Федеральное агентство по образованию

АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ГОУВПО «АмГУ»

Факультет математики и информатики

УТВЕРЖДАЮ

Зав. кафедрой МАиМ

________Т.В. Труфанова

7 мая 2007г.

СПЕЦКУРС И СПЕЦСЕМИНАР ВЫПУКЛЫЙ АНАЛИЗ И МАТЕМАТИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

Учебно – методический

комплекс дисциплины

для специальности

010101 – математика

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

Выпуклый анализ и математическое программирование. Учебно – методический комплекс дисциплины для студентов АмГУ очной формы обучения специальности 010101 «Математика».– Благовещенск: Амурский гос. ун–т, 2007. − с.

Учебно – методический комплекс дисциплины "Выпуклый анализ и математическое программирование" 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.РАБОЧАЯ ПРОГРАММА

РАБОЧАЯ ПРОГРАММА

по дисциплине **"Выпуклый анализ и математическое программирование"**

для специальности 010101–"Математика"

Курс 3 Семестр 5, 6.

Лекции 72 (36+36) час. Экзамен 6 семестр.
Практические (семинарские) занятия 72 (36+36) час. Зачет 5 семестр.

Лабораторные занятия (нет).

Самостоятельная работа 56 час.

Всего 200 час.

Составитель В.В.Сельвинский, зам. зав. кафедрой, доцент.
Факультет математики и информатики.

Кафедра математического анализа и моделирования.

2006 г.

ЦЕЛИ И ЗАДАЧИ ДИСЦИПЛИНЫ

Целью дисциплины является изучение основ математического программирования, современных методов исследования экстремальных задач линейного и выпуклого программирования. Задачи такого типа возникают в экономических расчетах функционирования больших производственных систем.

При изучении курса математического программирования студенты должны научиться ориентироваться в классификации задач и методах их решения, получить практические навыки в конкретных расчетах.

В ходе изучения курса студенты выполняют ряд расчетных работ (не менее двух работ в течение каждого семестра). Кроме этого предусмотрено проведение итоговых контрольных работ в каждом семестре.

СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Наименование тем, их содержание, объем в лекционных часах

5 семестр

1. Введение - 2 часа

Предмет математического программирования. О моделях. Вопросы классификации и специфики. Примеры математических моделей. Основные обозначения.

2. Элементы выпуклого анализа - 8 часов

Евклидово пространство. Выпуклые множества. Проекция. Теоремы отделимости. Конус. Теорема Фаркаша. Выпуклые функции. Сильная выпуклость функций.

3. Основы математического программирования - 8 часов

Задачи математического программирования. Возможные направления. Экстремальные свойства. Экстремальные свойства на выпуклых множествах. Достаточные условия оптимальности. Функция Лагранжа.Условия оптимальности.

4. Теория линейного программирования - 8 часов

Основные понятия. Основные теоремы. Алгебраическая характеристика угловой точки. Двойственные задачи со смешанными ограничениями. Канонический вид задачи линейного программирования.

5. Конечные методы решения задач линейного программирования –
–10 часов

Симплексный метод. Рекуррентные соотношения алгоритма симплексного метода (связь между параметрами последовательных итераций). Методы отыскания исходной угловой точки. Вырожденность. Метод возмущений. Замечание о применении симплексного метода для решения специальных классов задач линейного программирования. О модифицированном симплексном методе. Мультипликативное представление обратной матрицы. Двойственный симплексный метод. Решение двойственной задачи как оценки влияния. О применении двойственного симплексного метода к задачам с возрастающим количеством условий. Метод одновременного решения прямой и двойственной задач. Метод декомпозиции.

6 семестр

6. Метод штрафных функций - 4 часа

Описание метода. Теоремы о сходимости.

7. Вопросы устойчивости в математическом программировании - 8 часов

Корректные и некорректные задачи. Один класс корректных задач. Задачи с точными ограничениями.Метод регуляризации. Сходимость. Метод регуляризации (общий случай). Сходимость метода регуляризации (общий случай).

8. Методы одномерной минимизации - 4 часа

О задачах одномерной минимизации. Поиск отрезка, содержащего точку минимума. Методы Фибоначчи и золотого сечения. Метод парабол. Метод кубической аппроксимации. Метод касательных.

9. Релаксационные методы решения экстремальных задач.Методы безусловной минимизации - 6 часа

Вопросы сходимости и устойчивости. Релаксационные процессы. Вспомогательный аппарат. Теоремы об оценках. Методы спуска. Методы первого и второго порядка. Метод сопряженных направлений. Методы нулевого порядка.

10. Релаксационные методы решения экстремальных задач с ограничениями - 8 часов

Характеристика методов. Метод проекции градиента. Метод условного градиента. Метод возможных направлений. Способы отыскания допустимой точки. Методы случайного спуска.

11.Метод модифицированных функций Лагранжа - 6 часов

Метод множителей Лагранжа. Метод модифицированных функций Лагранжа. Взаимосвязь методов множителей Лагранжа и штрафных функций.

Практические занятия, их содержание и объем в часах

5 семестр

1. Общая постановка задач линейного программирования - 2 часа
2. Графический метод решения основной задачи - 6 часов
3. Симплексный метод решения канонической задачи- 10 часов
4. Метод искусственного базиса - 4 часа
5. Целочисленное линейное программирование - 6 часов
6. Вопросы устойчивости в математическом программировании - 6 часов
7. Контрольная работа. - 2 часа

6 семестр

8. Задачи нелинейного программирования - 2 часа

9. Методы возможных направлений - 6 часов
10. Градиентные методы решения задач - 8 часов
11. Методы штрафных и барьерных функций - 8 часов
12. Дискретное динамическое программирование - 10 часов
13. Контрольная работа. - 2 часа

Темы для типового расчета расчетно-графических работ

1. Двойственные задачи со смешанными ограничениями.
2. Симплексный метод.
3. Методы Фибоначчи и золотого сечения.
4. Метод множителей Лагранжа.

Вопросы к зачету, 5 семестр

1. Предмет математического программирования.
2. Классификация задач математического программирования.
3. Постановка задачи о рационе.
4. Постановка транспортной задачи.
5. Выпуклые множества: определения, свойства.
6. Проекция точки на множество. Теорема отделимости.
7. Теорема о представлении.
8. Теорема Фаркаша.
9. Выпуклые функции и их свойства.
10. Неравенство Иенсена.
11. Экстремальные свойства выпуклых функций.
12. Сильная выпуклость функций.
13. Основная задача математического программирования.
14. Вожможности направления, активные ограничения.
15. Условия регулярности множества.

Вопросы к экзамену, 6 семестр

1. Теорема Куна –Таккера (дифференцируемый случай)
2. Седловая точка и условия ее существования.
3. Функция Лагранжа.
4. Теорема Куна – Таккера (общий случай).
5. Двойственные задачи линейного программирования.
6. Теоремы двойственности.
7. Теоремы о решениях двойственных задач.
8. Алгебраическая характеристика угловой точки.
9. Симплексный метод.
10. Метод искусственного базиса.
11. Выраженность задачи линейного программирования.
12. Двойственный симплексный метод.

13. Двойственные задачи со смешанными ограничениями.
14. метод сопряженных направлений.
15. Метод множителей Лагранжа.

Самостоятельная работа студентов (56 часов)

1. Знакомство с периодической литературой по выпуклому анализу и математическому программированию.
2. Подготовка докладов по отдельным учебным вопросам.
3. Выполнение индивидуальных заданий (по 2 задания в каждом семестре).

Критерии допуска и сдачи зачета

Для сдачи зачета необходимо представить индивидуальные задания (2 задания), уметь объяснить выполнение составляющих их упражнений, а также ответить на один из теоретических вопросов из общего списка вопросов к зачету.

Критерии сдачи экзамена

Для сдачи экзамена необходимо уметь выполнять типовые упражнения, а также достаточно хорошо ориентироваться в теоретическом материале программы данной дисциплины.

УЧЕБНО-МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ

Основная литература

1. Карманов В.Г. Математическое программирование. -М.: Наука, 2002.
2. Ашманов С.А. Линейное программирование. -М.: Наука, 1981.
3. Сборник задач по математике для втузов. Под ред. Ефимова А.В. -М.: Наука, ч.4, 1994.

