

Министерство образования Республики Беларусь
Министерство образования и науки Российской Федерации

ГУВПО «Белорусско-Российский университет»

Кафедра «Автоматизированные системы управления»

Лабораторный практикум

для студентов специальности 1-53 01 02
«Автоматизированные системы обработки информации»

по дисциплине

«Системный анализ и исследование операций»

6 семестр

Разработала:
канд. техн. наук, ст.преп. кафедры АСУ
Мрочек Т.В.

Могилев 2011

СОДЕРЖАНИЕ

Лабораторная работа № 1	
Классификация систем. Разработка функциональной модели системы	3
Лабораторная работа № 2	
Основные понятия и определения теории оптимизации	10
Лабораторная работа № 3	
Методы нулевого порядка безусловной одномерной минимизации.	
Метод золотого сечения.....	17
Лабораторная работа № 4	
Методы нулевого порядка безусловной многомерной минимизации.	
Метод сопряженных направлений (метод Пауэлла)	22
Лабораторная работа № 5	
Методы первого порядка безусловной многомерной минимизации	
(градиентные методы). Метод Гаусса-Зейделя	26
Лабораторная работа № 6	
Методы второго порядка безусловной многомерной минимизации.	
Метод Ньютона.....	33
Лабораторная работа № 7	
Численные методы поиска условного экстремума. Метод штрафов	38
Лабораторная работа № 8.	
Методология решения неструктурированных проблем.	
Метод экспертного ранжирования.	
Методы парных и последовательных сопоставлений.....	44
Лабораторная работа № 9.	
Методология решения неструктурированных проблем.	
Ранжирование проектов методом парных сравнений.....	50
Лабораторная работа № 10.	
Принципы решения слабо структурированных проблем.	
Кластерный анализ	55
Список рекомендуемой литературы	67

Лабораторная работа № 1

Классификация систем. Разработка функциональной модели системы

2 часа

Цель работы:

1. Изучение понятийного аппарата системного анализа, критериев классификации систем.
2. Ознакомиться с назначением и возможностями использования методологии IDEF0.
3. Ознакомиться с процессом создания функциональной модели.

1. Классификация систем

Классификацию систем можно осуществить по разным критериям. Её часто жестко невозможно проводить и она зависит от цели и ресурсов. Приведем основные способы классификации (возможны и другие критерии классификации систем).

1. **По отношению системы к окружающей среде:**
 - о *открытые* (есть обмен с окружающей средой ресурсами);
 - о *изолированные* (не обмениваются с окружающей средой ни материально-информационными ресурсами, ни энергией. Процессы самоорганизации в них невозможны);
 - о *закрытые* (нет обмена ресурсами с окружающей средой).
2. **По происхождению системы (элементов, связей, подсистем):**
 - о *искусственные* (орудия, механизмы, машины, автоматы, роботы и т.д.);
 - о *естественные* (живые, неживые, экологические, социальные и т.д.);
 - о *виртуальные* (воображаемые и, хотя они в действительности реально не существующие, но функционирующие так же, как и в случае, если бы они реально существовали);
 - о *смешанные* (экономические, биотехнические, организационные и т.д.).
3. **По типу описания закона (законов) функционирования системы:**
 - о типа “*Черный ящик*” (неизвестен полностью закон функционирования системы; известны только входные и выходные сообщения системы);
 - о *непараметризованные* (закон не описан, описываем с помощью хотя бы неизвестных параметров, известны лишь некоторые априорные свойства закона);
 - о *параметризованные* (закон известен с точностью до параметров и его возможно отнести к некоторому классу зависимостей);
 - о типа “*Белый (прозрачный) ящик*” (полностью известен закон).
4. **По способу управления системой (в системе):**
 - о *управляемые извне системы* (без обратной связи, регулируемые, управляемые структурно, информационно или функционально);
 - о *управляемые изнутри* (самоуправляемые или саморегулируемые – программно управляемые, регулируемые автоматически, адаптируемые – приспособляемые с помощью управляемых изменений состояний и самоорганизующиеся – изменяющие во времени и в пространстве свою структуру наиболее оптимально, упорядочивающие свою структуру под воздействием внутренних и внешних факторов);
 - о *с комбинированным управлением* (автоматические, полуавтоматические, автоматизированные, организационные).

Под **регулированием** понимается коррекция управляющих параметров по наблюдениям за траекторией поведения системы - с целью возвращения системы в нужное состояние (на нужную траекторию поведения системы; при этом под траекторией системы понимается последовательность принимаемых при функционировании системы состояний системы, которые рассматриваются как некоторые точки во множестве состояний системы).

5. По характеру поведения:

- о *детерминированные* (система, состояние которой в будущем однозначно определяется ее состоянием в настоящий момент времени и законами, описывающими

переходы элементов и системы из одних состояний в другие. Составные части в детерминированной системе взаимодействуют точно известным образом);

о *вероятностные (стохастические)* (системы, поведение которых описывается законами теории вероятностей. Для вероятностной системы знание текущего состояния и особенностей взаимной связи элементов недостаточно для предсказания будущего поведения системы со всей определенностью. Для такой системы имеется ряд направлений возможных переходов из одних состояний в другие, т. е. имеется группа сценариев преобразования состояний системы, и каждому сценарию поставлена в соответствие своя вероятность);

о *игровые* (система, осуществляющая разумный выбор своего поведения в будущем. В основе выбора лежат оценки ситуаций и предполагаемых способов действий, выбираемых на основе заранее сформированных критериев, а также с учетом соображений неформального характера. Руководствоваться этими соображениями может только человек);

6. По степени сложности:

о *простые* (характеризуются небольшим количеством возможных состояний, их поведение легко описывается в рамках той или иной математической модели);

о *сложные* (отличаются разнообразием внутренних связей, но допускают их описание. Причем набор методов, привлекаемых для описания сложных систем, как правило, многообразен, т. е. для построения математической модели сложной системы применяются различные подходы и разные разделы математики);

о *очень сложные* (характеризуются большой разветвленностью связей и своеобразностью отношений между элементами. Многообразие связей и отношений таково, что нет возможности все их выявить и проанализировать);

7. По степени организованности:

о *хорошо организованные* (система, у которой определены все элементы, их взаимосвязь, правила объединения в более крупные компоненты, связи между всеми компонентами и целями системы, ради достижения которых создается или функционирует система. При этом подразумевается, что все элементы системы с их взаимосвязями между собой, а также с целями системы можно отобразить в виде аналитических зависимостей. При формулировании задачи принятия решения для хорошо организованной системы проблемная ситуация описывается в виде математического выражения, критерия эффективности, критерия функционирования системы, который может быть представлен сложным уравнением, системой уравнений, сложными математическими моделями, включающими в себя и уравнения, и неравенства, и т. п. Важно, что решение задачи при представлении ее в виде хорошо организованной системы осуществляется аналитическими методами с использованием моделей формализованного представления системы);

о *плохо организованные* (при представлении объекта в виде **плохо организованной системы** не ставится задача определить все учитываемые компоненты, их свойства и связи между собой, а также с целями системы. Для плохо организованной системы формируется набор макропараметров и функциональных закономерностей, которые будут ее характеризовать. Определение этих параметров и восстановление функциональных зависимостей осуществляется на основании некоторой выборочной информации, характеризующей исследуемый объект или процесс. Далее полученные оценки характеристик распространяют на поведение системы в целом. При этом предполагается, что полученный результат обладает ограниченной достоверностью и его можно использовать с некоторыми оговорками. Так, например, если результат получен на основании статистических наблюдений за функционированием системы на ограниченном интервале времени, т. е. на основании выборочных наблюдений, то его можно использовать с некоторой доверительной вероятностью);

о *самоорганизующиеся* (это системы, обладающие свойством адаптации к изменению условий внешней среды, способные изменять структуру при взаимодействии системы со средой, сохраняя при этом свойства целостности, системы, способные

формировать возможные варианты поведения и выбирать из них наилучшие. Эти особенности обусловлены наличием в структуре системы активных элементов, которые, с одной стороны, обеспечивают возможность адаптации, приспособления системы к новым условиям существования, с другой стороны, вносят элемент неопределенности в поведение системы, чем затрудняют проведение анализа системы, построение ее модели, формальное ее описание и, в конечном счете, затрудняют управление такими системами);

8. По реакции на возмущающее воздействие:

- о *активные* (выдают ответную реакцию на возмущающее воздействие среды);
- о *пассивные* (не оказывают ответного воздействия на среду).

2 Разработка функциональной модели системы

2.1 Общие сведения о методологии IDEF0

При изучении вопроса о структурировании функций системы следует использовать инструментальный концептуальный анализ систем, например SADT-диаграммы, методологию IDEF0. Для структурирования информационных потоков применяются методология IDEF1, диаграммы «сущность-связь». Для структурирования процессов применяются методология IDEF3 и т. д.

Создание современных информационных систем представляет собой сложнейшую задачу, решение которой требует применения специальных методик и инструментов. Неудивительно, что в последнее время среди системных аналитиков и разработчиков значительно вырос интерес к CASE-технологиям и инструментальным CASE-средствам, позволяющим максимально систематизировать и автоматизировать все этапы разработки программного обеспечения: CASE-средство верхнего уровня BPwin, поддерживающее методологии IDEF0, Erwin – CASE-средство, позволяющее осуществить прямое и обратное проектирование базы данных, поддерживает методологию IDEF1X. CASE-средство BPwin значительно облегчают задачу создания информационной системы, позволяя осуществить декомпозицию сложной системы на более простые с тем, чтобы каждая из них могла проектироваться независимо, и для понимания любого уровня проектирования достаточно было оперировать с информацией о немногих ее частях.

Стандарт IDEF0 предназначен для функционального моделирования. Его применение – это сравнительно новое направление, но уже достаточно популярное и заслужившее серьезное отношение к себе. В основе стандарта лежит понятие функции, под которой понимается управляемое действие над входными данными, осуществляющееся посредством определенного механизма, результатом его являются выходные данные.

Стандарт IDEF0 базируется на трех основных принципах:

1. Принцип функциональной декомпозиции – любая функция может быть разбита на более простые функции;
2. Принцип ограничения сложности – количество блоков от 2 до 8 (в BPwin) условие удобочитаемости;
3. Принцип контекста – моделирование делового процесса начинается с построения контекстной диаграммы, на которой отображается только один блок – главная функция моделирующей системы.

Специализированным средством создания IDEF0 диаграмм является BPwin. Пакет BPWin предназначен для функционального моделирования и анализа деятельности предприятия. Модель в BPWin представляет собой совокупность SADT-диаграмм, каждая из которых описывает отдельный процесс в виде разбиения его на шаги и подпроцессы. С помощью соединяющих дуг описываются объекты, данные и ресурсы, необходимые для выполнения функций. Имеется возможность для любого процесса указать стоимость, время и частоту его выполнения. Эти характеристики в дальнейшем могут быть просуммированы с целью вычисления общей стоимости затрат – таким образом выявляются узкие места технологических цепочек, определяются затратные центры. BPWin может импорти-

ровать фрагменты информационной модели из ERWin (при этом сущности и атрибуты информационной модели ставятся в соответствие дугам SADT-диаграммы). Генерация отчетов по модели может осуществляться в формате MS Word и MS Excel.

Результатом применения методологии SADT является модель, которая состоит из диаграмм, фрагментов текстов и глоссария, имеющих ссылки друг на друга. Диаграммы – главные компоненты модели, все функции и интерфейсы на них представлены как блоки и дуги. Место соединения дуги с блоком определяет тип интерфейса. Диаграммы строятся при помощи блоков (см. рис.1.1). Каждый блок описывает какое-либо законченное действие. Четыре стороны блока имеют различное предназначение:

- Слева отображаются входные данные – исходные ресурсы для описываемой блоком функции (исходная информация, материалы).
- Справа показываются выходные ресурсы – результирующие ресурсы (продукция или информация), полученные в результате выполнения описываемой блоком функции;
- Сверху управление – то, что воздействует на процесс выполнения описываемой блоком функции и позволяет влиять на результат выполнения действия (средства управления, люди);
- Механизм изображается снизу – это то, посредством чего осуществляется данное действие (станки, приборы, люди и т.д.).



Рис. 1.1 – Построение диаграммы BPwin

2.2 Иерархия диаграмм

Построение SADT-модели начинается с представления всей системы в виде простейшей компоненты – одного блока и дуг, изображающих интерфейсы с функциями вне системы. Поскольку единственный блок представляет всю систему как единое целое, имя, указанное в блоке, является общим. Это верно и для интерфейсных дуг – они также представляют полный набор внешних интерфейсов системы в целом. Затем блок, который представляет систему в качестве единого модуля, детализируется на другой диаграмме с помощью нескольких блоков, соединенных интерфейсными дугами. Эти блоки представляют основные подфункции исходной функции. Данная декомпозиция выявляет полный набор подфункций, каждая из которых представлена как блок, границы которого определены интерфейсными дугами. Каждая из этих подфункций может быть декомпозирована подобным образом для более детального представления.

Во всех случаях каждая подфункция может содержать только те элементы, которые входят в исходную функцию. Кроме того, модель не может опустить какие-либо элементы, т.е., как уже отмечалось, родительский блок и его интерфейсы обеспечивают контекст. К нему нельзя ничего добавить, и из него не может быть ничего удалено.

Модель SADT представляет собой серию диаграмм с сопроводительной документацией, разбивающих сложный объект на составные части, которые представлены в виде блоков. Детали каждого из основных блоков показаны в виде блоков на других диаграммах. Каждая детальная диаграмма является декомпозицией блока из более общей диа-

граммы. На каждом шаге декомпозиции более общая диаграмма называется родительской для более детальной диаграммы.

Дуги, входящие в блок и выходящие из него на диаграмме верхнего уровня, являются точно теми же самыми, что и дуги, входящие в диаграмму нижнего уровня и выходящие из нее, потому что блок и диаграмма представляют одну и ту же часть системы. Каждый блок на диаграмме имеет свой номер. Блок любой диаграммы может быть далее описан диаграммой нижнего уровня, которая, в свою очередь, может быть далее детализирована с помощью необходимого числа диаграмм. Таким образом, формируется иерархия диаграмм.

Для того, чтобы указать положение любой диаграммы или блока в иерархии, используются номера диаграмм. Например, A21 является диаграммой, которая детализирует блок 1 на диаграмме A2. Аналогично, A2 детализирует блок 2 на диаграмме A0, которая является самой верхней диаграммой модели.

Допустим, главный процесс – это создать курсовой проект. На входе этого процесса – исходные данные по заданию. В качестве управляющего воздействия выступают методическое пособие, ГОСТы, необходимые требования.

Механизм осуществления создания курсового проекта – программное обеспечение, с помощью которого представлен материал и разработан проект, и исполнитель проекта (студент).



Рис. 1.2 – Блок «Создать курсовой проект»

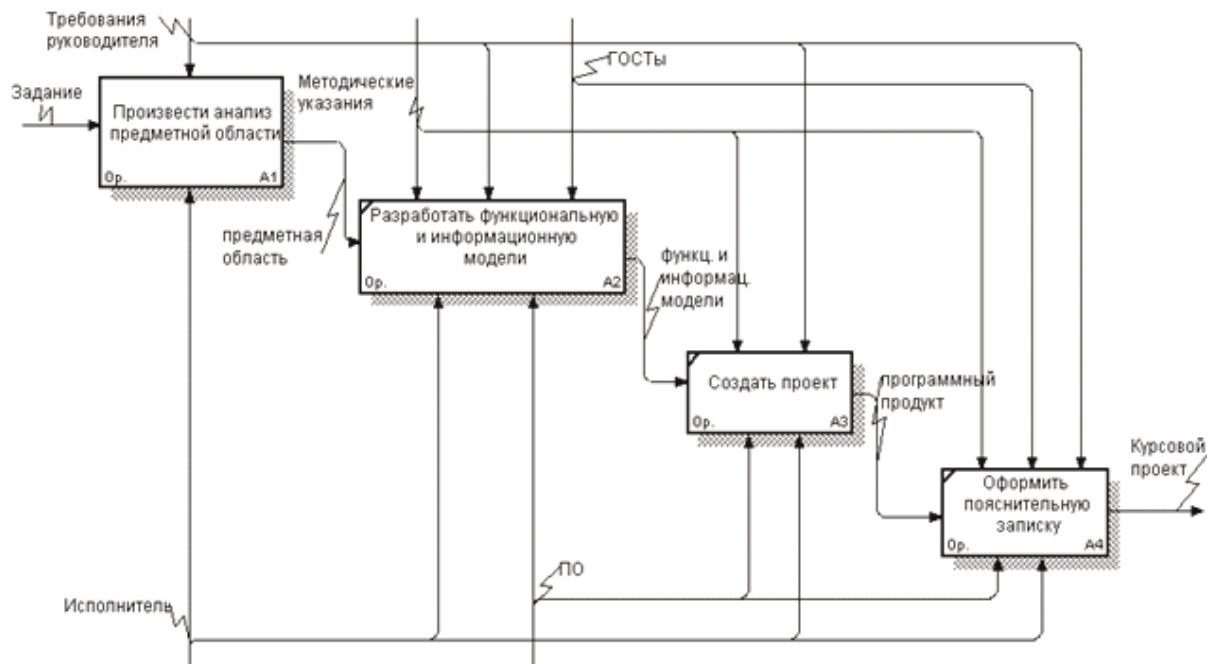


Рис. 1.2 – Блок «Создать курсовой проект»

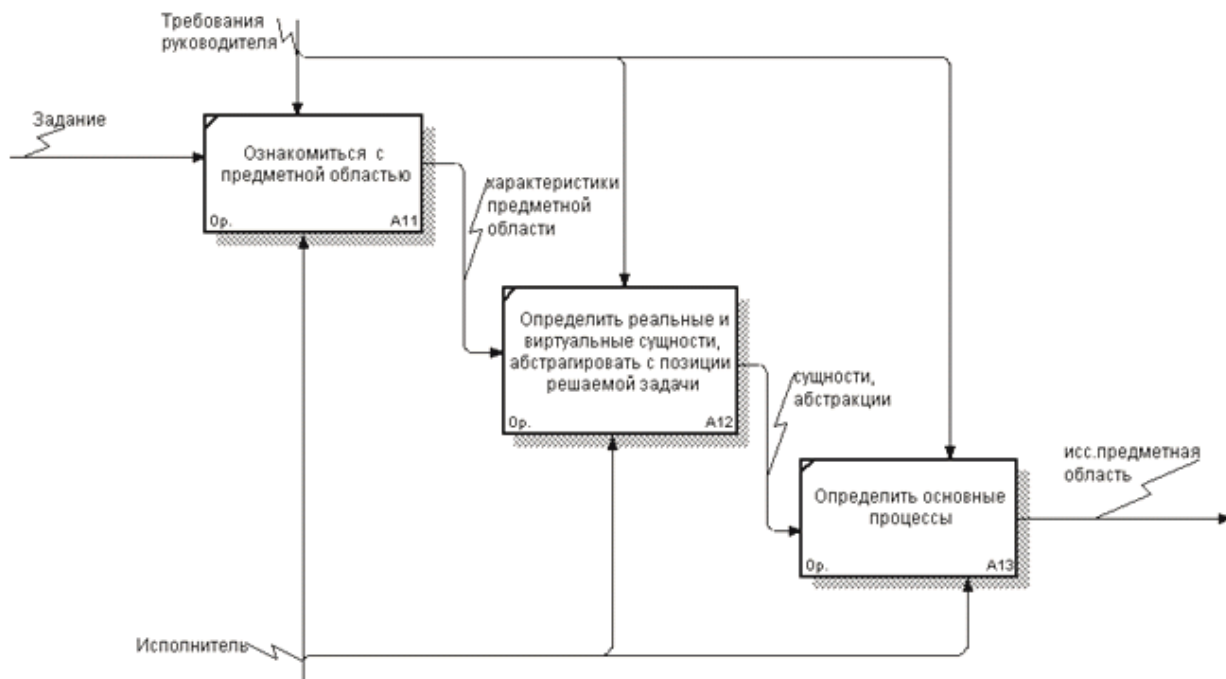


Рис. 1.3 – Декомпозиция блока «Произвести анализ предметной области»

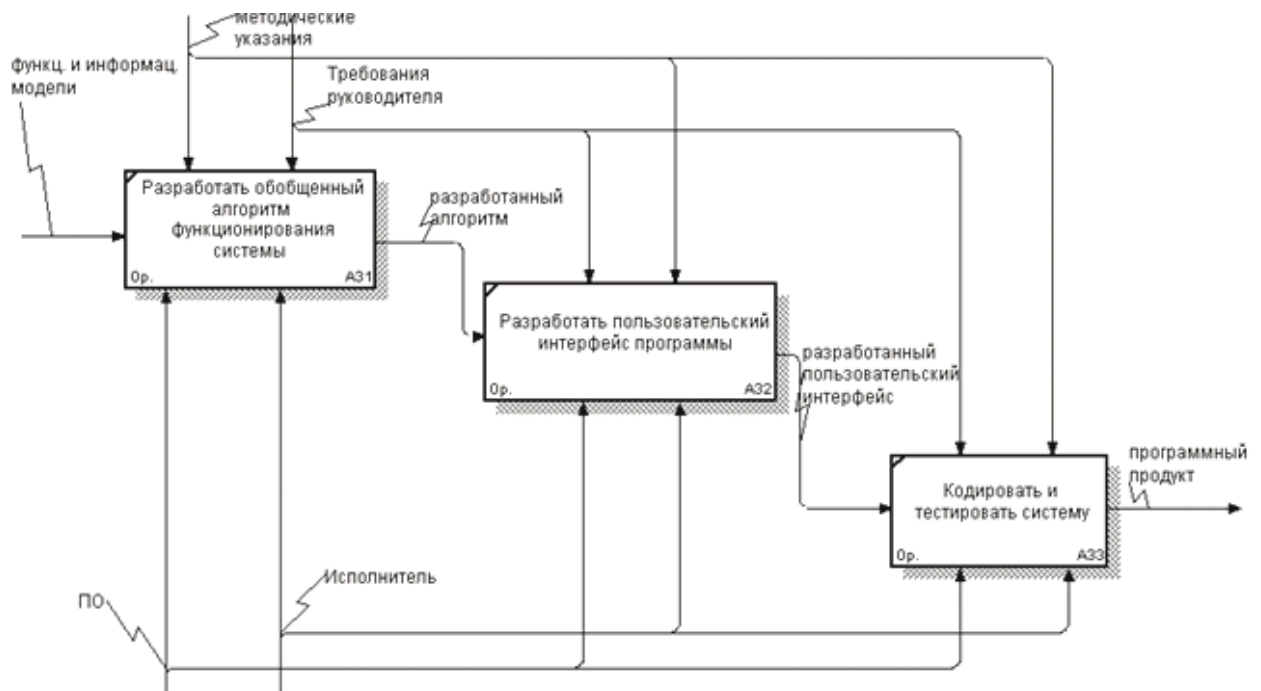


Рис. 1.4 – Декомпозиция блока «Создать проект»

Задание:

Необходимо:

- 1) классифицировать систему согласно варианту задания;
- 2) разработать функциональную модель компьютерной программы по учету особенностей работы анализируемой системы с помощью пакета BPwin. Модель должна содержать не менее трех уровней декомпозиции.

