лагранжа — нули. Этого не может быть. Значит, $\lambda_0 \neq 0$. Положим $\lambda_0 = \frac{1}{2}$. Тогда уравнение Эйлера перепишется в виде

$$\ddot{x} + x - \lambda_1 = 0.$$

Общее решение: $x = C_1 \sin t + C_2 \cos t + C_3$. Неизвестные константы C_1, C_2, C_3 находим из условий на концах и изопериметрического условия. Решение уравнения Эйлера с начальным условием $x(0) = 0$ есть функция вида $x = A \sin t + B(\cos t - 1)$. Отметим, что среди допустимых экстремалей всегда имеется допустимая экстремаль $\widehat{x}(t) \equiv 0$.

Условия экстремума 2-го порядка.

Условие Лежандра. $\widehat{L}_{\dot{x}\dot{x}} = 2 > 0$ — выполняется усиленное условие Лежандра в задаче на минимум.

Условие регулярности очевидным образом выполняется, поскольку функция $g(t) = -\frac{d}{dt} f_{1\dot{x}} + f_{1x} = 1$ на любом отрезке $[0;\, \tau]$, $[\tau;\, T_0]$ при любом τ линейно независима.

Условие Якоби. Выпишем вспомогательную задачу

$$J_0''(\widehat{x})[h, h] \to \min; \quad J_1'(\widehat{x})[h] = 0,\ h(t_0) = h(t_1) = 0 \iff$$

$$\int\limits_0^{T_0} (\dot{h}^2 - h^2)\, dt \to \min; \quad \int\limits_0^{T_0} h dt = 0,\ h(0) = h(T_0) = 0. \qquad (\widetilde{P})$$

Уравнение Эйлера по h для лагранжиана изопериметрической задачи $(\widetilde{P})$ $\widetilde{L} = \mu_0(\dot{h}^2 - h^2) + \mu_1 h$ с множителем Лагранжа $\mu_0 = \frac{1}{2}$

$$-\frac{d}{dt}\widetilde{L}_{\dot{h}} + \widetilde{L}_h = 0 \iff -2\mu_0(\ddot{h} + h) + \mu_1 = 0 \iff \ddot{h} + h - \mu_1 = 0,$$

является уравнением Якоби для исходной задачи (P). В квадратично-линейной задаче уравнение Якоби совпадает с уравнением Эйлера.

Решение h_0 однородного уравнения $\ddot{h} + h = 0$ с условиями $h_0(0) = 0$, $\dot{h}_0(0) = 1$ есть $\sin t$.

Решение h_1 неоднородного уравнения с $\mu_1 = 1$ $\ddot{x} + x - 1 = 0$ с условиями $h_1(0) = \dot{h}_1(0) = 0$ есть $1 - \cos t$.

Матрица $H(t)$ имеет вид

$$H(t) = \begin{pmatrix} h_0(t) & h_1(t) \\ \int\limits_0^t h_0 ds & \int\limits_0^t h_1 ds \end{pmatrix} = \begin{pmatrix} \sin t & 1 - \cos t \\ 1 - \cos t & t - \sin t \end{pmatrix}.$$

Таким образом, сопряженные точки — это решения уравнения

$$\det H(\tau) = 0 \iff \tau \sin \tau - \sin^2 \tau - 1 + 2\cos \tau - \cos^2 \tau = 0$$
$$\iff \tau \sin \tau = 2 - 2\cos \tau \iff 2\tau \sin \frac{\tau}{2} \cos \frac{\tau}{2} = 4 \sin^2 \frac{\tau}{2}$$
$$\iff \left[\begin{array}{l} \sin \dfrac{\tau}{2} = 0 \Leftrightarrow \tau = 2k\pi, \ k \in \mathbb{Z}, \\ \tau \cos \dfrac{\tau}{2} = 2 \sin \dfrac{\tau}{2} \Leftrightarrow \dfrac{\tau}{2} = \operatorname{tg} \dfrac{\tau}{2}. \end{array} \right.$$

Уравнение $\frac{\tau}{2} = \operatorname{tg} \frac{\tau}{2}$ имеет первый положительный корень больший 8. Поэтому ближайшая к нулю сопряженная точка: $\tau = 2\pi$.