Дополнительная литература

1. Акулич И.Л. Математическое программирование в примерах и задачах. -М.: Высш.шк., 1986.
2. Кузнецов Ю.Н. и др. Математическое программирование. -М.: Высш. шк., 1980.
3. Мухачева Э.А., Рубинштейн Г.Ш. Математическое программирование. Новосибирск, Наука, 1987.

2. ГРАФИК САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

5 семестр

1.Реферат по теме "Роль и место выпуклого анализа в современной математике» (2 нед.) 4 час

2. Индивидуальное задание (РГР) по теме "Выпуклые

множества и выпуклые функции" (9 нед.) 6 час
3. Индивидуальные домашние задания по каждой теме
практических занятий (еженед.) 30 ч.
4. Подготовка к зачету 16 ч.
Итого: 62 ч.

6 семестр

1. Индивидуальное задание (РГР) по теме "Метод возможных направлений в задачах нелинейного программирования " (9 нед.) 6 час
2. Индивидуальные домашние задания по каждой теме
практических занятий (еженед.) 30 ч.
3. Подготовка к экзамену 18 ч.
Итого: 60 ч.

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 называетсяадачей без условного оптимума.

Если $\varphi(\bar{x})$ линейная, $\varphi(x) = (\bar{c}, \bar{x}) = c_1 x_1 + \ldots + c_n x_n$ $X = \{\bar{x} : f_i(\bar{x}) \geq 0, i = \overline{1,m}\}$ $f_i(\bar{x})$ – линейная, то задачу относят к задачам линейного программирования.

Если $\varphi(\bar{x})$ интегральный функционал $\varphi(x) \sim [y(x)] = \int_{x_0}^{x_1} F(x, y(x), y`(x))dx$ 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

$$\sum_{j=1}^{n} x_{ij} \le a_i,$$
$$\sum_{i=1}^{m} x_{ij} = b_j,$$
$$x_{ij} \ge 0, \quad i = \overline{1,m}, \quad j = \overline{1,n}.$$

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 $(1,2,\ldots,j,\ldots,m)$ с заданным объемом потребления b_j в каждом пункте. Имеются n возможных пунктов производства $(1,2,\ldots,i,\ldots,n)$, 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

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

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 $f_i(x_i)$, как правило, монотонно возрастают и выпуклы вверх. Множество значений x_{ij}, удовлетворяющих ограничениям задачи, образует выпуклый многогранник, вершины которого являются точками локальных минимумов функции $L(x_{ij})$ (рис. 1). Отсюда и название подобных задач – многоэкстремальные.

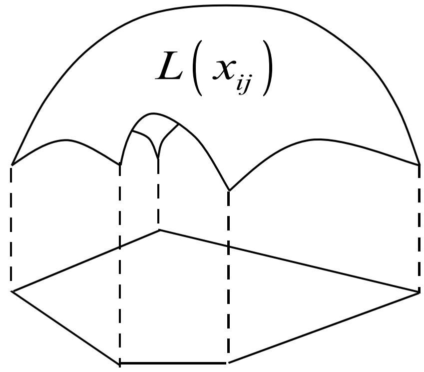

Рис. 1

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),

$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$ (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

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

Точка x называется *предельной точкой* последовательности.

Приведем еще несколько определений.

2.1.2. Множество

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

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 $z = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_m x_m, \alpha_i \ge 0 \ \ (i = \overline{1, m})$ 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 $P_X(v) \equiv p \in X$, что

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

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

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

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

то

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

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

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

откуда следует (2.2).

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

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

т.е. p является проекцией v на X. Δ

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

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

и

$$\|x-p\| \leq \|x-v\|$$

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

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

Действительно, 1) вытекает из того, что

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

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

$$\|P_X(y)-P_X(z)\|^2 \leq \langle P_X(z)-P_X(y), z-y \rangle \leq \|P_X(z)-P_X(y)\| \|z-y\|. \ \Delta$$

2.2.5 Гиперплоскостью в E_n называют множество вида

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

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

$$\left\{ x : \langle c, x \rangle \leq \lambda \right\}, \qquad \left\{ x : \langle c, x \rangle \geq \lambda \right\}.$$

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

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

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

$$\langle x, v - p \rangle \leq \langle p, v - p \rangle < \langle v, v - p \rangle, \qquad x \in X.$$

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).

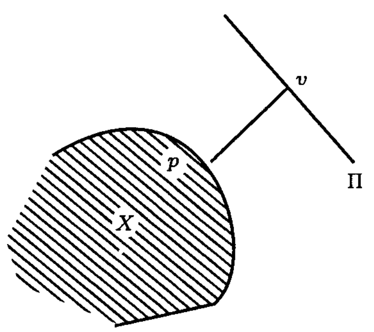

Рис. 2.1

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

$$\Pi = \{ x : \langle c, x \rangle = \lambda \},\ \lambda = \langle c, x^0 \rangle \text{ и для всех } x \in X \ \langle c, x \rangle \le \lambda .$$

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

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

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

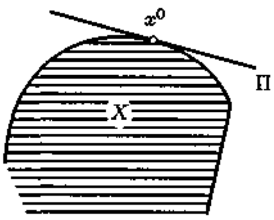

Рис.2

Итак, гиперплоскость

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

опорная. Δ

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, 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.

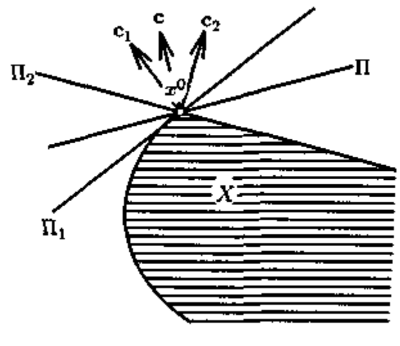

Рис. 3

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 Π, т.е. существует вектор $c \neq 0$ такой, что

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

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

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

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. Δ

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 $x_1, x_2, \ldots, x_N$ - угловые точки множества X_0 такие, что

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

Покажем, что $x_1, x_2, \ldots, x_N$ являются угловыми точками и для X. Предположим противное, т.е. что для некоторой точки x_i найдутся $x', x'' \in X,\ x' \neq x''$, 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

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

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

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

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

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

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

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 $\lim\limits_{n\to\infty} \alpha_n = \alpha$ и, следовательно, $y - \alpha a_k = \lim\limits_{n\to\infty}(y_n - \alpha_n a_k) = \lim\limits_{n\to\infty} \overline{y_n} \cong \overline{y_n} \in \overline{Y}$ 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 $\overline{y}_n \cong \lim\limits_{n\to\infty} \frac{1}{\alpha_n} \overline{y}_n = -a_k$. 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 $\langle v,x\rangle = \langle B^T u, x\rangle = \langle u, Bx\rangle \le 0$.

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

Рассмотрим конус

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

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

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

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

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

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 $K = \{x : Bx \le 0\}$.

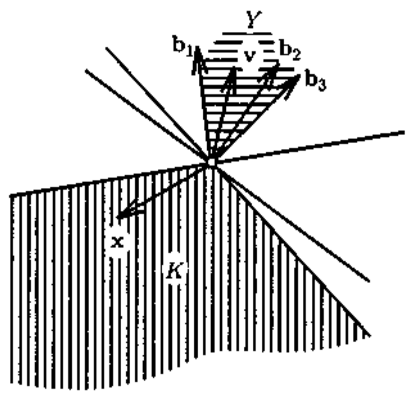

Рис. 4

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 $Y = \{y : y = B^T u, u \geq 0\}$ - горизонтальными линиями).