Варианты систем:

- | | |
|------------------------------|------------------------------|
| 1. Ювелирный магазин. | 9. Строительный кран. |
| 2. Грузовой автомобиль. | 10. Компьютер. |
| 3. Школа. | 11. Бумажная фабрика. |
| 4. Строительная фирма. | 12. Холодильник. |
| 5. Пружинные весы. | 13. Супермаркет. |
| 6. Молочный комбинат. | 14. Городской жилищный фонд. |
| 7. Электромеханические часы. | 15. Льнокомбинат. |
| 8. Механический динамометр. | |

Содержание отчета

- 1 Тема и цель работы.
- 2 Ответы на п.п. 1 и 2 задания.

Контрольные вопросы

- 1 Что такое цель, структура, система, подсистема, системность?
- 2 Перечислите основные критерии классификации систем.
- 3 Назначение и базовые принципы стандарта IDEF0.
- 4 Что представляет собой модель в BPWin? Перечислите основные компоненты диаграммы Bpwin.

Лабораторная работа № 2

Основные понятия и определения теории оптимизации

4 часа

Цель работы:

- 1) Изучить основные математические понятия, используемые при решении задач условной и безусловной оптимизации;
- 2) изучить типы данных и синтаксис основных операторов Java. Научиться преобразовывать типы данных;
- 3) изучить организацию консольного ввода и вывода, особенности использования функций класса Math;
- 4) изучить основы работы с массивами;
- 5) научиться создавать и выполнять простейшие программы на Java в среде NetBeans на примере решения хорошо структурированных задач системного анализа.

1 Основные математические понятия, используемые при решении задач условной и безусловной оптимизации

Матричное исчисление

Операции сложения и вычитания вводятся только для матриц одинаковых размеров.

Суммой (разностью) матриц $A_{m \times n}$ и $B_{m \times n}$ называется такая матрица $C_{m \times n} = A_{m \times n} \pm B_{m \times n}$, что $c_{ij} = a_{ij} \pm b_{ij}$, $i = \overline{1, m}$, $j = \overline{1, n}$.

Произведением AB матрицы $A_{m \times n}$ на матрицу $B_{n \times k}$ называется матрица $C_{m \times k}$, элементы которой c_{ij} находятся по формуле

$$c_{ij} = \sum_{s=1}^n a_{is} b_{sj}.$$

Из определения следует, что элемент c_{ij} равен сумме произведений элементов i -й строки матрицы A и соответствующих элементов j -го столбца матрицы B .

Пример: Даны матрицы

$$A = \begin{bmatrix} -1 & 2 & 3 & -5 \\ 1 & 0 & 4 & 3 \\ -2 & 3 & 0 & 0 \\ 1 & -1 & 3 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -3 \\ 1 & 4 \\ 5 & -2 \\ 3 & 4 \end{bmatrix}.$$

Выяснить, существуют ли произведения AB и BA , и, если они существуют, найти их.

Решение. Произведение AB существует, так как число столбцов матрицы A равно числу строк матрицы B . Тогда

$$\begin{aligned} C_{4 \times 2} = A_{4 \times 4} B_{4 \times 2} &= \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \\ c_{41} & c_{42} \end{bmatrix} = \\ &= \begin{bmatrix} -1 & 2 & 3 & -5 \\ 1 & 0 & 4 & 3 \\ -2 & 3 & 0 & 0 \\ 1 & -1 & 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ 1 & 4 \\ 5 & -2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 0 & -15 \\ 31 & 1 \\ -1 & 18 \\ 22 & -5 \end{bmatrix}. \end{aligned}$$

Здесь:

$$\begin{aligned}c_{11} &= (-1) \cdot 2 + 2 \cdot 1 + 3 \cdot 5 + (-5) \cdot 3 = 0, \\c_{12} &= (-1)(-3) + 2 \cdot 4 + 3(-2) + (-5) \cdot 4 = -15, \\c_{21} &= 1 \cdot 2 + 0 \cdot 1 + 4 \cdot 5 + 3 \cdot 3 = 31, \\c_{22} &= 1(-3) + 0 \cdot 4 + 4(-2) + 3 \cdot 4 = 1, \\c_{31} &= (-2) \cdot 2 + 3 \cdot 1 + 0 \cdot 5 + 0 \cdot 3 = -1, \\c_{32} &= (-2)(-3) + 3 \cdot 4 + 0(-2) + 0 \cdot 4 = 18, \\c_{41} &= 1 \cdot 2 + (-1) \cdot 1 + 3 \cdot 5 + 2 \cdot 3 = 22, \\c_{42} &= 1(-3) + (-1) \cdot 4 + 3(-2) + 2 \cdot 4 = -5.\end{aligned}$$

Произведение BA не существует, так как число столбцов матрицы B не равно числу строк матрицы A .

Определителем второго порядка $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$ называется

выражение (§ 12)

$$a_1 b_2 - a_2 b_1.$$

Определителем третьего порядка

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \quad (1)$$

называется выражение (§ 118)

$$a_1 b_2 c_3 - a_1 b_3 c_2 + b_1 c_2 a_3 - b_1 c_3 a_2 + c_1 a_2 b_3 - c_1 a_3 b_2, \quad (2)$$

или, что то же,

$$a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}. \quad (3)$$

Определителем четвертого порядка

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} \quad (1)$$

называется выражение

$$\Delta = a_1 A_1 + b_1 B_1 + c_1 C_1 + d_1 D_1, \quad (2)$$

где A_1, B_1, C_1, D_1 — алгебраические дополнения (§ 182) элементов a_1, b_1, c_1, d_1 , т. е.

$$\left. \begin{aligned} A_1 &= \begin{vmatrix} b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \\ b_4 & c_4 & d_4 \end{vmatrix}, & B_1 &= - \begin{vmatrix} a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \\ a_4 & c_4 & d_4 \end{vmatrix}, \\ C_1 &= \begin{vmatrix} a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \\ a_4 & b_4 & d_4 \end{vmatrix}, & D_1 &= - \begin{vmatrix} a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{vmatrix}. \end{aligned} \right\} \quad (3)$$

Минором какого-либо элемента называется определитель, получаемый из данного определителя вычеркиванием той строки и того столбца, на пересечении которых стоит элемент.

Пример. Минор элемента b_2 определителя есть

$$\begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix}, \text{ схема: } \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & \overline{b_2} & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Если в матрице размера $m \times n$ выбрать произвольно k строк и k столбцов, то элементы, принадлежащие выбранным строкам и столбцам, образуют квадратную матрицу порядка k , а определитель этой матрицы называется минором порядка k исходной матрицы.

Рангом матрицы A (обозначается одним из символов: r , r_A , $\text{rang } A$) называется наибольший порядок r отличных от нуля миноров матрицы A .

Если все миноры матрицы равны нулю, то ранг ее считают равным нулю.

Ранг матрицы выражается целым числом, заключенным между нулем и меньшим из чисел m, n , т.е. $0 \leq r \leq \min(m, n)$, где m — число столбцов, n — число строк матрицы.

Ранг матрицы равен нулю, если матрица является нулевой.

Для квадратной матрицы n -го порядка $r = n$ тогда и только тогда, когда матрица является невырожденной (т.е. $\det A \neq 0$).

Для нахождения ранга матрицы A применяют **метод окаймляющих миноров**. Пусть в матрице A найден минор M k -го порядка, отличный от нуля. Рассмотрим лишь те миноры $(k+1)$ -го порядка, которые содержат в себе (окаймляют) минор M . Если все они равны нулю, то ранг матрицы A равен k . В противном случае среди окаймляющих миноров найдется ненулевой минор $(k+1)$ -го порядка, и вся процедура повторяется.

Пример: Найти ранг матрицы A методом окаймляющих миноров, если

$$A = \begin{bmatrix} 2 & -1 & 3 & -2 & 4 \\ 4 & -2 & 5 & 1 & 7 \\ 2 & -1 & 1 & 8 & 2 \end{bmatrix}.$$

Решение. Вычисляем:

$$\begin{vmatrix} 2 & -1 \\ 4 & -2 \end{vmatrix} = -4 + 4 = 0, \quad \begin{vmatrix} -1 & 3 \\ -2 & 5 \end{vmatrix} = -5 + 6 = 1.$$

Найден минор второго порядка, отличный от нуля. Вычислим окаймляющие его миноры третьего порядка:

$$\begin{vmatrix} -1 & 3 & -2 \\ -2 & 5 & 1 \\ -1 & 1 & 8 \end{vmatrix} = -40 - 3 + 4 - 10 + 1 + 48 = 0,$$

$$\begin{vmatrix} -1 & 3 & 4 \\ -2 & 5 & 7 \\ -1 & 1 & 2 \end{vmatrix} = -10 - 21 - 8 + 20 + 12 + 7 = 0,$$

$$\begin{vmatrix} 2 & -1 & 3 \\ 4 & -2 & 5 \\ 2 & -1 & 1 \end{vmatrix} = -4 - 10 - 12 + 12 + 4 + 10 = 0.$$

Так как не существует миноров третьего порядка, отличных от нуля и окаймляющих минор второго порядка, отличный от нуля, то $r_A = 2$.

Основные табличные производные:

$$(x^\alpha)' = \alpha x^{\alpha-1}$$

(α — любое действительное число)

$$(x)' = 1$$

$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$$

$$(\sqrt{x})' = \frac{1}{2\sqrt{x}}$$

$$(\operatorname{tg} x)' = \frac{1}{\cos^2 x},$$

$$(\ln x)' = \frac{1}{x} \quad (x > 0),$$

$$(\operatorname{ctg} x)' = -\frac{1}{\sin^2 x},$$

$$(a^x)' = a^x \ln a,$$

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}} \quad (|x| < 1)$$

$$(\operatorname{arctg} x)' = \frac{1}{1+x^2},$$

$$(e^x)' = e^x,$$

$$(\arccos x)' = -\frac{1}{\sqrt{1-x^2}} \quad (|x| < 1),$$

$$(\operatorname{arccotg} x)' = -\frac{1}{1+x^2}.$$

$$(\sin x)' = \cos x,$$

$$(\cos x)' = -\sin x,$$

$$C' = 0$$

$$(\alpha u + \beta v)' = \alpha u' + \beta v' \quad (uv)' = u'v + uv'$$

$$(uvw)' = u'vw + uv'w + uvw' \quad \left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$$

(здесь C, α, β – постоянные, а u, v и w – функции от x , имеющие производные).

ПРИМЕР 1:

Найти y' , если $y = 2\sqrt[3]{x} + \frac{3}{x^2}$.

Перепишем $y = 2x^{\frac{1}{3}} + 3x^{-2}$.

$$y' = 2 \cdot \frac{1}{3} x^{-\frac{2}{3}} + 3 \cdot (-2) x^{-3}, \quad y' = \frac{2}{3\sqrt[3]{x^2}} - \frac{6}{x^3}.$$

ПРИМЕР 2:

Дано $y = x^3 \cos x$. Найти y' .

$$y' = x^3 (-\sin x) + 3x^2 \cos x, \quad \text{или} \quad y' = -x^3 \sin x + 3x^2 \cos x.$$

Градиентом функции многих переменных $f(X)$ называется вектор, составленный из первых частных производных функции по всем переменным:

$$\nabla f(X) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)^T$$

Градиент обозначается символом $\nabla f(X)$ (читается «набла $f(X)$ ») или $\text{grad } f(X)$.

Градиент – это вектор-столбец размерности $(n \cdot 1)$, где n – число переменных функции.

Свойства градиента:

- 1) градиент функции перпендикулярен касательной к линии уровня функции $f(X)$;
- 2) направление градиента есть направление наиболее быстрого роста функции.

Для нахождения градиента функции двух переменных $f(X)$ в заданной точке $X^0 = (x_1^0, x_2^0)$ необходимо:

- найти частные производные функции $f(X)$ и записать градиент функции

$$\nabla f(X) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right)^T$$

- вычислить значения частных производных функции в точке X^0 и составить полученный вектор градиента

$$\nabla f(X^0) = \left(\frac{\partial f}{\partial x_1} \Big|_{X=X^0}, \frac{\partial f}{\partial x_2} \Big|_{X=X^0} \right)^T$$

Матрицей Гессе называется квадратная матрица, составленная из вторых частных производных функции $f(X)$ по всем переменным, матрица имеет размерность $(n \cdot n)$, где n – число переменных функции:

$$H(X) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \dots & \frac{\partial^2 f}{\partial x_1 x_n} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} & \dots & \frac{\partial^2 f}{\partial x_2 x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_n x_1} & \frac{\partial^2 f}{\partial x_n x_2} & \dots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Пример:

Для функции $f(x) = x_1^2 + x_2^2$ необходимо:

а) вычислить и построить градиент в точках $x^0 = (0, 1)^T$, $x^1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$, $x^2 = (1, 0)^T$, $x^3 = (0, -1)^T$; б) найти матрицу Гессе.

Решение:

$$\nabla f(x) = (2x_1, 2x_2)^T, \quad \nabla f(x^0) = (0, 2)^T, \quad \nabla f(x^1) = (\sqrt{2}, \sqrt{2})^T, \quad \nabla f(x^2) = (2, 0)^T, \\ \nabla f(x^3) = (0, -2)^T, \quad H(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Пример:

Для функции $f(x) = x_1^2 + x_2^4$ вычислить градиент и найти матрицу Гессе в точках $x^0 = (0, 0)^T$, $x^1 = (1, 1)^T$.

Решение:

$$\nabla f(x) = (2x_1, 4x_2^3)^T, \quad H(x) = \begin{pmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{pmatrix}; \quad \nabla f(x^0) = (0, 0)^T, \quad H(x^0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}; \\ \nabla f(x^1) = (2, 4)^T, \quad H(x^1) = \begin{pmatrix} 2 & 0 \\ 0 & 12 \end{pmatrix}. \blacksquare$$

Собственные числа матрицы A – это корни характеристического уравнения вида: $\det(A - \lambda E) = 0$

Матрица A называется	если ее собственные числа
<i>положительно определенной</i> $A > 0$	положительны $\lambda_j > 0$
<i>отрицательно определенной</i> $A < 0$	отрицательны $\lambda_j < 0$
<i>положительно полуопределенной</i> $A \geq 0$	неотрицательны $\lambda_j \geq 0$
<i>отрицательно полуопределенной</i> $A \leq 0$	неположительны $\lambda_j \leq 0$
<i>знаконеопределенной</i> $A < > 0$	разного знака

Критерий Сильвестра (критерий знакоопределенности матрицы)

Матрица

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

является положительно определенной, если все ее диагональные миноры положительны;

отрицательно определенной, если они чередуют знак, начиная с минуса «-».

Здесь диагональные миноры:

$$\Delta_1 = a_{11} \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad \dots \quad \Delta_n = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

2 Основные понятия языка Java и основы работы в NetBeans изложены в приложении А, параграфы 1 – 11.

Задания

Часть 1. Матрицы

Необходимо найти ответы на все 6 поставленных вопросов:

- 1) Ввести с консоли $n = 3$ – размерность матрицы $A [n \times n]$. Задать значения элементов матрицы с помощью датчика случайных чисел;
- 2) Вычислить определитель матрицы A ;
- 3) Транспонировать матрицу и умножить на нее исходную матрицу A ;
- 4) Найти ранг матрицы A ;
- 5) Применить критерий Сильвестра и определить, является ли матрица A положительно определенной.
- 6) Найти максимальный элемент(ы) в матрице A и удалить из матрицы строку(и), в которой он (они) находятся.

Часть 2.

Для функции $f(x, y)$ вычислить градиент и найти матрицу Гессе в точке $x^0 = [0, 0]^T$.

Варианты заданий:

1. $f(x, y) = x^2 + y^3 - 2xy + x$
2. $f(x, y) = x^2 + y^{1/5} + 4x^2y$
3. $f(x, y) = x^2 + y^2 - 4x^2y^2 - y$
4. $f(x, y) = x^2 - 4xy + 4x^2 - 5y$
5. $f(x, y) = x^{2/3} + y^3 + 2xy + y^2$
6. $f(x, y) = x^{-2} - y^3 + 10xy - x$
7. $f(x, y) = x^4 + 2xy^3 - 2xy - 1/x$
8. $f(x, y) = x^{1/2} + y^4 - xy - 5x$
9. $f(x, y) = x^2 + y^{-3} - 2xy^{2/5}$
10. $f(x, y) = x^{-3} - 2y^3 - 6x^2y - 5$
11. $f(x, y) = x^2 + y^3 - 2xy$
12. $f(x, y) = x^{-3/2} + 2y^3 + 10xy^2$
13. $f(x, y) = 1/(x^2) + 10y^5 - 5x^2y^3$
14. $f(x, y) = x^{5/6} + y^2 - 20x^3y$
15. $f(x, y) = x^2 - y^{3/5} - 2x^2y^3 + 10$

Содержание отчета

- 1 Тема и цель работы.
- 2 Блок-схема и текст программы.
- 3 Результаты выполнения программы.

Контрольные вопросы

- 1 Что такое минор, ранг матрицы? Сущность метода окаймляющих миноров.
- 2 Что такое градиент, матрица Гессе? Что позволяет определить критерий Сильвестра?
- 3 Основные понятия языка Java и основы работы в NetBeans (приложение А, параграфы 1 – 11).

Лабораторная работа № 3

Методы нулевого порядка безусловной одномерной минимизации.

Метод золотого сечения

4 часа

Цель работы:

- 1) Изучить метод золотого сечения;
- 2) изучить создание классов, методов, конструкторов в Java.

Большинство известных методов одномерной минимизации применяется для класса унимодальных функций.

Методы одномерной минимизации широко применяются в методах первого и второго порядков для нахождения оптимальной величины шага. При этом левая граница начального интервала неопределенности, как правило, совпадает с началом координат, т.е. $a_0 = 0$.

Для эвристического выбора начального интервала неопределенности можно применить **алгоритм Свенна** [Swann W.H.]:

- 1) задать произвольно следующие параметры: x^0 – некоторую точку, $t > 0$ – величину шага. Положить $k = 0$;
- 2) вычислить значение функции в трех точках: $x^0 - t$; x^0 ; $x^0 + t$;
- 3) проверить условие окончания:
 - а) если $f(x^0 - t) \geq f(x^0) \leq f(x^0 + t)$, то начальный интервал неопределенности найден: $[a_0, b_0] = [x^0 - t, x^0 + t]$;
 - б) если $f(x^0 - t) \leq f(x^0) \geq f(x^0 + t)$, то функция не является унимодальной, а требуемый интервал неопределенности не может быть найден. Вычисления при этом прекращаются (рекомендуется задать другую начальную точку x^0);
 - в) если условие окончания не выполняется, то перейти к шагу 4;
- 4) определить величину Δ :
 - а) если $f(x^0 - t) \geq f(x^0) \geq f(x^0 + t)$, то $\Delta = t$; $a_0 = x^0$; $x^1 = x^0 + t$; $k = 1$;
 - б) если $f(x^0 - t) \leq f(x^0) \leq f(x^0 + t)$, то $\Delta = -t$; $b_0 = x^0$; $x^1 = x^0 - t$; $k = 1$;
- 5) найти следующую точку $x^{k+1} = x^k + 2^k \Delta$;
- 6) проверить условие убывания функции:
 - а) если $f(x^{k+1}) < f(x^k)$ и $\Delta = t$, то $a_0 = x^k$;
если $f(x^{k+1}) < f(x^k)$ и $\Delta = -t$, то $b_0 = x^k$;
в обоих случаях положить $k = k + 1$ и перейти к шагу 5;
 - б) если $f(x^{k+1}) \geq f(x^k)$, процедура завершается. При $\Delta = t$ положить $b_0 = x^{k+1}$, а при $\Delta = -t$ положить $a_0 = x^{k+1}$. В результате имеем $[a_0, b_0]$ – искомый начальный интервал неопределенности.

Метод золотого сечения

Постановка задачи: Требуется найти безусловный минимум функции $f(x)$ одной переменной, т.е. такую точку $x^* \in R$, что $f(x^*) = \min_{x \in R} f(x)$.

Для построения конкретного метода одномерной минимизации, работающего по принципу последовательного сокращения интервала неопределенности, следует задать правило выбора на каждом шаге двух внутренних точек. Конечно, желательно, чтобы одна из них всегда использовалась в качестве внутренней и для следующего интервала. Тогда число вычислений функции сократится вдвое и одна итерация потребует расчета только одного нового значения функции. В методе золотого сечения в качестве двух внутренних точек выбираются точки золотого сечения.

Точка производит "золотое сечение" отрезка, если отношение длины всего отрезка к большей части равно отношению большей части к меньшей.

На отрезке $[a_0, b_0]$ имеются две симметричные относительно его концов точки y_0 и z_0 :

$$\frac{b_0 - a_0}{b_0 - y_0} = \frac{b_0 - y_0}{y_0 - a_0} = \frac{b_0 - a_0}{z_0 - a_0} = \frac{z_0 - a_0}{b_0 - z_0} = \frac{1 + \sqrt{5}}{2} \approx 1,618.$$

Кроме того, точка y_0 производит золотое сечение отрезка $[a_0, z_0]$, а точка z_0 – отрезка $[y_0, b_0]$ (рис 3.1).

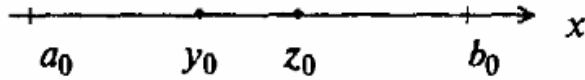


Рис. 3.1.

Стратегия поиска

Задается начальный интервал неопределенности $[a_0, b_0]$ и требуемая точность l . Алгоритм уменьшения интервала опирается на анализ значений функции в двух точках y и z . В качестве точек вычисления функции выбираются точки золотого сечения. Тогда с учетом свойств золотого сечения на каждой итерации, кроме первой, требуется только одно новое вычисление функции. Условия окончания процесса поиска стандартные: поиск заканчивается, когда длина Δ текущего интервала неопределенности оказывается меньше установленной величины l .

