МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В.Ломоносова
Механико-математический факультет

ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ И ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

проф. С.В.Конягин
1/2 года, 4 курс, отделение математики, 2 поток
Лекции записаны и набраны студентами В.Ю.Лёвиным и В.В.Осокиным.

Последняя редакция:
19.12.2005г.

Москва 2005

Содержание

1 Необходимые условия экстремума для гладких задач без ограничений.

Пусть задана функция $f : A \to \mathbb{R}$. Будем исследовать ее точки экстремумов ($f(x) \to extr$).

Определение 1.1 *$\widehat{x}$ - точка минимума (строгого минимума) функции f(x), если $\forall x \in A\ f(x) \geq f(\widehat{x})$ ($f(x) > f(\widehat{x})$ соответственно).*

Аналогично определяются точки максимума (строгого максимума) и точки локального минимума и максимума.

Напомним формулировку теоремы Ферма.

Теорема 1.1 *Пусть $A \subset \mathbb{R}$, $\widehat{x} \in intA$, $\widehat{x}$ - точка локального минимума f. Тогда из существования $f'(\widehat{x})$ следует $f'(\widehat{x}) = 0$*

Будем рассматривать линейное нормированное пространство (ЛНП) X над $\mathbb{R}$. Напомним свойства нормы:

- $\|x\| \geq 0$; $\|x\| = 0 \Leftrightarrow x = 0$
- $\|x + y\| \leq \|x\| + \|y\|$
- $\|\alpha x\| = |\alpha|\|x\|$

Напомним теперь определение линейного ограниченного функционала.

Определение 1.2 *линейный функционал x^* - отображение $x \to < x^*, x >$, линейное по x. Этот функционал называется ограниченным (и, значит, непрерывным), если $\exists c > 0 : \forall x \in X\ | < x^*, x > | \leq c\|x^*\|$, где $\|x^*\| := \sup_{x\neq 0} \frac{|<x^*,x>|}{\|x\|} = \sup_{\|x\|\leq 1} | < x^*, x > | < \infty$*

Определение 1.3 *пространство X^* линейных ограниченных функционалов над X называется сопряженным пространством.*

Пример: $X = \mathbb{R}^n$, $x = (x_1, ..., x_n)^T, \|x\| = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$. Тогда $X^* = \{x^* = (x_1^*, ..., x_n^*)\}, < x^*, x >= x^*x$.

Пусть X-ЛНП, $A \subset X$, $\widehat{x} \in intA$, $h \in X$.

Определение 1.4 *Вариацией по Лагранжу в точке $\widehat{x}$ называется предел $\delta f(\widehat{x}, h) := \lim_{\alpha\to 0} \frac{f(\widehat{x}+\alpha h)-f(\widehat{x})}{\alpha}$, если он существует для любого h.*

Если $X = \mathbb{R}, h = 1$, то $\delta f(\widehat{x}, h) = f'(\widehat{x})$

Если $X = \mathbb{R}^n, h = e_i = (0...1...0)^T$, то $\delta f(\widehat{x}, e_i) = \frac{\partial f}{\partial x_i}(\widehat{x})$

Приведем пример функции, у которой существуют частные производные в точке, но не существует вариации по Лагранжу:

$X = \mathbb{R}^2, f(x) = sgn(x_1x_2), \frac{\partial f}{\partial x_i}(0) = 0$

Определение 1.5 *Пусть X-ЛНП, $A \subset X, \widehat{x} \in intA, f : A \to \mathbb{R}$. Если $\exists x^* \in X^* : f(\widehat{x}+h) = f(\widehat{x}) + < x^*, h > + r(h)$, $\lim_{h\to 0} \frac{|r(h)|}{\|h\|} = 0$ $(*)$, то говорим, что f дифференцируема в точке $\widehat{x}$ по Фреше и $f'(\widehat{x}) = x^*$.*

Корректность этого определения вытекает из леммы

Лемма 1.1 *если $(*)$ выполнено, $h \in X$, то $\delta f(\widehat{x}, h) =< x^*, h >$.*

$\square$ Пусть $h \in X$, $h \neq 0$, $\alpha \neq 0$, $|\alpha|$ мал (т.е. $\widehat{x} + \alpha h \in A$). Тогда $\frac{f(\widehat{x}+\alpha h)-f(\widehat{x})}{\alpha} = \frac{1}{\alpha}(f(\widehat{x}) + < x^*, \alpha h > + r(\alpha h) - f(\widehat{x})) = \frac{\alpha < x^*, h>}{\alpha} + \frac{r(\alpha h)}{\|\alpha h\|}\frac{\|\alpha h\|}{\alpha}$. Но $\frac{r(\alpha h)}{\|\alpha h\|} \to 0, \alpha \to 0$; $\frac{\|\alpha h\|}{\alpha}$ - ограничено. Значит, $lim_{\alpha\to 0}\frac{f(\widehat{x}+\alpha h)-f(\widehat{x})}{\alpha} =< x^*, h >$ $\blacksquare$

Приведем пример, когда существует вариация по Лагранжу, но нет производной:

$$f(x) = \begin{cases} 0, x_1^2 \neq x_2 \vee x_1 = x_2 = 0 \\ 1, x_1^2 = x_2 \neq 0 \end{cases}$$

Здесь вариация по Лагранжу в 0 равна 0.

Определение 1.6 *Пусть X - ЛНП, $A \subset X$, $\hat{x} \in A, f : A \to \mathbb{R}$. $\hat{x}$ назовем точкой локального минимума (locmin), если $\exists \epsilon > 0 : \forall x \in A$ $(\|x - \hat{x}\| < \epsilon \Rightarrow f(x) \geq f(\hat{x}))$*

Необходимое условие экстремума:

Теорема 1.2 *Пусть X - ЛНП, $A \subset X$, $\hat{x} \in intA, f : A \to \mathbb{R}, \hat{x} - locextr, \forall h \in X$ $\exists \delta f(\hat{x}, h)$. Тогда $\delta f(\hat{x}, h) = 0$.*

Вывод: если есть вариация по Лагранжу, то вариация по всем направлениям равна 0.

$\square$ $\phi(\alpha) := f(\hat{x} + \alpha h)$, 0 - locextr для ϕ. По теореме Ферма $\phi'(0) = 0$. Но $\phi'(0) = \lim_{\alpha\to 0}\frac{\phi(\alpha)-\phi(0)}{\alpha} = \lim_{\alpha\to 0}\frac{f(\hat{x}+\alpha h)-f(\hat{x})}{\alpha} = \delta f(\hat{x}, h) = 0$ $\blacksquare$

Учитывая, что $< f'(\hat{x}), h >= \delta f(\hat{x}, h)$, получаем следствие (теорему Ферма для линейных нормированных подпространств):

Теорема 1.3 *Пусть X - ЛНП, $A \subset X, \hat{x} \in intA, f : A \to \mathbb{R}, \hat{x}$ - locextr, существует $f'(\hat{x})$. Тогда $f'(\hat{x}) = 0$*

Рассмотрим теперь следующую задачу (Аполлония): требуется найти расстояние от точки на плоскости до эллипса: $(\xi, \eta)^T \in \mathbb{R}^2, \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1; (x_1 - \xi)^2 + (x_2 - \eta)^2 \to min$. Теорема Ферма здесь не применима, так как минимум ищется по множеству, не содержащему внутренних (относительно $\mathbb{R}^2$) точек. Однако, теорему Ферма можно использовать, если задачу формализовать по-другому. Для этого перейдем в новые координаты:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a\cos\phi \\ b\sin\phi \end{pmatrix}$$

В них задача принимает следующий вид: $(a\cos\phi - \xi)^2 + (b\sin\phi - \eta)^2 \to min$. Здесь уже минимизируем по ϕ, причем ϕ пробегает всю прямую $\Rightarrow$ можно применить теорему Ферма.

2 Простейшая задача классического вариационного исчисления. Уравнение Эйлера.

Рассмотрим задачу

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

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

$U = intU \subset \mathbb{R}^{2n+1}, L : U \to \mathbb{R}$ — непрерывная функция, называемая интегрантом. Эта задача называется простейшей задачей классического вариационного исчисления (З).

Функция $x(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ называется допустимой, если $x(t_0) = x_0, x(t_1) = x_1$; $\forall t \in [t_0, t_1]$ $(t, x(t), \dot{x}(t)) \in U$.

Введем в пространстве $C^1([t_0, t_1], \mathbb{R}^n)$ норму следующим образом:

$$\|x(\cdot)\|_{C^1([t_0,t_1],\mathbb{R}^n)} = \max(\|x(\cdot)\|_{C([t_0,t_1],\mathbb{R}^n)}, \|\dot{x}(\cdot)\|_{C([t_0,t_1],\mathbb{R}^n)})$$

где $\|x(\cdot)\|_{C([t_0,t_1],\mathbb{R}^n)} = max_{t\in[t_0,t_1]}|x(t)|_{\mathbb{R}^n}$.

Определение 2.1 *Допустимая функция $\hat{x}(\cdot)$ доставляет слабый минимум в (З), если $\exists \epsilon > 0 : \forall$ допустимой функции $x(\cdot)$, такой что $\|x(\cdot) - \hat{x}(\cdot)\|_{C^1} < \epsilon$ верно $J(x(\cdot)) \geq J(\hat{x}(\cdot))$.*

Аналогично определяется слабый максимум.

Будем рассматривать случай $n = 1$. Получим необходимое условие экстремума:

Теорема 2.1 *В дополнение к условиям задачи (З) предположим, что $L, L_x, L_{\dot{x}}$ - непрерывные в U, $\hat{x}(\cdot)$ - слабый экстремум в (З). Тогда $\forall t \in [t_0, t_1]$ выполнено уравнение Эйлера $\frac{d}{dt}L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) = L_x(t, \hat{x}(t), \dot{\hat{x}}(t))$ (в сокращенной записи: $\frac{d}{dt}\hat{L}_{\dot{x}}(t) = \hat{L}_x(t)$).*

$\square$ Для доказательства можно использовать теорему Ферма. Хотя в нашем случае все производные существуют, будем проверять равенство нулю вариации Лагранжа. Заметим, что в (З) есть ограничения $x(t_0) = x_0, x(t_1) = x_1$, которые мешают применить теорему Ферма в явном виде. Но эта проблема легко решается за счет введения следующего ЛНП:

$X = C_0^1[t_0, t_1](:= C_0^1([t_0, t_1], \mathbb{R})) = \{x(\cdot) \in C^1[t_0, t_1] : x(t_0) = x(t_1) = 0\}$, $\|\cdot\|_{C_0^1} = \|\cdot\|_{C^1}$

Рассмотрим следующую функцию: $F(h(\cdot)) = J(\hat{x}(\cdot) + h(\cdot))$. Тогда (функция) 0 - locextr для функции F в пространстве $C_0^1[t_0, t_1]$, т.к. если $\hat{x}(\cdot)$ - слабый минимум, то при малых ϵ и $\|h(\cdot)\|_{C_0^1} < \epsilon$ $x(\cdot) = \hat{x}(\cdot) + h(\cdot)$ является допустимой $\Rightarrow J(x(\cdot)) \geq J(\hat{x}(\cdot)) \Leftrightarrow F(h(\cdot)) \geq F(0)$.

Мы знаем, что вариация F по Лагранжу в 0 равна 0. Получим отсюда требуемое уравнение Эйлера:

$$\delta F(0, h(\cdot)) = \lim_{\alpha\to 0} \frac{1}{\alpha}(F(\alpha h(\cdot)) - F(0)) = \{\hat{L}(t) := L(t, \hat{x}(t), \dot{\hat{x}}(t))\} =$$

$$= \lim_{\alpha \to 0} \frac{1}{\alpha} \left(\int_{t_0}^{t_1} L(t, \hat{x}(t) + \alpha h(t), \dot{\hat{x}}(t) + \alpha \dot{h}(t)) dt - \int_{t_0}^{t_1} \hat{L}(t) dt \right) =$$

$$= \{\text{учитывая, что сходимость равномерна по t}\} =$$

$$= \int_{t_0}^{t_1} \lim_{\alpha \to 0} \frac{1}{\alpha} \left(L(t, \hat{x}(t) + \alpha h(t), \dot{\hat{x}}(t) + \alpha \dot{h}(t)) - \hat{L}(t) \right) dt =$$

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

0 - locextr $\to$ $\forall h(\cdot) \in C_0^1[t_0, t_1]$ $\delta F(0, h) = 0$. Отсюда и интеграл выше равен нулю. Заметим, что из условий теоремы и уравнения Эйлера следует, что $\hat{L}_{\dot{x}}$ дифференцируема по t. Тогда для окончания доказательства теоремы достаточно доказать следующую лемму (Дюбуа-Реймона):

Лемма 2.1 *Пусть* $a(\cdot), b(\cdot) \in C[t_0, t_1]$, *причем* $\forall h(\cdot) \in C_0^1[t_0, t_1]$ $\int_{t_0}^{t_1} (a(t)h(t) + b(t)\dot{h}(t))dt = 0$. *Тогда* $\forall t$ $a(t) = \dot{b}(t)$

□ Положим $\dot{A}(t) = a(t)$, т.е. A - какая-либо первообразная от a. Тогда $\int_{t_0}^{t_1} a(t)h(t)dt = \int_{t_0}^{t_1} h(t)dA(t) = h(t)A(t)|_{t_0}^{t_1} - \int_{t_0}^{t_1} A(t)\dot{h}(t)dt \Rightarrow \int_{t_0}^{t_1} (b(t) - A(t))\dot{h}(t)dt = 0$. Выберем теперь $A(\cdot)$ и $h_0(\cdot)$ так, чтобы $\int_{t_0}^{t_1} A(t)dt = \int_{t_0}^{t_1} b(t)dt$, а $h_0(t) = \int_{t_0}^{t} (b(s) - A(s))ds$. При этом $h_0(t_0) = h_0(t_1) = 0 \Rightarrow h_0(\cdot) \in C_0^1$. Для таких A и h_0 имеем $0 = \int_{t_0}^{t_1} (b(t) - A(t))\dot{h}_0(t)dt = \int_{t_0}^{t_1} (b(t) - A(t))^2 dt$ и, следовательно, $b \equiv A$. Тогда $\dot{b}(t) = \dot{A}(t) = a(t)$ ■

На этом и заканчивается доказательство теоремы ■

Рассмотрим теперь вкратце векторный случай. Пусть $x(t) = (x_1(t), ..., x_n(t))^T, \hat{L}_x = (\hat{L}_{x_1}, ..., \hat{L}_{x_n})$. Тогда уравнение Эйлера имеет вид $\frac{d}{dt}\hat{L}_{\dot{x}} = \hat{L}_x$, что равносильно системе $\{\frac{d}{dt}\hat{L}_{\dot{x}_i}(t) = \hat{L}_{x_i}(t) (i \in \overline{1, n})\}$.

Задача 1 *доказать необходимое условие экстремума для произвольного n*

Указание: фиксируем $\hat{x}_1(\cdot), ..., \hat{x}_{i-1}(\cdot), \hat{x}_{i+1}(\cdot), ...$ Варьируем $x_i(\cdot)$.

Определение 2.2 *если* $\hat{x}(\cdot)$ *удовлетворяет уравнению Эйлера, то* $\hat{x}(\cdot)$ *называется экстремалью.*

Любая функция, которая доставляет минимум или максимум, является экстремалью. Обратное, вообще говоря, неверно.

3 Задача Больца. Условия трансверсальности.

Перейдем теперь к рассмотрению другой задачи, а именно задачи Больца (ЗБ). Пусть $U = intU \subset \mathbb{R}^{2n+1}, L \in C(U)(:= C(U, \mathbb{R})), V = intV \subset \mathbb{R}^{2n}, l \in C(V)$. Будем исследовать функционал Больца

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

L, как и раньше, называется интегрантом, l называется терминантом. Допустимая функция для (ЗБ) определяется аналогично тому, как она определялась в (З).

Определение 3.1 *допустимая $\hat{x}(\cdot)$ доставляет слабый минимум в (ЗБ), если $\exists \epsilon > 0 : \forall x(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ такой, что $\|x(\cdot) - \hat{x}(\cdot)\|_{C^1} < \epsilon$, верно $B(x(\cdot)) \geq B(\hat{x}(\cdot))$.*

(Заметим, что U и V - открытые $\Rightarrow$ интеграл определен и для x, близких к $\hat{x}$). Слабый максимум определяется аналогично. Получим необходимое условие экстремума:

Теорема 3.1 *Пусть $L, L_x, L_{\dot{x}}$ - непрерывные, l - непрерывно дифференцируема. Пусть $\hat{x}(\cdot)$ - слабый экстремум в (ЗБ). Тогда*

- $\forall t \in [t_0, t_1]$ $\frac{d}{dt}\hat{L}_{\dot{x}}(t) = \hat{L}_x(t)$
- *условия трансверсальности:* $\hat{L}_{\dot{x}}(t_0) = \hat{l}_{x(t_0)}$, $\hat{L}_{\dot{x}}(t_1) = -\hat{l}_{x(t_1)}$

(В векторном случае: $\forall i$ $\hat{L}_{\dot{x}_i}(t_0) = \hat{l}_{x_i(t_0)}$, $\hat{L}_{\dot{x}_i}(t_1) = -\hat{l}_{x_i(t_1)}$)

$\square$ Доказываем только для $n = 1$. Пусть $x_0 = \hat{x}(t_0), x_1 = \hat{x}(t_1), \hat{x}(\cdot)$ доставляет слабый минимум в (ЗБ). Рассмотрим снова задачу (З):

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

Заметим, что и в этой задаче $\hat{x}(\cdot)$ доставляет слабый минимум. Действительно, пусть $\|x(\cdot) - \hat{x}(\cdot)\|_{C^1} < \epsilon, x(t_0) = x_0, x(t_1) = x_1$. Тогда $J(x(\cdot)) - J(\hat{x}(\cdot)) = (B(x(\cdot)) - l(x_0, x_1)) - (B(\hat{x}(\cdot)) - l(x_0, x_1)) = B(x(\cdot)) - B(\hat{x}(\cdot)) \geq 0$.

Таким образом уравнение Эйлера выполнено. Осталось проверить условия трансверсальности. Пусть $h(\cdot) \in C^1[t_0, t_1]$.

$$\delta B(\hat{x}(\cdot), h(\cdot)) = \lim_{\alpha \to 0} \frac{1}{\alpha} \left(\int_{t_0}^{t_1} L(t, \hat{x}(t) + \alpha h(t), \dot{\hat{x}}(t) + \alpha \dot{h}(t) \right) dt +$$

$$+ l(\hat{x}(t_0) + \alpha h(t_0), \hat{x}(t_1) + \alpha h(t_1)) - \int_{t_0}^{t_1} \hat{L}(t)dt - l(\hat{x}(t_0), \hat{x}(t_1))) =$$

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

$$+\lim_{\alpha\to 0}\frac{1}{\alpha}(l(\hat{x}(t_0)+\alpha h(t_0),\hat{x}(t_1)+\alpha h(t_1))-l(\hat{x}(t_0),\hat{x}(t_1)))=$$

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

по теореме Ферма ($\forall h(\cdot)\in C^1[t_0,t_1]$). Проинтегрируем по частям: $\int_{t_0}^{t_1}\hat{L}_{\dot{x}}(t)\dot{h}(t)dt=\{\hat{L}_{\dot{x}}(t)$ непрерывно дифференцируема в силу формулы Эйлера $\}=\hat{L}_{\dot{x}}(t)h(t)|_{t_0}^{t_1}-\int_{t_0}^{t_1}\frac{d}{dt}\hat{L}_{\dot{x}}(t)h(t)dt$. Подставляя в равенство выше, получаем:

$$\hat{L}_{\dot{x}}(t)h(t)|_{t_0}^{t_1}+\hat{l}_{x(t_0)}h(t_0)+\hat{l}_{x(t_1)}h(t_1)=0.$$

$$-\hat{L}_{\dot{x}}(t_0)h(t_0)+\hat{L}_{\dot{x}}(t_1)h(t_1)+\hat{l}_{x(t_0)}h(t_0)+\hat{l}_{x(t_1)}h(t_1)=0\Leftrightarrow$$

$$\Leftrightarrow(\hat{l}_{x(t_0)}-\hat{L}_{\dot{x}}(t_0))h(t_0)+(\hat{L}_{\dot{x}}(t_1)+\hat{l}_{x(t_1)})h(t_1)=0$$

Возьмем теперь произвольную h, удовлетворяющую условиям

1. $h(t_1)=0$; $h(t_0)=1$. Тогда $\hat{L}_{\dot{x}}(t_0)=\hat{l}_{x(t_0)}$;
2. $h(t_0)=0$; $h(t_1)=1$. Тогда $\hat{L}_{\dot{x}}(t_1)=-\hat{l}_{x(t_1)}$.

На этом мы заканчиваем доказательство теоремы. ■

4 Интегралы импульса и энергии.

Пусть $L(t,x,\dot{x})=L(t,\dot{x})\Rightarrow$ уравнение Эйлера приобретает вид:

$$\frac{d}{dt}\hat{L}_{\dot{x}}(t)=0\Leftrightarrow\hat{L}_{\dot{x}}(t)\equiv const.$$

Введем обозначение: $p=\hat{L}_{\dot{x}}(t)$. Тогда $p(t)\equiv const$ - *интеграл импульса.*Теперь рассмотрим ситуацию вида: $L(t,x,\dot{x})=L(x,\dot{x})$ т.е нет зависимости от времени, $\hat{x}(\cdot)\in C^2[t_0,t_1]$. Тогда *интеграл энергии* $H(t)=\hat{L}_{\dot{x}}\dot{\hat{x}}-L\equiv const$. Покажем это:

□

$$\frac{d}{dt}H(t)=\frac{d}{dt}\left(\hat{L}_{\dot{x}}\dot{\hat{x}}-L(\hat{x},\dot{\hat{x}})\right)=$$

$$=\left(\frac{d}{dt}\hat{L}_{\dot{x}}\right)\dot{\hat{x}}+\hat{L}_{\dot{x}}\ddot{\hat{x}}-\hat{L}_x\dot{\hat{x}}-\hat{L}_{\dot{x}}\ddot{\hat{x}}=\left(\frac{d}{dt}\hat{L}_{\dot{x}}-\hat{L}_x\right)\dot{\hat{x}}=0■$$

Уравнение $H(t)\equiv const$ равносильно уравнению Эйлера, если только $\dot{\hat{x}}(t)\neq 0$.

Векторный случай аналогичен рассмотренному выше, только теперь

$$\hat{L}_{\dot{x}}\dot{\hat{x}}=\left(\hat{L}_{\dot{x}_1},\ldots,\hat{L}_{\dot{x}_n}\right)\begin{pmatrix}\dot{\hat{x}}_1\\ \vdots\\ \dot{\hat{x}}_n\end{pmatrix}$$

Заметим, что в векторном случае условие $H(t)\equiv const$ дает одно скалярное дифференциальное уравнение, в то время как векторное уравнение Эйлера есть система n скалярных уравнений.

Задача 2 *Положив $n = 1$ и $\hat{x} \in C^2[t_0, t_1]$, привести пример $L(x, \dot{x})$ такой, что $H(t) = const$, $\hat{x}(\cdot)$ - не экстремаль.*

Рассматривая интеграл энергии, мы полагаем, что $\hat{x}(\cdot) \in C^2[t_0, t_1]$, тогда как экстремали ищутся в более широком классе $C^1[t_0, t_1]$. Покажем обоснованность такого допущения:

Теорема 4.1 *Пусть $L = L(x, \dot{x}) \in C^2$, $\forall t$ $\hat{L}_{\dot{x}\dot{x}} > 0$ или $\forall t$ $\hat{L}_{\dot{x}\dot{x}} < 0$ (это* ***усиленное условие Лежандра****) и $\hat{x}(\cdot)$ - экстремаль. Тогда $\hat{x}(\cdot) \in C^2$*

$\square$ $p := L_{\dot{x}}$, $p = F(x, \dot{x})$, $p(\cdot) \in C^1$. Далее имеем: $\forall t$ $\hat{F}_{\dot{x}}(t) = \hat{L}_{\dot{x}\dot{x}} > 0$ или $\hat{F}_{\dot{x}}(t) = \hat{L}_{\dot{x}\dot{x}} < 0$. Следовательно мы можем воспользоваться теоремой о неявной функции и получить, что $\dot{x} = G(x, p)$, $G \in C^1$, $\dot{\hat{x}}(t) = G\left(\hat{x}(t), \hat{L}_{\dot{x}}(t)\right)$. Заметим, что $\hat{x}(t) \in C^1$, $\hat{L}_{\dot{x}}(t) \in C^1$, последнее имеет место в силу уравнения Эйлера. Таким образом $G\left(x\hat{(t)}, \hat{L}_{\dot{x}}(t)\right) \in C^1 \Rightarrow \dot{\hat{x}}(\cdot) \in C^1$ и $\hat{x}(\cdot) \in C^2$. $\blacksquare$

В векторном случае усиленное условие Лежандра означает, что для любого $t \in [t_0, t_1]$ оператор $L_{\dot{x}\dot{x}}$ положительно определен ($\left\langle \hat{L}_{\dot{x}\dot{x}} u, u \right\rangle > 0$ $\forall u \neq 0$) или для любого $t \in [t_0, t_1]$ оператор $L_{\dot{x}\dot{x}}$ отрицательно определен ($\left\langle \hat{L}_{\dot{x}\dot{x}} u, u \right\rangle < 0$ $\forall u \neq 0$). Первое неравенство в векторном виде записывается следующим образом:

$$\left\langle \begin{pmatrix} \hat{L}_{\dot{x}_1\dot{x}_1} & \dots & \hat{L}_{\dot{x}_1\dot{x}_n} \\ \vdots & & \vdots \\ \hat{L}_{\dot{x}_n\dot{x}_1} & \dots & \hat{L}_{\dot{x}_n\dot{x}_n} \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \right\rangle > 0$$

И в векторном случае если выполнено усиленное условие Лежандра, то экстремаль есть функция класса C^2.