Ответ. По теореме о необходимых и достаточных условиях экстремума квадратичного функционала следует, что при $T_0 < 2\pi$ функция $\widehat{x}(t) \equiv 0$ — единственная допустимая экстремаль, доставляющая абсолютный минимум, $S_{\text{absmin}} = 0$; при

$T_0 > 2\pi$ значение задачи $S_{\text{absmin}} = -\infty$. При $T_0 = 2\pi$ допустимые экстремали имеют вид $\widehat{x}(t) = C \sin t$ и можно показать, что они доставляют абсолютный минимум, $S_{\text{absmin}} = 0$.

Замечание. Рассмотренную задачу заменой $\dot{y} = x$, $y(0)=0$ $\left(y(t) = \int_0^t x(s)\,ds\right)$ можно свести к задаче со старшими производными:

$$\int\limits_0^{T_0} (\ddot{y}^2 - \dot{y}^2)\,dt \to \min;\ y(0) = \dot{y}(0) = y(T_0) = \dot{y}(T_0) = 0$$

$\left(y(T_0) = \int_0^{T_0} x\,ds = 0\right)$, а её можно исследовать методами исследования задач со старшими производными.

4 Задача со старшими производными

4.1 Постановка задачи

Рассмотрим задачу со старшими производными (для определенности задачу на минимум) в пространствах $C^n([t_0;\, t_1], \mathbb{R})$[13] и $PC^n([t_0;\, t_1], \mathbb{R})$[14]

$$J(x(\cdot)) = \int\limits_{t_0}^{t_1} L(t, x, \dot{x}, \ddot{x}, \ldots, x^{(n)})\, dt \to \min; \qquad (P)$$

$$x^{(k)}(t_0) = x_{k0},\ x^{(k)}(t_1) = x_{k1},\ k = 0, 1, \ldots, n-1.$$

Дадим определения слабого и сильного экстремума для задачи со старшими производными.

Определение 10. Говорим, что допустимая функция $\widehat{x}(\cdot) \in C^n([t_0;\, t_1])$ доставляет *слабый локальный минимум* в задаче (P), и пишем $\widehat{x}(\cdot) \in \operatorname{wlocmin} P$, если существует $\delta > 0$ такое, что $J(x(\cdot)) \geq J(\widehat{x}(\cdot))$ для любой допустимой функции $x(\cdot)$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^n([t_0;\, t_1])} < \delta$.

Определение 11. Говорим, что допустимая функция $\widehat{x}(\cdot) \in PC^n([t_0;\, t_1])$ доставляет *сильный локальный минимум* в задаче (P), и пишем $\widehat{x}(\cdot) \in \operatorname{strlocmin} P$, если существует $\delta > 0$ такое, что $J(x(\cdot)) \geq J(\widehat{x}(\cdot))$ для любой допустимой функции $x(\cdot)$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^{n-1}([t_0;\, t_1])} < \delta$.

Так как множество функций, среди которых доставляется сильный экстремум, шире, чем для слабого экстремума, то если функция $\widehat{x} \in C^n([t_0;\, t_1])$ доставляет сильный,

[13] Норма в пространстве $C^n([t_0;\, t_1], \mathbb{R})$ задается формулой: $\|y\|_n := \|y\|_{C^n([t_0;\, t_1])} := \max\left\{\|y\|_{C([t_0;\, t_1])}, \|\dot{y}\|_{C([t_0;\, t_1])}, \ldots, \|y^{(n)}\|_{C([t_0;\, t_1])}\right\}$.

[14] Напомним, что $PC^n([t_0;\, t_1], \mathbb{R})$ — пространство $n-1$ раз непрерывно дифференцируемых функций, n-я производная которых является кусочно-непрерывной функцией.

то она доставляет и слабый экстремум. Поэтому для функций $\widehat{x} \in C^{n-1}([t_0;\, t_1])$ необходимое условие слабого экстремума является необходимым условием сильного, а достаточное условие сильного экстремума является достаточным условием слабого.

$$\text{ДУ} \implies \widehat{x} \in \operatorname{strlocmin} P \implies \widehat{x} \in \operatorname{wlocmin} P \implies \text{НУ}$$

4.2 Условия Лежандра и Якоби

Будем далее предполагать, что интегрант L является $2n$ раз непрерывно дифференцируемой функцией в некоторой окрестности расширенного графика $\Gamma_{\widehat{x}\ldots\widehat{x}^{(n)}} = \big\{(t, \widehat{x}(t), \ldots, \widehat{x}^{(n)}(t)) \mid t \in [t_0;\, t_1]\big\}$. Пусть $\widehat{x} \in C^{2n}([t_0;\, t_1])$ — экстремаль в задаче (P), т. е. на ней выполнено уравнение Эйлера–Пуассона.

