КАЗАНСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ

ИНСТИТУТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ
И ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Кафедра прикладной математики

Н.Б. ПЛЕЩИНСКИЙ

ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ И ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

Учебное пособие

Казань – 2022

УДК 517.97
ББК 22.162

Публикуется по решению
учебно-методической комиссии ИВМиИТ
протокол №5 от 24 февраля 2022 года

Рецензенты:
доктор педагогических наук,
профессор кафедры информатики и прикладной
математики КНИТУ **Н.К. Нуриев**,
кандидат физико-математических наук,
доцент кафедры прикладной математики КФУ **Л.У. Бахтиева**

Плещинский Н.Б.
Вариационное исчисление и оптимальное управление: Учебное пособие / Н.Б. Плещинский. – Казань: Казанский федеральный университет, 2022. – 80 с.

Изложены основы классического и интегрального вариационного исчисления и элементы теории оптимального управления.

Для бакалавров и магистрантов математических институтов и факультетов, изучающих теорию и методы решения экстремальных задач.

Предисловие

Учебное пособие представляет собой расширенный и дополненный курс лекций, который автор читает студентам Института вычислительной математики и информационных технологий Казанского федерального университета.

Первая версия этого курса была разработана под влиянием книги Алексеева В.М., Тихомирова В.М., Фомина С.В. "Оптимальное управление", одно из достоинств которой – единый подход к различным задачам на экстремум. В дальнейшем добавлены новые разделы, написанные в основном по работам автора пособия.

Вариационное исчисление и оптимальное управление – близкие друг к другу разделы общей теории экстремальных задач. Неизвестными величинами в задачах ВИ и ОУ обычно являются функции одной или нескольких переменных.

В классическом ВИ необходимые условия экстремума получены в форме краевых задач для дифференциальных уравнений. В задачах интегрального вариационного исчисления необходимыми условиями экстремума являются интегральные уравнения. При наличии дополнительных ограничений используется правило множителей Лагранжа (принцип Лагранжа).

В задачах оптимального управления предполагается, что исследуемый процесс характеризуется двумя наборами величин – состоянием и управлением. Чаще всего состояние является решением некоторого дифференциального или интегрального уравнения, а управление входит в это уравнение как параметр. В пособии рассматриваются только самые простые задачи теории оптимального управления. Основное внимание уделено градиентному методу их решения. Ограниченный объем курса не позволяет обсудить многие важные и интересные разделы теории.

PNB

1. ВВЕДЕНИЕ. ПОСТАНОВКИ ЗАДАЧ

Обсудим общие подходы к постановке экстремальных задач, а затем познакомимся с некоторыми конкретными задачами ВИ и ОУ. Несмотря на разное содержание, все они могут рассматриваться как частные случаи общей экстремальной задачи. О том, как возникают экстремальные задачи и как они формализуются, подробно написано в первой главе книги [1].

1.1. Основные понятия и терминология

Абстрактная экстремальная задача ставится так:

$$f(x) \to \text{extr}, \quad x \in \mathfrak{x} \subset X.$$

В ее постановку входят три компонента: *множество* X, на котором определен вещественнозначный *функционал* $f(x)$ (критерий качества), и *ограничения*, которые выделяют подмножество $\mathfrak{x}$ в X.

Поиск экстремума – это поиск или максимума, или минимума. Предполагается, что в множестве X определены окрестности $U(x)$ элементов x. Тогда говорят, что элемент x^0 доставляет локальный минимум функционалу $f(x)$, если $f(x) \geq f(x^0) \quad \forall x \in U(x^0)$. Элемент x^0 доставляет локальный максимум функционалу $f(x)$, если $f(x) \leq f(x^0) \quad \forall x \in U(x^0)$. Иногда требуется найти только элемент x^0, иногда – значение $f(x^0)$ функционала на нем, а по умолчанию – и то, и другое.

Обычно в качестве множества X рассматривается некоторое нормированное пространство, то есть линейное пространство с нормой. Напомним, что в линейном пространстве определены операции сложения элементов и умножение элемента на число. В нормированном пространстве каждому элементу x поставлено в соответствие вещественное число $||x||$ (норма элемента), причем так, что 1) $||x|| \geq 0$; если $||x|| = 0$, то $x = 0$ (нулевой элемент), 2) $||\lambda x|| = |l| \cdot ||x||$;

3) $||x_1 + x_2|| \leq ||x_1|| + ||x_2||$. Окрестность элемента x^0 – это множество таких элементов x, что $||x - x^0|| < \varepsilon$. Если нужно освежить в памяти основы функционального анализа, рекомендуем учебник [2].

Пусть часть ограничений задана как равенство вида $F(x) = 0$, которому должны удовлетворять искомые элементы. Тогда постановка экстремальной задачи уточняется:

$$f(x) \to \mathrm{extr}, \quad F(x) = 0, \quad x \in \mathfrak{x} \subset X.$$

Здесь возможны два подхода: или нужно найти элементы, доставляющие экстремум функционалу при одном из дополнительных ограничений в виде равенства; или нужно найти решения уравнения $F(x) = 0$, доставляющие экстремум.

В задачах управления предполагается, что *состояние* x некоторого объекта зависит от *управления* u. Пару связанных друг с другом элементов (x, u) обычно называют процессом. В задачах оптимального управления нужно найти *оптимальный процесс* (x^0, u^0), то есть наилучшие состояние и управление. Иными словами, элементы, доставляющие экстремум заданному функционалу.

Поэтому общая постановка абстрактной задачи ОУ может быть такой:

$$f(x, u) \to \mathrm{extr}, \quad F(x, u) = 0, \quad (x, u) \in \mathcal{M} \subset X \times U.$$

Состояние и управление с формальной точки зрения равноправны: можно искать оптимальное управление, а по нему определять оптимальное состояние. Можно действовать наоборот, по найденному оптимальному состоянию вычислять соответствующее ему управление.

Зависимость между состоянием и управлением вида $F(x, u) = 0$ называют *уравнением состояний*. В экстремальных задачах математической физики в качестве уравнений состояний обычно рассматриваются *дифференциальные уравнения* с соответствующими граничными условиями, а также *интегральные уравнения*. Если уравнение состояний удается разрешить относительно одного из искомых элементов, то в некоторых случаях задача ОУ становится более простой.

Еще раз обратим внимание на возможность рассматривать задачу ОУ с двух позиций: или мы ищем процесс, доставляющий экстремум функционалу и при этом удовлетворяющий уравнению состояний, или мы хотим найти такое решение уравнения состояний, которое доставляет экстремум этому функционалу.

Различают *стационарные* и *нестационарные* задачи ОУ, то есть или состояние и управление не зависят от времени, или эта зависимость имеет место.

В задачах *с сосредоточенными параметрами* рассматриваются конечномерные пространства X и U, а в задачах *с распределенными параметрами* – бесконечномерные. В первом случае имеют дело с обыкновенными дифференциальными уравнениями, во втором случае – с уравнениями с частными производными.

Как уже было сказано, в задачах ВИ и ОУ искомыми элементами являются функции. Среди функционалов, определенных на множествах функций, выделяют *интегральные*, *терминальные* и *смешанные* функционалы. Например, в одной из задач классического вариационного исчисления

$$\int\limits_{t_0}^{t_1} L(t, x(t), x'(t))\, dt + l(x(t_0), x(t_1)) \to \mathrm{extr}$$

смешанный функционал состоит из двух слагаемых. Первое слагаемое – интегральный функционал, его подинтегральная функция зависит от значений искомой функции и ее производной. Второе слагаемое – терминальный функционал, его значения вычисляются по значениям искомой функции только в конечном числе точек (точнее, в двух точках). В этой задаче $X = C^1[t_0, t_1]$ – пространство функций, определенных на отрезке $[t_0, t_1]$ и имеющих непрерывную производную.

В дальнейшем, если это не указано явно, $X = C^1[t_0, t_1]$.

1.2. Задачи

Рассмотрим несколько конкретных постановок задач ВИ и ОУ.

Задача Дидоны (825 г. до н.э.)

Как отгородить веревкой заданной длины участок морского берега максимальной площади? Математически эта задача может быть сформулирована так:

$$\int\limits_{t_0}^{t_1} x(t)\, dt \to \sup, \quad \int\limits_{t_0}^{t_1} \sqrt{1 + [x'(t)]^2}\, dt = l,$$

$$x(t_0) = 0, \quad x(t_1) = 0, \quad x(\cdot) \in C^1[t_0, t_1].$$

Здесь границу участка суши определяет график функции $x(t)$. Первый интеграл – площадь области, второй интеграл – длина дуги кривой. Крайние точки искомой линии заданы.

Интуиция подсказывает, что веревку следует расположить по дуге окружности. А что, если линия берега – не прямая?

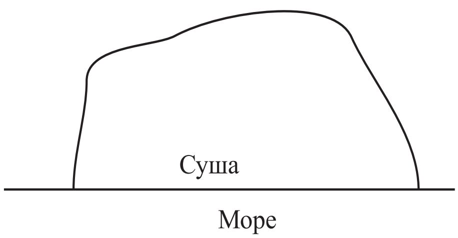

Рис. 1. Задача Дидоны

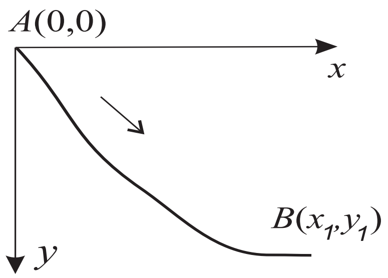

Рис 2. Задача о брахистохроне

Задача о брахистохроне (И. Бернулли, 1696)

Пусть материальное тело движется в вертикальной плоскости под действием силы тяжести вдоль некоторой гладкой кривой. Какой должна быть эта кривая, чтобы путь был пройден за минимальное время?

Предположим, что тело движется из точки $A(0,0)$ (начало координат) в точку $B(x_1, y_1)$. По закону Галилея в точке с ординатой $y(x)$ скорость тела $\sqrt{2gy(x)}$, g – ускорение свободного падения. Так как на участке от $(x, y(x))$ до $(x+dx, y(x)+dy)$ длина дуги $ds = \sqrt{dx^2+dy^2}$, то задача ставится так:

$$\int\limits_0^{x_1} \frac{\sqrt{1+[y'(x)]^2}}{\sqrt{2gy(x)}}\,dx \to \inf, \quad y(0)=0, \;\; y(x_1)=y_1 \;\; y(x) \in C^1[0,x_1].$$

Почему скорость тела в точке $(x, y(x))$ определяет выражение $\sqrt{2gy(x)}$? Так как $v'(t)=g$, то $v(t)=gt$ (начальная скорость нулевая). Так как $s'(t)=v(t)$, то $s(t)=gt^2/2$ (начальная точка – начало координат). Тогда $v^2(t)=g^2t^2=g\,2s(t)$.

Задача о распределении нагрузки вдоль упругой струны

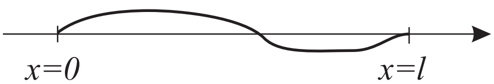

Рис. 3. Упругая струна

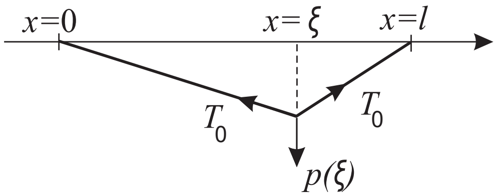

Рис 4. Сосредоточенная сила

Если на тонкую упругую невесомую нить действует сила, то нить отклоняется от положения равновесия. Как должна быть распределена нагрузка вдоль струны, чтобы ее отклонение было наиболее близким к заданному? Суммарная нагрузка на струну задана.

Эта задача формализуется следующим образом:

$$\int_0^l \Big[\int_0^l p(\xi)G(x,\xi)\,d\xi - u_0(x)\Big]^2 dx \to \min, \quad \int_0^l p(\xi)\,d\xi = P,$$

где

$$G(x,\xi) = \{x \le \xi\ :\ \frac{x(l-\xi)}{T_0 l};\ \ x \ge \xi\ :\ \frac{(l-x)\xi}{T_0 l}\}.$$

Здесь $p(\xi)$ – сила, действующая на струну, $G(x,\xi)$ – функция Грина, $u_0(x)$ – желаемое отклонение точек струны от положения равновесия, l – длина струны, T_0 – начальное натяжение струны.

Решение этой экстремальной задачи дано в п.3.3.

Поясним, почему отклонение струны в точке с координатой x вычисляется по формуле

$$u(x) = \int_0^l p(\xi)\,G(x,\xi)\,d\xi.$$

Пусть начальное натяжение струны T_0. Если в точке с координатой ξ приложена малая сила $p(\xi)$, то можно считать, что натяжение струны и ее длина не изменились. Пусть $u(\xi)$ – отклонение струны от положения равновесия в точке ξ (см. рис. 4). Проекции трех сил на вертикальную прямую дают равенство

$$T_0\,\frac{u(\xi)}{\xi} + T_0\,\frac{u(\xi)}{l-\xi} = p(\xi).$$

Отсюда $u(\xi) = p(\xi)\,\dfrac{\xi(l-\xi)}{T_0 l}$. Тогда для точки с координатой x (подобные треугольники)

$$u(x) = \frac{x}{\xi}\,u(\xi) \quad \text{при} \quad x \le \xi; \quad u(x) = \frac{l-x}{l-\xi}\,u(\xi) \quad \text{при} \quad x \ge \xi.$$

Аналогичная задача возникает, если нужно распределить нагрузку вдоль упругой мембраны, закрепленной по краю.

Оптимальное управление колебаниями струны

([3], Гл.1, §8)

Рассмотрим теперь нестационарную задачу. Управляя внешними силами нужно привести колеблющуюся струну в состояние, наиболее близкое к заданному. Состояние струны в каждой ее точке и в каждый момент времени определяют отклонение от положения равновесия и скорость. В курсе УМФ (уравнения математической физики) было показано, что процесс колебаний струны описывает уравнение с частными производными 2-го порядка вместе с граничными и начальными условиями. Тогда имеем задачу ОУ

$$\beta_0 \int_0^l [x(s,T,u) - x_0(s)]^2\, ds + \beta_1 \int_0^l [\frac{\partial x}{\partial t}(s,T,u) - x_1(s)]^2\, ds \to \min,$$

$$\frac{\partial^2 x}{\partial t^2} - a^2 \frac{\partial^2 x}{\partial s^2} = f(s,t), \quad 0 < s < l, \ \ 0 < t < T,$$

$$\frac{\partial x}{\partial s}(0,t) = p(t), \quad \frac{\partial x}{\partial s}(l,t) = 0, \quad x(s,0) = \varphi_0(s), \quad \frac{\partial x}{\partial t}(s,0) = \varphi_1(s).$$

В этой задаче управление – это сила $p(t)$, действующая на левый конец струны, и сила $f(s,t)$, распределенная вдоль струны. Заметим, что с помощью весовых коэффициентов β_0 и β_1 можно установить, что важнее в итоге: форма струны или ее скорость.

Оптимальное управление температурой стержня

([3], Гл.1, §7)

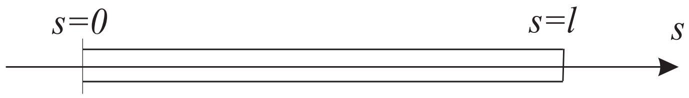

Рис. 5. Распространение тепла в стержне

Задача о распространении тепла по тонкому однородному стержню также рассматривалась в курсе УМФ. Температура $x(s,t)$ в точке с координатой s в момент времени t удовлетворяет уравнению теплопроводности. Если левый конец стержня изолирован, а через правый происходит обмен теплом с окружающей средой, то математическая модель процесса

$$\frac{\partial x}{\partial t} = a^2\frac{\partial^2 x}{\partial s^2} + f(s,t), \quad (s,t) \in Q = \{0 < s < l,\ 0 < t < T\},$$

$$\frac{\partial x}{\partial s}\Big|_{s=0} = 0,\ \ 0 < t \le T, \quad \frac{\partial x}{\partial s}\Big|_{s=l} = \nu[p(t) - x(l,t)],\ \ 0 < t \le T,$$

$$x|_{t=0} = \varphi(s),\ \ 0 \le s \le l.$$

При какой плотности источников тепла $f(s,t)$ внутри стержня и при какой температуре окружающей среды $p(t)$ распределение температуры в стержне будет наиболее близким к заданному распределению $x_0(s)$? Добавим условие экстремума

$$\int_0^l [x(s,T) - x_0(s)]^2\, ds \to \min.$$

Решение этой экстремальной задачи дано в п.6.3.

Оптимальное управление летательным аппаратом

Состояние движущегося объекта характеризуется пространственными координатами его центра масс, ориентацией относительно выбранной системы отсчета (если нужно) и скоростью (это вектор). Управление осуществляется, например, за счет изменения силы тяги – по абсолютному значению и по направлению. Имеют практическое значение самые различные постановки задач оптимального управления. В их числе:

– перевести объект в заданную точку пространства за минимальное время или при минимальных затратах топлива;

– обеспечить максимальную дальность или максимальную высоту полета;

– перехватить цель, то есть обеспечить режим полета, при котором траектория движения объекта пересечет траекторию движения цели (именно в тот момент, когда цель будем там).

Дополнительные ограничения в подобных задачах самые естественные. Например, объект не может опуститься ниже поверхности Земли; скорость его движения не может превысить технически возможную; масса топлива и расход его в единицу времени не могут быть отрицательными и так далее.

В теории ОУ особое внимание уделяется задачам управления движущимися объектами. Рекомендуем познакомиться с задачей о мягкой посадке космического аппарата на поверхность небесного тела [4], п.12.1.