5 Вариация интегрального функционала с подвижными концами.

Рекомендуемая литература: М. И. Зеликин "Оптимальное управление и вариационное исчисление."

Рассмотрим семейство функций вида: $\{x(t, \alpha) : t \in \triangle, \alpha \in (-\alpha_0, \alpha_0)\}$ и положим, что x, $\frac{\partial x}{\partial t}$, $\frac{\partial x}{\partial \alpha}$, $\frac{\partial^2 x}{\partial t \partial \alpha}$ $\exists$ и непрерывны. $t_0(\alpha), t_1(\alpha) \in (C^1[-\alpha_0, \alpha_0], \triangle)$.

Теперь можно определить интегральный функционал, зависящий от параметра α:

$$J(\alpha) = \int\limits_{t_0(\alpha)}^{t_1(\alpha)} L\left(t, x(t, \alpha), \dot{x}(t, \alpha)\right) dt$$

Предполагается также, что L, L_x, $L_{\dot{x}}$ - непрерывны и область определения L такая, что подинтегральное выражение определено для любого α.

Введем некоторые вспомогательные обозначения:

$$\hat{x}(t) := x(t, 0); \hat{t}_0 := t_0(0); \hat{t}_1 := t_1(0); x_0(\alpha) := x(t_0(\alpha), \alpha); x_1(\alpha) := x(t_1(\alpha), \alpha);$$

$$\hat{x}_0 := x_0(0); \hat{x}_1 := x_1(0); h(t) = \left.\frac{\partial x(t,\alpha)}{\partial \alpha}\right|_{\alpha=0}$$

Мы имеем

$$\dot{h}(t) = \left.\frac{\partial^2 x(t,\alpha)}{\partial\alpha\partial t}\right|_{\alpha=0} = \left.\frac{\partial^2 x(t,\alpha)}{\partial t\partial\alpha}\right|_{\alpha=0} = \left.\frac{\partial \dot{x}(t,\alpha)}{\partial \alpha}\right|_{\alpha=0}$$

Предположим теперь, что $\hat{x}$ - экстремаль. Интегрируя по частям так же, как мы делали при рассмотрении задачи Больца, получаем

$$J'(0) = \int\limits_{\hat{t}_0}^{\hat{t}_1} \left(\hat{L}_x(t)h(t) + \hat{L}_{\dot{x}}(t)\dot{h}(t)\right) dt + \hat{L}(\hat{t}_1)t_1'(0) - \hat{L}(\hat{t}_0)t_0'(0) =$$

$$= \hat{L}_{\dot{x}}(\hat{t}_1)h(\hat{t}_1) - \hat{L}_{\dot{x}}(\hat{t}_0)h(\hat{t}_0) + \hat{L}(\hat{t}_1)t_1'(0) - \hat{L}(\hat{t}_0)t_0'(0)$$

Заметим, что $x_i'(0) = \frac{d}{d\alpha}\left.(x(t_i(\alpha),\alpha))\right|_{\alpha=0} = \dot{\hat{x}}(\hat{t}_i)t_i'(0) + h(\hat{t}_i)$, где $i = 0, 1$

Следовательно $h(\hat{t}_i) = x_i'(0) - \dot{\hat{x}}(\hat{t}_i)t_i'(0) \Rightarrow$

$$J'(0) = \hat{L}_{\dot{x}}(\hat{t}_1)x_1'(0) - \hat{L}_{\dot{x}}(\hat{t}_1)\dot{\hat{x}}(\hat{t}_1)t_1'(0) - \hat{L}_{\dot{x}}(\hat{t}_0)x_0'(0) + \hat{L}_{\dot{x}}(\hat{t}_0)\dot{\hat{x}}(\hat{t}_0)t_0'(0) + \hat{L}(\hat{t}_1)t_1'(0) - \hat{L}(\hat{t}_0)t_0'(0) =$$

$$= \hat{L}_{\dot{x}}(\hat{t}_1)x_1'(0) - \hat{L}_{\dot{x}}(\hat{t}_0)x_0'(0) - \left(\hat{L}_{\dot{x}}(\hat{t}_1)\dot{\hat{x}}(\hat{t}_1) - \hat{L}(\hat{t}_1)\right)t_1'(0) + \left(\hat{L}_{\dot{x}}(\hat{t}_0)\dot{\hat{x}}(\hat{t}_0) - \hat{L}(\hat{t}_0)\right)t_0'(0) =$$

$$= p(\hat{t}_i)x_i'(0)\Big|_{i=0}^{1} - H(\hat{t}_i)t_i'(0)\Big|_{i=0}^{1}$$

Тем самым доказана теорема о вариации интегрального функционала, что в более короткой записи имеет вид:

$$dJ = pdx|_{\hat{t}_0}^{\hat{t}_1} - Hdt|_{\hat{t}_0}^{\hat{t}_1}.$$

6 Сильный экстремум в простейшей задаче классического вариационного исчисления. Теорема Вейерштрасса-Эрдмана.

Определение 6.1 $PC^1([t_0, t_1], \mathbb{R}^n)$ *- множество функций* $x : [t_0, t_1] \to \mathbb{R}^n$ *таких, что*

1. $x(\cdot)$ *- непрерывна.*
2. $\dot{x}(\cdot)$ *- существует и непрерывна везде за исключением, быть может, конечного числа точек.*
3. *точки несуществования* $\dot{x}(\cdot)$ *- точки разрыва первого рода.*

Легко заметить, что $C \subset PC^1$

Рассмотрим простейшую задачу классического вариационного исчисления для функций $x(\cdot) \in PC^1$:

$$J(x(\cdot)) = \int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \longrightarrow extr$$

$x(t_0) = x_0$, $x(t_1) = x_1$, $L : U \to \mathbb{R}$, $U = intU \subset \mathbb{R}^{2n+1}$ (З)

Определение 6.2 $x(\cdot) \in PC^1([t_0, t_1], \mathbb{R}^n)$ *- допустимая, если* $\forall t \in [t_0, t_1]$ $(t, x(t), \dot{x}(t)) \in U$

Определение 6.3 $\hat{x}(\cdot)$ *- сильный минимум, если* $\hat{x}(\cdot)$ *- допустимая и* $\exists \varepsilon > 0$ $\forall$ *допустимой* $x(\cdot) \in PC^1([t_0, t_1], \mathbb{R}^n)$ *такой, что* $||x(\cdot) - \hat{x}(\cdot)||_{C([t_0,t_1],\mathbb{R}^n)} < \varepsilon$ *верно* $J(x(\cdot)) \geq J(\hat{x}(\cdot))$.

Замечание 6.1 *Если функция* $\hat{x}(\cdot) \in C^1[t_0, t_1]$ *доставляет сильный экстремум, то она доставляет и слабый экстремум.*

$\square$ $||\cdot||_{C^1} \geq ||\cdot||_C$, т.е $||x(\cdot) - \hat{x}(\cdot)||_{C^1} < \varepsilon \Rightarrow ||x(\cdot) - \hat{x}(\cdot)||_C < \varepsilon$ $\blacksquare$

Теорема 6.1 *Пусть* $t_0 \leq \tilde{t}_0 < \tilde{t}_1 \leq t_1$

$$\tilde{J}(x(\cdot)) = \int\limits_{\tilde{t}_0}^{\tilde{t}_1} L(t, x(t), \dot{x}(t))dt \to extr$$

$x(\tilde{t}_0) = \tilde{x}_0$; $x(\tilde{t}_1) = \tilde{x}_1$ *(З2)*

Пусть $\hat{x}(\cdot)$ *- сильный экстремум в задаче (З) и* $x(\tilde{t}_0) = \tilde{x}_0$; $x(\tilde{t}_1) = \tilde{x}_1$. *Тогда* $\hat{x}(\cdot)$ *- сильный экстремум в (З2)*

$\square$ Положим для определенности, что $\hat{x}(\cdot)$ - сильный минимум в задаче (З). Покажем, что $\exists \varepsilon > 0 : \tilde{J}(\tilde{x}(\cdot)) \geq \tilde{J}(\hat{x}(\cdot))$ $\forall$ допустимой $\tilde{x}(\cdot) \in PC^1\left([\tilde{t}_0, \tilde{t}_1], \mathbb{R}^n\right)$ такой, что $||\tilde{x}(\cdot) - \hat{x}(\cdot)||_{C\left([\tilde{t}_0,\tilde{t}_1],\mathbb{R}^n\right)} < \varepsilon$.

Положим

$$x(t) = \begin{cases} \hat{x}(t), & t \in [t_0, \tilde{t}_0] \text{ или } t \in [t_1, \tilde{t}_1]; \\ \tilde{x}(t), & t \in [\tilde{t}_0, \tilde{t}_1]. \end{cases}$$

При этом

$$||x(\cdot) - \hat{x}(\cdot)||_{C([t_0,t_1],\mathbb{R}^n)} = max_{t\in[t_0,t_1]}|x(t) - \hat{x}(t)| = max_{t\in[\tilde{t}_0,\tilde{t}_1]}|x(t) - \hat{x}(t)| < \varepsilon$$

Следовательно $J(x(\cdot)) \geq J(\hat{x}(\cdot))$

$$\tilde{J}(\hat{x}(\cdot)) = J(\hat{x}(\cdot)) - \int\limits_{t_0}^{\tilde{t}_1} L(t, \hat{x}(t), \dot{\hat{x}}(t))dt - \int\limits_{\tilde{t}_1}^{t_1} L(t, \hat{x}(t), \dot{\hat{x}}(t))dt$$

$$\tilde{J}(\tilde{x}(\cdot)) = J(x(\cdot)) - \int\limits_{t_0}^{\tilde{t}_1} L(t, \hat{x}(t), \dot{\hat{x}}(t))dt - \int\limits_{\tilde{t}_1}^{t_1} L(t, \hat{x}(t), \dot{\hat{x}}(t))dt$$

Значит, $\tilde{J}(\tilde{x}(\cdot)) \geq \tilde{J}(\hat{x}(\cdot))$ Следовательно, $\hat{x}(\cdot)$ - сильный минимум в (З2). ■

Итак, доказаны 2 свойства:

1. $\hat{x} \in C^{'}[t_0, t_1]$ - сильный экстремум $\Rightarrow \hat{x}$ - слабый экстремум.
2. $\hat{x} \in PC^{'}[t_0, t_1]$ - сильный экстремум и $t_0 \leq \tilde{t}_0 < \tilde{t}_1 \leq t_1 \Rightarrow \hat{x}$ - сильный экстремум на $[\tilde{t}_0, \tilde{t}_1]$.

Теорема 6.2 *Пусть $\hat{x}$ - сильный экстремум $\Rightarrow \forall t \in (t_0, t_1)$*

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

(Здесь имеет место раздвоение уравнения Эйлера.)

□ $\hat{x}(\cdot) \in PC^1 \Rightarrow \exists\ \tilde{t}_0 < t : \hat{x}(\cdot) \in C^{'}[\tilde{t}_0, t]$. Рассмотрим задачу на $[\tilde{t}_0, t]$. $\hat{x}(\cdot)$ доставляет сильный экстремум в силу (2), следовательно, по (1) он доставляет и слабый экстремум, откуда, как мы знаем, следует уравнение Эйлера. Для предела справа рассуждения аналогичны. ■

Рассмотрим $P = \hat{L}_{\dot{x}}$ и $H = \hat{L}_{\dot{x}}\dot{\hat{x}} - L$.

Теорема 6.3 *(Вейерштрасса-Эрдмана) Пусть L, L_x, $L_{\dot{x}}$ - непрерывны и $\hat{x}(\cdot) \in PC^{'}[t_0, t_1]$ - доставляет сильный экстремум. $\Rightarrow$ Функции $P(\cdot)$, $H(\cdot) \in C[t_0, t_1]$*

□ Заметим, что в точках непрерывности $\dot{\hat{x}}(\cdot)$ это утверждение очевидно. При доказательстве $H(\cdot) \in C[t_0, t_1]$ будем предполагать, что $\hat{x}(\cdot) \in C^2$ левее и правее от рассматриваемой точки.

Положим $\hat{x}(\cdot) \in C^1[t_0, \tau]$, $\hat{x}(\cdot) \in C^1[\tau, t_1]$, т.е. τ - единственная точка разрыва производной, и рассмотрим два семейства функций:

$\left\{x^{(l)}(t, \alpha)\right\}$ – левое семейство функций (left).
$\left\{x^{(r)}(t, \alpha)\right\}$ – правое семейство функций (right).

Введем обозначения: $t_0^{(l)}(\alpha) = t_0$; $t_1^{(l)}(\alpha) = \tau + \alpha$; $x_0^{(l)}(\alpha) = x_0$; $x_1^{(l)}(\alpha) = \hat{x}(\tau)$; $x^{(l)}(t, 0) = \hat{x}(t), t \in [t_0, \tau]$. Аналогично для правого семейства $t_0^{(r)}(\alpha) = \tau + \alpha$; $t_1^{(r)}(\alpha) = t_1$; $x_0^{(r)}(\alpha) = \hat{x}(\tau)$; $x_1^{(r)}(\alpha) = x_1$; $x^{(r)}(t, 1) = \hat{x}(t), t \in [\tau, t_1]$ (см. рис 1а). Мы требуем чтобы у $x^{(\cdot)}(t, \alpha)$ были непрерывные частные и смешанные производные и, значит, для них выполняется теорема о вариации интегрального функционала.

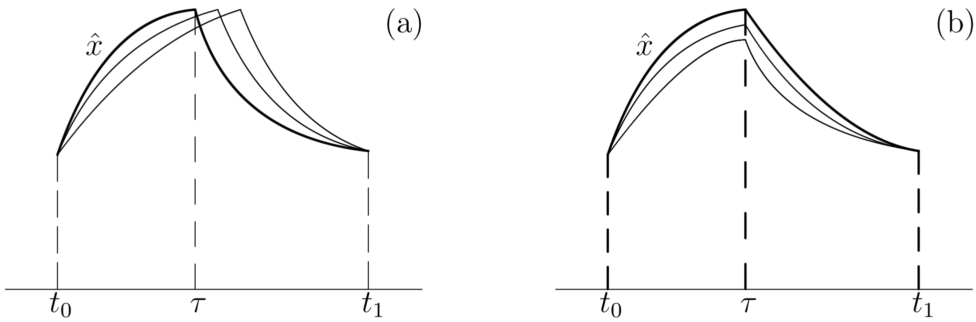


Рис. 1: Семейства функций для $P(\cdot)$ (a) и $H(\cdot)$ (b)

Теперь рассмотрим два интеграла:

$$J_{(l)}(\alpha) = \int\limits_{t_0^{(l)}(\alpha)}^{t_1^{(l)}(\alpha)} L(t, x^{(l)}(t,\alpha), \dot{x}^{(l)}(t,\alpha))dt;$$

$$J_{(r)}(\alpha) = \int\limits_{t_0^{(r)}(\alpha)}^{t_1^{(r)}(\alpha)} L(t, x^{(r)}(t,\alpha), \dot{x}^{(r)}(t,\alpha))dt;$$

Т.к. соответствующие абсциссы и ординаты совпадают, то семейства кривых "left" и "right" склеиваются, поэтому мы можем рассмотреть новое семейство вида:

$$x(t,\alpha) = \begin{cases} x^{(l)}(t,\alpha), \text{если } t \le \tau+\alpha; \\ x^{(r)}(t,\alpha), \text{если } t \ge \tau+\alpha; \end{cases}$$

При этом $\forall \alpha\ x(t,\alpha) \in PC^{'}[t_0, t_1]$ и $\|x(t,\alpha) - \hat{x}(t)\|_C \to 0$ при $\alpha \to 0$. Тогда

$$J_{(l)}(\alpha) + J_{(r)}(\alpha) = \int\limits_{t_0}^{t_1} L(t, x(t,\alpha), \dot{x}(t,\alpha))dt = J(x(\cdot,\alpha)) = F(\alpha)$$

Если $\hat{x}(\cdot)$ доставляет сильный экстремум, то функция $F(\alpha)$ достигает экстремума при $\alpha = 0$, следовательно по теореме Ферма $F^{'}(0) = 0 \Rightarrow F^{'}(0) = J^{'}_{(l)}(0) + J^{'}_{(r)}(0) = \{$по теореме о вариации интегрального функционала$\} = -H(\tau-0) + H(\tau+0)$. Итак, $H(\tau-0) = H(\tau+0)$ и непрерывность H показана. Правда пока неясно, существуют ли рассмотренные семейтсва "left" и "right". Зададим их явным образом:

$$x^{(l)}(t,\alpha) = \hat{x}(t) - \frac{t-t_0}{\tau+\alpha-t_0}\left(\hat{x}(\tau+\alpha) - \hat{x}(\tau)\right)$$

$$x^{(l)}(\tau+\alpha,\alpha) = \hat{x}(\tau+\alpha) - \frac{\tau+\alpha-t_0}{\tau+\alpha-t_0}\left(\hat{x}(\tau+\alpha) - \hat{x}(\tau)\right) = \hat{x}(\tau)$$

Легко проверить, что это семейство удовлетворяет всем условиям на гладкость. Здесь-то и понадобится условие $\hat{x} \in C^2$. Аналогично строится и семейство "right".

Следовательно мы показали, что для $\hat{x}(\cdot) \in C^2[t_0, \tau]$, $\hat{x}(\cdot) \in C^2[\tau, t_1]$ $H(\tau - 0) = H(\tau + 0)$.

Предположим теперь, что $\hat{x}(\cdot) \in C^1[t_0, \tau]$, $\hat{x}(\cdot) \in C^1[\tau, t_1]$ и докажем, что $p(\tau - 0) = p(\tau + 0)$. Снова будем строить 2 семейства кривых (рис. 1б): $t_0^{(l)}(\alpha) = t_0$; $t_1^{(l)}(\alpha) = \tau$; $x_0^{(l)}(\alpha) = x_0$; $x_1^{(l)}(\alpha) = \hat{x}(\tau) + \alpha$; $x^{(l)}(t, 0) = \hat{x}(t), t \in [t_0, \tau]$ и $t_0^{(r)}(\alpha) = \tau$; $t_1^{(r)}(\alpha) = t_1$; $x_0^{(r)}(\alpha) = \hat{x}(\tau) + \alpha$; $x_1^{(r)}(\alpha) = x_1$; $x^{(r)}(t, 0) = \hat{x}(t), t \in [\tau, t_1]$. Тогда получаем

$$0 = F^{'}(0) = p(\tau - 0) - p(\tau + 0)$$

$\Rightarrow p(\tau - 0) = p(\tau + 0)$. В явном виде семейство "left" задается следующим образом:

$$x^{(l)} = \hat{x}(t) + \alpha \frac{t - t_0}{\tau - t_0}.$$

Общий случай: пусть τ - не единственная точка разрыва производной. $\Rightarrow$ $\exists \left[\tilde{t}_0, \tilde{t}_1\right] \subset [t_0, t_1] : \hat{x}(\cdot) \in C^1\left[\tilde{t}_0, \tau\right]$, $\hat{x}(\cdot) \in C^1\left[\tau, \tilde{t}_1\right]$. $\hat{x}$ доставляет сильный экстремум на отрезке $\left[\tilde{t}_0, \tilde{t}_1\right] \Rightarrow$ задача сведена к уже разобранной выше. ■

Замечание 6.2 *Эта теорема работает и доказывается аналогично и для векторных функций.*

7 Необходимые и достаточные условия второго порядка для слабого экстремума в простейшей задаче классического вариационного исчисления.

Пусть L, L_x, $L_{\dot{x}}$, L_{xx}, $L_{x\dot{x}}$, $L_{\dot{x}\dot{x}}$ - непрерывны в $U \subset \mathbb{R}^3$ и $L = L(t, x, \dot{x})$. Рассмотрим задачу:

$$J(x(\cdot)) = \int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \longrightarrow min(max)$$

$$x(t_0) = x_0; x(t_1) = x_1;$$

Пусть $\hat{x}$ - слабый минимум. $h(\cdot) \in C_0^1[t_0, t_1]$ т.е $h(\cdot) \in C^1$ и $h(t_0) = h(t_1) = 0$.

Рассмотрим $J(\hat{x}(\cdot) + \alpha h(\cdot)) = F(\alpha)$, разложим L в ряд Тейлора:

$$L\left(t, \hat{x}(t) + \alpha h(t), \dot{\hat{x}}(t) + \alpha \dot{h}(t)\right) = \hat{L}(t) + \alpha \hat{L}_x(t)h(t) + \alpha \hat{L}_{\dot{x}}(t)\dot{h}(t) +$$

$$+ \frac{\alpha^2}{2}\hat{L}_{xx}(t)h^2(t) + \alpha^2 \hat{L}_{\dot{x}x}(t)h(t)\dot{h}(t) + \frac{\alpha^2}{2}\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + \underbrace{o(\alpha^2)}_{\text{равномерно по } t}$$

Следовательно

$$F(\alpha) = \int\limits_{t_0}^{t_1} \hat{L}(t)dt + \alpha \int\limits_{t_0}^{t_1} \left(\hat{L}_x(t)h(t) + \hat{L}_{\dot{x}}(t)\dot{h}(t)\right)dt +$$

$$+\frac{\alpha^2}{2}\int\limits_{t_0}^{t_1}\left(\hat{L}_{xx}(t)h^2(t)+2\hat{L}_{\dot{x}x}(t)h(t)\dot{h}(t)+\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t)\right)dt+o(\alpha^2)$$

Как мы знаем, $\int\limits_{t_0}^{t_1}\left(\hat{L}_x(t)h(t)+\hat{L}_{\dot{x}}(t)\dot{h}(t)\right)dt=\delta J(\hat{x}(\cdot),h(\cdot))=0$

Введем обозначения: $A(t):=\hat{L}_{xx}(t)$; $B(t):=\hat{L}_{\dot{x}x}(t)$; $C(t):=\hat{L}_{\dot{x}\dot{x}}$;

$$K(h(\cdot))=\int\limits_{t_0}^{t_1}\left(Ah^2+2Bh\dot{h}+C\dot{h}^2\right)dt$$

$\Rightarrow F(\alpha)=F(0)+\frac{\alpha^2}{2}K(h(\cdot))+o(\alpha^2)\Rightarrow\lim\limits_{\alpha\to 0}\frac{F(\alpha)-F(0)}{\alpha^2}=\frac{1}{2}K(h(\cdot))$

$\hat{x}(\cdot)$ - слабый минимум $\Rightarrow F(\alpha)\geq F(0)\Rightarrow$

$$0\leq\lim_{\alpha\to 0}\frac{F(\alpha)-F(0)}{\alpha^2}=\frac{1}{2}K(h(\cdot))$$

Мы получили необходимое условие слабого минимума (максимума):

$\forall h(\cdot)\in C_0^1[t_0,t_1]$ $K(h(\cdot))\geq 0$ (соответственно $K(h(\cdot))\leq 0$). Но это условие сложно проверять, позже получим более удобные условия.

Мы будем использовать следующую идею: рассмотрим функцию $\omega(\cdot)\in C^1[t_0,t_1]$

$$\int\limits_{t_0}^{t_1}\left(\omega(t)h^2(t)\right)'dt=\omega(t_1)h^2(t_1)-\omega(t_0)h^2(t_0)=0$$

$\Rightarrow$

$$\int\limits_{t_0}^{t_1}\left(\omega(t)h^2(t)\right)'dt=\int\limits_{t_0}^{t_1}\left(\dot{\omega}(t)h^2(t)+2\omega(t)h(t)\dot{h}(t)\right)dt=0$$

Следовательно, необходимое условие можно переписать следующим образом:

$$K(h(\cdot))=\int\limits_{t_0}^{t_1}\left(\left(A(t)+\dot{\omega}(t)\right)h^2(t)+2\left(B(t)+\omega(t)\right)h(t)\dot{h}(t)+C(t)\dot{h}^2(t)\right)dt\geq 0$$

Мы можем выбирать $w(\cdot)\in C^1[t_0,t_1]$ так, как нам нравится $\Rightarrow$ подберем $w(\cdot)$ так, чтобы $(B(t)+w(t))^2=(A(t)+\dot{w}(t))C(t)$ $(C(t)\neq 0)$. Это выполняется тогда и только тогда, когда $w(\cdot)$ - решение уравнения Риккати

$$\dot{w}(t)=\frac{(B(t)+w(t))^2}{C(t)}-A(t)$$

Тогда, как несложно проверить,

$$K(h(\cdot))=\int_{t_0}^{t_1}C(t)\left(\frac{B(t)+w(t)}{C(t)}h(t)+\dot{h}(t)\right)^2dt$$

Однако, $C(t)$ может обращаться в нуль на $[t_0,t_1]$; даже если этого не происходит, то решение уравнения Риккати на всем отрезке $[t_0,t_1]$ может не существовать. Эффективным оказывается использование локального решения уравнения Риккати.