Алгоритм метода золотого сечения

Шаг 1. Задать начальный интервал неопределенности $L_0 = [a_0, b_0]$, точность $l > 0$.

Шаг 2. Положить $k = 0$.

Шаг 3. Вычислить

$$y_0 = a_0 + \frac{3 - \sqrt{5}}{2}(b_0 - a_0); \quad z_0 = a_0 + b_0 - y_0, \quad \frac{3 - \sqrt{5}}{2} = 0,38196.$$

Шаг 4. Вычислить $f(y_k), f(z_k)$.

Шаг 5. Сравнить $f(y_k)$ и $f(z_k)$:

а) если $f(y_k) \leq f(z_k)$, то положить $a_{k+1} = a_k$, $b_{k+1} = z_k$ и $y_{k+1} = a_{k+1} + b_{k+1} - y_k$, $z_{k+1} = y_k$ (рис. 3.2, а). Перейти к шагу 6;

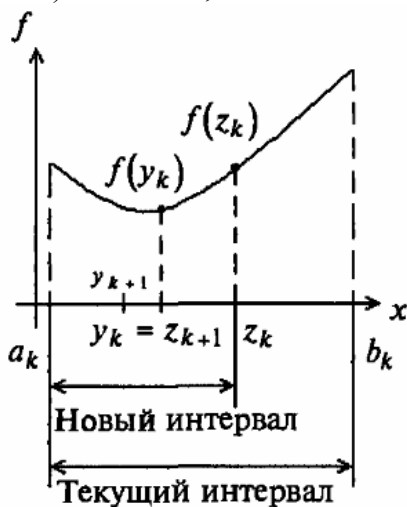
б) если $f(y_k) > f(z_k)$, то положить $a_{k+1} = y_k$, $b_{k+1} = b_k$ и $y_{k+1} = z_k$, $z_{k+1} = a_{k+1} + b_{k+1} - z_k$ (рис. 3.2, б).

Шаг 6. Вычислить $\Delta = |a_{k+1} - b_{k+1}|$ и проверить условие окончания:

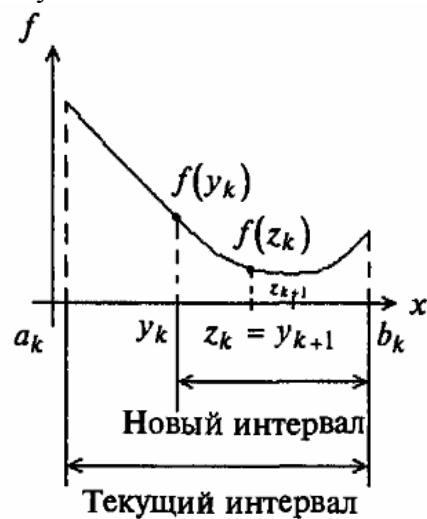
а) если $\Delta \leq l$, процесс поиска завершается и $x^* \in [a_{k+1}, b_{k+1}]$. В качестве приближенного решения можно взять середину последнего интервала:

$$x^* = \frac{a_{k+1} + b_{k+1}}{2};$$

б) если $\Delta > l$, положить $k = k + 1$ и перейти к шагу 4.



а



б

Рис. 3.2 – Графическая иллюстрация алгоритма метода золотого сечения

Сходимость метода

Для метода золотого сечения характеристика относительного уменьшения начального интервала неопределенности находится по формуле $R(N) = (0,618)^{N-1}$, где N – количество вычислений функции.

Пример. Найти минимум функции $f(x) = 2x^2 - 12x$ методом золотого сечения.

1. Найдем начальный интервал неопределенности методом Свенна:

а) зададим начальную точку $x^0 = 5$, шаг $t = 5$. Положим $k = 0$;

б) вычислим значение функции в трех точках: $x^0 - t = 0$; $x^0 = 5$; $x^0 + t = 10$:

$f(x^0 - t) = 0$; $f(x^0) = -10$; $f(x^0 + t) = 80$;

в) так как $f(x^0 - t) > f(x^0) < f(x^0 + t)$, то начальный интервал неопределенности найден: $L_0 = [0, 10]$. Зададим $N = 9$ так, чтобы L_0 содержал $N + 1 = 10$ равных подынтервалов.

Таким образом, начальный интервал неопределенности: $L_0 = [0, 10]$. Положим $l = 1$.

2. Положим $k = 0$.

3⁰. Вычислим $y_0 = a_0 + 0,382(b_0 - a_0) = 0 + 0,382 \cdot 10 = 3,82$;

$$z_0 = a_0 + b_0 - y_0 = 0 + 10 - 3,82 = 6,18.$$

4⁰. Вычислим $f(y_0) = -16,65$;

$$f(z_0) = 2,22.$$

5⁰. Сравним $f(y_0)$ и $f(z_0)$. Так как $f(y_0) < f(z_0)$, то $a_1 = a_0 = 0$, $b_1 = z_0 = 6,18$ (рис. 3.2, а); $y_1 = a_1 + b_1 - y_0 = 0 + 6,18 - 3,82 = 2,36$; $z_1 = y_0 = 3,82$.

6⁰. Получим $L_2 = [0; 6,18]$, $|L_2| = 6,18 > l = 1$. Положим $k = 1$ и перейдем к шагу 4.

4¹. Вычислим $f(y_1) = -17,18$ (новое вычисление), $f(z_1) = f(y_0) = -16,65$ (уже было вычислено на шаге 4⁰).

5¹. Сравним $f(y_1)$ и $f(z_1)$. Так как $f(y_1) < f(z_1)$, то $a_2 = a_1 = 0$, $b_2 = z_1 = 3,82$; $y_2 = a_2 + b_2 - y_1 = 0 + 3,82 - 2,36 = 1,46$; $z_2 = y_1 = 2,36$.

6¹. Получим $L_3 = [0; 3,82]$, $|L_3| = 3,82 > l = 1$. Положим $k = 2$ и перейдем к шагу 4.

4². Вычислим $f(y_2) = -13,25$ (новое вычисление), $f(z_2) = f(y_1) = -17,18$ (уже было вычислено на шаге 4¹).

5². Сравним $f(y_2)$ и $f(z_2)$. Так как $f(y_2) > f(z_2)$, то $a_3 = y_2 = 1,46$; $b_3 = b_2 = 3,82$; $y_3 = z_2 = 2,36$; $z_3 = a_3 + b_3 - y_2 = 1,46 + 3,82 - 2,36 = 2,92$.

6². Получим $L_4 = [1,46; 3,82]$, $|L_4| = 3,82 - 1,46 = 2,36 > l = 1$. Положим $k = 3$ и перейдем к шагу 4.

4³. Вычислим $f(y_3) = f(z_2) = -17,18$ (уже было вычислено на шаге 4²), $f(z_3) = -17,99$.

5³. Сравним $f(y_3)$ и $f(z_3)$. Так как $f(y_3) > f(z_3)$, то $a_4 = y_3 = 2,36$; $b_4 = b_3 = 3,82$; $y_4 = z_3 = 2,92$; $z_4 = a_4 + b_4 - y_3 = 2,36 + 3,82 - 2,92 = 3,26$.

6³. Получим $L_5 = [2,36; 3,82]$, $|L_5| = 3,82 - 2,36 = 1,46 > l = 1$. Положим $k = 4$ и перейдем к шагу 4.

4⁴. Вычислим $f(y_4) = f(z_3) = -17,99$ (было известно), $f(z_4) = -17,86$.

5⁴. Сравним $f(y_4)$ и $f(z_4)$. Так как $f(y_4) < f(z_4)$, то $a_5 = a_4 = 2,36$; $b_5 = z_4 = 3,26$; $y_5 = a_5 + b_5 - y_4 = 2,36 + 3,26 - 2,92 = 2,7$; $z_5 = y_4 = 2,92$.

6⁴. Получим $L_6 = [2,36; 3,26]$, $|L_6| = 3,26 - 2,36 = 0,9 < l = 1$, $x^* \in L_6$, $N = 6$,

$$x^* = \frac{2,36 + 3,26}{2} = 2,81.$$

Первые итерации поиска изображены на рис. 3.3.

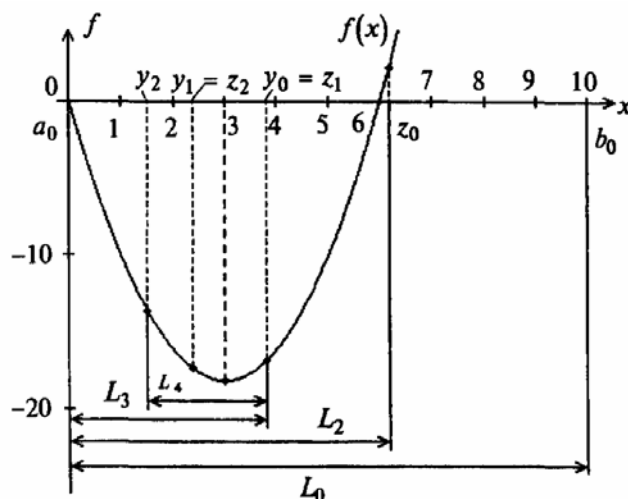


Рис. 3.3 – Первые итерации поиска

Основы создания классов, методов, конструкторов в Java изложены в приложении А, параграфы 1 – 12.

Задание

Определить класс «function». Создать метод для нахождения минимума функции $f(x)$ методом золотого сечения. Определить методы вывода результатов расчетов.

Определить, сколько итераций потребуется для достижения точности $l = 10^{-3}, 10^{-5}, 10^{-7}$ при одних и тех же начальных приближениях.

Варианты заданий:

1. $f(x) = 1/\sqrt{x} - x^2 + 4x^3 + \sin x$
2. $f(x) = x^5 + x^{1/5} - 20x^{-1/2} - \cos x$
3. $f(x) = \sqrt[4]{x} + 2x^2 - 2x^{1/7} + \operatorname{tg} x$
4. $f(x) = 15x^3 - 5/x + 10x^{-3/2} + \cos x$
5. $f(x) = x^{2/3} + x^{-3} + 2x - \sin x$
6. $f(x) = x^3 - x^{5/4} + 7\sqrt{x} + 2\sin x$
7. $f(x) = x^{-0.5} + 2x - 1/x - \operatorname{tg} x$
8. $f(x) = 2/x^2 + x^{1/4} - \sqrt{x} - 2\cos x$
9. $f(x) = x^7 + 3x^{1/3} - 2x + 3\sin x$
10. $f(x, y) = x^{-3} - 2x^3 - 6\sqrt[3]{x} - \cos x$
11. $f(x) = x^2 + x^3 - 2x^{3/5} - \operatorname{tg} x$
12. $f(x) = x^{0.8} + 2x^5 + 9\sqrt[3]{x} + \sin x$
13. $f(x) = 5/(x^2) + 10x^2 - 5x^3 - 2\sin x$
14. $f(x) = 4x + 3/x^2 - x^{-0.3} + 3\cos x$
15. $f(x) = x^2 - 2/\sqrt{x} - 2x^3 - 5\sin x$

Содержание отчета

- 1 Тема и цель работы.
- 2 Диаграмма классов и текст программы.
- 3 Результаты выполнения программы.

Контрольные вопросы

- 1 Сущность метода золотого сечения.
- 2 Основные изученные понятия языка Java (приложение А, параграфы 1 – 12).

Лабораторная работа № 4

Методы нулевого порядка безусловной многомерной минимизации. Метод сопряженных направлений (метод Пауэлла)

4 часа

Цель работы:

- 1) Изучить метод сопряженных направлений;
- 2) изучить перегрузку конструкторов в Java.

Постановка задачи

Требуется найти безусловный минимум функции $f(x)$ многих переменных, т.е. найти такую точку $x^* \in R^n$, что $f(x^*) = \min_{x \in R^n} f(x)$.

Определение. Пусть H – симметрическая матрица размера $n \times n$. Векторы $d_1, d_2, \dots, d_n$ называются H -сопряженными или просто *сопряженными*, если $d_i^T H d_j = 0$ при всех $i \neq j$.

Стратегия поиска

В методе сопряженных направлений (методе Пауэлла [Powell M.J.D.]) используется факт, что минимум квадратичной функции может быть найден не более чем за n шагов при условии, что поиск ведется вдоль сопряженных относительно матрицы Гессе направлений. Так как достаточно большой класс целевых функций может быть представлен в окрестности точки минимума своей квадратичной аппроксимацией, описанная идея применяется и для неквадратичных функций. Задается начальная точка и направления $d_1, d_2, \dots, d_n$, совпадающие с координатными. Находится минимум $f(x)$ при последовательном движении по $(n + 1)$ направлениям с помощью одного из методов одномерной минимизации. При этом полученная ранее точка минимума берется в качестве исходной для поиска по следующему направлению, а направление d_n используется как при первом ($d_0 = d_n$), так и последнем поиске. Находится новое направление поиска, сопряженное с d_n . Оно проходит через точки, полученные при первом и последнем поиске. Заменяется d_1 на d_2 , d_2 на d_3 и т.д. Направление d_n заменяется сопряженным направлением, после чего повторяется поиск по $(n + 1)$ направлениям, уже не содержащим старого направления d_1 . Для квадратичных функций последовательность n^2 одномерных поисков приводит к точке минимума (если все операции выполнены точно). Построение сопряженного направления для квадратичной функции при $n = 2$ изображено на рис. 4.1. Оно проходит через точки 1 и 3.

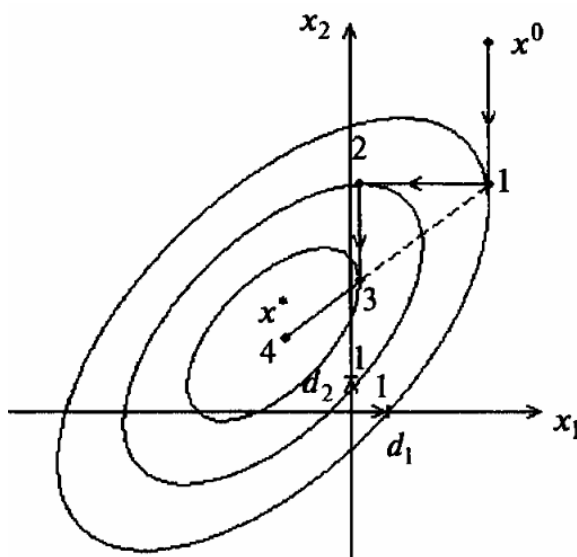


Рис. 4.1 – Построение сопряженного направления для квадратичной функции при $n = 2$

Алгоритм

Шаг 1. Задать начальную точку x^0 , число $\varepsilon > 0$ для окончания алгоритма, начальные направления поиска

$$d_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, d_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, d_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Положим $d_0 = d_n, i = 0, y^0 = x^0, k = 0$.

Шаг 2. Найти $y^{i+1} = y^i + t_i d_i$, где шаг t_i находится в результате поиска минимума функции $f(y^i + t_i d_i)$ по t_i одним из методов одномерной минимизации.

Шаг 3. Проверить выполнение условий:

а) если $i < n - 1$, положить $i = i + 1$ и перейти к шагу 2;

б) если $i = n - 1$, проверить успешность поиска по n первым направлениям.

Если $y^n = y^0$, то поиск завершить: $x^* \equiv y^n$, иначе положить $i = i + 1$ и перейти к шагу 2;

в) если $i = n$, проверить успешность поиска по n последним направлениям.

Если $y^{n+1} = y^1$, поиск завершить: $x^* \equiv y^{n+1}$, иначе перейти к шагу 4 для построения сопряженного направления.

Шаг 4. Положить $x^{k+1} = y^{n+1}$ и проверить условие окончания:

а) если $\|x^{k+1} - x^k\| < \varepsilon$, то поиск завершить: $x^* \equiv x^{k+1}$;

б) иначе положить: $\bar{d}_0 = \bar{d}_n = y^{n+1} - y^1$ (новое направление);

$\bar{d}_i = d_{i+1}, i = 1, \dots, n - 1$ (исключается старое направление).

Если при этом $\text{rang}(\bar{d}_1, \dots, \bar{d}_n) = n$, то новая система направлений линейно независима. В этом случае положить: $\bar{d}_i = d_i, i = 0, 1, \dots, n; k = k + 1, i = 0, y^0 = x^{k+1}$ и перейти к шагу 2.

Если $\text{rang}(\bar{d}_1, \dots, \bar{d}_n) < n$, то новая система направлений линейно зависима. Тогда следует продолжать поиск в старых направлениях. Для этого положить: $d_i = d_i, i = 0, 1, \dots, n; y^0 = x^{k+1}, k = k + 1, i = 0$ и перейти к шагу 2.

В данном алгоритме $\|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$ – евклидова норма вектора x .

Пример. Найти минимум функции $f(x) = 2x_1^2 + x_1 x_2 + x_2^2$ методом Пауэлла.

□ 1°. Зададим $x^0 = (0,5; 1)^T, d_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, d_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \varepsilon = 0,1$. Положим

$$d_0 = d_n = d_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}; y^0 = x^0 = (0,5; 1)^T, i = 0, k = 0.$$

2⁰. Получаем $y^1 = y^0 + t_0 d_0 = (0,5; 1)^T + t_0(0, 1)^T = (0,5; 1 + t_0)^T$. Найдем минимум функции $f(0,5; 1 + t_0) = 0,5 + 0,5(1 + t_0) + (1 + t_0)^2$ по t_0 . Применим метод квадратичной интерполяции, так как минимизируемая функция квадратичная (см. разд. 5.1). Получим $t_0 = -\frac{5}{4}$ (подробности опущены), $y^1 = (0,5; -0,25)^T$.

3⁰. Имеем $i = 0 < n - 1 = 1$, поэтому $i = i + 1 = 1$ и переходим к шагу 2.

2¹. Получаем $y^2 = y^1 + t_1 d_1 = (0,5; -0,25)^T + t_1(1, 0)^T = (0,5 + t_1; -0,25)^T$. Найдем минимум функции $f(0,5 + t_1; -0,25) = 2(0,5 + t_1)^2 - 0,25(0,5 + t_1) + 0,0625$ по t_1 . Аналогично шагу 2⁰ получаем $t_1 = -\frac{7}{16}$, $y^2 = \left(\frac{1}{16}, -\frac{1}{4}\right)^T$.

3¹. Имеем $i = 1 = n - 1$, $y^2 \neq y^0$. Поэтому $i = i + 1 = 2$ и переходим к шагу 2.

2². Получаем $y^3 = y^2 + t_2 d_2 = \left(\frac{1}{16}, -\frac{1}{4}\right)^T + t_2(0, 1)^T = \left(\frac{1}{16}, -\frac{1}{4} + t_2\right)^T$. Найдем безусловный минимум функции $f\left(\frac{1}{16}, t_2 - \frac{1}{4}\right) = \frac{1}{128} + \frac{1}{16}\left(t_2 - \frac{1}{4}\right) + \left(t_2 - \frac{1}{4}\right)^2$. Получаем $t_2 = \frac{7}{32}$, $y^3 = \left(\frac{1}{16}, -\frac{1}{32}\right)^T$.

3². Имеем $i = 2 = n$, $y^3 \neq y^1$. Перейдем к шагу 4.

4⁰. Находим $x^1 = y^3 = \left(\frac{1}{16}, -\frac{1}{32}\right)^T$. Так как $\|x^1 - x^0\| = 1,5 > \varepsilon$, положим $\bar{d}_0 = \bar{d}_n = \bar{d}_2 = y^3 - y^1 = \left(\frac{1}{16}, -\frac{1}{32}\right)^T - (0,5; -0,25)^T = \left(-\frac{7}{16}, \frac{7}{32}\right)^T$, $\bar{d}_1 = d_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Так как $\text{rang} \begin{pmatrix} 0 & -\frac{7}{16} \\ 1 & \frac{7}{32} \end{pmatrix} = 2 = n$, то новая система линейно независима. Положим

$d_0 = d_2 = \left(-\frac{7}{16}, \frac{7}{32}\right)^T$, $d_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $k = k + 1$, $i = 0$, $y^0 = x^1 = \left(\frac{1}{16}, -\frac{1}{32}\right)^T$ и перейдем к шагу 2.

2³. Получаем

$$y^1 = y^0 + t_0 d_0 = \left(\frac{1}{16}, -\frac{1}{32}\right)^T + t_0 \left(-\frac{7}{16}, \frac{7}{32}\right)^T = \left(\frac{1}{16} - \frac{7}{16}t_0, -\frac{1}{32} + \frac{7}{32}t_0\right)^T.$$

Найдем минимум функции $f\left(\frac{1}{16} - \frac{7}{16}t_0, -\frac{1}{32} + \frac{7}{32}t_0\right)^T = 2\left(\frac{1}{16} - \frac{7}{16}t_0\right)^2 + \left(\frac{1}{16} - \frac{7}{16}t_0\right)\left(\frac{7}{32}t_0 - \frac{1}{32}\right) + \left(\frac{7}{32}t_0 - \frac{1}{32}\right)^2$ по t_0 . Получаем $t_0 = \frac{1}{7}$, $y^1 = (0, 0)^T$.

3³. Имеем $i = 0 < n - 1 = 1$, поэтому $i = i + 1 = 1$ и перейдем к шагу 2.

2⁴. Получаем $y^2 = y^1 + t_1 d_1 = (0, 0)^T + t_1(0, 1)^T = (0, t_1)^T$. Найдем минимум функции $f(0, t_1) = t_1^2$ по t_1 . В результате находим $t_1 = 0$, $y^2 = y^1$.

3⁴. Имеем $i = 1 = n - 1$, $y^2 \neq y^0$, поэтому $i = i + 1 = 2$ и перейдем к шагу 2.

$$2^5. \text{ Получаем } y^3 = y^2 + t_2 d_2 = (0,0)^T + t_2 \left(-\frac{7}{16}, \frac{7}{32} \right)^T = \left(-\frac{7}{16} t_2, \frac{7}{32} t_2 \right)^T.$$

$$\text{Найдем минимум функции } f\left(-\frac{7}{16} t_2, \frac{7}{32} t_2\right) = \frac{49}{256} t_2^2 - \frac{49}{512} t_2^2 + \frac{49}{1024} t_2^2 = \frac{3 \cdot 49}{1024} t_2^2$$

по t_2 . Очевидно, $t_2 = 0$ и $y^3 = (0,0)^T$.