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:

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

либо имеет решение система:

$$v = B^T u \qquad u \geq 0. \tag{2.9}$$

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

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

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

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

$$\varphi(\alpha x + (1-\alpha) y) \leq \alpha \varphi(x) + (1-\alpha) \varphi(y). \tag{2.12}$$

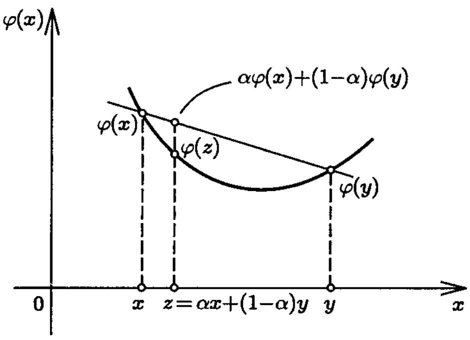

Рис. 2.5

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

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

$$\varphi(\alpha x + (1-\alpha) y) \geq \alpha \varphi(x) + (1-\alpha) \varphi(y) \tag{2.13}$$

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:

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

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

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

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

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

то

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

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 $x,y\in X$

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

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

$$\varphi^2(\alpha(1-\alpha)y) \le \alpha^2\varphi^2(x) + 2\alpha(1-\alpha)\varphi(x)\varphi(y) + (1-\alpha)^2\varphi^2(y) =$$
$$= \alpha\varphi^2(x) + (1-\alpha)\varphi^2(y) - \alpha(1-\alpha)(\varphi(x)-\varphi(y))^2 \le \alpha\varphi^2(x) + (1-\alpha)\varphi^2(y). \quad \Delta$$

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 $X^* \subset X$ 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 $\lambda \geq 0$ 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 $\lim\limits_{i\to\infty} s_{k_i} = s$. Для любого фиксированного $\lambda > 0$ найдется такой номер k_0, что для всех $k_i \geq k_0$ 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

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

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

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

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

Доказательство. Необходимость очевидна.

Достаточность. Пусть для некоторого α множество $X(\alpha)$ ограничено. Если $\beta \le \alpha$, 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 $\lambda \ge 0$. Поскольку множество $X(\alpha)$ ограничено, то найдется такое , что $x + \lambda s \notin X(\alpha)$ при всех $\lambda \ge \lambda_0$. Итак, для $z = x + \lambda_0 s$ будет $f(x) \le \alpha < f(z) \le \beta$, а для $v = v(\lambda) = x + \lambda \lambda_0 s$ при всех $\lambda > 0$ будет $f(v) \le \beta$. Так как $z = x + \lambda_0 s = \frac{1}{\lambda} v + (1 - \frac{1}{\lambda}) x$, то при $\lambda \ge 1$ из выпуклости функции $f(x)$ на выпуклом замкнутом множестве $X(\beta)$ следует неравенство $f(z) \le \frac{1}{\lambda} f(v) + (1 - \frac{1}{\lambda}) f(x)$, откуда $f(v) \ge \lambda (f(z) - f(x)) + f(x)$. Но $f(z) - f(x)$ - фиксированное положительное число, вследствие чего $f(v) \underset{\lambda \to +\infty}{\to} +\infty$, что противоречит условию $f(v) \le \beta$ 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 $\langle \varphi(x) - \varphi'(y), x - y \rangle \ge \rho\|x - y\|^2$;

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

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

Доказательство. а) Из определения сильной выпуклости следует

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

и в силу неравенства

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

соотношение а) становится справедливым.

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

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

Отсюда и из определения 2.5.2 получаем

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

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

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

Поэтому из б) получаем

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

т.е. неравенство в).

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

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

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

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

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 $\beta \in \left[0, \bar{\beta}\right]$.

Например, если $X = \{x : x \geq 0\}$, то в точке $x = 0$ любой вектор $-s \geq 0$, $s \neq 0$, 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 $\bar{s}$:

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

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

Очевидно, если x — внутренняя точка множества X, то любое направление $\bar{s}$ 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 $f_i(x) \leq 0$, для которых в этой точке выполняются равенства $f_i(x) = 0$. В связи с этим вводится следующее определение.

Ограничение $f_i(x) \leq 0$ 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* $\sigma \ge 0$*, 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. ∆

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 $\langle a_i, s\rangle \ge 0$. ∆

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*

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

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

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

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

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 $u_i \geq 0 (i = \overline{1,m})$, 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

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

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

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

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

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

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

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 $\langle f_i'(x), s\rangle \le 0, i \in I(x)$. 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*

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

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

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

имеет место неравенство

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\langle \varphi'(x), s\rangle \le 0. \tag{3.16}$$

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

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

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* $X = \{x : \langle a_i, x\rangle - b_i \le 0, i = \overline{1,m}\}$, *необходимо и достаточно, чтобы существовали такие числа* $u_i \ge 0 (i = \overline{1,m})$, *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*

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

выполнялось условие

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

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 $y \ge 0$. 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 $y \ge 0$, если

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

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

для всех $x \in \Gamma$ и $y \ge 0$.

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:

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

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

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

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

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

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

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 $f^T(x) = (f_1(x), f_2(x), \ldots, f_m(x))$. 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 д*ля 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* $x \in \Gamma, y \geq 0$, *то* x^* — *оптимальная точка задачи выпуклого программирования* (3.27).

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

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

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

$$\langle y, f(x^*) \rangle \geq \langle y^*, f(x^*) \rangle, \quad (3.30)$$

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

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

$$\langle y^*, f(x^*) \rangle \leq 0,$$

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

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

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

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

$$\varphi(x^*) \leq \varphi(x) - \left\langle y^*, f(x)\right\rangle \leq \varphi(x).$$

Но $x^* \in X$ (так как $x^* \in \Gamma$ и $f(x^*) \geq 0$), 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

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

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*

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

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

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

на множестве $x \ge 0, y \ge 0$.

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

Необходимость. Докажем эквивалентность условий теоремы 3.5.3:

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

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

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

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

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

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 $v^T = (v_1^*, v_2^*, \ldots, v_n^*)$ и $f^T(x) = (f_1(x), f_2(x), \ldots, f_m(x))$. 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 $x^* \in X$.Δ

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. Δ

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

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

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 $f(x) \geq 0, y \geq 0$, то

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

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

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

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)

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

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

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

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

будем называть *точкой Куна - Таккера.*

ТЕОРИЯ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

4.1. Основные понятия

4.1.1. Основная задача. Следующую экстремальную задачу будем называть *основной задачей линейного программирования:*

$$\min_{x \in R_1} \langle c, x \rangle, \qquad R_1 = \{ x : Ax \geq b, x \geq 0 \}. \tag{4.1}$$

Несколько позже будет показано, что любая задача линейного программирования может быть приведена к виду (4.1). Поэтому основные теоремы линейного программирования будут рассмотрены применительно к задаче (4.1), что никак не снижает общности рассмотрений.

4.1.2. Двойственность. Задачей, двойственной к основной задаче линейного программирования, будем называть задачу

$$\max_{y \in Q_1} \langle b, y \rangle,$$
$$Q_1 = \left\{ y A^T y \le c, y \ge 0 \right\}. \tag{4.2}$$

4.1.3. Эквивалентность. Две экстремальные задачи будем называть *эквивалентными,* если либо множества их решений совпадают, либо обе задачи не имеют решений.

Легко убедиться, что задача (4.1) будет двойственной к задаче (4.2). Для этого достаточно рассмотреть задачу, эквивалентную задаче (4.2):

$$\min_{y \in Q_1} \langle -b, y \rangle,$$
$$Q_1 = \left\{ y : -A^T y \ge -c, y \ge 0 \right\},$$

и построить, в соответствии с определением, к ней двойственную:

$$\max_{x \in R_1} \langle -c, x \rangle,$$
$$R_1 = \left\{ x : -Ax \le -b, x \ge 0 \right\}.$$

А это и есть задача (4.1).

Таким образом, задачи (4.1) и (4.2) взаимно двойственны.

4.1.4. Терминология. Всюду в настоящей книге будем придерживаться уже введенной терминологии (целевая функция, допусти мое множество, допустимая точка, оптимальная точка и т. д.).

Однако следует сказать, что терминология линейного программирования неоднозначна. В значительной степени это связано с экономическими приложениями, термины которых часто сохраняют и для задач линейного программирования.

Матрицу $A = [a_1, a_2, \ldots, a_n]$ называют *матрицей условий* задачи (4.1).

Ее столбцы $a_i (i = \overline{1, n})$ называют *векторами условий* задачи (4.1).