Теорема 7.1 $\hat{x}(\cdot)$ *- слабый минимум (слабый максимум)* $\Rightarrow \forall t \in [t_0,t_1]$ $\hat{L}_{\dot{x}\dot{x}}(t) \geq 0$ *(*$\hat{L}_{\dot{x}\dot{x}}(t) \leq 0$ *соответственно) (условие Лежандра)*

$\square$ Доказываем теорему для минимума. Пусть существует $\tau \in [t_0,t_1] : \hat{L}_{\dot{x}\dot{x}}(t) < 0$. Тогда существует и некоторая окрестность $V \ni \tau$: $\forall t \in V$ $C(t) = \hat{L}_{\dot{x}\dot{x}}(t) < 0$. В этой окрестности имеет смысл уравнение Риккати. Более того, по теореме существования из диффуров $\exists (\tilde{t}_0,\tilde{t}_1) : \tau \in (\tilde{t}_0,\tilde{t}_1)$: на $[\tilde{t}_0,\tilde{t}_1]$ существует решение уравнения Риккати. Тогда рассмотрим $h(\cdot) \in C_0^1[t_0,t_1]$: $h(t) \equiv 0$ при $t \leq \tilde{t}_0, t \geq \tilde{t}_1$. Имеем $C(t) < 0$ и

$$0 \leq K(h(\cdot)) = \int_{\tilde{t}_0}^{\tilde{t}_1} C(t) \left(\frac{B(t)+w(t)}{C(t)} h(t) + \dot{h}(t) \right)^2 dt \leq 0$$

Т.о. этот интеграл равен 0 и, значит, подинтегральное выражение равно 0. Следовательно, $-\frac{A(t)+w(t)}{C(t)} h(t) = \dot{h}(t)$ на $[\tilde{t}_0,\tilde{t}_1]$. Но произвольная h не обязана удовлетворять такому дифференциальному уравнению. Противоречие. ■

Запишем уравнение Эйлера для функционала $K(h(\cdot)) = \int_{t_0}^{t_1} (Ah^2 + 2Bh\dot{h} + C\dot{h}^2) dt$:

$$-\frac{d}{dt}\Big(2B(t)h(t) + 2C(t)\dot{h}(t)\Big) + 2A(t)h(t) + 2B(t)\dot{h}(t) = 0$$

(в векторном случае для последнего слагаемого имеем $B^T(t)$). Это уравнение называется уравнением Якоби. Пусть $h(\cdot)$ - его решение, $h(t_0) = 0, h(\cdot)$ - не тождественный ноль. Тогда

Определение 7.1 *точка* $\tau > t_0$ *- сопряженная (с точкой* t_0*), если* $h(\tau) = 0$

Докажем следующее необходимое условие слабого экстремума:

Теорема 7.2 *Пусть* $\dot{x}(\cdot)$ *- слабый минимум (максимум),* $\forall t \in [t_0,t_1]$ $\hat{L}_{\dot{x}\dot{x}}(t) > 0$ *(*$\hat{L}_{\dot{x}\dot{x}}(t) < 0$ *соответственно). Тогда* (t_0,t_1) *не содержит сопряженных точек.*

Аналогичная теорема верна и для многомерного случая. Правда тогда необходима проверка для произвольной функции $h(\cdot)$, тогда как в одномерном случае достаточно рассмотреть конкретную функцию в силу одномерности пространства решений дифуравнения Якоби.

$\square$ Снова доказываем для минимума и снова от противного. Пусть $\hat{h}(\cdot)$ - решение уравнения Якоби, $\hat{h}(t_0) = \hat{h}(\tau) = 0$, $\hat{h}(\cdot)$ - не тождественный ноль, $\tau \in (t_0,t_1)$. Положим

$$\tilde{K}(h(\cdot)) = \int_{t_0}^{\tau} (Ah^2 + 2Bh\dot{h} + C\dot{h}^2) dt$$

Заметим, что $\forall \alpha$ $\alpha\hat{h}(\cdot)$ является экстремалью для $\tilde{K}(\cdot)$, т.к. $\alpha\hat{h}(\cdot)$ - тоже решение уравнения Якоби. Значит, $\forall h(\cdot) \in C_0^1[t_0,\tau]$ $\delta\tilde{K}(\alpha\hat{h}(\cdot), h(\cdot)) = 0$. Положим $f(\alpha) = \tilde{K}(\alpha\hat{h}(\cdot))$. Тогда

$$f'(\alpha) = \lim_{\varepsilon \to 0} \frac{\tilde{K}(\alpha\hat{h}(\cdot) + \varepsilon\hat{h}(\cdot)) - \tilde{K}(\alpha\hat{h}(\cdot))}{\varepsilon} = \delta\tilde{K}(\alpha\hat{h}(\cdot), \hat{h}(\cdot)) = 0, f(0) = 0$$

Значит, $f \equiv 0$, $\tilde{K}(\hat{h}(\cdot)) = 0$ ■

Нам понадобится следующая лемма (о скруглении углов), которую мы приведем без доказательства:

Лемма 7.1 *рассмотрим задачу*

$$K(h(\cdot)) = \int_{t_0}^{t_1} M(t, h(t), \dot{h}(t))dt \to min$$

$$h(t_0) = h_0, h(t_1) = h_1, \tilde{h}(\cdot) \in C_0^1[t_0, t_1]$$

1) Если $M \in C([t_0, t_1] \times \mathbb{R}^2)$, $K(\tilde{h}(\cdot)) = \min\limits_{h(\cdot) \in C_0^1[t_0,t_1]} K(h(\cdot))$, *то*

$$K(\tilde{h}(\cdot)) = \min_{h(\cdot) \in PC_0^1[t_0,t_1]} K(h(\cdot))$$

2) Если U *— окрестность графика* $\{(t, \tilde{h}(t)) : t \in [t_0, t_1]\}$,

$$K(\tilde{h}(\cdot)) = \min_{\substack{h(\cdot) \in C_0^1[t_0,t_1] \\ \|h(\cdot) - \tilde{h}(\cdot)\|_{C[t_0,t_1]} < \varepsilon}} K(h(\cdot)), \text{ то}$$

$$K(\tilde{h}(\cdot)) = \min_{\substack{h(\cdot) \in PC_0^1[t_0,t_1] \\ \|h(\cdot) - \tilde{h}(\cdot)\|_{C[t_0,t_1]} < \varepsilon}} K(h(\cdot))$$

Из этой леммы получаем следующий полезный вывод: пусть $\tilde{h}(\cdot) \equiv 0$, $h(\cdot)$ доставляет слабый минимум, т.е $K(\tilde{h}(\cdot)) = \min\limits_{h(\cdot) \in C_0^1[t_0,t_1]} K(h(\cdot))$ Тогда

$$\min_{h(\cdot) \in PC_0^1[t_0,t_1]} K(h(\cdot)) = 0$$

Пусть теперь $\tilde{h}(t) = \begin{cases} \hat{h}(t), t \in [t_0, \tau] \\ 0, t > \tau \end{cases}$

Тогда и $K(\tilde{h}(\cdot)) = 0$, т.е. и на функции $\tilde{h}(\cdot)$ достигается сильный минимум.

$$K(h(\cdot)) = \int_{t_0}^{t_1} (A(t)h^2(t) + 2B(t)h(t)\dot{h}(t) + C(t)\dot{h}^2(t))dt$$

Мы имеем: $p(t) = 2B(t)\tilde{h}(t) + 2C(t)\dot{\tilde{h}}(t)$. $t > \tau \Rightarrow p(t) = 0$.

$$\lim_{t \to \tau, t < \tau} p(t) = 2B(\tau)\tilde{h}(\tau) + 2C(\tau)\dot{\tilde{h}}(\tau - 0) =$$

$$= \{\tilde{h} \text{ доставляет сильный минимум, по теореме Вейерштрасса-Эрдмана}\} =$$

$$= 2C(\tau)\dot{\hat{h}}(\tau) = 0$$

Значит, $\dot{\hat{h}}(\tau) = 0$ по условиям Лежандра. Но это невозможно, т.к. тогда $\hat{h}(\cdot) \equiv 0$ по теореме существования и единственности решения линейного однородного уравнения второго порядка.

Перейдем теперь к достаточным условиям.

Теорема 7.3 *Пусть $\hat{x}(\cdot)$ - допустимая экстремаль, $\forall t \in [t_0, t_1]$ $\hat{L}_{\dot{x}\dot{x}}(t) > 0$. Пусть на $(t_0, t_1]$ нет сопряженных точек. Тогда $\hat{x}(\cdot)$ доставляет слабый минимум.*

Заметим, что единственное отличие от необходимого условия - это то, что теперь требуется, чтобы и точка t_1 не была сопряженной.

■ Для доказательства потребуются следующие леммы:

Лемма 7.2 *Существует такое решение $h(\cdot)$ уравнения Якоби, что $\forall t \in [t_0, t_1]$ $h(t) > 0$*

□ Как мы знаем, существует решение h_0 уравнения Якоби: $\forall t \in (t_0, t_1]$ $h_0(t) > 0, h_0(t_0) = 0, \dot{h}_0(t_0) = 1$. Рассмотрим решение $h_1(\cdot)$ уравнения Якоби: $h_1(t_0) = 1, \dot{h}_1(t_0) = 0$. $\exists \tau \in (t_0, t_1) : \min\limits_{t \in [t_0, \tau]} h_1(t) > 0$. Положим $\delta = \min\limits_{t \in [\tau, t_1]} h_0(t) > 0$, $M = \max\limits_{t \in [\tau, t_1]} |h_1(t)|$. Выберем α так, чтобы $0 < \alpha < \frac{\delta}{M}$.

Тогда положим $h(t) = h_0(t) + \alpha h_1(t)$. При $t \in [t_0, \tau]$ имеем $h(t) > 0 + 0 = 0$, а при $t \in [\tau, t_1]$ $h(t) \geq h_0(t) - \alpha|h_1(t)| \geq \delta - \alpha M > 0$ ■

Найдем решение уравнения Риккати. По предыдущей лемме на всем отрезке определено отношение $\frac{\dot{h}(t)}{h(t)}$. Значит, имеет смысл следующая лемма:

Лемма 7.3 $w(t) = -B(t) - C(t)\frac{\dot{h}(t)}{h(t)}$ *- решение уравнения Риккати.*

□ Умножим равенство из условия слева и справа на $h(t)$:

$$w(t)h(t) = -B(t)h(t) - C(t)\dot{h}(t)$$

$$\frac{d}{dt}(w(t)h(t)) = \dot{w}(t)h(t) + w(t)\dot{h}(t)$$

С другой стороны в силу уравнения Якоби

$$\frac{d}{dt}(w(t)h(t)) = \frac{d}{dt}(-B(t)h(t) - C(t)\dot{h}(t)) = -A(t)h(t) - B(t)\dot{h}(t)$$

Приравнивая полученные выражения для $\frac{d}{dt}(w(t)h(t))$, получаем

$$(\dot{w}(t) + A(t))h(t) = (-w(t) - B(t))\dot{h}(t)$$

Следовательно

$$\dot{w}(t) = -\left(B(t)+w(t)\right)\frac{\dot{h}(t)}{h(t)} - A(t) = -(B(t)+w(t))\left(-\frac{B(t)+w(t)}{C(t)}\right) - A(t)$$

и, значит, $w(t)$ удовлетворяет уравнению Риккати. ■

Продолжим доказательство достаточного условия. Разложим $L(t,\hat{x}(t)+h(t),\dot{\hat{x}}(t)+\dot{h}(t))$ по формуле Тейлора:

$L(t,\hat{x}(t)+h(t),\dot{\hat{x}}(t)+\dot{h}(t)) = \hat{L}(t)+\hat{L}_x(t)h(t)+\hat{L}_{\dot{x}}(t)\dot{h}(t)+\frac{1}{2}\hat{L}_{xx}(t)h^2(t)+\hat{L}_{\dot{x}x}(t)h(t)\dot{h}(t)+\frac{1}{2}\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t)+o(h^2(t)+\dot{h}^2(t))$, последнее слагаемое равномерно по t при $h^2(t)+\dot{h}^2(t)\to 0$. Отсюда $L(t,\hat{x}(t)+h(t),\dot{\hat{x}}(t)+\dot{h}(t)) \geq \hat{L}(t)+\hat{L}_x(t)h(t)+\hat{L}_{\dot{x}}(t)\dot{h}(t)+\frac{1}{2}\hat{L}_{xx}(t)h^2(t)+\hat{L}_{\dot{x}x}(t)h(t)\dot{h}(t)+\frac{1}{2}\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t)-\delta(h^2(t)+\dot{h}^2(t))$ как только $\|h(\cdot)\|_{C^1[t_0,t_1]}<\varepsilon$

Итак, $J(\hat{x}(\cdot)+h(\cdot))-J(\hat{x}(\cdot))\geq$

$$\geq \frac{1}{2}\int_{t_0}^{t_1}(A(t)-2\delta)h^2(t)+2B(t)h(t)\dot{h}(t)+(C(t)-2\delta)\dot{h}^2(t)dt$$

Покажем, что это выражение неотрицательно. Вспоминая, что $\int_{t_0}^{t_1}\left(\dot{\omega}(t)h^2(t)+2\omega(t)h(t)\dot{h}(t)\right)dt=0$, получаем, что достаточно показать, что

$$\int_{t_0}^{t_1}(A(t)-2\delta+\dot{w}(t))h^2(t)+2(B(t)+w(t))h(t)\dot{h}(t)+(C(t)-2\delta)\dot{h}^2(t))dt\geq 0$$

Рессмотрим уравнение $\dot{w}(t)=\frac{(B(t)+w(t))^2}{C(t)-2\delta}-(A(t)-2\delta)$ — модификацию уравнения Риккати, пусть w - его решение. Тогда интеграл выше переписывается следующим образом:

$$\int_{t_0}^{t_1}(C(t)-2\delta)\left(\frac{A(t)-2\delta-\dot{w}(t)}{C(t)-2\delta}h(t)+\dot{h}(t)\right)^2dt$$

Несложно видеть, что при достаточно малых δ $(C(t)-2\delta \geq 0)$ и, следовательно, весь интеграл неотрицателен. Заметим, что существование решения у модифицированного уравнения Риккати для малых δ следует из существования решения у самого уравнения Риккати (которое соответствует значению $\delta=0$).

Итак, $J(\hat{x}(\cdot)+h(\cdot))-J(\hat{x}(\cdot))\geq 0$, и, значит, $\hat{x}$ действительно доставляет слабый минимум. Теорема доказана полностью. ■

8 Игольчатые вариации. Условие Вейерштрасса — необходимое условие сильного экстремума.

Пусть $V=intV\subset\mathbb{R}^{n+1}$, $L\in C\left(V\times\mathbb{R}^n\right)$; $\forall t\in[t_0,t_1]$ $(t,\hat{x}(t))\in V$, $\hat{x}(\cdot)\in C\left([t_0,t_1],\mathbb{R}\right)$ - экстремаль.

$$J(x(\cdot)) = \int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \to extr;$$

$x(t_0) = x_0$; $x(t_1) = x_1$;

Далее положим $n = 1$ и $\hat{x}(\cdot)$ - доставляет сильный минимум.

Теперь введем понятие игольчатой вариации: рассмотрим $\tau \in (t_0, t_1)$, $v \in \mathbb{R}$, $\alpha > 0$ - некий малый параметр. Пусть $u = \dot{\hat{x}}(\tau)$. Построим следующую функцию:

$$x(t, \alpha) = \begin{cases} \hat{x}(t), & t_0 \le t \le \tau - \alpha; \\ \hat{x}(t) + (v - u)(t - \tau + \alpha), & \tau - \alpha \le t \le \tau; \\ \hat{x}(t) + \frac{(v-u)\alpha}{t_1 - \tau}(t_1 - t), & \tau \le t \le t_1. \end{cases}$$

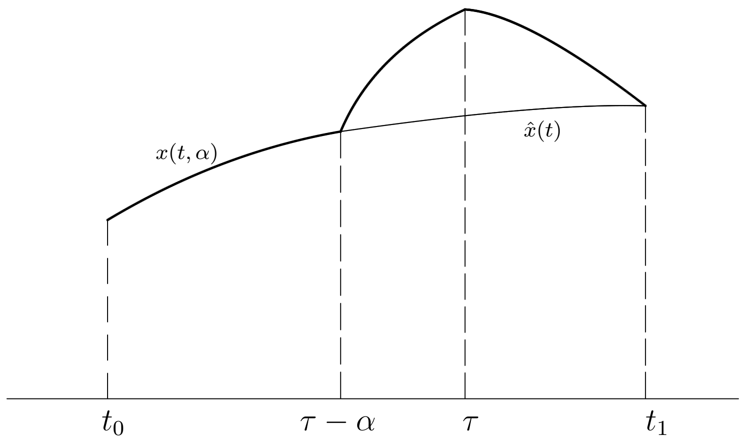


Рис. 2: Функция $x(t, \alpha)$

Заметим, что $\forall \alpha$ - это кусочно-непрерывно дифференцируемая функция. (причем значение этой функции в точке τ состыковано, т.е. $\hat{x}(t) + (v - u)(t - \tau + \alpha)|_{t=\tau} = \hat{x}(t) + \frac{(v-u)\alpha}{t_1 - \tau}(t_1 - t)|_{t=\tau} = \hat{x}(\tau) + (v - u)\alpha$.) По определению такого рода вариации называются игольчатыми.

Рассмотрим разность вида:

$$J(x(\cdot, \alpha)) - J(\hat{x}(\cdot)) = \underbrace{\int\limits_{t_0}^{\tau - \alpha} \Big(L(t, x(t, \alpha), \dot{x}(t, \alpha)) - L(t, \hat{x}(t), \dot{\hat{x}}(t)) \Big) dt}_{I_0} +$$

$$+ \underbrace{\int\limits_{\tau - \alpha}^{\tau} \Big(L(t, x(t, \alpha), \dot{x}(t, \alpha)) - L(t, \hat{x}(t), \dot{\hat{x}}(t)) \Big) dt}_{I_1} + \underbrace{\int\limits_{\tau}^{t_1} \Big(L(t, x(t, \alpha), \dot{x}(t, \alpha)) - L(t, \hat{x}(t), \dot{\hat{x}}(t)) \Big) dt}_{I_2}$$

Теперь оценим I_0, I_1, I_2.

$I_0 = 0$ т.к. по определению функции $x(t, \alpha)$, $L(t, x(t, \alpha), \dot{x}(t, \alpha)) - L(t, \hat{x}(t), \dot{\hat{x}}(t)) = L(t, \hat{x}(t), \dot{\hat{x}}(t)) - L(t, \hat{x}(t), \dot{\hat{x}}(t)) = 0$

Далее по формуле вариации интегрального функционала с подвижными концами имеем:

$$\frac{dI_2}{d\alpha}\bigg|_{\alpha=0} = p(t)\frac{dx}{d\alpha}\bigg|_{\alpha=0}\bigg|_{t=\tau}^{t=t_1} = -p(\tau)(v-u) \Rightarrow I_2 = -\hat{L}_{\dot{x}}(\tau)(v-u)\alpha + \bar{\bar{o}}(\alpha), \alpha \to 0$$

Воспользовавшись теоремой о среднем, получаем, что

$$I_1 = \alpha\left(L(\tilde{t}, x(\tilde{t},\alpha), \dot{x}(\tilde{t},\alpha)) - L(\tilde{t}, \hat{x}(\tilde{t}), \dot{\hat{x}}(\tilde{t}))\right), \tilde{t} \in (\tau-\alpha, \tau)$$

Замечаем, что $L(\tilde{t}, \hat{x}(\tilde{t}), \dot{\hat{x}}(\tilde{t})) \to L(\tau, \hat{x}(\tau), u), \alpha \to 0.$ и
$L(\tilde{t}, x(\tilde{t},\alpha), \dot{x}(\tilde{t},\alpha)) \to L(\tau, \hat{x}(\tau), v), \alpha \to 0.$

Следовательно, $I_1 = \alpha\left(L(\tau, \hat{x}(\tau), v) - L(\tau, \hat{x}(\tau), u)\right) + \bar{\bar{o}}(\alpha)$, при $\alpha \to 0 \Rightarrow$

$$J(x(\cdot,\alpha)) - J(\hat{x}(\cdot)) = \alpha\left(L(\tau, \hat{x}(\tau), v) - L(\tau, \hat{x}(\tau), u) - \hat{L}_{\dot{x}}(\tau)(v-u)\right) + \bar{\bar{o}}(\alpha), \alpha \to 0$$

Определение 8.1 *Функция Вейерштрасса:*

$$\mathcal{E}(t,x,u,v) = L(t,x,v) - L(t,x,u) - L_u(t,x,u)(v-u)$$

$$0 \le \lim_{\alpha\to 0+}\frac{J(x(\cdot,\alpha)) - J(\hat{x}(\cdot))}{\alpha} = L(\tau,\hat{x}(\tau),v) - L(\tau,\hat{x}(\tau),\dot{\hat{x}}(\tau)) - (v-\dot{\hat{x}}(\tau))\hat{L}_{\dot{x}}(\tau,\hat{x}(\tau),\dot{\hat{x}}(\tau)) =$$

$$= \mathcal{E}(\tau,\hat{x}(\tau),\dot{\hat{x}}(\tau),v)$$

Итак окончательно получаем, что

$$0 \le \mathcal{E}(\tau,\hat{x}(\tau),\dot{\hat{x}}(\tau),v)$$

Проверим, что это условие имеет место и на концах отрезка $[t_0,t_1]$:

$$\mathcal{E}(t_0,\hat{x}(t_0),\dot{\hat{x}}(t_0),v) = \lim_{\tau\to t_0}\mathcal{E}(\tau,\hat{x}(\tau),\dot{\hat{x}}(\tau),v) \ge 0$$

Аналогично

$$\mathcal{E}(t_1,\hat{x}(t_1),\dot{\hat{x}}(t_1),v) = \lim_{\tau\to t_1}\mathcal{E}(\tau,\hat{x}(\tau),\dot{\hat{x}}(\tau),v) \ge 0$$

Тем самым мы доказали следующую теорему:

Теорема 8.1 *Если $\hat{x}(\cdot) \in C^1[t_0,t_1]$ - доставляет сильный минимум (максимум), то $\forall t \in [t_0,t_1]$, $\forall v \in \mathbb{R}$ $\mathcal{E}(t,\hat{x}(t),\dot{\hat{x}}(t),v) \ge 0$ ($\forall t \in [t_0,t_1]$, $\forall v \in \mathbb{R}$ $\mathcal{E}(t,\hat{x}(t),\dot{\hat{x}}(t),v) \le 0$)*

Замечание 8.1 *Утвеждение верно и для $\hat{x}(\cdot) \in PC^1$.*

Замечание 8.2 *Утверждение верно и для $n \ge 1$ (векторный случай).*

Рассмотрим $L(t,x,\dot{x})$, пусть $\exists L_{\dot{x}\dot{x}}$. Мы имеем $\mathcal{E}(t,x,u,u)=0$;

$$\frac{\partial}{\partial v}\mathcal{E}(t,x,u,v)\Big|_{v=u} = L_v(t,x,v)|_{v=u} - L_u(t,x,u) = 0$$

Так как $\frac{\partial^2}{\partial v^2}\mathcal{E}(t,x,u,v) = L_{vv}(t,x,v)$ то условие $\mathcal{E}(t,\hat{x}(t),\dot{\hat{x}}(t),v) \geq 0\ \forall v$ влечет

$$L_{\dot{x}\dot{x}}(t,\hat{x}(t),\dot{\hat{x}}(t)) \geq 0$$

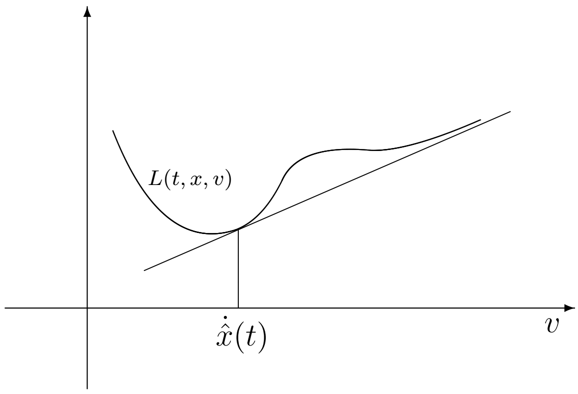


Рис. 3: При фиксированных t,x график $L(t,x,v)$ лежит выше касательной.

Получили условие Лежандра. Заметим, что условие Вейерштрасса говорит нам о том, что график функции лежит выше касательной, см. (рис.3). Это своего рода глобальное условие.

9 Элементы теории поля.

Рассмотрим простейшую задачу классического вариационного исчисления:

$$J(x(\cdot)) = \int\limits_{t_0}^{t_1} L(t,x(t),\dot{x}(t))dt \to extr$$

$x(t_0)=x_0$; $x(t_1)=x_1$. Положим $L \in C(V \times \mathbb{R}^n)$, $V = intV \subset \mathbb{R}^{n+1}$, $\hat{x}(\cdot)$ - экстремаль, $(t,\hat{x}(t)) \in V$

Рассмотрим некое семейство экстремалей: $\{x(t,\alpha)\}$ $t \in (t_0^{'},t_1^{'})$ $(t_0^{'} < t_0 < t_1 < t_1^{'})$, $\alpha -$ пробегает окрестность 0;

Пусть $x(t,0) = \hat{x}(t)$, т.е. семейство экстремалей содержит, в частности, и нашу экстремаль (при $\alpha = 0$)

Наложим на это семейство некоторые условия гладкости:

$$x, \dot{x}, \frac{\partial x}{\partial \alpha}, \frac{\partial \dot{x}}{\partial \alpha} - \text{непрерывны.}$$

Определение 9.1 $\hat{x}(\cdot)$ - *окружена полем экстремалей, если определена функция* $\alpha : V \to \mathbb{R}^n$ $(\alpha = \alpha(t, x(t)))$ *при этом* $x = x(t, \alpha) \Leftrightarrow \alpha = \alpha(t, x(t))$, *последнее означает, что через каждую точку* x *области* V *проходит ровно одна экстремаль.*

Пример 9.1

$$\int_0^1 \dot{x}^2 dt \to extr$$

$x(0) = 0$; $x(1) = 0$. *Найти соответствующее поле экстремалей.*

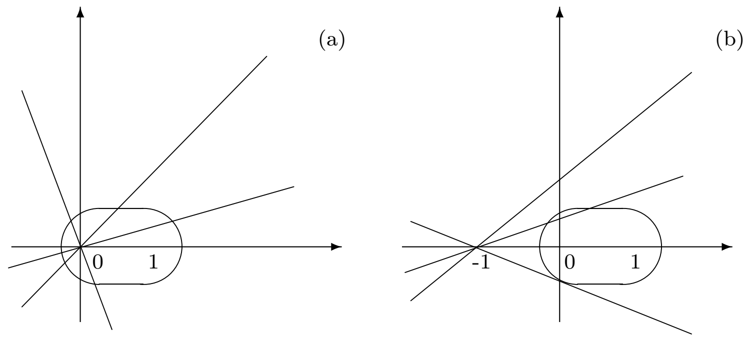


Рис. 4:

□ $\hat{x}(t) = 0$ - допустимая экстремаль. Теперь рассмотрим семейство экстремалей вида:

$$x(t, \alpha) = \alpha t;$$

Тогда на области выделенной на (рис. 8(a)) существует точка, через которую проходят все экстремали семейства, что противоречит определению поля. Рассмотрим другое семейство:

$$x(t, \alpha) = (t + 1)\alpha;$$

Тогда через каждую точку области на (рис. 8(b)) проходит ровно одна экстремаль. Поэтому это семейство можно рассматривать как поле экстремалей. ■

Будем далее требовать, чтобы $\alpha(\cdot) \in C^1$.