Модель Ланчестера динамики боя

При планировании боевых действий иногда рассматривают задачи оптимального управления следующего типа. Пусть $x_j(t)$ – количества боевых единиц сторон, участвующих в конфликте, $u_j(t)$ – скорости поступления резервов. Уравнения состояний вместе с начальными условиями в простейшем случае имеют вид

$$x_1'(t) = -a_2(t)\,x_2(t) + u_1(t), \quad x_1(t_0) = x_1^0,$$
$$x_2'(t) = -a_1(t)\,x_2(t) + u_2(t), \quad x_2(t_0) = x_2^0.$$

Естественные ограничения таковы, например, что $x_j(t) \geq 0$, $0 \leq u_j(t) \leq u_j^0$. Цели второй воюющей стороны могут быть разные:

$$x_2(t_1) \to \min, \quad x_1(t_1) \to \max \quad \text{или} \quad \int\limits_{t_0}^{t_1} u_2(t)\,dt \to \min.$$

Простейшая задача о быстродействии ([1], с.37)

Рассмотрим одну из исторически первых задач ОУ. Нужно перегнать тележку с одной позиции на другую за минимальное время при ограничении на силу тяги двигателя. Задача ставится так:

$$t_1 - t_0 \to \inf, \quad mx''(t) = u(t), \quad x(t_0) = x_0, \ \ x'(t_0) = x_1,$$

$$x(t_1) = a, \ \ x'(t_1) = 0, \ \ |u(t)| \le u_0 \ \ \forall t \in [t_0, t_1].$$

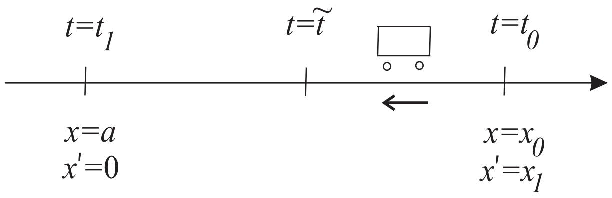

Рис. 6. Простейшая задача о быстродействии

Попробуем найти ее решение. Примем гипотезу: тележка должна двигаться на максимальной тяге в направлении указанной позиции, а в какой-то момент времени $\tilde{t}$ сила тяги должна изменить направление на противоположное.

Пусть точка a расположена левее точки x_0 на оси x. При движении от точки x_0 до точки переключения направления силы тяги уравнение движения $mx''(t) = -u_0$. Тогда $x'(t) = -\dfrac{u_0}{m}t + c_1$ и $x(t) = -\dfrac{u_0}{m}\dfrac{t^2}{2} + c_1 t + c_2$. Из условий при $t = t_0$ находятся значения $c_1 = x_1 + \dfrac{u_0}{m}t_0$ и $c_2 = x_0 - x_1 t_0 - \dfrac{u_0}{m}\dfrac{t_0^2}{2}$. Следовательно,

$$x(t) = -\frac{u_0}{2m}(t - t_0)^2 + x_1(t - t_0) + x_0.$$

При движении от точки переключения до точки a уравнение движения $mx''(t) = u_0$. Тогда $x'(t) = \dfrac{u_0}{m}t + c_1$ и $x(t) = \dfrac{u_0}{m}\dfrac{t^2}{2} + c_1 t + c_2$. Из условий при $t = t_1$ находятся значения $c_1 = -\dfrac{u_0}{m}t_1$ и $c_2 = a + \dfrac{u_0}{m}\dfrac{t_1^2}{2}$. Следовательно,

$$x(t) = \frac{u_0}{2m}(t_1 - t)^2 + a.$$

При $t = \tilde{t}$ найденные решения должны быть равны:

$$-\frac{u_0}{2m}(\tilde{t} - t_0)^2 + x_1(\tilde{t} - t_0) + x_0 = \frac{u_0}{2m}(t_1 - \tilde{t})^2 + a.$$

Отсюда нужно выразить t_1 через $\tilde{t}$ и искать $\tilde{t}$ так, чтобы это выражение было минимальным.

Задача о расширенном воспроизводстве ([5], с.9)

Пусть $x(t)$ – выручка от продажи продукции предприятия и $u(t)$ – ее доля, направляемая на расширение производства. Тогда уравнение состояний (с начальным условием)

$$\frac{dx}{dt} = \alpha u(t)x(t), \;\; \alpha > 0, \quad x(0) = a, \;\; a > 0.$$

Пусть расходы предприятия пропорциональны доходу: $\beta x(t), \;\; \beta > 0$. Ясно, что $0 \le u(t) \le 1 - \beta$.

Сумма налога назначается по остатку $x(t) - \beta x(t) - u(t)x(t)$. Поэтому чистая прибыль — также функция от остатка. Отсюда возникает следующая задача:

$$\Phi[x, u] = \int\limits_0^T \varphi[(1 - \beta - u(t))x(t)]\, dt \to \max, \quad 0 \le u(t) \le 1 - \beta.$$

* * *

Другие интересные задачи ВИ и ОУ рассмотрены в книгах [1], [4], [5], [6].

2. КЛАССИЧЕСКОЕ ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ

Исторически первой задачей вариационного исчисления была задача о брахистохроне. В задачах КВИ рассматриваются в первую очередь интегральные функционалы, подинтегральные функции которых зависят от искомых функций и их производных. Необходимыми условия экстремума являются граничные задачи для дифференциальных уравнений. .

Будем придерживаться следующих обозначений: $x(\cdot)$ – функция, а $x(t)$ – ее значение при аргументе t.

2.1. Простейшая задача КВИ

Простейшая задача КВИ состоит в следущем:

$$\mathrm{I}\,[x(\cdot)] = \int\limits_{t_0}^{t_1} L(t, x(t), x'(t))\, dt \to \mathrm{extr}, \quad x(t_0) = x_0, \;\; x(t_1) = x_1. \quad (1)$$

Норму в пространстве $C^1([t_0, t_1])$ зададим так:

$$||x(\cdot)|| = \max(\max_{t\in[t_0,t_1]} |x(t)|, \max_{t\in[t_0,t_1]} |x'(t)|).$$

Получим необходимое условие экстремума в задаче (1) на основании известного принципа: если отойти от точки экстремума, то значение функционала перестает быть оптимальным. В математическом анализе имеет место теорема Ферма: если функция достигает экстремума в точке и дифференцируема, то ее производная равна нулю в этой точке. Для функционалов, зависящих от функций, используется аналогичное понятие – вариация.

Чтобы дополнительные ограничения $x(t_0) = x_0, \quad x(t_1) = x_1$ всегда были выполнены, будем рассматривать приращения функций $x(\cdot)$ из подпространства $C_0^1[t_0, t_1]$ – множества непрерывно дифференцируемых функций, удовлетворяющих условиям $x(t_0) = 0$, $x(t_1) = 0$. Пространство C_0^1 – подпространство пространства C^1.

Пусть $h(\cdot) \in C_0^1$. *Первой вариацией* функционала $\mathrm{I}\,[x(\cdot)]$ называется

$$\delta\mathrm{I}\,[x(\cdot); h(\cdot)] = \lim_{\lambda\downarrow 0} \frac{\mathrm{I}\,[x(\cdot) + \lambda h(\cdot)] - \mathrm{I}\,[x(\cdot)]}{\lambda}$$

(если этот предел существует). Запись $\lambda \downarrow 0$ означает, что λ стремится к нулю справа, то есть оставаясь положительным. Ясно, что

$$\delta \mathrm{I}\left[x(\cdot); h(\cdot)\right] = \frac{d}{d\lambda} \mathrm{I}\left[x(\cdot) + \lambda h(\cdot)\right]\Big|_{\lambda=0}.$$

Легко видеть, что если $x^0(\cdot)$ – решение задачи (1) и первая вариация функционала существует, то $\delta \mathrm{I}\left[x^0(\cdot); h(\cdot)\right] = 0 \quad \forall h(\cdot) \in C_0^1$ (*необходимое условие экстремума*).

Докажем одно вспомогательное утверждение, которое часто называют основной леммой ВИ.

Лемма 1. (лемма Лагранжа)
Если $a(\cdot) \in C[t_0, t_1]$ *и*

$$\int\limits_{t_0}^{t_1} a(t)\, h(t)\, dt = 0 \quad \forall h(\cdot) \in C_0^1[t_0, t_1],$$

то $a(t) = 0\ \ \forall t \in [t_0, t_1]$.

Доказательство. От противного. Предположим, что $\exists \tilde{t} \mid a(\tilde{t}) > 0$. Тогда $\exists \varepsilon > 0 \ \mid\ a(t) > a(\tilde{t})/2 \quad \forall t \in (\tilde{t} - \varepsilon, \tilde{t} + \varepsilon) \subset (t_0, t_1)$.
Функция $\tilde{h}_\varepsilon(t) = 1 + \cos \dfrac{\pi(t - \tilde{t})}{\varepsilon}$ при $t \in (\tilde{t} - \varepsilon, \tilde{t} + \varepsilon)$
и $\tilde{h}_\varepsilon(t) = 0$ при $t \notin (\tilde{t} - \varepsilon, \tilde{t} + \varepsilon)$ такова, что

$$\int\limits_{t_0}^{t_1} a(t)\, \tilde{h}_\varepsilon(t)\, dt \neq 0.$$

Противоречие.

Если $\tilde{t} = t_0$ или $\tilde{t} = t_1$, то нужно рассматривать полуокрестности этой точки. •

Вычислим вариацию интегрального функционала. Пусть функция $L(t, x, y)$ и ее частные производные $\dfrac{\partial L}{\partial x}$, $\dfrac{\partial L}{\partial y}$ непрерывны. Тогда

$$\delta \mathrm{I}\left[x(\cdot); h(\cdot)\right] = \int\limits_{t_0}^{t_1} \Big\{ \frac{\partial L}{\partial x}(t, x(t), x'(t))\, h(t) + \frac{\partial L}{\partial y}(t, x(t), x'(t))\, h'(t) \Big\}\, dt.$$

Предположим дополнительно, что у функции $L(t,x,y)$ непрерывны 2-е частные производные. Проинтегрируем по частям и получим

$$\delta \mathrm{I}\,[x(\cdot);h(\cdot)] = \int\limits_{t_0}^{t_1}\{\frac{\partial L}{\partial x}(t,x(t),x'(t)) - \frac{d}{dt}\,\frac{\partial L}{\partial y}(t,x(t),x'(t))\}\,h(t)\,dt.$$

Теорема 1. (необходимое условие экстремума 1-го порядка) *Пусть функция $L(t,x,y)$ имеет непрерывные 2-е частные производные. Если функция $x^0(\cdot)$ – решение задачи (1), то она удовлетворяет дифференциальному уравнению Эйлера*

$$\frac{\partial L}{\partial x}(t,x(t),x'(t)) - \frac{d}{dt}\,\frac{\partial L}{\partial y}(t,x(t),x'(t)) = 0, \quad t \in (t_0,t_1), \qquad (2)$$

и краевым условиям $x(t_0) = x_0,\ \ x(t_1) = x_1$.

Доказательство. Приравняем нулю вариацию функционала, а затем применим лемму 1. •

Любое решение уравнения Эйлера (2) называют *экстремалью.* Помним, что теорема 1 дает только необходимое условие экстремума. Решение задачи (1) – экстремаль, но не любая экстремаль – решение.

Пример.

$$\int\limits_0^1 [x'(t)]^2\,dt \to \min, \quad x(0)=0, \quad x(1)=1, x(\cdot) \in C^1[0,1].$$

Ответ: $x^0(t) = t$. Заметим, что минимальное значение функционала достигается при $x'(t)=0$, но функция $x(t)=c$ не удовлетворяет дополнительным ограничениям.

Пример. Чтобы доказать, что кратчайшей линией, соединяющей две точки на плоскости, может быть только прямая, рассмотрим экстремальную задачу

$$\int\limits_{t_0}^{t_1}\sqrt{1+[x'(t)]^2}\,dt \to \min, \quad x(t_0)=x_0,\ \ x(t_1)=x_1, \quad x(\cdot) \in C^1[t_0,t_1].$$

Из теоремы 1 следует, если такая линия существует, то она – прямая. Но существует ли она?

Замечание. Достаточно считать, что функция L и ее производные $\dfrac{\partial L}{\partial x}$ и $\dfrac{\partial L}{\partial y}$ непрерывны. Чтобы убедиться в этом, докажем еще одно полезное вспомогательное утверждение.

Лемма 2. (лемма Дюбуа-Реймона)
Если $a_0(\cdot),\ a_1(\cdot) \in C([t_0, t_1])$ *и*

$$\int_{t_0}^{t_1} [a_0(t)\, h(t) + a_1(t)\, h'(t)]\, dt = 0 \quad \forall h(\cdot) \in C_0^1,$$

то $a_1(\cdot) \in C^1$ *и* $a_0(t) - a_1'(t) = 0\ \ \forall t \in (t_0, t_1)$.

Доказательство (по книге [7]).

Пусть $p(\cdot) \in C^1$, причем $p'(t) = a_0(t)$ и $\int_{t_0}^{t_1} p(t)\ dt = \int_{t_0}^{t_1} a_1(t)\ dt$. Тогда, после интегрирования по частям,

$$\int_{t_0}^{t_1} [a_1(t) - p(t)] h'(t)\ dt = 0 \quad \forall h(\cdot) \in C_0^1.$$

Пусть $q(\cdot) \in C^1$, причем $q'(t) = a_1(t) - p(t)$ и $q(t_0) = 0$. Из свойств функции $p(\cdot)$ следует, что $q(t_1) = 0$. Значит $q(\cdot) \in C_0^1$. Положим $h(\cdot) = q(\cdot)$ и получим, что

$$\int_{t_0}^{t_1} [a_1(t) - p(t)]^2\ dt = 0.$$

Отсюда следует, что $a_1(t) = p(t)$, т. е. $a_1(\cdot) \in C^1$ и $a_1'(t) = a_0(t)$. •

2.2. Задача Больца

В *задаче Больца*

$$\mathrm{B}\,[x(\cdot)] = \int_{t_0}^{t_1} L(t, x(t), x'(t))\, dt + l(x(t_0), x(t_1)) \to \mathrm{extr} \qquad (3)$$

функционал состоит из двух слагаемых: интегральной части и терминальной части. Дополнительных ограничений нет.

Теорема 2.(необходимое условие экстремума 1-го порядка в задаче Больца)
Пусть функции $L(t,x,y)$ и $l(x_0,x_1)$ имеют непрерывные производные. Если функция $x^0(\cdot)$ – решение задачи (3), то она удовлетворяет дифференциальному уравнению Эйлера (2) и условиям трансверсальности

$$\frac{\partial L}{\partial y}(t_0,x(t_0),x'(t_0))=\frac{\partial l}{\partial x_0}(x(t_0),x(t_1)),$$

$$\frac{\partial L}{\partial y}(t_1,x(t_1),x'(t_1))=-\frac{\partial l}{\partial x_1}(x(t_0),x(t_1)). \tag{4}$$

Доказательство. Вычислим вариацию функционала, но не будем предполагать, что $h(t_0)=0,\quad h(t_1)=0$, и приравняем ее нулю. Получим

$$\delta\text{B}\,[x(\cdot);h(\cdot)]=\int\limits_{t_0}^{t_1}\{\frac{\partial L}{\partial x}(t,x(t),x'(t))-\frac{d}{dt}\frac{\partial L}{\partial y}(t,x(t),x'(t))\}\,h(t)\,dt+$$

$$+\frac{\partial L}{\partial y}(t,x(t),x'(t))h(t)\Big|_{t_0}^{t_1}+\frac{\partial l}{\partial x_0}(x(t_0),x(t_1))h(t_0)+$$

$$+\frac{\partial l}{\partial x_1}(x(t_0),x(t_1))h(t_1)=0.$$

Если $h(\cdot)\in C_0^1$, то отсюда сразу следует (2). Если $h(t_0)\neq 0,\quad h(t_1)=0$ или $h(t_0)=0,\quad h(t_1)\neq 0$, то получим (4). $\bullet$

Латинское слово transversus – повернутый в сторону.

Пример.

$$\text{B}\,[x(\cdot)]=\int\limits_0^1[x'^2(t)-x(t)]\;dt+x^2(1)\to\text{extr},\quad x(\cdot)\in C^1[0,1].$$

В этом случае уравнения Эйлера имеет вид $2x'' + 1 = 0$ и условия трансверсальности $x'(0) = 0,\ x'(1) = -x(1)$. Отсюда следует, что решением задачи Больца может быть только функция $x^0(t) = \dfrac{3-t^2}{4}$.

Как узнать, действительно ли она является искомым решением? Нужно убедиться, что выражение $\mathrm{B}\,[x^0(\cdot)+h(\cdot)] - \mathrm{B}\,[x^0(\cdot)]$ не меняет знак. Одновременно выяснится, максимум или минимум доставляет функция $x^0(\cdot)$ (а вдруг экстремума вовсе нет?).

Должно получиться так:

$$\mathrm{B}\,[x^0(\cdot) + h(\cdot)] - \mathrm{B}\,[x^0(\cdot)] = \int\limits_0^1 [h'(t)]^2\,dt + h^2(1) \geq 0.$$

2.3. Изопериметрическая задача

Пусть

$$\mathrm{I}_{\,j}[x(\cdot)] = \int\limits_{t_0}^{t_1} L_j(t, x(t), x'(t))\,dt, \quad j = 0..m.$$

Рассмотрим *изопериметрическую задачу*

$$\mathrm{I}_{\,0}[x(\cdot)] \Longrightarrow \mathrm{extr}, \quad \mathrm{I}_{\,j}[x(\cdot)] = c_j,\ j = 1..m,$$

$$x(t_0) = x_0,\ \ x(t_1) = x_1, \quad x(\cdot) \in C^1[t_0, t_1]). \tag{5}$$

Теорема 3. (правило множителей Лагранжа)
Пусть функции $L_j,\ \dfrac{\partial L_j}{\partial x},\ \dfrac{\partial L_j}{\partial y}$ *непрерывны. Если* $x^0(\cdot)$ *– решение задачи (5), то найдутся такие числа* $\lambda_0, \lambda_1, \ldots, \lambda_m$ *(не равные нулю одновременно), что* $x^0(\cdot)$ *удовлетворяет уравнению Эйлера (2), составленному для функции*

$$L(t, x, y) = \sum_{j=0}^{m} \lambda_j\, L_j(t, x, y).$$

Если все функционалы $I_{\,j}[x(\cdot)]$ *не имеют стационарных точек, то можно считать, что* $\lambda_0 = 1$.

Доказательство проведем тем же способом, что в п. 1.4.5 книги [1]. Ограничимся случаем $m = 1$.

Если $\delta \mathrm{I}_1[x^0(\cdot); h(\cdot)] = 0 \quad \forall h(\cdot) \in C_0^1$ (то есть функция $x^0(\cdot)$ – стационарная точка функционала $\mathrm{I}_1[x(\cdot)]$), то можно выбрать $\lambda_0 = 0$ и $\lambda_1 = 1$.