Вектор b называют *вектором ограничений* задачи (4.1).

Допустимую точку $x \in R_1$ называют также *планом.* Иногда угловую точку множества R_1 называют *опорным планом,* а решение задачи линейного программирования (оптимальную точку) называют *оптимальным планом.*

4.1.5. Геометрическая интерпретация. В пространстве E_n множество R_1 можно рассматривать как пересечение полупространств (при $n = 2$ — полуплоскостей):

$$(Ax)_i \ge b_i, \quad i = \overline{1,m},$$
$$x_j \ge 0, \quad j = \overline{1,n}.$$

Рассмотрим $\langle c,x \rangle = \lambda$

— семейство параллельных гиперплоскостей (при $n = 2$ — параллельных прямых). Вектор $-c$ направлен в сторону убывания целевой функции. На рис. 4.1 изображен случай, когда множество ограничено, т.е. многоугольник.

Рассмотрим некоторую точку $x_0 \in R_1$. Ей соответствует значение целевой функции

$$\lambda_0 = \langle c, x_0 \rangle.$$

Теперь будем перемещать прямую

$\langle c,x \rangle = \lambda$

в направлении $-c$, т. е. в направлении убывания величины λ, до тех пор, пока не придем в такую точку $x^* \in R_1$, для которой значение λ минимально. Геометрический смысл задачи здесь очевиден.

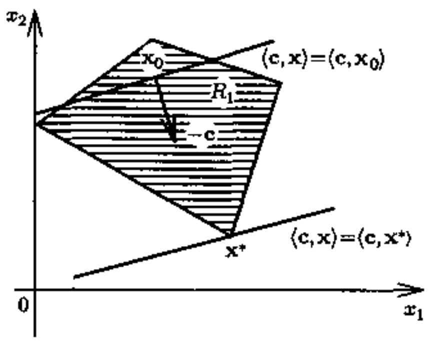

Рис 4

Легко изобразить на чертеже, когда R_1 неограниченно, но решение существует; когда решение существует, но не единственно; и, наконец, когда R_1 неограниченно и $\langle c,x \rangle$ неограниченна на R_1.

4.2. Основные теоремы

4.2.1 Сформулируем, во-первых, теорему, которая является частным случаем теоремы 3.6.9 (когда $\varphi(x)$ линейна) и поэтому в отдельном доказательстве не нуждается.

Будем обозначать функцию Лагранжа для задачи (4.1) так:

$$L_1(x,y) = \langle c,x\rangle + \langle y, b - Ax\rangle. \tag{4.3}$$

Теорема. *Для того чтобы точка* $x^* \in R$ *была оптимальной для основной задачи линейного программирования, необходимо и достаточно существования* $y^* \geq$ *такого, чтобы пара* x^*, y^* *была седловой точкой функции Лагранжа* $L_1(x^*,y) \leq L_1(x^*,y^*) \leq L_1(x,y^*)$ *в области* $x \geq 0, y \geq 0$, *т.е.*

$$L_1(x^*,y) \leq L_1(x^*,y^*) \leq L_1(x,y^*). \tag{4.4}$$

4.2.2. Из соотношения (3.33) следует, что необходимым и достаточным условием оптимальности x^* для задачи (4.1) является следующее представление вектора c:

$$-c = -\sum_{i \in I(x^*)} y_i^* a_i - \sum_{j \in J(x^*)} v_j^* e_j, \qquad y_i^* \geq 0, \qquad v_j^* \geq 0, \tag{4.5}$$

4.2.3. Теоремы двойственности.

Первая теорема двойственности. *Прямая и двойственная задачи либо обе имеют оптимальные точки* x^* *и* y^* *, причем*

$$\langle c, x^*\rangle = \langle b, y^*\rangle \tag{4-7}$$

либо обе их не имеют.

Доказательство. Рассмотрим функцию Лагранжа для задачи (4.2). Для этого запишем эту задачу в виде основной задачи линейного программирования, т. е. в виде (4.1).

Ясно, что задача

$$\max_{y \in Q_1} \langle b, y\rangle,$$
$$Q_1 = \{ y : A^T y \leq c, y \geq 0\},$$

эквивалентна задаче

$$\min_{y \in Q_1} \langle -b, y\rangle,$$
$$Q_1 = \{ y : -A^T y \geq -c, y \geq 0\}.$$

А для этой задачи, поскольку она записана в виде основной задачи, функцией Лагранжа будет

$$L_2(y,x) = -\langle -b, y\rangle + \langle x, -c + A^T y\rangle. \qquad (4.8)$$

Седловой точкой для $L_2(y,x)$ в области $y \ge 0$, $x \ge 0$ будет пара $\tilde{y}, \tilde{x}$ такая, что

$$L_2(\tilde{y}, x) \le L_2(\tilde{y}, \tilde{x}) \le L_2(y, \tilde{x}). \qquad (4.9)$$

Сравнивая (4.8) и (4.3), получим

$$L_2(y,x) = -[\langle c,x\rangle + \langle y, b - Ax\rangle] = -L_1(x,y) \qquad (4.10)$$

Из (4.10), (4.9) и (4.4) следует, что если x^*, y^* — седловая точка для $L_1(x,y)$, то y^*, x^* — седловая точка для $L_2(y,x)$, а значит, либо x^* и y^* оптимальны соответственно для задач (4.1) и (4.2), либо, когда седловая точка не существует, и задача (4.1), и задача (4.2) не имеют решений.

Наконец, равенство (4.7) следует из формул (4.6). ∆

Определения. Назовем условия в задачах (4.1) и (4.2)

$$(Ax)_j \ge b_j, \qquad y_j \ge 0, \qquad j = \overline{1,m},$$

парой двойственных условий. Аналогично условия

$$x_i \ge 0, \qquad (A^T y)_i \le c_i, \qquad i = \overline{1,n},$$

также являются парой двойственных условий.

Любое из условий называется *свободным,* если оно выполняется как строгое неравенство хотя бы для одного оптимального вектора.

Условие называется *закрепленным,* если оно выполняется как равенство для всех оптимальных векторов.

Вторая теорема двойственности. *Если прямая (а значит, и двойственная) задача разрешима, то в каждой паре двойственных условий одно является свободным, другое закрепленным.*

Доказательство. Утверждение, что свободному условию соответствует двойственное закрепленное, следует сразу из формул (4.6). Покажем, что закрепленному условию соответствует двойственное свободное. Пусть, например, условие $x_l \ge 0$ закрепленное: $x_l^* = 0$ для всех $x^* \in X^*$, где — X^* множество решений задачи (4.1). Построим решение $\overline{y}$ задачи (4.2), для которого $(A^T \overline{y})_l < c_l$. Воспользуемся индукцией по числу n.

При $n=1$ в прямой задаче

$$\min\langle c_1, x_1\rangle, \qquad c_1 \neq 0,$$
$$Ax_1 \geq b, \qquad x_1 \geq 0$$

(A — одностолбцовая матрица),

если $x_1^* = 0$, то $c_1 > 0$; в двойственной задаче

$$\max\langle b, y\rangle,$$
$$A^T y \leq c_1,$$
$$y \geq 0$$

$\bar{y} = 0$ является решением и $(A^T\bar{y})_1 \equiv A^T\bar{y} = 0 < c_1$.

Предположим, что для $n-1$ теорема верна. Построим $\bar{y}$ для размерности n. Рассмотрим вспомогательную задачу

$$\begin{aligned}&\min_{x\in R_1(\varepsilon)}\langle c, x\rangle,\\ &R_1(\varepsilon) = \{x : Ax \geq b, x \geq 0, x_l \geq \varepsilon > 0\}\end{aligned} \qquad \text{(A)}$$

1.Если $R_1(\varepsilon) = \varnothing$ при всех $\varepsilon > 0$, то $x_l = 0$ при всех $x \in R_1$, следовательно, задача (4.1) имеет размерность $n-1$, для которой теорема верна по предположению индукции.