Определение 9.2 *Поле экстремалей называется центральным полем экстремалей (ц.п.э.), если* $\exists (t_*, x_*)$ *такая, что* $\forall \alpha$ $x(t_*, \alpha) = x_*$. *(т.е. все экстремали проходят через одну точку.)*

Например, поле экстремалей из предыдущей задачи является центральным полем экстремалей. Очевидно, что $t_* \notin [t_0, t_1]$. Положим для определенности, что $t_* < t_0$

Определение 9.3 *Функцией наклона поля называется функция вида:*

$$u(t, x) = \dot{x}(t, \alpha(t, x)), u(\cdot) \in C(V)$$

Определение 9.4 *S - функцией называется функция вида:*

$$S(\tau,\xi)=\int\limits_{t_*}^{\tau} L(t,x(t,\alpha(\tau,\xi)),\dot{x}(t,\alpha(\tau,\xi)))dt;$$

Для нахождения дифференциала S-функции воспользуемся теоремой о дифференцировании интегрального функционала с подвижными концами. Фиксируем $\tau,\xi,\Delta\tau,\Delta\xi$. Рассмотрим семейство экстремалей $X(t,\beta)$, где β пробегает окрестность нуля. Положим

$$X(t,\beta)=x(t,\alpha(\tau+\beta\Delta\tau,\xi+\beta\Delta\xi)),$$

$$t_0(\beta)=\tau_*,t_1(\beta)=\tau+\beta\Delta\tau,$$
$$x_0(\beta)=x_*,x_1(\beta)=\xi+\beta\Delta\xi,$$

$$J(\beta)=\int_{t_0(\beta)}^{t_1(\beta)} L(t,X(t,\beta),\dot{X}(t,\beta))dt$$

По теореме о дифференцировании интегрального функционала с подвижными концами

$$J'(0)=L_{\dot{x}}(\tau,\xi,u(\tau,\xi))\Delta\xi-\left(L_{\dot{x}}(\tau,\xi,u(\tau,\xi))u(t,\xi)-L(\tau,\xi,u(\tau,\xi))\right)\Delta\tau$$

Значит,

$$dS=L_{\dot{x}}(\tau,\xi,u(\tau,\xi))d\xi-\left(L_{\dot{x}}(\tau,\xi,u(\tau,\xi))u(t,\xi)-L(\tau,\xi,u(\tau,\xi))\right)d\tau$$

$-\frac{\partial S}{\partial\tau}=H(\tau,\xi,\frac{\partial S}{\partial\xi})$, где $H(\tau,\xi,p)=pu(\tau,\xi)-L(\tau,\xi,u(\tau,\xi))$

Данное дифференциальное уравнение, которому удовлетворяет S-функция, называется уравнением Гамильтона-Якоби.

Пусть есть какая-то допустимая функция $x(\cdot)$. Найдем $J(x(\cdot))-J(\hat{x}(\cdot))$, считаем, что $\forall t\in[t_0,t_1]$ $(t,x(t))\in V$.

Имеем:

$$J(\hat{x}(\cdot))=\int_{t_0}^{t_1} L(t,x(t,0),\dot{x}(t,0))dt=$$

$$=\{\text{все экстремали нашего семейства проходят через } t_*\}=$$

$$=\int_{t_*}^{t_1} L(t,x(t,0),\dot{x}(t,0))dt-\int_{t_*}^{t_0} L(t,x(t,0),\dot{x}(t,0))dt=S(t_1,x_1)-S(t_0,x_0)=$$

$$=\int_{t_0}^{t_1} dS(t,x(t))=\{\xi=x(t)\Rightarrow d\xi=\dot{x}(t)dt\}=$$

$$\int_{t_0}^{t_1} -\left(L_{\dot{x}}(t,x(t),u(t,x(t))u(t,x(t)) - L(t,x(t),u(t,x(t)))\right) dt +$$

$$+L_{\dot{x}}(t,x(t),u(t,x(t)))\dot{x}(t)dt =$$

$$= \int_{t_0}^{t_1} \left(L(t,x(t),u(t,x(t))) + L_{\dot{x}}(t,x(t),u(t,x(t)))(\dot{x}(t) - u(t,x(t)))\right) dt$$

Далее имеем

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

$$-L_{\dot{x}}(t,x(t),u(t,x(t)))(\dot{x}(t) - u(t,x(t))))dt =$$

$$= \{\text{подинтегральное выражение - это функция Вейерштрасса}\} =$$

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

Полученная формула называется основной формулой Вейерштрасса.

Теорема 9.1 *Пусть $\hat{x}(\cdot)$ - допустимая экстремаль, окруженная центральным полем экстремалей (ц.п.э.) и пусть $\exists \delta > 0 : \forall t \in [t_0,t_1], \forall x : |x - \hat{x}(t)| < \delta$, $\forall u : |u - \dot{\hat{x}}(t)| < \delta$, $\forall v \in \mathbb{R}$ $\mathcal{E}(t,x,u,v) \geq 0$ (≤ 0), т.е. выполнено усиленное условие Вейерштрасса. Тогда $\hat{x}(\cdot)$ - сильный минимум (соответственно сильный максимум).*

$\square$ Воспользуемся тем, что u является непрерывной. Рассмотрим $\delta \to \varepsilon > 0$: если $|x - \hat{x}(t)| < |x - \hat{x}(t)| < \varepsilon$, то $|u(t,x) - u(t,\hat{x}(t))| < \delta$, что выполнено тогда и только

Проверим по определению. Пусть $x(\cdot)$ - допустимая функция.

$$\|x(\cdot) - \hat{x}(\cdot)\|_{C[t_0,t_1]} < \varepsilon \Rightarrow \forall t \;\; |x(t) - \hat{x}(t)| < \varepsilon \leq \delta, |u(t,x(t)) - \dot{\hat{x}}(t)| < \delta$$

По основной формуле Вейерштрасса $J(x(\cdot)) - J(\hat{x}(\cdot)) \geq 0$. Т.о. для $x(\cdot)$ из $C^1[t_0,t_1]$ теорема доказана. Ее верность для $x(\cdot)$ из $PC^1[t_0,t_1]$ получаем, воспользовавшись леммой о скруглении углов. $\blacksquare$

Следствие 9.1 *Пусть $\hat{x}(\cdot)$ окружена ц.п.э., $L(t,x,\dot{x})$ - выпуклая по $\dot{x}$. Тогда $\hat{x}(\cdot)$ - сильный минимум.*

$\square$ Действительно, из выпуклости следует, что $\mathcal{E}(t,x,u,v) \geq 0$. $\blacksquare$

Следующую теорему приведем без доказательства:

Теорема 9.2 *Пусть $L \in C^3(V \times \mathbb{R})$, выполнены усиленные условия Лежандра, Якоби, $\hat{x}(\cdot)$ - допустимая экстремаль. Тогда $\hat{x}(\cdot)$ можно окружить ц.п.э.*

Следствие 9.2 *При выполнении условий теоремы 9.2 и усиленного условия Вейерштрасса $\hat{x}(\cdot)$ - сильный экстремум.*

Заметим в заключение, что все сказанное в данном параграфе практически без изменений переносится на векторный случай, и сформулируем несколько обязательных задач.

Задача 3 $J(x(\cdot)) = \int_{t_0}^{t_1} L(t, \dot{x}(t))dt \to extr$, *$x(t_0) = x_0, x(t_1) = x_1$; $\hat{x}(\cdot) \in C^1[t_0, t_1]$ - допустимая экстремаль, $\forall t \in [t_0, t_1]$, $\forall v$ $\mathcal{E}(t, \hat{x}(t), \dot{\hat{x}}(t), v) \leq 0$. Тогда $\hat{x}(\cdot)$ - глобальный минимум (в классе допустимых функций).*

Задача 4 $J(x(\cdot)) = \int_{t_0}^{t_1} (A(t) + A_0(t)x(t) + A_1(t)\dot{x}(t) + A_{00}(t)x^2(t) + 2A_{01}(t)x(t)\dot{x}(t) + A_{11}\dot{x}^2(t))dt$; *пусть $\hat{x}(\cdot)$ - слабый минимум. Тогда $\hat{x}(\cdot)$ доставляет глобальный минимум.*

Задача 5 *Привести пример простейшей задачи к.в.и., такой что некоторая экстремаль $\hat{x}(\cdot)$ удовлетворяет усиленному условию Лежандра, условию Якоби, $\mathcal{E}(t, \hat{x}(t), \dot{\hat{x}}(t), v) \leq 0$ $\forall t, v$, но $\hat{x}(\cdot)$ не является сильным минимумом.*

□ Такой задачей является, например, следующая:

$$\int_0^1 (\dot{x}^2 - x\dot{x}^3)dt \to extr; x(0) = x(1) = 0; \hat{x}(t) = 0$$

■

10 Задача о брахистохроне.

В 1696 году Иоганн Бернулли в первом в истории математическом журнале “Acta eruditorum” (основанном в 1682 г.) опубликовал заметку “Problema novum, ad cujus solutionem Mathematici invitantur” — “Новая задача, решить которую приглашаются математики”, в котором предлагал вниманию математиков задачу о линии быстрейшего ската — *брахистохроне*. В этой задаче требуется определить линию, соединяющую две заданные точки A и B, не лежащие на одной вертикальной прямой, и обладающую тем свойством, что материальная точка под действием силы тяжести скатится по этой линии из точки A в точку B в кратчайшее время.

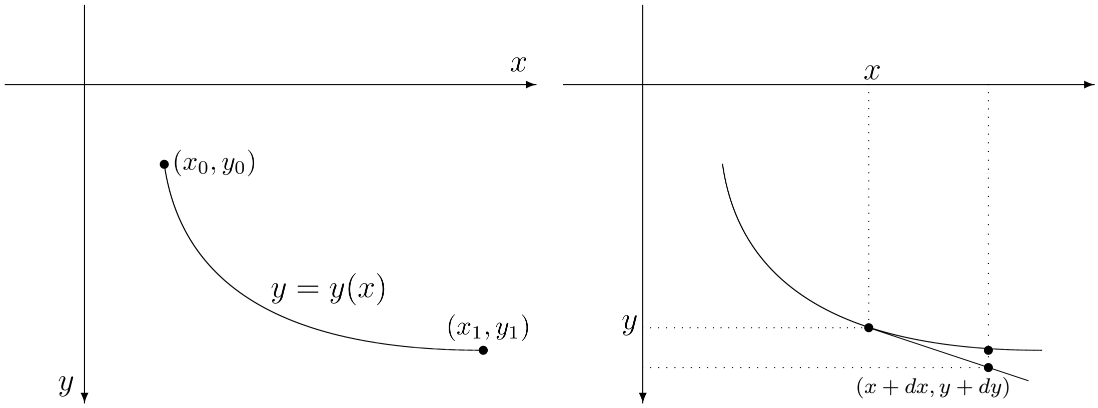


Рис. 5: Брахистохрон в координатах (x, y)

Введем декартову систему координат на плоскости (см. рис 6); пусть ось Ox соответствует нулевой скорости и направлена горизонтально, ось Oy направлена вертикально вниз. Пусть точки $A(x_0, y_0)$ и $B(x_1, y_1)$ имеют положительные ординаты, а $x_0 < x_1$. Соласно законам механики скорость движения материальной точки $\frac{ds}{dt} = \sqrt{2gy}$, откуда находим время, затрачиваемое на перемещение точки из положения A в положение B:

$$T[y(x)] = \frac{1}{\sqrt{2g}} \int_{x_0}^{x_1} \frac{\sqrt{1+\dot{y}^2}}{\sqrt{y}} dx; \quad y(x_0) = y_0, \quad y(x_1) = y_1.$$

Будем предполагать $y(\cdot)$ непрерывно дифференцируемой функцией. Обозначим $L(y, \dot{y}) = \frac{\sqrt{1+\dot{y}^2}}{\sqrt{y}}$. Тогда задача о брахистохроне записывается в виде экстремальной задачи (множитель $\dfrac{1}{\sqrt{2g}}$ отбрасывается):

$$J(y(*)) = \int_{x_0}^{x_1} L(y, \dot{y}) dx \to min; \quad y(x_0) = y_0, \quad y(x_1) = y_1 \qquad (1)$$

Уравнение Эйлера для экстремали y имеет вид

$$\frac{d}{dx}\left(\frac{1}{\sqrt{y}} z\right) = L_y, \quad z = \frac{\dot{y}}{\sqrt{1+\dot{y}^2}},$$

откуда $z(\cdot) \in C^1[x_0, x_1]$. Так как

$$\dot{y} = \frac{z}{\sqrt{1-z^2}},$$

то $\dot y(\cdot) \in C^1[x_0, x_1]$, или $y(\cdot) \in C^2[x_0, x_1]$.

Поскольку интегрант L в задаче (1) не содержит явно x и $y(\cdot) \in C^2[x_0, x_1]$, то уравнение Эйлера имеет первый интеграл $H = L - \dot y L_{\dot y} = C$, или в данном случае

$$\frac{\sqrt{1+\dot y^2}}{\sqrt{y}} - \frac{\dot y^2}{\sqrt{y(1+\dot y^2)}} = C,$$

откуда после упрощений будем иметь $\dfrac{1}{\sqrt{y(1+\dot y^2)}} = C$ или $y(1+\dot y^2) = C_1$.

Поскольку

$$0 = \frac{d}{dx}H = \dot y\left(\frac{d}{dx}L_{\dot y} - L_y\right) = 0,$$

следовательно, решения уравнения $y(1+\dot y^2) = C_1$ удовлетворяют уравнению Эйлера для тех x, для которых $\dot y(x) \neq 0$. Однако, если $y = const$ на некотором интервале (x', x''), то $\frac{d}{dx}L_{\dot y} = 0$, $L_y \neq 0$ на (x', x''), и y не является экстремалью.

Сделаем замену $\dot y = \operatorname{ctg} t$, где $(0, \pi) \ni t = t(x)$ — непрерывно дифференцируемая функция от x; тогда получим:

$$y = \frac{C_1}{1+\operatorname{ctg}^2 t} = C_1 \sin^2 t = \frac{C_1}{2}(1 - \cos 2t);$$

Если x выбрано так, что $y(x) < C_1$, то $\dot y \neq 0$, следовательно $t \neq \pi/2$,

$$\frac{dx}{dt} = \frac{dy}{dt}\frac{1}{\dot y} = \frac{C_1 \sin t \cos t}{\operatorname{ctg} t} = 2C_1 \sin^2 t = 2y.$$

Если же $y(x) = C_1$, то существует последовательность $\{x_n\}$ такая, что $x_n \to x$, $y(x_n) < C_1$. Тогда

$$\frac{dt}{dx}|_x = \lim_{n\to\infty} \frac{dt}{dx}|_{x_n} = \lim_{n\to\infty} \frac{1}{2y(x_n)} = \frac{1}{2y(x)}.$$

Значит, $\dfrac{dx}{dt} = 2y$ при всех t. Поэтому

$$x = C_1\big(t - \frac{\sin 2t}{2}\big) + C_2 = \frac{C_1}{2}(2t - \sin 2t) + C_2.$$

Сделаем подстановку $\tau = 2t$. Тогда в параметрической форме уравнение экстремали примет вид

$$\begin{cases} x = \dfrac{C_1}{2}(\tau - \sin\tau) + C_2, \\ y = \dfrac{C_1}{2}(1 - \cos\tau), \qquad (0 \le \tau \le 2\pi). \end{cases} \qquad (2)$$

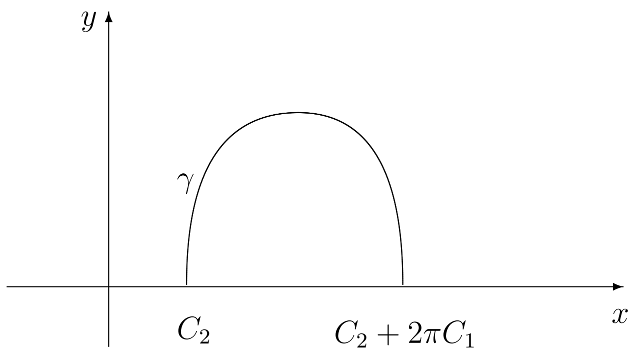


Рис. 6: Арка циклоиды.

Кривая, параметрические уравнения которой имеют вид (2) называется *аркой циклоиды.*

До исследования И. Бернулли циклоида появилась в работах Гюйгенса, доказавшего, что она является *изохронной*, т.е. такой кривой, для которой период колебания скользящей вдоль нее материальной точки не зависит от начального положения точки. Это дало повод И. Бернулли восхититься: "Природа всегда действует простейшим образом, как и в данном случае она с помощью одной и той же линии оказывает две различные услуги".

Теперь мы покажем, что допустимая экстремаль в задаче (1) существует и единственна.

Для этого достаточно, очевидно, показать, что через точки A и B проходит единственная арка циклоиды вида (2).

Рассмотрим случай, когда $y_0 \neq y_1$. Пусть для определенности $y_0 > y_1$. Проведем через точки A и B прямую l и обозначим через D точку пересечения прямой l с осью Ox. Введем параметр $\beta = |AB|/|BD|$.

Обозначим через γ арку циклоиды, параметрические уравнения которой

$$x = \frac{1}{2}(\tau - \sin\tau), \qquad y = \frac{1}{2}(1 - \cos\tau).$$

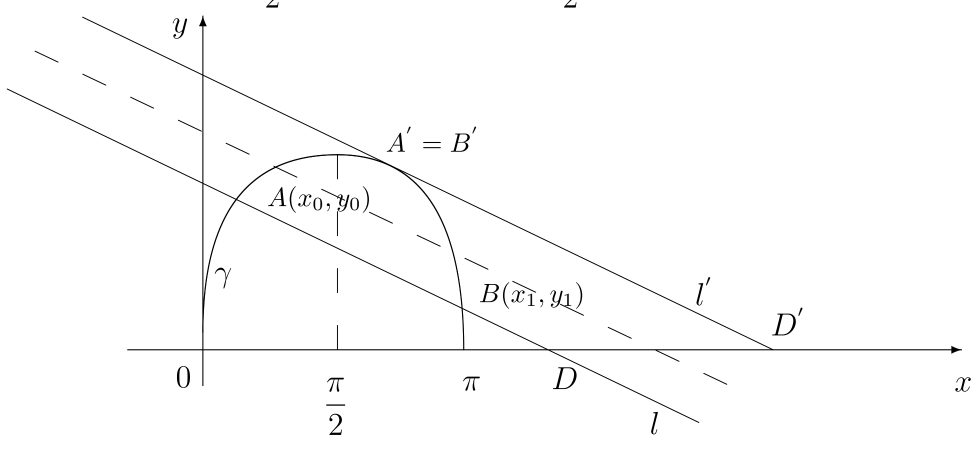


Рис. 7:

Проведем касательную l' к кривой γ параллельно l и обозначим через D' точку пересечения l' с осью Ox. Пусть $A' = B'$ — точка касания. Тогда $\beta' = |A'B'|/|B'D'| = 0$. Будем параллельно переносить прямую l' так, чтобы точка пересечения с осью Ox непрерывно приближалась к точке $(\pi/2, 0)$. Тогда, если A' и B' — точки пересечения l' с Ox, то абсцисса A' монотонно убывает, абсцисса B' монотонно возрастает, $|B'D'|$ монотонно убывает к нулю, $|A'B'|$ монотонно возрастает и ограничена. Поэтому величина β' будет непрерывно и монотонно стремиться к бесконечности и, следовательно, при некоторых однозначно определенных $A' = A^0$, $B' = B^0$, $D' = D^0$ примет значение β.

Преобразованием координат S_1 вида

$$x = c_1 x' + c_2, \qquad y = c_1 y' \qquad (c_1 > 0) \qquad (3)$$

переведем точку A' в точку A. При этом точка B' перейдет в точку B (в силу построения), а арка циклоиды γ перейдет в арку циклоиды, проходящую через точки A и B. Обратно, пусть через точки A и B проходит арка циклоиды. Преобразованием координат S_2 вида (3) переведем ее в "стандартную" арку γ. Преобразование $S_2 \circ S_1$ переводит точки A', B' и D' соответственно в точки A'', B'' и D'', которые принадлежат циклоиде, лежат на прямой, параллельной прямой l и при этом $|A''B''|/|B''D''| = \beta$. Из доказанного выше вытекает, что $A' = A''$, $B' = B''$, $D' = D''$ и, следовательно, преобразование S_2 есть обратное к S_1.

Итак, доказано, что в случае $y_0 \neq y_1$ существует и единственна арка циклоиды, проходящая через точки A и B.

В случае, когда $y_0 = y_1$ выпустим из точки $D((x_0+x_1)/2)$ лучи l_1 и l_2, проходящие через A и B соответственно. Из точки $D'(\pi/2, 0)$ выпустим лучи l_1' и l_2', колинеарные лучам l_1 и l_2. Пусть A' и B' — точки пересечения лучей l_1' и l_2' с кривой γ. Тогда положим $c_1 = |A'D'|/|AD|$, $c_2 = \pi/2 - D$ и будем далее рассуждать как при разборе случая $y_0 \neq y_1$.

Таким образом, для любых точек (x_0, y_0) и (x_1, y_1) таких, что $x_0 < x_1$, $y_0 > 0$, $y_1 > 0$, для задачи (1) существует единственная допустимая экстремаль $\hat{y}(\cdot)$. Из этого следует, что $\hat{y}(\cdot)$ доставляет абсолютный минимум в задаче (1). Действительно, экстремаль $\hat{y}(\cdot)$ определена в некоторой окрестности точки x_0; пусть $y(x_*) = y_*$ для некоторой точки $x_* < x_0$; тогда для любой точки (τ, λ), где $x_0 \leq \tau \leq x_1$, $\lambda > 0$, существует единственная экстремаль $y(\cdot)$, проходящая через точки (x_*, y_*) и (τ, λ) и формула $u(\tau, \lambda) = y'(\tau)$ определяет функцию наклона центрального поля, окружающего экстремаль $\hat{y}(\cdot)$. Если $y(\cdot)$— любая допустимая функция для задачи (1) (т.е. $y(\cdot) \in C^1[x_0, x_1]$, $y(x_0) = y_0$ и $y(x_1) = y_1$), то в силу основной формулы Вейерштрасса

$$J(y(\cdot)) - J(\hat{y}(\cdot)) = \int_{x_0}^{x_1} \mathcal{E}(x, y(x), u(x, y(x)), \dot{y}) dx \geq 0,$$

ибо вследствие выпуклости интегранта L по $\dot{y}$ функция Вейерштрасса $\mathcal{E}$ всюду неотрицательна.

11 Гладкая задача с ограничениями типа равенств.

Теорема 11.1 *(Правило Лагранжа для конечного числа ограничений типа равенств) Пусть X - ЛНП, $U = intU \subset X$, $f_i : U \to \mathbb{R}$, $(i = 1, \ldots, m)$ - непрерывные функции. $f_0(x) \to extr$, $f_i(x) = 0$ $(i = 1, \ldots, m)$; $\hat{x}$ - $locmin$, $\exists f_i'(\hat{x})$, $(i = 0, \ldots, m)$. Тогда $\exists \bar{\lambda} = (\lambda_0, \ldots, \lambda_m) \neq 0$ и функция (функция Лагранжа) $\mathcal{L}(x, \lambda) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ такая, что $\mathcal{L}_x(\hat{x}, \bar{\lambda}) = 0 \Leftrightarrow \left(\sum\limits_{i=0}^{m} \lambda_i f_i'(\hat{x}) = 0 \right)$*

$\square$ Для дальнейшего доказательства будем использовать следующую теорему:

Теорема 11.2 *(Теорема Брауэра или теорема о неподвижной точке) Пусть $B(0, r) = \left\{ x \in \mathbb{R}^d : \|x\| \leq r \right\}$ - шар радиуса r. $F : B(0, r) \to B(0, r)$ - непревное отображение. Тогда $\exists y \in B(0, r) : F(y) = y$.*

Следствие 11.1 *(Следствие об ε - сдвиге) Пусть $0 < \varepsilon < r$ и $G : B(0, r) \to \mathbb{R}^d$ - непрерывное отображение; $\forall y \in B(0, r)$ $\|G(y) - y\| \leq \varepsilon$, $\tilde{y} \in B(0, r - \varepsilon)$. Тогда $\exists y \in B(0, r) : G(y) = \tilde{y}$.*

$\square$ Рассмотрим отображение вида:

$$F(y) = y - G(y) + \tilde{y}$$

замечаем, что $F(y)$ - непрерывное, так как отображение $G(y)$ - непрерывное по условию. Кроме того,

$$\|F(y)\| \leq \|y - G(y)\| + \|\tilde{y}\| \leq \varepsilon + (r - \varepsilon) = r$$

Следовательно $F : B(0, r) \to B(0, r) \Rightarrow$ можно использовать теорему Брауэра. $\Rightarrow$ $\exists y \in B(0, r) : F(y) = y \Leftrightarrow G(y) = \tilde{y}$. $\blacksquare$

Теперь можно перейти к доказательству правила Лагранжа: без ограничения общности положим $\hat{x}$ - $locmin$ (заметим, что достаточно проводить доказательство только для $locmin$, т.к условие $f_0(x) \to locmax \Leftrightarrow -f_0(x) \to locmin$.) Положим также, что $f_0(\hat{x}) = 0$ (опять же, заметим, что можно положить $\tilde{f}_0(x) = f_0(x) - f_0(\hat{x})$ и далее рассмотреть задачу: $f_0 \to min$, $f_i = 0$ $(i = 1, \ldots, m)$)

Вспомним определение: $\hat{x}$ - доставляет локальный минимум $\Leftrightarrow$ $f_i(\hat{x}) = 0, (i = 0, \ldots, m)$, т.е $\hat{x}$ - допустима, и $\exists \varepsilon > 0 : \forall x : (\|x - \hat{x}\| \leq \delta,\ f_i(x) = 0\ (i = 1, \ldots, m)) \Rightarrow f_0(x) \geq 0$.

