Федеральное агентство по образованию

АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ГОУВПО «АмГУ»

Факультет математики и информатики

УТВЕРЖДАЮ

Зав. кафедрой МАиМ

________Т.В. Труфанова

7 мая 2007г.

МЕТОДЫ ОПТИМИЗАЦИИ

Учебно – методический

комплекс дисциплины

для специальности

010501 –прикладная математика и информатика

Составитель: **доцент В.В.Сельвинский**

Благовещенск

2007

ББК
С

Печатается по решению редакционно-издательского совета факультета математики и информатики Амурского государственного университета

В.В.Сельвинский

Методы оптимизации. Учебно – методический комплекс дисциплины для студентов АмГУ очной формы обучения специальности 010501 «Прикладная математика и информатика».– Благовещенск: Амурский гос. ун–т, 2007. − 90 с.

Учебно – методический комплекс дисциплины " Методы оптимизации" содержит рабочую программу дисциплины, план-конспект лекций, материалы для проведения практических занятий, контролирующие материалы для осуществления промежуточного и итогового контроля, справочный материал и библиографический список. Предназначен ведущим преподавателям и студентам, изучающим данную дисциплину.

1. ВЫПИСКА ИЗ ГОСУДАРСТВЕННОГО ОБРАЗОВАТЕЛЬНОГО СТАНДАРТА ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

Специальность 010501 – «Прикладная математика и информатика»

Квалификация – Системный программист

ОПДФ.08. Методы оптимизации: элементы выпуклого анализа, численные методы математического программирования, оптимальное управление, вариационное исчисление.

2.РАБОЧАЯ ПРОГРАММА

РАБОЧАЯ ПРОГРАММА

по дисциплине **«Методы оптимизации»**

для специальности 010501–«Прикладная математика и информатика»

Курс 4 Семестр 7

Лекции 32 час. Экзамен 7 семестр

Практические (семинарские) занятия 32 час. Зачет (нет)

Курсовая работа 7 семестр

Лабораторные занятия (нет)

Самостоятельная работа 38 час.

Всего 102 часа

Составитель: В.В.Сельвинский к.ф-м.н., доцент
Факультет математики и информатики.

Кафедра математического анализа и моделирования.

1. ЦЕЛИ И ЗАДАЧИ ДИСЦИПЛИНЫ, ЕЁ МЕСТО В УЧЕБНОМ ПРОЦЕССЕ

1.1 Цель преподавания учебной дисциплины

Дисциплина "Методы оптимизации" изучает математические модели естественнонаучных явлений, которые приводят к задачам отыскания экстремальных значений функционалов при заданных ограничениях на множестве допустимых решений.

Целью дисциплины является знакомство с методами исследования математических моделей различных процессов и явлений естествознания, изучение основных методов решения возникающих при этом математических задач, выяснение смысла полученного решения.

1.2. Перечень основных умений и навыков, приобретаемых при изучении дисциплины.

Дисциплина "Методы оптимизации" вырабатывает у студентов навыки построения математических моделей простейших физических явлений и решения (аналитического и численного) получающихся при этом математических задач. Студент должен свободно ориентироваться в основных разделах дисциплины, что включает:

классическое вариационное исчисление, уравнение Эйлера, условия второго порядка - Лежандра, Якоби; оптимальное управление, принцип максимума Понтрягина, методы решения задач линейного программирования, симплекс-метод, градиентные методы, метод Ньютона, методы сопряженных направлений.

1.3. Перечень дисциплин, необходимых для изучения данной дисциплины.

Дисциплина "Методы оптимизации" излагается на базе математического анализа, алгебры и аналитической геометрии, дифференциальных уравнений, в тесной связи с теорией функций комплексного переменного и с основами функционального анализа.

2. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

2.1. Федеральный компонент.

ОПДФ.08. Методы оптимизации: элементы выпуклого анализа, численные методы математического программирования, оптимальное управление, вариационное исчисление.

2.2. Наименование тем, их содержание, объем в часах лекционных занятий

Лекции – 32 часа

1. Предмет методов оптимизации. Общая структура задач оптимизации. Постановка транспортной задачи. Постановка задачи о рационе. – 2 часа

2. Выпуклые множества. Выпуклая комбинация точек. Проекция точки на множество. Теоремы отделимости. Теорема об опорной гиперплоскости. Теорема о разделяющей гиперплоскости. – 2 часа

3. Угловая точка множества. Теорема о представлении любой точки множества через его угловые точки. Конус. Теорема Фаркаша. Ребро конуса. Теорема о представлении любой точки конуса через его ребра. – 2 часа

4. Выпуклые функции. Выпуклость квадратичной функции. Свойства выпуклых функций. Неравенство Иенсена. Условие выпуклости дифференцируемой функции. – 2 часа

5. Возможные направления. Активные ограничения. Теорема об условии того, что данное направление является возможным. Свойство возможного направления. Экстремальные свойства. Система линейных неравенств относительно переменных $\boldsymbol{s}, \sigma$. Необходимое условие локального минимума целевой функции в данной точке в переменных $\boldsymbol{s}, \sigma$. Необходимое условие локального минимума целевой функции в случае линейной независимости градиентов активных ограничений. Условие дополняющей нежесткости. – 4 часа

6. Экстремальные свойства на выпуклых множествах. Условие регулярности Слейтера. Необходимое условие локального минимума целевой функции на множестве, регулярном по Слейтеру. Достаточное условие глобального минимума целевой функции. Теорема Куна-Таккера. – 2 часа

7. Седловая точка. Условия существования седловой точки. Функция Лагранжа. Достаточные условия оптимальности. Теорема Куна-Таккера о седловой точке функции Лагранжа. – 2 часа

8. Основная задача линейного программирования. Канонический вид задачи линейного программирования. Графический метод решения задачи линейного программирования. Симплекс-метод решения задачи линейного программирования. Метод искусственного базиса отыскания начальной угловой точки. – 2 часа

9. Элементы вариационного исчисления. Предмет вариационного исчисления. Вариация и ее свойства. Необходимое условие экстремума функционала. Сильный и слабый экстремум. – 2 часа

10. Метод вариаций в задачах с неподвижными границами. Основная лемма вариационного исчисления. Уравнение Эйлера. Частные случаи интегрируемости уравнения Эйлера. Задача о наименьшей поверхности вращения. Задача о брахистохроне. Функционалы, зависящие от производных более высокого порядка. Уравнение Эйлера-Пуассона. Функционалы, зависящие от функций нескольких независимых переменных. Уравнение Остроградского. – 2 часа

11. Вариационные задачи с подвижными границами. Простейшая задача с подвижными границами. Условие трансверсальности. – 2 часа

12. Поле экстремалей. Условие Якоби включения экстремали в поле экстремалей. Функция Вейерштрасса E(x,y,p,y'). Условие сильного и слабого экстремума для функционала. Условие Лежандра. – 2 часа

13. Вариационные задачи на условный экстремум. Случай геометрических связей. Случай неголономных связей. Изопериметрические задачи. – 4 часа

14. Постановка и методы решения задачи оптимального управления. Принцип максимума Понтрягина. – 2 часа.

2.3. Практические занятия, их содержание и объем в часах - 32 часа

1.Выпуклые множества – 2 часа

2.Выпуклые функции – 2 часа

3.Графический метод решения задач линейного программирования – 2 часа

4.Симплекс-метод решения задачи линейного программирования. – 2 часа

5.Метод искусственного базиса отыскания начальной угловой точки – 2 часа

6. Уравнение Эйлера. Частные случаи интегрируемости уравнения Эйлера. – 2 часа

7. Функционалы, зависящие от производных более высокого порядка. Уравнение Эйлера-Пуассона. – 2 часа

8. Функционалы, зависящие от нескольких независимых функций одной переменной. – 2 часа

9. Функционалы, зависящие от функций нескольких независимых переменных. Уравнение Остроградского. – 2 часа

10. Простейшая задача с подвижными границами. Условие трансверсальности. – 2 часа

11.Задача Больца

12. Достаточные условия экстремума функционалов. Условие Лежандра. – 2 часа

13. Вариационные задачи на условный экстремум. Случай геометрических связей. Случай неголономных связей.

14. Изопериметрические задачи. – 2 часа

15. Постановка и методы решения задачи оптимального управления. Принцип максимума Понтрягина. – 2 часа

16. Контрольная работа – 2 часа

2.4. Самостоятельная работа студентов – 38 часов

1. Курсовая работа.
2. Решение практических задач.

Темы для курсовых работ:

1. Исследование метода проекций градиента для решения условной задачи нелинейного программирования.
2. Исследование метода проекций градиента для решения нелинейной задачи с линейными ограничениями.
3. Использование метода сепарабельного программирования для поиска точек минимума (максимума) в задачах оптимизации нелинейного типа.
4. Исследование обобщенного метода приведенного градиента для поиска экстремальных точек в общих задачах нелинейного программирования.
5. Использование метода допустимых направлений Топкинса-Вейнота в нелинейных задачах условной оптимизации.
6. Методы штрафов (внутренние и внешние) для поиска экстремальных точек в задачах нелинейного программирования.
7. Исследование метода переменной метрики в задачах условной оптимизации.
8. Задачи геометрического программирования и методы их решения.
9. Исследование метода параметризации целевой функции задач математического программирования.

10. Исследование функций многих переменных на экстремум с помощью симплексного метода.

11. Использование выпуклого симплекс-метода для поиска решений в нелинейных условных задачах.

12. Использование конечного метода в задачах квадратичного программирования для поиска точек оптимума.

13. Исследование алгоритма с пошаговой комбинацией метода штрафов и субградиентного метода для условной задачи оптимизации.

14. Исследовать метод отсекающих плоскостей Келли (метод линеаризации) в общих задачах нелинейного программирования.

15. Исследование функций многих переменных на экстремум с помощью методов, использующих производные.

16. Использование квадратичной аппроксимации функции Лагранжа в задачах условной оптимизации.

17. Исследование методов, основанных на факторизации Холецкого для задач с квадратичной функцией без ограничений (квадратичная аппроксимация).

18. Исследование методов минимизации гладких функций без вычисления производных, использующих конечно-разностную аппроксимацию первых производных.

19. Исследование методов решения задач о наименьших квадратах.

20. Исследование методов поиска минимума задач с линейными ограничениями при ограничениях-равенствах (методы наискорейшего спуска, методы вторых производных, дискретные ньютоновские методы, квазиньютоновские методы).

Рекомендации по выполнению курсовой работы

Курсовая работа представляет собой теоретическое исследование или реферат по конкретной проблеме. Выполнение курсовой работы должно способствовать углубленному усвоению лекционного курса и приобретению навыков научного исследования. Курсовая работа базируется на изучении учебно-методических материалов, научной литературы.

Цель курсовой работы:

- приобрести навыки самостоятельной работы с литературой;

- научиться планировать свою работу и время, уметь анализировать по лученные результаты и правильно оформлять отчет по выполненной работе;

- приобрести умение кратко и четко докладывать результаты проделанной работы на отчетных студенческих научно-исследовательских конференциях, семинарах, заседаниях кафедры.

Написание курсовой работы осуществляется под руководством преподавателя- руководителя работы. Студент совместно с руководителем уточняет круг вопросов, подлежащих изучению, составляет план исследования, структуру работы, сроки выполнения ее этапов, определяет необходимую литературу и другие материалы (статистические отчеты, результаты экспериментов на предприятиях и т.п.).

Курсовая работа может иметь исследовательский или реферативный характер и должна включать в себя: обоснование актуальности темы и постановку задачи; краткий обзор литературы по изучаемому вопросу; описание метода решения задачи и полученные результаты, если таковые имеются; выводы и рекомендации по использованию полученных результатов, если таковые имеются.

Курсовая работа оформляется в соответствии со стандартом АмГУ (СТП АмГУ–05-97) и должна иметь следующую структуру:

- титульный лист;

- задание на выполнение работы;

- реферат;
- содержание;
- введение;
- основная часть;
- заключение;
- список использованных источников;
- приложения.

Объем курсовой работы зависит от предложенной темы и должен составлять не более 30 страниц, не включая приложения. Изложение должно достаточно подробно освещать самостоятельную работу студента. В связи с этим необходимо каждый раз делать ссылки на соответствующую литературу, в которой получен представленный в курсовой работе качественный вывод или результат. В заключении необходимо выделить самостоятельную часть работы.

При оценке работы учитываются содержание работы, ее актуальность, степень самостоятельности, оригинальность выводов и предложений, качество используемого материала, а также уровень грамотности (общий и специальный).

Курсовая работа должна быть защищена до сдачи экзамена. На защите студент должен кратко изложить содержание работы, дать исчерпывающие ответы на замечания рецензента и вопросы членов комиссии. Окончательная оценка курсовой работы выставляется комиссией по итогам защиты и качеству выполненной работы.

2.5. Вопросы к экзамену

1.Предмет методов оптимизации. Общая структура задач оптимизации.

2.Выпуклые множества. Выпуклая комбинация точек. Проекция точки на множество. Теоремы отделимости.

3.Теорема об опорной гиперплоскости.

4.Теорема о разделяющей гиперплоскости.

5.Угловая точка множества. Теорема о представлении любой точки множества через его угловые точки.

6.Конус. Теорема Фаркаша.

7.Ребро конуса. Теорема о представлении любой точки конуса через его ребра.

8.Выпуклые функции. Выпуклость квадратичной функции. Свойства выпуклых функций.

9.Неравенство Иенсена.

10.Условие выпуклости дифференцируемой функции.

11.Возможные направления. Активные ограничения. Теорема об условии того, что данное направление является возможным.

12.Система линейных неравенств относительно переменных $\boldsymbol{s}, \sigma$. Необходимое условие локального минимума целевой функции в данной точке в переменных $\boldsymbol{s}, \sigma$.

13.Необходимое условие локального минимума целевой функции в случае линейной независимости градиентов активных ограничений. Условие дополняющей нежесткости.

14.Условие регулярности Слейтера. Необходимое условие локального минимума целевой функции на множестве, регулярном по Слейтеру.

15. Достаточное условие глобального минимума целевой функции. Теорема Куна-Таккера.

16.Седловая точка. Условия существования седловой точки.

17.Функция Лагранжа. Достаточные условия оптимальности. Теорема Куна-Таккера о седловой точке функции Лагранжа.

18.Основная задача линейного программирования. Канонический вид задачи линейного программирования.

19.Графический метод решения задачи линейного программирования.
20.Симплекс-метод решения задачи линейного программирования.

21.Метод искусственного базиса отыскания начальной угловой точки.

22. Предмет вариационного исчисления. Вариация и ее свойства.

23. Необходимое условие экстремума функционала. Сильный и слабый экстремум.

24.Метод вариаций в задачах с неподвижными границами. Основная лемма вариационного исчисления.

25.Уравнение Эйлера.

26.Частные случаи интегрируемости уравнения Эйлера.

27.Задача о наименьшей поверхности вращения.

28.Задача о брахистохроне.

29.Функционалы, зависящие от производных более высокого порядка. Уравнение Эйлера-Пуассона.

30.Функционалы, зависящие от функций нескольких независимых переменных. Уравнение Остроградского.

31.Вариационные задачи с подвижными границами. Простейшая задача с подвижными границами. Условие трансверсальности.

32. Поле экстремалей. Условие Якоби включения экстремали в поле экстремалей.

33.Функция Вейерштрасса $E(x,y,p,y')$. Условие сильного и слабого экстремума для функционала.

34.Условие Лежандра.

35.Вариационные задачи на условный экстремум. Случай геометрических связей.

36. Вариационные задачи на условный экстремум. Случай неголономных связей.

37.Изопериметрические задачи.

38.Постановка и методы решения задачи оптимального управления.

39.Принцип максимума Понтрягина.

3. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

3.1. Основная литература

1.Карманов В.Г.; Математическое программирование: Учеб.пособие.-5-е изд. -М:Наука, 2001г.

2.Пантелеев А.В., Летова Т.А. Методы оптимизации в примерах и задачах.- М.: Высш.шк., 2002.- 544 с. (АмГУ)

3.Васильев О.В., Аргучинцев А.В. Методы оптимизации в задачах и упражнениях.- М.: Физматлит, 1999.- 208 с. (АмГУ)

3.2. Дополнительная литература

1.Сухарев А.Г., Тимохов А.В., Федоров В.В. Курс методов оптимизации. -М.: Наука, 1986.- 328 с.

2.Моисеев Н.Н., Иванилов Ю.П., Столяров Е.М.; Методы оптимизации. -М.:Наука;1978г.

3. Реклейтис Г., Рейвиндран А., Рэгсдел К. Оптимизация в технике. Пер. с англ. В 2-х кн.- М.: Мир, 1986.

4.Дегтярев Ю.И.; Методы оптимизации: Учеб.пособие; -М: Сов.радио,1980г.

3. ГРАФИК САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

В самостоятельную работу студентов входит курсовая работа, подготовка к текущим занятиям и подготовка к экзамену. Тематика и требования к содержанию и оформлению курсовой работы приведены в рабочей программе.

3. МАТЕРИАЛЫ ДЛЯ ЧТЕНИЯ ЛЕКЦИЙ

1 семестр

ВВЕДЕНИЕ

1.1 Задача оптимизации. Примеры математических моделей.

Под оптимизацией понимают процесс выбора наилучшего варианта из всех возможных.

Целевая функция $\varphi(\bar{x}) \to \min$, при $\bar{x} \in X \subset \bar{R}^n$ если $X \in R^n$, то задача называется задачей без условного оптимума.

Если $\varphi(\bar{x})$ линейная, $\varphi(x) = (\bar{c}, \bar{x}) = c_1x_1 + \ldots + c_nx_n$ $X = \{\bar{x} : f_i(\bar{x}) \geq 0, i = \overline{1,m}\}$ $f_i(\bar{x})$ – линейная, то задачу относят к задачам линейного программирования.

Если $\varphi(\bar{x})$ - интегральный функционал, $\varphi(x) \sim [y(x)] = \int_{x_0}^{x_1} F(x, y(x), y`(x))dx$, задача относится к вариационному исчислению.

Выделяют условную и безусловную задачи оптимизации. Безусловная задача заключается в нахождении max и min действительной функции от n действительных переменных и определении соответствующих значений аргумента на некотором множестве σ (n)-мерного пространства. Условные задачи подразумевают при своей формулировке задание некоторых ограничений на множество σ .

Приведем примеры некоторых математических моделей.

Задача о рационе. По заданному ассортименту продуктов при известном содержании в каждом из них питательных веществ известной стоимости продуктов составить рацион, удовлетворяющий необходимым потребностям с минимальными денежными затратами.

Пусть имеется n различных продуктов и m питательных веществ (например жиров, белков, углеводов, витаминов и др.). Обозначим через a_{ij} содержание (в единицах массы) j-го питательного вещества в единице массы i-го продукта; через b_i обозначим минимальную (в единицах массы) суточную потребность в j-м питательном веществе. Наконец, через x_i обозначим искомое суточное потребление i-го продукта. Очевидно, что $x_i \geq 0$.

Величина $\sum_{i=1}^{n} a_{ij}x_i$ есть общее содержание j-го питательного вещества в рационе, которое не должно быть меньше минимальной потребности b_j:

$$\sum_{i=1}^{n} a_{ij}x_j \geq b_j, \; j = \overline{1,m}.$$

Если c_i – стоимость единицы массы i-го продукта, то стоимость всего рациона определяет линейная форма $\sum_{i=1}^{n} c_i x_i$.

Итак, математическая формулировка задачи выбора рациона состоит в следующем:

найти

$$\min \sum_{i=1}^{n} c_i x_i$$

при условиях

$$\sum_{i=1}^{n} a_{ij}x_i \geq b_j, \quad j = \overline{1,m},$$

$$x_i \geq 0, \quad i = \overline{1,n}.$$

Эта задача является одной из типичных задач линейного программирования.

В постановке задачи вовсе не обязательно было указывать, что это задача о рационе. Достаточно ясно, что таким же образом могут быть сформулированы многочисленные задачи об оптимальных смесях (слово «смесь» здесь следует понимать в обобщенном смысле: это и собственно смесь, и сплав, и рацион, и т.д.).

Транспортная задача. Другим типичным примером задачи линейного программирования является транспортная задача. Требуется составить план перевозок одного груза таким образом, чтобы общая стоимость перевозок была минимальной.

Исходная информация:

a_i – количество единиц груза в i-м пункте отправления ($i = \overline{1,m}$);

b_j – потребность в j-м пункте назначения ($j = \overline{1,n}$) в единицах груза;

c_{ij} – стоимость перевозки единицы груза из i-го пункта в j-й.

Обозначим через x_{ij} планируемое количество единиц груза для перевозки из i-го пункта в j-й.

В принятых обозначениях:

$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$ – общая (суммарная) стоимость перевозок;

$\sum_{j=1}^{n} x_{ij}$ – количество груза, вывозимое из i-го пункта;

$\sum_{i=1}^{m} x_{ij}$ – количество груза, доставляемого в j-й пункт.

В простейшем случае должны выполняться следующие очевидные условия:

$$\sum_{j=1}^{n} x_{ij} = a_i, \quad i = \overline{1,m},$$

$$\sum_{i=1}^{m} x_{ij} = b_j, \quad j = \overline{1,n},$$

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j.$$

Таким образом, математической формулировкой транспортной задачи будет:

найти

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

при условиях

$$\sum_{j=1}^{n} x_{ij} = a_i,$$

$$\sum_{i=1}^{m} x_{ij} = b_i,$$

$$x_{ij} \geq 0, \quad i = \overline{1,m}, \quad j = \overline{1,n}.$$

Эта задача носит название замкнутой транспортной модели.

Заметим, что условие

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$$

является естественным условием разрешимости замкнутой транспортной задачи.

Более общей транспортной задачей является так называемая открытая транспортная модель:

найти

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

при условиях

$$\sum_{j=1}^{n} x_{ij} \leq a_i,$$

$$\sum_{i=1}^{m} x_{ij} = b_j,$$

$$x_{ij} \geq 0, \quad i = \overline{1,m}, \quad j = \overline{1,n}.$$

Ясно, что в этой задаче не предполагается, что весь груз, накопленный в i-м пункте, должен быть вывезен.

В схему транспортной задачи укладываются и некоторые другие задачи технико-экономического содержания, например, так называемая задача о выборе: задача о наиболее экономном (в смысле суммарных затрат времени) распределений n работ между m исполнителями при известном времени, затрачиваемом каждым исполнителем на каждой работе. Эта задача является частной моделью замкнутой транспортной задачи при $m = n$ и $a_i = b_j = 1$.