2.Пусть $R_1(\varepsilon) \neq \varnothing$ при всех достаточно малых $\varepsilon > 0$. Поскольку условие $x_l \geq 0$ закрепленное, а $R_1(\varepsilon) \subset R_1 = \{x : Ax \geq b, x \geq 0\}$, то

$$\min_{x\in R_1(\varepsilon)}\langle c, x\rangle = \varphi_\varepsilon^* > \varphi^* = \min_{x\in R_1}\langle c, x\rangle$$

(если предположить, что $\varphi_\varepsilon^* = \varphi^*$, то найдется $x_l^* \geq \varepsilon > 0$, что противоречит условию закрепленности).

Построим задачу, двойственную к (A):

$$\begin{aligned}&\max_{y\in Q}\left(\langle b, y\rangle + \varepsilon\, y_{m+1}\right),\\ &Q = \{y : A^T y + e_l y_{m+1} \leq c, \quad y \geq 0, \quad y_{m+1} \geq 0\}.\end{aligned} \qquad \text{(B)}$$

Пусть точка $\begin{bmatrix} \bar{y}(\varepsilon) \\ \bar{y}_{m+1}(\varepsilon) \end{bmatrix}$ является решением задачи (B). Покажем, что точка $\bar{y}(\varepsilon)$ будет решением задачи (4.2). Не ограничивая общности, считаем, что активными будут ограничения

$$\left(A^T \bar{y}(\varepsilon)\right)_i = c_i, \qquad i = \overline{1,k}, \qquad i \neq l, \qquad k \leq n,$$

$$\left(A^T \bar{y}(\varepsilon)\right)_l + \bar{y}_{m+1}(\varepsilon) = c_l,$$

$$\bar{y}_j(\varepsilon) = 0, \qquad j = \overline{1,p}, \qquad j \neq m+1, \qquad p \leq m.$$

По теореме Куна - Таккера (3.5.1) в решении $\begin{bmatrix} \bar{y}(\varepsilon) \\ \bar{y}_{m+1}(\varepsilon) \end{bmatrix}$ будет

$$\begin{bmatrix} b \\ \varepsilon \end{bmatrix} = \sum_{\substack{i=\overline{1,k} \\ i \neq l}} \lambda_i \begin{bmatrix} a_i \\ 0 \end{bmatrix} + \lambda_l \begin{bmatrix} a_l \\ 1 \end{bmatrix} - \sum_{j=\overline{1,p}} u_j \begin{bmatrix} e_j \\ 0 \end{bmatrix},$$

$$\lambda_j \geq 0, \;\; i = \overline{1,k}, \;\; i \neq l, \;\; \lambda_l > 0, \;\; u_j \geq 0, \;\; j = \overline{1,p}.$$

При $\varepsilon \to +0$ будет $\lambda_l \to +0$. Поэтому по теореме о замкнутости конуса (2.3.2) существуют такие неотрицательные числа $\lambda_i^* \geq 0, \; \left(i = \overline{1,k}, i \neq l\right), \lambda_l^*, u_l^* \geq 0 \; \left(j = \overline{1,p}, j \neq m+1\right)$, что

$$\begin{bmatrix} b \\ 0 \end{bmatrix} = \sum_{\substack{i=\overline{1,k} \\ i \neq l}} \lambda_i^* \begin{bmatrix} a_i \\ 0 \end{bmatrix} + \lambda_l^* \begin{bmatrix} a_l \\ 1 \end{bmatrix} - \sum_{j=\overline{1,p}} u_j^* \begin{bmatrix} e_j \\ 0 \end{bmatrix}, \text{ (C)}$$

Причем $\lambda_l^* = 0$. В свою очередь в точках $\bar{y}(\varepsilon)$ будет $\left(A^T \bar{y}(\varepsilon)\right)_l + \bar{y}_{m+1}(\varepsilon) = c_1$, а поскольку. Поэтому для задачи (4.2) в точках $\bar{y}(\varepsilon)$ активными будут только ограничения

$$\left(A^T \bar{y}(\varepsilon)\right)_i = c_i, \qquad i = \overline{1,k}, \qquad i \neq l, \qquad \bar{y}_j(\varepsilon) = 0, \qquad j = \overline{1,p}$$

Переписав (C) в виде

$$b = \sum_{\substack{i=\overline{1,k} \\ i \neq l}} \lambda_i^* a_i - \sum_{j=\overline{1,p}} u_j^* e_j$$

и учитывая предыдущее замечание, мы сразу попадаем в условия теоремы (3.5.1), откуда следует, что $\overline{y}(\varepsilon)$ — решение задачи (4.2) при $\varepsilon > 0$ достаточно малых. △

4.3. Алгебраическая характеристика угловой точки

4.3.1. Ранее было введено геометрическое определение угловой точки. Однако для того чтобы уметь находить (вычислять) угловую точку, необходима ее алгебраическая характеристика. Пусть, как и прежде,

$$I(x) = \{i : (Ax)_i = b_i\}, \qquad J(x) = \{j : x_i = 0\},$$
$$J = \{j = 1, 2, \ldots, n\}.$$

Рассмотрим систему уравнений

$$(Az)_i = b_i, \qquad i \in I(x), \qquad z_j = 0, \qquad j \in J(x). \quad (4.17)$$

Не умаляя общности, можно полагать

$$I(x) = \{j : i = 1, 2, \ldots, r\},$$
$$J(x) = \{j : j = k+1, k+2, \ldots, n\},$$

и тогда при $r = k$ система (4.17) будет квадратной.

4.3.2. Теорема. *Для того чтобы точка* $x \neq 0$ *являлась угловой точкой множества* R_1, *необходимо и достаточно, чтобы вектор* x *удовлетворял неособенной квадратной системе уравнений* (4.17).

Доказательство. Достаточность. Пусть $x \in R_1$ и существует матрица

$$B = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \ldots & \ldots & \ldots & \ldots \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{bmatrix}, \qquad \det B \neq 0, \quad (4.18)$$

такая, что

$$B\overline{x} = \overline{b}. \quad (4.19)$$

Здесь

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_k \end{bmatrix}, \qquad \bar{b} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_k \end{bmatrix}, \qquad x_{k+1} = x_{k+2} = \dots = x_n = 0,$$

$$x = \begin{bmatrix} \bar{x} \\ 0 \\ \dots \\ 0 \end{bmatrix}. \quad (4.20)$$

Предположим, что x не угловая точка, т. е. существуют $x' \in R_1, x'' \in R_1, x' \neq x'' \neq x$, такие, что

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

Для $j > k$

$$x_j = \alpha x'_j + (1-\alpha)x''_j = 0.$$

И так как $\alpha > 0, (1-\alpha) > 0, x'_j \geq 0, x''_j \geq 0$, то

$$x'_j = x''_j = 0, \qquad j = \overline{k+1, n}.$$

Далее,

$$B\bar{x}' \geq \bar{b}, \qquad B\bar{x}'' \geq \bar{b}, \qquad \alpha B\bar{x}' + (1-\alpha)\bar{x}'' = \bar{b},$$

поэтому

$$B\bar{x}' = B\bar{x}'' = \bar{b}$$

и поскольку $\det B \neq 0$, то $\bar{x}' = \bar{x}''$; следовательно, $x' = x''$, что противоречит предположению.

Необходимость. Пусть x — угловая точка множества R_1. Покажем, что хотя бы для одного i будет

$$(Ax)_i = b_i.$$

Предположим, что такого номера i не существует. Так как $x \neq 0$, то найдется такое j, что $x_j > 0$. Рассмотрим

$$(x')^T = (x_1, x_2, \dots, x_{j-1}, x_j + \varepsilon, x_{j+1}, \dots, x_n) \geq 0$$

и

$$(x'')^T = (x_1, x_2, \ldots, x_{j-1}, x_i - \varepsilon, x_{j+1}, \ldots, x_n) \geq 0.$$

Из предположения, что $Ax > b$, для достаточно малого ε следует

$$Ax' \geq b, \qquad Ax'' \geq b,$$

т. е. $x' \in R_1$, $x'' \in R_1$. Но

$$x = \frac{1}{2}x' + \frac{1}{2}x'',$$

что противоречит предположению (точка x по предположению угловая).