Теперь рассмотрим $Y = \left\{ \left(f_0'(\hat{x})[h], \ldots, f_m'(\hat{x})[h] \right) \in \mathbb{R}^{m+1}, h \in X \right\}$ - множество векторов длинны $m + 1$. Далее замечаем, что Y - подпространство в $\mathbb{R}^{m+1}$. Для доказательства последнего нужно проверить, что имеют место следующие факты:

1. $y_1, y_2 \in Y \Rightarrow y_1 + y_2 \in Y$
2. $y \in Y \Rightarrow \alpha y \in Y, \alpha \in \mathbb{R}$

Рассмотрим два случая:

1) $Y \neq \mathbb{R}^{m+1} \Rightarrow \exists \bar{\lambda} \in \mathbb{R}^{m+1} : \forall y \in Y < \bar{\lambda}, y >= 0$ (т.е вектор, ортогональный всем элементам подпространства Y) $\Leftrightarrow$

$$< (\lambda_0, \dots, \lambda_m), (f_0'(\hat{x})[h], \dots, f_m'(\hat{x})[h]) >= 0$$

Последнее равенство можно переписать в эквивалентном виде:

$$\sum_{i=0}^{m} \lambda_i f_i'(\hat{x})[h] = 0, \forall h \Leftrightarrow \sum_{i=0}^{m} \lambda_i f_i'(\hat{x}) = 0$$

А последнее означает равенство нулю производной функции Лагранжа в точке $\hat{x}$.

2) $Y = \mathbb{R}^{m+1} \Rightarrow \forall j = 0, \dots, m \ \exists h_j \in X :$

$$\left(f_0'(\hat{x})[h_j], \dots, f_m'(\hat{x})[h_j]\right) = (0, \dots, 1, \dots, 0)$$

$$f_i'(\hat{x})[h_j] = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Положим $y = (y_0, \dots, y_1)$.

$$\| \sum_{j=0}^{m} y_j h_j \| \leq \sum_{j=0}^{m} |y_j| \|h_j\| \ \leq \sum_{j=0}^{m} \|h_j\| \max_{j=0,\dots,m} |y_j| \leq \sum_{j=0}^{m} \|h_j\| \|y\|$$

Оценим величину $f_i(\hat{x} + \sum\limits_{j=0}^{m} y_j h_j)$ и положим $\|y\| \to 0$.

$$f_i(\hat{x} + \sum_{j=0}^{m} y_j h_j) = f_i(\hat{x}) + f_i'(\hat{x}) \left[\sum_{j=0}^{m} y_j h_j\right] + \bar{\bar{o}} \left(\| \sum_{j=0}^{m} y_j h_j \| \right) =$$

$$= 0 + \sum_{j=0}^{m} y_j \underbrace{f_i'(\hat{x})[h_j]}_{\delta_{ij}} + \bar{\bar{o}}(\|y\|) = y_i + \bar{\bar{o}}(\|y\|)$$

Следовательно

$$G(y) := \left(f_0(\hat{x} + \sum_{j=0}^{m} y_j h_j), \dots, f_m(\hat{x} + \sum_{j=0}^{m} y_j h_j) \right) = (y_0, \dots, y_m) + \bar{\bar{o}}(\|y\|) = y + \bar{\bar{o}}(\|y\|)$$

Пусть $r \sum\limits_{j=0}^{m} \|h_j\| < \delta$, $\|G(y) - y\| \leq \dfrac{\|y\|}{2}$ $(\|y\| < r)$

Итак, если $y \in B(0, r) \Rightarrow \|G(y) - y\| \leq \dfrac{\|y\|}{2} \leq \dfrac{r}{2} := \varepsilon$

Введем в рассмотрение вектор $\tilde{y} = (-\varepsilon, 0, \dots, 0) \Rightarrow \exists y \in B(0, r)$: $G(y) = \tilde{y} \Leftrightarrow$

$$f_0 \left(\hat{x} + \sum_{j=0}^{m} y_j h_j \right) = -\varepsilon$$

$$f_i\left(\hat{x}+\sum_{j=0}^{m} y_j h_j\right)=0$$

Кроме того,

$$\|\underbrace{\left(\hat{x}+\sum_{j=0}^{m} y_j h_j\right)}_{\bar{x}}-\hat{x}\| \le r\sum_{j=0}^{m}\|h_j\|<\delta$$

Теперь можно заметить, что наличие вектора $\bar{x}$ противоречит тому, что $\hat{x}-locmin$ поэтому пункт 2) не имеет места. Теорема доказана полностью. ■

12 Изопериметрическая задача.

$U = intU \subset \mathbb{R}^{2n+1}$; $f_i : U \to \mathbb{R}(i = 0, \dots, m)$

$$\mathcal{J}_0(x(\cdot)) = \int\limits_{t_0}^{t_1} f_0(t, x(t), \dot{x}(t))dt \to extr$$

$$\mathcal{J}_i(x(\cdot)) = \int\limits_{t_0}^{t_1} f_i(t, x(t), \dot{x}(t))dt = \alpha_i(i = 1, \dots, m)(1)$$

$$(2) = \begin{cases} x(t_0) = x_0 \\ x(t_1) = x_1 \end{cases}$$

$\hat{x}(\cdot)$ - допустима, если $\hat{x}(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$ и $\forall t \in [t_0, t_1]$ $(t, x(t), \dot{x}(t)) \in U$ (последнее необходимо для того, чтобы интегралы, записанные выше, были определены и выполнены условия (1),(2))

Теперь сформулируем необходимые условия локального экстремума для $n = 1$:

Теорема 12.1 *Пусть* $\hat{x}(\cdot) - locextr$, $f_i, (f_i)_x, (f_i)_{\dot{x}}$ *- непрерывны* $\forall i = (0, \dots, m)$ *Тогда* $\exists\bar{\lambda} = (\lambda_0, \dots, \lambda_m) \neq 0$ *такая, что функция* $L(t, x, \dot{x}) = \sum\limits_{i=0}^{m} \lambda_i f_i(t, x, \dot{x})$ *удовлетворяет уравнению Эйлера, т.е* $\forall t \in [t_0, t_1]$ $\dfrac{d}{dt}\hat{L}_{\dot{x}}(t) = \hat{L}_x(t)$

□ Рассмотрим отображение F_i : (окрестность 0 в $C_0^1[t_0, t_1]$) $\to \mathbb{R}$

$$F_0(h[\cdot]) = \mathcal{J}_0(\hat{x}(\cdot) + h(\cdot))$$
$$F_i(h[\cdot]) = \mathcal{J}_i(\hat{x}(\cdot) + h(\cdot)) - \alpha_i, (\alpha_0 = 0; i = (0, \dots, m))$$

$\hat{h}(\cdot) = 0 - locextr$ ($\hat{h}(\cdot)$ - обозначение нулевой функции.) $F_0(h(\cdot)) \to extr$ $F_i(h(\cdot)) = 0$ $(i = (1, \dots, m))$; Оценим $F_i(h(\cdot)) - F_i(0)$ при $h(\cdot) \in C_0^1[t_0, t_1]$ и $\|h\|_1 \to 0$, имеем:

$$F_i(h(\cdot)) - F_i(0) = \int\limits_{t_0}^{t_1} f_i(t, \hat{x}(t) + h(t), \dot{\hat{x}}(t) + \dot{h}(t))dt - \alpha_i - \int\limits_{t_0}^{t_1} f_i(t, \hat{x}(t), \dot{\hat{t}})dt + \alpha_i =$$

$$= \int\limits_{t_0}^{t_1} \left(f_i(t, \hat{x}(t), \dot{\hat{x}}(t)) + \left(\hat{f}_i\right)_x h(t) + \left(\hat{f}_i\right)_{\dot{x}} \dot{h}(t) + \bar{\bar{o}}(|h(t)| + |\dot{h}(t)|) - f_i(t, \hat{x}(t), \dot{\hat{x}}(t)) \right) dt =$$

$$= \int\limits_{t_0}^{t_1} \left(\left(\hat{f}_i\right)_x h + \left(\hat{f}_i\right)_{\dot{x}} \dot{h} \right) dt + \bar{\bar{o}}\|h\|_1$$

Следовательно $F_i^{'}(0)[h(\cdot)] = \int\limits_{t_0}^{t_1} \left(\left(\hat{f}_i\right)_x h + \left(\hat{f}_i\right)_{\dot{x}} \dot{h} \right) dt$

Пусть $\mathcal{L}(h(\cdot), \bar{\lambda}) = \sum\limits_{i=0}^{m} \lambda_i F_i(h(\cdot))$, тогда по принципу Лагранжа $\hat{\mathcal{L}}_{h(\cdot)}(0) = 0$

$L = \sum\limits_{i=0}^{m} \lambda_i f_i$,

$$\hat{\mathcal{L}}_{h(\cdot)}(0) = \sum_{i=0}^{m} \lambda_i F_i^{'}(0) = 0$$

$\Leftrightarrow$

$$\forall h(\cdot) \in C_0^1[t_0, t_1] \sum_{i=0}^{m} \lambda_i \int\limits_{t_0}^{t_1} \left(\left(\hat{f}_i\right)_x h + \left(\hat{f}_i\right)_{\dot{x}} \dot{h} \right) dt = 0$$

$\Leftrightarrow$

$$\forall h(\cdot) \in C_0^1[t_0, t_1] \sum_{i=0}^{m} \lambda_i \int\limits_{t_0}^{t_1} \left(\left(\sum_{i=0}^{m} \lambda_i \hat{f}_i \right)_x h + \left(\sum_{i=0}^{m} \lambda_i \hat{f}_i \right)_{\dot{x}} \dot{h} \right) dt = 0$$

$\Leftrightarrow$

$$\forall h(\cdot) \in C_0^1[t_0, t_1] \int\limits_{t_0}^{t_1} \left(\hat{L}_x h + \hat{L}_{\dot{x}} \dot{h} \right) dt = 0$$

Следовательно $\forall t \in [t_0, t_1]$ $\dfrac{d}{dt} \hat{L}_{\dot{x}}(t) = \hat{L}_x(t)$ ■

13 Задача с подвижными концами.

$U = intU \subset \mathbb{R}^{2n+1}$, $V = intV \subset \mathbb{R}^{2n+2}$; $L \in C(U, \mathbb{R})$, $\psi_i \in C^1(V, \mathbb{R})(i = 0, \ldots, m)$

$$\mathcal{J}(\cdot) = \int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt + \psi_0(t_0, x(t_0), t_1, x(t_1)) \to extr$$

$$\psi_i(t_0, x(t_0), t_1, x(t_1)) = 0, (i = (1, \ldots, m))$$

$t_0, t_1 \in int\triangle$, $\triangle$ - отрезок, $t_0 < t_1$, $x(\cdot) \in C^1(\triangle, \mathbb{R}^n)$

Тройка $(x(\cdot), t_0, t_1)$ - допустима, если $\forall t \in [t_0, t_1]$ $(t, x(t), \dot{x}(t)) \in U$, $(t_0, x(t_0), t_1, x(t_1)) \in V$, $\psi_i(t_0, x(t_0), t_1, x(t_1)) = 0$ $(i = 0, \ldots, m)$

Допустимая тройка $(\hat{x}(\cdot), \hat{t}_0, \hat{t}_1)$ $-$ $locmin$, если $\exists \varepsilon > 0$: $\forall$ допустимой тройки $(x(\cdot), t_0, t_1)$ такой, что $\|x(\cdot) - \hat{x}(\cdot)\|_{C^1(\triangle, \mathbb{R}^n)} < \varepsilon$, $|t_0 - \hat{t}_0| < \varepsilon$, $|t_1 - \hat{t}_1| < \varepsilon$ выполнено

$$\mathcal{J}(x(\cdot), t_0, t_1) \geq \mathcal{J}(\hat{x}(\cdot), \hat{t}_0, \hat{t}_1).$$

Теорема 13.1 *$(\hat{x}(\cdot), \hat{t}_0, \hat{t}_1)$ — слабый экстремум в задаче*

$$\mathcal{J}(x(\cdot), t_0, t_1) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt + \psi_0(t_0, x(t_0), t_1, x(t_1)) \to extr$$

$$\psi_i(t_0, x(t_0, t_1, x(t_1)) = 0 \ (i = 1, \ldots, m)$$

$L, L_x, L_{\dot{x}}$ — непрерывны ны $U, \psi_i \in C^1(V)$. Тогда

1) выполнено уравнение Эйлера $\frac{d}{dt}\hat{L}_{\dot{x}}(t) = \hat{L}_x(t)$

2) $\exists \overline{\lambda} = (\lambda_0, \ldots, \lambda_m) \neq 0$:

$$\mathcal{L}(x(\cdot), t_0, t_1, \overline{\lambda}) = \lambda_0 \mathcal{J}(x(\cdot), t_0, t_1) + \sum_{i=1}^{m} \lambda_i \psi_i(t_0, x(t_0), t_1, x(t_1)) =$$

$$= \int_{t_0}^{t_1} \lambda_0 L(t, x(t), \dot{x}(t))dt + \psi(t_0, x(t_0), t_1, x(t_1))$$

При этом выполняются следующие условия:

а) трансверсальности

$$\lambda_0 \hat{L}_{\dot{x}}(t_0) = \hat{\psi}_{x(t_0)}$$
$$\lambda_0 \hat{L}_{\dot{x}}(t_1) = -\hat{\psi}_{x(t_1)}$$

б) стационарности по t_0, t_1: $\hat{\mathcal{L}}_{t_0} = 0$, $\hat{\mathcal{L}}_{t_1} = 0$. Более подробно:

$$\frac{d}{dt_0}\mathcal{L}(\hat{x}(\cdot), t_0, \hat{t}_1, \overline{\lambda})|_{t_0=\hat{t}_0} = 0$$
$$\frac{d}{dt_1}\mathcal{L}(\hat{x}(\cdot), \hat{t}_0, t_1, \overline{\lambda})|_{t_1=\hat{t}_1} = 0$$

Задача 6 *Доказать теорему (n=1).*

Доказательство аналогично проведенному для изопериметрической задачи. Отличие лишь в том, что в данной задаче функционал действует в пространстве $X = C^1(\Delta) \times \mathbb{R}^2$ с нормой $\|(x(\cdot), t_0, t_1)\| = max(\|x(\cdot)\|_{C^1(\Delta)}, |t_0|, |t_1|)$.

Заметим также, что в рассмотренной задаче t_0 и t_1 могут быть фиксированы. Тогда просто нужно не брать производную по фиксированным концам.

Задача 7 *Привести пример задачи с подвижными концами, в которой обязательно $\lambda_0 = 0$.*

14 Задача с ограничениями типа равенств и неравенств.

Теорема 14.1 *Пусть X — линейное нормированное пространство (ЛНП), $U = intU \subset X$, $f_i \in C(U,\mathbb{R})$ $(i = 1,\dots,m')$ — некоторый набор функциий,*

$$A = \{x \in U : f_i(x) \le 0 \ \ (i = 1,\dots,m')\}$$
$$\hat{x} \in A, \ \ \exists f_i'(\hat{x}) = x_i^* \ \ (i = 1,\dots m')$$

Пусть также определены функции $f_0, f_{m'+1},\dots,f_m \in C(A,\mathbb{R})$ и выполнено условие "дифференцируемости" в точке $\hat{x}$:

$$\exists x_i^* \in X^*, f_i(x) = f_i(\hat{x}) + \langle x_i^*, x - \hat{x}\rangle + o(\|x - \hat{x}\|)$$
$$\text{при } x \in A, \ x \to \hat{x} \ (i = 0, m'+1,\dots,m)$$

Заметим, что x_i^* — как бы производная f_i в точке $\hat{x}$ $(i = 0, m'+1,\dots,m)$. Но говорить, что это производная, нельзя, т.к. область A может и не содержать никакую окрестность $\hat{x}$, как, например, на следующем рисунке:

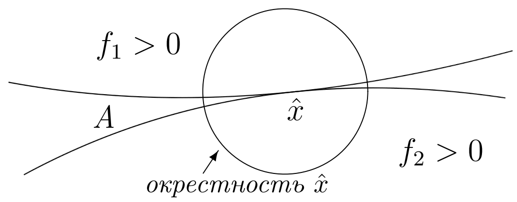


Рис. 8: область A заключена между двумя кривыми, которые касаются в точке $\hat{x}$

Далее рассматриваем экстремальную задачу $(*)$:

$$f_0(x) \to min,$$
$$f_i(x) \le 0 \ (i = 1,\dots,m'),$$
$$f_i(x) = 0 \ (i = m'+1,\dots,m)$$

Пусть $\hat{x}$ — locmin для $(*)$. Тогда $\exists\overline{\lambda} = (\lambda_0,\dots,\lambda_m) \ne 0$: выполнены следующие условия:

1) стационарности $\sum_{i=0}^m \lambda_i x_i^* = 0$;

2) неотрицательности $\lambda_i \ge 0$ $(i = 1,\dots,m')$;

3) дополняющей нежесткости $\lambda_i f_i(\hat{x}) = 0$ $(i = 1,\dots,m')$

Заметим, что здесь, конечно, можно было написать и $(i = 1\dots,m)$, но это и так выполняется для ограничений типа равенств.

Доказывать данную теорему пока не будем.

Нас будет интересовать лишь частный случай этой теоремы, а именно когда все функции f_i определены в окрестности $\hat{x}$. Переформулируем теорему для этого случая:

Теорема 14.2 *Пусть X — ЛНП, $U = intU \subset X$, $f_i \in C(U, \mathbb{R})$ $(i = 0, \dots, m)$; $\hat{x} \in U$ — такая, что $\exists f_i'(\hat{x}) = x_i^*$ $(i = 0, \dots, m)$ (заметим, что в этом случае все x_i^* — уже полноправные производные функции f_i), $\hat{x}$ — locmin в $(*)$.*

Тогда $\exists \overline{\lambda} = (\lambda_0, \dots, \lambda_m) \neq 0$: для функции Лагранжа $\mathcal{L}(x; \overline{\lambda}) = \sum_{i=0}^m \lambda_i f_i(x)$ выполнены следующие условия:

1) стационарности $\mathcal{L}_x(\hat{x}; \overline{\lambda}) = 0$ $(\Leftrightarrow \sum_{i=0}^m \lambda_i f_i'(\hat{x}) = 0)$;

2) неотрицательности $\lambda_i \geq 0$ $(i = 1, \dots, m')$;

3) дополняющей нежесткости $\lambda_i f_i(\hat{x}) = 0$ $(i = 1, \dots, m')$

В таком виде данная теорема очень полезна. Но и ее доказывать пока не будем. Вместо этого сформулируем одно полезное следствие, которое нам понадобится в скором будущем.

Рассмотрим пространство $X = \mathbb{R}^d = \{x = (x_1, \dots, x_d)\}$. Пусть

$$K = \{x : \ x \geq 0 \Leftrightarrow x_1 \geq 0, \dots, x_d \geq 0\}$$

Пусть также $\hat{x} = 0$, $\hat{x} \in U = intU$; рассмотрим $f : \ U \cap K \to \mathbb{R}$.

Определим одностороннюю производную в нуле (обозначать ее будем как обычную, но подразумевать одностороннюю): пусть

$$f(x) = f(0) + \langle x^*, x \rangle + o(\|x\|)$$
$$(x \in K, \ x \to 0)$$

Тогда $f'(0) = x^*$ будем называть односторонней производной в нуле.

Рассмотрим базисные вектора $e_j = (0, \dots, 0, 1, 0, \dots, 0)$, где 1 стоит на j-м месте. Тогда

$$\langle x^*, x \rangle = \sum_{i=1}^d x_i x_i^*, \quad x_j^* = \langle x^*, e_j \rangle$$

Сформулируем теперь упомянутое выше следствие:

Следствие 14.1 *Пусть $f_i \in C(U \cap K, \mathbb{R})$ $(i = 0, \dots, m)$. Рассмотрим задачу $(*)$:*

$$f_0(x) \to min$$
$$f_i(x) = 0 \ (i = 1, \dots, m), \ x \in U \cap K$$

Пусть $\hat{x}$ — locmin в $(*)$ и существует односторонняя производная $f_i'(\hat{x})$.

Тогда $\exists \overline{\lambda} = (\lambda_0, \dots, \lambda_m) \neq 0$:

1) для $\mathcal{L}(x, \overline{\lambda}) = \sum_{i=0}^m \lambda_i f_i(x)$ выполнено $\mathcal{L}_x(\hat{x}, \overline{\lambda}) \geq 0$ $(\Leftrightarrow \sum_{i=0}^m \lambda_i f_i'(\hat{x}) \geq 0$, это покомпонентное неравенство)

2) $\lambda_0 \geq 0$

□ Определим функции

$$\varphi_i(x) = -x_i \ (i = 1, \dots, d) \text{ — определены везде}$$

$$\varphi_i(x) = f_{i-d}(x) \ (i = d+1, \dots, d+m), \ \varphi_0(x) = f_0(x) \text{ —}$$
$$\text{— определены в } U \cap K$$

Тогда исходная задача записывается в следующем виде:

$$\varphi_0(x) \to min$$
$$\varphi_i(x) \le 0 \ (i = 1, \dots, d)$$
$$\varphi_i(x) = 0 \ (i = d+1, \dots, d+m)$$

Положим $x_i^* = \varphi_i'(0)$. Заметим, что для некоторых i это односторонняя производная, а для остальных — обычная, а именно:

$$x_0^* = f_0'(0)$$
$$x_i^* = f_{i-d}'(0) \ (i = d+1, \dots, d+m)$$
$$\langle x_i^*, x\rangle = -x_i \ (i = 1, \dots, d)$$

Теперь можем воспользоваться принципом Лагранжа:

$\exists \overline{\mu} = (\mu_0, \dots, \mu_{d+m}) \neq 0:$

1) $\sum_{i=0}^{d+m} \mu_i x_i^* = 0$

2) $\mu_i \ge 0 \ (i = 0, \dots, d)$

Заметим, что условие дополняющей нежесткости здесь не содержательно, поэтому мы его и не пишем.

Далее имеем

$$0 = \langle \sum_{i=0}^{m+d} \mu_i x_i^*, e_j\rangle = \langle \mu_0 x_0^* + \sum_{i=d+1}^{m+d} \mu_i x_i^*, e_j\rangle + \sum_{i=1}^{d} \mu_i \langle x_i^*, e_j\rangle =$$

$$= \{\langle x_i^*, e_j\rangle = \begin{cases} 0, i \neq j \\ -1, i = j \end{cases} \quad \text{т.к. это } i\text{-я координата вектора } e_j \text{ со знаком } -\} =$$

$$= \langle \mu_0 x_0^* + \sum_{i=d+1}^{m+d} \mu_i x_i^*, e_j\rangle - \mu_j$$

Отсюда, вспоминая, что $\mu_j \ge 0$, получаем:

$$\langle \mu_0 x_0^* + \sum_{i=d+1}^{m+d} \mu_i x_i^*, e_j\rangle \ge 0$$

j-я координата вектора $\mu_0 x_0^* + \sum_{i=d+1}^{m+d} \mu_i x_i^*$ не меньше нуля для любого j тогда и только тогда, когда $\mu_0 x_0^* + \sum_{i=d+1}^{m+d} \mu_i x_i^* \geq 0$. Здесь, напомним, $x_i^* = f'_{i-d}(0)$.

Положим $\mu_i = \lambda_{i-d}$, $\lambda_0 = \mu_0$. Тогда

$$\mu_0 x_0^* + \sum_{i=d+1}^{m+d} \mu_i x_i^* = \lambda_0 f'_0(0) + \sum_{i=d+1}^{d+m} \lambda_{i-d} f_{i-d}(0) \geq 0 \Leftrightarrow$$

$$\Leftrightarrow \sum_{i=1}^{m} \lambda_i f'_i(0) \geq 0$$

Т.о. искомый $\overline{\lambda}$ найден. Осталось показать, что $\overline{\lambda} = (\lambda_0, \ldots, \lambda_m) \neq 0$. Действительно, положим, $\overline{\lambda} = 0$. Но тогда $\forall j = 1, \ldots, d$ $\mu_j = 0$, т.к.

$$\langle \mu_0 x_0^* + \sum_{i=d+1}^{m+d} \mu_i x_i^*, e_j \rangle = \mu_j$$

Получили противоречие с условием $(\mu_0, \ldots, \mu_{m+d}) \neq 0$.

