ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ

Глава 1. Введение

1.1. Общая постановка

В общем случае в задачах вариационного исчисления задается класс D кривых или гиперповерхностей, описываемых функциями $x : G \subseteq R^n \to R^m$. Как правило, предполагается достаточная гладкость $x(s), s \in G$, а также накладываются некоторые дополнительные требования, в частности, на границе G. Далее на элементах $x(s)$ класса D определяется интегральный функционал

$$J[x] = \int_G F(s, x(s), ...)ds \tag{1}$$

подынтегральная функция которого может зависеть не только от s и $x(s)$, но и от производных $x(s)$ до некоторого порядка. Задача заключается в определении $x^* \in D$ для которого

$$J[x^*] = \min_{x \in D} J[x] \tag{2}$$

если глобальный минимум существует, или в определении локального минимума, тип которого зависит от вводимого понятия близости элементов в D. Кроме задач на минимум могут рассматриваться задачи на максимум.

1.2. Примеры задач вариационного исчисления

Задача с фиксированными концами

Пусть $G = [t_0, t_1]$, а множество допустимых кривых D имеет вид:

$$D = D_1 = \{x : x \in C^2(G), x : G \to R^1, x(t_0) = x_0, x(t_1) = x_1\} \tag{3}$$

$$J[x] = \int_{t_0}^{t_1} F(t, x(t), \dot{x}(t))dt \tag{4}$$

где функция F предполагается непрерывной вместе со своими частными производными до второго порядка по совокупности всех переменных.

Задача со свободными концами

Пусть $G = [t_0, t_1]$, а множество допустимых кривых D имеет вид:

$$D = D_2 = \{x : x \in C^2(G), x : G \to R^1\} \tag{5}$$

функционал $J[x]$ имеет вид (4) с теми же требованиями к функции F, что и в предыдущей задаче. Как видно из (5) ограничений на значения кривой на концах промежутка интегрирования, в отличие от задачи с фиксированными концами, не накладывается.

Задача со скользящими концами

Пусть $G = (-\infty, \infty)$ и заданы граничные кривые, определяемые соотношениями

$$\Psi_i(x, t) = 0, (i = 0, 1) \tag{6}$$

функции Ψ_i считаются достаточно гладкими по совокупности переменных.

Множество допустимых кривых D имеет вид:

$$\begin{aligned} D = D_3 = \{x : x \in C^2(G), x : G \to R^1; \exists t_i^x : \Psi_i(x(t_i^x), t_i^x) = 0, \\ (i = 0, 1); \dot{x}(t_i^x) \neq -(\Psi'_{it}(x(t_i^x), t_i^x)/\Psi'_{ix}(x(t_i^x), t_i^x))\} \end{aligned} \tag{7}$$

$$J[x] = \int_{t_0^x}^{t_1^x} F(t, x(t), \dot{x}(t))dt \tag{8}$$

где на функцию F накладываются прежние требования и $t_0^x < t_1^x$ – значения t, соответствующие пересечениям без касания кривой $x = x(t)$ с граничными кривыми.

Заметим, что в (8) пределы интегрирования зависят от кривой $x = x(t)$ и определяются ближайшими друг к другу точками её пересечения с граничными кривыми (6). Последнее условие в виде запрета равенства для производной $\dot{x}(t_i^x)$ в определении D_3 означает отсутствие касания кривой $x = x(t)$ с граничными кривыми в точках пересечения. Это требование необходимо для того, чтобы точка пересечения не исчезала при достаточно малом изменении кривой $x = x(t)$.

Глава 2. Слабый и сильный локальный минимумы

Определение 2.1. *Локальным минимумом функционала J из (4) на множестве кривых D из (3) или (5) пространства B назовем такую кривую $x^o \in D$ для которой $\exists \epsilon > 0$, что $\forall x \in D \cap O_\epsilon(x^o)$ выполняется*

$$J[x^o] \le J[x] \tag{9}$$

где

$$O_\epsilon(x^o) = \{x \in B : ||x^o - x|| \le \epsilon\}. \tag{10}$$

В зависимости от способа задания нормы можно получить различные трактовки понятия «локальный минимум». Введем нормы двух типов.

Определение 2.2. *Норму вида*

$$||x||_0 = \max_{t \in [t_0, t_1]} |x(t)| \tag{11}$$

определенную на линейном пространстве непрерывных на $[t0, t1]$ функций ($x \in C[t0, t1]$) назовем нормой нулевого порядка.