Пусть $h(\cdot)$, $g(\cdot) \in C_0^1$, причем $\delta \mathrm{I}_1[x^0(\cdot); g(\cdot)] \neq 0$. Рассмотрим функции

$$\varphi(\xi, \eta) = \mathrm{I}_0[x^0(\cdot) + \xi h(\cdot) + \eta g(\cdot)], \quad \psi(\xi, \eta) = \mathrm{I}_1[x^0(\cdot) + \xi h(\cdot) + \eta g(\cdot)],$$

зависящие от вещественных переменных ξ и η. Эти функции непрерывно дифференцируемы в окрестности точки $(0, 0)$, легко вычислить значения их частных производных в этой точке:

$$\frac{\partial \varphi}{\partial \xi}(0,0) = \delta I_0[x^0(\cdot); h(\cdot)], \quad \frac{\partial \varphi}{\partial \eta}(0,0) = \delta I_0[x^0(\cdot); g(\cdot)],$$

$$\frac{\partial \psi}{\partial \xi}(0,0) = \delta I_1[x^0(\cdot); h(\cdot)], \quad \frac{\partial \psi}{\partial \eta}(0,0) = \delta I_1[x^0(\cdot); g(\cdot)],$$

Их значения в точке $(0, 0) : \varphi(0, 0) = \mathrm{I}_0[x^0(\cdot)]$, $\psi(0, 0) = \mathrm{I}_1[x^0(\cdot)]$.

Предположим, что

$$\frac{D(\varphi, \psi)}{D(\alpha, \beta)}(0, 0) \neq 0 \quad \forall h(\cdot) \in C_0^1.$$

По теореме об обратном отображении в окрестности точки $\varphi(0, 0)$, $\psi(0, 0)$ существует непрерывно дифференцируемое отображение, обратное к отображению $(\xi, \eta) \mapsto (\varphi(\xi, \eta), \psi(\xi, \eta))$, переводящему точку $(0, 0)$ в точку $\varphi(0, 0)$, $\psi(0, 0)$. Тогда найдутся такие значения ξ', η', что или $\varphi(\xi', \eta') > \mathrm{I}_0[x^0(\cdot)]$, или $\varphi(\xi', \eta') < \mathrm{I}_0[x^0(\cdot)]$, но $\psi(\xi', \eta') = \mathrm{I}_1[x^0(\cdot)]$.

Тогда функция $x^0(\cdot)$ не может быть решением изопериметрической задачи. Поэтому

$$\frac{D(\varphi, \psi)}{D(\xi, \eta)}(0, 0) = 0 \quad \forall h(\cdot) \in C_0^1$$

или

$$\delta \mathrm{I}_0[x^0(\cdot); h(\cdot)] - \frac{\delta \mathrm{I}_0[x^0(\cdot); g(\cdot)]}{\delta \mathrm{I}_1[x^0(\cdot); g(\cdot)]}\, \delta \mathrm{I}_1[x^0(\cdot); h(\cdot)] = 0.$$

Следовательно, можно взять

$$\lambda_0 = 1, \quad \lambda_1 = -\frac{\delta \mathrm{I}_0[x^0(\cdot); g(\cdot)]}{\delta \mathrm{I}_1[x^0(\cdot); g(\cdot)]}. \bullet$$

Как отмечено в [1], идею доказательства теоремы 3 предложил Эйлер в 1744 году.

Пример.

$$\int\limits_0^1 [x'(t)]^2\, dt \to \mathrm{extr}, \quad \int\limits_0^1 x(t)\, dt = 0, \quad x(0) = 0, \quad x(1) = 1.$$

Ответ: единственная допустимая экстремаль $x^0(t) = 3t^2 - 2t$.

Пример. Задача Дидоны (см. п.1.2).

$$\int\limits_{t_0}^{t_1} x(t)\, dt \to \max, \ \int\limits_{t_0}^{t_1} \sqrt{1 + x'^2(t)}\, dt = l, \ x(t_0) = x_0, \ x(t_1) = x_1.$$

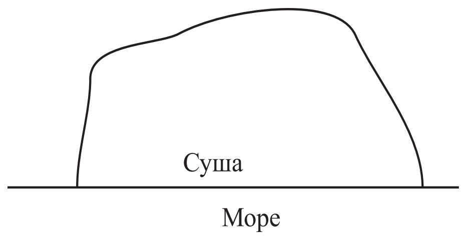

Рис. 1. Задача Дидоны

Пример. Какую форму примет тяжелая гибкая нерастяжимая нить длины l, концы которой закреплены в точках (x_0, y_0) и (x_1, y_1) ([8], с. 140)?

Форма нити должна быть такой, чтобы центр тяжести занял самое низкое положение. Тогда задача ставится так:

$$\int\limits_{x_0}^{x_1} y(x)\sqrt{1+[y'(x)]^2}\,dx \to \min, \quad \int\limits_{x_0}^{x_1} \sqrt{1+[y'(x)]^2}\,dx = l,$$

$$y(x_0)=y_0, \;\; y(x_1)=y_1, \quad y(\cdot)\in C^1[x_0,x_1].$$

Ее решение называют цепной линией.

2.4. Задача с подвижными концами

В простейшей задаче КВИ были заданы начало и конец графика искомой функции. В задаче с подвижными концами эти точки должны лежать на заданных линиях: $x(t_0)=\varphi_0(t_0),\; x(t_1)=\varphi_1(t_1)$, значения t_0 и t_1 также требуется найти. Такая задача возникает, если, например, по одной линии движется пусковая установка, а по другой линии движется цель. Нужно произвести пуск ракеты и поразить цель.

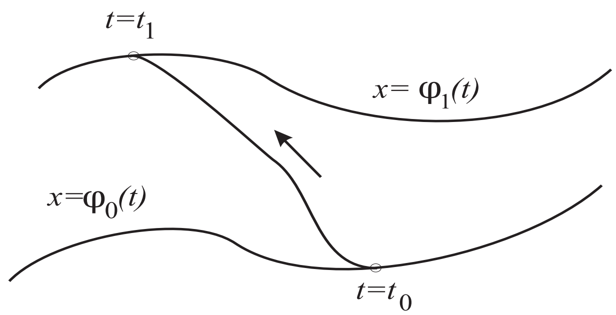

Рис. 2. Задача с подвижными концами

В общей постановке *задачи с подвижными концами* рассматривается конечное число дополнительных ограничений

$$\mathrm{I}\,[x(\cdot),t_0,t_1] = \int\limits_{t_0}^{t_1} L(t,x(t),x'(t))\,dt + \psi_0(t_0,x(t_0),t_1,x(t_1)) \to \mathrm{extr},$$

$$\psi_j(t_0,x(t_0),t_1,x(t_1)) = 0, \;\; j=1\,..\,m,$$

$$x(\cdot) \in C^1[t_0, t_1], \quad [t_0, t_1] \subset [a, b]. \tag{6}$$

Теорема 4. (необходимые условия экстремума в задаче с подвижными концами)
Пусть $(x^0(\cdot),\ t_0^0,\ t_1^0)$ *– решение задачи (6), функция* $L(t, x, y)$ *и ее частные производные* $\partial L/\partial x,\ \ \partial L/\partial y$ *непрерывны в окрестности множества*
$\{(t, x^0(t), x^{0\prime}(t) \ \ | \ \ t \in [a, b]\}$ *и функции* $\psi_j(t_0, x_0, t_1, x_1)$ *дифференцируемы в окрестности точки* $(t_0^0, x^0(t_0), t_1^0, x^0(t_1^0))$. *Тогда существуют такие ненулевые множители Лагранжа*
$\lambda = (\lambda_0, \lambda_1, \ .\ .\ , \lambda_m)$, *что для функции Лагранжа*

$$\mathcal{L}[x(\cdot), t_0, t_1, \lambda] = \lambda_0 \int\limits_{t_0}^{t_1} L(t, x(t), x'(t))\, dt + l(t_0, x(t_0), t_1, x(t_1)),$$

$$l(t_0, x_0, t_1, x_1) = \sum_{j=1}^{m} \lambda_j\, \psi_j(t_0, x_0, t_1, x_1)$$

выполнены:

1) условие стационарности по $x(\cdot)$ *– уравнение Эйлера (1)*

$$\frac{\partial L}{\partial x}(t, x^0(t), x^{0\prime}(t)) - \frac{d}{dt}\,\frac{\partial L}{\partial y}(t, x^0(t), x^{0\prime}(t)) = 0;$$

2) условия трансверсальности по $x(\cdot)$

$$\lambda_0\,\frac{\partial L}{\partial y}(t_0, x^0(t_0), x^{0\prime}(t_0)) = \frac{\partial l}{\partial x_0}(t_0^0, x^0(t_0^0), t_1^0, x^0(t_1^0)),$$

$$\lambda_0\,\frac{\partial L}{\partial y}(t_1, x^0(t_1), x^{0\prime}(t_1)) = -\frac{\partial l}{\partial x_1}(t_0^0, x^0(t_0^0), t_1^0, x^0(t_1^0));$$

3) условия стационарности по t_0, t_1

$$\frac{\partial \mathcal{L}}{\partial t_0}[x^0(\cdot), t_0^0, t_1^0, \lambda] = 0, \quad \frac{\partial \mathcal{L}}{\partial t_1}[x^0(\cdot), t_0^0, t_1^0, \lambda] = 0.$$

Идея доказательства основана на том, что необходимое условие экстремума функции нескольких переменных – равенство нулю всех ее частных производных. Для функции Лагранжа $\mathcal{L}[x(\cdot), t_0, t_1, \lambda]$ нужно приравнять нулю вариацию на функции $x(\cdot)$ и частные производные по t_0 и по t_1.

Пример.

$$\int\limits_0^T [x'^2(t) - x(t) + 1]\, dt \to \text{extr}, \quad x(0) = 0, \quad T\text{– неизвестно.}$$

Функция Лагранжа

$$\mathcal{L} = \int\limits_0^T \lambda_0 \, [x'^2(t) - x(t) + 1]\, dt + \lambda\, x(0).$$

Тогда единственная допустимая экстремаль $x^0(t) = -t^2/4 + t$ определена на отрезке $[0, 2]$.

2.5. Расширение класса искомых функций

Иногда пространство непрерывно дифференцируемых функций является слишком узким при решении задач КВИ. Экстремальная задача может иметь решение, которое не принадлежит этому пространству.

Пример Гильберта.

$$\int\limits_0^1 t^{2/3}[x'(t)]^2\, dt \to \min, \quad x(0) = 0, \quad x(1) = 1, \quad x(\cdot) \in C^1[0, 1].$$

В этой задаче уравнение Эйлера имеет решение $x(t) = c_1 t^{1/3} + c_2$, единственная допустимая экстремаль $x^0(t) = t^{1/3}$. Эта функция действительно доставляет минимум функционалу, но ее производная не является непрерывной в точке $t = 0$.

Рассмотрим еще два интересных примера.

Пример Вейерштрасса.

$$\int\limits_0^1 t^2[x'(t)]^2\,dt \to \min, \quad x(0)=0, \quad x(1)=1, \quad x(\cdot)\in C^1[0,1].$$

В этой задаче нет допустимых экстремалей, так как общее решение уравнения Эйлера $x(t)=c_1/t+c_2$. Никакое частное решение не удовлетворяет краевым условиям.

Пример.

$$\int\limits_0^{3\pi/2} (x'^2(t)-x^2(t))\,dt \to \min, \quad x(0)=x(3\pi/2)=0.$$

Общее решение уравнения Эйлера $x(t)=c_1\sin t+c_2\cos t$. Единственная допустимая экстремаль $x^0(t)\equiv 0$ не доставляет минимум. Интересно, почему?

Известны разные способы расширения класса искомых функций в задачах ВИ. Например, можно рассматривать функции, производные которых являются кусочно непрерывными. Линейное пространство таких функций обозначим $D^1[t_0,t_1]$. Очевидно, что $C^1[t_0,t_1]\subset D^1[t_0,t_1]$.

Окрестности, которые определяются с помощью нормы пространства C^1, называют *слабыми*. Экстремумы, которые определены с помощью слабых окрестностей, также называют *слабыми*. Норма

$$||x(\cdot)|| = \sup_{t\in[t_0,t_1]} |x(t)|$$

в пространстве D^1 называется *сильной*. Функция $x^0(\cdot)\in D^1[t_0,t_1]$ доставляет *сильный минимум* в задаче

$$I[x(\cdot)]\to\min, \quad x(t_0)=x_0, \quad x(t_1)=x_1,$$

если $I[x(\cdot)]\geq I[x^0(\cdot)]$ для всех функций $x(\cdot)\in D^1[t_0,t_1]$, $x(t_0)=x_0$, $x(t_1)=x_1$, принадлежащих сильной окрестности функции $x^0(\cdot)$.

Можно доказать, что если функция $x^0(\cdot) \in C^1[t_0, t_1]$ доставляет сильный локальный экстремум в экстремальной задаче, то она доставляет и слабый экстремум.

Теорема 5. (необходимое условие Вейерштрасса)
Если функция $x^0(\cdot) \in C^1[t_0, t_1]$ доставляет сильный минимум в задаче (1), то

$$L(\tau, x^0(\tau), x^{0\prime}(\tau) + \xi) - L(\tau, x^0(\tau), x^{0\prime}(\tau)) - \xi \frac{\partial L}{\partial y}(\tau, x^0(\tau), x^{0\prime}(\tau)) \geq 0$$

$$\forall \tau \in [t_0, t_1], \;\; \forall \xi \in R^1. \tag{7}$$

Функция, стоящая слева в неравенстве (7), называется функцией Вейерштрасса.

Полное доказательство теоремы имеется в книге [1], 1.4.4. Вводятся так называемые *игольчатые вариации*, то есть функции вида

$$h_\lambda(t) = \begin{cases} \xi\lambda + (t-\tau)\xi, & t \in [\tau - \lambda, \tau], \\ \xi\lambda - (t-\tau)\xi\sqrt{\lambda}, & t \in [\tau, \tau + \sqrt{\lambda}], \\ 0, & t \notin [\tau - \lambda, \tau + \sqrt{\lambda}], \end{cases}$$

$\lambda \geq 0, \quad \xi \in R^1, \quad \tau \in [t_0, t_1]$, и функция $\varphi(\lambda) = I[x(\cdot) + h_\lambda(\cdot)]$. Доказывается, что

$$\varphi(\lambda) - \varphi(0) = \lambda \left(L(\tau, x^0(\tau), x^{0\prime}(\tau) + \xi) - L(\tau, x^0(\tau), x^{0\prime}(\tau)) \right) =$$

$$-\xi \frac{\partial L}{\partial y}(\tau, x^0(\tau), x^{0\prime}(\tau)) + o(\lambda).$$

Тогда $\varphi'(0+0) \geq 0$. $\bullet$

В следующем разделе будет показано, как можно в задачах ВИ перейти к еще более широким пространствам.

* * *

Леонард Эйлер получил уравнение (1) другим способом. Он искал решение простейшей задачи как ломаную линию. Дифференциальное уравнение было получено предельным переходом из условий экстремума функции нескольких переменных (см. [9], 3.1).

В классическом вариационном исчислении рассматриваются также интегральные функционалы, подинтегральные выражения которых зависят от старших производных функции $x(\cdot)$; искомые функции могут зависеть от нескольких переменных, тогда уравнение Эйлера становится уравнением с частными производными; получены другие формы необходимых условий экстремума, а также достаточные условия экстремума, и многое другое.

В качестве дополнительной литературы рекомендуем книги [10], [12].

3. ИНТЕГРАЛЬНОЕ ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ

Если в задачах КВИ заменить операторы дифференцирования на интегральные операторы, то необходимые условия экстремума примут форму интегральных уравнений. Ниже будет показано, что задачи классического ВИ сводятся заменой искомой функции к задачам ИВИ с дополнительным изопериметрическим условием.

3.1. Простейшая задача ИВИ

Рассмотрим *интегральный функционал*

$$\mathrm{I}\,[x(\cdot)] = \int\limits_a^b f(t, x(t), (\mathrm{K}\,x)(t))\,dt,$$

где

$$(\mathrm{K}\,x)(t) = \int\limits_a^b k(\tau, t, x(\tau))\,d\tau, \quad t \in [a, b],$$

– *интегральный оператор Урысона.* Будем предполагать, что заданные функции $f(t,x,y)$, $k(\tau,t,x)$ непрерывны по совокупности переменных. Будем искать экстремумы функционала в пространстве непрерывных функций $C[a,b]$ с нормой

$$||x(\cdot)|| = \max_{t\in[a,b]} |x(t)|.$$

Назовем *простейшей задачей ИВИ* экстремальную задачу

$$\mathrm{I}\,[x(\cdot)] \to \mathrm{extr}, \quad x(\cdot) \in C[a,b]. \tag{1}$$

Лемма 1. (основная лемма ИВИ)
Пусть $a(\cdot) \in C[a,b]$. Если

$$\int\limits_a^b a(t)\,h(t)\,dt = 0 \quad \forall h(\cdot) \in C[a,b],$$

то $a(t) \equiv 0$ на $[a,b]$.