■

15 Задача Лагранжа.

Рассмотрим еще одну экстремальную задачу, задачу Лагранжа. Как мы увидим, все рассмотренные ранее задачи являются частными случаями задачи Лагранжа.

Итак, пусть $\Delta = [a, b] \subset \mathbb{R},\ a < b,\ x(\cdot) \in C^1(\Delta, \mathbb{R}^n)$ — так называемая *фазовая переменная*. Пусть задана функция $u(\cdot) \in C(\Delta, \mathbb{R}^r)$, которую назовем *управлением*. Пусть $t_0 < t_1,\ t_0, t_1 \in int\Delta$.

Рассмотрим четверку $\xi = (x(\cdot), u(\cdot), t_0, t_1)$ — *управляемый процесс*.

На самом деле нас интересует функция x, но непосредственно выбирать ее мы не можем. Все, что нам доступно — это "управление" функцией u, которая "косвенно влияет" на x. Отсюда и появилось название *управление*. Например, такая ситуация возникает при движении ракеты, когда мы можем управлять ускорением, косвенно влияя на скорость ракеты.

Далее имеем $\dot{x} = \varphi(t, x, u)$ — *дифференциальная связь*.

Также задано множество функционалов

$$\mathcal{B}_i(\xi) = \int_{t_0}^{t_1} f_i(t, x(t), u(t))dt + \psi_i(t_0, x(t_0), t_1, x(t_1))$$

Рассмотрим следующую экстремальную задачу:

$$\mathcal{B}_0(\xi) \to min$$
$$\mathcal{B}_i(\xi) \leq 0\ (i = 1, \ldots, m')$$
$$\mathcal{B}_i(\xi) = 0\ (i = m' + 1, \ldots, m)$$

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

При этом f_i, ψ_i, φ непрерывны.

ξ — *допустимый управляемый процесс*, если $\mathcal{B}_i(\xi)$ определены, определена $\varphi(t, x(t), u(t))$, $t \in [t_0, t_1]$ и выполнены ограничения типа неравенств, равенств и дифференциальная связь.

Заметим, что в данной задаче нет ограничений на управление в отличие от *задачи оптимального управления*, которую рассмотрим позже.

Допустимый управляемый процесс $\hat{\xi}$ — называется *оптимальным в слабом смысле*, если $\exists\ \varepsilon > 0$: для любого допустимого управляемого процесса ξ такого, что

$$\|x(\cdot) - \hat{x}(\cdot)\|_{C^1(\triangle, \mathbb{R}^n)} < \varepsilon;$$
$$\|u(\cdot) - \hat{u}(\cdot)\|_{C(\triangle, \mathbb{R}^r)} < \varepsilon;$$
$$|t_0 - \hat{t}_0| < \varepsilon;$$
$$|t_1 - \hat{t}_1| < \varepsilon;$$

выполнено $\mathcal{B}_0(\xi) \geq \mathcal{B}_0(\hat{\xi})$.

Теорема 15.1 *(необходимые условия оптимальности)*

Пусть f_i , φ — непрерывны вместе с производными по x, u, пусть также $\psi_i \in C^1$; Пусть $\hat{\xi}$ — оптимальный управляемый процесс. Тогда существует набор множителей Лагранжа, т.е. $\exists\ (\bar{\lambda}, p(\cdot)) \neq 0$, где $\lambda = (\lambda_0, \ldots, \lambda_m)$, $p(\cdot) \in C^1(\triangle, \mathbb{R}^n)$:

$$\mathcal{L}(\xi, \bar{\lambda}, p(\cdot)) = \sum_{i=0}^{m} \lambda_i \mathcal{B}_i(\xi) + \int\limits_{t_0}^{t_1} \langle p(\cdot), \dot{x}(\cdot) - \varphi(\cdot, x, u) \rangle \, dt =$$

$$= \int\limits_{t_0}^{t_1} \underbrace{L(t, x(t), \dot{x}(t), u(t))}_{\textit{лагранжан}} dt + \underbrace{l(t_0, x(t_0), t_1, x(t_1))}_{\textit{терминант}}$$

и выполняются следующие условия:

1) уравнение Эйлера : $\dfrac{d}{dt}\hat{L}_{\dot{x}}(t) = \hat{L}_x(t)\ \forall\ t \in [\hat{t}_0, \hat{t}_1]$;

2) условия трансверсальности $\hat{L}_{\dot{x}}(\hat{t}_i) = (-1)^i \cdot \hat{l}_{x(t_i)}$ $(i = 0, 1)$;

3) стационарность по подвижным концам : $\hat{\mathcal{L}}_{t_i} = 0$ $(i = 0, 1)$;

4) стационарность по управлению : $\hat{L}_u(t) = 0\ \forall\ t \in [\hat{t}_0, \hat{t}_1]$;

5) неотрицательность : $\lambda_i \geq 0$ $(i = 0, \ldots, m')$;

6) дополняющая нежесткость : $\lambda_i \mathcal{B}_i(\hat{\xi}) = 0$ $(i = 1, \ldots, m')$;

□ Без доказательства. ■

Заметим, что если t_0 или t_1 фиксированны, то мы просто не пишем условие стационарности по этому концу.

Покажем, что простейшая задача является задачей Лагранжа:

$$\int_{t_0}^{t_1} f(t, x(t), \dot{x}(t))dt \to min;$$

$$x(t_0) = x_0;$$
$$x(t_1) = x_1;$$

(если нет ограничений типа неравенств, то можно рассматривать эту задачу на экстремум.)

Теперь положим:

$$u := \dot{x},\ \xi = (x(\cdot), u(\cdot));$$

$$\mathcal{B}_0(\xi) = \int_{t_0}^{t_1} f(t, x(t), u(t))dt \to min;$$

$$\mathcal{B}_1(\xi) = x(t_0) - x_0 = 0;$$
$$\mathcal{B}_2(\xi) = x(t_1) - x_1 = 0;$$
$$\dot{x}(t) - u(t) = 0;$$

Следовательно произошло сведение к задаче Лагранжа.

по предыдущей теореме имеем:

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

$$+\lambda_1(x(t_0) - x_0) + \lambda_2(x(t_1) - x_1);$$

1) $\dot{p}(t) = \lambda_0 \hat{f}_x(t)$;

2) $p(t_0) = \lambda_1,\ p(t_1) = -\lambda_2$;

3) это условие мы не пишем, так как у нас нет подвижных концов.

4) $\lambda_0 \hat{f}_u(t) = p(t)$;

Заметим, что если $\lambda_0 = 0$, то $p(\cdot) = 0$ и $\lambda_1 = \lambda_2 = 0$, поэтому без ограничения общности положим, что $\lambda_0 = 1$ ($\neq 0$). Вывод уравнения Эйлера для простейшей задачи не представляет труда.

Задача 8 *Выписать необходимые условия экстремума для задачи:*

$$\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t), \ddot{x}(t))dt \to extr$$

(и имеет место полный набор условий первого порядка)

Подсказка: $L \in C^2$, в качестве фазовой переменной взять $(x_1, x_2) = (x, \dot{x})$, в качестве управления взять $u = \ddot{x}$ и в качестве уравнения дифференциальной связи положить:

$$\begin{cases} \dot{x}_1 - x_2 = 0 \\ \dot{x}_2 - u = 0 \end{cases}$$

16 Задача оптимального управления.

Пусть $\triangle = [a,b] \subset \mathbb{R},\ x(\cdot) \in PC^1(\triangle, \mathbb{R}^n)$ — *фазовая переменная.* $u(\cdot) \in PC(\triangle, \mathbb{R}^r)$ — *управление.* $\xi = (x(\cdot), u(\cdot), t_0, t_1)$ — *управляемый процесс;*

$$\mathcal{B}_i(\xi) = \int\limits_{t_0}^{t_1} f_i(t, x(t), u(t))dt + \psi_i(t_0, x(t_0), t_1, x(t_1));$$

$$\mathcal{B}_0(\xi) \to min;$$
$$\mathcal{B}_i(\xi) \leq 0 \ (i = 1, \dots, m^{'});$$
$$\mathcal{B}_i(\xi) = 0 \ (i = m^{'} + 1, \dots, m);$$
$$\dot{x}(t) - \varphi(t, x(t), u(t)) = 0 \text{ — уравнение дифф. связи в точках непрерывности } u(\cdot);$$
$$\forall\ t \in \triangle\ u(\cdot) \in U\ (U \subset \mathbb{R}^r - fixed) \text{ — ограничение на управление.}$$

Допустимый управляемый процесс $\hat{\xi}$ — называется *оптимальным в сильном смысле*, если $\exists\ \varepsilon > 0$: для любого допустимого управляющего процесса ξ :

$$\|x(\cdot) - \hat{x}(\cdot)\|_{C(\triangle, \mathbb{R}^n)} < \varepsilon;$$
$$|t_0 - \hat{t}_0| < \varepsilon;$$
$$|t_1 - \hat{t}_1| < \varepsilon;$$

выполнено $\mathcal{B}_0(\xi) \geq \mathcal{B}_0(\hat{\xi})$.

Теорема 16.1 *(необходимые условия оптимальности или принцип максимума Понтрягина)*

Пусть f_i , φ — непрерывны вместе с производными по x, u и $\psi_i \in C^1$; Пусть $\hat{\xi}$ — оптимальный управляемый процесс. Тогда существует набор множителей Лагранжа, т.е. $\exists\ (\bar{\lambda}, p(\cdot)) \neq 0$, где $\lambda = (\lambda_0, \dots, \lambda_m)$, $p(\cdot) \in PC^1(\triangle, \mathbb{R}^n)$:

$$\mathcal{L}(\xi, \bar{\lambda}, p(\cdot)) = \sum_{i=0}^{m} \lambda_i \mathcal{B}_i(\xi) + \int\limits_{t_0}^{t_1} \langle p(\cdot), \dot{x}(\cdot) - \varphi(\cdot, x, u) \rangle\, dt =$$

$$= \int\limits_{t_0}^{t_1} \underbrace{L(t, x(t), \dot{x}(t), u(t))}_{\text{лагранжан}} dt + \underbrace{l(t_0, x(t_0), t_1, x(t_1))}_{\text{терминант}}$$

и выполняются следующие условия:

1) уравнение Эйлера : $\dfrac{d}{dt}\hat{L}_{\dot{x}}(t) = \hat{L}_x$ *в точках непрерывности* $u(\cdot)$*;*

2) условия трансверсальности $\hat{L}_{\dot{x}}(\hat{t}_i) = (-1)^i \cdot \hat{l}_{x(t_i)}$ $(i = 0, 1)$*;*

3) стационарность по подвижным концам : $\hat{\mathcal{L}}_{t_0} = 0;\ \hat{\mathcal{L}}_{t_1} = 0$*;*

4) $\min\limits_{v \in U} L(t, \hat{x}(t), \dot{\hat{x}}(t), v) = L(t, \hat{x}(t), \dot{\hat{x}}(t), \hat{u}(t))$*, где t — точка непрерывности* $u(\cdot)$*;*

5) неотрицательность : $\lambda_i \geq 0$ $(i = 0, \dots, m^{'})$*;*

6) дополняющая нежесткость : $\lambda_i \mathcal{B}_i(\hat{\xi}) = 0$ $(i = 1, \dots, m^{'})$*;*

$\square$ Ниже мы будем доказывать эту теорему в одном частном случае. $\blacksquare$

17 Задача со свободным концом.

Пусть $t_0,\ t_1$ — fixed, $\xi = (x(\cdot), u(\cdot))$, $A = \{t \in [t_0, t_1] : u(\cdot)$ непрерывна в $t\}$

$$\mathcal{B}_0(\xi) = \int\limits_{t_0}^{t_1} f(t, x(t), u(t))dt + \psi(x(t_1)) \to \min;$$

$$x(t_0) = x_0;$$

$$\dot{x}(t) = \varphi(t, x(t), u(t)) \ \forall\ t \in A, u(t) \in U$$

В этом случае предыдущая теорема имеет довольно простое доказательство.

Мы покажем, что можно взять $\lambda_0 = 1$. Проверим, что существуют $p(\cdot)$ и $\overline{\lambda}$ такие, что

$$\mathcal{L}(\xi, \bar{\lambda}, p(\cdot)) = \int\limits_{t_0}^{t_1} \underbrace{(f(t, x(t), u(t)) + \langle p(t), \dot{x}(t) - \varphi(t, x(t), u(t))\rangle)}_{\text{лагранжан}} dt +$$

$$\underbrace{+\,\psi(x(t_1)) + \sum_{i=1}^{n} \lambda_i (x_i(t_0) - (x_0)_i)}_{\text{терминант}}$$

и выполняются условия:

1) уравнение Эйлера : $\dot{p}(t) = \hat{f}_x(t) - p(t)\hat{\varphi}_x(t)\ \forall\ t \in A$;

2) трансверсальность : $p_i(t_0) = \lambda_i\ (i = 1, \dots, n),\ p(t_1) = -\hat{\psi}_{x(t_1)}$;

3) это условие мы не пишем, т.к. подвижных концов нет;

4) $\forall\ t \in A\ L(t, \hat{x}(t), \dot{\hat{x}}(t), \hat{u}(t)) = \min\limits_{v \in U} L(t, \hat{x}(t), \dot{\hat{x}}(t), v) \Leftrightarrow$

$$\forall\ t \in A, \forall\ v \in U\ \ L(t, \hat{x}(t), \dot{\hat{x}}(t), v) \geq L(t, \hat{x}(t), \dot{\hat{x}}(t), \hat{u}(t))$$

Т.о. нам нужно показать, что

$$f(t, \hat{x}(t), v) + \left\langle p(t), \dot{\hat{x}}(t) \right\rangle - \langle p(t), \varphi(t, \hat{x}(t), v)\rangle \geq$$

$$\geq f(t, \hat{x}(t), \hat{u}(t)) + \left\langle p(t), \dot{\hat{x}}(t) \right\rangle - \langle p(t), \varphi(t, \hat{x}(t), \hat{u}(t))\rangle \quad (*)$$

□ Рассмотрим функцию $p(\cdot)$, она определена следующими условиями из пунктов 1) и 2):

$$\begin{cases} \dot{p}(t) = \hat{f}_x(t) - p(t)\hat{\varphi}_x(t)\ \forall\ t \in A \quad (1) \\ p(t_1) = -\hat{\psi}_{x(t_1)} \quad (2) \end{cases}$$

Как мы знаем из дифференциальных уравнений, такая задача имеет единственное решение на интервалах непрерывности $u(t)$. Покажем, что решение единственно и на всем $[t_0, t_1]$.

Рассмотрим разбиение отрезка $[t_0, t_1]$ вида $t_0 < \tau_1 < \tau_2 < \ldots < \tau_l < t_1$, где $\tau_i \quad (i = 1, \dots, l)$ — точки разрыва функции $u(\cdot)$. Тогда на каждом отрезке разбиения у выражения (1) непрерывная правая часть, поэтому из условий (1), (2)

$p(\cdot)$ определяется однозначно на отрезке $[\tau_l, t_1]$, далее на отрезке $[\tau_{l-1}, \tau_l]$ и т.д — в конце концов, на всем отрезке $[t_0, t_1]$ (рис. 9).

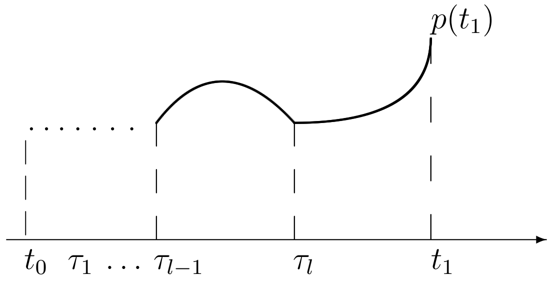


Рис. 9:

Теперь $\forall\ \tau \in A : \tau > t_0$, $v \in U$, $\alpha > 0$ — некоторое малое число, рассмотрим следующую игольчатую вариацию (семейство управляемых процессов):

$$u_{\tau,v,\alpha}(t) = u_\alpha(t) = \begin{cases} \hat{u}(t),\ t \notin (\tau - \alpha, \tau] \\ v,\ t \in (\tau - \alpha, \tau] \end{cases}$$

Пусть также $x_{\tau,v,\alpha}(t) = x_\alpha(t)$, где x_α — решение следующей задачи:

$$\begin{cases} \dot{x}_\alpha = \varphi(t, x_\alpha, u_\alpha),\ t \in A, t \neq \tau - \alpha, t \neq \tau \\ x_\alpha(t_0) = x_0 \end{cases}$$

Опять же, из дифференциальных уравнений следует, что решение x_α такой задачи существует и единственно, а именно:

Лемма 17.1 (об игольчатых вариациях)

1) при малом $\alpha > 0$ $\exists$ решение $x_\alpha(\cdot)$;

2) $x_\alpha(\cdot) \rightrightarrows \hat{x}(\cdot)$ при $\alpha \to 0+$;

3) $\lim_{\alpha \to 0+} \frac{x_\alpha(t) - \hat{x}(t)}{\alpha} = y(t)$ (сходимость равномерная на $[\tau, t_1]$)

Здесь $y(\cdot)$ — решение следующей задачи

$$\begin{cases} \dot{y}(t) = \hat{\varphi}_x(t) y(t)\ (t > \tau) \\ y(\tau) = \varphi(\tau, \hat{x}(\tau), v) - \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau)) \end{cases}$$

(решение такой задачи существует и единственно).

□ Без доказательства. ■

Для доказательства неравенства (∗) будем сравнивать значения исходного функционала $\mathcal{B}_0$ на $(\hat{x}(\cdot), \hat{u}(\cdot))$ и $(x_\alpha(\cdot), u_\alpha(\cdot))$. Начнем с нетерминантной части:

$$\int_{t_0}^{t_1} f(t, x_\alpha(t), u_\alpha(t))dt - \int_{t_0}^{t_1} f(t, \hat{x}(t), \hat{u}(t))dt =$$

$$= \int_{t_0}^{\tau-\alpha} (f(t, x_\alpha(t), u_\alpha(t)) - f(t, \hat{x}(t), \hat{u}(t)))dt +$$

$$+\int_{\tau-\alpha}^{\tau}(f(t,x_\alpha(t),u_\alpha(t))-f(t,\hat{x}(t),\hat{u}(t)))dt+$$

$$+\int_{\tau}^{t_1}(f(t,x_\alpha(t),u_\alpha(t))-f(t,\hat{x}(t),\hat{u}(t)))dt =: I_0+I_1+I_2 =$$

$$=\{x_\alpha(\cdot)=\hat{x}(\cdot), u_\alpha(\cdot)=\hat{u}(\cdot) \text{ на } [t_0,\tau-\alpha] \Rightarrow I_0=0\} = I_1+I_2$$

Для нахождения I_1 воспользуемся теоремой о среднем. Для этого выберем α таким малым, чтобы подинтегральная функция была непрерывна в $[\tau-\alpha,\tau]$.

Имеем

$$I_1=\alpha(f(\tilde{\tau},x_\alpha(\tilde{\tau}),u_\alpha(\tilde{\tau}))-f(\tilde{\tau},\hat{x}(\tilde{\tau}),\hat{u}(\tilde{\tau}))), \tilde{\tau}\in(\tau-\alpha,\tau]$$

Устремим $\alpha\to 0$. Тогда $\tilde{\tau}\to\tau$, $x_\alpha(\tilde{\tau})\to\hat{x}(\tau)$, $u_\alpha(\tilde{\tau})=v$, $\hat{x}(\tilde{\tau})\to\hat{x}(\tau), \hat{u}(\tilde{\tau})\to\hat{u}(\tau)$. Значит,

$$I_1=\alpha(f(\tau,\hat{x}(\tau),v)-f(\tau,\hat{x}(\tau),\hat{u}(\tau)))+o(\alpha)$$

Теперь разберемся с I_2.

$$f(t,x_\alpha(t),u_\alpha(t))-f(t,\hat{x}(t),\hat{u}(t))=\langle\hat{f}_x(t),(x_\alpha(t)-\hat{x}(t))\rangle+o(|x_\alpha(t)-\hat{x}(t)|)=$$

$$=\{\text{в силу леммы об игольчатых вариациях}\}=$$

$$=\langle\hat{f}_x(t),(\alpha y(t)+o(\alpha))\rangle+o(\alpha y(t)+o(\alpha))=\alpha\langle\hat{f}_x(t),y(t)\rangle+o(\alpha)$$

Итак,

$$I_2=\alpha\int_{\tau}^{t_1}\langle\hat{f}_x(t),y(t)\rangle dt+o(\alpha)$$

Но в таком представлении с I_2 работать неудобно, будем его дальше преобразовывать. Для этого положим $p(t_1)=-\mu$ и подсчитаем $\langle p(t_1),y(t_1)\rangle-\langle p(\tau),y(\tau)\rangle$ двумя способами.

С одной стороны,

$$\langle p(t_1),y(t_1)\rangle-\langle p(\tau),y(\tau)\rangle=-\langle\mu,y(t_1)\rangle-\langle p(\tau),\varphi(\tau,\hat{x}(\tau),v)-\varphi(\tau,\hat{x}(\tau),\hat{u}(\tau))\rangle$$

С другой по формуле Лейбница из уравнения Эйлера $\dot{p}(t)=\hat{f}_x(t)-p(t)\hat{\varphi}_x(t)$ имеем

$$\langle p(t_1),y(t_1)\rangle-\langle p(\tau),y(\tau)\rangle=\int_{\tau}^{t_1}(\langle\hat{f}_x(t),y(t)\rangle-\langle p(t)\hat{\varphi}_x(t),y(t)\rangle+\langle p(t)\hat{\varphi}_x(t),y(t)\rangle)dt=$$

$$=\int_{\tau}^{t_1}\langle\hat{f}_x(t),y(t)\rangle dt$$

Итак,

$$I_2=\alpha(-\langle\mu,y(t_1)\rangle-\langle p(\tau),(\varphi(\tau,\hat{x}(\tau),v)-\varphi(\tau,\hat{x}(\tau),\hat{u}(\tau))\rangle)+o(\alpha)$$

Разберемся и с терминантным членом: $x_\alpha(t_1)-\hat{x}(t_1)=\alpha y(t_1)+o(\alpha)$.

Тогда

$$\psi(x_\alpha(t_1))-\psi(\hat{x}(t_1))=\langle\hat{\psi}_{x(t_1)},\alpha y(t_1)\rangle+o(\alpha)$$

Теперь уже можем оценить $\mathcal{B}_0(x_\alpha(\cdot), u_\alpha(\cdot)) - \mathcal{B}_0(\hat{x}(\cdot), \hat{u}(\cdot))$:

$$0 \leq \mathcal{B}_0(x_\alpha(\cdot), u_\alpha(\cdot)) - \mathcal{B}_0(\hat{x}(\cdot), \hat{u}(\cdot)) = \alpha(f(\tau, \hat{x}(\tau), v) - f(\tau, \hat{x}(\tau), \hat{u}(\tau)) - \langle \hat{\psi}_{x(t_1)}, y(t_1)\rangle -$$

$$-\langle p(\tau), (\hat{\varphi}(\tau, \hat{x}(\tau), v) - \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau)))\rangle + \langle \hat{\psi}_{x(t_1)}, y(t_1)\rangle) + o(\alpha)$$

Значит, выражение, стоящее в скобках при α — неотрицательно. Но это и есть нужное нам условие оптимальности по $u(\cdot)$. ■

Рассмотрим теперь еще одну задачу.

$$\mathcal{B}_0(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} f(t, x(t), u(t))dt \to min$$
$$x(t_0) = x_0, x(t_1) = x_1$$

Ограничение на управление:

$$\dot{x}(t) = \varphi(t, x(t), u(t)), \quad t \in A, u(t) \in U$$

Докажем теорему Болтянского — принцип максимума Понтрягина для этой задачи. Имеем:

$$\mathcal{L} = \lambda_0 \mathcal{B}_0 + \sum_{i=1}^{n} \lambda_i (x_i(t_0) - (x_0)_i) + \sum_{i=1}^{n} \mu_i (x_i(t_1) - (x_1)_i) +$$
$$+ \int_{t_0}^{t_1} \langle p(t), \dot{x}(t) - \varphi(t, x(t), u(t))\rangle dt$$

Напишем необходимые условия экстремума:

1) $\dot{p}(t) = \lambda_0 \hat{f}_x(t) + \langle p(t), \hat{\varphi}_x(t)\rangle$ $(t \in A)$ (заметим, что здесь $\hat{\varphi}_x(t)$ — уже матрица)

2) $p_i(t_0) = \lambda_i$, $p_i(t_1) = -\mu_i$ $(i = 1 \ldots n)$

3) условий стационарности нет, т.к. нет подвижных концов

4) $\forall \tau \in A$ $\forall v \in U$

$$\lambda_0(f(\tau, \hat{x}(\tau), v) - f(\tau, \hat{x}(\tau), \hat{u}(\tau))) - \langle p(\tau), \varphi(\tau, \hat{x}(\tau), v) - \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau))\rangle \geq 0$$

5) $\lambda_0 \geq 0$

Здесь, как и в предыдущей задаче, снова будем использовать игольчатые вариации, но более сложные. Раньше была одна «иголка», т.е. один интервал, где мы что-то меняли. А теперь таких интервалов будет несколько (игольчатая вариация с пакетом иголок). Итак, положим

$$\overline{\tau} = (\tau_1, \ldots, \tau_l) \ (\tau_j \in A, \tau_i \neq \tau_j),$$

$$\overline{\alpha} = (\alpha_1, \ldots, \alpha_l) \geq 0, |\alpha| \text{ мал}$$

Определим

$$u_{\overline{\tau}, \overline{v}, \overline{\alpha}}(t) = u_{\overline{\alpha}}(t) = \begin{cases} v_j, \text{ если } t \in (\tau_j - \alpha_j, \tau_j] \\ \hat{u}(t) \text{ в противном случае} \end{cases}$$

Заметим, что α_j должны быть настолько малы, чтобы интервалы $(\tau_j - \alpha_j, \tau_j]$ лежали внутри $[t_0, t_1]$ и не пересекались.