3⁵. Так как $i = 2 = n$, $y^3 = y^1$, то процесс поиска минимума завершается:
 $x^* \equiv y^3 = (0,0)^T$; $f(x^*) = 0$. ■

Основные понятия языка Java и основы работы в NetBeans изложены в приложении А, параграф 13.

Задание

Решить задание из части 2 лабораторной работы № 2 методом Пауэлла.

Определить, сколько итераций потребуется для достижения точности $\varepsilon = 10^{-3}, 10^{-5}, 10^{-7}$ при одних и тех же начальных приближениях.

Реализовать перегрузку конструкторов следующим образом:

- 1) Создать конструктор, которому будет передаваться в качестве параметра только точность ε . Начальную точку и направление поиска задать в конструкторе константами;
- 2) создать конструктор, которому будет передаваться в качестве параметров точность и начальная точка. Направление поиска задать в конструкторе константой;
- 3) создать конструктор, которому будет передаваться в качестве параметров точность, начальная точка и направление поиска.

Содержание отчета

- 1 Тема и цель работы.
- 2 Диаграмма классов и текст программы.
- 3 Результаты выполнения программы.

Контрольные вопросы

- 1 Сущность метода сопряженных направлений.
- 2 Основные изученные понятия языка Java (приложение А, параграф 13).

Лабораторная работа № 5

Методы первого порядка безусловной многомерной минимизации (градиентные методы). Метод Гаусса-Зейделя

4 часа

Цель работы:

- 1) Изучить метод Гаусса-Зейделя;
- 2) изучить перегрузку методов в Java.

Постановка задачи

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т.е. найти такую точку $x^* \in R^n$, что $f(x^*) = \min_{x \in R^n} f(x)$.

Стратегия поиска

Стратегия метода Гаусса-Зейделя [Gauss-Seidel] состоит в построении последовательности точек $\{x^k\}$, $k = 0, 1, \dots$, таких, что $f(x^{k+1}) < f(x^k)$, $k = 0, 1, \dots$.

Точки последовательности $\{x^k\}$ вычисляются по правилу

$$x^{jk+1} = x^{jk} - t_k \left(\frac{\partial f(x)}{\partial x_{k+1}} \right)_{x=x^{jk}} \cdot e_{k+1}, \quad (6.6)$$

где j - номер цикла вычислений, $j = 0, 1, 2, \dots$; k - номер итерации внутри цикла, $k = 0, 1, \dots, n-1$; e_{k+1} - единичный вектор, $(k+1)$ -я проекция которого равна 1; точка x^{00} задается пользователем, величина шага t_k выбирается из условия

$$\varphi(t_k) = f \left(x^{jk} - t_k \left(\frac{\partial f(x)}{\partial x_{k+1}} \right)_{x=x^{jk}} \cdot e_{k+1} \right) \rightarrow \min_{t_k}.$$

Эта задача является задачей одномерной минимизации функции $\varphi(t_k) = f \left(x^{jk} - t_k \left(\frac{\partial f(x)}{\partial x_{k+1}} \right)_{x=x^{jk}} \cdot e_{k+1} \right)$ и может быть решена либо с использованием

условий $\frac{d\varphi}{dt_k} = 0$, $\frac{d^2\varphi}{dt_k^2} > 0$, либо численно с использованием методов одномерной минимизации, как задача $\varphi(t_k) \rightarrow \min_{t_k \in [a, b]}$ (см. разд. 5.1).

Если уравнение $\frac{d\varphi}{dt_k} = 0$ имеет высокую степень и корни его трудно определить, можно аппроксимировать функцию $\varphi(t_k)$ полиномом $P(t_k)$ второй или третьей степени и определить t_k^* из условий $\frac{dP}{dt_k} = 0$, $\frac{d^2P}{dt_k^2} > 0$.

При численном решении задачи определения величины шага степень близости найденного значения t_k к оптимальному значению t_k^* , удовлетворяющему

условиям $\frac{d\varphi}{dt_k} = 0$, $\frac{d^2\varphi}{dt_k^2} > 0$, зависит от задания интервала $[a, b]$ и точности методов одномерной минимизации [28].

Легко видеть, что при фиксированном j за одну итерацию с номером k изменяется только одна проекция точки x^{jk} , имеющая номер $k+1$, а в течение всего цикла с номером j , т.е. начиная с $k=0$ и кончая $k=n-1$, изменяются все n проекций точки x^{j0} . После этого точке x^{jn} присваивается номер $x^{j+1,0}$ и она берется за начальную точку для вычислений в $(j+1)$ -м цикле.

Расчет заканчивается в точке x^{jk} при выполнении по крайней мере одного из трех критериев окончания счета: $\|\nabla f(x^{jk})\| < \varepsilon_1$, или $k \geq M$, или двукратного выполнения неравенств $\|x^{jk+1} - x^{jk}\| < \varepsilon_2$, $|f(x^{jk+1}) - f(x^{jk})| < \varepsilon_2$. Здесь $\varepsilon_1, \varepsilon_2$ - малые положительные числа, M - предельное число циклов итераций.

Полученные в результате вычислений точки могут быть записаны как элементы последовательности $\{x^l\}$, где $l = n \cdot j + k$ - порядковый номер точки, т.е. $\{x^l\} = \{x^0 = x^{00}, x^1 = x^{01}, \dots, x^n = x^{0n} = x^{10}, x^{n+1} = x^{11}, x^{n+2} = x^{12}, \dots\}$.

Алгоритм

Шаг 1. Задать x^{00} , $\varepsilon_1 > 0$, $\varepsilon_2 > 0$; предельное число M циклов счета, кратное n , где n - размерность вектора x . Найти градиент $\nabla f(x)$.

Шаг 2. Задать номер цикла $j = 0$.

Шаг 3. Проверить условие $j \geq M$:

- а) если $j \geq M$, то расчет окончен и $x^* = x^{jk}$;
- б) если $j < M$, то перейти к шагу 4.

Шаг 4. Задать $k = 0$.

Шаг 5. Проверить условие $k \leq n-1$:

- а) если $k \leq n-1$, то перейти к шагу 6;
- б) если $k = n$, то положить $j = j+1$ и перейти к шагу 3.

Шаг 6. Вычислить $\nabla f(x^{jk})$.

Шаг 7. Проверить выполнение условия $\|\nabla f(x^{jk})\| < \varepsilon_1$:

- а) если условие выполнено, то расчет окончен и $x^* = x^{jk}$;
- б) если нет, то перейти к шагу 8.

Шаг 8. Вычислить t_k^* из условия

$$\varphi(t_k) = f\left(x^{jk} - t_k \left(\frac{\partial f(x)}{\partial x_{k+1}}\right)_{x=x^{jk}} \cdot e_{k+1}\right) \rightarrow \min_{t_k}.$$

Шаг 9. Вычислить $x^{jk+1} = x^{jk} - t_k^* \left(\frac{\partial f(x)}{\partial x_{k+1}}\right)_{x=x^{jk}} \cdot e_{k+1}$.

Шаг 10. Проверить выполнение условий

$$\|x^{jk+1} - x^{jk}\| < \varepsilon_2, \quad |f(x^{jk+1}) - f(x^{jk})| < \varepsilon_2:$$

а) если оба условия выполнены в двух последовательных циклах с номерами j и $j-1$, то расчет окончен, найдена точка $x^* = x^{jk+1}$;

б) если не выполняется хотя бы одно условие, положить $k = k+1$ и перейти к шагу 5.

Геометрическая интерпретация метода для $n=2$ приведена на рис. 5.1.

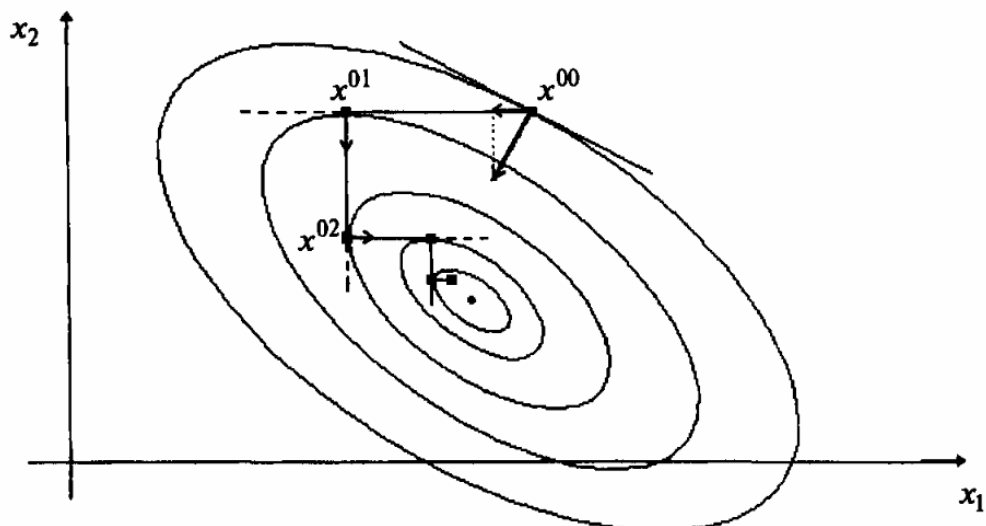


Рис. 5.1 – Геометрическая интерпретация метода Гаусса–Зейделя

Пример. Найти локальный минимум функции

$$f(x) = 2x_1^2 + x_1x_2 + x_2^2.$$

□ I. Определение точки x^{jk} , в которой выполнен хотя бы один из критериев окончания расчетов.

1. Зададим $x^{00}, \varepsilon_1, \varepsilon_2, M$: $x^{00} = (0,5; 1)^T$, $\varepsilon_1 = 0,1$; $\varepsilon_2 = 0,15$; $M = 10$. Найдем градиент функции в произвольной точке $\nabla f(x) = (4x_1 + x_2; x_1 + 2x_2)^T$.

2. Зададим $j = 0$.

3⁰. Проверим выполнение условия $j \geq M$: $j = 0 < 10 = M$.

4⁰. Зададим $k = 0$.

5⁰. Проверим выполнение условия $k \leq n-1$: $k = 0 < 1 = n-1$.

6⁰. Вычислим $\nabla f(x^{00})$: $\nabla f(x^{00}) = (3; 2,5)^T$.

7⁰. Проверим условие $\|\nabla f(x^{00})\| < \varepsilon_1$: $\|\nabla f(x^{00})\| = 3,9 > 0,1$.

8⁰. Определим величину шага t_0^* из условия

$$\varphi(t_0) = f\left(x^{j0} - t_0 \left(\frac{\partial f(x)}{\partial x_1}\right)_{x=x^{j0}} \cdot e_1\right) \rightarrow \min_{t_0}.$$

Воспользуемся формулой (6.6) при $k=0, j=0$: $x^{01} = x^{00} - t_0 \left(\frac{\partial f(x)}{\partial x_1}\right)_{x=x^{00}} \cdot e_1$.

Поскольку $\left(\frac{\partial f(x)}{\partial x_1}\right)_{x=x^{00}} = 4x_1 + x_2 \Big|_{x=x^{00}} = 2 + 1 = 3$, $e_1 = (1; 0)^T$, то $x^{01} = (0,5; 1)^T - t_0 \cdot 3 \cdot (1; 0)^T = (0,5 - 3t_0; 1)^T$ или $x_1^{01} = 0,5 - 3t_0$, $x_2^{01} = 1$. Подставляя полученные выражения в $f(x)$, имеем $\varphi(t_0) = 2(0,5 - 3t_0)^2 + (0,5 - 3t_0) \cdot 1 + 1$.

Из необходимого условия экстремума $\frac{d\varphi(t_0)}{dt_0} = 4 \cdot (0,5 - 3t_0) \cdot (-3) - 3 = 0$ или $36t_0 - 9 = 0$ находим $t_0^* = \frac{1}{4}$. Так как $\frac{d^2\varphi(t_0)}{dt_0^2} = 36 > 0$, то найденное значение шага обеспечивает минимум функции $\varphi(t_0)$ по t_0 .

Можно показать, что в силу структуры заданной функции $f(x)$ величина шага в направлении $-\left(\frac{\partial f(x)}{\partial x_1}\right)_{x=x^{jk}} \cdot e_1$ не зависит от x^k , является постоянной и равной $\frac{1}{4}$.

9⁰. Определим $x^{01} = x^{00} - t_0^* \left(\frac{\partial f(x)}{\partial x_1}\right)_{x=x^{00}} \cdot e_1$: $x^{01} = (-0,25; 1)^T$.

10⁰. Проверим условия $\|x^{01} - x^{00}\| < \varepsilon_2$, $|f(x^{01}) - f(x^{00})| < \varepsilon_2$:

$$\|x^{01} - x^{00}\| = 0,25 > 0,15, \quad |f(x^{01}) - f(x^{00})| = |0,875 - 2| = 1,125 > 0,15.$$

Полагаем $k = 1$, переходим к шагу 5.

5¹. Проверим условие $k \leq n - 1$: $k = 1 = n - 1$.

6¹. Вычислим $\nabla f(x^{01})$: $\nabla f(x^{01}) = (0; 1,75)^T$.

7¹. Проверим условие $\|\nabla f(x^{01})\| < \varepsilon_1$: $\|\nabla f(x^{01})\| = 1,75 > 0,1$.

8¹. Определим величину шага t_1^* из условия

$$\varphi(t_1) = f\left(x^{j1} - t_1 \left(\frac{\partial f(x)}{\partial x_2}\right)_{x=x^{j1}} \cdot e_2\right) \rightarrow \min_{t_1}.$$

Воспользуемся формулой (6.6) при $k = 1, j = 0$: $x^{02} = x^{01} - t_1 \left(\frac{\partial f(x)}{\partial x_2}\right)_{x=x^{01}} \cdot e_2$.

Поскольку $\left(\frac{\partial f(x)}{\partial x_2}\right)_{x=x^{01}} = x_1 + 2x_2 \Big|_{x=x^{01}} = -0,25 + 2 = 1,75$, $e_2 = (0; 1)^T$, то

$x^{02} = (-0,25; 1)^T - t_1 \cdot 1,75 \cdot (0; 1)^T = (-0,25; 1 - 1,75 \cdot t_1)^T$ или $x_1^{02} = -0,25$, $x_2^{02} = 1 - 1,75 \cdot t_1$. Подставляя полученные выражения в $f(x)$, имеем

$$\varphi(t_1) = 2(-0,25)^2 + (-0,25) \cdot (1 - 1,75 \cdot t_1) + (1 - 1,75 \cdot t_1)^2.$$

Из необходимого условия экстремума

$$\frac{d\varphi(t_1)}{dt_1} = 0,25 \cdot 1,75 + 2 \cdot (1 - 1,75 \cdot t_1) \cdot (-1,75) = 0 \quad \text{или} \quad 2 \cdot 1,75^2 t_1 - 1,75^2 = 0$$

находим $t_1^* = \frac{1}{2}$. Так как $\frac{d^2\varphi(t_1)}{dt_1^2} = 2 \cdot 1,75^2 > 0$, то найденное значение шага обеспечивает минимум функции $\varphi(t_1)$ по t_1 .

Можно показать, что в силу структуры функции $f(x)$ величина шага в направлении $-\left(\frac{\partial f(x)}{\partial x_2}\right)_{x=x^k} \cdot e_2$ остается постоянной и равной $\frac{1}{2}$.

$$9^1. \text{ Вычислим } x^{02} = x^{01} - t_1^* \left(\frac{\partial f(x)}{\partial x_2}\right)_{x=x^{01}} \cdot e_2: \quad x^{02} = (-0,25; 0,125)^T.$$

$$10^1. \text{ Проверим условия } \|x^{02} - x^{01}\| < \varepsilon_2, \quad |f(x^{02}) - f(x^{01})| < \varepsilon_2:$$

$$\|x^{02} - x^{01}\| = 0,875 > 0,15, \quad |f(x^{02}) - f(x^{01})| = |0,12 - 0,875| = 0,755 > 0,15.$$

Полагаем $k = 2$, переходим к шагу 5.

5². Проверим условие $k \leq n - 1$: $k = 2 = n$. Полагаем $j = 1, x^{10} = x^{02}$, переходим к шагу 3.

$$3^1. \text{ Проверим условие } j \geq M: j = 1 < 10 = M.$$

$$4^1. \text{ Зададим } k = 0.$$

$$5^3. \text{ Проверим условие } k \leq n - 1: k = 0 < 1 = n - 1.$$

$$6^3. \text{ Вычислим } \nabla f(x^{10}): \nabla f(x^{10}) = \nabla f(x^{02}) = (-0,875; 0,00)^T.$$

$$7^3. \text{ Проверим условие } \|\nabla f(x^{10})\| < \varepsilon_1: \|\nabla f(x^{10})\| = 0,875 > 0,1.$$

$$8^3. \text{ Полагаем } t_0^* = 0,25 \text{ (см. п. 8}^0\text{)}.$$

$$9^3. \text{ Вычислим } x^{11} = x^{10} - t_0^* \left(\frac{\partial f(x)}{\partial x_1}\right)_{x=x^{10}} \cdot e_1: \quad x^{11} = (-0,03; 0,125)^T.$$

$$10^3. \text{ Проверим условия } \|x^{11} - x^{10}\| < \varepsilon_2, \quad |f(x^{11}) - f(x^{10})| < \varepsilon_2:$$

$$\|x^{11} - x^{10}\| = 0,22 > 0,15, \quad |f(x^{11}) - f(x^{10})| = |0,013 - 0,1375| = 0,124 < 0,15.$$

Полагаем $k = 1$, переходим к шагу 5.

$$5^4. \text{ Проверим условие } k \leq n - 1: k = 1 = n - 1.$$

$$6^4. \text{ Вычислим } \nabla f(x^{11}): \nabla f(x^{11}) = (0,005; 0,22)^T.$$

$$7^4. \text{ Проверим условие } \|\nabla f(x^{11})\| < \varepsilon_1: \|\nabla f(x^{11})\| = 0,22 > 0,1.$$

$$8^4. \text{ Зададим } t_1^* = 0,5 \text{ (см. п. 8}^1\text{)}.$$

9⁴. Вычислим $x^{12} = x^{11} - t_1^* \left(\frac{\partial f(x)}{\partial x_2} \right)_{x=x^{11}} \cdot e_2$: $x^{12} = (-0,03; 0,015)^T$.

10⁴. Проверим условия $\|x^{12} - x^{11}\| < \varepsilon_2$, $|f(x^{12}) - f(x^{11})| < \varepsilon_2$:

$$\|x^{12} - x^{11}\| = 0,11 < 0,15, \quad |f(x^{12}) - f(x^{11})| = |0,0015 - 0,013| = 0,0115 < 0,15.$$

Полагаем $k = 2$ и переходим к шагу 5.

5⁵. Проверим условие $k \leq n - 1$: $k = 2 = n$. Полагаем $j = 2$, $x^{20} = x^{12}$ и переходим к шагу 3.

3². Проверим условие $j \geq M$: $j = 2 < 10 = M$.

4². Зададим $k = 0$.

5⁶. Проверим условие $k \leq n - 1$: $k = 0 < 1 = n - 1$.

6⁵. Вычислим $\nabla f(x^{20})$: $\nabla f(x^{20}) = (-0,105; 0)^T$.

7⁵. Проверим условие $\|\nabla f(x^{20})\| < \varepsilon_1$: $\|\nabla f(x^{20})\| = 0,105 > \varepsilon_1$.

8⁵. Зададим $t_0^* = 0,25$ (см. п. 8⁰).

9⁵. Вычислим $x^{21} = x^{20} - t_0^* \left(\frac{\partial f(x)}{\partial x_1} \right)_{x=x^{20}} \cdot e_1$: $x^{21} = (-0,004; 0,015)^T$.

10⁵. Проверим условия $\|x^{21} - x^{20}\| < \varepsilon_2$, $|f(x^{21}) - f(x^{20})| < \varepsilon_2$:

$$\|x^{21} - x^{20}\| = 0,026 < 0,15, \quad |f(x^{21}) - f(x^{20})| = |0,000197 - 0,0015| = 0,0013 < 0,15.$$

Условия $\|x^{jk+1} - x^{jk}\| < \varepsilon_2$, $|f(x^{jk+1}) - f(x^{jk})| < \varepsilon_2$ выполнены в двух последовательных циклах с номерами $j = 2$ и $j - 1 = 1$. Расчет окончен, найдена точка $x^{21} = (-0,004; 0,015)^T$; $f(x^{21}) = 0,000197$. На рис. 6.5 полученные точки последовательности $x^{00} \rightarrow x^{01} \rightarrow x^{02} = x^{10} \rightarrow x^{11} \rightarrow x^{12} = x^{20} \rightarrow x^{21}$ соединены сплошной линией. Очевидно, метод Гаусса-Зейделя сходится быстрее, чем метод покоординатного спуска.

11. Анализ точки x^{21} .

Точка x^{21} является найденным приближением точки глобального минимума $f(x)$, так как функция $f(x)$ строго выпуклая (см. пример 6.1). ■

Основные понятия языка Java и основы работы в NetBeans изложены в приложении А, параграф 14.

Задание

Реализовать перегрузку метода поиска минимума функции $f(x, y)$ из части 2 лабораторной работы № 2 методом Гаусса-Зейделя.

Определить, сколько итераций потребуется для достижения точности $\varepsilon = 10^{-3}, 10^{-5}, 10^{-7}$ при одних и тех же начальных приближениях.

Содержание отчета

- 1 Тема и цель работы.
- 2 Диаграмма классов и текст программы.
- 3 Результаты выполнения программы.

Контрольные вопросы

- 1 Сущность метода Гаусса-Зейделя.
- 2 Основные изученные понятия языка Java (приложение А, параграф 14).

Лабораторная работа № 6

Методы второго порядка безусловной многомерной минимизации. Метод Ньютона

4 часа

Цель работы:

- 1) Изучить метод Ньютона;
- 2) изучить наследование в Java.

Постановка задачи

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т.е. найти такую точку $x^* \in R^n$, что $f(x^*) = \min_{x \in R^n} f(x)$, $f(x) \in C^2$.