Это утверждение доказывается примерно тем же способом, что и лемма Лагранжа КВИ. Доказательство можно найти, например, в книге [11], с.232. •

Аналогом леммы Дюбуа-Реймона является

Лемма 2. *Пусть $c(\cdot), d(\cdot) \in C[a,b]$, $k(\cdot,\cdot) \in C([a,b]\times[a,b])$. Если*

$$\int\limits_a^b [c(t)h(t) + d(t)\int\limits_a^b h(\tau)k(\tau,t)\,d\tau]\,dt = 0 \quad \forall h(\cdot) \in C[a,b],$$

то

$$c(t) + \int\limits_a^b d(\tau)k(t,\tau)\,d\tau \equiv 0 \;\; \textit{на;}\;\; [a,b].$$

Чтобы это доказать, нужно изменить порядок интегрирования в левой части условия и применить лемму 1. •

Теорема 1. (необходимое условие экстремума в простейшей задаче ИВИ)
Пусть функции $k(\tau, t, x)$, $\dfrac{\partial k}{\partial x}$, $f(t, x, y)$, $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial f}{\partial y}$ *непрерывны по совокупности переменных. Если функция* $x^0(\cdot)$ *– решение задачи (1), то она удовлетворяет интегральному уравнению*

$$\frac{\partial f}{\partial x}(t, x(t), (\mathrm{K}\, x)(t)) + \int\limits_a^b \frac{\partial k}{\partial x}(t, \tau, x(t)) \, \frac{\partial f}{\partial y}(\tau, x(\tau), (\mathrm{K}\, x)(\tau)) \, d\tau = 0. \tag{2}$$

Доказательство. Вычислим первую вариацию функционала

$$\delta \mathrm{I}\, [x(\cdot); h(\cdot)] = \int\limits_a^b \frac{\partial f}{\partial x}(t, x(t), (\mathrm{K}\, x)(t)) \, h(t) \, dt +$$

$$+ \int\limits_a^b \frac{\partial f}{\partial y}(t, x(t), (\mathrm{K}\, x)(t)) \int\limits_a^b \frac{\partial k}{\partial x}(\tau, t, x(\tau)) \, h(\tau) \, d\tau \, dt.$$

Изменим порядок интегрирования во втором слагаемом и переобозначим переменные интегрирования. Тогда

$$\delta \mathrm{I}\, [x(\cdot); h(\cdot)] = \int\limits_a^b [\frac{\partial f}{\partial x}(t, x(t), (\mathrm{K}\, x)(t)) +$$

$$+ \int\limits_a^b \frac{\partial k}{\partial x}(t, \tau, x(t)) \, \frac{\partial f}{\partial y}(\tau, x(\tau), (\mathrm{K}\, x)(\tau)) \, d\tau] h(t) \, dt.$$

Приравняем нулю вариацию и применим лемму 1. (Лемму 2 можно было бы использовать еще раньше, до перестановки интегралов). •

Замечание. Как и ожидалось, аналогом уравнения Эйлера в простейшей задаче интегрального вариационного исчисления является интегральное уравнение. В отличие от классического вариационного исчисления, при выводе необходимого условия экстремума в простейшей задаче вместо интегрирования по частям используется перестановка интегралов в повторном интеграле. В постановке простейшей задачи ИВИ нет дополнительных условий вида $x(t_0) = x_0,\ x(t_1) = x_1$.

Утверждение теоремы 1 остается в силе для таких интегральных функционалов и для таких классов функций, когда 1) существует вариация функционала (1); 2) можно изменить порядок интегрирования в повторном интеграле; 3) сохраняется утверждение аналога основной леммы вариационного исчисления.

Следствие 1. *Если*

$$(\mathrm{K}\,x)(t) = \int\limits_a^b x(\tau)\,k(\tau,t)\,d\tau, \quad t \in [a,b],$$

то аналог уравнения Эйлера (2) принимает вид

$$\frac{\partial f}{\partial x}(t,x(t),(\mathrm{K}\,x)(t)) + \int\limits_a^b \frac{\partial f}{\partial y}(\tau,x(\tau),(\mathrm{K}\,x)(\tau))\,k(t,\tau)d\tau = 0, \quad t \in [a,b].$$

Замечание. Если в интегральном операторе K использовать интеграл с переменным верхним пределом (то есть $k(\tau,t) = 0$ при $\tau > t$), то в аналоге уравнения Эйлера будет стоять интеграл с переменным нижним пределом.

Следствие 2. *Если*

$$f(t,x,y) = A(t)x^2 + 2B(t)xy + C(t)y^2 - 2D(t)x - 2E(t)y,$$

то уравнение (2) становится линейным интегральным уравнением с симметричным ядром

$$A(t)x(t) + \int\limits_a^b x(\tau)\,L(\tau,t)\,d\tau = g(t), \quad t \in [a,b], \tag{3}$$

где

$$L(\tau,t) = B(\tau)\,k(t,\tau) + B(t)\,k(\tau,t) + \int\limits_a^b C(\xi)\,k(\tau,\xi)\,k(t,\xi)\,d\xi,$$

$$g(t) = D(t) + \int\limits_a^b E(\tau)\, k(t,\tau)\, d\tau.$$

В этом частном случае легко получить следующее утверждение.

Теорема 2. (достаточное условие экстремума) *Пусть функционал в простейшей задаче ИВИ квадратичный и интегральный оператор линейный. Если*

$$\int\limits_a^b [A(t)h^2(t) + 2B(t)h(t)(\mathrm{K}\,h)(t) + C(t)(\mathrm{K}\,h)^2(t)]\, dt \ge 0 \ \forall h(\cdot) \in C[a,b],$$

то любое решение интегрального уравнения (3) доставляет минимум в простейшей задаче.

Действительно,

$$\mathrm{I}\,[x(\cdot) + h(\cdot)] - \mathrm{I}\,[x(\cdot)] = 2\int\limits_a^b h(t)\Big[A(t)\, x(t) + \int\limits_a^b x(\tau)\, L(\tau,t)\, d\tau - g(t)\Big]\, dt +$$

$$+ \int\limits_a^b [A(t)\, h^2(t) + 2B(t)\, h(t)\, (\mathrm{K}\,h)(t) + C(t)\, (\mathrm{K}\,h)^2(t)]\, dt. \bullet$$

Замечание. Если интегральный оператор K имеет *вырожденное ядро*, то и интегральное уравнение (3) имеет вырожденное ядро. Напомним, что вырожденным ядром интегрального уравнения называют функцию вида

$$k(\tau,t) = \sum_{j=1}^m p_j(\tau)\, q_j(t).$$

Интегральное уравнение 2-го рода с вырожденным ядром легко сводится к системе линейных алгебраических уравнений.

3.2. Интегральные уравнения Фредгольма

Интегральным уравнением Фредгольма 2-го рода на отрезке называют уравнение

$$(\mathrm{M}\, x) = A(t)\, x(t) + \int\limits_a^b x(\tau)\, k(\tau, t)\, d\tau = f(t), \quad t \in [a, b].$$

Будем предполагать, что все рассматриваемые функции являются непрерывными. Интегральное уравнение

$$(\mathrm{M}\,'y) = A(t)\, y(t) + \int\limits_a^b y(\tau)\, k(t, \tau)\, d\tau = 0, \quad t \in [a, b],$$

называют уравнением *с транспонированным ядром* (или *союзным* с однородным уравнением Фредгольма).

Рассмотрим экстремальную задачу (это **пример**)

$$\mathrm{I}\,[x(\cdot)] = \int\limits_a^b [A(t)\, x(t) + \int\limits_a^b x(\tau)\, k(\tau, t)\, d\tau - f(t)]^2\, dt \to \min, \quad x(\cdot) \in C.$$

Эту задачу можно понимать так: найти функцию $x(\cdot)$, при которой левая часть интегрального уравнения $(\mathrm{K}\, x)(t) = f(t)$ наиболее близка (в смысле квадратического среднего) к его правой части. Условие теоремы 2 в данном случае выполняется. Поэтому функция $x(\cdot)$ является решением задачи тогда и только тогда, когда

$$A(t)\, [A(t)\, x(t) + \int\limits_a^b x(\tau)\, k(\tau, t)\, d\tau - f(t)] +$$

$$+ \int\limits_a^b [A(\xi)\, x(\xi) + \int\limits_a^b x(\tau)\, k(\tau, \xi)\, d\tau - f(\xi)]\, k(t, \xi)\, d\xi = 0$$

или

$$\mathrm{M}\,'((\mathrm{M}\, x)(t) - f(t)) = 0.$$

Так как функционал выпуклый, то решение $x^0(\cdot)$ экстремальной задачи существует. Следовательно, или $\mathrm{M}\, x^0 = f$, или функция

$y = \mathrm{M}\, x^0 - f$ является нетривиальным решением транспонированного однородного уравнения $\mathrm{M}\,' y = 0$. Это – *первая теорема Фредгольма* (альтернатива Фредгольма).

Рассмотрим подробнее случай, когда транспонированное однородное уравнение $\mathrm{M}\,' y = 0$ имеет ненулевое решение $y(\cdot)$. Тогда, как было установлено выше, $\mathrm{M}\, x^0 = f + y$. Это уравнение разрешимо тогда и только тогда, когда

$$\int_a^b [f(t) + y(t)]\, y_j(t)\, dt = 0, \quad j = 1..n,$$

где $y_1(\cdot), \ldots, y_n(\cdot)$ – полная система независимых решений однородного транспонированного уравнения (*третья теорема Фредгольма*). Будем считать, что эта система решений ортономирована. Поэтому $y(\cdot) = c_1 y_1(\cdot) + \ldots + c_n y_n(\cdot)$, где

$$c_j = \int_a^b y(t)\, y_j(t)\, dt, \quad j = 1..n.$$

Тогда

$$c_j = -\int_a^b f(t)\, y_j(t)\, dt, \quad j = 1..n.$$

Следовательно, минимальное значение функционала

$$\min \mathrm{I} = \mathrm{I}\,[x^0(\cdot)] = \int_a^b y^2(t)\, dt = \sum_{j=1}^n c_j^2 = \sum_{j=1}^n (\int_a^b f(t)\, y_j(t)\, dt)^2.$$

Таким образом, значения интегралов

$$\int_a^b f(t)\, y_j(t)\, dt$$

не только определяют, разрешимо или нет неоднородное линейное интегральное уравнение при заданной правой части, но и показывают, насколько можно приблизить левую часть уравнения к правой части, если уравнение решений не имеет.

3.3. Изопериметрическая задача

Пусть

$$\mathrm{I}_j[x(\cdot)] = \int\limits_a^b f_j(t, x(t), (K_j x)(t))\, dt, \quad j = 0..m.$$

Рассмотрим изопериметрическую задачу

$$\mathrm{I}_0[x(\cdot)] \to \mathrm{extr}, \quad \mathrm{I}_j[x(\cdot)] = c_j,\ j = 1..m, \quad x(\cdot) \in C[a, b]. \qquad (4)$$

Теорема 3. (правило множителей Лагранжа)
Пусть все заданные функции и их частные производные непрерывны. Если $x^0(\cdot)$ – решение задачи (4), то найдутся такие числа $\lambda_0, \lambda_1, \ldots, \lambda_m$ (не равные нулю одновременно), что $x^0(\cdot)$ удовлетворяет интегральному уравнению (2), составленному для функции

$$f(t, x, y) = \sum_{j=0}^{m} \lambda_j\, f_j(t, x, y).$$

Если функционалы $I_j[\cdot]$, $j = 1..m$ не имеют стационарных точек, то можно считать, что $\lambda_0 = 1$.

Доказательство. Утверждение теоремы следует из общего правила множителей Лагранжа (см. п.2.3). •

Пример. Распределение нагрузки вдоль струны.

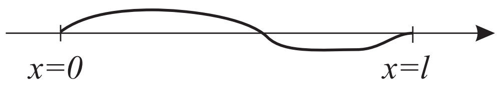

Рис. 1. Упругая струна

Вернемся к задаче из п.1.2. Пусть концы упругой струны закреплены в точках $x = 0$ и $x = l$. Если вдоль струны распределена сила $p(\cdot)$, то отклонение струны от положения равновесия (рассматриваются только малые отклонения)

$$u(x) = \int\limits_0^l G(x, \xi)\, p(\xi)\, d\xi,$$

$$G(x,\xi)=\left\{\frac{x(l-\xi)}{T_0 l},\ \ 0\le x\le\xi,\quad \frac{(l-x)\xi}{T_0 l},\ \ \xi\le x\le l\right\},$$

T_0 – натяжение ненагруженной струны.

Будем искать такое распределение силы $p(\cdot)$, при котором струна примет форму $u(\cdot)$, наиболее близкую к заданной $u_0(\cdot)$, при условии, что суммарная нагрузка на струну постоянна:

$$\int_0^l [u(x)-u_0(x)]^2\,dx\to\min,\qquad \int_0^l p(\xi)\,d\xi=P.$$

Если подставим выражение $u(\cdot)$ через $p(\cdot)$ в функционал, то получим изопериметрическую задачу ИВИ. В этой задаче функционал $\int_0^l p(\xi)\,d\xi$ не имеет стационарных точек.

В соответствии с утверждением теоремы 3, необходимые и достаточные условия экстремума имеют вид

$$\frac{\lambda}{2}+\int_0^l p(\xi)H(\xi,x)\,d\xi=\int_0^l u_0(\xi)G(x,\xi)\,d\xi,\quad \int_0^l p(\xi)\,d\xi=P,$$

где λ – множитель Лагранжа,

$$H(\xi,x)=\int_0^l G(x,t)\,G(t,\xi)\,dt,\quad \xi,x\in[0,l].$$

Интегральное уравнение Фредгольма 1-го рода вместе с дополнительным условием могут быть решены численно методом регуляризации А.Н.Тихонова.

Пример. Распределение нагрузки вдоль мембраны.

Пусть упругая мембрана закреплена вдоль сторон прямоугольника $[0,a]\times[0,b]$. Ее отклонение от положения равновесия под действием распределенной силы $p(\cdot,\cdot)$ является решением задачи

$$\frac{\partial^4 u}{\partial x^4}+2\frac{\partial^4 u}{\partial x^2\partial y^2}+\frac{\partial^4 u}{\partial y^4}=p(x,y),$$

$$u=0,\ \ \frac{\partial^2 u}{\partial x^2}=0 \ \text{ при } \ x=0,\ \ x=a, \quad u=0,$$
$$\frac{\partial^2 u}{\partial y^2}=0 \ \text{ при } \ y=0,\ \ y=b.$$

Это решение может быть записано в виде

$$u(x,y)=\frac{4}{\pi^4 ab}\int\limits_0^a\int\limits_0^b G(\xi,\eta,x,y)\,p(\xi,\eta)\,d\xi\,d\eta,$$

$$G(\xi,\eta,x,y)==\Big[\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{\sin\dfrac{m\pi x}{a}\ \sin\dfrac{n\pi y}{b}\ \sin\dfrac{m\pi\xi}{a}\ \sin\dfrac{n\pi\eta}{b}}{(\dfrac{m^2}{a^2}+\dfrac{n^2}{b^2})^2}\Big]$$

(функция Грина определяет прогиб мембраны под действием сосредоточенной единичной силы, приложенной в точке (ξ,η)). Тогда исходная задача может быть сформулирована как изопериметрическая задача ИВИ

$$\int\limits_0^a\int\limits_0^b[u(x,y)-u_0(x,y)]^2\,dx\,dy\to\min,\qquad \int\limits_0^a\int\limits_0^b p(x,y)\,dx\,dy=P.$$

3.4. Классы искомых функций

Исследуем связь между задачами КВИ и задачами ИВИ.

Рассмотрим простейшую задачу КВИ

$$\mathrm{I}\,[x(\cdot)]\equiv\int\limits_{t_0}^{t_1}L(t,x(t),x'(t))\,dt\to\mathrm{extr},$$

$$x(t_0)=x_0,\ \ x(t_1)=x_1,\quad x(\cdot)\in C^1[t_0,t_1].$$

Введем новую искомую функцию $y(\cdot)=x'(\cdot)$, тогда

$$x(t)=\int\limits_{t_0}^{t}y(\tau)\,d\tau+x_0,\quad t\in[t_0,t_1]. \tag{5}$$

Простейшая задача КВИ равносильна изопериметрической задаче ИВИ

$$\int\limits_{t_0}^{t_1} L(t, \int\limits_{t_0}^{t} y(\tau)\, d\tau + x_0, y(t))\, dt \to \mathrm{extr},$$

$$\int\limits_{t_0}^{t_1} y(\tau)\, d\tau = x_1 - x_0, \quad y(\cdot) \in C[t_0, t_1].$$

Второй функционал не имеет стационарных точек, поэтому можно выбрать множитель Лагранжа $\lambda_0 = 1$. Тогда необходимое условие экстремума

$$\frac{\partial L}{\partial y}(t, \int\limits_{t_0}^{t} y(\tau)\, d\tau + x_0, y(t)) + \int\limits_{t}^{t_1} \frac{\partial L}{\partial x}(\tau, \int\limits_{t_0}^{\tau} y(\xi)\, d\xi + x_0, y(\tau))\, d\tau + \lambda = 0.$$

При $t = t_1$ найдем значение

$$\lambda = -\frac{\partial L}{\partial y}(t_1, \int\limits_{t_0}^{t_1} y(\tau)\, d\tau + x_0, y(t_1)) = -\frac{\partial L}{\partial y}(t_1, x_1, y(t_1)).$$

Вернемся к старой искомой функции. Получим необходимое условие экстремума (интегральный аналог уравнения Эйлера)

$$\frac{\partial L}{\partial y}(t, x(t), x'(t)) - \frac{\partial L}{\partial y}(t_1, x_1, x'(t_1)) + \int\limits_{t}^{t_1} \frac{\partial L}{\partial x}(\tau, x(\tau), x'(\tau\, d\tau = 0.$$

Здесь достаточно считать, что функция $L(t, x, y)$ имеет непрерывные первые частные производные.

Если потребовать дополнительно, что функция $L(t, x, y)$ имеет непрерывные вторые частные производные, то это уравнение можно переписать в виде

$$\int\limits_{t}^{t_1} [\frac{\partial L}{\partial x}(\tau, x(\tau), x'(\tau)) - \frac{d}{d\tau}\frac{\partial L}{\partial y}(\tau, x(\tau), x'(\tau))]\, d\tau = 0.$$

Продифференцируем по t и получим получить классическое уравнение Эйлера – необходимое условие экстремума в простейшей задаче КВИ.

Таким образом, любая задача КВИ приводится к задаче ИВИ с дополнительным изопериметрическим условием.

Уточним, в каких классах можно рассматривать искомые функции.

Пусть равенство (5) устанавливает соответствие между функциями $x(\cdot)$ из некоторого пространства X и функциями $y(\cdot)$ из некоторого пространства Y. Исследуем, как согласованы друг с другом различные способы задания окрестностей элементов в X и Y.