При наших допущениях относительно гладкости интегранта L функционал $J(x)$ имеет вторую производную по Фреше в точке $\widehat{x}$ на пространстве функций $h \in C^n([t_0;\, t_1])$:

$$J''(\widehat{x})[h, h] = \int\limits_{t_0}^{t_1} \left(\sum_{i,j=0}^{n} \widehat{L}_{x^{(i)}x^{(j)}} h^{(j)} h^{(i)} \right) dt.$$

Важную роль играет коэффициент $\widehat{L}_{x^{(n)}x^{(n)}}$ при $(h^{(n)})^2$.

Говорим, что на экстремали $\widehat{x}$ выполнено *условие Лежандра*, если $\widehat{L}_{x^{(n)}x^{(n)}}(t) \geq 0 \ \forall\ t \in [t_0;\, t_1]$ и *усиленное условие Лежандра*, если $\widehat{L}_{x^{(n)}x^{(n)}}(t) > 0 \ \ \forall\ t \in [t_0;\, t_1]$.

Уравнение Эйлера–Пуассона $\sum\limits_{k=0}^{n} (-1)^k \dfrac{d^k}{dt^k} \widetilde{L}_{h^{(k)}} = 0$ для интегранта $\widetilde{L} = \widetilde{L}(t, h, \dot{h}, \ldots, h^{(n)}) := \sum\limits_{i,j=0}^{n} \widehat{L}_{x^{(i)}x^{(j)}} h^{(j)} h^{(i)}$ называется *уравнением Якоби* для исходной задачи на экстремали $\widehat{x}$. Подставляя $\widetilde{L}$ перепишем уравнение следующим образом:

$$\sum_{k=0}^{n} (-1)^k \frac{d^k}{dt^k} \left(\sum_{j=0}^{n} \widehat{L}_{x^{(k)}x^{(j)}} h^{(j)} \right) = 0$$

Для квадратичного функционала

$$J(x) = \int\limits_{t_0}^{t_1} \Big(\sum_{k=0}^{n} A_k(t) \big(x^{(k)}(t)\big)^2 \Big) dt,$$

интегрант $\widetilde{L} = 2 \sum\limits_{k=0}^{n} A_k (h^{(k)})^2$ и уравнение Якоби приобретает более простой вид:

$$\sum_{k=0}^{n} (-1)^k \frac{d^k}{dt^k} \big(A_k h^{(k)}\big) = 0.$$

Пусть далее на $\widehat{x}$ выполнено усиленное условие Лежандра. Точка τ называется *сопряженной к точке* t_0, если существует нетривиальное решение уравнения Якоби $h(\cdot)$, для которого $h^{(k)}(t_0) = h^{(k)}(\tau) = 0$, $k = 0, 1, \ldots, n-1$. Говорят, что на $\widehat{x}$ 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.

Дадим аналитический способ нахождения сопряженных точек. Уравнение Якоби — линейное дифференциальное уравнение $2n$-го порядка, которое (из-за усиленного условия Лежандра) можно разрешить относительно старшей производной. Пусть $h_1(\cdot), \ldots, h_n(\cdot)$ — фундаментальная система решений уравнения Якоби. Выберем её из условий $H(t_0) = 0$, а $H^{(n)}(t_0)$ — невырожденная матрица, где

$$H(t) = \begin{pmatrix} h_1(t) & \ldots & h_n(t) \\ \ldots & \ldots & \ldots \\ h_1^{(n-1)}(t) & \ldots & h_n^{(n-1)}(t) \end{pmatrix},$$

$$H^{(n)}(t) = \begin{pmatrix} h_1^{(n)}(t) & \ldots & h_n^{(n)}(t) \\ \ldots & \ldots & \ldots \\ h_1^{(2n-1)}(t) & \ldots & h_n^{(2n-1)}(t) \end{pmatrix}.$$

Покажем, что точка τ является сопряженной к t_0 тогда и только тогда, когда матрица $H(\tau)$ является вырожденной. Действительно, условие $\det H(\tau) = 0$ равносильно тому, что столбцы матрицы $H(\tau)$ являются линейно зависимыми. Это означает, что существуют числа $\alpha_1, \ldots, \alpha_n$, не все равные нулю, такие что для функции $h = \sum\limits_{i=1}^{n} \alpha_i h_i$ выполняются граничные условия $h^{(k)}(\tau) = 0$, $k = 0, 1, \ldots, n-1$. Граничное условие $H(t_0) = 0$ равносильно тому, что $h_i^{(k)}(t_0) = 0$, $i = 1, \ldots, n$, $k = 0, 1, \ldots, n-1$. Следовательно, $h^{(k)}(t_0) = 0$, $k = 0, 1, \ldots, n-1$, т. е. τ — сопряжённая точка.