2) Пусть $(Ax)_i = b_i$ для $i = \overline{1,r}$ и $x_j = 0$ для $j = \overline{k+1,n}$. Обозначим

$$\bar{a}_p = \begin{bmatrix} a_{1p} \\ a_{2p} \\ \ldots \\ a_{rp} \end{bmatrix}, \quad b' = \begin{bmatrix} b_1 \\ b_2 \\ \ldots \\ b_r \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} \bar{a}_1, \bar{a}_2, \ldots, \bar{a}_k \end{bmatrix}.$$

В этих обозначениях $\bar{A}\bar{x} = b'$ и $\bar{x} > 0$.

Докажем, что $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_k$ линейно независимы (в этом случае $k \leq r$).

Предположим противное, т. е. что существует $\bar{x}' \neq 0$ такой, что $\bar{A}\bar{x}' = 0$. Возьмем

$$x_1 = \begin{bmatrix} \bar{x} + \varepsilon\,\bar{x}' \\ 0 \\ \ldots \\ 0 \end{bmatrix}, \qquad x_2 = \begin{bmatrix} \bar{x} - \varepsilon\,\bar{x}' \\ 0 \\ \ldots \\ 0 \end{bmatrix}.$$

Легко видеть, что $x_1 \in R_1$, $x_2 \in R_1$ при малых ε. Но

$$x = \frac{1}{2}x_1 + \frac{1}{2}x_2,$$

что противоречит предположению. Итак, $k \leq r$.

Вычеркнув $r - k$ строк из $\bar{A}$, получим матрицу B, для которой выполняются (4.18)-(4.20).Δ

4.3.3. Следствие. *Число угловых точек множества* R_1 *конечно.*

Это утверждение очевидно, поскольку число неособенных квадратных клеток (подматриц) матрицы условий конечно. $\triangle$

4.3.4. Базис. Если точка $x^T = \left(\bar{x}^T, 0, \dots, 0\right), \bar{x} > 0$, угловая, то систему линейно независимых векторов $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_k$ в представлении

$$\bar{b} = \sum_{i=1}^{k} x_i \bar{a}_i, \qquad x_i > 0, \qquad i = \overline{1,k},$$

называют *базисом* угловой точки, а матрицу

$$B = \left[\bar{a}_1, \bar{a}_2, \dots, \bar{a}_k\right]$$

— *матрицей базиса* угловой точки.

4.4. Двойственные задачи со смешанными ограничениями

4.4.1. Задачи линейного программирования наиболее общего вида — это задачи со смешанными ограничениями, когда допустимое множество задается системой равенств и неравенств, а часть переменных (или все переменные) может быть свободна от ограничений.

Введем обозначения:

$$I = \{i : i = 1, 2, \dots, m\}, \qquad J = \{j : j = 1, 2, \dots, n\}.$$

Пусть I_1 означает часть совокупности индексов I, т. е. $I_1 \subset I$, а $I_2 = I \setminus I_1$. Аналогично, $J_1 \subset J, J_2 = J \setminus J_1$.

Пару задач

$$\begin{aligned} &\min\langle c, x\rangle, \\ &(Ax)_i \geq b_i, \ \ i \in I_1, \\ &(Ax)_i = b_i, \ \ i \in I_1, \\ &x_j \geq 0, \ \ j \in J_1, \end{aligned} \qquad (4.21)$$

И

$$
\begin{aligned}
&\max\langle b, y\rangle,\\
&(A^T y)_i \le c_j,\quad j \in J_1,\\
&(A^T y)_j = c_j,\quad j \in J_2,\\
&y_i \ge 0,\quad i \in I_1,
\end{aligned}
\qquad (4.22)
$$

называют *двойственным задачами со смешанными ограничениями.* Таким же образом, как и в случае задач (4.1) и (4.2), легко убедиться во взаимной двойственности задач (4.21) и (4.22).

4.4.2. Приведение к эквивалентным задачам основного вида. Запишем задачи (4.21) и (4.22) в следующей форме:

$$
\begin{aligned}
&\min\left[\langle c_1, x_1\rangle + \langle c_2, x_2\rangle\right],\\
&A_{11}x_1 + A_{12}x_2 \ge b_1,\\
&A_{21}x_1 + A_{22}x_2 = b_2,\\
&x_1 \ge 0,
\end{aligned}
\qquad (4.23)
$$

$$
\begin{aligned}
&\max\left[\langle b_1, y_1\rangle + \langle b_2, y_2\rangle\right],\\
&A_{11}^T y_1 + A_{21}^T y_2 \le c_1,\\
&A_{12}^T y_1 + A_{22}^T y_2 = c_2,\\
&y_1 \ge 0.
\end{aligned}
\qquad (4.24)
$$

Обозначения $x_1, x_2, c_1, c_2, b_1, b_2, A_{11}$ и т. д. естественны; например x_1 — вектор, составленный из тех компонент x_j вектора x, для которых $j \in J_1$; этому x_1 соответствует c_1, и т. д.

Понятны и обозначения клеток матрицы

$$
A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.
$$

Для приведения будем каждое равенство (например, $\varphi = 0$) заменять двумя неравенствами ($\varphi \ge 0$ и $-\varphi \ge 0$ или $\varphi \le 0$ и $-\varphi \le 0$), а вместо переменной u, на которую не наложены условия неотрицательности, будем вводить неотрицательные переменные $\bar{u}$ и $\bar{\bar{u}}$ следующим образом: $u = \bar{u} - \bar{\bar{u}}$, где

$\bar{u} = \max\{u, 0\} \ge 0$ и $\bar{\bar{u}} = \max\{-u, 0\} \ge 0$.

Итак, сделаем следующие замены переменных:

$$x_2 = \bar{x}_2 - \bar{\bar{x}}_2, \qquad y_2 = \bar{y}_2 - \bar{\bar{y}}_2,$$

$$\bar{x}_{j2} = \max\{x_{j2}, 0\}, \quad j \in J_2, \qquad \bar{y}_{i2} = \max\{y_{i2}, 0\}, \quad i \in I_2,$$

$$\bar{\bar{x}}_{j2} = \max\{-x_{j2}, 0\}, \quad j \in J_2, \qquad \bar{\bar{y}}_{i2} = \max\{-y_{i2}, 0\}, \quad i \in I_2.$$

Вводя эти переменные и заменяя равенства соответствующими системами неравенств, приведем задачи (4.23) и (4.24) к следующему виду:

$$\begin{aligned}
&\min\left[\langle c_1, x_1\rangle + \langle c_2, \bar{x}_2\rangle - \langle c_2, \bar{\bar{x}}_2\rangle\right],\\
&A_{11}x_1 + A_{12}\bar{x}_2 - A_{12}\bar{\bar{x}}_2 \ge b_1,\\
&A_{21}x_1 + A_{22}\bar{x}_2 - A_{22}\bar{\bar{x}}_2 \ge b_2,\\
&-A_{21}x_1 - A_{22}\bar{x}_2 + A_{22}\bar{\bar{x}}_2 \ge -b_2,\\
&x_1 \ge 0, \quad \bar{x}_2 \ge 0, \quad \bar{\bar{x}}_2 \ge 0,
\end{aligned} \tag{4.25}$$

$$\begin{aligned}
&\min\left[\langle b_1, y_1\rangle + \langle b_2, \bar{y}_2\rangle - \langle b_2, \bar{\bar{y}}_2\rangle\right],\\
&A_{11}^T y_1 + A_{12}^T \bar{y}_2 - A_{12}^T \bar{\bar{y}}_2 \ge c_1,\\
&A_{21}^T y_1 + A_{22}^T \bar{y}_2 - A_{22}^T \bar{\bar{y}}_2 \ge c_2,\\
&-A_{21}^T y_1 - A_{22}^T \bar{y}_2 + A_{22}^T \bar{\bar{y}}_2 \ge -c_2,\\
&y_1 \ge 0, \quad \bar{y}_2 \ge 0, \quad \bar{\bar{y}}_2 \ge 0,
\end{aligned} \tag{4.26}$$

Ясно, что (4.25) и (4.26) образуют двойственную пару задач линейного программирования основного вида.