Пусть далее $x_{\overline{\tau},\overline{v},\overline{\alpha}}(t) = x_{\overline{\alpha}}(t)$, где $x_{\overline{\alpha}}$ — решение задачи

$$\begin{cases} \dot{x}_{\overline{\alpha}}(t) = \varphi(t, x_{\overline{\alpha}}(t), u_{\overline{\alpha}}(t)) \text{ в тех точках, где } u_{\overline{\alpha}} \text{ непрерывна} \\ x_{\overline{\alpha}}(t_0) = x_0 \end{cases}$$

Трудность по сравнению с предыдущей задачей в том, что из условий не вытекает, что $x_{\overline{\alpha}}$ в правом конце попадает куда нужно, т.к. теперь правый конец не является свободным.

Лемма 17.2 (об игольчатых вариациях)

1) $|\alpha|$ мал $\Rightarrow \exists$ решение $x_{\overline{\alpha}}(\cdot)$

2) $x_{\overline{\alpha}}(\cdot) \rightrightarrows \hat{x}(\cdot)$ при $\overline{\alpha} \to 0$

3)

$$\frac{\partial x_{\overline{\alpha}}(t_1)}{\partial \alpha_j}(0) = y_j(t_1)$$

где y_j — решение дифференциального уравнения

$$\begin{cases} \dot{y}_j(t) = \hat{\varphi}_x(t) y_j(t) \ (t \geq \tau_j) \\ y_j(\tau_j) = \varphi(\tau_j, \hat{x}(\tau_j), v_j) - \varphi(\tau_j, \hat{x}(\tau_j), \hat{u}(\tau_j)) \end{cases}$$

$$x_{\overline{\alpha}}(t_1) = \hat{x}(t_1) + \sum_{j=1}^{l} \frac{\partial x_{\overline{\alpha}}(t_1)}{\partial \alpha_j}(0)\alpha_j + o(|\overline{\alpha}|) \ (\overline{\alpha} \to 0)$$

4) $x_{\overline{\alpha}}(t_1)$ непрерывна по $\overline{\alpha}$, когда $|\overline{\alpha}|$ мал

Лемма 17.3 (о вариации интегрального функционала)

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

$$p_i(t_1) = -\mu_i \ (i = 1, \dots, n), \ \mu = (\mu_1, \dots, \mu_n)$$

$$g(\overline{\alpha}) = \int_{t_0}^{t_1} \lambda_0 f(t, x_{\overline{\alpha}}(t), u_{\overline{\alpha}}(t)) dt$$

Тогда

1)

$$\frac{\partial g(\overline{\alpha})}{\partial \alpha_j}(0) = \lambda_0 (f(\tau_j, \hat{x}(\tau_j), v_j) - f(\tau_j, \hat{x}(\tau_j), \hat{u}(\tau_j))) -$$

$$- \langle p(\tau_j), \varphi(\tau_j, \hat{x}(\tau_j), v_j) - \varphi(\tau_j, \hat{x}(\tau_j), \hat{u}(\tau_j)) \rangle - \mu y_j(t_1)$$

$$g(\overline{\alpha}) = g(0) + \sum_{j=1}^{l} \alpha_j \frac{\partial g(\overline{\alpha})}{\partial \alpha_j}(0) + o(|\alpha|) \ (|\alpha| \to 0)$$

2) g — непрерывна по $\overline{\alpha}$

Итак, имеем следующую задачу с ограничениями типа равенств и неравенств.

$$\mathcal{B}_0(x_{\overline{\alpha}}(\cdot)) \to min$$

$$x_{\overline{\alpha}}(t_1) = 0,\ \overline{\lambda} \geq 0$$

В силу следствия 14.1 принципа Лагранжа в задаче с ограничениями типа равенств и неравенств существует набор множетелей Лагранжа $\bar{\mu} = (\mu_0, \ldots, \mu_n),\ \mu_0 \geq 0$ таких, что

$$\sum_{i=0}^{n} \mu_i g_i^{'}(0) \geq 0$$

последнее в координатах можно записать, как

$$\forall j\ \sum_{i=0}^{n} \mu_i \frac{\partial g_i}{\partial \alpha_j}(0) \geq 0$$

Этим свойством мы будем пользоваться для того, чтобы получить условие оптимальности по управлению.

Положим $\lambda_0 = \mu_0$, пусть $\mu = (\mu_1, \ldots, \mu_n)$. В качестве $p(\cdot)$ рассмотрим функцию, удовлетворяющую уравнению Эйлера

$$\dot{p}(t) = \lambda_0 \hat{f}_x(t) - p(t)\hat{\varphi}_x(t) \qquad (t \in A)$$

Добавив к нему условие трансверсальности, получаем задачу Коши

$$\begin{cases} \dot{p}(t) = \lambda_0 \hat{f}_x(t) - p(t)\hat{\varphi}_x(t) \qquad (t \in A) \\ p(t_1) = -\mu \qquad (p_i(t_1) = -\mu_i \quad \forall i) \end{cases}$$

решение которой, как известно, существует и единственно. Т.о. все необходимые условия экстремума выполняются автоматом, кроме условия оптимальности по управлению. Проверим его.

По лемме 17.3 о вариации интегрального функционала имеем

$$\lambda_0 \frac{\partial g_0}{\partial \alpha_j}(0) = \lambda_0 (f(\tau_j, \hat{x}(\tau_j), v_j) - f(\tau_j, \hat{x}(\tau_j), \hat{u}(\tau_j))) -$$

$$-p(\tau_j)(\varphi(\tau_j, \hat{x}(\tau_j), v_j) - \varphi(\tau_j, \hat{x}(\tau_j), \hat{u}(\tau_j))) - \mu y_j(t_1)$$

где y — решение следующих задач из леммы 17.2 об игольчатых вариациях:

$$\begin{cases} \dot{y}_j = \hat{\varphi}_x y_j \\ y_j(\tau_j) = \varphi(\tau_j, \hat{x}(\tau_j), v_j) - \varphi(\tau_j, \hat{x}(\tau_j), \hat{u}(\tau_j)) \end{cases}$$

Поскольку $\dfrac{\partial g_i}{\partial \alpha_j}(0) = (y_j)_i(t_1)$, где $(y_j)_i$ — i-ая координата вектора y_j, имеем

$$\sum_{i=1}^{n} \mu_i \frac{\partial g_i}{\partial \alpha_j}(0) = \mu y_j(t_1)$$

Значит,

$$\sum_{i=0}^{n} \mu_i \frac{\partial g_i}{\partial \alpha_j}(0) = \lambda_0(f(\tau_j, \hat{x}(\tau_j), v_j) - f(\tau_j, \hat{x}(\tau_j), \hat{u}(\tau_j)))-$$

$$-p(\tau_j)(\varphi(\tau_j, \hat{x}(\tau_j), v_j) - \varphi(\tau_j, \hat{x}(\tau_j), \hat{u}(\tau_j)))$$

Как отмечалось выше, из следствия принципа Лагранжа можно сделать вывод, что последнее выражение должно быть ≥ 0.

Напомним, что для доказательства оптимальности по управлению нам нужно найти такие множители Лагранжа $\overline{\mu} = (\mu_0 = \lambda_0, \mu_1, \ldots, \mu_n),\ \mu_0 \geq 0$, что

$$\dot{p} = \lambda_0 \hat{f}_x - p\hat{\varphi}_x, \quad p(t_1) = -\bar{\mu}\ (= -(\mu_1, \ldots, \mu_n))$$

и $\forall \tau \in A,\ \forall v \in U$ имеет место неравенство:

$$\lambda_0(f(\tau, \hat{x}(\tau), v) - f(\tau, \hat{x}(\tau), \hat{u}(\tau)))-$$

$$-p(\tau)(\varphi(\tau, \hat{x}(\tau), v) - \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau))) \geq 0 \qquad (1)$$

Казалось бы, именно это мы и получили. Но на самом деле для каждого j (т.е. для каждого набора $(\tau_j, v_j),\ j = 1, \ldots, l;\ \tau_i \neq \tau_j$ при $i \neq j$) найдены свои множители Лагранжа, которые, вообще говоря, могут и не совпадать. Покажем, как можно выбрать множители Лагранжа, единые для любых наборов (τ, v).

Заметим для начала, что для полученного $\overline{\mu}$ можно провести нормировку и считать, что $\|\bar{\mu}\| = 1$, т.е. $\sum\limits_{i=0}^{n} \mu_i^2 = 1$. Последнее можно сделать заменой $\bar{\mu} \to \dfrac{\bar{\mu}}{\|\bar{\mu}\|}$ заметив, что при умножении всех множетелей Лагранжа на одно и то же число никакие равенства не изменятся.

Пусть $\bar{\tau} = (\tau_1, \ldots, \tau_l)$, где $\tau_1, \ldots, \tau_l$ не обязательно различны. Очевидно, $\exists\ \bar{\tau}^\nu \to \bar{\tau}$ при $\nu \to \infty$ (то есть $\tau_j^\nu \to \tau_j\ \forall j$), такие что $\forall \nu\ \tau_1^\nu, \ldots, \tau_l^\nu$ различны. Значит, как показано выше, по набору $(\bar{\tau}^\nu, v)$ можно выбрать $\bar{\mu}^\nu = (\mu_0^\nu, \ldots, \mu_n^\nu),\ \|\bar{\mu}^\nu\| = 1,\ \mu_0^\nu \geq 0$

В силу компактности последнего множества $\exists$ последовательность $\{\nu_i\}$: $\bar{\mu}^{\nu_i} \to \bar{\mu}$ $(i \to \infty)$. Переходя к пределу при $i \to \infty$, получаем неравенство (1) уже для произвольных наборов $(\tau, v) = (\tau_j, v_j)$, а не только с $\tau_i \neq \tau_j,\ i \neq j$. Теперь выберем единый $\bar{\mu}$ для всех наборов (τ, v).

$\forall \tau \in A\ \forall v \in U$ рассмотрим множество

$$M_{\tau,v} = \{\bar{\mu} = (\mu_0, \ldots, \mu_n) : \mu_0 \geq 0, \|\bar{\mu}\| = 1,$$

$$\text{соответствующая функция } p(\cdot) \text{ удовлетворяет неравенству } (1)\}$$

Далее заметим, что $\forall\ (\tau_1, v_1), \ldots, (\tau_l, v_l)$

$$\bigcap_{j=1}^{l} M_{\tau_j, v_j} \neq \varnothing$$

Также легко заметить, что $\forall\ (\tau, v)\ M_{\tau,v}$ — замкнуто $\Rightarrow$ компактно. Далее воспользуемся леммой из функционального анализа:

Лемма 17.4 (о центрированной системе) *Пусть имеется метрическое пространство и система его компактных подмножеств, причем пересечение конечного числа любых подмножеств системы не является пустым. Тогда пересечение всех подмножеств системы не пусто.*

В нашем случае получаем, что

$$\exists \bar{\mu} \in \bigcap_{(\tau,v)} M_{\tau,v} \neq \varnothing$$

и искомые множители Лагранжа $\bar{\mu} = (\mu_0, \dots, \mu_n)$, единые для всех наборов (τ, v), найдены.

18 Уравнение Беллмана и принцип максимума.

Задача оптимального быстродействия формулируется следующим образом:

$$T \to \min; \qquad (1)$$

$$x(0) = x_0; \quad g(x(T)) = 0; \qquad (2)$$

$$\dot{x}(t) = \varphi(x(t), u(t)); \qquad (3)$$

$$u(t) \in U. \qquad (4)$$

При этом $x(\cdot)$, $u(\cdot)$, $g(\cdot)$— вектор-функции со значениями в $\mathbb{R}^n$, $\mathbb{R}^r$, $\mathbb{R}^s$ соответственно. Уравнение (3) предполагается выполненным в точках непрерывности управления u. Требование (1) равносильно задаче минимизации T. Переменная t трактуется как время, а функция $x(t)$ — как движение материальной точки в пространстве $\mathbb{R}^n$. Мы обозначаем

$$\mathcal{M} = \{x : g(x) = 0\}.$$

Таким образом, задача оптимального быстродействия есть задача достижения точкой в $\mathbb{R}^n$, вначале находящейся в положении x_0, множества $\mathcal{M}$ за минимальное время, при том что в каждый момент времени имеются ограничения на вектор скорости движения точки в зависимости от ее текущего положения; эти ограничения определяются условиями (3) и (4).

Принцип максимума пишется при следующих предположениях гладкости: функция $\varphi(x, u)$ предполагается непрерывной и имеющей непрерывную производную по x, а функция $g(x)$ предполагается непрерывно дифференцируемой.

Согласно принципу максимума, если $(\hat{x}(\cdot), \hat{u}(\cdot), \hat{T})$ — оптимальный управляемый процесс, то найдутся множители Лагранжа $\lambda = (\lambda_0, \dots, \lambda_{n+s})$ и $p(\cdot) \in KC^1([0, T] \to (\mathbb{R}^n)^*)$ такие, что $(\lambda, p(\cdot)) \neq 0$ и для функции Лагранжа

$$\mathcal{L}(x(\cdot), u(\cdot), \lambda, p(\cdot)) = \int_0^T (\lambda_0 + p(t)(\dot{x}(t) - \varphi(x(t), u(t))))dt + \mu(x(0) - x_0) + \nu g(x(T)),$$

где $\mu = (\lambda_1, \dots, \lambda_n)$, $\nu = (\lambda_{n+1}, \dots, \lambda_{n+s})$, выполнены условия стационарности по x,T, оптимальности по u и неотрицательности λ_0:

1) уравнение Эйлера $\dot{p}(t) = -\hat{\varphi}_x(t)p(t)$;

2) условия трансверсальности $p(0) = \mu, \quad p(\hat{T}) = -\nu\hat{g}_{x(T)}$;

3) условие оптимальности по u: $-p(t)\hat{\varphi}(t) = \max\limits_{u\in U}(-p(t)\varphi(\hat{x}(t), u))$;

4) условие стационарности по подвижному концу: $\lambda_0 + \nu\hat{g}_{x(T)}\dot{\hat{x}}(\hat{T}) = 0$;

5) условие неотрицательности: $\lambda_0 \geq 0$.

При этом 1) и 3) выполнены в точках непрерывности управления u.

Предположим, что для любого $x \in \mathcal{M}$ мы имеем $\mathrm{Im} g_x(x) = \mathbb{R}^s$. В этом случае $\mathcal{M}$ является гладкой поверхностью размерности $n - s$. Если при этом $h \in \mathbb{R}^n$— касательный вектор к $\mathcal{M}$ в точке x, то мы имеем

$$\frac{d}{d\alpha} g(x + \alpha h)|_{\alpha=0} = 0,$$

или $g_x(x)h = 0$. В частности, если $h \in \mathbb{R}^n$— касательный вектор к $\mathcal{M}$ в точке $\hat{x}(\hat{T})$, то $\hat{g}_{x(T)}h = 0$. Второе условие трансверсальности влечет

$$p(\hat{T})h = -\nu\hat{g}_{x(T)}h = 0.$$

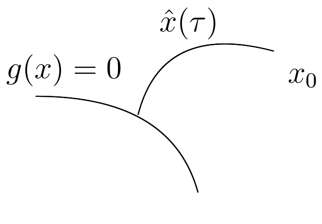
Следовательно, $p(\hat{T})$ ортогонален поверхности $\mathcal{M}$.

Лемма 18.1

$$\hat{\hat{x}}(t) = \hat{x}(t + \tau);$$
$$\hat{\hat{u}}(t) = \hat{u}(t + \tau);$$
$$\hat{\hat{T}} = \hat{T} - \tau$$

$\square$ Пусть $\tilde{x}(t) = \hat{x}(t + \tau)$; $\tilde{u}(t) = \hat{u}(t + \tau)$; $\tilde{T} = \hat{T} - \tau$ Далее имеем:

$$\hat{\hat{T}} \leq \hat{T} - \tau$$



Допустим, что есть допустимый процесс и $\tilde{T} < \hat{T} - \tau$. Тогда рассмотрим

$$x(t) = \begin{cases} \hat{x}(t) & \text{t} \leq \tau \\ \tilde{x}(t - \tau) & \text{t} > \tau \end{cases}$$

$$u(t) = \begin{cases} \hat{u}(t) & \text{t} \leq \tau \\ \tilde{u}(t - \tau) & \text{t} > \tau \end{cases}$$

Получаем, что $g(x(\tilde{T} + \tau)) = 0$; $T = \tilde{T} + \tau < \hat{T}$. $(x(\cdot), u(\cdot), T)$— допустимый процесс, $\tau \in (0, T) \Rightarrow \hat{\hat{T}} \geq \hat{T} - \tau$ $\blacksquare$

Предположим, что для любого $x_0 \in \mathbb{R}^n$ в (1) достигается глобальный минимум, равный $T(x_0)$; функция $\omega(x_0) = -T(x_0)$ называется *функцией Беллмана* . Если ω дифференцируема по Фреше в точке $x_0 \in \mathbb{R}^n \backslash \mathcal{M}$, то справедливо *уравнение Беллмана*

$$\max_{v\in U}\left(\frac{\partial\omega}{\partial x}(x_0)\varphi(x_0,v)\right)=1. \quad (5)$$

При этом максимум в (5) достигается для $v=u_0$, где $u_0=\hat{u}(0)$, где $(\hat{x}(\cdot),\hat{u}(\cdot),\hat{T})$— оптимальный управляемый процесс в задаче оптимального быстродействия.

Пусть $\omega(\hat{x}(t))=t-\hat{T}\Rightarrow$ можно записать

$$1=\frac{d\omega(\hat{x}(t))}{dt}(0)=\frac{\partial\omega}{dx}(x_0)\dot{\hat{x}}(0)=\frac{d\omega}{dx}(x_0)\varphi(\hat{x}(0),\hat{u}(0))=\frac{\partial\omega}{\partial x}(x_0)\varphi(x_0,u_0)$$

Пусть $v\in U$. Рассмотрим решение задачи Коши

$$\dot{x}=\varphi(x,v),\ x(0)=x_0$$

Тогда $\omega(x(t))\leq t-\hat{T}$. (Иначе мы бы могли составить допустимый процесс, фазовая переменная которого склеена из функций $x(\cdot)$ на отрезке $[0,t]$ и фазовой переменной $\tilde{x}$, оптимальной для задачи оптимального быстродействия с $\tilde{x}(0)=x(t)$. При этом мы получим допустимый процесс с $T<\hat{T}$.)

Следовательно,

$$1\geq\frac{d\omega(x(t))}{dt}(0)=\frac{\partial\omega}{dx}(x_0)\dot{x}(0)=\frac{d\omega}{dx}(x_0)\varphi(\hat{x}(0),v)=\frac{\partial\omega}{\partial x}(x_0)\varphi(x_0,v)$$

Тем самым мы проверили уравнение Беллмана.

Предположим теперь, что $\omega\in C^2(\mathbb{R}^n\setminus\mathcal{M})$. Фиксируем $x_0\in\mathbb{R}^n\setminus\mathcal{M}$. Пусть $(\hat{x}(\cdot),\hat{u}(\cdot),\hat{T})$—соответствующий оптимальный прцесс, $u_0=\hat{u}(0)$, $F(x)=\frac{\partial\omega}{\partial x}(x)\varphi(x,u_0)$. В силу уравнения Беллмана, $F(x_0)=1$ и $F(x)\leq 1$ для любого x. Таким образом, функция $F(x)$ достигает максимума в точке x_0. Значит, $\frac{\partial F}{\partial x}(x_0)=0$. Пусть x_j $(j=1,\ldots,n)$ есть j- ая координата вектора x. Для $j=1,\ldots,n$ мы имеем

$$0=\frac{\partial F}{\partial x_j}(x_0)=\sum_{i=1}^{n}\frac{\partial^2\omega}{\partial x_i x_j}(x_0)\varphi_i(x_0,u_0)+\sum_{i=1}^{n}\frac{\partial\omega}{\partial x_i}(x_0)\frac{\partial\varphi_i}{\partial x_j}(x_0,u_0),$$

что, используя дифференцирование по векторному аргументу, можно переписать как

$$0=\frac{\partial^2\omega}{\partial^2 x}(x_0)\varphi(x_0,u_0)+\frac{\partial\omega}{\partial x}(x_0)\frac{\partial\varphi}{\partial x}(x_0,u_0). \quad (6)$$

Обозначим $\psi(x)=\frac{\partial\omega}{\partial x}(x)$, $p(t)=\psi(\hat{x}(t))$. Заметим, что $x_0=\hat{x}(0)$, $\varphi(x_0,u_0)=\hat{\varphi}(0)=\dot{\hat{x}}(0)$, поэтому (6) может быть переписана в виде

$$0=\frac{\partial\psi}{\partial x}(\hat{x}(0))\frac{d\hat{x}}{dt}(0)+\psi(x_0)\hat{\varphi}_x(0)=\dot{p}(0)+p(0)\hat{\varphi}_x(0).$$

Таким образом, условие 1) в принципе максимума проверено для $t=0$. Уравнение Беллмана доет условие 3) оптимальности по u также для $t=0$.

Проверка выполнения условий 1) и 3) в произвольной точке $t=\tau$, управления $\hat{u}(\cdot)$ в которой непрерывно, сводится к случаю $t=0$. Действительно, пусть $(\hat{x}(\cdot),\hat{u}(\cdot),\hat{T})$— оптимальный управляемый процесс, $0<\tau<T$, $\tilde{x}_0=\hat{x}(\tau)$. Рассмотрим задачу оптимального быстродействия, в которой начальное условие заменено на $x(0)=\tilde{x}_0$.

Известно, что дуга оптимальной траектории, соединяющая произвольную ее точку с концом, является сама по себе оптимальной траекторией.

Поэтому оптимальным управляемым процессом для новой задачи будет тройка

$(\hat{\hat{x}}(\cdot), \hat{\hat{u}}(\cdot), \hat{\hat{T}})$, где $\hat{\hat{x}}(\cdot) = \hat{x}(\cdot + \tau)$, $\hat{\hat{u}}(\cdot) = \hat{u}(\cdot + \tau)$, $\hat{\hat{T}} = \hat{T} - \tau$. Соответствующая функция $\tilde{p}(\cdot)$ определяется формулой

$$\tilde{p}(t) = \psi(\hat{\hat{x}}(t)) = \psi(\hat{x}(t+\tau)) = p(t+\tau).$$

Следовательно,

$$\dot{p}(\tau) + p(\tau)\hat{\varphi}_x(\tau) = \dot{\tilde{p}}(0) + \tilde{p}(0)\hat{\tilde{\varphi}}_x(0) = 0$$

по доказанному, и уравнение 1) из принципа максимума в точке $t = \tau$ проверено; аналогично проверяется соотношение 3). Можно показать, что при подходящем выборе множителей Лагранжа λ условия 2), 4), 5) также будут выполнены.

Задача 9 *Привести пример задачи оптимального управления, в которой функция p не является непрерывно дифференцируемой.*

Указание. Рассмотрите задачу

$$\int_0^2 (-xu^2) \to min; x(0) = 0, x(2) = 1,$$

$$\dot{x} = u,\ 0 \leq u(t) \leq 1.$$

Задача 10 *Привести пример задачи оптимального управления*

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

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

такой, что при применении к ней принципа максимума Понтрягина необходимо брать $\lambda_0 = 0$.

Указание. Рассмотрите задачу

$$\int_0^1 (-\sqrt{u})dt \to min; x(0) = 0, x(1) = 0,$$

$$\dot{x} = u, u(t) \geq 0.$$

Задача 11 *Докажите, что для задачи оптимального управления*

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

$$\dot{x} = u, u(t) \in \mathbb{R}^n$$

при применении к ней принципа максимума Понтрягина мы имеем $\lambda_0 \neq 0$.

19 Оптимальный выбор существует. Доказано Филипповым.

До сих пор в рассматриваемых задачах не поднимался вопрос, существует ли вообще оптимальный процесс. В данном разделе покажем, что уже при достаточно малых ограничениях оптимальный процесс существует. Начнем с вспомогательных утверждений из функционального анализа.

Теорема 19.1 (отделимости) *Пусть X — ЛНП, $A \subset X, B \subset X$ — непустые выпуклые, $A \cap B = \oslash$, $intB \neq \oslash$. Тогда существует функционал $x^* \in X^*, \|x^*\| = 1:$*

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

$\square$ без доказательства $\blacksquare$

Следствие 19.1 *Пусть $A \subset X$ — непусто, выпукло, замкнуто, $x_0 \notin A$. Тогда $\exists x^* \in X^*,\ \|x^*\| = 1:$*

$$(sup_{x \in A} \langle x^*, x \rangle) < (\langle x^*, x_0 \rangle)$$

(т.е. имеет место строгая отделимость).

$\square$ Положим $B = \{x \in X:\ \|x - x_0\| \leq \varepsilon\}$, при достаточно малом ε $A \cap B = \oslash$. B — шар и, значит, выпуклое множество, имеет внутреннюю точку (например, x_0). Значит, по теореме отделимости $\exists\ x^* \in X^*,\ \|x^*\| = 1 : (sup_{x \in A} \langle x^*, x \rangle) \leq (inf_{x \in B} \langle x^*, x \rangle)$.