4.3 Условия экстремума II порядка

4.3.1 Необходимые условия слабого экстремума

Теорема 4.1. *Пусть функция $\widehat{x} \in C^{2n}([t_0;\, t_1], \mathbb{R})$ доставляет слабый локальный минимум в задаче (P) ($\widehat{x} \in \operatorname{wlocmin} P$), интегрант L непрерывно дифференцируем $n+2$ раза в некоторой окрестности расширенного графика $\Gamma_{\widehat{x} \ldots \widehat{x}^{(n)}}$. Тогда на $\widehat{x}$ выполняются*

a) уравнение Эйлера–Пуассона:

$$\sum_{k=0}^{n} (-1)^k \frac{d^k}{dt^k} \widehat{L}_{x^{(k)}}(t) = 0 \quad \forall\ t \in [t_0;\, t_1];$$

b) условие Лежандра: $\widehat{L}_{x^{(n)} x^{(n)}}(t) \geq 0\ \forall\ t \in [t_0;\, t_1]$;

c) если на экстремали выполнено усиленное условие Лежандра, то выполняется и условие Якоби.

4.3.2 Достаточные условия сильного экстремума

Теорема 4.2. *Пусть функция* $\widehat{x} \in C^{2n}([t_0;\, t_1], \mathbb{R})$ *— допустимая экстремаль в задаче* (P)*, интегрант* $L \in C^{n+2}(V \times \mathbb{R})$*, где* $V \subset \mathbb{R}^{n+1}$ *— некоторая окрестность графика* $\big\{(t, \widehat{x}(t), \ldots, \widehat{x}^{(n-1)}(t)) \mid t \in [t_0;\, t_1]\big\}$*, на* $\widehat{x}$ *выполнены усиленные условия Лежандра и Якоби, интегрант* L *является выпуклым по* $x^{(n)}$ *на* V*. Тогда* $\widehat{x}$ *доставляет сильный локальный минимум в задаче со старшими производными* (P)*.*

Доказательство. Наметим схему доказательства теоремы.
А) *Построение центрального поля экстремалей.* Расписывая уравнение Эйлера–Пуассона

$$\sum_{k=0}^{n} (-1)^k \frac{d^k}{dt^k} L_{x^{(k)}}(t, x, \dot{x}, \ldots, x^{(n)}) = 0$$

и используя усиленное условие Лежандра, сводим его к уравнению $2n$-го порядка, которое (из-за усиленного условия Лежандра) можно разрешить относительно старшей производной. Поэтому можно построить n-параметрическое семейство функций $x(t, \lambda) = x(t, \lambda_1, \ldots, \lambda_n)$, удовлетворяющих уравнению Эйлера–Пуассона с начальными данными

$$x^{(k)}(t_*, \lambda) = \widehat{x}^{(k)}(t_*),\ x^{(k+n)}(t_*, \lambda) = \widehat{x}^{(k+n)}(t_*) + \lambda_{k+1}, \qquad k = 0, 1, \ldots, n-1, \quad (*)$$

где t_* — некоторая точка в окрестности точки t_0, $t_* < t_0$ и $x(t, 0) = \widehat{x}(t)$.

Обозначим

$$x_\lambda(t, \lambda)\Big|_{\lambda=0} = h(t) \quad \Big(h_i(t) = \frac{\partial x(t, \lambda)}{\partial \lambda_i}\Big|_{\lambda=0},\ i = 1, \ldots, n\Big).$$

Рассмотрим матрицу

$$H(t) = \begin{pmatrix} h_1(t) & \ldots & h_n(t) \\ \ldots & \ldots & \ldots \\ h_1^{(n-1)}(t) & \ldots & h_n^{(n-1)}(t) \end{pmatrix}.$$

Тогда из условий $(*)$ получим, что $H(t_*) = 0$, $H^{(n)}(t_*) = I$.