Стратегия поиска

Стратегия метода Ньютона [Newton I.] состоит в построении последовательности точек $\{x^k\}, k = 0, 1, \dots$, таких, что $f(x^{k+1}) < f(x^k)$, $k = 0, 1, \dots$. Точки последовательности вычисляются по правилу

$$x^{k+1} = x^k + d^k, \quad k = 0, 1, \dots, \quad (7.1)$$

где x^0 - задается пользователем, а направление спуска d^k определяется для каждого значения k по формуле

$$d^k = -H^{-1}(x^k) \nabla f(x^k). \quad (7.2)$$

Выбор d^k по формуле (7.2) гарантирует выполнение требования $f(x^{k+1}) < f(x^k)$ при условии, что $H(x^k) > 0$. Формула (7.2) получена из следующих соображений:

1. Функция $f(x)$ аппроксимируется в каждой точке последовательности $\{x^k\}$ квадратичной функцией $F_k = f(x^k) + (\nabla f(x^k), d^k) + \frac{1}{2}(d^k, H(x^k)d^k)$.

2. Направление d^k определяется из необходимого условия экстремума первого порядка: $\frac{dF_k}{dd^k} = 0$. Таким образом, при выполнении требования $H(x^k) > 0$ последовательность является последовательностью точек минимумов квадратичных функций F_k , $k = 0, 1, \dots$ (рис. 7.1). Чтобы обеспечить выполнение требования $f(x^{k+1}) < f(x^k)$, $k = 0, 1, \dots$, даже в тех случаях, когда для каких-либо значений матрица Гессе $H(x^k)$ не окажется положительно определенной, рекомендуется для соответствующих значений k вычислить точку x^{k+1} по методу градиентного спуска $x^{k+1} = x^k - t_k \nabla f(x^k)$ с выбором величины шага t_k из условия $f(x^k - t_k \nabla f(x^k)) < f(x^k)$.

Построение последовательности $\{x^k\}$ заканчивается в точке x^k , для которой $\|\nabla f(x^k)\| < \varepsilon_1$, где ε_1 - заданное малое положительное число, или при $k \geq M$ (M - предельное число итераций), или при двукратном одновременном выполнении двух неравенств $\|x^{k+1} - x^k\| < \varepsilon_2$, $|f(x^{k+1}) - f(x^k)| < \varepsilon_2$, где ε_2 - малое положительное число. Вопрос о том, может ли точка x^k рассматриваться как найденное приближение искомой точки минимума, решается путем проведения дополнительного исследования, которое описано ниже.

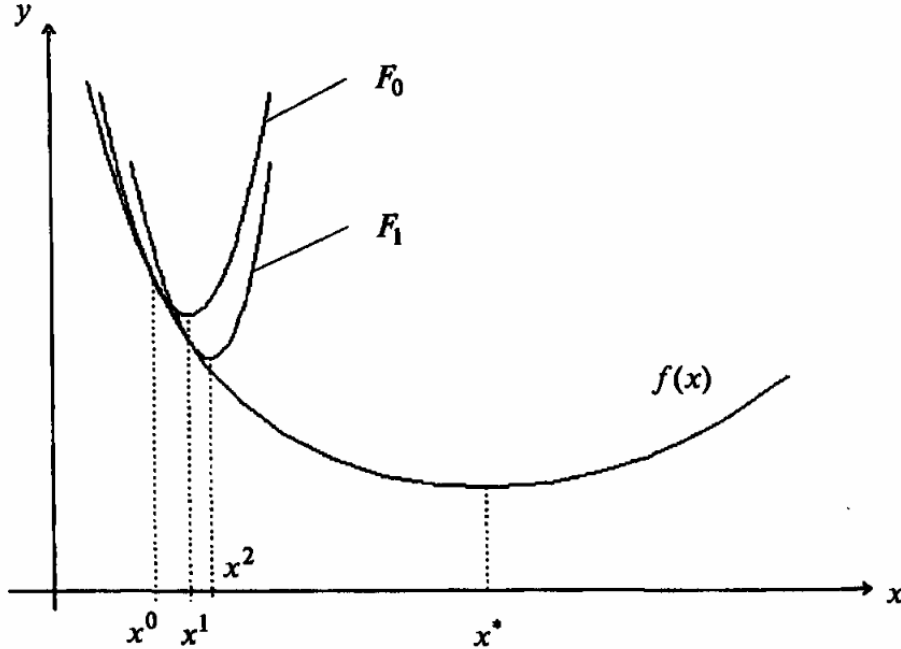


Рис. 7.1

Алгоритм

Шаг 1. Задать x^0 , $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M - предельное число итераций. Найти градиент $\nabla f(x)$ и матрицу Гессе $H(x)$.

Шаг 2. Положить $k = 0$.

Шаг 3. Вычислить $\nabla f(x^k)$.

Шаг 4. Проверить выполнение критерия окончания $\|\nabla f(x^k)\| \leq \varepsilon_1$:

- а) если неравенство выполнено, то расчет окончен и $x^* = x^k$;
- б) в противном случае перейти к шагу 5.

Шаг 5. Проверить выполнение неравенства $k \geq M$:

- а) если неравенство выполнено, расчет окончен и $x^* = x^k$;
- б) если нет, перейти к шагу 6.

Шаг 6. Вычислить матрицу $H(x^k)$.

Шаг 7. Вычислить матрицу $H^{-1}(x^k)$.

Шаг 8. Проверить выполнение условия $H^{-1}(x^k) > 0$:

- а) если $H^{-1}(x^k) > 0$, то перейти к шагу 9;
- б) если нет, то перейти к шагу 10, положив $d^k = -\nabla f(x^k)$.

Шаг 9. Определить $d^k = -H^{-1}(x^k) \nabla f(x^k)$.

Шаг 10. Найти точку $x^{k+1} = x^k + t_k d^k$,

положив $t_k = 1$, если $d^k = -H^{-1}(x^k) \nabla f(x^k)$,

или выбрав t_k из условия $f(x^{k+1}) < f(x^k)$, если $d^k = -\nabla f(x^k)$.

Шаг 11. Проверить выполнение условий

$$\|x^{k+1} - x^k\| < \varepsilon_2, \quad |f(x^{k+1}) - f(x^k)| < \varepsilon_2:$$

а) если оба условия выполнены при текущем значении k и $k = k - 1$, то расчет окончен, $x^* = x^{k+1}$;

б) в противном случае положить $k = k + 1$ и перейти к шагу 3.

Сходимость

Утверждение 7.1. Пусть $f(x)$ дважды непрерывно дифференцируемая сильно выпуклая функция с константой $l > 0$ на R^n и удовлетворяет условию

$$\|H(x) - H(y)\| \leq L \|x - y\| \quad \forall x, y \in R^n,$$

где $L > 0$, а начальная точка такова, что $\|\nabla f(x^0)\| \leq 8 \frac{l^2}{L}$, т.е.

$$\|\nabla f(x^0)\| = \frac{8l^2 q}{L}, \quad (7.3)$$

где $q \in (0,1)$. Тогда последовательность $\{x^k\}$ сходится к точке минимума с квадра-

тичной скоростью $\|x^k - x^*\| \leq \frac{4lq^2}{L} [39]$.

З а м е ч а н и я 7.1.

1. Сходимость метода Ньютона доказана лишь для сильно выпуклых функций и для достаточно хорошего начального приближения, определяемого условием (7.3), практическое использование которого крайне затруднено, так как постоянные l и L , как правило, неизвестны или требуют трудоемкого исследования для их определения. Поэтому при практическом использовании метода Ньютона следует:

а) анализировать матрицу $H(x^k)$ на выполнение условия $H(x^k) > 0$ $\forall k = 0, 1, \dots$ и заменять формулу $x^{k+1} = x^k - H^{-1}(x^k) \nabla f(x^k)$ на формулу $x^{k+1} = x^k - t_k \nabla f(x^k)$ в случае его невыполнения;

б) производить анализ точки x^k с целью выяснения, является ли она найденным приближением искомой точки x^* .

2. При решении задачи поиска безусловного максимума формула (7.2) не изменяется, так как в этом случае $H(x^k) < 0$.

Процедура решения задачи

1. Используя алгоритм Ньютона, найти точку x^k , в которой выполняется по крайней мере один критерий окончания расчета.

2. Так как $f(x) \in C^2$, то осуществить проверку выполнения достаточных условий минимума $H(x^k) > 0$. Если условие выполнено, то точка x^k может рассматриваться как найденное приближение точки минимума x^* . Проверку выполнения достаточных условий минимума можно заменить проверкой функции $f(x)$ на выпуклость.

Пример. Найти локальный минимум функции

$$f(x) = 2x_1^2 + x_1x_2 + x_2^2.$$

□ I. Определение точки x^k , в которой выполняется по крайней мере один критерий окончания расчетов.

1. Зададим $x^0, \varepsilon_1, \varepsilon_2, M$: $x^0 = (0,5; 1)^T$; $\varepsilon_1 = 0,1$; $\varepsilon_2 = 0,15$; $M = 10$. Найдем градиент функции $\nabla f(x) = (4x_1 + x_2; x_1 + 2x_2)^T$ и матрицу Гессе $H(x) = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$.

2. Положим $k = 0$.

3⁰. Вычислим $\nabla f(x^0)$: $\nabla f(x^0) = (3; 2,5)^T$.

4⁰. Проверим выполнение условия $\|\nabla f(x^0)\| \leq \varepsilon_1$: $\|\nabla f(x^0)\| = 3,9 > 0,1$.

Переходим к шагу 5.

5⁰. Проверим выполнение условия $k \geq M$: $k = 0 < 10$. Переходим к шагу 6.

6⁰. Вычислим $H(x^0)$: $H(x^0) = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$.

7⁰. Вычислим $H^{-1}(x^0)$: $H^{-1}(x^0) = \begin{pmatrix} \frac{2}{7} & -\frac{1}{7} \\ -\frac{1}{7} & \frac{4}{7} \end{pmatrix}$.

8⁰. Проверим выполнение условия $H^{-1}(x^0) > 0$. Так как $\Delta_1 = \frac{2}{7} > 0$, $\Delta_2 = \frac{1}{7} > 0$, то согласно критерию Сильвестра $H^{-1}(x^0) > 0$ (см. § 2).

9⁰. Определим $d_0 = -H^{-1}(x^0)\nabla f(x^0) = -\begin{pmatrix} \frac{2}{7} & -\frac{1}{7} \\ -\frac{1}{7} & \frac{4}{7} \end{pmatrix} \begin{pmatrix} 3 \\ 2,5 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -1 \end{pmatrix}$.

10⁰. Вычислим $x^1 = \left(\frac{1}{2}, 1\right)^T + \left(-\frac{1}{2}, -1\right)^T = (0, 0)^T$.

11⁰. Проверим выполнение условий $\|x^1 - x^0\| < \varepsilon_2$, $|f(x^1) - f(x^0)| < \varepsilon_2$:

$$\|x^1 - x^0\| = 1,12 > 0,15; \quad |f(x^1) - f(x^0)| = 2 > 0,15.$$

Полагаем $k = 1$, переходим к шагу 3.

3¹. Вычислим $\nabla f(x^1)$: $\nabla f(x^1) = (0, 0)^T$.

4¹. Проверим выполнение условия $\|\nabla f(x^1)\| < \varepsilon_1$: $\|\nabla f(x^1)\| = 0 < 0,1$. Расчет

окончен. Заметим, что в точке x^1 выполняется необходимое условие первого порядка, поэтому она является стационарной точкой.

II. Анализ точки x^1 .

Функция $f(x) = 2x_1^2 + x_1x_2 + x_2^2$ является строго выпуклой, так как ее матрица вторых производных $H = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} > 0$ в силу того, что $\Delta_1 = 4 > 0$, $\Delta_2 = 7 > 0$.

Найденная точка $x^1 = (0, 0)^T$ есть точка локального и одновременно глобального минимума $f(x)$.

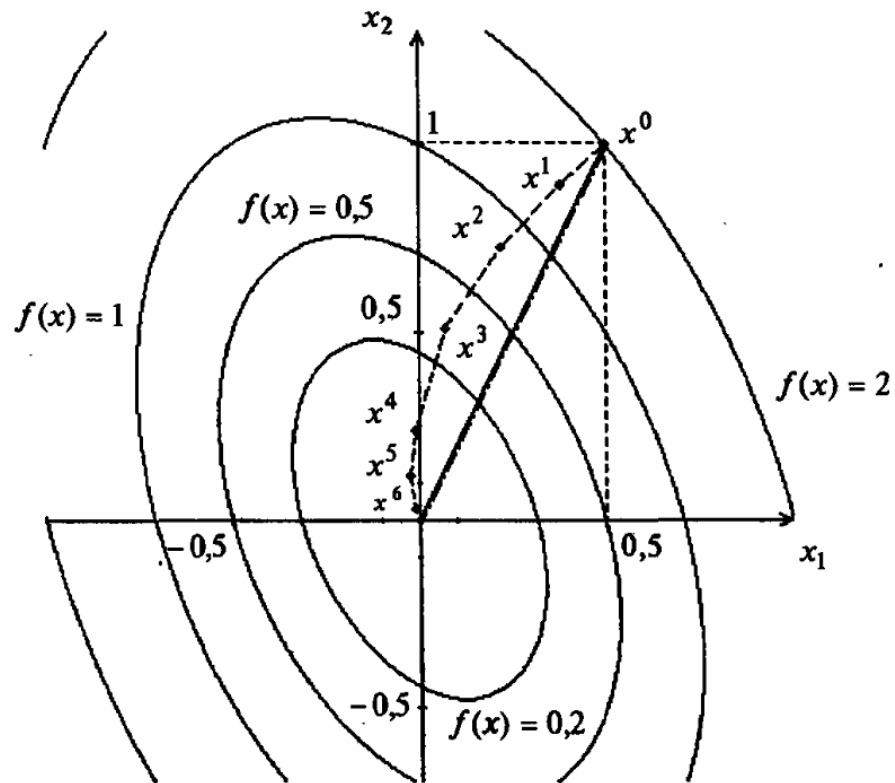


Рис. 7.2

На рис. 7.2 траектория спуска изображена сплошной линией.

Основные понятия языка Java и основы работы в NetBeans изложены в приложении А, параграф 15.

Задание

Создать производный класс от базового класса, реализующего нахождение градиента и матрицы Гессе из лабораторной работы № 2. В производном классе реализовать поиск минимума функции методом Ньютона.

Определить, сколько итераций потребуется для достижения точности $\varepsilon = 10^{-3}, 10^{-5}, 10^{-7}$ при одних и тех же начальных приближениях.

Содержание отчета

- 1 Тема и цель работы.
- 2 Диаграмма классов и текст программы.
- 3 Результаты выполнения программы.

Контрольные вопросы

- 1 Сущность метода Ньютона.
- 2 Основные изученные понятия языка Java (приложение А, параграф 15).

Лабораторная работа № 7

Численные методы поиска условного экстремума

Метод штрафов

4 часа

Цель работы:

- 1) Изучить метод штрафов;
- 2) изучить шаблон Bridge в Java.

Постановка задачи

Даны дважды непрерывно дифференцируемые целевая функция $f(x) = f(x_1, \dots, x_n)$ и функции ограничений $g_j(x) = 0$, $j = 1, \dots, m$; $g_j(x) \leq 0$, $j = m+1, \dots, p$, определяющие множество допустимых решений X .

Требуется найти локальный минимум целевой функции на множестве X , т.е. такую точку $x^* \in X$, что

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

$$\text{где } X = \left\{ x \left| \begin{array}{l} g_j(x) = 0, \quad j = 1, \dots, m; \quad m < n \\ g_j(x) \leq 0, \quad j = m+1, \dots, p \end{array} \right. \right\}.$$

Стратегия поиска

Идея метода заключается в сведении задачи на условный минимум к решению последовательности задач поиска безусловного минимума *вспомогательной функции*:

$$F(x, r^k) = f(x) + P(x, r^k) \rightarrow \min_{x \in R^n},$$

где $P(x, r^k)$ - *штрафная функция*, r^k - параметр штрафа, задаваемый на каждой k -й итерации. Это связано с возможностью применения эффективных и надежных методов поиска безусловного экстремума, изложенных в гл. 2.

Штрафные функции конструируются, исходя из условий:

$$P(x, r^k) = \begin{cases} 0, & \text{при выполнении ограничений,} \\ > 0, & \text{при невыполнении ограничений,} \end{cases}$$

причем при невыполнении ограничений и $r^k \rightarrow \infty$, $k \rightarrow \infty$ справедливо $P(x, r^k) \rightarrow \infty$. Чем больше r^k , тем больше штраф за невыполнение ограничений. Как правило, для ограничений типа равенств используется квадратичный штраф (рис. 9.1, а), а для ограничений типа неравенств - квадрат срезки (рис. 9.1, б):

$$P(x, r^k) = \frac{r^k}{2} \left\{ \sum_{j=1}^m [g_j(x)]^2 + \sum_{j=m+1}^p [g_j^+(x)]^2 \right\},$$

где $g_j^+(x)$ - срезка функции:

$$g_j^+(x) = \max \{0, g_j(x)\} = \begin{cases} g_j(x), & g_j(x) > 0, \\ 0, & g_j(x) \leq 0. \end{cases}$$

Начальная точка поиска задается обычно вне множества допустимых решений X . На каждой k -й итерации ищется точка $x^*(r^k)$ минимума вспомогательной функции $F(x, r^k)$ при заданном параметре r^k с помощью одного из методов безусловной минимизации. Полученная точка $x^*(r^k)$ используется в качестве начальной на следующей итерации, выполняемой при возрастающем значении параметра штрафа. При неограниченном возрастании r^k последовательность точек $x^*(r^k)$ стремится к точке условного минимума x^* .

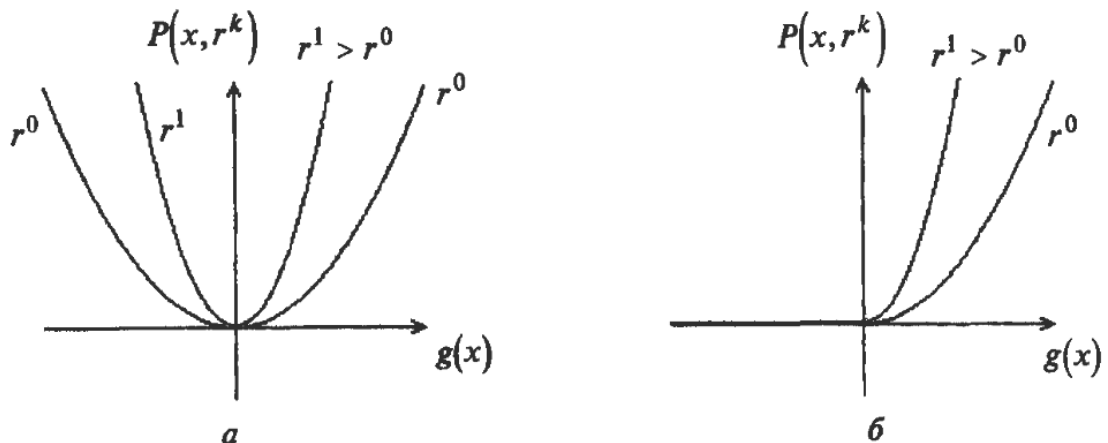


Рис. 9.1
Алгоритм

Шаг 1. Задать начальную точку x^0 ; начальное значение параметра штрафа $r^0 > 0$; число $C > 1$ для увеличения параметра; малое число $\varepsilon > 0$ для остановки алгоритма. Положить $k = 0$.

Шаг 2. Составить вспомогательную функцию

$$F(x, r^k) = f(x) + \frac{r^k}{2} \left\{ \sum_{j=1}^m [g_j(x)]^2 + \sum_{j=m+1}^p [g_j^+(x)]^2 \right\}.$$

Шаг 3. Найти точку $x^*(r^k)$ безусловного минимума функции $F(x, r^k)$ по x с помощью какого-либо метода (нулевого, первого или второго порядка):

$$F(x^*(r^k), r^k) = \min_{x \in R^n} F(x, r^k).$$

При этом задать все требуемые выбранным методом параметры. В качестве начальной точки взять x^k . Вычислить $P(x^*(r^k), r^k)$.

Шаг 4. Проверить условие окончания:

а) если $P(x^*(r^k), r^k) \leq \varepsilon$, процесс поиска закончить:

$$x^* = x^*(r^k), \quad f(x^*) = f(x^*(r^k));$$

б) если $P(x^*(r^k), r^k) > \varepsilon$, положить: $r^{k+1} = C r^k$, $x^{k+1} = x^*(r^k)$, $k = k + 1$

и перейти к шагу 2.

Сходимость

Утверждение 9.1. Пусть x^* - локально единственное решение задачи поиска условного минимума, а функции $f(x)$ и $g_j(x)$ непрерывно дифференцируемы в окрестности точки x^* . Тогда для достаточно больших r^k найдется точка $x^*(r^k)$ локального минимума функции $F(x, r^k)$ в окрестности x^* и $x^*(r^k) \rightarrow x^*$ при $r^k \rightarrow \infty$ [34].

З а м е ч а н и я 9.1.

1. Так как сходимость метода обеспечивается при $r^k \rightarrow \infty$, то возникает вопрос о том, нельзя ли получить решение исходной задачи в результате однократного поиска безусловного минимума вспомогательной функции с параметром r^k , равным большому числу, например 10^{20} . Однако такая замена последовательного решения вспомогательных задач не представляется возможной, так как с ростом r^k функция $F(x, r^k)$ приобретает ярко выраженную овражную структуру. Поэтому скорость сходимости любого метода безусловной минимизации к решению $x^*(r^k)$ резко падает, так что процесс его определения заканчивается, как правило, значительно раньше, чем будет достигнута заданная точность, и, следовательно, полученный результат не дает возможности судить об искомом решении x^* .

2. Точки $x^*(r^k)$ в алгоритме - это точки локального минимума функции $F(x, r^k)$. Однако функция $F(x, r^k)$ может быть неограниченной снизу и процедуры методов безусловной минимизации могут расходиться. Это обстоятельство необходимо учитывать при программной реализации.

3. В методах штрафных функций имеется тесная связь между значениями параметров штрафа и множителями Лагранжа для регулярной точки минимума (см. разд. 3.4):

$$\lambda_j(r^k) = r^k g_j[x^*(r^k)], \quad j = 1, \dots, m;$$

$$\lambda_j(r^k) = r^k g_j^+[x^*(r^k)], \quad j \in J_a;$$

$$\lim_{r^k \rightarrow \infty} \lambda_j(r^k) = \lambda_j^*, \quad j = 1, \dots, m; \quad j \in J_a.$$

4. Обычно выбирается $r^0 = 0,01; 0,1; 1$, а $C \in [4, 10]$. Иногда начинают с $r^0 = 0$, т.е. с задачи поиска безусловного минимума.