А. Пусть $X = C^1[t_0, t_1]$ и окрестности в пространстве определяются с помощью нормы

$$||x(\cdot)||_{C^1} = \max\{\max_{t\in[t_0,t_1]} |x(t)|, \max_{t\in[t_0,t_1]} |x'(t)|\}$$

(слабые окрестности на языке КВИ). В этом случае $Y = C[t_0, t_1]$ и стандартная норма

$$||y(\cdot)||_C = \max_{t\in[t_0,t_1]} |y(t)|$$

задает в пространстве Y окрестности, согласованные с окрестностями в X. Это следует из неравенств:

$$||y(\cdot) - y^0(\cdot)||_C \leq ||x(\cdot) - x^0(\cdot)||_{C^1},$$

$$||x(\cdot)-x^0(\cdot)||_{C^1} = \max\{\max_{t\in[t_0,t_1]} |\int\limits_{t_0}^{t} [y(\tau)-y^0(\tau)]d\tau|, \max_{t\in[t_0,t_1]} |y(t)-y^0(t)|\} \leq$$

$$\leq \max(t_1 - t_0, 1)\ ||y(\cdot) - y^0(\cdot)||_C.$$

Если же окрестности в пространстве $X = C^1[t_0, t_1]$ задает неравенство $||x(\cdot)-x_0(\cdot)||_C < \varepsilon$ (сильные окрестности), то можно определить окрестности в пространстве $Y = C[t_0, t_1]$ так:

$||y(\cdot) - y^0(\cdot)||_C < \varepsilon$. В этом случае если $y(\cdot)$ – решение экстремальной задачи в Y, то $x(\cdot)$ – решение экстремальной задачи в X, но не наоборот.

Б. Аналогичная ситуация имеет место, если $X = AC[t_0, t_1]$ (пространство абсолютно непрерывных функций) и $Y = L_1(t_0, t_1)$. Пусть

$$||y(\cdot) - y^0(\cdot)||_{L_1} = \int\limits_{t_0}^{t_1} |y(t) - y^0(t)|\, dt,$$

$$||x(\cdot) - x^0(\cdot)||_{AC} = |x(t_0) - x^0(t_0)| + ||x'(\cdot) - (x^0)'(\cdot)||_{L_1}.$$

Определенные таким способом окрестности согласованы, так как

$$||y(\cdot) - y^0(\cdot)||_{L_1} = ||x'(\cdot) - (x^0)'(\cdot)||_{L_1} \le ||x(\cdot) - x^0(\cdot)||_{AC},$$

$$||x(\cdot) - x^0(\cdot)||_{AC} = ||y(\cdot) - y^0(\cdot)||_{L_1}.$$

Еще раз рассмотрим **пример Гильберта**

$$\int\limits_0^1 t^{2/3}[x'(t)]^2\, dt \to \min, \quad x(0) = 0, \quad x(1) = 1,$$

с помощью которого было установлено, что в задачах КВИ пространство C^1 является слишком узким.

Введем новую искомую функцию $y(t) = x'(t)$ и будем искать решения изопериметрической задачи ИВИ в пространстве AC. Необходимое условие экстремума

$$2t^{2/3}y(t) + \lambda = 0,$$

из изопериметрического условия находим $\lambda = -2/3$ и тогда
$y(t) = \dfrac{1}{3}t^{-2/3}$. Следовательно, $x(t) = t^{1/3}$ – функция из L_1.

3.5. Аналог задачи Больца

Добавим к интегральному функционалу простейшей задачи (1) *терминальную часть*

$$\mathrm{T}\,[x(\cdot)] = F(x(t_1), ..., x(t_n), \int\limits_a^b k_1(\tau, x(\tau))\, d\tau, ..., \int\limits_a^b k_m(\tau, x(\tau))\, d\tau),$$

где $t_1, \ldots, t_n$ – фиксированные точки на отрезке $[a, b]$ и $k_1(\tau, x), \ldots$ $k_m(\tau, x)$ – заданные непрерывные функции. Аналогом *задачи Больца* КВИ является следующая задача:

$$\mathrm{B}\,[x(\cdot)] = \mathrm{I}\,[x(\cdot)] + \mathrm{T}\,[x(\cdot)] \to \mathrm{extr}, \quad x(\cdot) \in C[a, b]. \tag{6}$$

Теорема 3. (необходимое условие экстремума для аналога задачи Больца)

Пусть функции f, k, F, k_j, $j = 1..m$ *и их частные производные* $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial f}{\partial y}$, $\dfrac{\partial k}{\partial x}$, $\dfrac{\partial F}{\partial x_j}$, $j = 1..n$, $\dfrac{\partial F}{\partial y_j}$, $j = 1..m$ *непрерывны по совокупности переменных. Если функция* $x^0(\cdot)$ – *решение задачи (6), то она удовлетворяет интегральному уравнению*

$$\frac{\partial f}{\partial x}(t) + \int\limits_a^b \frac{\partial f}{\partial y}(s)\frac{\partial k}{\partial x}(t, s, x(t))\, ds + \sum_{l=1}^{m} \frac{\partial F}{\partial y_l}\,\frac{\partial k_l}{\partial x}(t, x(t)) = 0 \tag{7}$$

и выполняются условия

$$\frac{\partial F}{\partial x_j} = 0, \quad j = 1..n, \tag{8}$$

здесь для краткости записи обозначено

$$\frac{\partial f}{\partial \xi}(t) = \frac{\partial f}{\partial \xi}(t, x(t), \int\limits_a^b k(\tau, t, x(\tau))\, d\tau), \quad \xi = x \;\; \textit{или} \;\; y,$$

$$\frac{\partial F}{\partial \eta} = \frac{\partial F}{\partial \xi}(x(t_1), ..., x(t_n), \int\limits_a^b k_1(\tau, x(\tau))\, d\tau, ..., \int\limits_a^b k_m(\tau, x(\tau))d\tau) = 0,$$

$$\eta = x_j,\ j = 1..n,\ \ \text{или}\ y_l,\ l = 1..m.$$

Доказательство. Вычислим вариацию функционала в задаче (6) и приравняем ее нулю. Лемма 1 допускает следующую модификацию.

Лемма 1′.
Пусть $a(\cdot) \in C[a,b]$. *Если*

$$\int\limits_a^b a(t)\,h(t)\,dt = 0 \quad \forall h(\cdot) \in C[a,b]\ \mid\ h(t_j) = 0,\ j = 1..n,$$

то $a(t) \equiv 0$.

Отсюда следует интегральное уравнение (7). Если же рассматривать такие функции $h(\cdot)$, что $h(t_j) \neq 0$ только для одной из точек t_j, то получим условия (8). •

Рассмотрим два частных случая аналога задачи Больца.

А. Если терминальная часть функционала Больца не зависит от переменных $y_j,\ j = 1..m$, то необходимые условия экстремума (7) и (8) становятся независимыми.

Следствие 3. (первый частный случай)
Пусть выполнены условия теоремы 3. Непрерывная функция $x(\cdot)$ *доставляет локальный минимум (максимум) функционалу*

$$B(x(\cdot)) = I(x(\cdot)) + F(x(t_1), ..., x(t_n))$$

тогда и только тогда, когда
1) функция $x(\cdot)$ *доставляет локальный минимум (максимум) функционалу* $I(x(\cdot))$ *и*
2) вектор $(x(t_1), ..., x(t_n))$ *доставляет локальный минимум (максимум) функции* $F(x_1, ..., x_n)$.

Б. Пусть терминальная часть функционала Больца не зависит от переменных $x_j,\ j = 1..n$.

Следствие 4. (второй частный случай)
Пусть выполнены условия теоремы 3 и, кроме того, $\dfrac{\partial F}{\partial x_j} = 0 \ \forall j = 1..n$. *если функция* $x^0(\cdot)$ *– решение задачи (6), то она удовлетворяет интегральному уравнению*

$$\frac{\partial f}{\partial x}(t) + \int\limits_a^b \frac{\partial f}{\partial y}(s)\frac{\partial k}{\partial x}(t, s, x(t)\, ds + \sum_{j=1}^m \frac{\partial F}{\partial y_j}\frac{\partial k_j}{\partial x}(t, x(t)) = 0.$$

Замечание. Во втором частном случае в интегральном уравнении не содержатся значения $x(t_j)$, $j = 1..n$. Поэтому можно расширить класс искомых функций до пространства $L(a, b)$.

3.6. Задача с подвижными концами

Рассмотрим задачу с подвижными концами КВИ

$$\int\limits_{t_0}^{t_1} L(t, x(t), x'(t))\, dt \to \mathrm{extr},$$

$$x(t_0) = \varphi_0(t_0), \quad x(t_1) = \varphi_1(t_1), \quad x(\cdot) \in C^1[t_0, t_1], \tag{9}$$

здесь значения t_0, t_1 неизвестны (предполагаем, что эти точки принадлежат некоторому достаточно широкому интервалу $\Delta \subset R^1$). Функции $\varphi_0(\cdot)$ и $\varphi_1(\cdot)$ заданы. Смысл дополнительных ограничений в том, что концы $x(t_0)$ и $x(t_1)$ искомой траектории движутся по линиям $x = \varphi_0(t)$ и $x = \varphi_1(t)$.

Перейдем к задаче ИВИ

$$\int\limits_{t_0}^{t_1} L(t, \int\limits_{t_0}^{t} y(\tau)\, d\tau + \varphi_0(t_0), y(t))\, dt \to \mathrm{extr},$$

$$\int\limits_{t_0}^{t_1} y(\tau)\, d\tau = \varphi_1(t_1) - \varphi_0(t_0), \ \ y(\cdot) \in C[t_0, t_1].$$

Составим функцию Лагранжа (второй функционал не имеет стационарных точек)

$$\mathcal{L} = \int\limits_{t_0}^{t_1} L(t, \int\limits_{t_0}^{t} y(\tau)\, d\tau + \varphi_0(t_0), y(t))\, dt + \lambda[\int\limits_{t_0}^{t_1} y(\tau)\, d\tau - \varphi_1(t_1) + \varphi_0(t_0)],$$

искомыми величинами являются $y(\cdot),\ t_0,\ t_1$. Приравняем нулю производные по всем этим переменным.

1). Если $\delta\mathcal{L}[y(\cdot); h(\cdot)] = 0$, то

$$\frac{\partial L}{\partial y}(t, \ldots) + \int\limits_{t}^{t_1} \frac{\partial L}{\partial x}(\tau, \ldots)\, d\tau + \lambda = 0,$$

отсюда при $t = t_1$ получим

$$\lambda = -\frac{\partial L}{\partial y}(t_1, \ldots).$$

Следовательно,

$$\frac{\partial L}{\partial y}(t, \int\limits_{t_0}^{t} y(\tau)\, d\tau + \varphi_0(t_0), y(t)) - \frac{\partial L}{\partial y}(t_1, \varphi_1(t_1), y(t_1)) +$$

$$+ \int\limits_{t}^{t_1} \frac{\partial L}{\partial x}(\tau, \ldots)\, d\tau = 0. \qquad (10)$$

2). Если $\dfrac{\partial \mathcal{L}}{\partial t_0} = 0$, то

$$\int\limits_{t_0}^{t_1} \frac{\partial \mathrm{L}}{\partial x}(t, \ldots)[\varphi_0'(t_0) - y(t_0)]\, dt -$$

$$-L(t_0, \varphi_0(t_0), y(t_0)) - \frac{\partial L}{\partial y}(t_1, \ldots)[\varphi_0'(t_0) - y(t_0)] = 0.$$

Запишем равенство (10) при $t = t_0$

$$\frac{\partial L}{\partial y}(t_0, \varphi_0(t_0), y(t_0)) - \frac{\partial L}{\partial y}(t_1, \varphi_1(t_1), y(t_1)) +$$

$$+\int\limits_{t_0}^{t_1} \frac{\partial L}{\partial x}(\tau, \int\limits_{t_0}^{\tau} y(\xi)\, d\xi + \varphi_0(t_0), y(\tau))\, d\tau = 0.$$

Отсюда следует, что

$$L(t_0, \varphi_0(t_0), y(t_0)) + [\varphi_0'(t_0) - y(t_0)] \frac{\partial L}{\partial y}(t_0, \varphi_0(t_0), y(t_0)) = 0. \quad (11)$$

3). Если $\dfrac{\partial \mathrm{L}}{\partial t_1} = 0$, то с помощью таких же рассуждений получим

$$L(t_1, \varphi_1(t_1), y(t_1)) + [\varphi_1'(t_1) - y(t_1)] \frac{\partial L}{\partial y}(t_1, \varphi_1(t_1), y(t_1)) = 0. \quad (12)$$

Теорема 4. (необходимые условия экстремума в задаче с подвижными концами)
Если $y(\cdot)$, t_0, t_1 – решение задачи (9), то выполняются условия (10), (11) и (12).

* * *

Этот раздел написан по работам [13], [14]. В [15] установлено, когда условия экстремума интегрального функционала общего вида имеют форму интегрального уравнения.

4. ПРОИЗВОДНЫЕ ОТОБРАЖЕНИЙ И УСЛОВИЯ ЭКСТРЕМУМА

Вариационное исчисление основано на том факте, что если функция доставляет экстремум функционалу, то вариация функционала обращается в нуль на этой функции. В современном функциональном анализе построена более совершенная теория дифференцирования отображений в нормированных пространствах. Автор пособия надеется, что читателю знакомы понятия: линейное пространство, нормированное пространство, отображение, линейный оператор, ограниченный оператор, непрерывный оператор...

4.1. Производные отображений

Пусть X, Y – линейные нормированные пространства, $x^0 \in X$, U – окрестность x^0 в X, отображение $f\,:\,U \to Y,\quad U \subset D(f)$.

1. Пусть $h \in X$. Если существует предел

$$\lim_{\lambda \downarrow 0} \frac{f(x^0 + \lambda\, h) - f(x^0)}{\lambda},$$

то он называется *производной отображения* f *на элементе* x^0 *по направлению* h и обозначается $f'(x^0;h)$ (здесь $f'(x^0;h)$ – элемент пространства Y).

2. Если $\forall h \in X$ существует $f'(x^0;h)$, то отображение $\delta f(x^0;\cdot)\ :\ h \in X \mapsto f'(x^0;h) \in Y$ называется *первой вариацией отображения* f *на элементе* x^0, а при $f'(x^0;-h) = -f'(x^0;h)$ – первой вариацией по Лагранжу.

3. Если A – такое линейное непрерывное отображение (оператор), что

$$f(x^0 + \lambda h) = f(x^0) + \lambda A h + o(\lambda) \ \text{ при } \ \lambda \to 0 \ \ \forall h \in X,$$

то A – *производная Гато* (или *слабая производная*) *отображения* f *на элементе* x^0; обозначается $f'_G(x^0)$.

Это определение равносильно тому, что первая вариация $\delta f(x^0;\cdot)$ является линейным и непрерывным отображением.

Здесь, как обычно, $o(\lambda)$ такая функция со значениями в Y, что $o(\lambda)/\lambda \to 0$ при $\lambda \to 0$.

4. Если $A\,:\,X \to Y$ – такое линейное непрерывное отображение, что

$$f(x^0 + h) = f(x^0) + Ah + o(||h||) \quad \text{при} \quad ||h|| \to 0,$$

то A – *производная Фреше* (или *сильная производная*) *отображения* f *на элементе* x^0; обозначается $f'(x^0)$.

Отображение дифференцируемо по Фреше тогда и только тогда, когда

$$\forall \varepsilon > 0 \;\; \exists \delta(\varepsilon) > 0 \;\; | \;\; ||f(x^0 + h) - f(x^0) - f'(x^0)h|| < \varepsilon ||h||$$

$$\text{при} \;\; ||h|| < \delta.$$

5. Отображение $f : X \to Y$ называется *строго дифференцируемым* на элементе x^0, если найдется такое линейное непрерывное отображение $A : X \to Y$, что

$$\forall \varepsilon > 0 \;\; \exists \delta(\varepsilon) > 0 \;\; | \;\; ||f(x') - f(x'') - A(x' - x'')|| < \varepsilon ||x' - x''||$$

$$\text{при} \;\; ||x' - x^0|| < \delta, \;\; ||x'' - x^0|| < \delta.$$

Ясно, что A – производная Фреше.

Если отображение дифференцируемо по Фреше на элементе x^0, оно непрерывно на этом элементе. Если отображение строго дифференцируемо на элементе x^0, оно непрерывно в окрестности этого элемента. Эти утверждения, как и следующая теорема, следуют непосредственно из определений.

Теорема 1.
Каждое определение дифференцируемости отображения является более сильным, чем предыдущее, но не наоборот.

Пример. Пусть $f(x) = b \;\; \forall x \in X$. Тогда $f'(x^0) =?$

Пример. Пусть A – линейный непрерывный оператор. У аффинного оператора $Ax+b$ производная Гато совпадает с A. Как частный случай имеем: производная линейного непрерывного оператора совпадает с ним самим.

Пример. Пусть $A : X \to Y$ – линейный непрерывный оператор, X – банахово пространство, Y – гильбертово пространство с естественной нормой $||y||^2 = (y, y)$. Вычислим производную функционала $f(x) = ||Ax - b||^2$:

$$f(x^0+h)-f(x^0) = (Ax^0+Ah-b, Ax^0+Ah-b)-(Ax^0-b, Ax^0-b) =$$

$$= 2(Ax^0 - b, Ah) + (Ah, Ah) = 2(A^*(Ax^0 - b), h) + ||Ah||^2 \ldots$$

Пример. Дифференцируемо ли отображение

$$f[x(\cdot)] = \int\limits_c^d (\int\limits_a^b x(\tau)k(\tau, t)\, d\tau - f(t))^2\, dt$$

в пространствах C и L_2?

4.2. Правила дифференцирования

Легко проверить, что сумма отображений дифференцируема, если дифференцируемо каждое слагаемое. Точно также, скалярный множитель можно выносить за знак производной.

Теорема 2. (о суперпозиции отображений)
Пусть X, Y, Z – нормированные пространства, U – окрестность x^0 в X, V – окрестность y^0 в Y, $\varphi : U \to V$, $\varphi(x^0) = y^0$, $\psi : V \to Z$, $f = \psi \circ \varphi : U \to Z$ – суперпозиция отображений. Если ψ дифференцируемо по Фреше на y^0, а φ дифференцируемо на x^0 в смысле 1. – 4., то и f дифференцируемо в смысле 1. – 4. При этом

$$f'(x^0) = \psi'(y^0) \circ \varphi'(x^0) \quad \textit{или} \quad f'(x^0; h) = \psi'(y^0)(\varphi'(x^0; h)).$$

Если отображения φ и ψ строго дифференцируемы, то и отображение f строго дифференцируемо.