Функция $x(t, \lambda)$ удовлетворяет уравнению Эйлера–Пуассона для любого λ. Продифференцируем это уравнение по λ в точке $\lambda = 0$:

$$\frac{\partial}{\partial \lambda_i} \sum_{k=0}^{n} (-1)^k \frac{d^k}{dt^k} L_{x^{(k)}}(t, x(t,\lambda), \dot{x}(t,\lambda), \ldots, x^{(n)}(t,\lambda))\Big|_{\lambda=0} = 0 \Rightarrow$$

$$\sum_{k=0}^{n} (-1)^k \frac{d^k}{dt^k} \sum_{j=0}^{n} \widehat{L}_{x^{(k)}x^{(j)}}(t) \frac{\partial}{\partial \lambda_i} x^{(j)}(t,\lambda)\Big|_{\lambda=0} = 0 \Rightarrow$$

$$\sum_{k=0}^{n} (-1)^k \frac{d^k}{dt^k} \sum_{j=0}^{n} \widehat{L}_{x^{(k)}x^{(j)}}(t) h_i^{(j)}(t) = 0, \ i = 1, \ldots, n.$$

Получили, что функции h_i удовлетворяют уравнению Якоби. Из усиленного условия Якоби следует, что при t_* достаточно близком к t_0 матрица $H(t)$ невырождена $\forall\ t \in [t_0; t_1]$. Как и при построении центрального поля в п. 1.4.6 по теореме об обратной функции существует экстремаль $x(t, \lambda)$, для которой

$$x^{(k)}(\tau, \lambda) = \xi_k, \ k = 0, 1, \ldots, n-1,$$

причём по паре (τ, ξ), $\xi = (\xi_0, \ldots, \xi_{n-1})$, где $|\widehat{x}^{(k)}(\tau) - \xi_k|$ — мало, величина $\lambda(\tau, \xi)$ определяется однозначно.

Функция наклона поля

$$u(\tau, \xi) := \frac{d^n}{dt^n} x(t, \lambda(\tau, \xi))\Big|_{t=\tau}.$$

В) *S-функция и её дифференциал.* Положим

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

Дифференцируя S-функцию, приходим к формулам

$$\frac{\partial S(\tau,\xi)}{\partial \tau} = L(\tau,\xi,u(\tau,\xi)) - \sum_{k=0}^{n-2} p_k(\tau,\xi)\xi_{k+1} - p_{n-1}(\tau,\xi)u(\tau,\xi),$$

$$\frac{\partial S(\tau,\xi)}{\partial \xi_k} = p_k(\tau,\xi),$$

где

$$p_k(\tau,\xi) := L_{x^{(k+1)}}(\tau,\xi,u(\tau,\xi)) - \frac{d}{dt} L_{x^{(k+2)}}(t,\xi,u(\tau,\xi))\Big|_{t=\tau} + \ldots$$

$$+(-1)^{n-k-1}\frac{d^{n-k-1}}{dt^{n-k-1}} L_{x^{(n)}}(t,\xi,u(\tau,\xi))\Big|_{t=\tau},\ k=0,1,\ldots,n-1.$$

C) *Основная формула Вейерштрасса.* По основной формуле Вейерштрасса

$$J(x) - J(\widehat{x}) = \int\limits_{t_0}^{t_1} \mathcal{E}(t,x,\dot{x},\ldots,x^{(n-1)},u(t,x,\ldots,x^{(n-1)}),x^{(n)})\,dt,$$

где $\mathcal{E}(t,\overline{x},u,x^{(n)}) = f(t,\overline{x},x^{(n)}) - f(t,\overline{x},u) - (x^{(n)}-u)L_{x^{(n)}}(t,\overline{x},u)$, $\overline{x} = (x,\dot{x},\ldots,x^{(n-1)})$. Поскольку L выпукло по $x^{(n)}$ в окрестности траектории, то $\mathcal{E}(t,\overline{x},u,x^{(n)}) \geq 0\ \forall\ (u,x^{(n)}) \in \mathbb{R}\times\mathbb{R}$. Значит, $J(x) \geq J(\widehat{x})$, т. е. функция $\widehat{x}$ доставляет сильный локальный минимум. $\square$

4.3.3 Квадратичный функционал

Рассмотрим задачу со старшими производными для квадратичного функционала, имеющего “диагональную” форму:

$$J(x(\cdot)) = \int\limits_{t_0}^{t_1} \left(\sum_{k=0}^{n} A_k(t)(x^{(k)}(t))^2 \right) dt \to \min;$$
$$x^{(k)}(t_0) = x_{k0},\ x^{(k)}(t_1) = x_{k1},\ k = 0, 1, \ldots, n-1. \quad (P')$$