5. При решении задач процедура расчетов завершается при некотором конечном значении параметра штрафа r^k . При этом приближенное решение, как правило, не лежит в множестве допустимых решений, т.е. ограничения задачи не выполняются. Это является одним из недостатков метода. С ростом параметра штрафа r^k генерируемые алгоритмом точки приближаются к решению исходной задачи извне множества допустимых решений. Поэтому обсуждаемый метод иногда называют *методом внешних штрафов*.

6. На практике для получения решения исходной задачи с требуемой точностью достаточно бывает решить конечное (относительно небольшое) число вспомогательных задач. При этом нет необходимости решать их точно, а информацию, полученную в результате решения очередной вспомогательной задачи, обычно удается эффективно использовать для решения следующей.

Пример. Найти минимум в задаче

$$\begin{aligned} f(x) &= x^2 - 4x \rightarrow \min, \\ g_1(x) &= x - 1 \leq 0. \end{aligned}$$

□ 1. В поставленной задаче $m = 0$ (ограничения-равенства отсутствуют), $p = 1$. Решим задачу аналитически при произвольном параметре штрафа r^k , а затем получим решение последовательности задач поиска безусловного минимума.

2. Составим вспомогательную функцию:

$$F(x, r^k) = x^2 - 4x + \frac{r^k}{2} [\max\{0, (x - 1)\}]^2.$$

3. Найдем безусловный минимум функции $F(x, r^k)$ по x с помощью необходимых и достаточных условий (см. § 2):

$$\frac{\partial F(x, r^k)}{\partial x} = \begin{cases} 2x - 4 = 0, & x - 1 \leq 0, \\ 2x - 4 + r^k(x - 1) = 0, & x - 1 > 0. \end{cases}$$

Отсюда $x^* = 2$, но при этом не удовлетворяется условие $x^* - 1 \leq 0$, а также

$$x^*(r^k) = \frac{4 + r^k}{2 + r^k}.$$

В табл. 9.1 приведены результаты расчетов при $r^k = 1, 2, 10, 100, 1000, \infty$, а на рис. 9.2 дана графическая иллюстрация.

Таблица 9.1

k	r^k	$x^*(r^k)$	$F(x^*(r^k), r^k)$	$\lambda(r^k) = r^k g^+(x^*(r^k))$
0	1	$\frac{5}{3}$	-3,66	$\frac{2}{3}$
1	2	$\frac{3}{2} = 1,5$	-3,5	1
2	10	$\frac{7}{6} = 1,1666$	-3,166	$\frac{10}{6}$
3	100	$\frac{52}{51} = 1,0196$	-3,019	$\frac{100}{51}$
4	1000	$\frac{502}{501} = 1,00199$	-3,002	$\frac{1000}{501}$
5	∞	1	-3	2

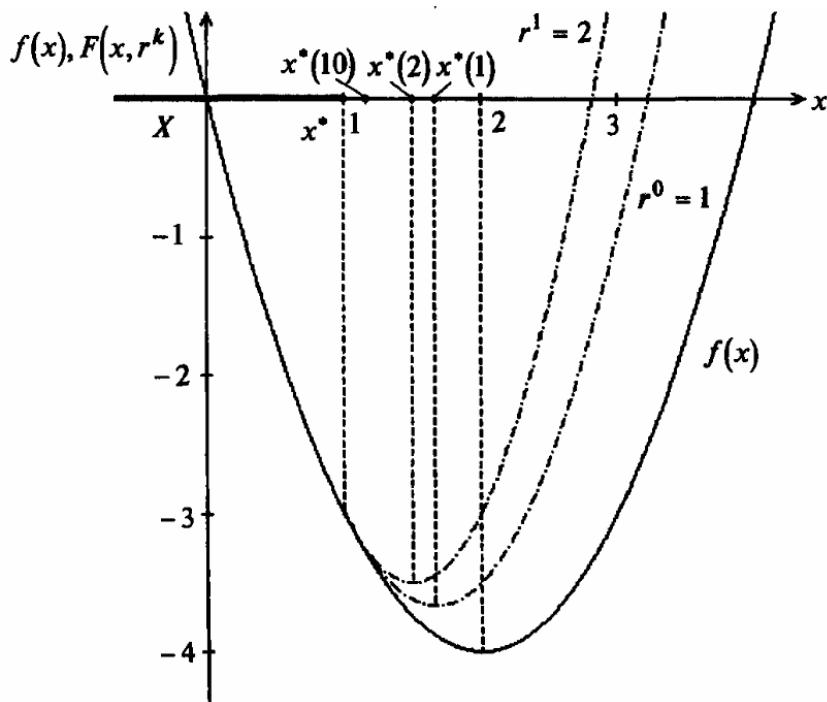


Рис. 9.2

Так как $\frac{\partial^2 F(x^*(r^k), r^k)}{\partial x^2} = 2 + r^k > 0$ при $r^k \geq 0$, то достаточные условия минимума $F(x, r^k)$ удовлетворяются. При $r^k \rightarrow \infty$ имеем

$$x^* = \lim_{r^k \rightarrow \infty} \frac{4 + r^k}{2 + r^k} = 1, \quad f(x^*) = -3.$$

Основные понятия языка Java и основы работы в NetBeans изложены в приложении А, параграф 16.

Задание

Найти минимум функции $f(x)$ в задаче условной оптимизации методом штрафов.

Определить, сколько итераций потребуется для достижения точности $\varepsilon = 10^{-3}, 10^{-5}, 10^{-7}$ при одних и тех же начальных приближениях.

При реализации программы использовать шаблон Bridge для отделения класса «Функция» от класса «Штрафная функция».

Варианты:

1. $f(x) = x_1 + x_2^2 + 3 \rightarrow \min,$
 $g_1(x) = x_1 - x_2 = 2.$
2. $f(x) = 2x_1^2 + x_2 + 3x_1 \rightarrow \min,$
 $g_1(x) = x_1 + 2x_2 \leq 1.$
3. $f(x) = 1/x_1 - x_2^2 + 3x_2 \rightarrow \min,$
 $g_1(x) = x_1 - x_2 - 1 = 0.$

4. $f(x) = x_1^2 + x_2^2 + 7x_2 \rightarrow \min,$
 $g_1(x) = x_1 - x_2 - 4 \leq 0.$
- $f(x) = -5x_1 - 2x_2^2 + 5x_1 \rightarrow \min,$
5. $g_1(x) = x_1 - x_2 \leq 2,$
 $g_2(x) = x_2 - 1 \leq 0.$
6. $f(x) = x_1^3 + x_2^4 + x_2 \rightarrow \min,$
 $g_1(x) = 3x_1 - x_2 = 0.$
7. $f(x) = 4x_1^3 + x_2^3 + 7x_1 \rightarrow \min,$
 $g_1(x) = x_1 + 3x_2 = 0.$
8. $f(x) = -6x_1^3 + 2x_2^4 + 2x_1 \rightarrow \min,$
 $g_1(x) = x_1 + 10x_2 - 1 = 0.$
- $f(x) = 2/x_1 - 4x_2^2 - 6x_1 \rightarrow \min,$
9. $g_1(x) = 2x_1 - 2x_2 - 1 = 0,$
 $g_2(x) = x_1 - x_2 \leq 0.$
10. $f(x) = 5/x_1^3 - 5/x_2^2 + x_2 \rightarrow \min,$
 $g_1(x) = x_1 - 5x_2 \leq 0.$
11. $f(x) = 7x_1^2 - 7x_2^2 + 7 \rightarrow \min,$
 $g_1(x) = 2x_1 - 3x_2 = 0.$
- $f(x) = 6/x_1^2 + 8x_2^2 + 10/x_2 \rightarrow \min,$
12. $g_1(x) = x_1 - 4x_2 \leq 0,$
 $g_2(x) = x_1 - 1 \leq 0.$
13. $f(x) = x_1 + x_2^3 + x_1x_2 \rightarrow \min,$
 $g_1(x) = 3x_1 - x_2 \leq 2.$
14. $f(x) = 2x_1^3 + 2/x_2^2 - x_2 \rightarrow \min,$
 $g_1(x) = x_1 - 10x_2 \leq 1.$
15. $f(x) = 8/x_1^2 + 8x_2^2 + 8x_2 \rightarrow \min,$
 $g_1(x) = x_1 - 2x_2 \leq 1,$
 $g_2(x) = 5 - 2x_2 \leq 0.$

Содержание отчета

- 1 Тема и цель работы.
- 2 Диаграмма классов и текст программы.
- 3 Результаты выполнения программы.

Контрольные вопросы

- 1 Сущность метода штрафов.
- 2 Основные изученные понятия языка Java (приложение А, параграф 16).

Лабораторная работа № 8.
Методология решения неструктурированных проблем.
Метод экспертного ранжирования.
Методы парных и последовательных сопоставлений

2 часа

Цель работы:

- 1) Изучить метод экспертного ранжирования, методы парных сравнений, последовательных сопоставлений;
- 2) изучить файлы и потоки ввода-вывода в Java.

Метод экспертного ранжирования

Пусть имеется m экспертов $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_m$ и n целей $Z_1, Z_2, \dots, Z_n$.

Каждый эксперт проводит оценку целей, пользуясь 10-бальной шкалой, причем оценки могут быть как целыми, так и дробными. В этих условиях веса целей определяются следующим образом:

1. Составляется матрица оценок экспертов:

$\mathcal{E}_j/Z_i$	Z_1	Z_2	...	Z_n
$\mathcal{E}_1$	S_{11}	S_{12}	...	S_{1n}
$\mathcal{E}_2$	S_{21}	S_{22}	...	S_{2n}
...	...	...	...	...
$\mathcal{E}_m$	S_{m1}	S_{m2}	...	S_{mn}

2. $0 \leq p_{ji} \leq 10$ ($j = 1, m, i = 1, n$)
3. Составляется матрица нормированных оценок:

$$\omega = p_{ji} / \sum p_{ji} \quad (j = 1, m, i = 1, n)$$

4. Вычисляются искомые веса целей:

$$\omega_i = \sum \omega_{ij} / \sum \sum \omega_{ij} \quad (i = 1, n) \quad \sum \omega_i = 1.$$

Пример:

Найдем веса целей для случая $m = 2$ и $n = 6$

1. Матрица оценок экспертов:

$\mathcal{E}_j/Z_i$	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
$\mathcal{E}_1$	10	7	9	3	4	5
$\mathcal{E}_2$	8	6	10	4	2	7

2. Матрица нормированных оценок:

$\mathcal{E}_j/Z_i$	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
$\mathcal{E}_1$	10/38	7/38	9/38	3/38	4/38	5/38
$\mathcal{E}_2$	8/38	6/38	10/38	4/38	2/38	7/38

3. Искомые веса целей:

$$\omega_1 = (10/38 + 8/38)/2 = 0,239;$$

$$\omega_2 = \dots = 0,173;$$

$$\omega_3 = \dots = 0,254;$$

$$\omega_4 = \dots = 0,093;$$

$$\omega_5 = \dots = 0,079;$$

$$\omega_6 = \dots = 0,16.$$

Метод парных сравнений

Эксперт проводит оценку целей Z_1, Z_2, Z_n .

Эксперт осуществляет парные сравнения целей во всех возможных сочетаниях. В каждой паре выделяется наиболее предпочтительная цель. И это предпочтение выражается с помощью оценки по какой-либо шкале.

Одна из возможных модификаций метода состоит в следующем:

1. составляется матрица бинарных предпочтений, в которой предпочтение целей выражается с помощью булевых переменных;
2. определяется цена каждой цели путем суммирования булевых переменных по соответствующей строке матрицы.
3. Обработка матрицы оценок позволяет найти веса целей, характеризующие их относительную важность.

Пример:

Эксперт проводит оценку 4-х целей, которые связаны с решением транспортной проблемы.

Z_1 – построить метрополитен

Z_2 – приобрести 2-хэтажный автобус

Z_3 – расширить транспортную сеть

Z_4 – ввести скоростной трамвай

1. Составим матрицу бинарных предпочтений:

Z_i / Z_j	Z_1	Z_2	Z_3	Z_4
Z_1		1	1	1
Z_2	0		0	0
Z_3	0	1		1
Z_4	0	1	0	

2. Определим цену каждой цели (складываем по строкам)

$$C_1=3; C_2=0; C_3=2; C_4=1.$$

Эти числа уже характеризуют важность объектов. Нормируем, т.к. этими числами неудобно пользоваться.

3. Нормированные веса целей:

$$v_1 = 3/6 = 0,5; v_2 = 0; v_3 = 0,33; v_4 = 0,17$$

Проверка: сумма v_i должна равняться 1.

Получаем следовательно порядок предпочтения целей:

$$Z_1, Z_3, Z_4, Z_2$$

Метод последовательных сравнений

Одна из возможных модификаций метода состоит в следующем:

1. Все цели располагаются в виде массива в порядке убывания их важности и назначаются предварительные оценки целей. При этом первая цель массива получает оценку 100, а остальным целям ставятся в соответствие оценки, отражающие их важность.
2. Первая цель массива сравнивается со всеми возможными комбинациями ниже стоящих целей по 2. В случае необходимости оценка первой цели корректируется. Вторая цель массива сравнивается со всеми возможными комбинациями ниже стоящих целей по 2. В случае необходимости оценка 2-ой цели корректируется и т.д.
3. Производится запись скорректированных оценок и расчет на их основе весов целей.

Пример:

Эксперт проводит оценку 4-х целей, которые связаны с решением транспортной проблемы (см. пример выше).

Расположим цели в виде массива и назначим предварительные оценки Z_1, Z_3, Z_4, Z_2 .

Выставляем баллы: $p_1 = 100, p_3 = 60, p_4 = 40, p_2 = 10$

Выполним сравнение целей и корректировку их оценок

$Z_1(Z_3 \otimes Z_4)$

$Z_1(Z_3 \otimes Z_2)$

$Z_1(Z_4 \otimes Z_2)$

$Z_3(Z_4 \otimes Z_2)$

(т.е. цель Z_1 сравниваем с комбинацией Z_3 и Z_4)...

Возможно, что построить метрополитен лучше, чем 3 и 4, но 3+4 дают 100, поэтому корректируем оценки: $p_1 = 125; p_3 = 60$;

Запишем скорректированные оценки и вычислим веса целей:

$p_1 = 125; p_3 = 60; p_4 = 40; p_2 = 10$;

$v_1 = 125/\text{сумма всех оценок} = 0,54; v_3 = 0,25; v_4 = 0,17; v_2 = 0,04$

Проверка: сумма v_i должна равняться 1.

Получаем порядок предпочтения целей: Z_1, Z_3, Z_4, Z_2

Основные понятия языка Java и основы работы в NetBeans изложены в приложении А, параграф 17.

Варианты заданий:

Примечание: К задаче необходимо составить матрицу попарного сравнения альтернатив.

Задачи, имеющие четные номера, решаются в NetBeans с применением метода экспертного ранжирования и метода парных сопоставлений.

Задачи, имеющие нечетные номера, решаются в NetBeans с применением метода экспертного ранжирования и метода последовательных сопоставлений.

Реализовать сохранение матриц оценок экспертов в файл и чтение из файла.

Задача 1. Группа экспертов из 3-х человек проводит оценку четырех вариантов по строительству торгового центра:

- 1 достроить одноэтажное неиспользуемое помещение в центральном районе города;
- 2 построить новый супермаркет, требующий крупных капиталовложений, с выгодным расположением;
- 3 построить супермаркет за чертой города, с небольшими затратами;
- 4 построить торговый центр на окраине города, район оснащен развитой транспортной сетью и паркингом.

Найти оптимальный вариант проекта.

Задача 2. Компания устраивает рекламную акцию по поводу открытия нового магазина. Группа экспертов собрана для выбора наилучшей программы для привлечения потенциальных покупателей:

- 1 пригласить популярную поп-группу;
- 2 предложить скидку каждому покупателю;
- 3 раздавать призы;
- 4 устроить лотерею.

Задача 3. Анализ результатов экономической деятельности предприятия показал его неспособность функционировать на рынке. Пригласили 4-х экспертов для помощи руководству принять решение о выходе из сложившейся ситуации. Рассматриваются следующие варианты:

- 1 Ликвидировать предприятие
- 2 Выставить на продажу
- 3 Объявить банкротом
- 4 Провести санацию

Выяснить оптимальный путь дальнейшего развития предприятия.

Задача 4. Для решения проблемы, связанной с невозможностью предоставления жилья иногородним студентам была создана группа экспертов из 4-х человек. Предложено несколько альтернатив:

- 1 Построить новое общежитие
- 2 Снять многоквартирный дом и частично оплачивать жилье
- 3 Назначить доплату незаселенным студентам.

Определить наилучшую альтернативу.

Задача 5. Правительство приняло решение выделить денежные средства из бюджета наиболее важному социальному объекту. Для выделения самого приоритетного была создана комиссия из 5-и экспертов и рассмотрены предложенные варианты:

- 1 Разбить парк отдыха
- 2 Построить теннисные корты
- 3 Построить телебашню
- 4 Реконструировать центральную площадь
- 5 Построить библиотеку

Определить наилучшую альтернативу.

Задача 6. Для распространения информации об услугах и продукции лаборатории и дополнительной рекламы группа из 3-х экспертов группа экспертов из отдела маркетинга должны оценить наиболее эффективный вариант мероприятия:

- 1 заказ мультимедиа презентации;
- 2 выпуск периодического печатного издания;
- 3 выпуск ежегодных каталогов услуг;
- 4 создание корпоративного сайта.

Найти наиболее эффективный и выгодный вариант рекламной акции.

Задача 7. Для продвижения товаров и услуг на рынке холдингу необходимо провести дополнительные рекламные мероприятия. Эксперт из отдела сбыта проводит анализ четырех вариантов решения этого вопроса :

- 1 Создание интернет-магазина;
- 2 Введение круглосуточного режима работы, увеличение кадров;
- 3 Открытие еще одного филиала;
- 4 Усилить рекламу в СМИ;

Определить наиболее информативный способ расширения и рекламы.

Задача 8. В результате эффективного использования иностранных инвестиций и грамотной политики предприятие получило значительную прибыль. Руководитель и инвесторы утвердили эксперта для решения проблемы выбора объекта, которому будут выделены средства на развитие. Эксперту предложены следующие цели:

- 1 Строительство физкультурного клуба для сотрудников на территории предприятия;
- 2 Заказ проекта корпоративного сайта;
- 3 Инвестирование крупного строительного проекта.

Определить наилучшую альтернативу

Задача 9. В результате успешной деятельности банка и востребования его услуг руководство стоит перед проблемой организации дальнейшего бесперебойного предос-

тавления услуг населению, расширения, привлечения новых клиентов. Для этого эксперту поручено определить наиболее удачный вариант решения вопроса:

- 1 Открытие дополнительного филиала в городе;
 - 2 Приобретение здания необходимого размера для перемещения банка и его расширения;
 - 3 Введение круглосуточного режима работы, увеличение кадров.
- Определить наилучшую альтернативу

Задача 10. Из республиканского и местных бюджетов выделены средства в сферу здравоохранения эксперт проводит оценку наиболее нуждающейся и важной сферы медицины для получения субсидий.

- 1 Замена всего оборудования, отработавшего нормативный срок, на новое;
 - 2 Установка дорогостоящего современного оборудования в специализированных центрах и диспансерах;
 - 3 Открытие поликлиник в густонаселенных микрорайонах;
 - 4 Строительство наркологического центра.
- Определить наиболее важную цель

Задача 11. Компания «Перспект» хочет получить максимальную прибыль. Для этого руководство пригласило 3 экспертов, для выбора наилучшей альтернативы из предложенных:

- 1 открытие собственного производства;
 - 2 увеличение затрат на рекламу;
 - 3 расширение рынка сбыта;
 - 4 снижение цен с целью увеличения оборота.
- Определить наилучшую альтернативу

Задача 12. За перевыполнение плана руководство компании хочет наградить работников. Для этого эксперту поручено определить наиболее удачный вариант решения вопроса:

- 1 выдать разовую прибыль;
 - 2 устроить корпоративную вечеринку;
 - 3 дать оплачиваемый отпуск;
 - 4 увеличить зарплату.
- Определить наилучшую альтернативу.

Задача 13. За отличную учебу родители решили поощрить своих детей. Для этого они пригласили 4 экспертов для выбора наилучшего варианта:

- 1 увеличение карманных денег;
 - 2 путевку в санаторий;
 - 3 разрешить гулять до 23:00
- Определить наилучшую альтернативу.

Задача 14. Руководство университета решило поспособствовать культурному обогащению учащихся.

Для этого руководство пригласило 3 экспертов, для выбора наилучшей альтернативы из предложенных:

- 1 бесплатные билеты в театр;
 - 2 бесплатные билеты на выставку;
 - 3 бесплатные билеты в кино;
- Определить наилучшую альтернативу.

Задача 15. На основе отчетных данных деятельности предприятия руководство признало необходимость дополнительной рекламы товаров. Рекламный отдел предложил не-

сколько вариантов рекламных мероприятий, различающихся как по стоимости так и по эффективности. Эксперт отдела проводит оценку предложенных целей:

- 1 Реклама в ведущих печатных изданиях республики и в радиоэфире;
 - 2 Рекламный ролик выпускаемой продукции на ведущем телеканале страны в дорогое эфирное время;
 - 3 Реклама в интернете (почтовые рассылки, баннерная реклама, электронные публикации);
 - 4 Установка выставочных стендов в главных торговых центрах крупных городов на длительный срок;
 - 5 Спонсировать какое-либо значимое общественное мероприятие (например, выступления известной команды КВН, выставку «Моторшоу», парк детских аттракционов).
- Определить наилучшую альтернативу.

Содержание отчета

- 1 Тема и цель работы.
- 2 Диаграмма классов и текст программы с комментариями.
- 3 Результаты выполнения программы.

Контрольные вопросы

- 1 Классификация и общая характеристика методов экспертных оценок.
- 2 Сущность метода экспертного ранжирования.
- 3 Сущность методов парных и последовательных сопоставлений.
- 4 Основные изученные понятия языка Java (приложение А, параграф 17).

Лабораторная работа № 9.

Методология решения неструктурированных проблем.

Ранжирование проектов методом парных сравнений

2 часа

Цель работы:

- 1) Освоить ранжирование проектов методом парных сравнений;
- 2) изучить способы оценки согласованности между ранжировками экспертов;
- 3) изучить шаблон Factory в Java.