Заметим, что решения транспортных задач обладают свойствами целочисленности при целочисленных значениях величин a_i, b_j, и поэтому эти задачи относятся к задачам линейного программирования.

Задача о режиме работы энергосистемы. В качестве примера задачи выпуклого программирования рассмотрим простейшую среди задач об оптимальном ведении режима работы энергосистемы.

Рассматривается изолированная энергосистем, состоящая из теплоэлектростанций, связанных линиями передач с узлом, в котором сосредоточена нагрузка. Ставится задача распределения активных мощностей между электростанциями в заданный момент времени. Распределение осуществляется по критерию минимизации суммарных топливных затрат на генерацию активной мощности.

Обозначим через x_i активную мощность, генерируемую на i-й станции. Мощности x_i заключены в пределах α_i и β_i, определяемых техническими условиями: $\alpha_i \le x_i \le \beta_i$. Кроме того, должно соблюдаться условие баланса мощностей, т.е. генерируемая общая мощность должна соответствовать потребляемой мощности P с учетом общих потерь π в линиях передач:

$$\sum_{i=1}^{m} x_i = P + \pi .$$

Топливные затраты на генерацию мощности x_i представляют собой функцию $T_i(x_i)$, выпуклую на отрезке $[\alpha_i, \beta_i]$.

Таким образом, задача принимает вид:

найти

$$\min \sum_{i=1}^{m} T_i(x_i)$$

при условиях

$$\sum_{i=1}^{m} x_i = P + \pi,$$
$$\alpha_i \le x_i \le \beta_i, \quad i = \overline{1,m}.$$

Построенная модель является типичной задачей выпуклого программирования с линейными ограничениями. Решение этой задачи дает весьма грубое приближение к действительно оптимальному режиму работы энергосистемы. В реальной ситуации нельзя считать всю нагрузку сосредоточенной в одном узле, а следует рассматривать n узлов. Кроме того, потери в системе, естественно, не являются константой, а зависят от величин передаваемых мощностей и параметров линий передач.

В качестве следующего приближения можно рассматривать задачу, в которой π является билинейной функцией $x_{ij}\left(i = \overline{1,m}, j = \overline{1,n}\right)$, где параметры управления x_{ij} означают количество активной мощности, передаваемое из i-й станции в j-й узел.

Очевидно, что в этой новой модели условия будут содержать нелинейности ($\pi\left(x_{ij}\right)$ в уравнении баланса).

Эта задача также является задачей выпуклого программирования, но более сложного типа, чем предыдущая.

Примером многоэкстремальной задачи является простейшая задача о размещении.

Задача о размещении. Даны m пунктов потребления $(1,2,...,j,...,m)$ с заданным объемом потребления b_j в каждом пункте. Имеются n возможных пунктов производства $(1,2,...,i,...,n)$, причем для каждого i-го пункта известна зависимость стоимости производства f_i от объема производства x_i. (Предполагается, что в стоимость производства $f_i(x_i)$ включены капитальные затраты.) наконец, задана матрица транспортных расходов a_{ij} (a_{ij} – стоимость перевозки единицы продукции из i-го пункта производства в j-й пункт потребления). Требуется найти такие объемы перевозок x_{ij} из i-го в j-й пункт и такие объемы производства $x_i = \sum_j x_{ij}$, которые минимизируют суммарные расходы; иначе говоря, ищется

$$\min\left[L\left(x_{ij}\right) = \sum_{i,j} a_{ij}x_{ij} + \sum_i f_i\left(x_i\right)\right]$$

при условиях

$$\sum_i x_{ij} = b_j,\quad x_{ij} \geq 0.$$

Поскольку себестоимость единицы продукции обычно убывает при увеличении объема производства, то функции $f_i\left(x_i\right)$, как правило, монотонно возрастают и выпуклы вверх. Множество значений x_{ij}, удовлетворяющих ограничениям задачи, образует выпуклый многогранник, вершины которого являются точками локальных минимумов функции $L\left(x_{ij}\right)$ (рис. 1). Отсюда и название подобных задач – многоэкстремальные.

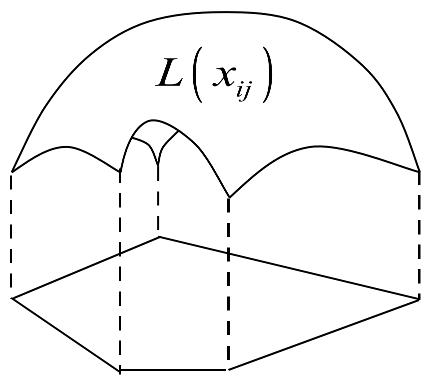

Рис. 1

2. ЭЛЕМЕНТЫВЫПУКЛОГО АНАЛИЗА

2.1. Евклидово пространство. Выпуклые множества

2.1.1. Мы будем иметь дело с функциями, определенными на множествах конечномерного евклидова пространства E_n.

Совокупность всех наборов $x^T = (x_1, x_2, ..., x_n)$ n вещественных чисел $(x_1, x_2, ..., x_n)$ называют *евклидовым пространством* размерности n, если выполняются следующие условия. Пусть $x \in E_n$, $y \in E$ и α - вещественное число. Тогда

$x + y = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)$ (сложение),

$\alpha x = (\alpha x_1, \alpha x_2, ..., \alpha x_n)$ (умножение на число),

$\langle x, y\rangle = \sum\limits_{i=1}^{n} x_i y_i$ (скалярное произведение).

Наборы $x^T = (x_1, x_2, ..., x_n)$ называют *точками (векторами),* а числа $(x_1, x_2, ..., x_n)$ - их *координатами.*

В евклидовом пространстве введено понятие *евклидовой нормы* (длины вектора)

$$\|x\| = \langle x, x\rangle^{1/2},$$

для которой справедливы следующие соотношения:

$$\|x\| \geq 0,$$

причем $\|x\| = 0$ тогда и только тогда, когда $x = 0$;

$$\|\alpha x\| = |\alpha| \|x\|$$

$$\|x + y\| \leq \|x\| + \|y\|$$

Евклидова норма и скалярное произведение связаны между собой неравенством Коши - Буняковского

$$\langle x, y\rangle \leq \|x\| \cdot \|y\|$$

Евклидова норма порождает в E_n сходимость. Будем говорить, что *последовательность* $\{x_m\}$ точек из E_n *сходится к точке* x при $m \to \infty$, т. е. $\lim\limits_{m\to\infty} x_m = x$, если

$$\lim_{m\to\infty} \|x_m - x\| = 0$$

Точка x называется *предельной точкой* последовательности.

Приведем еще несколько определений.

2.1.2. Множество

$U_\varepsilon(x) = \{y : \|y - x\| \le \varepsilon\}$

будем называть ε *-окрестностью* точки x.

2.1.3. Множество $X \subseteq E_n$ называют *замкнутым,* если оно содержит все свои предельные точки, т. е. такие точки, что любой окрестности каждой из них принадлежит бесконечно много точек из X.

2.1.4. Точка $x \in X$ называется *внутренней точкой* множества, X если существует такая ее окрестность, все точки которой принадлежат множеству X.

2.1.5. Точка $x \in X$ называется *граничной точкой* множества, X если в любой ее окрестности содержатся как точки, принадлежащие множеству X, так и точки, не принадлежащие этому множеству. Множество $G(X)$, состоящее из всех граничных точек множества X, называется *границей* множества X.

2.1.6. Множество X n-мерного евклидова пространства E_n называется *выпуклым,* если вместе с любыми двумя точками $x \in X$ и $y \in X$ ему принадлежит и соединяющий их отрезок $[x, y]$

Выпуклость множества X означает, что из $x, y \in X$ следует $z = \alpha x + (1 - \alpha)y \in X$ для всех $0 \le \alpha \le 1$, например, выпуклы отрезок, полупрямая, прямая, круг, треугольник, полуплоскость и вся плоскость.

Легко видеть, что если множество X задается системой линейных равенств и неравенств:

$X = \{x : Ax \ge a, Bx = b\}$,

где A, B — матрицы, то оно выпукло и замкнуто (последнее в силу линейности и, следовательно, непрерывности преобразований A и B). Читателю предоставляется возможность самостоятельно убедиться в том, что пересечение любого числа выпуклых множеств выпукло.

2.1.7. Точка z называется *выпуклой комбинацией* точек $x_1, x_2, ..., x_m$, если $z = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_m x_m, \alpha_i \geq 0 \ \ (i = \overline{1,m})$ и $\alpha_1 + \alpha_2 + ... + \alpha_m = 1$.

2.1.8. *Выпуклое множество X содержит все выпуклые комбинации своих точек.*

Доказательство (по индукции). В соответствии с определением 2.1.6 множество X содержит выпуклую комбинацию любых двух своих точек. Предположим, что X содержит выпуклые комбинации любых $m - 1 (m > 1)$ своих точек. Рассмотрим выпуклую комбинацию $z = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_m x_m$ любых m точек из X. Очевидно, что существует хотя бы один номер i такой, что $\alpha_i < 1$. Не умаляя общности, можно полагать, что $\alpha_1 < 1$. Тогда $y = \beta_2 x_2 + ... + \beta_m x_m$ при $\beta_i = \dfrac{\alpha_i}{1 - \alpha_1}$ является выпуклой комбинацией точек $x_2, ..., x_m$, и по индуктивному предположению $y \in X$. Поскольку $z = \alpha_1 x_1 + (1 - \alpha_1) y$ и $\alpha_1 \in [0,1]$, то по п. 2.1.6 из выпуклости X следует, что $x \in X$. Δ

2.2 Проекция. Теоремы отделимости

2.2.1 Проекция точки v на множество X называют такую точку $P_X(v) \triangleq p \in X$, что

$$\|p - v\| = \inf_{x \in X} \|x - v\| \triangleq \rho(v, X) = \rho \tag{2.1}$$

При этом $\rho = \rho(v, X)$ называют расстоянием от точки v до множества X.

2.2.2 Для любого замкнутого множества X и любой точки v существует точка $p \in X$, являющаяся проекцией v на X. Если, кроме того, множество X выпуклое, то точка p единственная.

Доказательство. Если $v \in X$, то очевидно, что $p = v$ и $\rho = 0$. Пусть точка v внешняя относительно $X : v \notin X$. Согласно определению нижней грани существует последовательность $\{x_k\}$, $x_k \in X$, такая, что

$$\lim_{k\to\infty}\|x_k - v\| = \rho .$$

Так как $\{x_k\}$ ограничена, то существует подпоследовательность $\{x_{k_i}\}$ такая, что

$$\lim_{i\to\infty} x_{k_i} = p .$$

Поскольку множество X выпуклое, то точка $z = \frac{1}{2}p' + \frac{1}{2}p''$ принадлежит X. Точки v, p', p'' и z лежат в одной плоскости, и из равнобедренности треугольника с вершиной в точке v, основанием $[p', p'']$ и высотой $[v, z]$ следует $\|z - v\| < \rho$, что противоречит определению ρ. Δ

2.2.3 Теорема. Для того чтобы точка $p \in X$ была проекцией точки v на выпуклое множество X, необходимо и достаточно, чтобы для всех $x \in X$ выполнялось неравенство

$$\langle x - p, v - p\rangle \le 0 . \qquad (2.2)$$

Доказательство. Пусть p - проекция точки v на X. Возьмем произвольную точку $x \in X$ и рассмотрим $z = \alpha x + (1-\alpha)p$. В силу выпуклости X для любого $\alpha \in [0,1]$ точка z принадлежит X. Так как

$$\|z - v\|^2 = \alpha^2\|x - p\|^2 + 2\alpha\langle x - p, p - v\rangle + \|p - v\|^2 ,$$

$$\|z - v\|^2 \ge \|p - v\|^2 ,$$

то

$$\alpha^2\|x - p\|^2 + 2\alpha\langle x - p, p - v\rangle \ge 0.$$

Поскольку это неравенство справедливо для всех $\alpha \in [0,1]$, то

$$\langle x - p, p - v\rangle \ge 0 ,$$

откуда следует (2.2).

Пусть теперь справедливо (2.2). Тогда для любого $x \in X$ будет

$$\|x - v\|^2 = \|(x - p) + (p - v)\|^2 = \|x - p\|^2 + 2\langle x - p, p - v\rangle + \|p - v\|^2 \ge \|x - p\|^2 ,$$

т.е. p является проекцией v на X. Δ

2.2.4 1) Для любого $x \in X$ выполняются соотношения

$$(x - v, x - p) = (x - p, x - p) + (p - v, x - p) \ge \|x - p\|^2$$

и

$\|x-p\| \le \|x-v\|$

2) для любых $y,z \in E_n$ справедливо неравенство

$\|P_X(y)+P_X(z)\| \le \|y-z\|$.

Действительно, 1) вытекает из того, что

$\|x-v\|^2 = \|(x-p)+(p-v)\|^2 = \|x-p\|^2 + 2\langle x-p, p-v\rangle + \|p-v\|^2 \ge \|x-p\|^2$.

Далее , из (2.2) следует $\langle P_X(y)-y, P_X(z)-P_X(y)\rangle \ge 0$ и $\langle P_X(z)-z, P_X(y)-P_X(z)\rangle \ge 0$. Суммируя эти неравенства, получаем $\langle P_X(y)-y-P_X(z)+z, P_X(z)-P_X(y)\rangle \ge 0$, откуда, пользуясь неравенством Коши - Буняковского, получаем условие 2):

$\|P_X(y)-P_X(z)\|^2 \le \langle P_X(z)-P_X(y), z-y\rangle \le \|P_X(z)-P_X(y)\|\|z-y\|$. Δ

2.2.5 Гиперплоскостью в E_n называют множество вида

$\Pi = \{x : \langle c,x\rangle = \lambda\}$,

где $c \ne 0$. В пространстве E_n гиперплоскость определяет два полупространства:

$\{x : \langle c,x\rangle \le \lambda\}$, $\{x : \langle c,x\rangle \ge \lambda\}$.

2.2.6 Теорема отделимости. Для любого выпуклого и замкнутого множества X и любой точки v, не принадлежащей множеству X, существует такая гиперплоскость Π, что

$$\langle c,v\rangle = \lambda \tag{2.3}$$

и для всех $x \in X$

$$\langle c,x\rangle < \lambda \tag{2.4}$$

Доказательство. Пусть p - проекция точки v на X. Рассмотрим гиперплоскость

$\Pi = \{x : \langle c,x\rangle = \lambda,\ c = v,\ \lambda = \langle c,v\rangle\}$,

для которой выполняется (2.3) из неравенства (2.2) следует

$\langle x, v-p\rangle \le \langle p, v-p\rangle < \langle v, v-p\rangle, \quad x \in X$.

Правое неравенство следует из (2.1) и из того, что $\rho > 0$. И, окончательно,

$$\langle c,x\rangle = \langle v-p,x\rangle < \langle v-p,v\rangle = \langle c,v\rangle = \lambda ,$$

т.е. имеет место (2.4). Δ

Замечание. Очевиден геометрический смысл теоремы: существует проходящая через точку v гиперплоскость Π такая, что X лежит в одном из полупространств, определяемых Π (рис. 2.1).

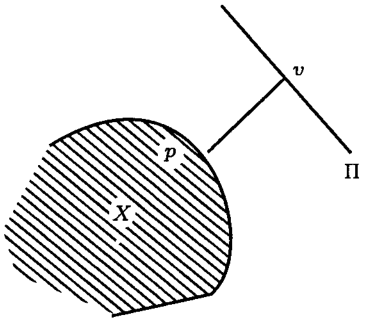

Рис. 2.1

2.2.7 Теорема об опорной гиперплоскости. В любой граничной точке x^0 выпуклого множества X существует опорная гиперплоскость, т.е. существует $c \neq 0$ и λ такие, что $\Pi = \{x : \langle c,x\rangle = \lambda\}$, $\lambda = \langle c,x^0\rangle$ и для всех

$$x \in X \quad \langle c,x\rangle \le \lambda .$$

Доказательство. Рассмотрим последовательность точек $\{v_k\}$, внешних относительно $\overline{X}$ (замыканий X) и таких, что

$$\lim_{k\to\infty} v_k = x^0 .$$

По теореме 2.2.6 для каждой v_k существует

$$\Pi_k = \{x : \langle c_k,x\rangle = \lambda_k\} ,$$

где $\lambda_k = \langle c_k,v_k\rangle$ и $\langle c_k,x\rangle < \lambda_k$ для всех $x \in \overline{X}$. Не умаляя общности моно полагать $\|c_k\| = 1$. Не меняя обозначений, будем считать, что

$$\lim_{k\to\infty} c_k = c .$$

Переходя к пределу в соответствиях, определяющих Π_k, получим

$$\langle c,x^0\rangle = \lim_{k\to\infty}\langle c_k,v_k\rangle \cong \lambda$$

и $\langle c,x\rangle \le \lambda$ для всех $x \in X$.

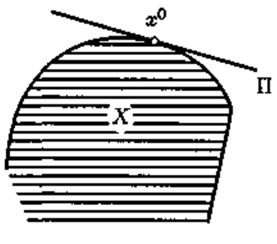

Рис 2

Итак, гиперплоскость

$$П = \{x : \langle c, x\rangle = \lambda\}$$

опорная. ∆

Замечание. Легко убедиться, что если в точке x^0 существует касательная гиперплоскость, то она совпадает с опорной (рис 2.2), и в этом случае опорная гиперплоскость единственна. Однако понятие опорной гиперплоскости значительно шире понятия касательной гиперплоскости. На рис 2.3 изображен случай,

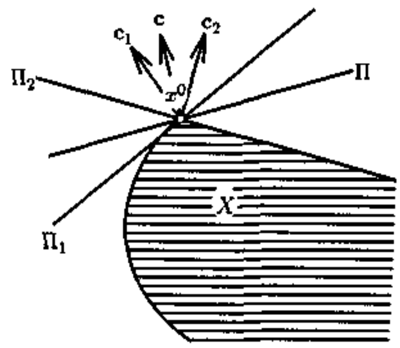

Рис. 3

когда в точке x^0 не существует касательной и в то же время в ней существует опорные прямые, причем в качестве вектора c здесь может быть выбран любой вектор, лежащий «между» c_1 и c_2.

2.2.8 Теорема о разделяющей гиперплоскости. Если множество X_0 внутренних точек выпуклого множества X непусто и не пересекается с выпуклым множеством Y $(X_0 \cap Y = \varnothing)$, то для множеств X и Y существует разделяющая гиперплоскость $П$, т.е. существует вектор $c \neq 0$ такой, что

$$\langle c, y\rangle \le \langle c, x\rangle$$

для всех $y \in Y$ и $x \in X$.

Доказательство. Множество

$$Z = \{z : z = y - x, y \in Y, x \in X_0\}$$

Выпукло, и $z = 0$ не является его внутренней точкой. Тогда из теоремы 2.2.6 и 2.2.7 следует существование $c \neq 0$ такого, что

$$\langle c, z\rangle = \langle c, y - x\rangle \le \langle c, 0\rangle = 0$$

для всех $y \in Y$ и $x \in X_0$. Это неравенство остается справедливым и для всех $y \in Y$ и $x \in X$, поскольку предельный переход не нарушает нестрогих неравенств. ∆

замечание. Обратим внимание на то, что требование непустоты множества X_0 существенно, поскольку в формулировке теоремы имеется в виду внутренность множества относительно пространства E_n. Так, например, очевидно, что в трехмерном пространстве ось z и плоскость $z = 0$ неразделимы в указанном выше смысле, хотя и не имеют общих внутренних точек.

2.2.9 *Определение*. Точка x множества X называется угловой (или крайней) точкой, если в X не существует таких точек x' и x'', $x' \neq x''$, что

$$x = \alpha x' + (1 - \alpha) x''$$

при некотором $\alpha \in (0,1)$.

Например, для круга любая точка ограничивающей его окружности являются угловой. Угловыми точками являются все вершины выпуклого многогранника.

2.2.10 *Теорема (о представлении)*. Любая точка x^0 выпуклого, замкнутого, ограниченного множества X может быть представлена в виде выпуклой комбинации конечного числа угловых точек этого множества.

Доказательство. (по индукции наименьшей размерности n пространства E_n, содержащего множество X).

Если $n = 1$, то X является отрезком, и утверждение теоремы очевидно.

Предположим, что для $n = k - 1$ теорема справедлива. Пусть теперь $X \subset E_k$. Рассмотрим два случая.

1) x^0 - граничная точка X. Построим в этой точке гиперплоскость, опорную к X:

$$\pi = \left\{ x : \langle c, x \rangle = \langle c, x^0 \rangle \right\}.$$

Множество $X_0 = X \cap \pi$ как пересечение выпуклого, замкнутого ограниченного множества X с выпуклым, замкнутым множеством π само выпукло, замкнуто и ограничено и, кроме того, существует $(k - 1)$-мерное подпространство, содержащее X_0 (поскольку $X_0 \subset \pi$). По предположению индукции для $x^0 \in X_0$ найдутся $x_1, x_2, ..., x_N$ - угловые точки множества X_0 такие, что

$$x^0 = \sum_{i=1}^{N} \alpha_i x_i, \qquad \alpha_i \ge 0, \qquad i = \overline{1,N},$$

$$\sum_{i=1}^{N} \alpha_i = 1.$$

Покажем, что $x_1, x_2, ..., x_N$ являются угловыми точками и для X. Предположим противное, т.е. что для некоторой точки x_i найдутся $x', x'' \in X$, $x' \ne x''$, и $\alpha \in (0,1)$ такие, что

$$x_i = \alpha x' + (1-\alpha)x''.$$

Так как $x_i \in X_0 \subset \pi$, то

$$\langle c, x_i \rangle = \langle c, x^0 \rangle,$$

и поскольку гиперплоскость π опорная к X, то

$$\langle c, x' \rangle \le \langle c, x^0 \rangle, \qquad \langle c, x'' \rangle \le \langle c, x^0 \rangle.$$

Из того, что $0 < \alpha < 1$, следует

$$\langle c, x' \rangle = \tfrac{1}{\alpha}\left[\langle c, x_i \rangle - (1-\alpha)\langle c, x'' \rangle\right] \ge \tfrac{1}{\alpha}\left[\langle c, x^0 \rangle - (1-\alpha)\langle c, x^0 \rangle\right] = \langle c, x^0 \rangle.$$

Последние условия показывают, что $x' \in \pi$ (так как $\langle c, x' \rangle = \langle c, x^0 \rangle$); но $x' \in X_0$. А тогда наше предположение противоречит тому, что x_i - угловая точка X_0.