Доказательство. Пусть отображение ψ дифференцируемо по Фреше. Тогда

$$\psi(y) = \psi(y^0) + \psi'(y^0)(y - y^0) + \alpha(y)||y - y^0||, \quad \lim_{y \to y^0} \alpha(y) = 0.$$

Если существует

$$\varphi'(x^0;h)=\lim_{\lambda\downarrow 0}\frac{\varphi(x^0+\lambda h)-\varphi(x^0)}{\lambda},$$

то

$$\lim_{\lambda\downarrow 0}\frac{f(x^0+\lambda h)-f(x^0)}{\lambda}=\lim_{\lambda\downarrow 0}\frac{\psi(\varphi(x^0+\lambda h))-\psi(y^0)}{\lambda}=$$

$$=\lim_{\lambda\downarrow 0}\frac{\psi'(y^0)(\varphi(x^0+\lambda h)-y^0)+\alpha(\varphi(x^0+\lambda h))\,||\varphi(x^0+\lambda h)-y^0||}{\lambda}=$$

$$=\psi'(y^0)(\varphi'(x^0;h)).$$

В курсе МА (математического анализа) была доказана теорема о среднем (формула конечных приращений Лагранжа): если функция $f\ :\ [a,b]\to R^1$ непрерывна на $[a,b]$ и дифференцируема в (a,b), то найдется $c\in(a,b)\quad|\quad f(b)-f(a)=f'(c)(b-a)$. Самое важное ее следствие: $|f(b)-f(a)|\le\max\limits_{c\in[a,b]}=|f'(c)|\,|b-a|$. Аналогичное утверждение можно доказать и для более общих отображений.

Отрезок в линейном пространстве определяется так:
$[a,b]=\{x\mid x=a+t(b-a),\ \ 0\le t\le 1\}$.

Теорема 3.
Пусть X, Y – нормированные пространства и открытое множество $U\subset X$ содержит отрезок $[a,b]$. Если отображение $f\ :\ U\to Y$ дифференцируемо по Гато на каждом элементе $x\in[a,b]$, то

$$||f(b)-f(a)||\le\sup_{c\in(a,b)}||f'_G(c)||\cdot||b-a||.$$

Следствие.
Пусть выполнены условия теоремы 3 и $\Lambda\in\mathcal{L}(X,Y)$. Тогда

$$||f(b)-f(a)-\Lambda(b-a)||\le\sup_{c\in(a,b)}||f'_G(c)-\Lambda||\cdot||b-a||.$$

4.3. Дифференцирование в произведениях пространств

Пусть X, Y, Z – нормированные пространства, $x^0 \in X$, U – окрестность x^0 в X, $f : X \to Y \times Z$, $U \subset D(f)$. Иными словами, $f = (g, h) \mid g : X \to Y$, $h : X \to Z$ – вектор-отображение, которое $x \in X$ ставит в соответствие $(g(x), h(x)) \in Y \times Z$.

Теорема 4.
Отображение $f = (g, h)$ дифференцируемо на элементе x^0 в смысле 1. – 5. тогда и только тогда, когда отображения g и h дифференцируемы на x^0 в том же смысле. При этом

$$f'(x^0; h) = (g'(x^0; h), h'(x^0; h)) \quad \dots \quad f'(x^0) = (g'(x^0), h'(x^0)).$$

Доказательство: все следует из определений. •

Пусть X, Y, Z – нормированные пространства, $x^0 \in X$, $y^0 \in Y$, U – окрестность (x^0, y^0) в $X \times Y$, $f : X \times Y \to Z$, $U \subset D(f)$.

Если отображение $x \mapsto f(x, y^0)$ дифференцируемо на элементе x^0 (по Гато, по Фреше или строго), то его производная называется *частной производной* по x отображения f на элементе (x^0, y^0) и обозначается $\dfrac{\partial f}{\partial x}(x^0, y^0)$. Аналогично определяется $\dfrac{\partial f}{\partial y}(x^0, y^0)$.

Теорема 5. (о полном дифференциале)
Пусть отображение $f : X \times Y \to Z$ имеет на каждом элементе $(x, y) \in X \times Y$ частные производные по Гато $\dfrac{\partial f}{\partial x}(x, y)$ и $\dfrac{\partial f}{\partial y}(x, y)$.

Если отображения $(x, y) \mapsto \dfrac{\partial f}{\partial x}(x, y)$ и $(x, y) \mapsto \dfrac{\partial f}{\partial y}(x, y)$ непрерывны на элементе (x^0, y^0), то отображение f строго дифференцируемо на (x^0, y^0) и

$$f'(x^0, y^0)(\xi, \eta) = \frac{\partial f}{\partial x}(x^0, y^0)\,\xi + \frac{\partial f}{\partial y}(x^0, y^0)\,\eta.$$

Доказательство. Пусть $\varepsilon > 0$. Выберем $\delta > 0$ так, чтобы 1) $\Pi = \{(x, y) \mid ||x - x^0|| < \delta, \;\; ||y - y^0|| < \delta\} \subset U$ и 2)

$$||\frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(x^0, y^0)|| < \varepsilon, \quad ||\frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(x^0, y^0)|| < \varepsilon.$$

Пусть $(x', y'),\ (x'', y'') \in U$. Тогда

$$\Delta = f(x', y') - f(x'', y'') - \frac{\partial f}{\partial x}(x^0, y^0)(x' - x'') - \frac{\partial f}{\partial y}(x^0, y^0)(y' - y'') =$$

$$= f(x', y') - f(x'', y') - \frac{\partial f}{\partial x}(x^0, y^0)(x' - x'') +$$

$$+ f(x'', y') - f(x'', y'') - \frac{\partial f}{\partial y}(x^0, y^0)(y' - y'').$$

Элемент (x'', y') также принадлежит окрестности U, и каждый из отрезков $[(x', y'), (x'', y')]$, $[(x'', y'), (x'', y'')]$ лежат в U. Отображения $x \mapsto f(x, y')$ и $y \mapsto f(x'', y)$ дифференцируемы по Гато. По теореме о среднем

$$||\Delta|| \le \sup_{\xi \in [x', x'']} ||\frac{\partial f}{\partial x}(\xi, y') - \frac{\partial f}{\partial x}(x^0, y^0)|| \cdot ||x' - x''|| +$$

$$+ \sup_{\eta \in [y', y'']} ||\frac{\partial f}{\partial y}(x', \eta) - \frac{\partial f}{\partial y}(x^0, y^0)|| \cdot ||y' - y''|| \le$$

$$\le \varepsilon\, ||x' - x''|| + \varepsilon\, ||y' - y''||. \;\bullet$$

4.4. Производные интегральных функционалов

Теперь можно строго доказать, что все рассмотренные в разделах 2 и 3 функционалы дифференцируемы, если соответствующие функции непрерывны и имеют непрерывные частные производные.

Рассмотрим функционал из простейшей задачи КВИ

$$I[x(\cdot)] = \int\limits_{t_0}^{t_1} L(t, x(t), x'(t))\, dt.$$

Чтобы обосновать дифференцируемость этого функционала, запишем его в виде суперпозиции нескольких отображений.

Пусть U – открытое множество в $R^3 = R^1 \times R^1 \times R^1$, $L : U \to R^1$, частные производные $\dfrac{\partial L}{\partial x}$ и $\dfrac{\partial L}{\partial y}$ непрерывны в U. Обозначим

$$V = \{x(\cdot) \in C^1[t_0, t_1] \ | \ (t, x(t), x'(t)) \in U \ \ \forall t \in [t_0, t_1]\}.$$

Пусть

$$D \ : \ x(\cdot) \mapsto x'(\cdot),$$

$$W \ : \ x(\cdot),\, y(\cdot) \mapsto (x(\cdot), y(\cdot)),$$

$$N \ : \ w(\cdot) \mapsto g(\cdot) \ | \ g(t) = L(t, w(t)),$$

$$J \ : \ g(\cdot) \mapsto \int\limits_{t_0}^{t_1} g(t)\, dt.$$

Тогда функционал I представляет собой суперпозицию этих отображений.

Функционал

$$F[x(\cdot)] = \int\limits_a^b f(t, x(t), \int\limits_a^b k(\tau, t, x(\tau))\, d\tau)\, dt$$

также можно представить в виде суперпозиции нескольких отображений.

4.5. Условия экстремума функционалов

Пусть X – нормированное (топологическое) пространство, $x^0 \in X$, U – окрестность x^0 в X, функционал $f \ : \ U \to R^1, \quad U \subset D(f)$.

Говорят, что элемент x^0 *доставляет минимум* функционалу $f(\cdot)$, если $f(x) \geq f(x^0) \quad \forall x \in U$. Обозначение: $x^0 \in \mathrm{loc}\,\mathrm{min} f(\cdot)$.

Теорема 6. (необходимое условие экстремума 1-го порядка) *Если $x^0 \in loc\ min f(\cdot)$ и функционал $f(\cdot)$ имеет на элементе x^0 или производную по направлению h, или 1-ю вариацию по Лагранжу, или производную Гато (Фреше), то или $f'(x^0;h) \geq 0$, или $\delta f(x^0;h) = 0 \quad \forall h \in X$, или $f'_G(x^0) = 0$ ($f'(x^0) = 0$).*

Доказательство. Все следует из определений дифференцируемости, см. п. 4.1.

Замечание. Эта теорема – аналог теоремы Ферма классического МА: если дифференцируемая функция имеет экстремум в точке, то ее производная обращается в нуль в этой точке.

Определим *вторые производные* отображений. Для производных по направлению

$$f''(x^0;h) = \lim_{\lambda \downarrow 0} \frac{f'(x^0 + \lambda\, h; h) - f'(x^0;h)}{\lambda}.$$

(Мы пишем одно h в производной, но на самом деле их два.)

Лемма. (формула Тейлора)
Если отображение $f(\cdot)$ имеет производную Фреше порядка n, то

$$f(x^0 + h) = f(x^0) + f'(x^0)\, h + \frac{1}{2!}\, f''(x^0)\,(h,h) + \ldots$$

$$\ldots + \frac{1}{n!}\, f^{(n)}(x^0)\,(h,h,\ldots,h) + \alpha(h)\, ||h||^n, \qquad \lim_{h \to 0} \alpha(h) = 0.$$

Доказательство проводится по индукции.

Для вариаций отображений

$$\delta^n f(x^0;h) = \frac{d^n}{d\lambda^n}\, f(x^0 + \lambda h)|_{\lambda=0}.$$

Теорема 7. (необходимое условие экстремума 2-го порядка) *Пусть* $x^0 \in loc\ min f(\cdot)$. *Если функционал* $f(\cdot)$ *имеет на элементе* x^0 *или 2-ю вариацию по Лагранжу, или 2-ю производную Гато (Фреше), то*
или $\delta^2 f(x^0; h) \geq 0 \quad \forall h \in X$, *или* $f''_G(x^0)(h,h) \geq 0$ $(f''(x^0)(h,h) \geq 0)$ $\forall h \in X$.

Доказательство, как и в классическом анализе, основано на формуле Тейлора.

Теорема 8. (достаточное условие экстремума 2-го порядка) *Пусть функционал* $f(\cdot)$ *имеет на элементе* x^0 *2-ю производную Фреше. Если* $f'(x^0) = 0$ *и найдется такое число* $\alpha > 0$, *что*

$$f''(x^0)(h,h) \geq \alpha\, ||h||^2 \quad \forall h \in X,$$

то $x^0 \in loc\ min f(\cdot)$.

Доказательство также проводится с помощью формулы Тейлора.

Пример. Пусть X, Y – гильбертовы пространства, $A : X \to Y$ – линейный непрерывный оператор, $f(x) = ||Ax - y||^2 = (Ax - y, Ax - y)$. Вычислим производные этого функционала.

Так как

$$f(x+h) - f(x) = 2(Ax - y, Ah) + (Ah, Ah)$$

и

$$(Ax - y, Ah) = (A^*(Ax - y), h), \qquad (Ah, Ah) = ||Ah||^2 \leq ||A||^2 ||h||^2,$$

то можно считать, что $f'(x) = 2A^*(Ax - y)$. Тогда 2-я производная равна $2A^*A$.

4.6. Правило множителей Лагранжа

Метод множителей Лагранжа уже использовался в п.2.3 и п. 3.3 при решении изопериметрических задачи ВИ. Распространим этот метод на более сложные задачи. Начнем с наиболее простого конечномерного случая.

Пусть функции $f_j : R^n \to R^1, \quad j = 0..m \ (m < n)$. Рассмотрим конечномерную экстремальную задачу с ограничениями в виде равенств

$$f_0(x) \to \text{extr}, \quad f_j(x) = 0, \ \ j = 1..m. \quad (x \in R^n)$$

Теорема 9.
Пусть функции $f_j(\cdot), \ \ j = 0..m$, имеют непрерывные частные производные в окрестности точки x^0. Если x^0 – решение экстремальной задачи, то найдется такой ненулевой вектор (множители Лагранжа) $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, что для функции Лагранжа

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

выполняется условие (условие стационарности)

$$\frac{\partial \mathcal{L}}{\partial x}(x^0, \lambda) = 0 \quad \text{или} \quad \sum_{j=0}^{m} \lambda_j f_j'(x^0) = 0$$

$$\text{или} \quad \sum_{j=0}^{m} \lambda_j \frac{\partial f_j}{\partial x_k}(x^0) = 0, \ \ k = 1..n.$$

Доказательство. От противного. Предположим, что условие стационарности не выполняется:

$$\sum_{j=0}^{m} \lambda_j f_j'(x^0) \neq 0.$$

Тогда векторы $f_j'(x^0)$ линейно независимы и ранг матрицы

$$\{\frac{\partial f_j}{\partial x_k}(x^0)\}_{j=0..m,\, k=1..n}$$

равен $m+1$. Следовательно, существует не равный нулю минор порядка $m+1$.

Можно считать, что этот минор расположен в левом верхнем углу матрицы. А еще можно считать, что $f_0(x^0)=0$ (иначе введем новую искомую функцию $\widetilde{f}(x)=f_0(x)-f_0(x^0)$).

Обозначим $\check{x}=(x_1,\ldots,x_{m+1})$ и $\hat{x}=(x_{m+2},\ldots,x_n)$. В окрестности U фиксированной точки $\check{x}^0=(x_1^0,\ldots,x_{m+1}^0)$ рассмотрим отображение

$$F\ :\ U\to R^{m+1}\ \ |\ \ F(\check{x})=(f_0(\check{x},\hat{x}^0),\ldots,f_m(\check{x},\hat{x}^0)).$$

Это отображение непрерывно дифференцируемо и $F(\check{x}^0)=0$.

По теореме об обратной функции в конечномерных пространствах существует обратное отображение F^{-1}, определенное в окрестности точки $0\in R^{m+1}$ и дифференцируемое в этой окрестности. Для y из такой окрестности выполняется неравенство $|F^{-1}(y)-F^{-1}(0)|\leq L|y-0|$ или $|F^{-1}(y)-\check{x}^0|\leq L|y|$.

Пусть ε – достаточно малое число (любого знака). Пусть $y_\varepsilon=(\varepsilon,0,\ldots,0)$ и $\check{x}_\varepsilon=F^{-1}(y_\varepsilon)$. Тогда $F(\check{x}_\varepsilon)=(\varepsilon,0,\ldots,0)$ и $|\check{x}_\varepsilon-\check{x}^0|\leq L|\varepsilon|$.

Обозначим $x_\varepsilon=(\check{x}_\varepsilon,\hat{x}^0)$. Тогда $f_0(x_\varepsilon)=\varepsilon$ и $f_j(x_\varepsilon)=0,\ \ j=1..m$. Получилось, что x^0 не доставляет локальный экстремум в задаче, так как в окрестности этой точки есть точки x_ε, удовлетворяющие ограничениям в виде равенств, но значение $f_0(x_\varepsilon)$ может быть как положительным, так и отрицательным. $\bullet$

Замечание 1. Производные можно понимать как по Гато или Фреше, так и как 1-е вариации по Лагранжу.

Замечание 2. Если векторы $f_j'(x^0),\ \ j=1..m$ линейно независимы, то $\lambda_0\neq 0$.

Пример.

$$xyz\to \mathrm{extr},\ \ x+y+z=0,\ \ x^2+y^2+z^2=1.$$

В случае бесконечномерных пространств дело обстоит существенно сложнее.

Пусть X, Y – нормированное пространство, отображение $F : X \to Y$, функционалы $f_j : X \to R^1, \;\; j = 0..m$. Рассмотрим экстремальную задачу с ограничениями типа равенств и неравенств:

$$f_0(x) \to \min, \quad f_j(x) \le 0, \;\; j = 1..m, \quad F(x) = 0, \quad x \in X.$$

Теорема 10.
Пусть X, Y – банаховы пространства, функционалы $f_j(\cdot), \quad j = 0..m$ и отображение $F(\cdot)$ строго дифференцируемы в окрестности точки x^0, множество значений оператора $F'(x^0)$ – замкнутое подпространство в Y.

Если x^0 – решение экстремальной задачи, то существуют такие вектор $\lambda \in R^{m+1}$ и функционал $y^* \in Y^*$ (множители Лагранжа), не равные нулю одновременно, что для функции Лагранжа

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

выполняются:
1) условие стационарности

$$\frac{\partial \mathcal{L}}{\partial x}(x^0, \lambda, y^*) = 0,$$

2) условия дополняющей нежесткости $\lambda_j \, f_j(x^0) = 0, \;\; j = 1..m$,
3) условия неотрицательности $\lambda_j \ge 0, \;\; j = 0..m$.

Дополнение к теореме. Если $f_j(x) = 0$, то условия 2) и 3) отсутствуют.

Доказательство можно найти в книге [4].

Пример (все равно конечномерный :-)).

$$x_1^2 + x_2^2 + x_3^2 \to \min, \quad 2x_1 - x_2 + x_3 \le 5, \quad x_1 + x_2 + x_3 = 3.$$

Функция Лагранжа

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

Условия стационарности

$$1)\;\; \lambda_0 2x_1 + 2\lambda_1 + \lambda_2 = 0, \;\; \lambda_0 2x_2 - \lambda_1 + \lambda_2 = 0, \;\; \lambda_0 2x_3 + \lambda_1 + \lambda_2 = 0,$$

условия дополняющей нежесткости

$$2) \qquad \lambda_1\,(2x_1 - x_2 + x_3 - 5) = 0,$$

условия неотрицательности

$$3) \qquad \lambda_0 \geq 0, \quad \lambda_1 \geq 0$$

и ограничения задачи

$$4) \qquad 2x_1 - x_2 + x_3 \leq 5, \quad x_1 + x_2 + x_3 = 3$$

образуют систему равенств и неравенств, которая имеет единственное решение $x_1 = x_2 = x_3 = -\lambda_2 = 1$, $\lambda_1 = 0$, $\lambda_0 = 0.5$ или какое-то другое положительное число.