Сущность метода:

Пусть имеется m экспертов $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_m$ и n проектов $P_1, P_2, \dots, P_n$, подлежащих оценке. Рассмотрим метод экспертных оценок, позволяющий ранжировать проекты по их важности.

1. Эксперты осуществляют попарное сравнение проектов, оценивая их важность в долях единицы.
2. Находятся оценки C_i , характеризующие предпочтение одного из проектов над всеми прочими проектами, для чего складываются все оценки по каждому проекту. Находится сумма оценок всех проектов $\sum C_i$ (по столбцу)
3. Вычисляются нормированные веса проектов
4. Полученные веса позволяют ранжировать проекты по их важности.

Пример.

1. Четыре эксперта осуществляют попарное сравнение проектов, оценивая их важность в долях единицы (таблица 9.1). Для определенности будем считать, что 4 эксперта оценивают важность 4-х проектов P_1, P_2, P_3, P_4 .

Таблица 9.1

$\{\mathcal{E}_j\}$	$\pi_1 - \pi_2$		$\pi_1 - \pi_3$		$\pi_1 - \pi_4$		$\pi_2 - \pi_3$		$\pi_2 - \pi_4$		$\pi_3 - \pi_4$	
$\mathcal{E}_1$	0,4	0,6	0,65	0,35	0,5	0,5	0,6	0,4	0,7	0,3	0,6	0,4
$\mathcal{E}_2$	0,3	0,7	0,55	0,45	0,6	0,4	0,7	0,3	0,6	0,4	0,6	0,4
$\mathcal{E}_3$	0,4	0,6	0,5	0,5	0,7	0,3	0,6	0,4	0,6	0,4	0,5	0,5
$\mathcal{E}_4$	0,5	0,5	0,5	0,5	0,6	0,4	0,5	0,5	0,7	0,3	0,7	0,3
Σ	1,6	<u>2,4</u>	2,2	<u>1,8</u>	2,4	<u>1,6</u>	<u>2,4</u>	<u>1,6</u>	<u>2,6</u>	<u>1,4</u>	<u>2,4</u>	<u>1,6</u>

2. Находятся оценки, характеризующие предпочтение одного из проектов над всеми прочими проектами

Складываются все оценки по первому проекту: $f^{(1)} = 1,6 + 2,2 + 2,4 = 6,2$

Складываются все оценки по второму проекту: $f^{(2)} = 2,4 + 2,4 + 2,6 = 7,4$

Складываются все оценки по третьему проекту: $f^{(3)} = 1,8 + 1,6 + 2,4 = 5,8$

Складываются все оценки по четвертому проекту: $f^{(4)} = 1,6 + 1,4 + 1,6 = 4,6$

3. Вычисляются нормированные веса проектов:

$$\omega_1 = 0,26; \omega_2 = 0,31; \omega_3 = 0,24; \omega_4 = 0,19$$

4. Полученные веса позволяют ранжировать проекты по их важности.

P_2, P_1, P_3, P_4 – результат решения.

Метод парных сравнений очень прост и он позволяет исследовать большее количество объектов (по сравнению, например, с методом рангов) и с большей точностью.

Согласованность между ранжировками двух экспертов можно определить с помощью *коэффициента ранговой корреляции Спирмэна*:

$$\rho = 1 - \frac{6 \sum_{i=1}^n (x_{ij} - x_{ik})^2}{n(n^2 - 1)} = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)},$$

где x_{ij} – ранг, присвоенный i -му объекту j -ым экспертом;

x_{ik} – ранг, присвоенный i -му объекту k -ым экспертом;

d_i – разница между рангами, присвоенными i -му объекту;

n – количество факторов (проектов).

Величина ρ может изменяться в диапазоне от -1 до $+1$. При полном совпадении оценок коэффициент равен единице. Равенство коэффициента минус единице наблюдается при наибольшем расхождении в мнениях экспертов.

Кроме того, расчет коэффициента ранговой корреляции может применяться как способ оценки взаимоотношений между каким-либо фактором и результативным признаком (реакцией) в тех случаях, когда признаки не могут быть измерены точно, но могут быть упорядочены.

В этом случае значение коэффициента Спирмэна может быть интерпретировано подобно значению коэффициента парной корреляции. Положительное значение свидетельствует о прямой связи между факторами, отрицательное – об обратной, при этом, чем ближе абсолютное значение коэффициента к единице, тем теснее связь.

Когда необходимо определить согласованность в ранжировках большого (более двух) числа экспертов, рассчитывается так называемый **коэффициент конкордации (коэффициент Кендалла)** – общий коэффициент ранговой корреляции для группы, состоящей из m экспертов:

$$W = \frac{12 \cdot S}{m^2(n^2 - n)},$$

где

$$S = \sum_{i=1}^n \left(\sum_{j=1}^m x_{ij} - \frac{1}{2} m(n+1) \right)^2$$

Заметим, что вычитаемое в скобках представляет собой не что иное, как среднюю сумму рангов (при суммировании для каждого объекта), полученных i объектами от экспертов.

Коэффициент W изменяется в диапазоне от 0 до 1. Его равенство единице означает, что все эксперты присвоили объектам одинаковые ранги. Чем ближе значение коэффициента к нулю, тем менее согласованными являются оценки экспертов. Если коэффициент конкордации больше 0,5, то обычно согласованность считается достаточной.

Основные понятия языка Java и основы работы в NetBeans изложены в приложении А, параграф 18.

Задания

Примечание: К задаче, решаемой в NetBeans, необходимо составить матрицу попарного сравнения альтернатив. Эксперты оценивают выгодность проектов в долях единицы. Необходимо оценить согласованность между ранжировками экспертов с помощью *коэффициента ранговой корреляции Спирмэна* либо с помощью *коэффициента конкордации (коэффициента Кендалла)*.

Реализовать шаблон Factory для создания строк матрицы попарного сравнения альтернатив.

Задача 1. Крупная компания по ремонту автомобилей решает расширить свою деятельность посредством легального импорта автомобилей. Для этого необходимо определить социальную группу, для которой их поставлять и, следовательно, цены и марки автомобилей. Для этого проводится маркетинговое исследование населения, результаты которого оценивают четыре эксперта:

Z_1 – импортировать дорогие и редкие «заокеанские» автомобили для обеспеченных клиентов (2010-2011 гг выпуска, дорогие запчасти)

Z_2 – импортировать дорогие европейские марки (более дешевые запчасти)

Z_3 – организовать доставку, ориентируясь на среднюю стоимость (2003-2008 гг)

Z_4 – закупать доступные автомобили, б/у (1995-2002 гг)

Определить наиболее предпочтительный вариант

Задача 2. Компания по производству бытовой техники принимает решение расширить производство, для чего необходимы некоторые денежные средства. Чтобы грамотнее выбрать источник финансирования финансовый директор приглашает группу экспертов из 5-ти человек, которые оценивают ниже перечисленные варианты:

Z_1 – привлечь инвесторов

Z_2 – взять кредит в банке/ у финансового консультанта

Z_3 – создать совместное предприятие

Z_4 – выпустить коммерческое предприятие

Определить наиболее предпочтительный вариант

Задача 3. Издательство «ЭКМО» принимает решение о пополнение своего ассортимента за счет нового журнала/газеты. Был проведен анализ существующих изданий, в итоге появились некоторые новые идеи, которые были представлены для оценки группе из 2-х экспертов

Z_1 – разносторонний политический обозреватель

Z_2 – «страны и континенты», туризм, путешествие

Z_3 – экстремальные виды спорта

Z_4 – новинки в сфере искусства (кино, музыка, живопись)

Z_5 – трудоустройство

Определить наиболее предпочтительный вариант

Задача 4. Руководство компании мобильной связи МТС, желая увеличить количество абонентов, объявила конкурс среди сотрудников на лучшую акцию по достижению поставленной цели. Наиболее перспективные проекты были вынесены на суд 2-х экспертов:

Z_1 – бесплатные разговоры внутри сети

Z_2 – подарки каждому новому(3-му, 4-му...) абоненту

Z_3 – льготы тем, кто привел друзей

Z_4 – телефоны в рассрочку

Z_5 – каждому новому абоненту – 60 минут звонков на город бесплатно

Определить наиболее предпочтительный вариант

Задача 5. Городское управление решило организовать мероприятие по озеленению центра города. Для этого собирается группа экспертов в составе 4-х человек для выбора наилучшего проекта из предложенных:

Z_1 – разбить клумбы с цветами;

Z_2 – посадить деревья;

Z_3 – поставить искусственные деревья;

Z_4 – повесить на столбы кашпо с цветами.

Определить наиболее предпочтительный вариант

Задача 6. Руководство университета выделило деньги на техническое переоснащение аудиторий. Для этого администрация пригласила группу экспертов из 5-ти человек, которые оценивают ниже перечисленные варианты:

Z_1 – купить новые компьютеры;

Z_2 – купить новое оборудование в лингвистический кабинет;

Z_3 – оснастить аудитории новыми телевизорами;

Определить наиболее предпочтительный вариант

Задача 7. Руководству университета выделили деньги на обеспечение иногородних студентов жильем. Для чего разработали несколько проектов решения данной проблемы и предложили их комиссии из 4-х экспертов:

Z_1 – построить новое общежитие;

Z_2 – сделать пристройку к общежитию;

Z_3 – арендовать многоэтажный дом.

Определить наиболее предпочтительный вариант.

Задача 8. Компании требуются кадры в новый отдел. Чтобы грамотнее выбрать будущих работников директор приглашает группу экспертов из 3-х человек, которые оценивают ниже перечисленные варианты:

Z_1 – переманить лучших сотрудников конкурентов;

Z_2 – отправить сотрудников других отделов на курсы повышения квалификации;

Z_3 – провести конкурс среди студентов, заканчивающих ВУЗ.

Определить наиболее предпочтительный вариант.

Задача 9. Документооборот предприятия увеличивается с каждым днем. Для более эффективной работы предприятия решили внедрить систему электронного документооборота. Для выбора самой подходящей системы создали группу экспертов из 5 человек.

Предстоит выбрать:

Z_1 – российская система Евфрат-документооборот;

Z_2 – европейская Lotus Notes;

Z_3 – создание новой системы, приспособленной к данному предприятию;

Определить наиболее предпочтительный вариант.

Задача 10. Собрана группа экспертов в составе 2-х человек для выбора объекта инвестирования. Были предложены проекты:

1 Минский тракторный завод

2 Минский завод холодильников «Атлант»

3 Кондитерская фабрика «Витьба»

4 Минская швейная фабрика «Элема»

5 Минский молочный завод «ГМЗ №1»

Определить наиболее перспективный объект инвестирования.

Задача 11. На основе отчетных данных руководство предприятия признало необходимость дополнительной рекламы товаров.

Три эксперта рекламного отдела проводят оценку предложенных вариантов:

1 Реклама в печатных изданиях республики и в радиоэфире;

2 Рекламный ролик выпускаемой продукции;

3 Реклама в интернете;

4 Реклама в радиоэфире.

Определить наиболее предпочтительный вариант.

Задача 12. Управление текстильного предприятия ОАО «Камволь» безуспешно пыталось выйти из долговой «ямы» и предприятию в будущем грозит банкротство. При-

знана необходимость радикальных мер для решения денежного вопроса. Выбрана группа экспертов (4 человека) для выбора самого лучшего варианта погашения долгов:

- 1 Выставить несколько крупных ведомственных объектов на аукцион
- 2 Банковское кредитование, которое позволило бы начать хозяйственную деятельность с прибылью, но не достаточное для погашения огромных долгов.
- 3 Получение разрешения Министерства финансов отсрочить погашение части долгов.

Задача 13. К 60-летию освобождения РБ предложен ряд мероприятий для ветеранов ВОВ и труда. Группа экспертов (5 человек) собрана для выбора наилучшего и наиболее необходимого нововведения.

- 1 Предоставить участникам ВОВ бесплатные авиабилеты в экономическом классе, что даст возможность посетить места боевой славы, встретиться с друзьями.
- 2 Открыть сеть магазинов «Ветеран» со сниженной на 10% торговой надбавкой
- 3 Обеспечить льготную продажу в кредит, организовать обслуживание с предварительным заказом и доставкой на дом

Задача 14. Компания «Строим вместе» решила увеличить выпуск продукции. Для этого 4 экспертам поручено определить наиболее удачный вариант решения проблемы:

- 1 ввести новую поточную линию;
- 2 увеличить количество рабочих;
- 3 построить новый завод.

Задача 15. Иностранная фирма инвестировала деньги белорусскому предприятию. Руководство предприятия собрало группу из 3-х экспертов для выбора наилучшего варианта:

- 1 вложить деньги в рекламу продукции;
- 2 закупить новую технику;
- 3 отремонтировать износившееся оборудование;
- 4 повысить зарплату руководящему составу.

Содержание отчета

- 1 Тема и цель работы.
- 2 Диаграмма классов и текст программы.
- 3 Результаты выполнения программы.

Контрольные вопросы

- 1 Сущность ранжирования проектов методом парных сравнений.
- 2 Когда применяется коэффициент ранговой корреляции Спирмэна? А коэффициент конкордации?
- 3 Для чего нужно оценивать согласованность между ранжировками экспертов?
- 4 Основные изученные понятия языка Java (приложение А, параграф 18).

Лабораторная работа № 10.

Принципы решения слабо структурированных проблем.

Кластерный анализ

4 часа

Цель работы:

- 1) Изучить основные понятия кластерного анализа;
- 2) изучить метод К средних, метод максимина;
- 3) изучить шаблон Expert в Java.

Методы кластерного анализа предназначены для разделения множества анализируемых объектов и явлений на кластеры, т.е. на группы объектов, схожих друг с другом по каким-либо признакам.

Пример. Анализируется информация о девяти инвестиционных фондах (Ф1, Ф2,...,Ф9). Показатели, характеризующие деятельность фондов, приведены в табл.10.1.

Таблица 10.1 – Информация о девяти инвестиционных фондах (Ф1, Ф2,...,Ф9)

Фонд	Ф1	Ф2	Ф3	Ф4	Ф5	Ф6	Ф7	Ф8	Ф9
Прибыль за анализируемый период, тыс. ден.ед.	16476	17081	13827	13187	11793	16728	10386	15145	15596
Экспертная оценка риска, баллы	4	4	5	1	3	4	2	7	4

Примечание. Оценки риска заданы экспертом по десятибалльной шкале: 1 – минимальный риск, 10 – максимальный.

Требуется выделить группы фондов, имеющих сходные значения показателей.

В данной задаче имеется девять объектов (М=9). Для их описания используется два признака (N=2). Решение этой задачи различными методами рассматривается ниже.

Подготовка данных для кластерного анализа**1. Меры различия**

Как отмечено выше, задача кластерного анализа состоит в разделении множества анализируемых объектов на группы объектов, сходных друг с другом по каким-либо признакам. При этом необходимо учитывать, что каждый объект, как правило, описывается несколькими признаками. Эти признаки обычно различаются по размерности (измеряются в разных единицах) и по диапазону значений (одни признаки выражаются большими числами, другие – малыми). Некоторые признаки могут указываться в виде балльных оценок (например, по 10- или 100-балльной шкале). В некоторых случаях объекты описываются качественными (словесными) признаками: для описания объектов используются оценки “отлично”, “хорошо”, “часто”, “редко” и т.д. Такое разнообразие оценок затрудняет сопоставление объектов и делает невозможным получение оценки различия между объектами в виде одного числа. Поэтому, прежде чем применять какие-либо методы кластерного анализа, необходимо выполнить нормирование признаков объектов. В результате нормировки все значения признаков объектов должны быть безразмерными (т.е. не должны измеряться в каких-либо единицах) и находиться в некотором ограниченном диапазоне (например, от нуля до единицы). Существует несколько методов нормирования. Обычно применяются следующие методы:

1) **деление на максимальное значение:** значения признака для всех объектов делятся на максимальное значение этого признака. Результатом являются безразмерные величины, находящиеся в диапазоне от нуля до единицы;

2) **стандартизация:** из каждого значения признака вычитается среднее значение данного признака, полученная разность делится на стандартное отклонение данного признака. Результатом являются безразмерные величины, большинство из которых принимает значения в диапазоне от -3 до 3 .

При небольшом количестве анализируемых объектов обычно применяется деление на максимальное значение, при большом количестве объектов – стандартизация.

Если для описания объектов используются качественные (словесные) оценки, то следует перейти от таких оценок к числовым величинам (например, балльным экспертным оценкам), а затем выполнить нормировку на основе одного из рассмотренных способов.

Рассмотрим пример, приведенный выше (анализ информации об инвестиционных фондах). В этом примере объекты описываются двумя признаками. Первый из этих признаков (прибыль) измеряется в денежных единицах, второй (оценка риска) – в баллах. Кроме того, первый признак может принимать большие значения (десятки тысяч), а второй – выражается числами от 1 до 10. На основе таких оценок невозможно получить какую-либо величину, характеризующую различие между объектами. Поэтому требуется нормировка значений признаков.

1) Выполним нормирование, используя **деление на максимальное значение**. Для признака “прибыль” максимальное значение равно 17 081, для признака “оценка риска” – 7 (см. табл.10.2). Для нормирования разделим каждое значение признака на соответствующее максимальное значение. Результаты приведены в табл.10.2.

Таблица 10.2 – Нормирование с использованием **деления на максимальное значение**

Фонд	Ф1	Ф2	Ф3	Ф4	Ф5	Ф6	Ф7	Ф8	Ф9
Прибыль за анализируемый период	0,96	1,00	0,81	0,77	0,69	0,98	0,61	0,89	0,91
Экспертная оценка риска	0,57	0,57	0,71	0,14	0,43	0,57	0,29	1,00	0,57

2) Выполним нормирование на основе **стандартизации**. Для этого необходимо сначала найти среднее значение и стандартное отклонение каждого признака. Обозначим средние значения признаков как $\bar{A}_i$, а стандартные отклонения – как σ_i , $i=1, \dots, M$ (где M – количество признаков), N – количество объектов. Эти величины находятся по следующим формулам:

$$\bar{A}_i = \frac{1}{N} \sum_{j=1}^N X_{ij}, \quad (10.1)$$

$$\sigma_i = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (X_{ij} - \bar{A}_i)^2}. \quad (10.2)$$

Для данного примера $\bar{A}_1=14468,78$; $\bar{A}_2=3,78$; $\sigma_1=2333,68$; $\sigma_2=1,72$.

Как указано выше, нормирование выполняется следующим образом: из каждого значения признака (табл.10.2) вычитается среднее значение данного признака, полученная разность делится на стандартное отклонение. Результаты нормирования приведены в табл.10.3.

Таблица 10.3 – Нормирование на основе стандартизации

Фонд	Ф1	Ф2	Ф3	Ф4	Ф5	Ф6	Ф7	Ф8	Ф9
Прибыль за анализируемый период	0,86	1,12	-0,28	-0,55	-1,15	0,97	-1,75	0,29	0,48
Экспертная оценка риска	0,13	0,13	0,71	-1,62	-0,45	0,13	-1,04	1,88	0,13

Здесь, например, значение признака “прибыль за анализируемый период” для фонда Ф1 получено следующим образом: $(16476-14468,78)/2333,68=0,86$.

Во всех последующих расчетах будут использоваться только нормированные значения признаков.

Чтобы принять решение о том, можно ли считать некоторые объекты достаточно сходными и отнести их к одному кластеру, необходимо использовать некоторую числовую меру различия между объектами. Обычно в качестве такой меры различия используется *евклидово расстояние*. Значение евклидова расстояния между некоторыми объектами X_j и X_k определяется по следующей формуле:

$$D(X_j, X_k) = \sqrt{\sum_{i=1}^M (X_{ij} - X_{ik})^2}. \quad (10.3)$$

Найдем, например, евклидово расстояние между объектами X_1 и X_2 из рассматриваемого примера, т.е. меру различия между фондами Ф1 и Ф2. Для расчета евклидова расстояния будем использовать нормированные значения признаков, полученные путем деления на максимальное значение (табл.10.3):

$$D(X_1, X_2) = \sqrt{(0,96 - 1)^2 + (0,57 - 0,57)^2} = 0,04.$$

Если признаки различны по важности (т.е. различия по одним признакам необходимо учитывать в большей степени, по другим – в меньшей), то в качестве меры различия используется взвешенное евклидово расстояние:

$$D(X_j, X_k) = \sqrt{\sum_{i=1}^M W_i \cdot (X_{ij} - X_{ik})^2}, \quad (10.4)$$

где $W_i, i=1, \dots, M$ – веса признаков (чем важнее признак, тем больше его вес). Они могут определяться, например, на основе методов экспертного анализа. Обычно используются значения весов, удовлетворяющие следующему условию: $W_1 + W_2 + \dots + W_M = 1$.

В некоторых случаях применяются также следующие меры различия:

- **расстояние city-block (манхэттенское):**

$$D(X_j, X_k) = \sum_{i=1}^M |X_{ij} - X_{ik}|; \quad (10.5)$$

Это расстояние рассчитывается как среднее разностей по координатам. В большинстве случаев эта мера расстояния приводит к результатам, подобным расчетам расстояния евклида. Однако, для этой меры влияние отдельных выбросов меньше, чем при использовании *евклидова расстояния*, поскольку здесь координаты не возводятся в квадрат.

- **чебышевское расстояние:**

$$D(X_j, X_k) = \max_i |X_{ij} - X_{ik}|. \quad (10.6)$$

Это расстояние стоит использовать, когда необходимо определить два объекта как “различные”, если они отличаются по какому-то одному измерению.

Смысл всех мер различия следующий: чем больше различаются значения признаков, описывающих объекты, тем большее значение принимают меры различия. Объекты с небольшими значениями мер различия должны относиться к одному кластеру, с большими – к разным.

2. Разделение объектов на заданное число кластеров

Разделение объектов на заданное число кластеров может производиться с использованием **метода К средних, метода максимина.**

Метод К средних (K-Means Cluster)

Метод предназначен для разделения объектов на заданное число кластеров.