4.4.3. Нетрудно убедиться, что задачи (4.23) и (4.24) эквивалентны задачам (4.25) и (4.26), а именно:

1) если

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ и } \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

оптимальны для (4.23) и (4.24), то

$$\begin{bmatrix} x_1 \\ \bar{x}_2 \\ \bar{x}_2 \end{bmatrix} \text{ и } \begin{bmatrix} y_1 \\ \bar{y}_2 \\ \bar{\bar{y}}_2 \end{bmatrix}$$

оптимальны для (4.25) и (4.26), и обратно;

2)если допустимое множество для (4.23) пусто, то пусто допустимое множество для (4.25), и обратно;

3)если целевая функция задачи (4.23) неограниченна снизу на допустимом множестве, то неограниченна снизу и целевая функция задачи (4.25), и обратно.

Утверждения 1)-3) обосновываются элементарными рассуждениями, которые (в случае необходимости) читатель может провести самостоятельно.

Из эквивалентности этих задач вытекает, что все теоремы, доказанные для задач (4.25) и (4.26), остаются справедливыми и для задач (4.23) и (4.24).

4. МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО И ИТОГОВОГО КОНТРОЛЯ

Вопросы к зачету по СК «Выпуклый анализ и математическое программирование» 5-ый семестр

1. Предмет математического программирования.
2. Классификация задач математического программирования.
3. Постановка задачи о рационе.
4. Постановка транспортной задачи.
5. Выпуклые множества: определение, свойства.
6. Проекция точки на множество. Теорема отделимости.
7. Теорема о представлении.
8. Теорема Фаркаша.
9. Выпуклые функции и их свойства.
10. Неравенство Иенсена.
11. Экстремальные свойства выпуклых функций.

12. Сильная выпуклость функций.
13. Основная задача математического программирования.
14. Возможные направления, активные ограничения.
15. Условия регулярности множества.
16. Теорема Куна-Таккера (дифференцируемый случай).
17. Седловая точка и ее существование.
18. Функция Лагранжа.
19. Теорема Куна-Таккера (общий случай).
20. Двойственные задачи линейного программирования.
21. Теоремы двойственности.
22. Теоремы о решениях двойственных задач.

Экзаменационные билеты
по спецкурсу «Выпуклый анализ и математическое программирование»
6-ой семестр

Билет 1

1. Предмет математического программирования.
2. Теорема Куна-Таккера (дифференцируемый случай).
3. Сумма средств $S = 300$ *тыс.руб.* выделяется предприятию в течение $N = 3$ лет. Прибыль, получаемая предприятием в результате выделения ему средств u в течение k–го года, составляет $J_k(u)$ $(k=1,2,3)$: $J_1(u)= J_2(u)=24 - 6\,(u-2)^2$ *(тыс.руб.)*, $J_3(u)=16 - 4\,(u-2)^2$ *(тыс.руб.)*. Распределить выделяемые средства по годам таким образом, чтобы суммарная прибыль предприятия за эти 3 года была максимальной.
4.

Билет 2

1. Классификация задач математического программирования.
2. Седловая точка и ее существование.
3. Решить задачу нелинейного программирования методом штрафных функций:

$$f(x)=x_1^2 + x_2^2 - 5x_1 - 10x_2 \to \min,$$
$$9x_1 + 8x_2 - 72 \le 0,$$
$$x_1 + 2x_2 - 10 \le 0.$$

Задачи для самостоятельного решения

Задачи оптимизации

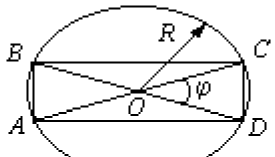

Рис.1

1.Найти стороны прямоугольника, вписанного в окружность радиуса R и имеющего наибольшую площадь S (рис.1).

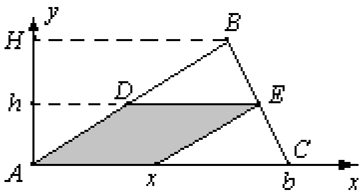

Рис.2

2.(Задача Евклида) В заданный треугольник ABC с высотой H и основанием b вписать параллелограмм наибольшей площади, стороны которого параллельны двум сторонам треугольника (рис.2).

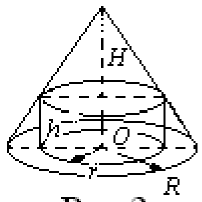

Рис.3

3.В прямой круговой конус вписан прямой круговой цилиндр так, что основания конуса и цилиндра лежат в одной плоскости (рис.3). Найти наибольшую возможную часть объема конуса, занятую таким цилиндром.

4.Луч света, переходя из одной однородной среды в другую, падает на границу раздела этих сред под углом α_1 к направлению нормали и после преломления на этой границе составляет с нормалью угол α_2 (рис.4). Требуется найти соотношение между этими углами.

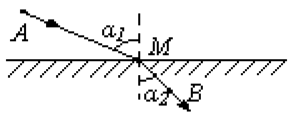

Рис.4

5.Как из прямоугольной листовой заготовки с соотношением сторон 1:2 вырезать круговой сектор, из которого можно было бы изготовить коническую воронку наибольшего объема?

Методы одномерной минимизации

6.Проверьте, являются ли унимодальными следующие функции:

а) $f(x)=x^2-2x-1$ на отрезках [0; 2], [1,5; 2];

б) $f(x)=\left| |x-1| -1\right|$ на отрезках [-3;3], [-3; 1], [1; 3], [0; 2]. (с. 93, 2.2)

7.Имеет ли функция $f(x)=\begin{cases} x^4 \, Sin\dfrac{1}{x}, & x \neq 0; \\ 0, & x=0, \end{cases}$ минимум в точке $x=0$, выполняется ли в этой точке необходимое, достаточное условия экстремума?

8.Для каких унимодальных функций метод золотого сечения приводит к цели за меньшее количество итераций, чем метод Ньютона?

9.Какой из методов: золотого сечения, Ньютона, кубической интерполяции - окажется более эффективным, если производные вычисляются приближенно через разность значений функции в близких точках?

10. Минимизируйте функции $f(x)=(x-1)^4$, $y(x)=(x-1)^2Sinx$ на отрезке [-2; 3] с помощью метода золотого сечения.

11. Минимизируйте функцию $f(x)=x\cdot arctgx-0,5\cdot ln(1+x^2)$ на отрезке [-6; 6] методом Ньютона. Выбирая различные начальные приближения, найдите какое-либо значение x_0, при котором метод начнет расходиться.

12. Минимизируйте функцию $f(x)=(x-1)^4$ на отрезке[0,5; 2] и функцию $g(x)= x\cdot Sin(1/x)$ на отрезке [0,2; 1] методами дихотомии и золотого сечения, а также с помощью оптимального последовательного поиска, градиентного метода и метода Ньютона. Сравните эти методы.

Индивидуальное задание 1
Минимизация выпуклых функций

1. Установите, являются ли выпуклыми множества:

а) $\Omega = \{(x_1,x_2)\in \boldsymbol{R}^2 : 2x_1 + x_2 \le 2, 2x_1 - x_2 \ge -2, x_2 \ge 0\}$;

б) $\Omega = \{(x_1,x_2)\in \boldsymbol{R}^2 : x_1 - x_2 \le 2,\ x_1^2 + x_2^2 \le 4\}$;

в) $\Omega = \{(x_1,x_2,x_3)\in \boldsymbol{R}^3 : x_3 \ge x_1^2 + x_2^2\}$;

г) $\Omega = \{(x_1,x_2,x_3)\in \boldsymbol{R}^3 : x_3 \le x_1^2 + x_2^2\}$.

2. Проверьте, какие из указанных функций являются выпуклыми в $\boldsymbol{R}^n$:

а) $f(x_1,x_2) = x_1^2 + 2x_1x_2 - 10x_1 + 5x_2$;

б) $f(x_1,x_2) = x_1^4 + x_2^4 + x_1^2 + x_2^2 + x_1^2x_2^2$;

в) $f(x_1,x_2) = x_1 \cdot e^{-(x_1+x_2)}$;

г) $f(x_1,x_2,x_3) = 2x_1^2 + x_1x_2 + x_2^2 + 2x_3^2 - 6x_1x_3$;

д) $f(x_1,x_2,x_3) = exp(x_1^2 + x_2^2 + x_3^2)$.