2) Пусть теперь x^0 - внутренняя точка множества X. Проведем через x^0 прямую l. Пересечение $l \cap X$ является отрезком с концами $\bar{x}$ и $\bar{\bar{x}}$, принадлежащими границе множества X, и поскольку x^0 - внутренняя точка для X, то существует $\alpha \in (0,1)$ такое, что

$$x^0 = \alpha \bar{x} + (1-\alpha)\bar{\bar{x}}.$$

Поскольку для граничных точек $\bar{x}$ и $\bar{\bar{x}}$ теорема верна, то она верна и для x^0. Действительно, для граничных точек имеют место соотношения

$$\bar{x} = \sum_{i=1}^{N_1} \beta_i y_i, \qquad \sum_{i=1}^{N_1} \beta_i = 1, \qquad \beta_i \ge 0, \qquad i = \overline{1,N_1},$$

$$\bar{\bar{x}} = \sum_{i=1}^{N_2} \gamma_i z_i, \qquad \sum_{i=1}^{N_2} \gamma_i = 1, \qquad \gamma_i \ge 0, \qquad i = \overline{1,N_2},$$

где все y_i и z_i - угловые точки множества X. Тогда

$$x^0 = \sum_{i=1}^{N_1} \alpha \beta_i \gamma_i + \sum_{i=1}^{N_2} (1-\alpha) \gamma_i z_i,$$

откуда и следует утверждение теоремы. Δ

2.3. Конус. Теорема Фаркаша

2.3.1 Множество K называется конусом, если из $x \in K$ следует $\lambda x \in K$ для всех $\lambda > 0$. Например, все пространство E_n, как и всякое подпространство, является конусом. Неотрицательный ортант $\{x : x \geq 0\}$ - также конус. Очевидно, что конусами являются множества

$$\{x : Ax \leq 0\}, \qquad \{y : y = Ax, x \geq 0\}.$$

Следующий факт будет использован при доказательстве теоремы Фаркаша.

2.3.2 *Теорема*. Множество

$$Y = \{y : y = Ax, x \geq 0\}$$

Замкнуто.

Доказательство. Пусть $A = [a_1, a_2, ..., a_m]$. Докажем утверждение индукцией по числу m. При $m = 1$ множество Y является полупрямой, и, следовательно, оно замкнуто. Предположим, что для $m = k - 1$ конус $\overline{Y}$, порожденный векторами $a_1, a_2, ..., a_{k-1}$, замкнут.

1) Если конусу Y принадлежат векторы $-a_1, -a_2, ..., -a_k$, то он является подпространством размерности, не превышающей k, и, следовательно, замкнутым множеством.

2) Предположим, что хотя бы один из векторов $-a_k \notin Y$. Всякий $y \in Y$ представим в виде $y = \overline{y} + \alpha a_k, \alpha \geq 0$, где $\overline{y} \in \overline{Y}$. Рассмотрим последовательность $\{y_n\}$, сходящуюся к y. В силу предыдущего $y_n = \overline{y_n} + \alpha_n a_k$, $\alpha_n \geq 0$, для всех номеров n. Если последовательность $\{\alpha_n\}$ ограничена, то не умаляя общности, можно считать, что $\lim_{n \to \infty} \alpha_n = \alpha$ и, следовательно,

$y-\alpha a_k=\lim\limits_{n\to\infty}\left(y_n-\alpha_n a_k\right)=\lim\limits_{n\to\infty}\overline{y_n}\equiv\overline{y_n}\in\overline{Y}$ в силу замкнутости $\overline{Y}$. Значит, $y=\overline{y}+\alpha a_k\in Y$. Итак, если последовательность $\{\alpha_n\}$ ограничена, то множество Y замкнуто.

Предположим теперь, что $\alpha_n\underset{n\to\infty}{\to}+\infty$. Тогда $\frac{1}{\alpha_n}y_n=\frac{1}{\alpha_n}\overline{y}_n+a_k$. Так как $\lim\limits_{n\to\infty}\frac{1}{\alpha_n}y_n=0$, то $\overline{y}_n\equiv\lim\limits_{n\to\infty}\frac{1}{\alpha_n}\overline{y}_n=-a_k$. Но $\overline{y}_n\in Y$ в силу замкнутости последнего, и, следовательно, $-a_k\in Y$. Противоречие. Δ

2.3.3 Пусть заданы матрица B размерности $m\times n$ и вектор $v\in E_n$.

Теорема Фаркаша. Неравенство $\langle v,x\rangle\le 0$ выполняется для всех $x\in\{x:Bx\le 0\}$ в том и только том случае, если существует такой вектор $u\ge 0$, что $v=B^T u$.

Доказательство. Достаточность. Пусть выполняются соотношения $u\ge 0$ и $v=B^T u$. Тогда для любого $x\in\{x:Bx\le 0\}$ будет $\langle v,x\rangle=\left\langle B^T u,x\right\rangle=\langle u,Bx\rangle\le 0$.

Необходимость. Пусть для всех $x\in\{x:\ Bx\le 0\}$ справедливо $\langle v,x\rangle\le 0$.

Рассмотрим конус

$$Y=\left\{y:y=B^T u,u\ge 0\right\}.$$

Если $v\in$, то теорема доказана. Предположим, что $v\notin Y$. Множество Y выпукло (2.1.6) и замкнуто (2.3.2), поэтому по теореме 2.2.6 существует вектор $c\ne 0$ такой, что

$$\langle c,y\rangle<\langle c,v\rangle \qquad (2.5)$$

для всех $y\in Y$.

Так как $\lambda y\in Y$ при всех $\lambda\ge 0$, то из (2.5) получаем, что $\lambda\langle c,y\rangle<\langle c,v\rangle$ при всех $\lambda\ge 0$. А значит, $\langle c,y\rangle\le 0$. Но

$$\langle c,y\rangle=\left\langle c,B^T u\right\rangle=\langle u,Bc\rangle\le 0.$$

И так как это имеет место для всех $u\ge 0$, то

$$Bc\le 0. \qquad (2.6)$$

Но $y=0\in Y$, поэтому из (2.5) следует

$\langle c, v\rangle > 0$. (2.7)

Взяв $x = c$, из (2.6) и (2.7) получаем противоречие условиям теоремы. Δ

Полезно будет привести геометрическое истолкование теоремы Фаркаша. Пусть

$$B = \begin{bmatrix} b_1^T \\ b_2^T \\ b_3^T \end{bmatrix}$$

и $K = \{x : Bx \le 0\}$.

Конус K есть совокупность всех векторов x, которые образуют с каждым из векторов b_1, b_2, b_3 неострые углы (на рис. 2.4 конус K заштрихован вертикальными линиями, конус $Y = \{y : y = B^T u, u \ge 0\}$ - горизонтальными линиями).

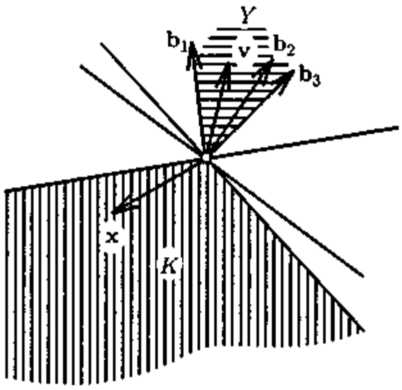

Рис. 4

Теперь очевиден геометрический смысл теоремы: чтобы для любого $x \in K$ угол между v и x был неострым, необходимо и достаточно, чтобы v принадлежал конусу Y.

2.3.4 *Следствие*. Для любой матрицы B и любого вектора v имеет место следующая альтернатива: либо имеет решение система:

$$Bx \ge 0, \qquad \langle v, x\rangle < 0, \tag{2.8}$$

либо имеет решение система:

$v = B^T u \qquad u \ge 0$. (2.9)

Доказательство. Если справедливо (2.9), то из теоремы Фаркаша вытекает, что для любого $x \in \{x \mid Bx \ge 0\}$ будет $\langle v, x\rangle \ge 0$, и, следовательно, система (2.8) неразрешима. Предположим теперь, что система (2.9) неразрешима. Тогда по теореме Фаркаша для всех $x \in \{x : Bx \ge 0\}$ не будет выполняться условие $\langle v, x\rangle \ge 0$, а следовательно, $\langle v, x\rangle < 0$ для некоторого $x \in \{x : Bx \ge 0\}$. Δ

2.3.5 Доказанное следствие допускает простую модификацию, полезную для ряда приложений.

Следствие Для любой матрицы B и любого вектора v имеет место следующая альтернатива: либо имеет решение система:

$$Bx \geq 0, \qquad x \geq 0, \qquad \langle v, x \rangle < 0, \tag{2.10}$$

либо имеет решение система:

$$v \geq B^T u, \qquad u \geq 0. \tag{2.11}$$

Доказательство. Запишем систему (2.10) в виде (2.8):

$$\overline{B}x \geq 0, \qquad \langle v, x \rangle < 0.$$

Здесь $\overline{B} = \begin{bmatrix} B \\ E \end{bmatrix}$, E - единичная матрица. Утверждение следствия с очевидностью вытекает из примечания следствия 2.3.4 к матрице $\overline{B}$. Δ

2.3.6 Теорема о представлении. Рассмотрим конус $S = \{x : Bx = 0, x \geq 0\}$. Вектор x конуса S назовем *ребром*, если в S не существует таких x', x'', $x' \neq x''$, что при некоторых $\lambda_1 > 0$, $\lambda_2 > 0$ будет выполняться равенство $x = \lambda_1 x' + \lambda_2 x''$.

Теперь рассмотрим гиперплоскость

$$L = \left\{ x : \sum_{i=1}^{n} x_i = 1, \;\; x_i \geq 0, \;\; i = \overline{1, n} \right\}.$$

Очевидно, что множество $Q = S \cap L$ выпукло и ограничено.

Если $x_1^*, x_2^*, \ldots, x_k^*$ - угловые точки множества Q, то векторы $\alpha_1 x_1^*, \alpha_2 x_2^*, \ldots, \alpha_k x_k^*$, $\alpha_i \geq 0$ $(i = \overline{1, k})$, являются ребрами конуса S.

Действительно, если $z = \alpha_j x_j^*$ не является ребром, то найдутся такие $x' \in S$, $x'' \in S$, что $z = \lambda_1 x' + \lambda_2 x''$ при некоторых $\lambda_1, \lambda_2 > 0$. Из определения множества Q следует, что найдутся такие $\beta_1, \beta_2 > 0$, при которых $y' = \beta_1 x' \in Q$, $y'' = \beta_2 x'' \in Q$. Но

$$x_j^* = \frac{1}{\alpha_j} z = \gamma_1 y' + \gamma_2 y'', \qquad \gamma_1 = \frac{\lambda_1}{\alpha_j \beta_1}, \qquad \gamma_2 = \frac{\lambda_2}{\alpha_j \beta_2},$$

$$\sum_{i=1}^{k} \frac{1}{\alpha_j} z_i = 1, \qquad \sum_{i=1}^{k} y_i' = 1, \qquad \sum_{i=1}^{k} y_i'' = 1;$$

поэтому $\gamma_1 + \gamma_2 = 1$, что противоречит предположению о том, что точка x_j^* угловая для множества Q.

Доказанное утверждение и теорема 2.2.10 позволяют сформулировать теорему о представлении.

Теорема. Любой вектор x, принадлежащий конусу S, может быть представлении в виде положительных комбинаций его ребер, т.е. найдутся такие числа $\lambda_1 \geq 0, \lambda_2 \geq 0, \ldots, \lambda_k \geq 0$, что $x = \sum_{i=1}^{k} \lambda_i x_i^*$.

2.4 Выпуклые функции

2.4.1 *Определение*. Функцию $\varphi(x)$, определенную на выпуклом множестве X, называют *выпуклой*, если для любых $x, y \in X$ и всех $\alpha \in [0,1]$ выполняется неравенство

$$\varphi(\alpha x + (1-\alpha) y) \leq \alpha \varphi(x) + (1-\alpha) \varphi(y) . \qquad (2.12)$$

На рис. 2.5 изображена выпуклая функция. Очевидно, что каждая точка любой хорды графика функции φ либо лежит над графиком, либо принадлежит ему.

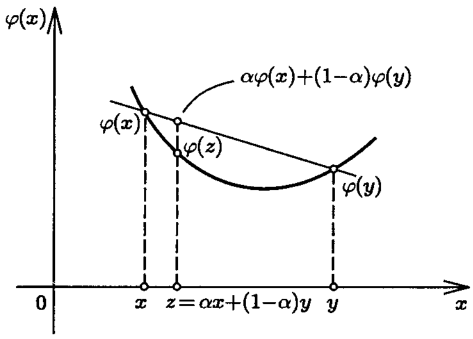

Рис. 2.5

2.4.2 *Вогнутой функцией* называют такую функцию $\varphi(x)$, для которой функция $-\varphi(x)$ выпукла. Таким образом, если

$$\varphi(\alpha x + (1-\alpha) y) \geq \alpha \varphi(x) + (1-\alpha) \varphi(y) \qquad (2.13)$$

для любых $x, y \in X$ и всех $\alpha \in [0,1]$, то $\varphi(x)$ - вогнутая функция.

2.4.3 Если для любого $\alpha \in (0,1)$ неравенство (2.12) строгое, то функцию $\varphi(x)$ называют *строго выпуклой*.

2.4.4 Примером выпуклой функции служит квадратичная функция с положительно определенной матрицей.

Поскольку этим свойством квадратичных функций мы будем пользоваться в дальнейших рассмотрениях, докажем его как самостоятельный результат.

Теорема. Для того чтобы квадратичная функция

$$\varphi(x) = \langle x, Bx\rangle + \langle p, x\rangle$$

Была выпуклой, необходимо и достаточно, чтобы симметрическая матрица B была положительно определенной.

Доказательство.

$$\varphi(\alpha x + (1-\alpha)y) = \alpha^2\langle x, Bx\rangle + 2\alpha(1-\alpha)\langle x, Bx\rangle + (1-\alpha)^2\langle y, By\rangle + \alpha\langle p, x\rangle +$$
$$+ (1-\alpha)\langle x-y, B(x-y)\rangle$$

И при $\alpha \in (0,1)$ будет

$$\alpha(1-\alpha)\langle x-y, B(x-y)\rangle \geq 0$$

в том и только том случае, когда B положительно определена. Отсюда и из определения выпуклости следует утверждение теоремы. Δ

Замечание. Очевидно, что для строгой выпуклости квадратичной функции $\varphi(x)$ необходима и достаточна строгая положительная определенность матрицы B.

2.4.5 *Теорема*. Для любой выпуклости функции $\varphi(x)$, определенной на выпуклом множестве X, и любого числа λ множество $Z = \{x \in X : \varphi(x) \leq \lambda\}$ выпукло.

Доказательство. В соответствии с определением 2.1.6 достаточно показать, что из $x, y \in Z$ следует $z = \alpha x + (1-\alpha)y \in Z$ для всех $\alpha \in [0,1]$. Из выпуклости множества X следует, что $z \in X$. Воспользовавшись неравенством (2.12), получаем:

$$\varphi(z) = \varphi(\alpha x + (1-\alpha) y) \le \alpha \varphi(x) + (1-\alpha)\varphi(y) \le \alpha\lambda + (1-\alpha)\lambda = \lambda. \ \Delta$$

Замечание. Очевидно, что если $\varphi(x)$ - вогнутая функция, то множество $Z = \{x \in X : \varphi(x) \ge \lambda\}$ выпукло.

2.4.6 *Неравенство Иенсена*. Если $\varphi(x)$ выпукла на выпуклом множестве X и

$$z = \sum_{i=1}^{m} \alpha_i x_i, \qquad \sum_{i=1}^{m} \alpha_i = 1,$$

$$\alpha_i \ge 0, \ \ x_i \in X, \ \ i = \overline{1,m},$$

то

$$\varphi\left(\sum_{i=1}^{m} \alpha_i x_i\right) \le \sum_{i=1}^{m} \alpha_i \varphi(x_i). \qquad , \tag{2.14}$$

Доказательство (по индукции). При $m=1$ неравенство (2.14) очевидно. Предположим, что (2.14) справедливо для $m-1, m>1$. Из 2.1.8 следует, что $z \in X$. Если $\alpha_m = 1$, то $\alpha_1 = \alpha_2 = \ldots = \alpha_{m-1} = 0$, и в (2.14) будет равенство. Если $0 \le \alpha_m < 1$, то из выпуклости и индуктивности предположения следует

$$\varphi(z) = \varphi\left(\sum_{i=1}^{m} \alpha_i x_i\right) = \varphi\left((1-\alpha_m)\sum_{i=1}^{m-1} \frac{\alpha_i}{1-\alpha_m} x_i + \alpha_m x_m\right) \le (1-\alpha_m)\varphi\left(\sum_{i=1}^{m-1} \frac{\alpha_i}{1-\alpha_m} x_i\right) + \alpha_m \varphi(x_m) \le$$

$$\le (1-\alpha_m)\sum_{i=1}^{m-1} \frac{\alpha_i}{1-\alpha_m}\varphi(x_i) + \alpha_m \varphi(x_m) = \sum_{i=1}^{m} \alpha_i \varphi(x_i)$$

. Δ

2.4.7 Приведем (без доказательства) следующее важное свойство выпуклых функций.

Выпуклая функция $\varphi(x)$, определенная на выпуклом множестве X, непрерывна в каждой внутренней точке этого множества и имеет в каждой внутренней точке производную по любому направлению s. $(\|s\| = 1)$:

$$\frac{\partial \varphi(x)}{\partial s} = \lim_{\lambda \to +0} \frac{\varphi(x + \lambda s) - \varphi(x)}{\lambda}.$$

2.4.8 В методе штрафных функций находят применение следующие два свойства выпуклых функций.

Если $\chi(x)$ выпукла на выпуклом множестве X, то выпукла на X и функция

$$\varphi(x) = \max\{\chi(x), 0\}.$$

В самом деле, для любых $x, y \in X$

$$\begin{aligned}&\varphi(\alpha x + (1-\alpha)y) = \max\{\chi(\alpha x + (1-\alpha)y), 0\} \le \max\{\alpha\chi(x) + (1-\alpha)\chi(y), 0\} \le \\ &\le \alpha \max\{\chi(x), 0\} + (1-\alpha)\max\{\chi(y), 0\} = \alpha\varphi(x) + (1-\alpha)\varphi(y)\end{aligned}. \quad \Delta$$

Если $\varphi(x)$ выпукла и неотрицательна на выпуклом множестве X, то будет выпукла на X и функция $\varphi^2(x)$.

Действительно, в силу выпуклости $\varphi(x)$ и ее неотрицательности имеем

$$\begin{aligned}&\varphi^2(\alpha(1-\alpha)y) \le \alpha^2\varphi^2(x) + 2\alpha(1-\alpha)\varphi(x)\varphi(y) + (1-\alpha)^2\varphi^2(y) = \\ &= \alpha\varphi^2(x) + (1-\alpha)\varphi^2(y) - \alpha(1-\alpha)(\varphi(x) - \varphi(y))^2 \le \alpha\varphi^2(x) + (1-\alpha)\varphi^2(y)\end{aligned}. \quad \Delta$$

2.4.9 Важное свойство выпуклых дифференцируемых функций, которым мы будем часто пользоваться, устанавливает следующее утверждение.

Функция $\varphi(x)$, дифференцируема на выпуклом множестве X, выпукла в том и только том случае, если для любых $x \in X$ и $y \in X$ будет

$$\langle \varphi'(x), y - x \rangle \le \varphi(y) - \varphi(x). \tag{2.15}$$

Для вогнутой функции

$$\langle \varphi'(x), y - x \rangle \ge \varphi(y) - \varphi(x).$$

Доказательство. Пусть $\varphi(x)$ выпукла. Тогда для любых $x, y \in X, x \ne y$, и всех α таких, что $0 < \alpha \le 1$, справедливо неравенство:

$$\varphi(x + \alpha(y - x)) \le \varphi(x) + \alpha[\varphi(y) - \varphi(x)],$$

откуда

$$\|y - x\| \frac{\varphi(x + \beta s) - \varphi(x)}{\beta} \le \varphi(y) - \varphi(x),$$

где $s = \dfrac{y - x}{\|y - x\|}$, а $\beta = \alpha\|y - x\|$.

Переходя к пределу в последнем неравенстве при $\beta \to 0$, получим

$$\frac{\partial \varphi(x)}{\partial s}\|y-x\| \le \varphi(y)-\varphi(x),$$

т.е.

$$\langle \varphi'(x), s\rangle \|y-x\| = \langle \varphi'(x), y-x\rangle \le \varphi(y)-\varphi(x).$$

Пусть теперь выполняется условие (2.15). Рассмотрим точку $z=\alpha x+(1-\alpha)y$ при $0 \le \alpha \le 1$. Так как $z \in X$, то

$$\langle \varphi'(z), x-z\rangle \le \varphi(x)-\varphi(z),$$
$$\langle \varphi'(z), y-z\rangle \le \varphi(y)-\varphi(z).$$

Умножив первое неравенство на α, второе – на $(1-\alpha)$ и сложив полученные неравенства, имеем

$$0=\langle \varphi'(z), 0\rangle \le \alpha\varphi(x)+(1-\alpha)\varphi(y)-\varphi(z). \ \Delta$$

2.4.10 Экстремальные свойства. Рассмотрим задачу отыскания точки x^* выпуклого множества X, в которой выпуклая функция $\varphi(x)$, определенная на X, достигает минимального значения, - задачу отыскания оптимальной точки $x^*=\arg\min\{\varphi(x): x \in X\}$.

2.4.11 *Теорема*. Если выпуклы функция $\varphi(x)$ и множество X, то любая точка $x^* \in X$, являющаяся точкой локального минимума, будет оптимальной для задачи минимизации функции $\varphi(x)$ на множестве X.

Доказательство. Предположим, что x^* не является оптимальной точкой, т.е. найдется точка $x' \in X$ такая, что

$$\varphi(x') < \varphi(x^*).$$

Рассмотрим точки вида

$$x=\alpha x'+(1-\alpha)x^*, \qquad \alpha \in [0,1].$$

Так как множество X выпукло, то $x \in X$. Далее, из выпуклости $\varphi(x)$ и из предположения об x' следует

$$\varphi(x)=\varphi(\alpha x'+(1-\alpha)x^*) \le \alpha\varphi(x')+(1-\alpha)\varphi(x^*) < \alpha\varphi(x^*)+(1-\alpha)\varphi(x^*)=\varphi(x^*),$$

т.е.

$\varphi(x) < \varphi(x^*)$.