Теорема 4.3. *Пусть в задаче* (P') *функции* $A_k \in C^k([t_0;\, t_1])$, $k = 0, 1, \ldots, n$, *и выполнено усиленное условие Лежандра* $(\widehat{L}_{x^{(n)}x^{(n)}}(t) = 2A_n(t) > 0\ \forall\ t \in [t_0;\, t_1])$. *Тогда*

a) если выполнено усиленное условие Якоби, то допустимая экстремаль $\widehat{x}$ *существует, единственна и доставляет абсолютный минимум;*

b) если не выполнено условие Якоби, то $S_{\text{absmin}} = -\infty$.

Доказательство. a) Пусть выполнено усиленное условие Якоби. Докажем, что тогда допустимая экстремаль $\widehat{x}$ существует, единственна и доставляет абсолютный минимум.

Для квадратичного функционала по формуле Тейлора имеет место равенство

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

Поскольку $\widehat{x}$ — допустимая экстремаль в задаче, то $J'(\widehat{x})[h] = 0$ $\forall\ h \in C_0^n([t_0;\, t_1])$ (это эквивалентно уравнению Эйлера–Пуассона). Значит,

$$J(\widehat{x} + h) = J(\widehat{x}) + J(h) \quad \forall\ h \in C_0^n([t_0;\, t_1]). \qquad (*)$$

Заметим, что уравнения Эйлера–Пуассона и Якоби для задачи (P') совпадают. Предположим, выполнено усиленное условие Якоби. Тогда существует и единственна экстремаль $\widehat{x}$, являющаяся решением уравнений Эйлера–Пуассона и Якоби с краевыми условиями $x^{(k)}(t_0) = x_{k0}$, $x^{(k)}(t_1) = x_{k1}$,

$k = 0, \ldots, n-1$. Действительно, если бы $\overline{x}$ была бы другой допустимой экстремалью, то $h = \widehat{x} - \overline{x}$ было бы нетривиальным решением уравнения Якоби с нулевыми краевыми условиями, а это противоречит усиленному условию Якоби.

Поскольку уравнение Эйлера–Пуассона для квадратичного функционала J является однородным уравнением, то функция $\overline{h} \equiv 0$ будет экстремалью. Окружим её центральным полем экстремалей. Тогда для функции $h \in C_0^n([t_0;\, t_1])$ по основной формуле Вейерштрасса для квадратичного функционала (5.6)

$$J(\widehat{x} + h) - J(\widehat{x}) \overset{(*)}{=} J(h) = J(h) - J(\overline{h} \equiv 0) =$$

$$= \int\limits_{t_0}^{t_1} \Big(L(t, h, \dot{h}, \ldots, h^{(n)}) - L(t, h, \dot{h}, \ldots, h^{(n-1)}, u(t,h)) -$$

$$-(h^{(n)} - u(t,h)) L_{h^{(n)}}(t, h, \dot{h}, \ldots, h^{(n-1)}, u(t,h)) \Big) dt =$$

$$= \int\limits_{t_0}^{t_1} \mathcal{E}(t, h, \ldots, h^{(n-1)}, u, h^{(n)}) dt \overset{(5.6)}{=} \int\limits_{t_0}^{t_1} \frac{L_{h^{(n)}h^{(n)}}}{2} (h^{(n)} - u)^2 dt \geq 0,$$

поскольку по усиленному условию Лежандра $L_{h^{(n)}h^{(n)}}(t) = 2A_n(t) > 0$. Значит, $\widehat{x} \in \operatorname{absmin} P'$.

b) Пусть не выполнено условие Якоби. Докажем, что тогда $S_{\text{absmin}} = -\infty$.
Функция $\overline{h} \equiv 0$ не доставляет абсолютный минимум в задаче

$$J(h) = \frac{1}{2}J''(h) = \int\limits_{t_0}^{t_1} \left(\sum_{k=0}^{n} A_k (h^{(k)})^2 \right) dt \to \min; \quad \begin{array}{l} h^{(k)}(t_0) = 0, \\ h^{(k)}(t_1) = 0, \end{array}$$
$$k = 0, 1, \ldots, n-1 \quad (P'')$$