Определение 2.3. *Норму вида*

$$||x||_1 = \max_{t \in [t_0, t_1]} |x(t)| + \max_{t \in [t_0, t_1]} |\dot{x}(t)| \tag{12}$$

определенную на линейном пространстве непрерывно дифференцируемых на $[t0, t1]$ функций ($x \in C^1[t0, t1]$) назовем нормой первого порядка.

Аналогично можно ввести понятие нормы k-го порядка.

Определение 2.4. *Локальный минимум, понимаемый в смысле нормы нулевого порядка из (11) называется сильным локальным минимумом функционала, а понимаемый в смысле нормы первого порядка из (12) — слабым локальным минимумом.*

Заметим, что всякий сильный локальный минимум одновременно является и слабым, однако обратное неверно.

Глава 3. Метод вариаций Лагранжа

Пусть выделена некоторая допустимая кривая $\hat{x} \in D$. Определим для нее класс пробных функций $M(\hat{x})$:

$$\begin{gathered} M(\hat{x}) = \{\eta : \text{где } \eta : G \to R^1 \text{ и } \forall \alpha \in R^1, \\ \text{если } |\alpha| \text{ – достаточно мало, то } \hat{x} + \alpha \cdot \eta \in D\} \end{gathered} \tag{13}$$

Таким образом, при $\eta \in M(\hat{x})$ и достаточно малом по модулю α изменение кривой $\hat{x}(t)$ с помощью добавки $\alpha \cdot \eta(t)$ не выводит измененную кривую $\hat{x} + \alpha \cdot \eta$ из допустимого множества D.

Изменение, добавляемое к функции $\hat{x}$, называют вариацией кривой $\delta\hat{x}$. В нашем случае $\delta\hat{x} = \alpha \cdot \eta$. Кривую $x = \hat{x} + \delta\hat{x}$ называют проварьированной кривой.

При использовании параметрических вариаций вида $\delta\hat{x} = \alpha \cdot \eta$ значение функционала J на проварьированной кривой может быть рассмотрено как функция, зависящая от параметра вариации α:

$$Q(\alpha) = J[x + \alpha\eta] \tag{14}$$

Сделанные ранее предположения о задаче позволяют представить приращение функционала при варьировании кривой в виде разложения:

$$J[x + \alpha\eta] - J[x] = Q'(0)\alpha + Q''(0)\frac{\alpha^2}{2} + o(\alpha^2) \tag{15}$$

Определение 3.1. *Первой вариацией функционала на кривой x в направлении пробной функции $\eta \in M(x)$ назовем главную линейную часть приращения функционала:*

$$\delta_\eta J(x, \alpha) = Q'(0) \cdot \alpha \tag{16}$$

Второй вариацией функционала назовем главную квадратичную часть его приращения:

$$\delta^2_{\eta\eta} J(x, \alpha) = Q''(0) \cdot \frac{\alpha^2}{2} \tag{17}$$

Теорема 1. *Если x^o — локальный минимум функционала (слабый или сильный), то верно следующее:*

1. $\forall \eta \in M(x^o) : \delta_\eta J(x^o, \alpha) = 0$ *– необходимое условие первого порядка;*
2. $\forall \eta \in M(x^o) : \delta^2_{\eta\eta} J(x^o, \alpha) = 0$ *– необходимое условие второго порядка.*

Доказательство 1. *Пусть это не так и первое утверждение неверно. Тогда найдется пробная функция $\hat{\eta} \in M(x^o)$, что $Q'(0) \neq 0$. Для определенности можно считать, что $Q'(0) < 0$. Тогда из (15) следует, что при любых достаточно малых значениях $\alpha > 0$ выполнится*

$$J[x^o + \alpha\hat{\eta}] < J[x^o] \tag{18}$$

Однако

$$||x^o + \alpha\hat{\eta} - x^o||_1 = \alpha||\eta||_1 \to 0 \tag{19}$$

при $\alpha \to 0$, что означает принадлежность кривой $x^o + \alpha\hat{\eta}$ любой слабой окрестности x^o при достаточно малом α. Это противоречит слабой оптимальности кривой x^o. Таким образом, первое утверждение доказано для слабого локального минимума. Однако, поскольку сильный локальный минимум одновременно является слабым, то утверждение 1 справедливо и для сильного локального минимума.

Аналогично доказывается второе утверждение.