Но это противоречит условию, что x^* - точка локального минимума, поскольку при малых α точка x находится в достаточно малой окрестности точки x^*. Δ

2.4.12 *Теорема*. Если выпуклы функция $\varphi(x)$ и множество X, то множество оптимальных точек

$$X^* = \left\{x^* \in X : \varphi(x^*) = \mu = \min_{x \in X} \varphi(x)\right\} = Arg\min\{\varphi(x) : x \in X\}$$

выпукло.

Доказательство. Пусть $x', x'' \in X^*$. Поскольку $X^* \subseteq X$ и X - выпуклое множество, то для любого $\lambda \in [0,1]$ будет

$$z = \lambda x' + (1-\lambda) x'' \in X,$$

а в силу выпуклости $\varphi(x)$ имеем

$$\varphi(z) = \varphi(\lambda x' + (1-\lambda) x'') \le \lambda \varphi(x') + (1-\lambda)\varphi(x'') = \lambda\mu + (1-\lambda)\mu = \mu .$$

Кроме того, $\varphi(z) \ge \mu$, поэтому $\varphi(z) = \mu$, т.е. $z \in X^*$. Δ

2.4.13 Если $\varphi(x)$ строго выпукла на выпуклом множестве X и точка $x^* \in X$ оптимальна, т.е.

$$\mu = \varphi(x^*) = \min_{x \in X} \varphi(x),$$

то для всех $x \in X$ и $x \ne x^*$ будет

$$\varphi(x) > \varphi(x^*),$$

и, значит, точка x^* единственна.

Доказательство. Предположим, что найдется точка $x' \in X$, $x' \ne x^*$, такая, что

$$\varphi(x') = \varphi(x^*) = \mu .$$

Тогда для любого $\alpha \in (0,1]$ точка $x = \alpha x' + (1-\alpha) x^*$ принадлежит множеству X, и в силу строгой выпуклости функции $\varphi(x)$ будет

$$\varphi(x) = \varphi(\alpha x' + (1-\alpha) x^*) < \alpha \varphi(x') + (1-\alpha)\varphi(x^*) = \alpha\mu + (1-\alpha)\mu = \mu ,$$

что противоречит оптимальности точки x^*. Δ

2.4.14 Одно свойство вогнутых функций. Если $\varphi(x)$ - вогнутая функция и $\varphi(x) > \varphi(y)$, то для любой точки $z \in [x, y]$ такой что $\varphi(x) > \varphi(z) > \varphi(y)$, справедливо неравенство

$$\frac{\varphi(z) - \varphi(y)}{\|z - y\|} \ge \frac{\varphi(x) - \varphi(z)}{\|x - z\|}.$$

Доказательство. Предположим, что найдется $\alpha \in (0,1)$ такое что для $z = \alpha x + (1-\alpha)y$ будет

$$\frac{\varphi(z) - \varphi(y)}{\|z - y\|} < \frac{\varphi(x) - \varphi(z)}{\|x - z\|}.$$

Так как $\|z - y\| = \dfrac{\alpha}{1-\alpha}\|x - z\|$, то

$$\frac{\varphi(z) - \varphi(y)}{\|z - y\|} = \frac{\varphi(z) - \varphi(y)}{\|x - z\|}\left(\frac{1-\alpha}{\alpha}\right) < \frac{\varphi(x) - \varphi(z)}{\|x - z\|},$$

откуда $(1-\alpha)[\varphi(z) - \varphi(y)] < \alpha[\varphi(x) - \varphi(z)]$. Таким образом, приходим к неравенству

$$\varphi(z) < \alpha\varphi(x) + (1-\alpha)\varphi(y),$$

противоречащему вогнутости функции $\varphi(x)$. Δ

2.4.15 *Теорема*. Если множество X выпукло, замкнуто и неограниченно, то существует такое направление s, $\|s\| = 1$, что $x + \lambda s \in X$ для всех $\lambda \ge 0$ и всех $x \in X$.

Доказательство. Так как множество X неограниченно, то найдется последовательность $\{x_k\} \subset X, \|x\| \to \infty, k \to \infty$. Рассмотрим фиксированную точку $x \in X \setminus \{x_k\}$ и направление $s_k = \dfrac{1}{\|x_k - x\|}(x_k - x$. Очевидно существование такой подпоследовательности $\{s_{k_i}\}$, что $\lim_{i\to\infty} s_{k_i} = s$. Для любого фиксированного $\lambda > 0$ найдется такой номер k_0, что для всех $k_i \ge k_0$ будет $0 < \dfrac{\lambda}{\|x_{k_i} - x\|} < 1$.

Так как $x + \lambda s_{k_i} = \dfrac{\lambda}{\|x_{k_i} - x\|} x_k + \left(1 - \dfrac{\lambda}{\|x_{k_i} - x\|}\right) x \in X$

для всех $k_i \geq k_0$, то $\lim\limits_{i\to\infty}(x+\lambda s_{k_i}) = x+\lambda s$ принадлежит X в силу замкнутости множества X.

Далее, пусть y - произвольный фиксированный элемент из X. Так как $v_k = x+\lambda_k\lambda_s \in X$ для всех $\lambda \geq 1, \lambda_k > 1$, то

$$\frac{1}{\lambda_k}v_k + (1-\frac{1}{\lambda_k})y = \frac{1}{\lambda_k}(x-y)+y+\lambda s \in X,$$

и при $\lambda_k \to \infty$, в силу замкнутости множества X получаем $y+\lambda s \in X$. Δ

2.4.16 *Теорема*. Если функция $f(x)$ выпукла и непрерывна на выпуклом замкнутом множестве X, то для ограниченности множества

$$X(\beta) = \{x \in X : f(x) \leq \beta\}$$

при любом β необходимо и достаточно существования числа α, при котором множество $X(\alpha) = \{x \in X : f(x) \leq \alpha\}$ непусто и ограничено.

Доказательство. Необходимость очевидна.

Достаточность. Пусть для некоторого α множество $X(\alpha)$ ограничено. Если $\beta \leq \alpha$, то $X(\beta) \subset X(\alpha)$; следовательно, $X(\beta)$ - ограниченное множество.

Предположим, что найдется такое $\beta > \alpha$, что множество $X(\beta)$ неограниченно. Тогда в фиксированной точке $x \in X(\alpha) \subset X(\beta)$ существует такое направление s, $\|s\| = 1$, что $x+\lambda s \in X(\beta)$ при всех $\lambda \geq 0$. Поскольку множество $X(\alpha)$ ограничено, то найдется такое, что $x+\lambda s \notin X(\alpha)$ при всех $\lambda \geq \lambda_0$. Итак, для $z = x+\lambda_0 s$ будет $f(x) \leq \alpha < f(z) \leq \beta$, а для $v = v(\lambda) = x+\lambda\lambda_0 s$ при всех $\lambda > 0$ будет $f(v) \leq \beta$. Так как $z = x+\lambda_0 s = \frac{1}{\lambda}v + (1-\frac{1}{\lambda})x$, то при $\lambda \geq 1$ из выпуклости функции $f(x)$ на выпуклом замкнутом множестве $X(\beta)$ следует неравенство $f(z) \leq \frac{1}{\lambda}f(v) + (1-\frac{1}{\lambda})f(x)$, откуда $f(v) \geq \lambda(f(z)-f(x)) + f(x)$. Но $f(z)-f(x)$ - фиксированное положительное число, вследствие чего $f(v) \underset{\lambda\to+\infty}{\to} +\infty$, что противоречит условию $f(v) \leq \beta$ при всех $\lambda > 0$. Δ

2.5 Сильная выпуклость функций

2.5.1 Как мы могли видеть ранее, если для выпуклой функции $\varphi(x)$ существует точка локального минимума на выпуклом и замкнутом множестве X, то она является оптимальной, а для строго выпуклой функции эта точка вдобавок и единственна. Подчеркнем, что эти утверждения справедливы лишь в предложении существования точки локального минимума.

Рассмотрим теперь класс функций, для которых на любом непустом замкнутом множестве всегда существует точка минимума, и если вдобавок это множество выпукло, то эта точка единственна.

2.5.2 *Определение*. Функцию $\varphi(x)$, определенную на некотором множестве X, будем называть *сильно выпуклой*, если существует константа $\rho > 0$ такая, что для любых $x, y \in X$ таких, что $[x, y] \subset X$ и для любого $\alpha \in [0,1]$ будет выполняться неравенство

$$\varphi(\alpha x + (1-\alpha)y) \le \alpha\varphi(x) + (1-\alpha)\varphi(y) - \alpha(1-\alpha)\rho\|x-y\|^2.$$

Величину ρ в дальнейшем будем называть параметром сильной выпуклости.

2.5.3 Пример сильно выпуклой функции. Рассмотрим квадратичную функцию

$$\varphi(x) = \langle x, Bx\rangle + \langle p, x\rangle,$$

где B - строго положительно определенная матрица. Сильная выпуклость следует из соотношения

$$\varphi(\alpha x + (1-\alpha)y) = \alpha\varphi(x) + (1-\alpha)\varphi(y) - \alpha(1-\alpha)\langle x-y, B(x-y)\rangle,$$

поскольку

$$\langle x-y, B(x-y)\rangle \ge \lambda\|x-y\|^2,$$

где λ - наименьшее собственное число матрицы B.

Укажем на некоторые свойства сильно выпуклых функций.

2.5.4 *Теорема* Если функция $\varphi(x)$ сильно выпукла и непрерывна на выпуклом и замкнутом множестве X, то для любой точки $y \in X$ множество

$$X_0 = \{x \in X : \varphi(x) \le \varphi(y)\}$$

ограничено и существует единственная точка

$$x^* = \arg\min\{\varphi(x) : x \in X\}.$$

Доказательство. Зафиксируем произвольную точку $y \in X$. В силу непрерывности функции $\varphi(x)$ для $\rho > 0$ найдется такое $\varepsilon > 0$, что $|\varphi(x) - \varphi(y)| \le \rho$ при всех $x \in U = \{x \in X : \|x - y\| \le \varepsilon\}$. Таким образом, $\varphi(x) \ge \varphi(y) - \rho$ при всех $x \in U$.

Пусть $x \in X \setminus U$. Тогда $\alpha = \varepsilon \|x - y\|^{-1} < 1$. Из 2.5.2 получаем $\alpha\varphi(x) \ge \varphi(y + \alpha(x - y)) - (1 - \alpha)\varphi(y) + \alpha(1 - \alpha)\rho\|x - y\|^2$. Но $z = y + \alpha(x - y) \in U$, поскольку $\|z - y\| \le \varepsilon$, и, следовательно, $\varphi(y + \alpha(x - y)) - \varphi(y) \ge -\rho$. Поэтому

$$\alpha\varphi(x) \ge -\rho + \alpha\varphi(y) + \alpha(1 - \alpha)\rho\|x - y\|^2,$$

или

$$\varphi(x) \ge \varphi(y) + (1 - \alpha)\rho\|x - y\|^2 - \tfrac{\rho}{\alpha} = \varphi(y) - \rho(\varepsilon + \tfrac{1}{\varepsilon})\|x - y\| + \rho\|x - y\|^2.$$

Таким образом, если множество X неограниченно, то при $\|x\| \to \infty$ будет $\varphi(x) \to \infty$. Теперь нетрудно убедиться, что множество X_{01} ограничено. В самом деле, если это не так, то найдется последовательность $\{x_k\} \subset X_0$ такая, что $\|x_k\| \underset{k \to \infty}{\to} \infty$. Но тогда найдется номер $k_0 = k_0(y)$, начиная с которого будет $\varphi(x_k) > \varphi(y), k \ge k_0$, и, следовательно, $x_k \notin X_0$ при $k \ge k_0$.

Существование x^* очевидно, так как из непрерывности $\varphi(x)$ на ограниченном и замкнутом множестве X_0 и из определения X_0 следует, что

$$x^* = \arg\min\{\varphi(x) : x \in X_0\} = \arg\min\{\varphi(x) : x \in X\}.$$

Единственность же точки x^* следует из того, что сильно выпуклая функция является в то же время строго выпуклой. Δ

2.5.5 *Теорема*. Если $\varphi(x)$ сильно выпукла на выпуклом и замкнутом множестве X, то:

а) для всех $x \in X$ справедливо неравенство

$$\|x - x^*\|^2 \le \tfrac{2}{\rho}(\varphi(x) - \varphi(x^*)).$$

Если при этом $\varphi(x) \in C^1(X)$, то:

б) для всех $x, y \in X$ будет $\langle \varphi'(x) - \varphi'(y), x - y \rangle \geq \rho \|x - y\|^2$;

в) $\|x - x^*\| \leq \frac{1}{\rho} \|\varphi'(x)\|$;

г) $0 \leq \varphi(x) - \varphi(x^*) \leq \frac{1}{\rho} \|\varphi'(x)\|^2$.

Доказательство. а) Из определения сильной выпуклости следует

$$\varphi\left(\frac{1}{2}x + \frac{1}{2}x^*\right) \leq \frac{1}{2}\varphi(x) + \frac{1}{2}\varphi(x^*) - \frac{1}{4}\rho \|x - x^*\|^2,$$

и в силу неравенства

$$\varphi(x^*) \leq \varphi\left(\tfrac{1}{2}x + \tfrac{1}{2}x^*\right)$$

соотношение а) становится справедливым.

б) Так как функция $\varphi(x)$ выпукла, то

$$\varphi(x) - \varphi(y) \leq \langle \varphi'(x), x - y \rangle.$$

Отсюда и из определения 2.5.2 получаем

$$\begin{aligned} &\tfrac{1}{4}\rho \|x - y\|^2 \leq \tfrac{1}{2}\left[\varphi(x) - \varphi\left(\tfrac{1}{2}x + \tfrac{1}{2}y\right)\right] + \tfrac{1}{2}\left[\varphi(y) - \varphi\left(\tfrac{1}{2}x + \tfrac{1}{2}y\right)\right] \leq \\ &\leq \tfrac{1}{4}\langle \varphi'(x), x - y \rangle + \tfrac{1}{4}\langle \varphi'(y), y - x \rangle = \tfrac{1}{4}\langle \varphi'(x) - \varphi'(y), x - y \rangle \end{aligned}.$$

в) В точке x^* минимума $\varphi(x)$ на X выполняется для всех $x \in X$ неравенство

$$\langle \varphi'(x^*), x - x^* \rangle \geq 0,$$

Поэтому из б) получаем

$$\rho \|x - x^*\|^2 \leq \langle \varphi'(x) - \varphi'(x^*), x - x^* \rangle \leq \langle \varphi'(x), x - x^* \rangle \leq \|\varphi'(x)\| \|x - x^*\|,$$

т.е. неравенство в).

г) Из выпуклости $\varphi(x)$ и из в) имеем

$$0 \leq \varphi(x) - \varphi(x^*) \leq \langle \varphi'(x), x - x^* \rangle \leq \|\varphi'(x)\| \|x - x^*\| \leq \frac{1}{\rho} \|\varphi'(x)\|^2. \ \Delta$$

6 семестр

ОСНОВЫ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

В этой главе будут рассмотрены условия существования локальных экстремумов дифференцируемых функций на допустимых множествах весьма общего вида, а также условия существования глобальных экстремумов (минимумов) в задачах выпуклого программирования.

3.1. Задачи математического программирования

3.1.1. Основная задача математического программирования.

Рассмотрим множество

$$X = \left\{ x : f_i(x) \leq 0, i = \overline{1,m} \right\}, \tag{3.1}$$

где $f_i(x)(i = \overline{1,m})$ - заданные скалярные функции. Пусть скалярная функция $\varphi(x)$ определена на множестве X.

Задачу минимизации функции $\varphi(x)$ на множестве X будем называть *основной задачей математического программирования*.

3.1.2 Форма записи. Условимся, что запись

$$\varphi(x) \to \min, \qquad x \in X, \tag{3.2}$$

или, ей эквивалентные,

$$\min\left\{ \varphi(x) : x \in X \right\}, \qquad \min_{x \in X} \varphi(x),$$

будут означать, что ставится задача:

1) либо найти оптимальную точку $x^* \in X$:

$$\varphi(x^*) = \min_{x \in X} \varphi(x);$$

2) либо, если не существует такой точки x^*, найти

$$\varphi^* = \inf_{x \in X} \varphi(x);$$

3) либо убедиться, что $\varphi(x)$ - неограниченная снизу на множестве X функция;

4) либо убедиться в том, что $X = \varnothing$.

3.1.3.Если множество X выпукло и выпукла функция $\varphi(x)$, то задачу (3.2) называют *задачей выпуклого программирования.*

3.1.4.**Основная задача выпуклого программирования.** Из 2.4.5 и 2.1.6 вытекает, что для выпуклости множества X (см. (3.1)) достаточно, чтобы функции $f_i(x)(i = \overline{1,m})$ были вогнутыми. Если в задаче выпуклого программирования (3.2) все функции $f_i(x)$ вогнуты, а $\varphi(x)$ выпукло, то будем называть ее *основной задачей выпуклого программирования.*

3.1.5.**Терминология.** Множество X в задаче (3.2) называют *допустимым множеством,* точки этого множества — *допустимыми точками,* а неравенства $f_i(x) \geq 0$, определяющие допустимое множество, *ограничениями.* Точку $x^* = \arg\min\{\varphi(x) : x \in X\}$ называют *решением* или *оптимальной точкой,* а иногда *точкой глобального минимума.* Наконец, точку x, в которой выполняются необходимые условия локального минимума функции $\varphi(x)$ на множестве X, будем называть *стационарной.*

3.2. Возможные направления

3.2.1. Возможные направления. Понятие возможного направления занимает важное место в математическом программировании. В этой главе возможные направления играют вспомогательную роль, однако в дальнейшем они приобретут самостоятельное значение.

Определение. Направление $-s \neq 0$ в точке $x \in X$ называется *возможным,* если существует такое число $\bar{\beta} > 0$, что $x - \beta s \in X$ для всех $\beta \in [0, \bar{\beta}]$.

Например, если $X = \{ x : x \geq 0\}$, то в точке $x = 0$ любой вектор $\vec{s} \geq 0$, $s \neq 0$, задает возможное направление, а в точке

$$x = \begin{bmatrix} x_1 = 0 \\ x_2 > 0 \\ \dots \\ x_n > 0 \end{bmatrix}$$

возможным является любое направление $\vec{s}$:

$$s = \begin{bmatrix} s_1 \geq 0 \\ s_2 \\ \dots \\ s_n \end{bmatrix}$$

где $s_s, s_2, \dots, s_n$ — произвольные числа, $s \neq 0$.

Очевидно, если x — внутренняя точка множества X, то любое направление $\vec{s}$ в этой точке является возможным.

В следующих разделах будут выяснены условия существования возможных направлений.

3.2.2. Активные ограничения. Очевидно, что на экстремальные свойства целевой функции $\varphi(x)$ в некоторой достаточно малой окрестности фиксированной точки $x \in X$ влияют лишь те ограничения $f_i(x) \geq 0$, для которых в этой точке выполняются равенства $f_i(x) = 0$. В связи с этим вводится следующее определение.

Ограничение $f_i(x) \geq 0$ называется *активным* в фиксированной точке $x \in X$, если $f_i(x) = 0$.

3.2.3. В дальнейшем нам понадобится совокупность индексов активных ограничений в точке $x \in X$

$I(x) = \{ i : f_i(x) = 0 \}$.

3.2.4.**Предположения.** Всюду в этой главе будем предполагать функции $\varphi(x)$ и непрерывно $f_i(x) (i = \overline{1,m})$ дифференцируемыми.

3.2.5.Пусть s — некоторый n-мерный вектор, а σ — некоторая скалярная величина.

Для фиксированного $x \in X$ рассмотрим следующую систему неравенств относительно переменных s и σ :

$$\langle f_i'(x), s \rangle + \sigma \le 0, \qquad i \in I(x). \quad (3.3)$$

При определенных условиях решения этой системы определяют возможные направления в точке x.

3.2.6. *Если* s $(\|s\| \ne 0)$ *удовлетворяет системе* (3.3) *при некотором* $\sigma > 0$*, то направление* $-s$ *является возможным в точке* $x \in X$.

Доказательство. Естественно предполагать, что $I(x) \ne \varnothing$, поскольку в противном случае точка x является внутренней точкой множества X, и поэтому любое направление $-s$ является возможным. Если $i \notin I(x)$, то, $f_i(x) > 0$ и малое перемещение из точки x по любому направлению, в частности по направлению $-s$, не нарушит этого неравенства. Пусть $i \in I(x)$. Предположим, что $f_i(x - \beta s) < 0$ для любого сколь угодно малого $\beta > 0$, т. е. направление $-s$ не является возможным. Так как $f_i(x) = 0$, то для любого $\beta > 0$ будет

$$\frac{1}{\beta}[f_i(x) - f_i(x - \beta s)] > 0,$$

а значит (см. (2.4.7)),

$$\lim_{\beta \to 0} \frac{1}{\beta}[f_i(x) - f_i(x - \beta s)] = \langle f_i'(x), s \rangle \ge 0,$$

что противоречит условию (3.3) при $\sigma > 0$. Δ

3.2.7. *Если направление* $-s$ *в точке* $s \in X$ *является возможным, то существует такое* $\sigma \le 0$*, что пара* s*,* σ *удовлетворяет системе* (3.3).

Доказательство. Предположим, что найдется хотя бы один такой номер $i \in I(x)$, для которого $\langle f_i'(x), s\rangle > 0$. Так как $-f_i(x-\beta s) = f_i(x) - f_i(x-\beta s) = \beta \langle f_i'(x), s\rangle + o(\beta)$, то $f_i(x-\beta s) < 0$ при достаточно малых значениях $\beta > 0$, т. е. $x - \beta s \notin X$, что противоречит предположению о том, что направление $-s$ возможное. Δ

3.2.8. *Если множество* X *задается системой линейных неравенств:*

$$X = \left\{ x : f_i(x) = \langle a_i, x\rangle - b_i \le 0, \qquad i = \overline{1,m} \right\},$$

то условия

$$\langle a_i, s\rangle \ge 0, \qquad i \in I(x),$$

являются необходимыми и достаточными для того, чтобы направление $-s$ *было возможным в точке* $x \in X$.

Доказательство. Нужно выяснить условия, при которых точка $y = x - \beta s$ хотя бы для достаточно малых $\beta > 0$ будет принадлежать множеству X:

$$\langle a_i, y\rangle - b_i = \langle a_i, x\rangle - b_i - \beta \langle a_i, s\rangle.$$

Если $i \notin I(x)$, то $\langle a_i, x\rangle - b_i > 0$, и при достаточно малых β будет $\langle a_i, y\rangle - b_i \le 0$.