(если $\overline{h} \equiv 0 \in \operatorname{absmin} P''$, то по теореме о необходимых условиях слабого минимума, выполнено условие Якоби). Значит, $S_{\operatorname{absmin} P''} < 0$. Поэтому существует функция $h \in C_0^n([t_0; t_1])$ такая, что $J''(h) < 0$. Возьмем произвольную допустимую в задаче (P') функцию x, например, линейную функцию. Тогда $x + \lambda h$ также допустимая функция и по формуле Тейлора

$$J(x + \lambda h) = J(x) + \lambda J'(x)[h] + \frac{1}{2}\lambda^2 J''(h) \to -\infty$$

при $\lambda \to +\infty$, т. е. $S_{\operatorname{absmin} P'} = -\infty$. $\square$

4.4 Пример

$$J(x(\cdot)) = \int\limits_0^\pi (\ddot{x}^2 - \dot{x}^2)\,dt \to \text{extr};\ x(0) = \dot{x}(0) = \dot{x}(\pi) = 0,\ x(\pi) = 1.$$

Решение. Условие экстремума 1-го порядка — *уравнение Эйлера–Пуассона*:

$$\frac{d^2}{dt^2} L_{\ddot{x}} - \frac{d}{dt} L_{\dot{x}} + L_x = 0 \iff x^{(4)} + \ddot{x} = 0.$$

Характеристическое уравнение $k^4 + k^2 = 0 \Leftrightarrow k_{1,2} = \pm i$, $k_{3,4} = 0$. Общее решение: $x(t) = C_1 \sin t + C_2 \cos t + C_3 t + C_4$, $\dot{x}(t) = C_1 \cos t - C_2 \sin t + C_3$. Неизвестные константы $C_1, \ldots, C_4$ находим из условий на концах:

$x(0) = 0 \Rightarrow C_2 + C_4 = 0 \Rightarrow C_4 = -C_2$,

$\dot{x}(0) = 0 \Rightarrow C_1 + C_3 = 0 \Rightarrow C_3 = -C_1 \Rightarrow$

$x = C_1(\sin t - t) + C_2(\cos t - 1)$,

$x(\pi) = 1 \Rightarrow C_1(-\pi) + C_2(-2) = 1$,

$\dot{x}(\pi) = 0 \Rightarrow C_1(-2) = 0 \Rightarrow C_1 = 0 \Rightarrow C_2 = -\frac{1}{2}$.

Допустимая экстремаль $\widehat{x} = \dfrac{1 - \cos t}{2}$.

Исследование. Возьмем допустимую функцию x. Тогда для функции $h = x - \widehat{x}$ ($\Leftrightarrow x = \widehat{x} + h$) выполняется условие $h \in C_0^2([0;\, \pi])$. Для квадратичного функционала

$$J(\widehat{x} + h) - J(\widehat{x}) = J(h) = \int\limits_0^\pi (\ddot{h}^2 - \dot{h}^2)\,dt.$$

Непосредственно из выписанного равенства невозможно понять, доставляет ли $\widehat{x}$ экстремум, так как знак разности неопределен. Проведем дальнейшее исследование с помощью условий II порядка.

Условие Лежандра. $\widehat{L}_{\ddot{x}\ddot{x}} = 2 > 0$ — выполняется усиленное условие Лежандра в задаче на минимум.

Условие Якоби. Для квадратичного функционала уравнение Якоби совпадает с уравнением Эйлера–Пуассона:

$$h^{(4)} + \ddot{h} = 0.$$

Общее решение:

$h = C_1 \sin t + C_2 \cos t + C_3 t + C_4 \Rightarrow$
$\dot{h} = C_1 \cos t - C_2 \sin t + C_3,$
$\ddot{h} = -C_1 \sin t - C_2 \cos t,$
$\dddot{h} = -C_1 \cos t + C_2 \sin t.$

Ищем решения уравнения Якоби $h_1(\cdot)$, $h_2(\cdot)$ с начальными условиями

$$H(0) = \begin{pmatrix} h_1(0) & h_2(0) \\ \dot{h}_1(0) & \dot{h}_2(0) \end{pmatrix} = 0, \quad \ddot{H}(0) = \begin{pmatrix} \ddot{h}_1(0) & \ddot{h}_2(0) \\ \dddot{h}_1(0) & \dddot{h}_2(0) \end{pmatrix} = I.$$