Т.к. $\|x^*\| = 1$, $\exists y \in X:\ \|y\| \leq 1,\ \langle x^*, y \rangle > \frac{1}{2}$. Рассмотрим $x_1 = x_0 - \varepsilon y \in B$. Имеем:

$$\langle x^*, x_1 \rangle = (\langle x^*, x_0 \rangle - \varepsilon \langle x^*, y \rangle) < (\langle x^*, x_0 \rangle - \frac{\varepsilon}{2}),$$

а $(inf_{x \in B} \langle x^*, x \rangle) \leq (\langle x^*, x_1 \rangle)$. Т.о. следствие доказано. $\blacksquare$

Следствие 19.2 (следствия 19.1) *Пусть $A \subset \mathbb{R}^n$ — непусто, выпукло, замкнуто, $t_0 < t_1$, $y(\cdot):\ [t_0, t_1] \to A$ — интегрируемая по Лебегу функция. Тогда*

$$\frac{1}{t_1 - t_0} \int_{t_0}^{t_1} y(t)dt \in A$$

$\square$ От противного. Пусть $\frac{1}{t_1 - t_0} \int_{t_0}^{t_1} y(t)dt = x_0 \notin A$. Тогда в силу следствия 19.1 $\exists x^* \in X^*,\ \|x^*\| = 1, \exists u \in \mathbb{R} : \forall x \in A\ (\langle x^*, x \rangle) \leq u,\ (\langle x^*, x_0 \rangle) > u$.

Имеем:

$$u < (\langle x^*, x_0 \rangle) = \frac{1}{t_1 - t_0} \langle x^*, \int_{t_0}^{t_1} y(t)dt \rangle = \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \langle x^*, y(t) \rangle dt \leq$$

$$\leq \{y(t) \in A\} \leq \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} udt = u$$

Противоречие. $\blacksquare$

Пусть далее $u : [t_0, t_1] \to \mathbb{R}^r$ — измерима, $x : [t_0, t_1] \to \mathbb{R}^n$ — абсолютно непрерывна.

Теорема 19.2 (Филиппова) *Рассмотрим задачу оптимального быстродействия*

$$T \to min, x(0) = x_0, x(T) = x_1$$

$$\dot{x}(t) = \varphi(x(t), u(t)) \text{ — в точках существования } \dot{x}(t), u(t) \in U$$

Пусть выполнены условия

1) $\varphi, \dot{\varphi}_x(\cdot)$ *— непрерывны;*

2) $\exists c > 0\ \forall x,\ \forall u \in U\ (\langle x, \varphi(x,u)\rangle) \le c(\|x\|^2 + 1)$*;*

3) $\forall x\ \varphi(x, U) := \bigcup\limits_{u \in U} \{\varphi(x,u)\}$ *— выпукло,* U *— замкнуто, ограничено;*

4) $\exists(\tilde{x}(\cdot), \tilde{u}(\cdot), \tilde{T})$ *— допустимый процесс.*

Тогда существует оптимальный процесс.

Лемма 19.1 *Пусть* $x(\cdot) : [0, T] \to \mathbb{R}^n$*— абсолютно непрерывная функция, и* $(\langle x(t), \dot{x}(t)\rangle) \le c(\|x\|^2 + 1)$*. Тогда* $\forall\ t \in [0, T]\ \|x(t)\| \le e^{ct}(\|x(0)\| + 1)$*.*

□ Рассмотрим функцию

$$y(t) = \|x(t)\|^2 + 1 = \langle x(t), x(t)\rangle + 1;$$

$$\dot{y}(t) = 2\,\langle x(t), \dot{x}(t)\rangle + 1;$$

Следовательно можно записать, что $\dot{y}(t) \le 2cy(t)$. Теперь рассмотрим функцию

$$z(t) = e^{-2ct} y(t);$$

$$\dot{z}(t) = e^{-2ct}\dot{y}(t) - 2ce^{-2ct}y(t) \le 0;$$

Заметим, что $z(t)$— абсолютно непрерывная функция, ее производная меньше или равна нуля, поэтому можно сделать вывод о том, что это невозрастающая функция. Следовательно $\forall\ t \in [0, T]\ z(t) \le z(0)\ \Rightarrow\ y(t) \le e^{2ct}y(0)\ \Rightarrow$

$$\|x(t)\|^2 + 1 \le e^{2ct}(\|x(0)\|^2 + 1)$$

Поэтому

$$\|x(t)\| \le e^{ct}(\|x(0)\|^2 + 1)^{\frac{1}{2}}.$$

■

Можно заметить, что если $\exists\ \dot{x}(t)$, тогда можно положить, что $\dot{x}(t) = \varphi(x(t), u(t))$ $(0 \le t \le \tau)$. Рассмотрим произвольное $\tilde{u} \in U$, и пусть $\dot{x}(t) = \varphi(x(t), \tilde{u})$ при $t \ge \tau$. Тогда здесь будут выполнены все условия теоремы существования и единственности решения данного обыкновенного дифференциального уравнения , т.к имеет место пункт 1) теоремы Филлипова. Теорема существования и единственности — локальная теорема, т.е предполагается, что решение существует и единственно в некоторой окрестности точки, но в условиях теоремы Филлипова в силу леммы 10.1 $\|x(t)\|$ не может возрастать неограниченно при ограниченном t, поэтому решение продолжается неограниченно.

Теперь докажем теорему Филлипова:

□ Рассмотрим все допустимые процессы $(x(\cdot), u(\cdot), T)$ такие, что $T \le \tilde{T}$. Это множество не пусто, так как имеет место пункт 4) условия теоремы.

Пусть $\hat{T} = \inf T$. Далее $\forall\ j \geq 1$ существует процесс $(\tilde{x}_j(t), \tilde{u}_j(t), \tilde{T})$ такой, что $\tilde{T}_j \leq \hat{T} + \dfrac{1}{j}$ и $\tilde{T}_j \leq \tilde{T}$. Продолжим $\tilde{x}_j(\cdot)$, $\tilde{u}_j(\cdot)$ на $[0, \tilde{T}]$. Заметим, что

$$\forall\ t\ \|\tilde{x}_j(t)\| \leq e^{c\tilde{t}}(\|x(0)\| + 1);$$

$$\|\dot{\tilde{x}}_j(t)\| = \|\varphi(\tilde{x}_j(t), \tilde{u}_j(t))\| \leq M;$$

Т.е имеет место равномерная ограниченность нормы. Заметим, что так как абсолютно непрерывная функция является интегралом от своей производной, то имеет место неравенство:

$$\|\tilde{x}_j(t_1) - \tilde{x}_j(t_2)\| \leq M|t_1 - t_2|,\ \ t_1, t_2 \in [0, \tilde{T}];$$

поэтому по теореме Арцела существует подпоследовательность

$$\tilde{x}_{j_m}(\cdot) \rightrightarrows \hat{x}(\cdot)$$

Введем обозначения:

$$x_m := \tilde{x}_{j_m};$$

$$u_m := \tilde{u}_{j_m};$$

$$T_m := \tilde{T}_{j_m};$$

Тогда в новых обозначениях $x_m(\cdot) \rightrightarrows \hat{x}(\cdot)$ на $[0, \tilde{T}]$, $T_m \to \hat{T}$, $\hat{x}(0) = \lim\limits_{m\to\infty} x_m(0) = x_0$. Далее

$$\|x_m(\hat{T}) - x_1\| = \|x_m(\hat{T}) - x_m(T_m)\| \leq M|\hat{T} - T_m| \to 0\ (m \to \infty);$$

Следовательно, в силу равномерной сходимости,

$$\hat{x}(\hat{T}) = \lim_{m\to\infty} x_m(\hat{T}) = x_1;$$

Пусть $t_0 \in [0, \hat{T}]$ и существует $\dot{\hat{x}}(t_0)$. Рассмотрим множество

$$V = \varphi(\hat{x}(t_0), U) = \{\varphi(\hat{x}(t_0), u):\ \ u \in U\};$$

по условию теоремы это множество является замкнутым, ограниченным, выпуклым. Следовательно V— выпуклое компактное множество. Рассмотрим $\varepsilon > 0$, и пусть V_ε— ε окрестность множества V, т.е

$$V_\varepsilon = \{y \in \mathbb{R}^n:\ \exists\ z \in V, \|y - z\| \leq \varepsilon\}$$

По этому $\varepsilon > 0$ подберем такое δ, что $0 < \delta$, и

$$\|\tilde{x}_1\| \leq e^{c\tilde{t}}(\|x(0)\| + 1);$$

$$\|\tilde{x}_2\| \leq e^{c\tilde{t}}(\|x(0)\| + 1);$$

$$\|\tilde{x}_1 - \tilde{x}_2\| \leq \delta$$

влекут $\forall\ u \in U$

$$\|\varphi(\tilde{x}_1, u) - \varphi(\tilde{x}_2, u)\| \leq \varepsilon$$

В силу равномерной непрерывности функции φ на компакте, такое δ существует, и его можно выбрать таким, что

$$\|x_m(\cdot) - \hat{x}(\cdot)\| \le \frac{\delta}{2};$$

Введем обозначение $\eta = \dfrac{\delta}{2M}$, тогда если $|t - t_0| \le \eta$, то

$$\|x_m(t) - x_m(t_0)\| \le \frac{\delta}{2}$$

$|t - t_0| \le \eta \;\Rightarrow \|x_m(t) - \hat{x}(t_0)\| \le \delta$ Рассмотрим

$$\|\varphi(x_m(t), u_m(t)) - \varphi(\hat{x}(t_0), u_m(t_0))\| \le \varepsilon$$

$$\varphi(\hat{x}(t_0), u_m(t_0)) \in V$$

Следовательно, из последних двух выражений вытекает, что

$$\varphi(x_m(t), u_m(t)) \in V_\varepsilon$$

Запишем соотношение:

$$\frac{x_m(t) - x_m(t_0)}{t - t_0} = \frac{1}{t - t_0} \int\limits_{t_0}^{t} \dot{x}_m(s) ds$$

Теперь воспользуемся следствием 19.2 и получим, что

$$\frac{x_m(t) - x_m(t_0)}{t - t_0} \in V$$

Следовательно

$$\frac{\hat{x}(t) - \hat{x}(t_0)}{t - t_0} \in V_\varepsilon \quad \text{при } |t - t_0| \le \eta;$$

$$\dot{\hat{x}}(t_0) = \lim_{t \to t_0} \frac{\hat{x}(t) - \hat{x}(t_0)}{t - t_0} \in V_\varepsilon;$$

Так как последнее верно для любого $\varepsilon > 0$, то из этого следует, что $\dot{\hat{x}}(t_0) \in V$, другими словами

$$\dot{\hat{x}}(t_0) = \varphi(\hat{x}(t_0), u) \qquad u \in U;$$

$\hat{u}(t_0) := u$— такое управление, что во всех точках недифференцируемости u имеет место условие дифференциальной связи. Во всех этих точках такую функцию можно выбрать измеримой. Этот факт обусловлен леммой об измеримости, которую мы доказывать не будем. Поэтому по модулю этой леммы теорема Филлипова доказана. ■

Задача 12 *Доказать, что условие выпуклости в теореме Филлипова существенно.*

Указание: рассмотреть задачу

$$T \to \min;$$

$$x(0) = -1; \; x(T) = 0;$$

$$y(0) = y(T) = 0;$$

$$\dot{x} = -y^2 + u^2; \; \dot{y} = u; \; |u| \le 1.$$

20 Теорема Куна—Таккера—Каруша.

Теорема 20.1 (Куна—Таккера—Каруша) *Пусть X—ЛНП, $A \subset X$— выпукло, $f_0, \ldots, f_m : A \to \mathbb{R}$— выпуклые функции;*

$$f_0(x) \to \min;$$

$$f_i(x) \le 0 \ (i = 1, \ldots, m) \qquad (з)$$

Пусть $\hat{x} \in A$— решение (з). Тогда

1) существует $\bar{\lambda} = (\lambda_0, \ldots, \lambda_m) \ne \bar{0}$;

а) $\min\limits_{x \in A} \mathcal{L}(x, \bar{\lambda}) = \mathcal{L}(\hat{x}, \bar{\lambda})$, где

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

б) $\lambda_i \ge 0 \ (i = 0, \ldots, m)$;

в) $\lambda_i f_i(\hat{x}) = 0 \ (i = 1, \ldots, m)$;

2) если $\lambda_0 > 0$, то условия а)—в)—достаточны;

3) если $\exists \ \bar{x} \in A, \ f_i(\bar{x}) < 0 \ (i = 1, \ldots, m)$ (это условие Слейтера), то $\lambda_0 > 0$.

$\square$ Без ограничения общности считаем, что $f_0(\hat{x}) = 0$. Тогда в общем случае положим:

$$\tilde{f}_0(x) = f_0(x) - f_0(\hat{x})$$

и рассмотрим задачу

$$\tilde{f}_0 \to \min;$$

$$f_i \le 0 \ (i = 1, \ldots, m).$$

Рассмотрим пространство

$$Y = \{(y_0, \ldots, y_m)\} = \mathbb{R}^{m+1}$$

$$B = \{(y_0, \ldots, y_m) < 0 : \ \forall \ i \ y_i < 0\}$$

Заметим, что B— выпукло, $\mathrm{int}B \ne \varnothing$ Пусть

$$C = \{(y_0, \ldots, y_m) : \ \exists \ x \in A \ f_0(x) \le y_0; \ldots; f_m(x) \le y_m\}$$

Исследуем на выпуклость множество C. Рассмотрим $y, \ y' \in C$, $\alpha \ge 0, \ \beta \ge 0 : \ \alpha + \beta = 1$ Тогда условие выпуклости множества C равносильно условию

$$\alpha y + \beta y' \in C \qquad (*)$$

Проверим условие (*): имеем, существуют $x, \ x' : \ \forall \ i \ f_i(x) \le y_i, \ f_i(x') \le y_i'$ Тогда

$$f_i(\alpha x + \beta x') \le \alpha f_i(x) + \beta f_i(x') \le \alpha y_i + \beta y' \ \Rightarrow (*)$$

Итак, мы показали, что C— выпукло. Эаметим, что $B \cap C = \varnothing$, так как в противном случае, если $y \in B \cap C$, то существует $x : \ f_i(x) \le y_i < 0$, и мы получаем противоречие с экстремальностью функции $\hat{x}(\cdot)$.

Пусть существует $\bar{\lambda} \in Y \setminus \{0\} : \ \forall\, y \in B, \ \forall\, z \in C$ имеет место неравенство:

$$(\langle \bar{\lambda}, y \rangle) \leq (\langle \bar{\lambda}, z \rangle)$$

или, другими словами,

$$\forall\, y \in B, \ \forall\, z \in C \ \sum_{i=0}^{m} \lambda_i y_i \leq \sum_{i=0}^{m} \lambda_i z_i$$

где λ_i—те самые множители Лагранжа, которые мы ищем.

Пусть $z_i = f_i(x)$ $(x \in A)$, тогда $\forall\, y \in B$

$$\sum_{i=0}^{m} \lambda_i y_i \leq \sum_{i=0}^{m} \lambda_i f_i(x)$$

Зафиксируем $i_0, \ x$, тогда

$$\lim_{y_{i_0} \to \infty} \lambda_{i_0} y_{i_0} \leq \sum_{i=0}^{m} \lambda_i f_i(x) - \sum_{i \neq i_0, 0 \leq i \leq m} \lambda_i y_i$$

Если $\lambda_{i_0} < 0$, то $\lim = \infty$— противоречие. Поэтому считаем, что $\forall\, i \ \lambda_i \geq 0$

$$0 = \lim_{y \to 0, \ y \in B} \sum_{i=0}^{m} \lambda_i y_i \leq \sum_{i=0}^{m} \lambda_i f_i(x)$$

Следовательно $\sum\limits_{i=0}^{m} \lambda_i f_i(x) \geq 0$ Возьмем $x = \hat{x}$. Рассмотрим $\sum_{i=0}^{m} \lambda_i f_i(\hat{x})$. $\lambda_i \geq 0$, $f_i \leq 0$ $\Rightarrow \sum \leq 0$. Но, как мы знаем, $\sum \geq 0$. Значит, $\sum = 0$ и $\forall i \neq 0 \ \lambda_i f_i(\hat{x}) = 0$ (условие дополняющей нежесткости).

Итак, $\sum_{i=0}^{m} \lambda_i f_i(x)$ достигает минимума в точке $\hat{x}$ и мы доказали I часть теоремы.

II часть: если $\lambda_0 > 0$, условия достаточны.

Пусть $x \in A$ — допустимая точка. Тогда

$$\lambda_0 f_0(x) \geq \sum_{i=0}^{m} \lambda_i f_i(x) \geq \sum_{i=0}^{m} \lambda_i f_i(\hat{x}) = \lambda_0 f_0(\hat{x})$$

(из условия неотрицательности λ_i, достижения минимума функции Лагранжа в точке $\hat{x}$ и условия дополняющей нежесткости).

Т.о. $f_0(x) \geq f_0(\hat{x})$, и мы доказали достаточность.

III часть: $\exists \overline{x} \in A : \ \forall i = 1, \ldots, m \ f_i(\overline{x}) < 0 \ \Rightarrow \lambda_0 > 0$

Будем доказывать от противного. Пусть $\lambda_0 = 0$. Тогда по доказанному $\sum_{i=1}^{m} \lambda_i f_i(\overline{x}) \geq 0$. При этом $\lambda_i \geq 0$, $f_i(\overline{x}) < 0$. Значит, $\lambda_i f_i(\overline{x}) \leq 0$. Но среди λ_i есть хотя бы одно $\neq 0$, откуда следует, что $\lambda_i f_i(\overline{x}) < 0$. Противоречие. ■

21 Доказательство принципа Лагранжа для задачи с ограничениями типа равенств и неравенств в частном случае.

Следующее утверждение было нами получено в качестве следствия из принципа Лагранжа, который в общем случае мы не доказали. Здесь мы приведем независимое доказательство этого следствия.

Следствие 21.1 (следствие 14.1 из принципа Лагранжа) *Пусть* $f_i \in C(U \cap K, \mathbb{R})$ $(i = 0, \ldots, m)$. *Рассмотрим задачу* $(*)$:

$$f_0(x) \to min$$
$$f_i(x) = 0 \ (i = 1, \ldots, m), \ x \in U \cap K$$

Пусть $\hat{x}$ *— locmin в* $(*)$ *и существует односторонняя производная* $f_i'(\hat{x})$.

Тогда $\exists \overline{\lambda} = (\lambda_0, \ldots, \lambda_m) \neq 0$:

1) для $\mathcal{L}(x, \overline{\lambda}) = \sum_{i=0}^{m} \lambda_i f_i(x)$ *выполнено* $\mathcal{L}_x(\hat{x}, \overline{\lambda}) \geq 0$ $(\Leftrightarrow \sum_{i=0}^{m} \lambda_i f_i'(\hat{x}) \geq 0$, *это покомпонентное неравенство*)

2) $\lambda_0 \geq 0$

☐ Без ограничения общности положим $f_0(0) = 0$. Рассмотрим линейный оператор: $\Lambda h = (f_1'(0)[h], \ldots, f_m'(0)[h])$, $\Lambda : \mathbb{R}^d = X \to \mathbb{R}^m$.

I случай: $\Lambda X = Y \neq \mathbb{R}^m$, тогда существует вектор $(\lambda_1, \ldots, \lambda_m) \neq 0$, ортогональный пространству Y. Значит, $\forall h \in X$ $\sum_{i=1}^{m} \lambda_i f_i'(0)[h] = 0$. Положив $\lambda_0 = 0$, получаем отсюда, что $\forall h \in X$ $\sum_{i=0}^{m} \lambda_i f_i'(0)[h] = 0$. Т.о. мы нашли λ_i : производная $\mathcal{L}_x = 0$ и тем самым теорема в этом случае доказана.

II случай: $\Lambda X = \mathbb{R}^m$. Рассмотрим $h = (h_1, \ldots, h_d) \in \mathbb{R}^d$, $h_0 = -f_0'(0)[h]$. Определим

$$A_k = \{h : \ h_i > 0 \ (i = k, \ldots, d), \ h \in Ker\Lambda\} \quad (k = 0, \ldots, d+1)$$

В частности, $A_{d+1} = Ker\Lambda$. Отсюда $0 \in A_{d+1} \neq \oslash$. Очевидно, что $A_0 \subset A_1 \subset \ldots \subset A_{d+1}$

Лемма 21.1 (основная) $A_0 = \oslash$

☐ Рассмотрим $h^j \in X$: $f_i'(0)[h^j] = \delta_{ij}$, т.е. $\Lambda[h^j] = (0, \ldots, 0, 1, 0, \ldots, 0)$, 1 стоит на j-й позиции. Доказывать будем от противного. Предположим, существует $\tilde{h} \in A_0$: $\tilde{h}_0 > 0, \ \tilde{h}_i > 0 \ (i = 1, \ldots, d), \ \tilde{h} \in Ker\Lambda$. Несложно заметить, что, умножив на любое положительное число, он все равно останется в A_0. Поэтому $\tilde{h}$ можно выбрать таким образом, что

$$\tilde{h}_0 - \sum_{j=1}^{m} |h_0^j| > 0 \ \text{ и } \ \tilde{h}_i - \sum_{j=1}^{m} |h_i^j| > 0, \quad i = 1, \ldots, d$$

Пусть $r > 0$ — мало. Определим отображение

$$G(y) = (f_1(r\tilde{h} + \sum_{j=1}^{m} y_j h^j), \ldots, f_m(r\tilde{h} + \sum_{j=1}^{m} y_j h^j)),\ \|y\| \leq r$$

Можно показать, что $G(y) = y + o(r),\ r \to 0$. Отсюда $\|G(y) - y\| \leq r$ для достаточно малого r (пользуемся следствием об ε-сдвиге). $\exists y :\ \|y\| \leq r, G(y) = 0$.

Пусть $h = r\tilde{h} + \sum_{j=1}^{m} y_j h^j$. Тогда $f_i(h) = 0\ (i = 1, \ldots, m)$ и

$$f_0(h) = \underbrace{(-\tilde{h}_0 - \sum_{j=1}^{m} y_j h_0^j)}_{\text{производная в } 0} r + o(r) < 0$$

для достаточно малого r. Пришли к противоречию с тем, что минимум достигается в 0. Т.о. лемма доказана. ■

Итак, $\oslash = A_0 \subset A_1 \subset \ldots \subset A_{d+1} \neq \oslash$. Значит, $\exists k :\ A_k = \oslash,\ A_{k+1} \neq \oslash$. Это самое k и рассмотрим.

Лемма 21.2

$$h = 0 \text{ — \textit{решение задачи}}$$
$$-h_k \to \min;$$
$$-h_{k+1} \leq 0, \ldots, -h_d \leq 0, \quad h \in Ker\Lambda$$

□ Предположим, что утверждение неверно.

$\exists h :\ h_k > 0\ (\neq 0),\ h_i \geq 0\ (i = k+1, \ldots, d), h \in Ker\Lambda$

$A_{k+1} \neq \oslash \Rightarrow \exists \tilde{h} :\ \ \tilde{h}_i > 0\ (i = k+1, \ldots, d),\ \tilde{h} \in Ker\Lambda$

Рассмотрим вектор $h + \alpha\tilde{h},\ \alpha > 0$ — мало. $h_k + \alpha\tilde{h}_k > 0$, $h_i + \alpha\tilde{h}_i > 0$ при $i > k,\ h + \alpha\tilde{h} \in Ker\Lambda$. Значит, $h + \alpha\tilde{h} \in A_k$, и мы получили противоречие с тем, что $A_k = \oslash$ ■

Перейдем, наконец, к доказательству теоремы. Задача про h — выпукла, выполнены условия Слейтера $\Rightarrow$ можно применить теорему Куна—Таккера—Каруша.

$$\exists(\mu_k, \ldots, \mu_d) \geq 0, \mu_k = 1 :\quad \min_{h \in Ker\Lambda}(-\sum_{i=k}^{d} \mu_i h_i) = 0$$

Мы имеем $\sum_{i=k}^{d} \mu_i h_i \leq 0$. Далее, если $h \in \Lambda$, то и $-h \in \Lambda$, $\sum_{i=k}^{d} \mu_i(-h_i) \leq 0$, т.е. $\sum_{i=k}^{d} \mu_i h_i \geq 0$. Поэтому $\forall h \in Ker\Lambda\ \ \sum_{i=k}^{d} \mu_i h_i = 0$. Положим $\mu_i = 0,\ i < k$. Тогда

$$\Lambda h^1 = \Lambda h^2 \Rightarrow \sum_{i=0}^{d} \mu_i h_i^1 = \sum_{i=0}^{d} \mu_i h_i^2$$

Значит, существует оператор $T \in (\mathbb{R}^m)^*$:

$$\sum_{i=0}^{d} \mu_i h_i = T(\Lambda h)$$

(значение оператора Λh однозначно определяет такую сумму). Заметим, что такое отображение линейно.

$$\Lambda h = (f_1'(0)[h], \dots, f_m'(0)[h])$$

$$\exists \lambda_1, \dots, \lambda_m : \sum_{i=0}^{d} \mu_i h_i = \sum_{i=1}^{m} \lambda_i f_i'(0)[h]$$

Пусть $\lambda_0 = \mu_0 \geq 0$

$$\sum_{i=1}^{m} \lambda_i f_i'(0)[h] - \mu_0 h_0 = \sum_{i=1}^{d} \mu_i h_i \ \forall h$$

При этом $\mu_0 = \lambda_0, h_0 = -f_0'(0)[h]$. Значит,

$$\sum_{i=0}^{m} \lambda_i f_i'(0)[h] = \sum_{i=1}^{d} \mu_i h_i \ \forall h$$

Нам нужно было проверить, что левая часть равенства выше неотрицательна. Но правая часть неотрицательна при $h \geq 0$ и тем самым теорема доказана. ■

-- ЗЕ ЕНД --