Если $i \in I(x)$, то $\langle a_i, y\rangle - b_i = -\beta \langle a_i, s\rangle$, и для выполнения неравенства $\langle a_i, y\rangle - b_i \le 0$ при $\beta > 0$ необходимо и достаточно, чтобы было $\langle a_i, s\rangle \ge 0$. Δ

3.3. Экстремальные свойства

В последующих параграфах этой главы рассматриваются основы теории математического программирования: доказываются теоремы существования локальных экстремумов и теоремы существования решений задач математического программирования.

В этом параграфе рассматриваются условия существования стационарных точек основной задачи математического программирования, т. е. (в соответствии с принятой терминологией) необходимые условия существования локальных минимумов.

Еще раз подчеркнем, что остаются в силе предположения 3.2.4 о непрерывной дифференцируемости функций $\varphi(x)$ и $f_i(x)(i=\overline{1,m})$.

3.3.1. Следующая система линейных неравенств относительно переменных s и σ играет важную роль в дальнейших рассмотрениях:

$$\langle f_i'(x), s\rangle + \sigma \le 0, \qquad i \in I(x), \tag{3.4}$$

$$-\langle \varphi'(x), s\rangle + \sigma \le 0. \tag{3.5}$$

3.3.2. Теорема. *Для того чтобы точка* $x \in X$ *являлась точкой локального минимума функции* $\varphi(x)$ *на множестве* X, *необходимо, чтобы в каждой паре* s, σ *удовлетворяющей системе* (3.4), (3.5), *было*

$$\sigma \le 0. \tag{3.6}$$

Доказательство. Пусть x — точка локального минимума. Предположим, что найдется такая пара s, σ удовлетворяющая системе (3.4), (3.5), что $\sigma > 0$. Согласно 3.2.6 направление является $-s$ возможным в точке x. В силу непрерывности $\varphi(x)$ и предположения, что $\langle \varphi'(x), s\rangle \ge \sigma > 0$, для достаточно малого $\beta > 0$ будет $\langle \varphi'(x-\beta s), s\rangle > 0$ и $x-\beta s \in X$. Но по теореме о среднем

$$\varphi(x) - \varphi(x-\beta s) = \beta\langle \varphi'(x-\theta\beta s), s\rangle > 0 \qquad 0 \le \theta \le 1.$$

Таким образом, в локальной окрестности точки локального минимума x нашлась точка $y = x - \beta s$ такая, что $\varphi(y) < \varphi(x)$. Противоречие. Δ

3.3.3. Следуя принятой терминологии (см. 3.1.5), теорему 3.3.2 можно переформулировать следующим образом: если в каждой паре s, σ, удовлетворяющей системе (3.4), (3.5), выполняется условие (3.6), то x является стационарной точкой основной задачи математического программирования.

3.3.4. Следствие. *Если точка* $x \in X$ *локального минимума функции* $\varphi(x)$ *на множестве* X *такова, что* $I(x) = \varnothing$*, то*

$$\varphi'(x) = 0.$$

Доказательство. Предположим, что $\varphi'(x) \neq 0$. Заметим, что x — внутренняя точка множества X. Так как $I(x) = \varnothing$, то в системе (3.4), (3.5) остается лишь одно неравенство (3.5). Пара $s = \varphi'(x)$, $\sigma = \langle s, s \rangle > 0$ удовлетворяет условию (3.5), но при этом нарушается условие (3.6) теоремы 3.3.2. Противоречие. Δ

3.3.5. *Теорема. Если в точке* $x \in X$ *локального минимума функции* $\varphi(x)$ *на множестве* X *система векторов* $f_i'(x), i \in I(x)$*, линейно независима, то найдутся такие числа* $u_i \geq 0, i \in I(x)$*, что*

$$\varphi'(x) = \sum_{i \in I(x)} u_i f_i'(x). \tag{3.7}$$

Доказательство. Согласно теореме 3.3.2 выполняются условия (3.4), (3.5), (3.6). Условие (3.6) формально запишем так:

$$\langle 0, s \rangle + 1 \cdot \sigma \leq 0, \tag{3.8}$$

и применим к системе (3.4), (3.5), (3.8) теорему Фаркаша: найдутся такие $v_i \geq 0, i \in I(x)$, и $v_0 \geq 0$, что

$$0 = \sum_{i \in I(x)} v_i f_i'(x) - v_0 \varphi'(x), \tag{3.9}$$

$$1 = \sum_{i \in I(x)} v_i + v_0. \tag{3.10}$$

Поскольку предположение $v_0 = 0$ противоречит линейной независимости векторов $f_i'(x), i \in I(x)$ (в силу (3.9) и (3.10)), то $v_0 > 0$, и, следовательно, при $u_i = \frac{1}{v_0} v_i, i \in I(x)$, из условия (3.9) получаем (3.7). Δ

3.3.6. *Замечание*. Иногда теорему 3.3.5 формулируют следующим образом: если найдутся такие числа $u_i \ge 0 (i = \overline{1,m})$, что

$$\varphi'(x) = \sum_{i=1}^{m} u_i f_i'(x), \qquad \sum_{i=1}^{m} u_i f_i(x) = 0,$$

то x является стационарной точкой основной задачи математического программирования.

Справедливость этих условий очевидна: достаточно положить $u_i = 0$ при $i \notin I(x)$.

3.3.7. Геометрическая интерпретация. Условиям (3.7) можно дать следующее геометрическое истолкование. Заметив, что векторы $-f_i'(x), i \in I(x)$, являются внешними нормалями в точке x к граничным поверхностям $f_i(x) = 0$, и переписав условие (3.7) в виде

$$-\varphi'(x) = \sum_{i \in I(x)} u_i \left[-f_i'(x) \right], \qquad u_i \ge 0, \qquad i \in I(x),$$

можем сказать: условие (3.7) означает, что антиградиент можно представить в виде линейной комбинации с неотрицательными коэффициентами внешних нормалей к ограничениям в точке x. Другими словами, антиградиент принадлежит конусу, натянутому на внешние нормали к ограничениям в точке x.

3.4. Экстремальные свойства на выпуклых множествах

3.4.1. Условие регулярности. В случае выпуклости множества $X = \left\{ x : f_i(x) \le 0, i = \overline{1,m} \right\}$ условия линейной независимости векторов $f_i'(x)$, соответствующих активным ограничениям, в предыдущей теореме можно заменить более просто проверяемым, а именно так называемым *условием регулярности.* Существуют различные условия регулярности ограничений; здесь будут рассмотрены следующие условия.

3.4.2. **Условие регулярности**. Если для каждого $1 \le i \le m$ существует такая точка $x_i \in X$, что

$$f_i(x_i) > 0, \tag{3.11}$$

то говорят, что множество X удовлетворяет *условию регулярности.*

3.4.3. Условие регулярности Слейтера. Существует такая точка $x \in X$, что для всех $i = \overline{1,m}$ будет

$$f_i(x_i) > 0. \tag{3.12}$$

Легко доказывается эквивалентность условий (3.11) и (3.12). Очевидно, что из (3.12) следует (3.11).

Пусть теперь выполняется (3.11). Выберем

$$x = \sum_{i=1}^{m} \alpha_i x_i, \qquad \sum_{i=1}^{m} \alpha_i = 1, \qquad \alpha_i \ge 0, \qquad i = \overline{1,m}.$$

Тогда (3.12) вытекает из неравенства Иенсена 2.4.6 для вогнутых функций $f_i(x)$.

3.4.4.Теорема. *Если функции $f_i(x)$ вогнуты, множество $X = \{x : f_i(x) \ge 0,\ i = \overline{1,m}\}$ регулярно по Слейтеру, а точка $x \in X$ является точкой локального минимума функции $\varphi(x)$ на множестве X, то найдутся такие числа $u_i \ge 0, i \in I(x)$, что*

$$\varphi'(x) = \sum_{i \in I(x)} u_i f_i'(x).$$

Доказательство. Повторяя доказательство теоремы 3.3.5, приходим к соотношениям

$$v_i \geq 0,\ \ i \in I(x),\ \ v_0 \geq 0,$$
$$0 = \sum_{i \in I(x)} v_i f_i'(x) - v_0 \varphi'(x), \tag{3.13}$$

$$1 = \sum_{i \in I(x)} v_i + v_0 \tag{3.14}$$

Для завершения доказательства осталось убедиться в том, что $v_0 > 0$. Предположим, что $v_0 = 0$. Из (3.14) в этом случае следует, что хотя бы одно $v_l > 0, l \in I(x)$. Из регулярности множества X вытекает существование такой точки $z \in X$, что $f_i(z) > 0 (i = \overline{1,m})$. Тогда направление $-s = z - x$ будет возможным. Так как $f_l(x)$ — вогнутая функция, то из 2.4.9 получаем

$$-\langle f_l'(x), s\rangle \geq f_l(z) - f_l(x) > 0.$$

Умножим равенство (3.13) скалярно на $-s$:

$$0 = \sum_{i \in I(x)} v_i \langle f_i'(x), s\rangle. \tag{3.15}$$

Поскольку $-s$ — возможное направление в точке x, то из теоремы 3.2.7 следует, что $\langle f_i'(x), s\rangle \leq 0, i \in I(x)$. Итак, все слагаемые в правой части равенства (3.15) одного знака, а одно из них, а именно l*-е,* заведомо отлично от нуля: $-v_l \langle f_l'(x), s\rangle > 0$, что противоречит равенству нулю всей суммы. Δ

Замечание. В случае выпуклости функций $f_i(x) (i = \overline{1,m})$ требование регулярности становится излишним.

Действительно, *если функции* $f_i(x)$ *выпуклы, то равенство*

$$\varphi'(x) = \sum_{i \in I(x)} u_i f_i'(x), \qquad u_i \geq 0,$$

является необходимым условием того, что x *является точкой локального минимума функции* $\varphi(x)$ *на множестве* X.

Справедливость этого утверждения будет следовать из теоремы Фаркаша, если мы покажем, что для любого $u \in E_n$ такого, что

$$\left\langle f_i'(x), u \right\rangle \geq 0, \qquad i \in I(x),$$

имеет место неравенство

$$\left\langle \varphi'(x), u \right\rangle \geq 0$$

Поскольку длина вектора u не влияет на выполнение указанного неравенства, то достаточно установить справедливость последнего неравенства для векторов достаточно мглой длины. Пусть $u = z - x$ и z лежит в такой достаточно малой окрестности точки x, что $f_j(z) > 0$, $j \in \{1,...,n\} \setminus I(x)$. Точка z принадлежит X, поскольку

$$f_i(z) = f_i(z) - f_i(x) \geq \left\langle f_i'(x), z - x \right\rangle \geq 0, \ i \in I(x).$$

Если $\left\langle \varphi'(x), z - x \right\rangle < 0$, то по формуле Тейлора получим

$$\varphi(z) - \varphi(x) = \left\langle \varphi'(x), z - x \right\rangle + o\left(\|z - x\|\right) < 0,$$

так как по предположению величина $\|z - x\|$ достаточно мала.

Последнее противоречит условию, что x является точкой локального минимума.

3.4.5. *Теорема. Если функции* $f_i(x)$ *вогнуты, замкнутое множество регулярно по Слейтеру, а точка* $x \in X$ *является точкой локального минимума функциина* $\varphi(x)$ *множестве* X*, то*

$$x = p\left(x - \varphi'(x)\right).$$

Здесь $p(v)$ означает проекцию точки на v множество X.

Доказательство. Пусть y — произвольная точка множества X. Направление $-s = y - x$ является возможным в точке x. Из теоремы 3.4.4 получаем, что

$$\left\langle \left(x - \varphi'(x)\right) - x, y - x \right\rangle = \left\langle \varphi'(x), s \right\rangle = \sum_{i \in I(x)} u_i \left\langle f_i'(x), s \right\rangle \leq 0.$$

По теореме 2.2.3 отсюда следует, что x является проекцией точки $x-\varphi'(x)$ на множество X. Δ

3.4.6. Случай линейных ограничений. Если ограничения, задающие допустимое множество, линейны, то предыдущая теорема справедлива без предположения регулярности множества X.

Теорема. Если функции $f_i(x)(i=\overline{1,m})$ *линейны, а точка*

$$x\in X=\left\{x: f_i(x)=\left\langle a_i,x\right\rangle-b_i\le 0, i=\overline{1,n}\right\}$$

является точкой локального минимума функции $\varphi(x)$ *на множестве* X*, то найдутся такие числа* $u_i\ge 0, i\in I(x)$*, что*

$$\varphi'(x)=\sum_{i\in I(x)}u_i a_i.$$

Доказательство. Пусть $\varepsilon>0$ — столь малое положительное число, что для всех точек, принадлежащих окрестности

$$U_\varepsilon(x)=\left\{y\in X:\|y-x\|\le\varepsilon\right\}$$

Точки x, будет

$$\varphi(y)\ge\varphi(x)$$

Рассмотрим произвольную точку $z\ne x$ множества X. В силу выпуклости множества X будет $x-\beta(x-z)\in U_\varepsilon(x)$ для всех $\beta\in\left(0,\overline{\beta}\right]$ при

$$\overline{\beta}=\min\left\{1,\frac{\varepsilon}{\|z-x\|}\right\}$$

Поэтому

$$\lim_{\beta\to+0}\frac{\varphi(x)-\varphi\left(x-\beta(x-z)\right)}{\beta}=\left\langle\varphi'(x),x-z\right\rangle\le 0.$$

Если положить $s=x-z$, то направление $-s$ будет возможным в точке x, и последнее неравенство примет вид

$$\langle \varphi'(x), s \rangle \le 0. \qquad (3.16)$$

Поскольку $-s$ — любое возможное направление в точке x, то из теоремы 3.2.8 следует, что неравенство (3.16) должно выполняться для всех s, удовлетворяющих неравенствам

$$\langle a_i, s \rangle \le 0, \qquad i \in I(x) \qquad (3.17)$$

Применяя к условиям (3.16) и (3.17) теорему 2.3.3, сразу же получаем утверждение теоремы. Δ

3.4.7. Поскольку при выводе условия (3.16) мы пользовались лишь свойством выпуклости множества X, то можно сформулировать следующее условие стационарности точки x в задаче выпуклого программирования: *для того чтобы точка x выпуклого множества X являлась точкой локального минимума функции $\varphi(x)$ на множестве X, необходимо, чтобы в этой точке производные по всем возможным направлениям были неотрицательными:*

$$\frac{\partial \varphi(x)}{\partial(-s)} \ge 0.$$

3.5. Достаточные условия оптимальности

3.5.1. *Теорема. Для того чтобы точка $x \in X$ была точкой глобального минимума основной задачи выпуклого программирования* 3.1.4, *достаточно существования таких чисел $u_i \ge 0, i \in I(x)$, что*

$$\varphi'(x) = \sum_{i \in I(x)} u_i f_i'(x).$$

Доказательство. Вследствие выпуклости множества X направление $-s = y - x$ является возможным в точке x при любом $y \in X$. Из теоремы 3.2.7 следует, что $\langle f_i'(x), s \rangle \le 0, i \in I(x)$, и поскольку функция $\varphi(x)$ выпукла, то, пользуясь (2.15), получаем соотношение

$$\varphi(x) - \varphi(y) \le \langle \varphi'(x), s \rangle = \left\langle \sum_{i \in I(x)} u_i f_i'(x), s \right\rangle = \sum_{i \in I(x)} u_i \langle f_i'(x), s \rangle \le 0,$$

справедливое для любого $y \in X$. Δ

3.5.2. Теоремы 3.4.4 и 3.5.1 можно объединить в одну теорему о критерии оптимальности.

Теорема Куна-Таккера (дифференцируемый случай). *Если функции* $f_i(x)$ *вогнуты, функция* $\varphi(x)$ *выпукла, множество* $X = \{x : f_i(x) \geq 0, i = \overline{1,m}\}$ *регулярно по Слейтеру, то для оптимальности точки* $x \in X$ *необходимо и достаточно существования таких чисел* $u_i \geq 0 \left(i = \overline{1,m}\right)$, *что*

$$\varphi'(x) = \sum_{i=1}^{m} u_i f_i'(x), \qquad \sum_{i=1}^{m} u_i f_i(x) = 0.$$

Для того чтобы убедиться в справедливости этой теоремы, достаточно вспомнить, что точка локального минимума выпуклой функции является оптимальной (теорема 2.4.11). Δ

3.5.3. Часто в допустимом множестве X выделяют ограничения неотрицательности допустимых точек:

$$X = \{x : f_i(x) \geq ,i = \overline{1,m}, x \geq 0\}.$$

В этом случае теорему Куна - Таккера формулируют следующим образом: *если функции* $f_i(x)$ *вогнуты, функция* $\varphi(x)$ *выпукла, множество* X *регулярно по Слейтеру, то для оптимальности точки* $x \in X$ *необходимо и достаточно существования таких чисел* $u_i \geq 0(i = \overline{1,m})$ *и* $v_j \geq 0(j = \overline{1,n})$, *что*

$$\varphi'(x) = \sum_{i=1}^{m} u_i f'(x) + \sum_{j=1}^{n} v_j e_j,$$

$$\sum_{i=1}^{m} u_i f_i(x) = 0, \qquad \sum_{j=1}^{n} v_j x_j = 0.$$

Здесь $e_j (j = \overline{1,n})$ — j-й координатный вектор: $e_j^T = (0,...,1,...,0)$, единица стоит на j-*и* месте.

3.5.4. Случай линейных ограничений (критерий оптимальности). Теоремы 3.4.6 и 3.5.1 можно также объединить в одну теорему о критерии оптимальности: *для того чтобы точка* x *была точкой глобального минимума выпуклой функции* $\varphi(x)$ *на множестве* $X = \{x : \langle a_i, x\rangle - b_i \geq 0, i = \overline{1,m}\}$, *необходимо и достаточно, чтобы существовали такие числа* $u_i \geq 0 (i = \overline{1,m})$, *что*

$$\varphi'(x) = \sum_{i=1}^{m} u_i a_i, \qquad \sum u_i\left(\langle a_i, x\rangle - b_i\right) = 0.$$

3.5.5. *Теорема. Для того чтобы точка* $x \in X$ *была точкой глобального минимума основной задачи выпуклого программирования* 3.1.4, *достаточно, чтобы для всех* s, *удовлетворяющих системе*

$$\langle f_i'(x), s\rangle \leq 0, \qquad i \in I(x), \tag{3.18}$$

выполнялось условие

$$\langle \varphi'(x), s\rangle \leq 0. \tag{3.19}$$

Доказательство. Применяя к условиям (3.18) и (3.19) теорему 2.3.3 (Фаркаша), сразу же приходим к условиям теоремы 3.5.1. Δ

3.6. Функция Лагранжа. Условия оптимальности

3.6.1. Седловая точка. Рассмотрим n-мерный вектор x, принадлежащий некоторому выпуклому множеству Γ, и m-мерный неотрицательный вектор $y \geq 0$. Пусть функция $L(x, y)$ выпукла по x на множестве Γ и вогнута по y на неотрицательном ортанте.

Определение. Пара x^*, y^* называется *седловой точкой* функции $L(x, y)$ на множестве всех $x \in \Gamma$ и $y \geq 0$, если

$$x^* \in \Gamma, \quad y^* \geq 0,$$

$$L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*) \tag{3.20}$$

для всех $x \in \Gamma$ и $y \geq 0$.

Соотношение (3.20) можно записать также следующим образом:

$$L(x^*, y^*) = \min_{x \in \Gamma} \max_{y \geq 0} L(x, y) = \max_{y \geq 0} \min_{x \in \Gamma} L(x, y).$$

3.6.2. Условия существования седловой точки. Для дальнейших целей нам достаточно рассмотреть случаи, когда $\Gamma = E_n^+ = \{ x : x \geq 0 \}$ и $\Gamma = E_n$.

Пусть функция $L(x, y)$ выпукла по x для всех $x \geq 0$, вогнута по y для всех $y \geq 0$ и непрерывно дифференцируема по x и по y.

Теорема. Для того чтобы пара $x^*, y^* (x^* \geq 0, y^* \geq 0)$ была седловой точкой функции $L(x, y)$ в области $x^* \geq 0, y^* \geq 0$, необходимо и достаточно выполнения

$$\frac{\partial L^*}{\partial x} \geq 0, \tag{3.21}$$

$$\left\langle x^*, \frac{\partial L^*}{\partial x} \right\rangle = 0, \tag{3.22}$$

$$x^* \geq 0, \tag{3.23}$$

$$\frac{\partial L^*}{\partial y} \leq 0, \tag{3.24}$$

$$\left\langle y^*, \frac{\partial L^*}{\partial y} \right\rangle = 0, \tag{3.25}$$

$$y^* \geq 0, \tag{3.26}$$

где

$$\frac{\partial L^*}{\partial x} \equiv \left. \frac{\partial L(x, y)}{\partial x} \right|_{\substack{x = x^* \\ y = y^*}}, \qquad \frac{\partial L^*}{\partial y} \equiv \left. \frac{\partial L(x, y)}{\partial y} \right|_{\substack{x = x^* \\ y = y^*}}.$$

Доказательство. Запишем условия (3.21) – (3.26) в эквивалентной форме:

$$\frac{\partial L^*}{\partial x_i} \geq 0, \qquad i = \overline{1, n}, \tag{3.21'}$$

$$x_i^* \frac{\partial L^*}{\partial x^*} = 0, \qquad i = \overline{1, n}, \tag{3.22'}$$

$$x_i^* \geq 0, \qquad i = \overline{1, n}, \tag{3.23'}$$

$$\frac{\partial L^*}{\partial y_j} \leq 0, \qquad j = \overline{1, m}, \tag{3.24'}$$

$$y_j^* \frac{\partial L^*}{\partial y_j} = 0, \qquad j = \overline{1, m}, \tag{3.25'}$$

$$y_j^* \geq 0, \qquad j = \overline{1, m}. \tag{3.26'}$$

Необходимость. Пусть выполняются условия (3.20) при $\Gamma = \{x : x \geq 0\}$.В частности, отсюда следует $L(x_i, y^*) \equiv L(x_1^*, ..., x_{i-1}^*, x_i, x_{i+1}^*, ..., x_n^*, y^* \geq L(x^*, y^*)$ для всех $x_i \geq 0$, т. е. точка x_i^* является точкой минимума выпуклой функции одного переменного $L(x_i, y^*)$ на полупрямой $x_i \geq 0$. Условия (3.21')-(3.23') и являются необходимыми условиями минимума (в частности, локального минимума) при $x_i \geq 0$ для функции одного переменного (поскольку либо x_i^* — внутренняя точка полуоси $x_i \geq 0$, и тогда $\frac{\partial L^*}{\partial x_i} = 0$, либо $x_i^* = 0$, и тогда $\frac{\partial L^*}{\partial x_i} \geq 0$). Аналогично, пользуясь вогнутостью функции $L(x, y)$ по y, легко доказать справедливость условий (3.24')-(3.26').