Рассуждения были такие. Если $\lambda_0 = 0$, то уравнения 1) имеют только нулевое решение. Этого быть не может. Следовательно $\lambda_0 \neq 0$. Удобно взять $\lambda_0 = 1/2$.

Если $\lambda_1 \neq 0$, то из 1) имеем $x_1 = -2\lambda_1 - \lambda_2$, $x_2 = \lambda_1 - \lambda_2$, $x_3 = -\lambda_1 - \lambda_2$. Условие 2) дает $2x_1 - x_2 + x_3 = 5$. Тогда получается, что $\lambda_1 < 0$. Этого также быть не может. Значит $\lambda_1 = 0$.

* * *

Этот раздел написан в основном по главе 2 книги [1].

5. ОПТИМИЗАЦИЯ В БЕСКОНЕЧНОМЕРНЫХ ПРОСТРАНСТВАХ

Для поиска экстремумов функций, зависящих от конечного числа переменных, предложено много эффективных алгоритмов (см., например, [16]). Случай бесконечномерных пространств существенно сложнее. Но некоторые хорошо известные методы решения задач оптимизации естественным образом переносятся на этот случай.

5.1. Градиентные методы

Пусть функционал $I \,:\, X \to R^1$ дифференцируем по Фреше на элементе $x^0 \in X$, то есть $I(x^0 + h) = I(x^0) + I'(x^0)\, h + o(||h||)$. Здесь $I'(x^0)$ – тоже функционал, причем линейный и непрерывный. Его значение на элементе h обозначено $I'(x^0)\, h$, также часто пишут $< h,\, I'(x^0) >$.

Если $X = R^n$ (конечномерное пространство), то $I'(x^0)$ – градиент функционала на элементе x^0. Это вектор (по определению), который показывает направление наибольшего возрастания функционала (или просто функции n переменных). Можно построить последовательность элементов x^k, сходящуюся к решению x^0 экстремальной задачи на минимум: x^1 – начальное приближение, $x^{k+1} = x^k - \alpha_k\, I'(x^k)$. Такую последовательность называют *минимизирующей последовательностью.* Но в бесконечномерном пространстве все не так просто.

Предположим, что X – гильбертово пространство (вещественное или даже комплексное). По теореме Ф. Рисса каждому линейному непрерывному функционалу $f(\cdot)$ на X соответствует такой единственный элемент $y \in X$, что $< x,\, f >= (x,\, y)$. Кроме того, $||f|| = ||y||$. Так вот, в бесконечномерном пространстве *градиентом* функционала $f(\cdot)$ называют этот самый элемент y. В дальнейшем вместо y будем писать $I'(x^0)$. Поэтому для решения экстремальных задач в бесконечномерных пространствах можно использовать практически такие же методы и алгоритмы, что и в конечномерном случае. Рассмотрим некоторые из них.

1. *Градиентный метод* используется при решении экстремальных задач без ограничений

$$I(x) \to \min, \quad x \in X,$$

здесь X – гильбертово пространство.

Алгоритм такой: выбирается начальный элемент x^1, а следующие элементы минимизирующей последовательности вычисляются по формуле

$$x^{k+1} = x^k - \alpha_k \, I'(x^k), \quad k = 1, 2, ...,$$

пока не будет достигнута требуемая точность. Здесь $I'(x^k)$ – градиент функционала на элементе x^k.

Числа α_k могут быть определены различными способами, в том числе:

– *примитивно*, то есть из условия монотонности $I(x^{k+1}) < I(x^k)$. Например, можно задать некоторое значение α, положить $\alpha_k = \alpha$, а если на k-м шаге условие монотонности не выполняется, нужно уменьшать это число;

– *априорно*, то есть так, что

$$\alpha_k > 0, \quad \sum_{k=1}^{+\infty} \alpha_k = +\infty, \quad \sum_{k=1}^{+\infty} \alpha_k^2 < +\infty.$$

Например, $\alpha_k = 1/(k^\lambda)$, $1/2 < \lambda < 1$;

– *оптимально*, когда α_k – решение задачи

$$I(x^k - \alpha_k \, I'(x^k)) \to \min$$

(это метод скорейшего спуска).

2. *Метод проекции градиента* применяют при решении задач с ограничениями

$$I(x) \to \min, \quad x \in \mathfrak{x} \subset X,$$

где $\mathfrak{x}$ – выпуклое замкнутое подмножество в X.

На каждом шаге алгоритма новое приближение к решению задачи нужно спроектировать на множество ограничений,

$$x^{k+1} = P_{\mathfrak{X}}[x^k - \alpha_k\, I'(x^k)], \quad k = 1, 2, ...$$

Здесь $P_{\mathfrak{X}}[x]$ – проекция элемента x на множество $\mathcal{X}$, то есть такой элемент $p \in \mathcal{X}$, что

$$||x - p|| = \inf_{q \in \mathfrak{X}} ||x - q||.$$

Например, если $\mathfrak{x} = \{x \in X \mid ||x - \tilde{x}|| \le R\}$ – замкнутый шар, то

$$P_{\mathfrak{X}}(x) = \{\ ||x - \tilde{x}|| \le R\ :\ x; \quad \text{иначе} \ \ :\ \tilde{x} + r\,(x - \tilde{x})/||x - \tilde{x}||\}.$$

3. *Метод условного градиента* также часто используется, он основан на следующем алгоритме:

$$x^{k+1} = x^k + \alpha_k\,(\overline{x}^k - x^k), \quad 0 \le \alpha_k \le 1,$$

где $\overline{x}^k$ – решение задачи

$$< \overline{x}^k - x^k,\, I'(x^k) >= \inf_{x \in \mathfrak{X}} < x - x^k,\, I'(x^k) >$$

($\mathfrak{x}$ – ограниченное выпуклое замкнутое множество).

4. *Метод Ньютона-Канторовича* отличается тем, что элемент $\overline{x}^k$ ищется из условия $I_k(\overline{x}^k) = \inf\limits_{x \in \mathfrak{X}} I_k(x)$, где

$$I_k(x) =< x - x^k,\, I'(x^k) > + \frac{1}{2} < x - x^k,\, I''(x^k)(x - x^k) > .$$

Если $\mathfrak{x} = X$ и оператор $I''(x^k)$ имеет обратный, то $x^{k+1} = x^k - \alpha_k\,(I''(x^k))^{-1} I'(x^k)$.

Доказаны теоремы о сходимости градиентных методов, получены оценки скорости их сходимости [16].

5.2. Задача оптимального управления с линейным уравнением состояний и квадратичным функционалом

Рассмотрим упрощенный одномерный вариант одной задачи оптимального управления ([3], с. 31).

Постановка задачи следующая:

$$I = [x(t_1) - x_1]^2 \to \min,$$

$$x'(t) = a(t)\,x(t) + b(t)\,u(t) + f(t), \quad t_0 < t < t_1, \quad x(t_0) = x_0,$$

$$u(t) \in \mathcal{U} \quad \forall t \in [t_0, t_1].$$

Здесь $u(\cdot)$ – управление, $x(\cdot)$ – состояние (оно зависит от управления, $x(\cdot) = x[u(\cdot)](\cdot)$). Поэтому и функционал зависит от управления, $I = I[u(\cdot)]$.

Смысл задачи состоит в следующем: задано начало траектории $x = x(t)$; нужно найти такое управление, что соответствующая ему траектория заканчивается наиболее близко к заданной точке.

Чтобы применить метод проекции градиента, нужно научиться вычислять градиент!

План такой: дадим приращение $\Delta u(\cdot)$ управлению $u(\cdot)$ и исследуем соответствующее ему приращение функционала

$$\Delta I = I[u(\cdot) + \Delta u(\cdot)] - I[u(\cdot)].$$

Если окажется, что

$$\Delta I = \int\limits_{t_0}^{t_1} G(\tau)\,\Delta u(\tau)\,d\tau + o(||\Delta u(\cdot)||),$$

то функция $G(\cdot)$ и есть градиент функционала.

Обозначим приращение состояния $\Delta x(\cdot)$,

$$\Delta x(\cdot) = x[u(\cdot) + \Delta u(\cdot)](\cdot) - x[u(\cdot)](\cdot).$$

Эта функция является решением задачи Коши

$$\Delta x'(t) = a(t)\,\Delta x(t) + b(t)\,\Delta u(t), \quad x(t_0) = 0.$$

Тогда (вспомним, как это неоднократно делали при изучении курса ДУ :-))

$$\Delta x(t) = \int\limits_{t_0}^{t} [a(\tau)\,\Delta x(\tau) + b(\tau)\,\Delta u(\tau)]\,d\tau$$

и

$$|\Delta x(t)| \le A \int\limits_{t_0}^{t} |\Delta x(\tau)|\,d\tau + B \int\limits_{t_0}^{t} |\Delta u(\tau)|\,d\tau \le B\,|\Delta u(\tau)|\,e^{A(t-t_0)}$$

(лемма об интегральных неравенствах), здесь $A = \max\limits_{t\in[t_0,t_1]} a(t)$, $B = \max\limits_{t\in[t_0,t_1]} b(t)$.

Рассмотрим теперь приращение функционала

$$\Delta I = \{x[u(\cdot)+\Delta u(\cdot)](t_1) - x_1\}^2 - \{x[u(\cdot)](t_1) - x_1\}^2 =$$

$$= \{x[u(\cdot)+\Delta u(\cdot)](t_1) - x[u(\cdot)](t_1) - 2\,x_1\}\,\Delta x(t_1) =$$

$$= 2\{x[u(\cdot)](t_1) - x_1\}\,\Delta x(t_1) + [\Delta x(t_1)]^2,$$

чтобы привести выражение к нужному виду мы вычли и добавили $x[u(\cdot)](t_1)$. Второе слагаемое в силу полученного выше неравенства и есть $o(||\Delta u(\cdot)||)$.

Теперь преобразуем первое слагаемое.

Пусть $y(\cdot)$ – решение задачи Коши

$$y'(t) = -a(t)\,y(t), \quad y(t_1) = 2\{x[u(\cdot)](t_1) - x_1\}.$$

Тогда первое слагаемое

$$y(t_1)\,\Delta x(t_1) = y(t_1)\,\Delta x(t_1) - y(t_0)\,\Delta x(t_0) = \int\limits_{t_0}^{t_1} [y(\tau)\,\Delta x(\tau)]'\,d\tau =$$

$$= \int_{t_0}^{t_1} [y'(\tau)\,\Delta x(\tau) + y(\tau)\,\Delta x'(\tau)]\,d\tau = \int_{t_0}^{t_1} y(\tau)\,b(\tau)\,\Delta u(\tau)\,d\tau$$

(почему именно так, скоро станет ясно). Следовательно, градиент функционала равен $y(\cdot)\,b(\cdot)$.

В итоге алгоритм решения задачи оптимального управления такой:
– задаем начальное приближение $u(\cdot)$;
– находим $x(\cdot)$ как решение задачи Коши и, следовательно, $x(t_1)$;
– находим $y(\cdot)$ как решение задачи Коши;
– вычисляем градиент функционала $I'[u(\cdot)](\cdot) = y(\cdot)\,b(\cdot)$;
– находим новое управление $u(\cdot)$ по методу проекции градиента;
– если точность вычислений не достигнута, возвращаемся на второй шаг.

5.3. Оптимальное управление температурой стержня

Пусть вдоль отрезка $[0, l]$ оси s расположен тонкий проводящий тепло стержень, $x(s, t)$ – температура в точке (в сечении) стержня с координатой s в момент времени t. Рассмотрим экстремальную задачу

$$\frac{\partial x}{\partial t} = a^2\,\frac{\partial^2 x}{\partial s^2} + f(s,t), \quad (s,t) \in Q = \{0 < s < l,\ 0 < t < T\},$$

$$\frac{\partial x}{\partial s}(0,t) = 0, \quad \frac{\partial x}{\partial s}(l,t) = \nu\,[p(t) - x(l,t)], \quad 0 < t \le T,$$

$$x(s,0) = \varphi(s), \quad 0 \le s \le l,$$

$$\int_0^l [x(s,T) - y(s)]^2\,ds \to \min.$$

Условия этой задачи таковы: температура стержня (состояние) – решение уравнения теплопроводности, левый его конец теплоизолирован, только через правый конец происходит обмен теплом с окружающей средой, начальное распределение температуры вдоль стержня задано. Нужно найти такое управление – плотность источников тепла внутри стержня $f(s,t)$ и температуру окружающей среды $p(t)$, что в момент времени T распределение температуры вдоль стержня будет наиболее близким к заданному распределению.

Дополнительные ограничения на управление такие:

$$p_{min} \le p(t) \le p_{max}, \quad \iint\limits_Q |f(s,t)|^2 \, ds \, dt \le F.$$

Будем считать, что управление $u = (p, f)$ – непрерывные функции (хотя они могут быть интегрируемы с квадратом; если имеются точки разрыва, то нужно рассматривать обобщенное решение уравнение теплопроводности). Скалярное произведение в пространстве управлений

$$(u_1,\, u_2) = \int\limits_0^T p_1(t)\, p_2(t)\, dt + \iint\limits_Q f_1(s,t)\, f_2(s,t)\, ds\, dt.$$

Докажем, что функционал в этой задаче оптимального управления дифференцируем и вычислим его градиент ([3]).

Приращению управления $\Delta u = (\Delta p,\, \Delta f)$ соответствуют приращение состояния Δx и, следовательно, приращение функционала ΔI.

Приращение состояния является решением задачи

$$\frac{\partial(\Delta x)}{\partial t} = a^2\, \frac{\partial^2(\Delta x)}{\partial s^2} + \Delta f,$$

$$\frac{\partial(\Delta x)}{\partial s}(0,t) = 0, \quad \frac{\partial \Delta x}{\partial s}(l,t) = \nu\,[\Delta p - \Delta x(l,t)], \quad \Delta x(s,0) = 0.$$

Приращение функционала

$$\Delta I = \int\limits_0^l [x(s,T) + \Delta x(s,T) - y(s)]^2\, ds - \int\limits_0^l [x(s,T) - y(s)]^2\, ds =$$

$$= \int\limits_0^l 2[x(s,T) - y(s)]\,\Delta x(s,T)\,ds + \int\limits_0^l (\Delta x(s,T)) : 2\,ds.$$

Можно показать, что второе слагаемое есть $o(||\Delta u||)$.

Пусть функция $\psi(s,t)$ – решение вспомогательной задачи

$$\frac{\partial \psi}{\partial t} = -a^2\,\frac{\partial^2 \psi}{\partial s^2}, \quad \frac{\partial \psi}{\partial s}(0,t) = 0, \;\; \frac{\partial \psi}{\partial s}(l,t) = -\nu\,\psi(l,t),$$

$$\psi(s,T) = 2[x(s,T) - y(s)].$$

Первое слагаемое в приращении функционала нужно представить в виде

$$\int\limits_0^T G(t)\,\Delta p(t)\,dt + \iint\limits_Q H(s,t)\,\Delta f(s,t)\,ds\,dt,$$

тогда градиент – пара функций $(G(t),\ H(s,t))$.

Сделаем следующее:

$$\int\limits_0^l 2[x(s,T) - y(s)]\,\Delta x(s,T)\,ds = \int\limits_0^l \psi(s,T)\,\Delta x(s,T)\,ds =$$

$$= \int\limits_0^l (\int\limits_0^T \frac{\partial}{\partial t}(\psi\,\Delta x)\,dt)\,ds = \int\limits_0^l \int\limits_0^T [\frac{\partial \psi}{\partial t}\,\Delta x + \psi\,\frac{\partial(\Delta x)}{\partial t}]\,dt\,ds =$$

$$= \int\limits_0^l \int\limits_0^T [-a^2\,\frac{\partial^2 \psi}{\partial s^2}\,\Delta x \mp a^2\,\frac{\partial \psi}{\partial s}\,\frac{\partial(\Delta x)}{\partial s} + \psi\,a^2\,\frac{\partial^2(\Delta x)}{\partial s^2} + \psi\,\Delta f]\,dt\,ds =$$

$$= a^2 \int\limits_0^T \int\limits_0^l (\frac{\partial}{\partial s}[-\frac{\partial \psi}{\partial s}\,\Delta x + \psi\,\frac{\partial(\Delta x)}{\partial s}]\,ds)\,dt + \iint\limits_Q \psi\,\Delta f\,ds\,dt.$$

Так как

$$\int\limits_0^l \frac{\partial}{\partial s}[-\frac{\partial \psi}{\partial s}\,\Delta x + \psi\,\frac{\partial(\Delta x)}{\partial s}]\,ds =$$

$$= -\frac{\partial \psi}{\partial s}\,\Delta x + \psi\,\frac{\partial(\Delta x)}{\partial s}\,\Big|_{s=0}^{s=l} = \ldots = \psi(l,t)\,\nu\,\Delta p,$$

то градиент равен $(a^2\,\nu\,\psi(l,t),\ \psi(s,t))$.

В итоге алгоритм решения задачи оптимального управления следующий:
– задаем начальное приближение $p(t)$, $f(s,t)$;
– решаем краевую задачу и находим $x(s,t)$ и $x(s,T)$;
– решаем краевую задачу и находим $\psi(s,t)$;
– вычисляем градиент $I'[u(\cdot)](\cdot) = (a^2\,\nu\,\psi(l,t),\, \psi(s,t))$;
– находим новое управление $p(t)$, $f(s,t)$ (не выходя из допустимого множества);
– если точность вычислений не достигнута, возвращаемся на второй шаг.

6. ЗАДАЧИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

В задачах линейного программирования отыскиваются экстремумы линейных функционалов на множествах, заданных линейными равенствами и неравенствами.

Классические задачи линейного программирования обычно рассматриваются в конечномерных пространствах. Например, прямая задача в стандартной форме

$$< c, x > \to \max, \quad Ax \le b, \quad x \ge 0 \tag{1}$$

и двойственная задача

$$< b, y > \to \min, \quad A'y \ge c, \quad y \ge 0 \tag{2}$$

являются конечномерными. Здесь $c, x \in R^n$, A — матрица $m \times n$, $b, y \in R^m$, A' — транспонированная матрица, x и y — искомые векторы.