Константы C_1, C_2, C_3, C_4 находим из условий на концах:

$$\begin{cases} h_1(0) = 0, \\ \dot{h}_1(0) = 0, \\ \ddot{h}_1(0) = 1, \\ \dddot{h}_1(0) = 0, \end{cases} \iff \begin{cases} C_2 + C_4 = 0, \\ C_1 + C_3 = 0, \\ -C_2 = 1, \\ -C_1 = 0, \end{cases} \iff \begin{cases} C_1 = 0, \\ C_2 = -1, \\ C_3 = 0, \\ C_4 = 1, \end{cases} \Longrightarrow$$

$h_1(t) = 1 - \cos t.$

Аналогично ищем функцию h_2 с условиями

$$\begin{cases} h_2(0) = 0, \\ \dot{h}_2(0) = 0, \\ \ddot{h}_2(0) = 0, \\ \dddot{h}_2(0) = 1, \end{cases} \iff \begin{cases} C_2 + C_4 = 0, \\ C_1 + C_3 = 0, \\ -C_2 = 0, \\ -C_1 = 1, \end{cases} \iff \begin{cases} C_1 = -1, \\ C_2 = 0, \\ C_3 = 1, \\ C_4 = 0, \end{cases} \Longrightarrow$$

$h_2(t) = t - \sin t.$

Следовательно, $H(t) = \begin{pmatrix} h_1 & h_2 \\ \dot{h}_1 & \dot{h}_2 \end{pmatrix} = \begin{pmatrix} 1 - \cos t & t - \sin t \\ \sin t & 1 - \cos t \end{pmatrix}.$

Сопряженные точки — нули уравнения

$$\det H(\tau) = 0 \iff \det\begin{pmatrix} 1-\cos\tau & \tau-\sin\tau \\ \sin\tau & 1-\cos\tau \end{pmatrix} = 0 \iff$$

$$(1-\cos\tau)^2 = \sin\tau(\tau-\sin\tau) \iff$$

$$1 - 2\cos\tau + \cos^2\tau = \tau\sin\tau - \sin^2\tau \iff 2(1-\cos\tau) = \tau\sin\tau$$

$$\iff 2\sin^2\frac{\tau}{2} = \tau\sin\frac{\tau}{2}\cos\frac{\tau}{2} \iff \sin\frac{\tau}{2} = 0,\ \operatorname{tg}\frac{\tau}{2} = \frac{\tau}{2}.$$

Ближайшая к нулю сопряженная точка: $\tau = 2\pi$. На полуинтервале $(0;\,\pi]$ нет сопряженных точек, следовательно, выполняется усиленное условие Якоби.

По теореме об условиях 2-го порядка для квадратичного функционала функция $\widehat{x} \in \mathrm{absmin}$,

$$S_{\mathrm{absmin}} = J(\widehat{x}) = \int\limits_0^\pi \left(\ddot{\widehat{x}}^2 - \dot{\widehat{x}}^2\right) dt =$$

$$= \frac{1}{4}\int\limits_0^\pi (\cos^2 t - \sin^2 t)\, dt = \frac{1}{4}\int\limits_0^\pi \cos 2t\, dt = 0.$$

Покажем, что $S_{\mathrm{absmax}} = +\infty$. Возьмем последовательность допустимых функций $x_n = \widehat{x} + nh$, $h \in C_0^2([t_0;\,t_1])$, например, $h(t) = \sin^2 t = \frac{1-\cos 2t}{2}$. Тогда $\dot{h} = \sin 2t$, $\ddot{h} = 2\cos 2t$,

$$J(x_n) = J(\widehat{x}) + J(nh) = J(\widehat{x}) + n^2\int\limits_0^\pi (\ddot{h}^2 - \dot{h}^2)\, dt =$$

$$= J(\widehat{x}) + n^2\int\limits_0^\pi (4\cos^2 2t - \sin^2 2t)\, dt =$$

$$= J(\widehat{x}) + n^2\int\limits_0^\pi \left(4\cdot\frac{1+\cos 2t}{2} - \frac{1-\cos 2t}{2}\right) dt =$$

$$= J(\widehat{x}) + n^2\left(\frac{4}{2} - \frac{1}{2}\right) \to +\infty \text{ при } n \to +\infty.$$

Ответ. $\widehat{x} = \dfrac{1-\cos t}{2} \in \mathrm{absmin}$, $S_{\mathrm{absmin}} = 0$; $S_{\mathrm{absmax}} = +\infty$.