Достаточность. Пусть выполняются условия (3.21)-(3.26). Поскольку $L(x, y)$ выпукла по x при $x \geq 0$, то, пользуясь неравенством из 2.4.9, получаем

$$L(x,y^*) \geq L(x^*,y^*) + \left\langle x - x^*, \frac{\partial L^*}{\partial x} \right\rangle.$$

Отсюда и из (3.21)-(3.23) получаем

$$L(x^*,y^*) \leq L(x,y^*), \qquad x \geq 0.$$

Аналогично доказывается и левое неравенство в (3.20). Δ

3.6.3. Если $\Gamma = E_n$, то аналогичными рассуждениями нетрудно убедиться, что седловая точка будет определяться условием

$$\frac{\partial L^*}{\partial x} = 0$$

и соотношениями (3.24)-(3.26).

3.6.4.Функция Лагранжа. Пусть $f(x)$ — m-мерный вектор $f^T(x) = (f_1(x), f_2(x), ..., f_m(x))$. Рассмотрим следующую задачу выпуклого программирования:

$$\min\{\varphi(x) : x \in X\}, \qquad X = \{x \in \Gamma : f(x) \geq 0\}. \quad (3.27)$$

Здесь Γ — выпуклое и замкнутое множество, $\varphi(x)$ — выпуклая функция, а все функции $f_i(x)$ вогнутые. Заметим, что при $\Gamma = E_n$ задача (3.27) является (согласно принятой терминологии) основной задачей выпуклого программирования.

Определение. Функцию

$$L(x,y) = \varphi(x) - \langle y, f(x) \rangle, \quad (3.28)$$

определенную при всех $x \in E_n$ и $y \geq 0$, называют *функцией Лагранжа для* задачи выпуклого программирования (3.27).

3.6.5. В задачах классического анализа об условном экстремуме (задачах, в которых допустимое множество задается системой уравнений) важную роль играет метод множителей Лагранжа: решение исходной задачи ищется среди стационарных точек функции $L(x, y)$ — точек, удовлетворяющих системе уравнений

$$\frac{\partial L}{\partial x} = 0, \qquad \frac{\partial L}{\partial y} = 0.$$

В задачах выпуклого (и, в частности, линейного) программирования функции Лагранжа также отводится важное место: при весьма общих предположениях задача выпуклого программирования сводится к отысканию седловых точек функции Лагранжа.

3.6.6. Достаточные условия оптимальности.

Теорема. *Если пара* x^*, y^* *является седловой точкой функции Лагранжа* (3.28) *на множестве* $x \in \Gamma, y \ge 0$, *то* x^* *— оптимальная точка задачи выпуклого программирования* (3.27).

Доказательство. Из (3.28) и определения седловой точки (3.20) получаем неравенства

$$\varphi(x^*) - \left\langle y, f(x^*) \right\rangle \le \varphi(x^*) - \left\langle y, f(x^*) \right\rangle \le \varphi(x) - \left\langle y^*, f(x) \right\rangle, \quad (3.29)$$

справедливые для всех $x \in \Gamma, y \ge 0$. Из левого неравенства следует

$$\left\langle y, f(x^*) \right\rangle \ge \left\langle y^*, f(x^*) \right\rangle, \quad (3.30)$$

а поскольку $y^* \ge 0$ и это неравенство справедливо для любого $y \ge 0$, то $f(x^*) \ge 0$.

В частности, (3.30) имеет место и для $y = 0$, поэтому

$$\left\langle y^*, f(x^*) \right\rangle \le 0,$$

а следовательно (так как $y^* \ge 0$ и $f(x) \ge 0$),

$\langle y^* f(x^*) \rangle = 0$. (3.31)

Если $x \in X$, то из (3.27) следует, что $f(x) \ge 0$, и поэтому для $x \in X$ будет

$\langle y^*, f(x) \rangle \ge 0.$ (3.32)

Так как неравенство (3.29) выполняется для всех $x \in \Gamma$ и, в частности, для $x \in X$, то из правого неравенства (3.29) и из (3.31) и (3.32) получаем для всех $x \in X$ неравенства

$$\varphi(x^*) \le \varphi(x) - \langle y^*, f(x) \rangle \le \varphi(x).$$

Но $x^* \in X$ (так как $x^* \in \Gamma$ и $f(x^*) \ge 0$), и, следовательно, x^* —оптимальная точка. Δ

Замечание. При доказательстве теоремы нигде не использовались ни свойства выпуклости функции $\varphi(x)$ и множества Γ, ни вогнутость функций $f_i(x)(i = \overline{1,m})$, ни какие-либо свойства гладкости. Таким образом, наличие седловой точки x^*, y^* функции Лагранжа определяет оптимальность точки x^* для общей задачи математического программирования. Сразу же подчеркнем: обратное утверждение, что из оптимальности точки x^* следует существование седловой точки x^*, y^* функции Лагранжа, справедливо лишь для задачи выпуклого программирования и вдобавок при условии регулярности допустимого множества. Это и есть известная теорема Куна - Таккера. Ниже эта теорема будет доказана в предположении непрерывной дифференцируемости функций $\varphi(x)$ и $f_i(x)(i = \overline{1,m})$ как очевидное следствие теорем 3.5.2 и 3.6.2.

3.6.7. Теорема Куна - Таккера. Пусть допустимое множество задачи выпуклого программирования имеет вид

$$X = \{x : f_i(x) \geq 0, i = \overline{1,m}, x \geq 0\}.$$

Будем предполагать, что выпуклая функция $\varphi(x)$ и вогнутые функции $f_i(x)(i = \overline{1,m})$ непрерывно дифференцируемы.

Теорема Куна - Таккера. *Если в задаче выпуклого программирования*

$$\min\{\varphi(x) : x \in X\}, \qquad X = \{x : f_i(x) \geq 0,\ i = \overline{1,m},\ x \geq 0\},$$

множество X *обладает свойством регулярности* 3.4.2, *то необходимым и достаточным условием оптимальности точки* $x^* \in X$ *является существование такого* $y^* \geq 0$, *чтобы пара* x^*, y^* *являлась седловой точкой функции Лагранжа*

$$L(x, y) = \varphi(x) - \langle y, f(x)\rangle$$

на множестве $x \geq 0, y \geq 0$.

Доказательство. Достаточность следует из теоремы 3.6.6 при $\Gamma = \{x : x \geq 0\}$.

Необходимость. Докажем эквивалентность условий теоремы 3.5.3:

$$\varphi'(x^*) = \sum_{i=1}^{m} y_i^* f_i'(x^*) + \sum_{j=1}^{n} v_j^* e_j, \qquad (3.33)$$

$$\sum_{i=1}^{m} y_i^* f_i(x^*) = 0, \qquad (3.34)$$

$$y_i^* \geq 0, \qquad i = \overline{1,m}, \qquad (3.35)$$

$$\sum_{j=1}^{n} v_j^* x_j^* = 0, \qquad (3.36)$$

$$v_j^* \geq 0, \qquad j = \overline{1,n}, \qquad (3.37)$$

и условий теоремы 3.6.2, определяющих седловую точку функции $L(x, y)$:

$$\frac{\partial L^*}{\partial x} \geq 0, \quad (3.38)$$

$$\left\langle x^*, \frac{\partial L^*}{\partial x} \right\rangle = 0, \quad (3.39)$$

$$x^* \geq 0, \quad (3.40)$$

$$\frac{\partial L^*}{\partial y} \leq 0, \quad (3.41)$$

$$\left\langle y^*, \frac{\partial L^*}{\partial y} \right\rangle = 0, \quad (3.42)$$

$$y^* \geq 0. \quad (3.43)$$

Эквивалентность будем обозначать символом ~.

Пусть $v^T = (v_1^*, v_2^*, ..., v_n^*)$ и $f^T(x) = (f_1(x), f_2(x), ..., f_m(x))$. Заметим, что

$$\frac{\partial L^*}{\partial x} = \varphi'(x^*) - \sum_{i=1}^{m} y_i^* f_i'(x^*) \cong v,$$
$$\frac{\partial L^*}{\partial y} = -f(x^*).$$

Очевидно,

(3.38) ~ (3.37),

(3.39) ~ (3.36),

(3.42) ~ (3.34),

(3.43) ~ (3.35).

Условия (3.40) и (3.41) означают, что $x^* \in X$.∆

3.6.8. Для основной задачи выпуклого программирования 3.1.4 теорема Куна - Таккера формулируется следующим образом: *если в основной задаче выпуклого программирования множество* $X = \{x : f_i(x) \le 0, \quad i = \overline{1,m}\}$ *обладает свойством регулярности* 3.4.2, *то необходимыми (достаточным условием оптимальности точки* $x^* \in X$ *является существование такого* $y^* \ge 0$, *чтобы пара* x^*, y^* *являлась седловой точкой функции Лагранжа на множестве* $x \in E_n$ *и* $y^* \ge 0$.

Справедливость этой теоремы следует из пп. 3.5.2 и 3.6.3.

3.6.9. Случай линейных ограничений. *Если функция* $\varphi(x)$ *выпукла, а допустимое множество задается линейными ограничениями,*

$$X = \{x : \langle a_i, x\rangle - b_i \le 0, i = \overline{1,m}, x \ge 0\},$$

то для оптимальности точки $x^* \in X$ *необходимо и достаточно существование такого* $y^* \ge 0$, *чтобы пара* x^*, y^* *являлась седловой точкой функции Лагранжа на множестве* $x \ge 0, y \ge 0$.

Справедливость этой теоремы следует из пп. 3.5.4 и 3.6.2. ∆

3.6.10. Существуют различные варианты теоремы Куна - Таккера, различные ее обобщения и применения в теории экстремальных задач. Эта теорема лежит в основе теории двойственности математического программирования.

Теорема Куна - Таккера находит также применение в численных методах решения задач математического программирования. Она позволяет исходную задачу заменить задачей отыскания седловой точки функции Лагранжа, т. е. задачей вида

$$\min_{x \in \Gamma} \max_{y \geq 0} L(x, y).$$

"Простые" ограничения этой задачи позволяют применять для ее решения методы, во многом схожие с численными методами безусловной оптимизации, достаточно хорошо изученные и апробированные. Замечание. Теорема Куна - Таккера перестает быть содержательной в случае вырожденности, например, когда

$$\varphi'(x^*) \equiv 0, \qquad f_i'(x^*) \equiv 0, \qquad i = \overline{1,m}.$$

В то же время существует широкий класс задач такого рода. В добавлении (см. Д.3) приводятся необходимые и достаточные условия оптимальности в вырожденных задачах.

3.6.11. Двойственность. Введем функцию $g(x) = \sup_{y \geq 0} L(x, y).$ Очевидно, что основная задача математического программирования может быть записана в виде

$$g(x) \to \min, \qquad x \in X = \{x : f(x) \leq 0\},$$

поскольку $g(x) = \varphi(x)$, если $x \in X$.Эту задачу принято называть *прямой*.

Обозначим $h(y) = \inf_{x \in E_n} L(x, y)$.Задачу

$$h(y) \to \min, \qquad y \in E_m^+ = \{y : y \geq 0\},$$

называют *двойственной,* а переменные $y_1, y_2, \ldots, y_m$ — *двойственными переменными.*

3.6.12. *Теорема.* 1) $\varphi(x) \geq h(y)$ *для всех* $x \in X, y \in E_m^+$.

2) *Если выполняются условия теоремы Кун - Таккера (п.* 3.6.7), *а пара* x^*, y^* *является седловой точкой функции Лагранжа, то*

$$y^* = \arg\max\{h(y) : y \in E_m^+\} \text{ и } \varphi(x^*) = h(y^*).$$

3) *Если* $\varphi\left(x^*\right) = h\left(y^*\right)$ *для любых* $x^* \in X, y^* \in E_m^+$, *то*

$$x^* = \arg\min\{\varphi(x) : x \in X\}, \qquad y^* = \arg\min\{h(y) : y \in E_m^+\}.$$

Доказательство. 1) Так как $f(x) \le 0, y \ge 0$, то

$$\varphi(x) \ge \varphi(x) - \langle y, f(x)\rangle = L(x,y) \ge \inf_{x \in E_n} L(x,y) = h(y).$$

2) Для всех $y \ge 0$ справедливо соотношение

$$h(y^*) = \inf_{x \in E_n} L(x,y^*) = L(x^*,y^*) \ge L(x^*,y) \ge \inf_{x \in E_n} L(x,y) = h(y),$$

поэтому y^* — решение двойственной задачи. Но $L(x^*,y^*) = \varphi(x^*)$, так как $\langle y^*, f(x^*)\rangle = 0$ (см. п. 3.31), поэтому $\varphi(x^*) = h(y^*)$.

3) В силу 1)

$$\varphi(x) \ge h(y^*) = \varphi(x^*) \ge h(y)$$

для любых $x \in X, y \in E_m^+$, следовательно, x^* — решение прямой задачи, а y^* — двойственной.

Определение. Точку $v^* = (x^*, y^*)$, удовлетворяющую условиям

$$\nabla\varphi(x^*) - \sum_{i=1}^{m} y_i^* \nabla f_i(x^*) = 0, \quad y_i^* \ge 0, \quad y_i^* f_i(x^*) = 0 \quad, i = \overline{1,m},$$

будем называть *точкой Куна - Таккера.*

Часть 2

1. **Функциональные пространства $C^k(D)$**

Множество функций $f(x)$, определенных на некотором отрезке $[a,b] \subset R$ и имеющих на этом отрезке непрерывные производные порядка k образуют линейное пространство $C^k[a,b]$. Например: пространство непрерывных функций $C^0[a,b]$; пространство функций, имеющих непрерывную первую производную $C^1[a,b]$ и так далее. Элементы этого пространства часто обозначают $f(\cdot)$ или просто f. Здесь точка в круглых скобках указывает на наличие аргумента, но его несущественную роль в обозначении элемента. Если в этих пространствах определить следующие нормы:

в $C^0[a,b]$– норму

$$\|f\|_0 = \max_{x \in [a,b]} |f(x)|,$$

в $C^k[a,b]$ – норму

$$\|f\|_k = \max_i \left\{ \|f\|_0, \|f'\|_0, \ldots, \|f^{(k)}\|_0 \right\} = \max_i \left\{ \|f^{(i)}\|_0 \right\} \qquad (i = 0,1,\ldots,k),$$

то эти пространства будет полными нормированными, или банаховыми. На основе указанной нормы можно определить расстояние между «точками» этого пространства – некоторыми функциями $f(x)$ и $g(x)$

$$\rho_k(f, g) = \|f - g\|_k .$$

Пусть $y_0(x) \in C^k[a,b]$ – некоторая функция. Множество функций $y(x) \in C^k[a,b]$, для которых $\rho_k(y, y_0) < \delta$ называют δ– окрестностью функции $y_0(x)$. Так как имеет место последовательное включение пространств $C^0[a,b] \supset C^1[a,b] \supset \ldots \supset C^k[a,b]$, в пространстве $C^k[a,b]$ можно использовать нормы более широких пространств $C^m[a,b]$ $(m<k)$. В этом случае δ – окрестность будет содержать функции более широкого класса; принято δ – окрестность по норме $\|\cdot\|_0$ называть сильной δ– окрестностью, а по норме $\|\cdot\|_1$ – слабой δ– окрестностью. Согласно определению, слабая δ– окрестность всегда содержится в сильной δ– окрестности. Заметим, что пространство $C^k[a,b]$ по отношению к норме $\|\cdot\|_m$ $(m < k)$ перестает быть банаховым. Это следует, например, из того факта, что не всякая непрерывная функция, являющаяся предельной для последовательности дифференцируемых функций, будет обязательно дифференцируемой.

Отображение $I{:}D \to R$ множества $D \subset C^k[a,b]$ на множество действительных чисел R называют функционалом и обозначают $I[y(\cdot)]$; множество D называют множеством допустимых функций для функционала I. Функционал $I[y(\cdot)]$ называется непрерывным по норме $\|\cdot\|_k$ в «точке» $y_0(\cdot) \in D$, если для любого $\varepsilon > 0$ можно указать $\delta = \delta(y_0, \varepsilon) > 0$, такое что, как только $\rho_k(y, y_0) < \delta$ и $y(\cdot) \in D$, то $|I[y(\cdot)] - I[y_0(\cdot)]| < \varepsilon$. Для любой $y_0(\cdot) \in D$ разность $\delta y = y - y_0$ называют вариацией функции $y_0(\cdot)$; здесь $y = y(\cdot) \in D$. Таким образом, вариация функции сама является функцией из того же пространства, то есть $\delta y = \delta y(\cdot) \in C^k[a,b]$. Будем говорить, что «точка» $y(\cdot)$ является внутренней для множества D, если она входит в него вместе с некоторой своей δ– окрестностью.

Рассмотрим приращение функционала в произвольной внутренней «точке» $y(\cdot) \in D$

$$\Delta I = \Delta I[y(\cdot), \delta y] = I[y(\cdot) + \delta y] - I[y(\cdot)].$$

Если в приращении функционала можно выделить линейную часть по отношению к вариации δy, то есть представить

$$\Delta I = L\,[y(\cdot), \delta y] + o\,(\|\delta y\|_k),$$

где $L\,[y(\cdot), \delta y]$ – линейный функционал относительно δy, а $o(\|\delta y\|_k)$, – величина, имеющая более высокий порядок малости, чем δy при $\|\delta y\|_k \to 0$, то линейная часть $L\,[y(\cdot), \delta y]$ называется дифференциалом функционала $I[y(\cdot)]$, а функционал считается дифференцируемым по Фреше (Фреше́ Морис Рене – французский математик).

Рассмотрим однопараметрическое семейство функций

$$y(x, \alpha) = y(x) + \alpha\,\delta y,$$

где $\alpha \in R$ – параметр, $y(\cdot) \in D$ – внутренняя «точка». Первой вариацией $\delta I[y(\cdot), \delta y]$ функционала I в «точке» $y(\cdot)$ называют предел

$$\delta I[y,\delta y] = \lim_{\alpha \to 0} \frac{I[y+\alpha\delta y] - I[y]}{\alpha} = \frac{\partial}{\partial\alpha} I[y+\alpha\delta y]\bigg|_{\alpha=0}.$$

Первую вариацию также называют дифференциалом Гато (Гато́ Рене Эжен – французский математик). Заметим, что из дифференцируемости по Фреше функционала I следует существование его первой вариации, которая в этом случае совпадает с дифференциалом, то есть $\delta I[y,\delta y] = L\,[y,\delta y]$. Однако существование первой вариации функционала I еще не означает его дифференцируемости в соответствующей «точке» $y(\cdot)$, подобно тому, как существование производной по любому направлению в данной точке для функции многих переменных еще не означает ее дифференцируемости в этой точке.

Вариации функционала I более высокого порядка можно определить как коэффициенты в разложении в ряд Тейлора в окрестности $\alpha=0$ функции $\varphi(\alpha)=I[y+\alpha\delta y]$, в которую превращается функционал на семействе функций $y(x,\alpha)=y(x)+\alpha\delta y$:

$$\Delta\varphi(\alpha) = I[y+\alpha\delta y] - I[y] = \alpha\delta I + \frac{\alpha^2}{2!}\delta^2 I + \ldots + \frac{\alpha^k}{k!}\delta^k I + \ldots;$$

здесь $\delta^k I = \dfrac{d^k\varphi(\alpha)}{d\alpha^k}\bigg|_{\alpha=0} = \dfrac{\partial^k I[y+\alpha\delta y]}{\partial\alpha^k}\bigg|_{\alpha=0}$ – вариация функционала порядка k.

Конечно, здесь функция $\varphi(\alpha)$ должна быть дифференцируема соответствующее число раз.

Говорят, что функционал $I[y]$, определенный на линейном пространстве $C^1[a,b]$, достигает сильного минимума в «точке» $y^*(\cdot)\in C^1[a,b]$, если найдется такая сильная ε – окрестность функции $y_*(x)$, что для любой функции $y(x)$ из этой окрестности выполняется неравенство $I[y^*(\cdot)]\le I[y(\cdot)]$. Минимум называется слабым, если все функции $y(x)$ берутся только из слабой ε – окрестности. Аналогично определяются сильный и слабый максимумы. Все максимумы и минимумы объединяются общим названием экстремумы.

Поскольку всякая функция, принадлежащая слабой ε – окрестности функции $y^*(x)$, заведомо входит в ее же сильную ε – окрестность, то всякий сильный экстремум одновременно является и слабым.

Говорят, что функционал $I[y(\cdot)]$, определенный на множестве D кривых $y(x)$, достигает на кривой $y^*(x)$ глобального минимума, если

$$I[y^*(x)]\le I[y(x)] \quad \forall y(x)\in D;$$

аналогично определяется глобальный максимум.

Методы решения вариационных задач основаны на следующем утверждении.

Теорема (Необходимое условие экстремума функционала). Если функционал $I[y]$, достигает экстремума во внутренней точке $y^*(x)$ своей области определения и в этой точке существует первая вариация $\delta I[y^*(\cdot),\delta y]$, то $\delta I[y^*(\cdot),\delta y]=0$.

При выводе необходимых условий экстремума для различных постановок вариационных задач применяется следующая важная теорема.

Теорема (Основная лемма вариационного исчисления). Если для каждой непрерывной функции $\delta y(x)$

$$\int_a^b A(x)\delta y(x)dx = 0,$$

где функция $A(x)$ непрерывна на отрезке $[a,b]$, то $A(x) \equiv 0$ на том же отрезке.

Замечания. 1. Утверждение основной леммы вариационного исчисления не изменится, если на функцию $\delta y(x)$ наложить следующее ограничение $\delta y(a) = \delta y(b) = 0$.

2. Все изложенное в этом разделе без существенных изменений переносится на функционалы $I[y(x)] = I[y_1(x),\ldots, y_n(x)]$, зависящие от вектор-функции $y(x) = (y_1(x),\ldots, y_n(x))^T$ одной переменной, а также на функционалы, зависящие от функций нескольких переменных.

Приведем ряд примеров, иллюстрирующих свойства функционалов и функциональных пространств.

Пример 1. Показать, что функционал

$$I[y(\cdot)] = \int_a^b \sqrt{1 + (y'(x))^2}\,dx, \tag{1}$$

определенный на множестве функций $y = y(x)$, непрерывных вместе с первой производной на отрезке $[a, b]$, непрерывен в пространстве $C^1[a,b]$, но не является непрерывным в $C^0[a,b]$.