Как перенести теорию линейного программирования на абстрактный бесконечномерный случай? Можно ли рассматривать задачи (1) и (2) как результат аппроксимации двойственных бесконечномерных задач линейного программирования?

6.1. Двойственные пространства и двойственные операторы

Пространства X и X' называются *двойственными пространствами* (двойственной парой), если задан невырожденный билинейный функционал (билинейная форма) $< \cdot, \cdot >: X \times X' \to R$.

Билинейный функционал называется невырожденным, если из $< x, x' >= 0 \; \forall x' \in X'$ следует $x = 0$ и из $< x, x' >= 0 \; \forall x \in X$ следует $x' = 0$. Это выполняется тогда и только тогда, когда существует такой элемент $x' \in X'$, что $< x, x' >\neq 0$ для всех $x \in X$, $x \neq 0$, и существует такой элемент $x \in X$, что $< x, x' >\neq 0$ для всех $x' \in X'$, $x' \neq 0$.

Оба пространства в двойственной паре равноправны. Пространство может быть двойственно само себе. Каждое вещественное линейное пространство со скалярным произведением двойственно само себе, если билинейная форма $< x, x' >= (x, x')$.

Линейное пространство X и алгебраически сопряженное пространство X^+ (множество всех линейных функционалов на X) образуют двойственную пару с $< x, x^+ >= x^+(x)$. Линейное пространство X и сопряженное пространство X^* (множество всех линейных непрерывных функционалов на X) также образуют двойственную пару. Но симметрия пространств в такой двойственной паре нарушена.

Пусть X, X' и Y, Y' — двойственная пара линейных пространств, $A : X \to Y$ — линейный оператор, причем его область определения $D(A) = X$. Оператор $A' \; : \; Y' \to X'$ называется *двойственным оператором* к A, если

$$< x, A'y' >=< Ax, y' > \quad \forall x \in X, \;\; \forall y' \in Y'.$$

Но в теории линейного программирования удобно пользоваться другим определением.

Операторы $A : X \to Y'$ и $A' : Y \to X'$ назовем *взаимно двойственными*, если

$$< x, A'y >=< y, Ax > \quad \forall x \in X, \ \ \forall y \in Y.$$

Здесь нет противоречия, всего лишь используются другие обозначения. В общем случае оператор, двойственный к линейному, может и не существовать.

Если $X = X' = C[a, b]$, то можно использовать билинейную форму

$$< x, x' >= \int\limits_a^b x(t)\, x'(t)\, dt.$$

Пусть $X = X' = Y = Y' = C[a, b]$ и обе билинейные формы определены именно так. Легко проверить, что если $k(\tau, t)$ – непрерывная функция, то операторы

$$A \,:\, x(t) \mapsto x(t) + \int\limits_a^b x(\tau)\, k(\tau, t)\, d\tau$$

и

$$A' \,:\, y(t) \mapsto y(t) + \int\limits_a^b y(\tau)\, k(t, \tau)\, d\tau$$

взаимно двойственны.

6.2. Бесконечномерное линейное программирование

Пусть X, X' и Y, Y' – двойственная пара линейных упорядоченных пространств, и $x'_0 \in X'$, $y'_0 \in Y'$ – их фиксированные элементы. Рассмотрим две двойственные задачи абстрактного линейного программирования.

Пусть $A : X \to Y'$ – линейный оператор и $A' : Y \to X'$ – оператор, двойственный к A. Прямая задача линейного программирования и двойственная к ней задача могут быть сформулированы так:

$$< x, x_0' >\to \max, \quad Ax \le y_0', \quad x \ge 0, \tag{3}$$

$$< y, y_0' >\to \min, \quad A'y \ge x_0', \quad y \ge 0. \tag{4}$$

Легко видеть, что двойственные задачи конечномерного линейного программирования (1) и (2) являются частным случаем этой общей схемы.

Теорема 1. *Если x и y – допустимые элементы, то выполняется неравенство $< x, x_0' >\le< y, y_0' >$. Если x и y – допустимые элементы и $< x^0, x_0' >=< y^0, y_0' >$, то x^0 и y^0 – решения задач (3) и (4).*

Доказательство. Если $x_0' \le A'y$, то $< x, x_0' >\le< x, A'y > \quad \forall x \in P_X^+$. С другой стороны, если $Ax \le y_0'$, то $< y, Ax >\le< y, y_0' > \forall y \in P_Y^+$. Но $< x, A'y >=< y, Ax >$ по определению двойственного оператора. Далее, пусть x, y — произвольные допустимые элементы. Тогда $< x, x_0' >\le< y^0, y_0' >=< x^0, x_0' >\le< y, y_0' >$. $\bullet$

Пример. Рассмотрим пару задач интегрального линейного программирования. Прямая задача имеет вид

$$\int\limits_\alpha^\beta c(\tau)x(\tau)\,d\tau \to \max,$$

$$\tag{5}$$

$$\int\limits_\alpha^\beta A(t,\tau)x(\tau)\,d\tau \le b(t),\ t \in [\gamma,\delta], \quad x(\tau) \ge 0,\ \tau \in [\alpha,\beta]$$

и двойственная задача –

$$\int\limits_\gamma^\delta b(t)y(t)\,dt \to \min,$$

$$\int\limits_{\gamma}^{\delta} A(\tau,t)y(t)\,dt \geq c(\tau),\ \tau\in[\alpha,\beta], \quad y(t)\geq 0,\ t\in[\gamma,\delta]. \tag{6}$$

Достаточно предположить, что все функции в формулах (5) и (6) непрерывны.

Еще один **пример** – задача о распределении силы вдоль упругой струны. Предположим, что струна закреплена в точках $x = 0$ и $x = l$. Как известно, ее отклонение от положения равновесия под действием распределенной силы $p(x)$

$$u(x) = \int\limits_0^l G(x,t)\,p(t)\,dt,$$

где $G(x,t)$ — функция Грина (см. п. 1.2). Необходимо, чтобы каждая точка струны вышла за границу заданной области. Пусть полная нагрузка должна быть минимальной, а распределение силы вдоль струны неизвестно. Тогда имеем задачу интегрального линейного программирования

$$\int\limits_0^l p(t)\,dt \to \min, \quad \int\limits_0^l G(x,t)p(t)\,dt \geq q(x), \quad p(x)\geq 0.$$

Канонические формы прямой и двойственной задач линейного программирования имеют вид

$$< x, x_0' > \to \max, \quad Ax = y_0', \quad x \geq 0,$$

$$< y, y_0' > \to \min, \quad A'x = x_0', \quad y \geq 0.$$

Легко преобразовать задачи в стандартной форме в задачи в канонической форме. Новые переменные можно добавить следующим образом.

Пусть $U = X \times Y'$ и $V = Y \times X'$ — новые пространства искомых элементов. Пространства $U' = X' \times Y$ и $V' = Y' \times X$ являются двойственными к ним, причем двойственность задается билинейными формами

$$< u, u' >=< x, x' > + < y, y' >=< v, v' > .$$

Пусть $u'_0 = (x'_0, 0)$ и $v'_0 = (y'_0, 0)$ — фиксированные элементы. Построим операторы $B : U \to V'$ и $C : V \to U'$ так, что $B : (x, y') \mapsto (Ax + y', 0)$ и $C : (y, x') \mapsto (A'y - x', 0)$. Тогда канонические задачи

$$< u, u'_0 >\to \max, \quad Bu = v'_0, \quad u \geq 0,$$

$$< v, v'_0 >\to \min, \quad Cv = u'_0, \quad v \geq 0$$

эквивалентны исходным прямой и двойственной задачам. Заметим, что операторы B и C не являются взаимно двойственными.

6.3. Аппроксимация и интерполяция пространств и операторов

Пространство $\overline{X}$ назовем *аппроксимирующим* пространство X, если заданы такие *операторы аппроксимации и интерполяции* $T_X : X \to \overline{X}$ и $S_X : \overline{X} \to X$, что $T_X\, S_X = I$ (здесь I – тождественный оператор). Если аппроксимация нетривиальная, то $S_X\, T_X \neq I$. Если пространства $\overline{X}$, $\overline{Y}$ аппроксимируют пространства X, Y, то любой оператор $\overline{A} : \overline{X} \to \overline{Y}$ назовем оператором, *аппроксимирующим* точный оператор A. Аналогично, если задан оператор $\overline{A} : \overline{X} \to \overline{Y}$, то любой оператор $\widetilde{A} : X \to Y$ является оператором, *интерполирующим* оператор $\overline{A}$. Операторы $\overline{A}^0 = T_Y\, A\, S_X$ и $\widetilde{A}^0 = S_Y\, \overline{A}\, T_X$ назовем операторами естественной аппроксимации и естественной интерполяции.

Пусть X, X' и Y, Y' — двойственные пары пространств, $A : X \to Y$ и $A' : Y' \to X'$ — двойственные операторы. Предположим, что пространства $\overline{X}$ и $\overline{Y}$ аппроксимируют пространства X и Y, операторы T_X, S_X, T_Y, S_Y — операторы аппроксимации и интерполяции.

Наиболее естественный способ аппроксимации двойственных пространств и операторов состоит в следующем. Пусть $\overline{X}'$ и $\overline{Y}'$ — пространства, двойственные к $\overline{X}$ и $\overline{Y}$ соответственно.

Теорема 2.

I. *Если операторы T_X и S_X имеют двойственные операторы, то пространство $\overline{X}'$ аппроксимирует пространство X'.*

II. *Если оператор $\overline{A}'$ существует, то он аппроксимирует оператор A'.*

III. *Если $\overline{A'} = \overline{A}'$, то $(\overline{A}^0)' = \overline{(A')}^0$ и $(\widetilde{A}^0)' = \widetilde{(A')}^0$.*

Доказательство можно найти в [17].

Первое утверждение теоремы можно сформулировать еще и так: пространство, двойственное к аппроксимирующему пространству, аппроксимирует пространство, двойственное к точному пространству.

Заметим, что из равенства $\overline{X'} = \overline{X}'$ не следует, что операции аппроксимации и двойственности коммутируют. Это равенство означает только, что пространство $\overline{X}'$ выбрано в качестве аппроксимации пространства X' среди всех возможных аппроксимаций.

Равенство $\overline{A'} = \overline{A}'$ также означает, что оператор $\overline{A}'$ выбран из операторов, аппроксимирующих оператор A. Можно выбрать другой аппроксимирующий оператор.

6.4. Аппроксимация задач линейного программирования

Предположим теперь, что операторы аппроксимации и интерполяции сохраняют порядок в линейных пространствах. Например, если $T_X : X \to \overline{X}$, то из $x \geq 0$ следует, что $\overline{x} = T_X x \geq 0$.

Задачу линейного программирования, аппроксимирующую прямую задачу (3), можно сформулировать следующим образом:

$$< \overline{x}, S'_X x'_0 > \to \max, \quad T_{Y'} A S_X \overline{x} \le T_{Y'} y'_0, \quad \overline{x} \ge 0. \tag{7}$$

Здесь элементы точного пространства заменены на элементы аппроксимирующего пространства, и оператор естественной аппроксимации $\overline{A}^0 = T_{Y'} A S_X$ выбран в качестве аппроксимации оператора $\overline{A} : \overline{X} \to \overline{Y'}$.

К задаче (7) можно прийти следующим образом. Будем искать решение задачи (3) в виде $x = S_X \overline{x}$. Тогда

$$< S_X \overline{x}, x'_0 > \to \max, \quad A S_X \overline{x} \le y'_0, \quad S_X \overline{x} \ge 0.$$

Заменим теперь оператор S_X в функционале на двойственный оператор, а затем спроектируем первое неравенство на пространство $\overline{Y}'$ и заменим второе неравенство на $\overline{x} \ge 0$.

Теорема 3. *При естественной аппроксимации задача, двойственная к задаче, аппроксимирующей прямую задачу, аппроксимирует двойственную задачу.*

Доказательство. Повторим приведенные выше рассуждения для двойственной задачи (4). Будем искать ее решения в виде $y = S_Y \overline{y}$. Из

$$< S_Y \overline{y}, y'_0 > \to \min, \quad A' S_Y \overline{y} \ge x'_0, \quad S_Y \overline{y} \ge 0$$

следует, что

$$< \overline{y}, S'_Y y'_0 > \to \min, \quad T_{X'} A' S_Y \overline{y} \ge T_{X'} x'_0, \quad \overline{y} \ge 0. \tag{8}$$

Легко видеть, что задачи (7) и (8) взаимно двойственны. •

Построим теперь задачи, аппроксимирующие задачи (5) и (6). Пусть $\tau_j \in [\alpha, \beta],\ j = 1..N$ и $t_k \in [\gamma, \delta],\ k = 1..M$ — узлы аппроксимации. Интегралы заменим на суммы

$$\int\limits_\alpha^\beta f(\tau)\, d\tau \approx \sum_{j=1}^N r_j\, f(\tau_j), \quad \int\limits_\gamma^\delta g(t)\, dt \approx \sum_{k=1}^M s_k\, g(t_k),$$

здесь r_j и s_k — (неотрицательные) коэффициенты квадратурных формул. Обозначим $x_j = r_j\, x(\tau_j)$, $y_k = s_k\, y(t_k)$. Тогда из (5) следует, что

$$\sum_{j=1}^{N} c(\tau_j)\, x_j \to \max,$$

$$\sum_{j=1}^{N} A(t_k,\tau_j)\, x_j \le b(t_k), \quad k = 1..M, \quad x_j \ge 0, \quad j = 1..N$$

и из (6) следует, что

$$\sum_{k=1}^{M} b(t_k)\, y_k \to \min,$$

$$\sum_{k=1}^{M} A(t_k,\tau_j)\, y_k \ge c(\tau_j), \quad j = 1..N, \quad y_k \ge 0, \quad k = 1..M.$$

В классическом случае допустимое множество представляет собой выпуклое многогранное множество в конечномерном пространстве. Множества элементов, удовлетворяющих уравнениям $Ax = y'_0$ и $A'y = x'_0$ легко можно описать в случае канонической задачи. Аналогичная ситуация имеет место и в бесконечномерном случае.

Пусть X, X' и Y, Y' — двойственные пары линейных пространств. Линейный оператор $K : X \to Y'$ назовем конечномерным, если существуют линейно независимые элементы $x'_1, \ldots, x'_n$ в пространстве X' и линейно независимые элементы $y'_1, \ldots, y'_n$ в пространстве Y' такие, что

$$K : x \mapsto \sum_{j=1}^{n} < x, x'_j > \, y'_j \quad \forall\, x \in X.$$

Легко доказать, что любой конечномерный линейный оператор имеет двойственный оператор.

Пусть $Y' = X$ и $X' = Y$. Через

$$T : x \mapsto \sum_{j=1}^{n} < x, x'_j > \, x_j, \qquad T' : x' \mapsto \sum_{j=1}^{n} < x_j, x' > \, x'_j$$

обозначим линейные конечномерные операторы, образующие двойственную пару. Если $A = I - T$, то решение уравнения $x - Tx = x_0$ имеет вид

$$x = x_0 + \sum_{j=1}^{n} \alpha_j x_j,$$

где числа α_j удовлетворяют системе уравнений

$$\alpha_k - \sum_{j=1}^{n} \alpha_j < x_j, x'_k >=< x_0, x'_k >, \quad k = 1 \,.\,.\, n.$$

Двойственное уравнение $x' - T'x' = x'_0$ имеет аналогичное свойство. Поэтому в таком случае задачи линейного программирования могут быть сведены к конечномерным задачам.

* * *

Последний раздел учебного пособия написан по работе [18].

Литература

[1] Алексеев В.М., Тихомиров В.М., Фомин С.В. Оптимальное управление. – М.: Наука, 1979. – 432 с.

[2] Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. – М.: Наука, 1972. – 496 с.

[3] Васильев Ф.П. Методы решения экстремальных задач. – М.: Наука, 1981. – 400 с.

[4] Галеев Э.М., Тихомиров В.М. Краткий курс теории экстремальных задач. – М.: Изд-во Моск. ун-та, 1989. – 204 с.

[5] Матвеев А.С., Якубович В.А. Оптимальные системы управления: Обыкновенные дифференциальные уравнения. Специальные задачи. – СПб.: Изд-во С-П. ун-та, 2003. – 540 с.

[6] Серазетдинов Т.К. Оптимальные системы с распределенными параметрами. – М.: Наука, 1977. – 480 с.

[7] Галеев Э.М. Курс лекций по вариационному исчислению и оптимальному управлению. – М.: Изд-во мехмата МГУ, 1996. – 160 с.

[8] Лаврентьев М.А., Люстерник Л.А. Курс вариационного исчисления. – М.-Л.: ГИТТЛ, 1950. – 296 с.

[9] Коша А. Вариационное исчисление. – М.: Высш. шк., 1983. – 280 с.

[10] Эльсгольц Л.Э. Дифференциальные уравнения и вариационное исчисление. – М.: Наука, 1969. – 424 с.

[11] Карташев А.П., Рождественский Б.Л. Обыкновенные дифференциальные уравнения и основы вариационного исчисления. – М.: Наука, 1979. – 288 с.

[12] Буслаев В.С. Вариационное исчисление. – Л.: Изд-во Ленингр. ун-та, 1980. – 288 с.

[13] Pleshchinskii N.B. Zur Optimierung einer Klasse von Integral-functionalen im Raum der stetigen Funktionen // Math. Nachr. – 1988. – 136. – С.69-79.

[14] Плещинский Н.Б. Об одном аналоге задачи Больца // Вестник МГУ. Сер.15. – 1989. – №2. – С.18-23.

[15] Плещинский Н.Б. Интегральные уравнения как необходимое условие экстремума интегральных функционалов // Вестник МГУ. Сер. 15. – 1988. – №4. – С.63-64.

[16] Васильев Ф.П. Численные методы решения экстремальных задач. – М.: Наука, 1980. – 520 с.

[17] Плещинский Н.Б. Абстрактные приближенные схемы (специальный курс): учебное пособие. – Казань: Казанск. ун-т, 2012. – 80 с.

[18] Pleshchinskii N.B. The infinite-dimensional linear programming problems and their approximation // In "Linear Programming – New Frontiers in Theory and Applications Ed. Zoltan Mann. – Nova Science: New York, USA, 2011. – P.121-132.

Содержание