3. Докажите, что если $f_i(\boldsymbol{x}),\ i = \overline{1,m}$ – выпуклые (строго выпуклые) функции на выпуклом множестве $\Omega\subset \boldsymbol{R}^n$, то функция $g(\boldsymbol{x}) = \max\limits_{i=\overline{1,m}} f_i(\boldsymbol{x})$ выпукла (строго выпукла) на Ω.

4. Докажите, что если $g(\boldsymbol{x})$– выпуклая скалярная функция в $\boldsymbol{R}^n$, то функция $f(\boldsymbol{x})= -1/ g(\boldsymbol{x})$ является выпуклой на множестве $\boldsymbol{\Omega} = \{\boldsymbol{x}\in \boldsymbol{R}^n: g(\boldsymbol{x})>0\}$.

5.Покажите, что произведение выпуклых функций необязательно выпукло. Существуют ли подклассы выпуклых функций, Замкнутые по отношению к умножению?

6. Пусть $\Omega\subset \boldsymbol{R}^n$ – выпуклое множество. Докажите, что функция $f(\boldsymbol{x}) = \inf\limits_{y\in\Omega}\|\boldsymbol{x} - \boldsymbol{y}\|$ выпукла на $\boldsymbol{\Omega}$. Запишите функцию $f(\boldsymbol{x})$ в явном виде, если $\Omega = \{(x_1,x_2)\in \boldsymbol{R}^2 : x_1^2 + x_2^2 \le 1\}$.

7. Пусть $f(\boldsymbol{x})$ – выпуклая (строго выпуклая) функция в $\boldsymbol{R}^n$ и $f(\boldsymbol{x}+\kappa\boldsymbol{p})\ge \ge f(\boldsymbol{x})$ при $\boldsymbol{x}\in \boldsymbol{R}^n$ и $\kappa\in[0,\delta]$, где $\delta >0$. Покажите, что функция $\varphi_x(\kappa)=f(\boldsymbol{x}+\kappa\boldsymbol{p})$ является неубывающей (возрастающей) функцией переменного κ.

8. Покажите, что если $f(\boldsymbol{x})$ – выпуклая функция в $\boldsymbol{R}^n$, то функция $g(\boldsymbol{y}) = \inf\limits_{Ax=y} f(\boldsymbol{x})$, где A – матрица размера $m\times n$, выпукла в $\boldsymbol{R}^m$.

9. С помощью необходимых и достаточных условий экстремума выделите среди стационарных точек заданных функций те, которые являются

точками локального максимума или локального минимума: а) $f(x_1,x_2) = (x_1^2 - x_2)^2 + (x_1 - 1)^2$; б) $f(x_1,x_2) = (x_1 - x_2^2)^2 + (x_1 - 1)^2$. (АГЗ - с. 162, 3.10)

10. Найдите и классифицируйте стационарные точки следующих функций:

а) $f(x_1,x_2) = x_1^3 - x_1x_2 + x_2^2 - 2x_1 + 3x_2 - 4$;

б) $f(x_1,x_2) = 2x_1^3 + 4x_1x_2^3 - 10x_1x_2 + x_2^2$

в) $f(x_1,x_2) = 2x_1^3 + 4x_1x_2^3 - 10x_1x_2 + x_2^3$;

г) $f(x_1,x_2) = 2x_1^3 + 4x_1x_2^2 - 10x_1x_2 + x_2^2$;

д) $f(x_1,x_2) = 2x_1^2 + 4x_1x_2^3 - 10x_1x_2 + x_2^2$.

11. Запишите функции, двойственные позиномам $2x_1^2x_2 + x_2/x_1 + 3/(x_1x_2^2)$ и $x_1x_3/(2x_2) + 2x_2/x_1 + x_1/(4x_3) + x_3$. Укажите дополнительные условия, которым должны удовлетворять аргументы этих функций.

12. Решая двойственную задачу, найдите точки локального минимума и наименьшее значение следующих позиномов (коэффициенты a,b,c положительны): а) $ax_2/x_1 + bx_1^2 + c/(x_1x_2^2)$; б) $ax_2^2/x_1^2 + bx_1^2x_2 + c/(x_1^2x_2^4)$; в) $ax_2^2/x_1^2 + bx_1^{5/4}x_2^{1/4} + c/(x_1x_2^5)$.

Индивидуальное задание 2
Решение задач нелинейного программирования.

Аналитические методы

1. Минимизируйте функцию $f(x_1,x_2) = x_1 - x_2$ при ограничении $x_1^2 + x_2^2 = 1$. Найдите стационарные точки и точки минимума. Проанализируйте поведение функции Лагранжа в окрестностях найденных точек. Классифицируйте найденные точки (точки минимума, максимума, седловые точки для функции Лагранжа по переменным $\boldsymbol{x}$, λ).

2. Решите задачу $\begin{cases} (x_1+1)^2 + (x_2-3)^2 \to min; \\ x_1^2 + x_2^2 = 1 \end{cases}$ и проверьте решение графически.

3. Решите задачу $\begin{cases} (x_1+1)^2 + (x_2-3)^2 \to min; \\ x_1 + 2x_2 = 2 \end{cases}$ и проверьте решение графически.

4.Решите задачу $\begin{cases} (x_1+1)^2 + (x_2-3)^2 \to min; \\ x_1 + \alpha x_2 + \beta \le 0 \end{cases}$ и проверьте решение графически при значениях а) $\alpha = 1, \beta = 0$; б) $\alpha = -1, \beta = 1$; в) $\alpha = 0, \beta = 0$.

5. Путем перехода к задаче геометрического программирования минимизируйте функцию $f(x_1,x_2) = \dfrac{a}{x_1x_2} + b\sqrt{x_1 + x_2}$.

Численные методы

6. Используя метод условного градиента, найдите решение задачи квадратичного программирования с целевой функцией $f(x_1,x_2) = 5x_1^2 + 2x_1x_2 + 5x_2^2 - 4\sqrt{2}(5x_1 + x_2) - 8$ на множестве $0 \le x_1 \le 1$, $0 \le x_2 \le 2$, выбирая различные начальные точки: $(0,\ 0)$, $(0,\ 1)$, $(1,\ 2)$. В условии прекращения итераций $\left|x^k - x^{k-1}\right| < \varepsilon$ положить $\varepsilon = 10^{-2}$. Проверьте полученный результат, применив метод проекции точки на множество и выбрав начальную точку $(0,\ 0)$. Результаты расчетов проиллюстрируйте графически.

7. Решите задачу на условный экстремум

$$\begin{cases} 2x_1^2 + x_2^2 - 4x_1 - 2x_2 \to \min; \\ \sqrt{5}\,x_1 + x_2^2 = 5, \end{cases}$$

выбрав начальную точку $\boldsymbol{x}^0 = (0, -\sqrt{5})$, с помощью метода проекции точки на множество и метода проекции антиградиента. Используйте значение параметра точности $\varepsilon = 10^{-2}$. Дайте сравнительный анализ этих алгоритмов по числу шагов итерационного процесса, необходимых для достижения минимума целевой функции с заданной точностью. Покажите различие процессов минимизации этими методами в случае ограничений $\sqrt{5}\,x_1 + x_2^2 = 5$ и $\sqrt{5}\,x_1 + x_2^2 \le 5$. Дайте графическую иллюстрацию полученных результатов.

8. Решите задачу нелинейного программирования

$$\begin{cases} 4x_1^2 + x_2^2 + 2x_2 \to \min; \\ x_1^2 + x_2^2 - 4x_2 \le 0, \\ -x_1 + x_2 - 2 \le 0 \end{cases}$$

методом возможных направлений. Графически проиллюстрируйте полученные результаты. Используя метод проекции антиградиента и метод возможных направлений, решите задачу квадратичного программирования.

5. КАРТА КАДРОВОЙ ОБЕСПЕЧЕННОСТИ ДИСЦИПЛИНЫ

Лекторы: доцент, к.ф.-м.н. Сельвинский В.В

Практические занятия: доцент, к.ф.-м.н. Сельвинский В.В.,

ОГЛАВЛЕНИЕ