Решение. Вначале покажем, что функционал (1) непрерывен в пространстве $C^1[a,b]$. Действительно, имеем соотношения

$$\left|I[y] - I[y_0]\right| = \left|\int_a^b \left(\sqrt{1 + y'^2} - \sqrt{1 + y_0'^2}\right)dx\right| \le \int_a^b \frac{|y' + y_0'|\,|y' - y_0'|}{\sqrt{1 + y'^2} + \sqrt{1 + y_0'^2}}\,dx.$$

Из непрерывности производных y' и y_0' следует, что

$$\frac{|y' + y_0'|}{\sqrt{1 + y'^2} + \sqrt{1 + y_0'^2}} \le M.$$

Далее, из определения нормы в пространстве $C^1[a,b]$ заключаем, что

$$\max_{x\in[a,b]}|y' - y_0'| = \|y' - y_0'\|_0 \le \|y - y_0\|_1.$$

По заданному $\varepsilon > 0$ выберем $\delta = \varepsilon/(M(b-a))$. Тогда для всех $y(\cdot)\in C^1[a,b]$ и таких, что $\|y - y_0\|_1 < \delta$, имеем

$$\left|I[y] - I[y_0]\right| \le M \max_{x\in[a,b]}|y' - y_0'|\,(b - a) \le M(b - a)\|y - y_0\|_1 < \varepsilon,$$

а это означает, что функционал (1) непрерывен в пространстве $C^1[a,b]$.

Функционал (1) не будет непрерывным в пространстве $C^0[a,b]$, так как он не ограничен в любой сильной δ - окрестности любой «точки» $y(\cdot)\in C^1[a,b]$ ввиду неограниченности всей совокупности значений производных функций $y(x)\in C^1[a,b]$.

Пример 2. Показать, что функционал

$$L(f) = \int_a^b \alpha(x)f(x)dx, \tag{2}$$

где $\alpha(x)$ - непрерывная фиксированная функция, является линейным в пространстве $C^0[a,b]$.

Решение. Аддитивность этого функционала очевидна. Покажем его непрерывность. Учитывая, что функция $\alpha(x)$ ограничена ($|\alpha(x)| < M$), оценим модуль разности; имеем

$$|L(f) - L(f_1)| \le \int_a^b |\alpha(x)|\,|f(x) - f_1(x)|dx \le$$

$$\le M \max_{x\in[a,b]} |f(x) - f_1(x)|(b-a) \le M\|f - f_1\|_0 (b-a) < \varepsilon\ ,$$

как только норма $\|f - f_1\|_0 < \dfrac{\varepsilon}{M(b-a)}$. А это означает, что функционал (2) непрерывен.

Функционал (1) не является линейным в пространстве $C^0[a,b]$, ибо для него не выполнены условия непрерывности и аддитивности. Этот же функционал не будет линейным и в пространстве $C^1[a,b]$; хотя он и непрерывен, но не является аддитивным.

Пример 3. Найти расстояния $\|y - y_0\|_0$, $\|y - y_0\|_1$ между кривыми $y(x)=x^2$ и $y_0(x)=x^3$ в пространствах $C^0[0,1]$ и $C^1[0,1]$.

Решение. Найдем расстояние в пространстве $C^0[0,1]$:

$$\rho_0(x,y) = \|y - y_0\|_0 = \max_{x\in[0,1]} |x^2 - x^3|.$$

Рассмотрим функцию $f(x)= x^2 - x^3$. Из необходимого условия экстремума $f'(x) = 0$ получаем $2x - 3x^2=0$, или $x_1=0$, $x_2=2/3$. Сравнивая значения функции $f(x)$ в точках экстремума и на концах промежутка $[0,1]$, $f(0)=f(1)=0$, устанавливаем искомое расстояние:

$$\rho_0(x,y) = |f(x)|_{x=2/3} = 4/27\,.$$

Найдем расстояние в пространстве $C^1[0,1]$:

$$\rho_1(x,y) = \|y - y_0\|_1 = \max\{\|f\|_0, \|f'\|_0\} = \max\{\|x^2 - x^3\|_0, \|2x - 3x^2\|_0\}.$$

Исследуя дополнительно функцию $g(x)= 2x - 3x^2$ на экстремум, из условия $g'(x)=0$ получаем $2 - 6x= 0$, или $x_3=1/3$ – стационарная точка. Сравнивая значения функции $g(x)$ в стационарной точке и на концах отрезка $[0,1]$, $g(0)=0$, $g(1/3)= 1/3$, $g(1)= -1$, устанавливаем

$$\|g\|_0 = \|2x - 3x^2\|_0 = |2x - 3x^2|_{x=1} = 1,$$

то есть

$$\rho_1(x,y) = \max\{4/27, 1\} = 1\,.$$

Пример 4. Найти первую вариацию функционала

$$I[y(\cdot)] = \int_a^b y^2(x)\,dx\,.$$

Решение. Сначала найдем вариацию функционала как линейную часть его приращения

$$\Delta I = \int_a^b [y(x) + \delta y(x)]^2 dx - \int_a^b y^2(x)\,dx = \int_a^b 2y(x)\cdot\delta y(x)dx + \int_a^b [\delta y(x)]^2\,dx\,.$$

Заметим, что первое слагаемое линейно относительно вариации $\delta y(x)$; второе слагаемое имеет более высокий порядок малости при $\|\delta y\|_0 \to 0$. Действительно

$$\int_a^b [\delta y(x)]^2\, dx \le \int_a^b [\max_{x\in[a,b]}|\delta y(x)|]^2\, dx = [\max_{x\in[a,b]}|\delta y(x)|]^2 (b-a) = (b-a)\|\delta y\|_0^2 .$$

Таким образом

$$\delta I = \int_a^b 2y(x)\cdot\delta y(x)dx .$$

Найдем вариацию функционала другим способом. Рассмотрим изменение функционала на семействе функций $y(x,\alpha)=y(x)+\alpha\,\delta y(x)$:

$$I[y(x)+\alpha\,\delta y(x)] = \int_a^b [y(x)+\alpha\,\delta y(x)]^2 dx .$$

Тогда

$$\delta I = \left.\frac{\partial I[y(x)+\alpha\,\delta y(x)]}{\partial\alpha}\right|_{\alpha=0} = \left.\int_a^b 2[y(x)+\alpha\,\delta y(x)]\delta y(x)dx\right|_{\alpha=0} = \int_a^b 2y(x)\delta y(x)dx .$$

Конечно, оба способа приводят к одному результату.

Пример 5. Доказать, что на кривой $y^*(x)=x$ функционал

$$I[y(\cdot)] = \int_0^1 y'^2(x)\,dx;\ \ y(0)=0,\ \ y(1)=1$$

достигает глобального минимума.

Решение. Очевидно, функция $y^*(x)=x\in C^1[0,1]$. Рассмотрим вариации $\delta y(x)\in C^1[0,1]$, удовлетворяющие условиям $\delta y(0)=\delta y(1)=0$. Исследуем приращение функционала:

$$I[y^*(x)+\delta y(x)] - I[y^*(x)] = \int_0^1 [y^{*\prime}(x)+\delta y'(x)]^2 dx - \int_0^1 [y^{*\prime}(x)]^2 dx =$$

$$= 2\int_0^1 y^{*\prime}(x)\delta y'(x)dx + \int_0^1 [\delta y'(x)]^2 dx = 2\delta y(x)\Big|_0^1 + \int_0^1 [\delta y'(x)]^2 dx \ge 0,$$

так как $y^{*\prime}(x)=1$. Поскольку кривая $y(x)=y^*(x)+\delta y(x)\in C^1[0,1]$ произвольна и $I[y(x)]=I[y^*(x)+\delta y(x)]\ge I[y^*(x)]$, то на функции $y^*(x)=x$ достигается глобальный минимум.

Пример 6. Доказать, что на кривой $y^*(x)=0$ функционал

$$I[y(\cdot)] = \int_0^\pi y^2(x)[3-y'^2(x)]dx;\ \ y(0)=y(\pi)=0$$

достигает слабого минимума.

Решение. Так как $I[y^*(x)]=0$, то согласно определению требуется доказать, что существует $\varepsilon>0$ такое, что для всех $y(x)$, удовлетворяющих условию

$$\|y(x)-y^*(x)\|_1 = \max_{x\in[0,\pi]}\left\{\|y(x)-y^*(x)\|_0, \|y'(x)-y^{*\prime}(x)\|_0\right\} =$$

$$= \max_{x\in[0,\pi]}\left\{\|y(x)\|_0, \|y'(x)\|_0\right\} < \varepsilon ,$$

справедливо неравенство $I[y(x)]\ge I[y^*(x)]=0$.

Пусть $\varepsilon = 1$, тогда для всех кривых из ε – окрестности первого порядка кривой $y^*(x) \equiv 0$ выполняются условия:

$$\max_{x\in[0,\pi]} |y(x)| < \varepsilon = 1, \ \max_{x\in[0,\pi]} |y'(x)| < \varepsilon = 1.$$

Поэтому $0 \le y^2(x) < 1$, $3 - y'^2(x) > 0$ и

$$I[y(\cdot)] = \int_0^{\pi} y^2(x)[3 - y'^2(x)]dx \ge 0,$$

что и требовалось доказать. Следовательно, на кривой $y^*(x) \equiv 0$ функционал достигает слабого минимума.

Исследуем функционал на наличие сильного минимума. При $\varepsilon = 1$ ε – окрестность нулевого порядка кривой $y^*(x) \equiv 0$ образуют кривые, удовлетворяющие условию

$$\|y(x) - y^*(x)\|_0 = \max_{x\in[0,\pi]} |y(x) - y^*(x)| = \max_{x\in[0,\pi]} |y(x)| < \varepsilon = 1 \quad .$$

Рассмотрим последовательность функций $y_n(x) = \frac{1}{2}\left(\frac{x}{\pi}\right)^n$ из этой ε – окрестности. Нетрудно заметить, что функционал на этих функциях принимает следующие значения

$$I[y_n(\cdot)] = \int_0^{\pi} y_n^2(x)[3 - y_n'^2(x)]dx = \frac{1}{4}\int_0^{\pi}\left(\frac{x}{\pi}\right)^{2n}[3 - \frac{n^2}{4\pi^2}\left(\frac{x}{\pi}\right)^{2n-2}]dx =$$

$$= \frac{48\pi^2 n - 2n^3 - n^2 - 12\pi^2}{16\pi(2n+1)(4n-1)},$$

которые становятся отрицательными, начиная уже с $n > n_0 = 15$. Аналогичные рассуждения справедливы при других значениях ε. Следовательно, на кривой $y^*(x) \equiv 0$ функционал не достигает сильного минимума.

2.1. Простейшая задача вариационного исчисления. Уравнение Эйлера

Простейшая задача вариационного исчисления формулируется следующим образом. Пусть функция $F(x,y,z)$ имеет непрерывные частные производные до второго порядка включительно по всем своим аргументам. Требуется среди всех функций $y(x) \in C^1[a,b]$, удовлетворяющих граничным условиям

$$y(a) = y_0, \ y(b) = y_1, \tag{1}$$

найти ту функцию, на которой достигается слабый экстремум функционала

$$I[y(\cdot)] = \int_a^b F(x, y, y')dx. \tag{2}$$

Другими словами, простейшая задача вариационного исчисления состоит в отыскании на множестве всех гладких кривых, проходящих через точки $M_0(a, y_0)$ и $M_1(b, y_1)$, той кривой, на которой функционал (2) достигает слабого экстремума (рис.1).

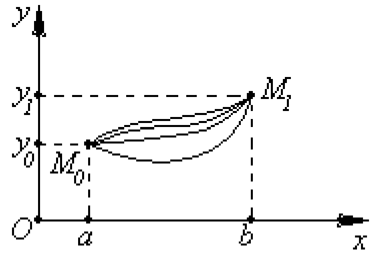

Рис. 1. Схема графиков семейства допустимых функций.

При решении простейшей задачи вариационного исчисления используется теорема, которая следует из необходимого условия экстремума, $\delta I=0$.

Теорема 1. Для того чтобы функционал (2) достигал на функции $y(x)\in C^1[a,b]$ слабого экстремума, необходимо, чтобы эта функция удовлетворяла уравнению Эйлера

$$F_y - \frac{d}{dx}F_{y'} = 0\,. \tag{3}$$

Решения (интегральные кривые) уравнения (3) называются экстремалями функционала (2).

Уравнение (3) в развернутом виде записывается следующим образом:

$$y''(x)F_{y'y'} + y'(x)F_{yy'} + F_{xy'} - F_y = 0\,.$$

Если $F_{y'y'} \neq 0$, то оно представляет собой обыкновенное дифференциальное уравнение второго порядка, поэтому его общее решение зависит от двух произвольных постоянных, которые находятся с помощью граничных условий (1).

Отметим что, так как всякий сильный экстремум функционала является и слабым, то теорема 1 дает необходимое условие и сильного экстремума функционала (2). Кроме того, так как абсолютный экстремум функционала (2) на множестве

$$G = \{\, y(x)\in C_1[a,b] \mid y(a) = y_0,\, y(b) = y_1 \} \tag{4}$$

является и локальным экстремумом (сильным и слабым), то теорема 1 определяет необходимое условие абсолютного экстремума функционала (2) на множестве (4).

Таким образом, решение краевой задачи

$$F_y - \frac{d}{dx}F_{y'} = 0,\;\; y(a) = y_0,\;\; y(b) = y_1, \tag{5}$$

позволяет найти все кривые возможного экстремума функционала (2) на множестве функций (4).

В отличие от задачи с начальными условиями, $y(a)=y_0,\ y'(a)=y'_0$ (задача Коши), задача с граничными условиями (1) может не иметь решение или иметь множество решений, даже если в окрестности точки $x=a$ выполняется теорема существования и единственности задачи Коши. Кроме этого, само уравнение Эйлера (3) может вырождаться в дифференциальное уравнение первого порядка ($F_{y'y'}\neq 0$) или даже вообще не быть дифференциальным ($F_{y'}=f(x)$).

Рассмотрим частные случаи интегрирования уравнения Эйлера.

1. $F=F(x,y)$, то есть подынтегральная функция в функционале (2) не зависит от y'. Уравнение (3) в этом случае принимает вид

$$F_y(x,y)=0.$$

Это конечное (не дифференциальное) уравнение, его решение $y=y(x)$ не содержит произвольных постоянных и, следовательно, удовлетворяет условиям (1) только в исключительных случаях. В остальных случаях задача отыскания экстремума функционала (2) не имеет решения в классе функций $y(x)\in C^1[a,b]$.

Пример. Найти функцию $y(x)\in C^1[0,b]$, на которой достигается экстремум функционала

$$I[y]=\int_0^b y(y-2x^2)dx,$$

удовлетворяющую граничным условиям $y(0)=0,\ y(b)=y_1$.

Решение. Уравнение Эйлера принимает вид $2y-2x^2=0$, отсюда $y=x^2$. Граничные условия удовлетворяются только, если $y_1=b^2$. В противном случае задача не имеет решения в пространстве $C^1[a,b]$.

2. $F(x,y)=M(x,y)+y'\cdot N(x,y)$, то есть подынтегральная функция в функционале (2) линейно зависит от y'. Уравнение (3) в этом случае принимает вид

$$M_y(x,y)-N_x(x,y)=0. \qquad (6)$$

Это уравнение не является дифференциальным, а его решение может не удовлетворять граничным условиям. Это означает, что в пространстве $C^1[a,b]$ экстремали функционала (2) отсутствуют, и исходная задача имеет решение в исключительных случаях. Заметим, что если условие (6) выполняется в некоторой области, имеющей граничные точки (a,y_0) и (b,y_1), то значение функционала не зависит от вида кривой $y(x)\in C^1[a,b]$. В этом случае функционал (2) можно рассматривать как криволинейный интеграл от дифференциальной формы

$$M(x,y)dx+N(x,y)dy,$$

для которой (6) является условием полного дифференциала. Исходная задача на отыскание экстремалей теряет смысл.

3. Функция F зависит только от y': $F = F(y')$. Уравнение Эйлера принимает вид $F_{y'y'} \cdot y'' = 0$, а его решение $y(x)=C_1x+C_2$. Таким образом, в данном случае экстремалями функционала $I[y(x)]$ являются всевозможные прямые.

Пример. Найти функцию $y(x)\in C^1[0,1]$, на которой достигается экстремум функционала,

$$I[y]=\int_0^1 (y')^2 e^{\cos y'} dx,$$

удовлетворяющую граничным условиям $y(0)=0,\ y(1)=-4$.

Решение. Подынтегральная функция зависит только от y', $F(y') = (y')^2 e^{\cos y'}$. Поэтому семейство экстремалей представляет собой двухпараметрическое семейство прямых, $y(x)=C_1x+C_2$. Используя граничные условия, получаем $C_1=-4,\ C_2=0$. Таким образом, искомой экстремалью является функция $y(x)=-4x$.

4. Функция F не зависит от y, т.е. $F = F(x,y')$. Тогда уравнение (3) записывается в виде $\frac{d}{dx}F_{y'}(x,y') = 0$, откуда получаем первый интеграл уравнения Эйлера $F_{y'}(x,y') = C_1$, т.е. дифференциальное уравнение первого порядка, решив которое, найдем экстремали функционала.

Пример. Материальная точка перемещается из точки $A(1,0)$ в точку $B(2,1)$ со скоростью $v=x$. Найти кривую, по которой время движения будет минимальным.

Решение: Используя известное кинематическое выражение $v = \frac{ds}{dt}$, где $ds = \sqrt{(dx)^2 + (dy)^2}$ – длина элемента дуги траектории точки, получаем дифференциальное соотношение для t: $dt = \frac{ds}{x}$. Поэтому время, затраченное на прохождение дуги кривой $y=y(x)$ $(1\le x\le 2)$, определяется с помощью интеграла

$$t[y] = \int_1^2 \frac{\sqrt{1+y'^2}}{x} dx,$$

представляющего собой функционал, в котором рассматриваемые кривые $y(x)$ удовлетворяют условиям $y(1)=0,\ y(2)=1$. Подынтегральная функция $F(x,y,y') = \frac{\sqrt{1+y'^2}}{x}$ не зависит от y, поэтому из $F_{y'} = C_1 = const$ имеем равенство

$$\frac{y'}{x\sqrt{1+y'^2}} = \frac{1}{C_1}.$$

Определяя отсюда y':

$$y' = \pm\frac{x}{\sqrt{C_1^2 - x^2}}$$

и интегрируя, находим экстремали

$$y = C_2 \pm \sqrt{C_1^2 - x^2}\text{, или } (y - C_2)^2 + x^2 = C_1^2 .$$

Из граничных условий $y(1)=0$ и $y(2)=1$ для определения C_1 и C_2 получаем систему

$$C_2^2 + 1 = C_1^2, \quad (1 - C_2)^2 + 4 = C_1^2 .$$

Отсюда находим, что $C_1=\sqrt{5}$, $C_2=2$ и уравнение искомой экстремали есть окружность $x^2 + (y - 2)^2 = 5$ с центром в точке $(0,2)$ радиуса $\sqrt{5}$.

Из физических соображений ясно, что максимума для времени движения по различным кривым не существует и функция $y = 2 - \sqrt{- 5 - x^2}$ дает минимум функционалу $t[y]$.

5. Функция F не зависит явно от x, т.е. $F = F(y,y')$. Уравнение Эйлера принимает вид

$$F_y - F_{yy'} \cdot y' - F_{y'y'} \cdot y'' = 0 ,$$

или (после умножения обеих частей этого равенства на y')

$$\frac{d}{dx}(F - y'F_{y'}) = 0 ,$$

откуда получаем первый интеграл уравнения Эйлера

$$F - y'F_{y'} = C_1 .$$

Это дифференциальное уравнение первого порядка можно проинтегрировать, разрешив его относительно y' и разделив переменные, или путем введения параметра. [17, Еф-Сб-4, стр. 115]

Пример. Среди кривых, соединяющих две точки $M_0(a,y_0)$ и $M_1(b,y_1)$, найти ту, которая при вращении вокруг оси Ox образует поверхность наименьшей площади.

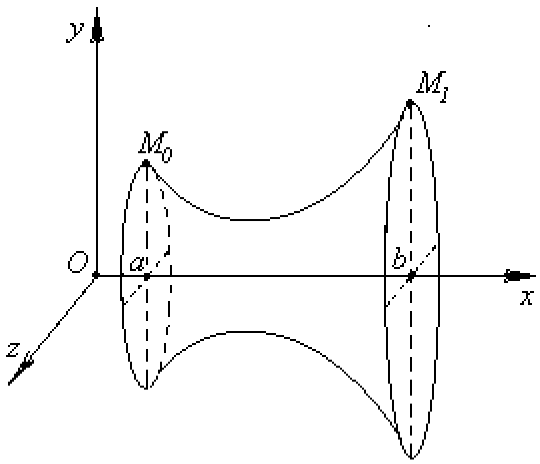

Рис. 2. Форма поверхности вращения наименьшей площади.

Решение: Площадь поверхности вращения вокруг оси Ox задается функционалом

$$I[y]=2\pi\int_a^b y\sqrt{1+y'^2}\,dx$$

причем допустимые кривые $y(x)$ удовлетворяют условию $y(a)=y_0$, $y(b)=y_1$. Подынтегральная функция не зависит от x, поэтому можем воспользоваться первым интегралом уравнения Эйлера $F-y'F_{y'}=C_1$, который в данном случае принимает вид

$$y\sqrt{1+y'^2}-\frac{yy'^2}{\sqrt{1+y'^2}}=C_1 \quad \text{, или } y=C_1\sqrt{1+y'^2}\ .$$

После элементарных преобразований получаем отсюда уравнение

$$C_1dy/\sqrt{y^2-C_1^2}=dx\,,$$

интегрируя которое, имеем

$$ln\left|\frac{y}{C_1}+\sqrt{\left(\frac{y}{C_1}\right)^2-1}\right|=\frac{x}{C_1}+C_2\,.$$

Разрешая полученное равенство относительно y, придем к уравнению цепной линии

$$y=C_1Ch(x/C_1+C_2).$$

Постоянные C_1 и C_2 находим из системы

$$y_0=C_1Ch(a/C_1+C_2),\ \ y_1=C_1Ch(b/C_1+C_2), \qquad (5)$$

которая может иметь одно, два или ни одного решения. Дальнейшие исследования показывают, что если система (5) не имеет решения, а также при достаточно малых отношениях $y_i/(b\text{-}a)$ $(i=0,1)$, множество значений площади фигур вращения имеют инфимум, равный

$$\pi(a^2+b^2),$$

который не достигается в пространстве функций $C^1[a,b]$. При достаточно больших отношениях $y_i/(b\text{-}a)$ $(i=0,1)$ и когда система (5) имеет два решения, на ближней к оси x кривой достигается локальный максимум, а на дальней кривой – абсолютный минимум.