Определение 3.2. *Экстремалями функционала J (стационарными кривыми) называют такие допустимые кривые $\bar{x} \in D$ для которых*

$$\forall \eta \in M(\bar{x}) : \delta_\eta J(\bar{x}, \alpha) = 0 \tag{20}$$

Теорема 2. ***Лемма Лагранжа.*** *Пусть $f(t) \in C[t_0, t_1]$, т.е. f – непрерывна на отрезке, и*

$$M = \{\eta : \eta \in C^k[t_0, t_1], (k \geq 2), \eta(t_0) = \eta(t_1) = 0\} \tag{21}$$

Если при этом $\forall \eta \in M$:

$$\int_{t_0}^{t_1} f(t)\eta(t)dt = 0 \tag{22}$$

то $f(t) \equiv 0$ для $t \in [t_0, t_1]$.

Доказательство 2. *Пусть утверждение неверно, т.е. $\exists \bar{t} \in [t_0, t_1]$, что $f(\bar{t}) = c \neq 0$ (не ограничивая общности будем считать, что $c > 0$). При этом (из непрерывности f) следует, что такое $\bar{t}$ найдется и в (t_0, t_1). Тогда существует δ–окрестность $\bar{t}$ в которой $f(t) > c/2$.*

Построим функцию $\bar{\eta}(t)$, гладкую до порядка $k > 2$, чтобы

$$\bar{\eta}(t) = \begin{cases} 1, t \in O_{\delta/2}(\bar{t}), \\ 0, t \notin O_{\delta}(\bar{t}), \end{cases} \tag{23}$$

и при этом была обеспечена неотрицательность $\bar{\eta}(t)$ для всех значений аргумента. Построенная функция $\bar{\eta}(t) \in M$.

Справедливы оценки

$$\int_{t_0}^{t_1} f(t)\bar{\eta}(t)dt = \int_{O_\delta(\bar{t})} f(t)\bar{\eta}(t)dt \geq c/2 \int_{O_{\delta/2}(\bar{t})} \bar{\eta}(t)dt = \frac{c}{2}\delta > 0 \tag{24}$$

Возникшее противоречие доказывает, что лемма верна.

Глава 4. Необходимые условия оптимальности

4.1. Условия оптмиальности первого порядка

Теорема 3. *(условия оптимальности первого порядка для задачи с закрепленными концами). Для того, чтобы в задаче с закрепленными концами кривая $x = x^*(t) \in C^2[t_0, t_1]$ являлась экстремалью функционала (4) необходимо и достаточно, а для того, чтобы она являлась экстремумом (минимумом или максимумом) необходимо, чтобы функция $x = x^*(t)$:*

1. *являлась решением дифференциального уравнения Эйлера:*

$$\frac{\delta F}{\delta x}(t, x, \dot{x}) - \frac{d}{dt}\Big(\frac{\delta F}{\delta \dot{x}}(t, x, \dot{x})\Big) = 0 \tag{25}$$

2. *удовлетворяла граничным условиям:*

$$x^*(t_i) = x_i, (i = 0, 1) \tag{26}$$

Теорема 4. *(условия оптимальности первого порядка для задачи со свободными концами). Для того, чтобы в задаче с закрепленными концами кривая $x = x^*(t) \in C^2[t_0, t_1]$ являлась экстремалью функционала (4) необходимо и достаточно, а для того, чтобы она являлась экстремумом (минимумом или максимумом) только необходимо, чтобы функция $x = x^*(t)$:*

1. *являлась решением дифференциального уравнения Эйлера (25)*

2. *удовлетворяла так называемым естественным граничным условиям*

$$\frac{\delta F}{\delta \dot{x}}(t_i, x(t_i), \dot{x}(t_i)) = 0, (i = 0, 1) \tag{27}$$

Теорема 5. *(условия оптимальности первого порядка для задачи со скользящими концами). Для того, чтобы в задаче со свободными концами кривая $x = x^*(t) \in C^2[t_0, t_1]$, имеющая неособые пересечения с граничными кривыми, являлась экстремалью необходимо и достаточно, а для того, чтобы она являлась экстремумом (минимумом или максимумом) только необходимо, чтобы функция $x = x^*(t)$:*

1. *являлась решением дифференциального уравнения Эйлера (25)*

2. *выполнялись так называемые граничные условия трансверсальности*

$$\frac{\delta F}{\delta \dot{x}}(t_i, x_i, \dot{x}_i) - \frac{\Psi'_{ix}(x_i, t_i) F(t_i, x_i, \dot{x}_i)}{.\Psi'_{it}(x_i, t_i) + \dot{x}_i \Psi'_{ix}(x_i, t_i)} = 0, (i = 0, 1) \qquad (28)$$