В общем случае приходится привлекать известные методы решения из теории дифференциальных уравнений.

Пример 4. Исследовать на экстремум функционал

$$I[y]=\int_1^2(x^2y'^2+12y^2)dx,\quad y(1)=1,y(2)=8.$$

Решение: Уравнение Эйлера для данного функционала имеет вид

$$x^2y''+2xy'-12y=0.$$

Линейные уравнения такого типа в теории дифференциальных уравнений называются также уравнениями Эйлера. Его решение ищем в виде $y=x^\lambda$. Найдем производные $y'=\lambda x^{\lambda-1}, y''=\lambda(\lambda-1)x^{\lambda-2}$; подставив их в уравнение Эйлера, получим

$$x^{\lambda}(\lambda^2+\lambda-12)=0.$$

Для определения λ имеем характеристическое уравнение

$$\lambda^2+\lambda-12=0,$$

корни которого $\lambda_1=3$ и $\lambda_2=-4$. Общее решение уравнения Эйлера имеет вид

$$y=c_1x^3+c_2x^{-4}.$$

Из граничных условий $y(1)=1, y(2)=8$ для определения постоянных C_1 и C_2 получаем систему

$$C_1+C_2=1, \qquad 8C_1+C_2/16=8.$$

Отсюда находим $C_1=1$, $C_2=0$. Следовательно, $y=x^3$ есть экстремаль данного функционала. В этом примере экстремаль $y=x^3$ реализует минимум функционала.

2.2. **Функционалы, зависящие от производных более высокого порядка**

Одним из обобщений простейшей задачи вариационного исчисления является задача на экстремум функционала $J[y(\cdot)]$, зависящего от производных высших порядков функции $y(x)$:

$$J[y(\cdot)]=\int_a^b F(x,y,\ldots,y^{(n)})dx, \tag{7}$$

где функция $F(x,y,\ldots,y^{(n)})$ имеет непрерывные частные производные вплоть до $(n+1)$-го порядка по всем аргументам, а $y(x)\in C^n[a,b]$.

Граничные условия в этой задаче имеют вид

$$\begin{aligned} &y(a)=y_0, \quad y^{(1)}(a)=y_0^{(1)},\ldots,y^{(n)}(a)=y_0^{(n)},\\ &y(b)=y_1, \quad y^{(1)}(b)=y_1^{(1)},\ldots,y^{(n)}(b)=y_1^{(n)}. \end{aligned} \tag{8}$$

В рассматриваемой задаче теорема 1 обобщается следующей теоремой.

Теорема 2. Для того чтобы функционал (7) достигал на функции $y(x)\in C^n[a,b]$ локального экстремума, необходимо, чтобы эта функция удовлетворяла уравнению Эйлера-Пуассона

$$F_y-\frac{d}{dx}F_{y^{(1)}}+\frac{d^2}{dx^2}F_{y^{(2)}}-\ldots(-1)^n\frac{d^n}{dx^n}F_{y^{(n)}}=0\,. \tag{9}$$

Подобно случаю простейшей задачи вариационного исчисления, решения уравнения (9) (экстремали функционала (7)), удовлетворяющие граничным условиям (8), являются кривыми возможного абсолютного экстремума этого функционала на множестве

$$G=\left\{y(x)\in C^n[a,b]\;\middle|\; y(a)=y_0,\ldots,y^{(n)}(a)=y_0^{(n)},\ y(b)=y_1,\ldots,y^{(n)}(b)=y_1^{(n)}\right\}.$$

2.3. **Функционалы, зависящие от нескольких функций одной переменной**

Другим обобщением простейшей задачи вариационного исчисления является задача отыскания экстремума функционала, зависящего от нескольких функций. Здесь требуется найти совокупность функций $y_1(x),\ldots, y_n(x)\in C^1[a,b]$, удовлетворяющих граничным условиям

$$y_k(a)= y_{k0},\ y_k(b)= y_{k1}\ \ (k=1,\ldots,n)$$

и доставляющих экстремум функционалу

$$I[y_1(x),\ldots,y_n(x)] = \int_a^b F(x,y_1(x),\ldots,y_n(x),y_1'(x),\ldots,y_n'(x))dx\,, \qquad (10)$$

где функция $F(x,y_1(x),\ldots,y_n(x),y_1'(x),\ldots,y_n'(x))$ имеет непрерывные частные производные до второго порядка включительно по всем своим аргументам.

В этом случае необходимое условие экстремума функционала (10) приводит к следующей теореме.

Теорема 3. Для того чтобы набор функций $y_1(x),\ldots, y_n(x)\in C^1[a,b]$ доставлял слабый экстремум функционалу (10), необходимо, чтобы эти функции удовлетворяли системе уравнений Эйлера

$$F_{y_k} - \frac{d}{dx}F_{y_k'} = 0, \qquad k = 1,\ldots,n \qquad (12)$$

Уравнения (12) представляют собой систему n дифференциальных уравнений второго порядка. Общее решение системы в общем случае содержит $2n$ произвольных постоянных, которые однозначно определяются из граничных условий.

3.1. Простейшая задача с подвижными границами. Условие трансверсальности

В задачах вариационного исчисления с подвижными границами в отличие от ранее рассмотренных задач граничные условия на функцию $y(x),\ x\in [a,b]$ на концах отрезка $[a,b]$ не зафиксированы.

Простейшая задача вариационного исчисления с подвижными границами состоит в определении функции $y(x)\in C^1[a,b]$ и точек $x_0,\ x_1\in[a,b],\ x_0<x_1$, для которых функционал

$$J[y(x)] = \int_{x_0}^{x_1} F(x,y,y')dx \qquad (13)$$

достигает слабого экстремума при условиях

$$y(x_0)=\varphi_0(x_0),\ \ y(x_1)=\varphi_1(x_1). \qquad (14)$$

Здесь $\varphi_0(x),\varphi_1(x)\in C^1[a,b]$, $F(x,y,z)$ - заданные функции и $F(x,y,z)$ имеет непрерывные производные до второго порядка включительно по всем аргументам.

Замечание. Напомним, что слабым называется локальный экстремум в пространстве $C^1[a,b]$. В задаче с подвижными границами на кривой $y^*(x)$ с абсциссами концов x_0^* и x_1^* функционал (13) достигает локального экстремума в $C^1[a,b]$, если существует число $\varepsilon>0$ такое, что для всех кривых $y(x)\in$

$C^1[a,b]$ и точек x_0 и x_1, удовлетворяющих неравенствам $\|y^* - y\|_1 < \varepsilon$, $|x_0^* - x_0| < \varepsilon$, $|x_1^* - x_1| < \varepsilon$, справедливо $J[y^*(x)] \le J[y(x)]$ (локальный минимум) или $J[y^*(x)] \ge J[y(x)]$ (локальный максимум).

Задачу (13)-(14) можно сформулировать и следующим образом. Пусть на плоскости заданы гладкие кривые γ_0: $y=\varphi_0(x)$ и γ_1: $y=\varphi_1(x)$, $x\in[a,b]$. Требуется найти такую гладкую кривую $y=y(x)$, которая соединяет какую – либо точку кривой γ_0 с какой – либо точкой кривой γ_1 и доставляет слабый экстремум функционалу (13).

Приведем обобщение теоремы 1 для простейшей задачи вариационного исчисления с подвижными границами.

Теорема 4. Для того чтобы функционал (13) достигал на функции $y(x)\in C^1[a,b]$ слабого экстремума при условиях (14), необходимо, чтобы эта функция удовлетворяла уравнению Эйлера

$$F_y - \frac{d}{dx}F_{y'} = 0$$

и условию трансверсальности

$$\left[F + (\varphi_0' - y')F_{y'}\right]_{x=x_0} = 0, \quad \left[F + (\varphi_1' - y')F_{y'}\right]_{x=x_1} = 0 \quad . \tag{15}$$

Таким образом, для определения экстремалей в простейшей задаче с подвижными границами необходимо найти общее решение $y(x,C_1,C_2)$ уравнения Эйлера, после чего из условий (15) и уравнений

$$y(x_0,C_1,C_2)=\varphi_0(x_0), \quad y(x_1,C_1,C_2)=\varphi_1(x_1) \tag{16}$$

определить постоянные C_1 и C_2 и концы отрезка $[x_0,x_1]$.

Если на одном из концов кривой $y(x)$ задано обычное граничное условие ($y(a)=y_0$ или $y(b)=y_1$), то условие трансверсальности (15) следует записать только для другого конца кривой.

Частным случаем задачи с подвижными границами является задача, в которой задана абсцисса одного из концов кривой $y(x)$, например $x_2=b$, но граничное условие для $x=b$ отсутствует. Это означает, что граничная точка $(b,y(b))$ кривой $y(x)$ может перемещаться по вертикальной прямой $x=b$, и вместо второго условия трансверсальности (15) следует записать естественное граничное условие

$$\left[F_{y'}\right]_{x=b} = 0. \tag{17}$$

К задачам вариационного исчисления с подвижными границами относится и задача Больца, состоящая в определении функции $y(x)\in C^1[a,b]$, доставляющей слабый экстремум функционалу

$$J[y(x)] = \int_a^b F(x,y,y')dx + f(y(a),y(b)), \tag{19}$$

где $f(u,v)$ - заданная функция, имеющая непрерывные производные по u и v.

Необходимое условие экстремума функционала (19) формулируется следующим образом.

Теорема 5. Для того чтобы функционал (19) достигал на функции $y(x)\in C^1[a,b]$ слабого экстремума, необходимо, чтобы эта функция удовлетворяла уравнению Эйлера

$$F_y - \frac{d}{dx}F_{y'} = 0$$

и условию трансверсальности для задачи Больца

$$\left[F_{y'} - \frac{\partial f}{\partial y(a)}\right]_{x=a} = 0, \quad \left[F_{y'} - \frac{\partial f}{\partial y(b)}\right]_{x=b} = 0. \qquad (20)$$

Условия (20) используются для определения постоянных C_1 и C_2 из общего решения $y(x,C_1,C_2)$ уравнения Эйлера.

Пример 1. Найти экстремали функционала

$$I[y] = \int_0^{\pi}(y^2 - y'^2 - 2y\sin x)dx$$

при условии, что левый конец закреплен $(y(0)=0)$, а правый перемещается по прямой $x=\pi$.

Решение: На значение экстремали $y(x)$ в правом конце $x=\pi$ не накладывается никаких условий, поэтому для отыскания экстремали следует найти решение уравнения Эйлера $y'' + y = \sin x$ при естественном граничном условии $F_{y'}\big|_{x=\pi} = y'\big|_{x=\pi} = 0.$ Общее решение уравнение Эйлера записывается в виде

$$y = C_1\cos x + C_2\sin x - \frac{1}{2}x\cos x.$$

Тогда из условия $y(0)=0$ находим $C_1 = 0$, а из условия $y'(\pi) = 0$ получаем уравнение

$$y'\big|_{x=\pi} = \left(C_2\cos x - \frac{\cos x}{2} + \frac{x\sin x}{2}\right)\Bigg|_{x=\pi} = -C_2 + \frac{1}{2} = 0,$$

откуда $C_2=1/2$. Следовательно, экстремалью является кривая

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

Экстремали с угловыми точками.

Односторонние вариации.

Достаточные условия экстремума. Поле экстремалей.

Функция Вейерштрасса E(x,y,p,y').

Преобразование уравнений Эйлера к каноническому виду.

Вариационные задачи на условный экстремум. Связи вида $f(x,y_1,...,y_n) = 0$.

Вариационные задачи на условный экстремум. Связи вида $f(x,y_1,...,y_n,y_1',...,y_n') = 0$.

Изопериметрические задачи.

Прямые методы в вариационных задачах. Конечно-разностный метод Эйлера.

Метод Ритца. Метод Канторовича.

Зависимость функции и множества, на котором она максимизируется, от параметра. Оптимизация процессов, линейных относительно управления, без ограничений на управление.

Задачи с ограничениями на управление.

Необходимые условия оптимального управления в форме Лагранжа-Понтрягина.

Непрерывные процессы. Многошаговые процессы (без ограничений на управление, с ограничениями на управление).

Метод Гамильтона-Якоби-Беллмана (динамическое программирование). Непрерывные процессы. Многошаговые процессы.

5. МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО И ИТОГОВОГО КОНТРОЛЯ

Экзаменационные билеты

Билет 1

1. Вариация функционала и ее свойства.
2. Условие Якоби включения экстремали в поле экстремалей.
3. Исследовать на экстремум функционал:

$$v[y(x)] = \int_1^2 y'(1 + x^2 y')dx; \; y(1) = 3, \; y(2) = 5.$$

4. Найти экстремали изопериметрической задачи:

$$v[y(x)] = \int_0^\pi y \, Sin \, x \, dx; \; \int_0^\pi y'^2 dx = 3\pi / 2; \; y(0) = 0, \; y(\pi) = \pi .$$

Билет 2

1. Необходимое условие экстремума функционала.
2. Функция Вейерштрасса E(x,y,p,y'). Условие сильного и слабого экстремума для функционала.
3. Найти геодезические линии круглого цилиндра $r = R$.

Указание: решение удобно искать в цилиндрических координатах r, φ, z .

4. Найти экстремали функционала в задаче с подвижными границами:

$$v[y(x)] = \int_0^{x_1} y'^2 dx; \; y(0) = 0, \; y(x_1) = -x_1 - 1.$$

Задачи для самостоятельного решения

1.Найдите расстояние между функциями $y_1(x)=x^2$, $y_2(x)= x^3$ по норме пространства: а) $C^0[0,1]$; б) $C^1[0,1]$.

2.Найдите расстояние между функциями $y_1(x)=xe^{-x}$, $y_2(x)=0$ по норме пространства: $C^0[0,2]$; б) $C^1[0,2]$.

3.Найдите расстояние между функциями $y_1(x)=x$, $y_2(x)=lnx$ по норме пространства: $C^0[e^{-1},e]$; б) $C^1[e^{-1},e]$.

4.Покажите, что функционал $I[y]=\int_0^1 (y-y')dx$, определенный на $C^1[0,1]$ с нормой $\|\cdot\|_1$, является непрерывным на функции $y_0(x)=x^3$.

5.Покажите, что функционал $I[y]=\int_0^1 (y')^2 dx$, определенный на $C^1[0,1]$, разрывен на функции $y_0(x)\equiv 0$ в случае нормы $\|\cdot\|_0$, но непрерывен на этой функции в случае нормы $\|\cdot\|_1$.

6.Покажите, что функционал $I[y]=\int_0^1 x^3\sqrt{1+y^2}dx$, определенный на пространстве $C[0,1]$, непрерывен на функции $y_0(x)=x^2$ по норме $\|\cdot\|_0$.

7.Докажите, что любой линейный непрерывный функционал в нормированном пространстве является дифференцируемым. Запишите его дифференциал.

8.Докажите, что функционал $I[y]=\int_a^b y^2 dx$, определенный в $C^0[a,b]$, является всюду дифференцируемым. Запишите его дифференциал.

9.Проверьте, являются ли дифференцируемыми следующие функционалы: а) $I[y]=y(a)$ в $C^0[a,b]$; б) $I[y]=y(a)$ в $C^1[a,b]$; в) $I[y]=|y(a)|$ в $C^0[a,b]$; г) $I[y]=\sqrt{1+y'(a)}$, в $C^1[a,b]$.

10.Найдите первую вариацию функционала, определенного на нормированном пространстве непрерывно дифференцируемых функций: а) $I[y]=\int_0^1 x^2\sqrt{1+y^2}dx$; б) $I[y]=\int_{-1}^1 (y'e^y+xy^2)dx$; в) $I[y]=\int_0^\pi y'Sinydx$; г) $I[y]=y^2(0)+\int_0^1 (xy+(y')^2)dx$.

11.Найти приращение и вариацию следующих функционалов:

а) $I[y]=\int_{-1}^e (yy'+xy'^2)dx$, если $y=\ln x,\ \delta y=\dfrac{\alpha(x-1)}{e-1}$;

б) $I[y]=\int_0^\pi y'^2 Sinxdx$, если $y=Sinx,\ \delta y=\alpha Cosx$.

Найдите все экстремали функционала $I[y]$, удовлетворяющие заданным краевым условиям (№№12-20):

12.$I[y]=\int_0^{\pi/2} ((y')^2-y^2)dx,\ y(0)=0,\ y(\pi/2)=1$. Отв. $y(x)=Sinx$.

13.$I[y]=\int_0^1 ((y')^2+12xy)dx,\ y(0)=0,\ y(1)=1$. Отв. $y(x)=x^3$.

14.$I[y]=\int_\pi^{2\pi} (4(y')^2-7yy'-y^2)dx,\ y(\pi)=0,\ y(2\pi)=0$. Отв. $y(x)=0$

15.$I[y]=\int_0^{\pi/8}(16y^2+(y')^2+2y(Sin2x+16x))dx,\ y(0)=0,\ y(\pi/8)=-\pi/8$.

Отв. $y(x)=\dfrac{\sqrt{2}Sh4x}{40Sh(\pi/2)}-\dfrac{1}{20}Sin2x-x$.

16.$I[y]=\int_{\pi/4}^{\pi/2}(3x^2y^2+Cos\,y+y'(2x^3y-xSin\,y))dx,\quad y(\pi/4)=0,\ y(\pi/2)=1$

. Отв. $y(x)\in C^1[\pi/4,\pi/2]$

17.$I[y]=\int_2^4(x(y')^4-2y(y')^3)dx,\ y(2)=4,\ y(4)=5$. Отв. $y(x)=0.5x+3$

18.$I[y]=\int_0^1((y')^2-2x^6y'-2xy)dx,\ y(0)=0,\ y(1)=-1/6$. Отв. $y(x)=2\cdot x^7/7-x^3/3-5x/42$

19.$I[y]=\int_0^1 tg\,y'dx,\ y(0)=0,\ y(1)=2$. Отв. $y(x)=2x$

20.$I[y]=\int_0^1((y')^2+\dfrac{2xy}{1+x^2})dx,\ y(0)=0,\ y(1)=-1$. Отв. $y(x)=05xln(x^2+1)-x+arctgx-0.25(2ln2-\pi)$

21.Покажите, что функционал

$$I[y]=\int_a^b(p(x)y'+q(x)y+r(x))\,dx,$$

где $p(x)\in C^1[a,b]$, $q(x)$, $r(x)\in C[a,b]$, не имеет экстремумов.

22.Покажите, что для всякого дифференциального уравнения

$$y''=\varphi(x,y,y')$$

с дважды непрерывно дифференцируемой правой частью $\varphi(x,y,y')$ можно найти такую функцию $f(x,y,y')$, что решения этого уравнения будут экстремалями функционала $\int_a^b f(x,y,y')dx$.

Найдите все экстремали заданного функционала, удовлетворяющие заданным краевым условиям:

а) $I[y]=\int_0^1(120\,xy-(y'')^2)\,dx$, $y(0)=0,\ y(1)=1,\ y'(0)=0,\ y'(1)=6;$

г) $I[y]=\int_0^1((y'')^2+y^2-2yx^2)\,dx$, $y(0)=0,\ y(1)=1,\ y'(0)=0,\ y'(1)=2;$

23.Среди всех функций класса $C^2[0,\pi]$, удовлетворяющих граничным условиям $y(0)=y(\pi)=0,\ y'(0)=y'(\pi)=1$, найти такую, которая реализует экстремум функционала $I[y]=\int_0^{\pi}(16y^2-(y'')^2+x^2)\,dx$.

24.Найти экстремали заданных функционалов:

$$I[y]=\int_0^1 ((y'')^2+2(y')^2+y^2)\,dx, \quad y(0)=y(1)=0,\ y'(0)=1,\ y'(1)=-Sh1.$$

25. Найти все экстремали функционала $J[y]$,удовлетворяющие граничным условиям.

26. $J(y)=\int_0^1 y''^2dx;\quad y(0)=y(1)=y'(1)=0,\ y'(0)=1.$

27. $J(y)=\int_0^1 (48y-y''^2)dx;\quad y(0)=y'(0)=0,\ y(1)=1, y'(1)=4.$

28. $J(y)=\int_0^1 (y''^2-24xy)dx;\quad y(0)=y'(0)=0,\ y(1)=1/5, y'(1)=1.$

29.

$$J(y)=\int_0^{\pi/2} (y''^2-y'^2)dx;\quad y(0)=y'(0)=0,\ y(\pi/2)=\pi/2,\ y'(\pi/2)=0.$$

30. $J(y)=\int_0^b (y''^2+y'^2)dx;\quad y(0)=y'(0)=y(b)=y'(b)=0.$

31. $J(y)=\int_0^1 e^{-x}y''^2dx;\quad y(0)=0,\ y'(0)=1,\ y(1)=e,\ y'(1)=2e.$

32.

$$J(y)=\int_0^1 (x+1)^3y''^2dx;\quad y(0)=1,\ y'(0)=-1,\ y(1)=1/2,\ y'(1)=-1/4.$$

33.

$$J(y)=\int_0^{\pi/2} (y''^2-y^2+x^2)dx;\quad y(0)=1,\ y'(0)=y(\pi/2)=0,\ y'(\pi/2)=-1.$$

34.

$$J(y)=\int_0^1 y'''^2dx;\quad y(0)=y'(0)=y''(0)=0,\ y(1)=1, y'(1)=4, y''(1)=12.$$

35.

$$J(y)=\int_0^1 (y'''^2+y''^2)dx;\quad y(0)=y''(0)=0,\ y'(0)=1, y(1)=y''(1)=Sh1,$$

$y'(1)=Ch1.$

36. $J(y)=\int_0^\pi (y'''^2-y''^2)dx;\quad y(0)=y'(0)=y''(0)=0,\ y(\pi)=\pi, y'(\pi)=2,$
$y''(\pi)=0.$

37. $J(y)=\int_0^\pi (y'''^2-y'^2)dx;\quad y(0)=y'(0)=y''(0)=0, y(\pi)=y''(\pi)=Sh\pi,$
$y'(\pi)=Ch\pi+1.$

6. КАРТА КАДРОВОЙ ОБЕСПЕЧЕННОСТИ ДИСЦИПЛИНЫ

Лекторы: доцент, к.ф.-м.н. Сельвинский В.В.

Практические занятия: доцент, к.ф.-м.н. Сельвинский В.В.

ОГЛАВЛЕНИЕ