Принцип работы метода следующий. На основе имеющейся информации о предметной области **задается количество кластеров (К)**. При этом указывается также содержательный смысл каждого кластера (например, объекты с высоким значением некоторого признака, объекты со средним значением некоторых признаков и т.д.). Для каждого кластера выбирается объект-*прототип* (представитель), т.е. объект, наиболее подходящий для данного класса по значениям признаков. Находится первоначальный вариант разделения объектов на кластеры: каждый объект относится к кластеру, представляемому *ближайшим* объектом-прототипом. Затем в каждом кластере находится новый прототип со средними (для данного кластера) значениями признаков. Снова выполняется отнесение каждого объекта к кластеру, представляемому ближайшим прототипом. Процедура повторяется до получения окончательного разбиения, т.е. до тех пор, пока на двух последовательных итерациях метода будет получено одинаковое разбиение.

Приведем пошаговый **алгоритм** разбиения объектов на заданное число кластеров на основе метода К средних.

1. Номер итерации алгоритма принимается равным нулю: $s = 0$.
2. Задается количество кластеров (К). Для каждого кластера выбирается первоначальный объект-прототип: P_k^0 , $k=1, \dots, K$.
3. Выполняется переход к очередной итерации алгоритма: $s = s + 1$.
4. Находятся расстояния от каждого из анализируемых объектов до каждого из объектов-прототипов. Выполняется отнесение каждого объекта к ближайшему кластеру, т.е. к кластеру, для которого расстояние между этим объектом и прототипом кластера минимально.
5. В каждом кластере определяется новый объект-прототип: P_k^s , $k = 1, \dots, K$. Значение каждого признака этого объекта-прототипа определяется как среднее арифметическое значений этого признака для всех объектов, входящих на текущей итерации в данный кластер.
6. Если объекты-прототипы всех кластеров на данной и на предыдущей итерации совпадают (т.е. выполняется условие $P_k^s = P_k^{s-1}$, $k=1, \dots, K$), то алгоритм завершается. Если на данной итерации получено разбиение объектов, отличное от предыдущего, то выполняется возврат к шагу 3.

Рассмотрим реализацию приведенного алгоритма на примере (анализ информации об инвестиционных фондах). Во всех расчетах будут использоваться нормированные значения признаков объектов, полученные путем деления на максимальное значение признака (табл.10.3).

Пусть предполагается, что инвестиционные фонды можно в основном разделить на три группы: 1) фонды с низким риском и низкой прибылью; 2) фонды с высоким риском и высокой прибылью; 3) фонды со средними значениями обоих показателей. Таким образом, предполагается, что число кластеров равно трем ($K=3$). Пошаговый алгоритм разбиения объектов на заданное число кластеров на основе метода К средних выполняется в следующем порядке.

1. Номер итерации принимается равным нулю: $s = 0$.
2. Задается количество кластеров: $K=3$. Для каждого кластера требуется выбрать прототип. Пусть на основе анализа показателей (табл.10.2) эксперт указал, что наиболее характерным представителем первого кластера (фонды с низким риском и низкой прибылью) является фонд, обозначенный как Ф4. В качестве характерного представителя второго кластера (фонды с высоким риском и высокой прибылью) экспертом указан фонд Ф8, а

в качестве представителя третьего кластера (фонды со средними значениями обоих показателей) – фонд Ф9. Таким образом, $P_1^0 = X_4 = (0,77; 0,14)$, $P_2^0 = X_8 = (0,89; 1)$, $P_3^0 = X_9 = (0,91; 0,57)$.

3. Выполняется **переход к очередной итерации: $s = 1$** .

4. Находятся расстояния от каждого из анализируемых объектов до каждого из прототипов по формуле (10.3). Эти расстояния приведены в табл.10.4.

Например, расстояние между объектом X_1 и прототипами кластеров найдено следующим образом:

$$D(X_1, P_1^0) = \sqrt{(0,96 - 0,77)^2 + (0,57 - 0,14)^2} = 0,47,$$

$$D(X_1, P_2^0) = \sqrt{(0,96 - 0,89)^2 + (0,57 - 1)^2} = 0,44,$$

$$D(X_1, P_3^0) = \sqrt{(0,96 - 0,91)^2 + (0,57 - 0,57)^2} = 0,05.$$

По найденным расстояниям выполняется отнесение каждого объекта к кластеру, представленному ближайшим прототипом. Например, для объекта X_1 расстояние до прототипа P_1^0 составляет 0,47, до прототипа P_2^0 - 0,44, до P_3^0 - 0,05. Таким образом, ближайшим к объекту X_1 оказался прототип третьего кластера, поэтому X_1 относится к этому кластеру. Результаты деления объектов на кластеры приведены в табл.10.4 (последняя строка).

Таблица 10.4 – Результаты деления объектов на кластеры методом К средних

Объект	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
P_1^0	0,47	0,49	0,57	0,00	0,30	0,48	0,22	0,87	0,45
P_2^0	0,44	0,44	0,30	0,87	0,60	0,44	0,76	0,00	0,43
P_3^0	0,05	0,09	0,17	0,45	0,26	0,07	0,41	0,43	0,00
Кластер	3	3	3	1	3	3	1	2	3

Таким образом, получено следующее разбиение объектов на кластеры: к первому кластеру относятся объекты X_4 и X_7 , ко второму – только объект X_8 , к третьему – объекты $X_1, X_2, X_3, X_5, X_6, X_9$.

5. В каждом кластере определяется новый объект-прототип. Значение каждого признака этого объекта-прототипа определяется как среднее арифметическое значений этого признака для всех объектов, входящих на текущей итерации в данный кластер. Для данного примера значения признаков объекта-прототипа первого кластера находятся как средние арифметические значений признаков объектов X_4 и X_7 :

$$P_1^1 = \left(\frac{0,77 + 0,61}{2}; \frac{0,14 + 0,29}{2} \right) = (0,69; 0,22).$$

Во втором кластере находится только один объект (X_8), поэтому он и становится прототипом этого класса: $P_2^1 = (0,89; 1,00)$.

Значения признаков объекта-прототипа третьего кластера находятся как средние арифметические значений признаков объектов $X_1, X_2, X_3, X_5, X_6, X_9$:

$$P_3^1 = \left(\frac{0,96 + 1 + 0,81 + 0,69 + 0,98 + 0,91}{6}; \frac{0,57 + 0,57 + 0,71 + 0,43 + 0,57 + 0,57}{6} \right) = (0,89; 0,57)$$

6. Выполняется сравнение прототипов, полученных на данной и на предыдущей итерациях. На данной итерации получены прототипы $P_1^1 = (0,69; 0,22)$, $P_2^1 = (0,89; 1,00)$, $P_3^1 = (0,89; 0,57)$. На предыдущей итерации прототипы были следующими: $P_1^0 = (0,77; 0,14)$, $P_2^0 = (0,89; 1)$, $P_3^0 = (0,91; 0,57)$. Прототипы не совпадают, поэтому требуется следующая итерация. Выполняется возврат к шагу 3.

3. Выполняется **переход к очередной итерации: $s = 2$** .

4. Находятся расстояния от каждого из анализируемых объектов до каждого из прототипов по формуле (10.3). По найденным расстояниям выполняется отнесение каждого объекта к кластеру, представленному ближайшим прототипом. Расстояния, а также результаты разделения объектов на кластеры приведены в табл.10.5.

Например, расстояние между объектом X_1 и прототипами кластеров найдено следующим образом:

$$D(X_1, P_1^1) = \sqrt{(0,96 - 0,69)^2 + (0,57 - 0,22)^2} = 0,45,$$

$$D(X_1, P_2^1) = \sqrt{(0,96 - 0,89)^2 + (0,57 - 1)^2} = 0,44,$$

$$D(X_1, P_3^1) = \sqrt{(0,96 - 0,89)^2 + (0,57 - 0,57)^2} = 0,07.$$

Таким образом, объект X_1 относится к третьему кластеру.

Таблица 10.5 – Результаты разделения объектов на кластеры

Объект	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
P_1^1	0,45	0,47	0,51	0,11	0,21	0,46	0,11	0,81	0,42
P_2^1	0,44	0,44	0,30	0,87	0,60	0,44	0,76	0,00	0,43
P_3^1	0,07	0,11	0,16	0,45	0,25	0,09	0,40	0,43	0,02
Кластер	3	3	3	1	1	3	1	2	3

Таким образом, получено следующее разбиение объектов на кластеры: к первому кластеру относятся объекты X_4 , X_5 и X_7 , ко второму – только объект X_8 , к третьему – объекты X_1 , X_2 , X_3 , X_6 , X_9 .

5. В каждом кластере определяется новый объект-прототип:

$$P_1^2 = \left(\frac{0,77 + 0,69 + 0,61}{3}; \frac{0,14 + 0,43 + 0,29}{3} \right) = (0,69; 0,29),$$

$$P_2^2 = (0,89; 1,00),$$

$$P_3^2 = \left(\frac{0,96 + 1 + 0,81 + 0,98 + 0,91}{5}; \frac{0,57 + 0,57 + 0,71 + 0,57 + 0,57}{5} \right) = (0,93; 0,60).$$

6. Выполняется сравнение прототипов, полученных на данной и на предыдущей итерациях. На данной итерации получены прототипы $P_1^2 = (0,69; 0,29)$, $P_2^2 = (0,89; 1,00)$, $P_3^2 = (0,93; 0,63)$, на предыдущей – прототипы $P_1^1 = (0,69; 0,22)$, $P_2^1 = (0,89; 1,00)$, $P_3^1 = (0,89; 0,57)$. Таким образом, прототипы не совпадают. Требуется следующая итерация. Выполняется возврат к шагу 3.

3. Выполняется **переход к очередной итерации: $s = 3$** .

4. Находятся расстояния от каждого из анализируемых объектов до каждого из прототипов по формуле (10.3), и выполняется отнесение каждого объекта к кластеру, представленному ближайшим прототипом. Результаты приведены в табл.10.6.

Таблица 10.6 – Результаты разделения объектов на кластеры

Объект	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉
P ₁ ²	0,39	0,42	0,44	0,17	0,14	0,41	0,08	0,74	0,36
P ₂ ²	0,44	0,44	0,30	0,87	0,60	0,44	0,76	0,00	0,43
P ₃ ²	0,04	0,07	0,17	0,49	0,29	0,06	0,45	0,40	0,04
Кластер	3	3	3	1	1	3	1	2	3

Таким образом, получено следующее разбиение объектов на кластеры: к первому кластеру относятся объекты X₄, X₅ и X₇, ко второму – только объект X₈, к третьему – объекты X₁, X₂, X₃, X₆, X₉.

5. В каждом кластере определяется новый объект-прототип (как и на предыдущих итерациях): $P_1^3 = (0,69; 0,29)$, $P_2^3 = (0,89; 1,00)$, $P_3^3 = (0,93; 0,60)$.

6. Выполняется сравнение прототипов, полученных на данной и на предыдущей итерациях. Прототипы совпадают: $P_1^3 = P_1^2$, $P_2^3 = P_2^2$, $P_3^3 = P_3^2$. Это означает, что получено окончательное разбиение объектов на кластеры.

Примечание. В качестве признака окончания алгоритма (т.е. окончательного разбиения) можно использовать совпадение разбиения на двух последовательных итерациях.

Таким образом, результаты разделения инвестиционных фондов на группы (кластеры) оказались следующими. К первой группе (фонды с низким риском и низкой прибылью) относятся фонды Ф4, Ф5, Ф7. Ко второй группе (фонды с высоким риском и высокой прибылью) можно отнести только фонд Ф8. В третью группу (фонды со средними значениями обоих показателей) входят фонды Ф1, Ф2, Ф3, Ф6, Ф9.

Достоинства метода К средних:

- простота использования;
- быстрота использования;
- понятность и прозрачность алгоритма.

Недостатки метода К средних:

• алгоритм слишком чувствителен к выбросам, которые могут искажать среднее. Возможным решением этой проблемы является использование модификации алгоритма – алгоритм k-медианы (Алгоритм РАМ (partitioning around Medoids));

• алгоритм может медленно работать на больших базах данных. Возможным решением данной проблемы является использование выборки данных.

Метод максимина

Метод предназначен для разделения объектов на кластеры, причем количество кластеров **заранее неизвестно**; оно определяется автоматически в процессе разбиения объектов.

Принцип работы метода следующий. Выбирается один из объектов (любой); он становится прототипом первого кластера. Находится объект, наиболее удаленный от выбранного; он становится прототипом второго кластера. Все объекты распределяются по двум кластерам; каждый объект относится к кластеру, представленному ближайшим прототипом. Затем в каждом из кластеров находится объект, *наиболее удаленный* от своего прототипа. Если расстояние между этим объектом и прототипом кластера оказывается значительным (превышающим некоторую предельную величину), то объект становится новым прототипом, т.е. образуется новый кластер. После этого распределение объектов по кластерам выполняется заново. Процесс продолжается, пока не будет получено такое разбиение на кластеры, при котором расстояние от каждого объекта до прототипа кластера не будет превышать заданную предельную величину.

Приведем пошаговый **алгоритм** реализации метода максимина.

1. Выбирается любой из объектов, например, первый в списке объектов (X_1). Он становится прототипом первого кластера: $P_1 = X_1$. Количество кластеров принимается равным единице: $K=1$.

2. Определяются расстояния от объекта P_1 до всех остальных объектов: $D(P_1, X_j)$, $j=1, \dots, N$. Определяется объект, наиболее удаленный от P_1 , т.е. объект X_f , для которого выполняется условие: $D(P_1, X_f) = \max_j D(P_1, X_j)$. Этот объект становится прототипом

второго кластера: $P_2 = X_f$. Количество кластеров принимается равным двум: $K=2$.

3. Определяется пороговое расстояние. Оно принимается равным *половине* расстояния между прототипами P_1 и P_2 : $T = D(P_1, P_2) / 2$. Эта величина будет использоваться для проверки условия окончания алгоритма.

4. Находятся расстояния от каждого из анализируемых объектов до каждого из имеющихся объектов-прототипов. Выполняется отнесение каждого объекта к ближайшему кластеру, т.е. кластеру, для которого расстояние между этим объектом и прототипом кластера минимально.

5. В каждом кластере определяется объект, наиболее удаленный от прототипа своего кластера. Обозначим эти объекты как Y_k , $k=1, \dots, K$ (здесь k – номер кластера, K – количество кластеров).

6. Для каждого из наиболее удаленных объектов, найденных на шаге 5, проверяется условие: $D(P_k, Y_k) < T$, $k=1, \dots, K$. Если это условие выполняется *для всех кластеров*, то алгоритм *завершается*. Если для некоторого объекта Y_k это условие не выполняется, то он становится прототипом нового кластера, и количество кластеров увеличивается на единицу ($K=K+1$). В результате этого шага количество кластеров K увеличивается на число, равное количеству новых кластеров.

7. Находится новое пороговое расстояние. Оно определяется как половина среднего арифметического всех расстояний между прототипами:

$$T = \frac{\sum_{i=1}^{K-1} \sum_{j=i+1}^K D(P_i, P_j)}{K \cdot (K-1)}. \quad (10.7)$$

8. Выполняется возврат к шагу 4.

Таким образом, окончательным является разбиение, для которого во всех кластерах расстояние от прототипа кластера до каждого из объектов, входящих в этот кластер (даже до самого удаленного), не превышает некоторой предельной величины (порогового расстояния).

Рассмотрим реализацию приведенного алгоритма на примере (анализ информации об инвестиционных фондах).

1. Первый из объектов принимается в качестве прототипа первого кластера: $P_1 = X_1 = (0,96; 0,57)$. Количество кластеров принимается равным единице: $K = 1$.

2. Определяются расстояния от прототипа P_1 до всех остальных объектов по формуле (10.3). Эти расстояния приведены в табл.10.7.

Таблица 10.7 – Расстояния от прототипа P_1 до всех остальных объектов

Объект	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
$P_1 = X_1$	0,00	0,04	0,21	0,47	0,30	0,02	0,45	0,44	0,05

Здесь, например, расстояние между прототипом P_1 и объектом X_2 найдено следующим образом: $D(P_1, X_2) = \sqrt{(0,96 - 1)^2 + (0,57 - 0,57)^2} = 0,04$.

Из таблицы 10.7 видно, что наиболее удаленным от прототипа P_1 является объект X_4 . Он становится прототипом второго кластера: $P_2 = X_4 = (0,77; 0,14)$. Количество кластеров становится равным двум: $K=2$.

3. Определяется пороговое расстояние: $T = D(P_1, P_2) / 2 = 0,235$.

4. Находятся расстояния от каждого из анализируемых объектов до каждого из имеющихся объектов-прототипов. По найденным расстояниям выполняется отнесение каждого объекта к кластеру, представленному ближайшим прототипом. Расстояния, а также результаты разбиения объектов на кластеры приведены в табл.10.8.

Таблица 10.8 – Результаты разбиения объектов на кластеры методом максимина

Объект	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
$P_1 = X_1$	0,00	0,04	0,21	0,47	0,30	0,02	0,45	0,44	0,05
$P_2 = X_4$	0,47	0,49	0,57	0,00	0,30	0,48	0,22	0,87	0,45
Кластер	1	1	1	2	2	1	2	1	1

Примечание. Если вычислить расстояния от объекта X_5 до прототипов P_1 и P_2 с большей точностью (до трех знаков), то они окажутся следующими: $D(P_1, X_5) = 0,304$, $D(P_2, X_5) = 0,301$. Поэтому объект X_5 отнесен ко второму кластеру.

5. В каждом кластере определяется объект, наиболее удаленный от прототипа *своего* кластера. Из таблицы 10.8 видно, что в первом кластере таким объектом является X_8 , во втором – X_5 . Таким образом, $Y_1 = X_8$, $Y_2 = X_5$.

6. Расстояния между наиболее удаленными объектами (найденными на шаге 5) и объектами-прототипами сравниваются с пороговым расстоянием. В данном примере наиболее удаленным объектом первого кластера оказался объект X_8 . Расстояние между этим объектом и прототипом P_1 равно 0,44; оно превышает пороговое расстояние (0,235). Таким образом, можно считать, что объект X_8 существенно отличается от прототипа своего кластера. Поэтому он становится прототипом нового кластера (P_3), и количество кластеров увеличивается на единицу ($K=3$). Во втором кластере наиболее удаленным объектом является X_5 . Расстояние между этим объектом и прототипом P_2 равно 0,3; оно превышает пороговое расстояние. Поэтому объект X_5 также становится прототипом нового кластера (P_4), и количество классов увеличивается еще на единицу ($K=4$).

7. Находится новое пороговое расстояние. Оно определяется как половина среднего арифметического всех расстояний между прототипами. В данном случае имеется четыре прототипа (X_1 , X_4 , X_8 , X_5). Требуется найти расстояния между ними: $D(X_1, X_4) = 0,47$, $D(X_1, X_8) = 0,44$, $D(X_1, X_5) = 0,3$, $D(X_4, X_8) = 0,87$, $D(X_4, X_5) = 0,3$, $D(X_8, X_5) = 0,6$. Новое пороговое расстояние находится по формуле (10.7) следующим образом:

$$T = \frac{0,47 + 0,44 + 0,3 + 0,87 + 0,3 + 0,6}{4 \cdot 3} = 0,25.$$

8. Выполняется возврат к шагу 4.

4. Находятся расстояния от каждого из анализируемых объектов до каждого из имеющихся объектов-прототипов. По найденным расстояниям выполняется отнесение каждого объекта к кластеру, представленному ближайшим прототипом. Результаты приведены в табл.10.9.

Таблица 10.9 – Результаты отнесения каждого объекта к кластеру, представленному ближайшим прототипом

Объект	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉
P ₁ = X ₁	0,00	0,04	0,21	0,47	0,30	0,02	0,45	0,44	0,05
P ₂ = X ₄	0,47	0,49	0,57	0,00	0,30	0,48	0,22	0,87	0,45
P ₃ = X ₈	0,44	0,44	0,30	0,87	0,60	0,44	0,76	0,00	0,43
P ₄ = X ₅	0,30	0,34	0,30	0,30	0,00	0,32	0,16	0,60	0,26
Кластер	1	1	1	2	4	1	4	3	1

5. В каждом кластере определяется объект, наиболее удаленный от прототипа *своего* кластера. Из табл. 10.9 видно, что в первом кластере таким объектом является X₃, в четвертом – X₇. Во второй и третий кластеры входит по одному объекту (X₄ и X₈ соответственно). Таким образом, Y₁=X₃, Y₂=X₄, Y₃=X₈, Y₄=X₇.

6. Расстояния между наиболее удаленными объектами (найденными на шаге 5) и объектами-прототипами сравниваются с пороговым расстоянием. В данном примере D(P₁,Y₁)=0,21, D(P₂,Y₂)=0, D(P₃,Y₃)=0, D(P₄,Y₄)=0,16. Все эти расстояния меньше порогового (T=0,25). Таким образом, можно считать, что во всех кластерах объекты достаточно близки к своим прототипам, т.е. имеют сходные значения признаков. Алгоритм завершается.

Таким образом, результаты разделения инвестиционных фондов на группы (кластеры) оказались следующими. К первому кластеру относятся фонды Ф1, Ф2, Ф3, Ф6, Ф9, ко второму – фонд Ф4, к третьему – Ф8, к четвертому - Ф5 и Ф7. Интерпретация этих результатов возможна только на основе анализа, выполняемого специалистами в соответствующей предметной области (в данном случае – экспертами-экономистами). Проанализировав показатели фондов, приведенные в табл.10.2, можно предложить следующую интерпретацию полученного разбиения. Первый кластер включает фонды со средней и высокой прибылью и со средними оценками риска. Второй кластер, включающий только фонд Ф4, соответствует средней прибыли и самому низкому риску. Третий кластер (фонд Ф8) соответствует достаточно высокой прибыли и самому высокому риску. Четвертый кластер включает фонды с низкой прибылью и невысоким (ниже среднего) риском.

Основные понятия языка Java и основы работы в NetBeans изложены в приложении А, параграф 19.

Задания

Варианты заданий 1-15 по кластерному анализу даны в табл. 1, а значения показателей производственно-хозяйственной деятельности предприятий машиностроения приведены в табл. 2.

Задание выполняется в NetBeans с использованием метода К средних и метода максимина с использованием шаблона Expert.

Рассматриваются следующие показатели:

- Y₁ - производительность труда;
- Y₂ - индекс снижения себестоимости продукции;
- Y₃ - рентабельность;
- X₄ - трудоемкость единицы продукции;
- X₅ - удельный вес рабочих в составе ППП;
- X₆ - удельный вес покупных изделий;

X7- коэффициент сменности оборудования;
 X8- премии и вознаграждения на одного работника;
 X9 - удельный вес потерь от брака;
 X10 - фондоотдача;
 X11 - среднегодовая численность ППП;
 X12 - среднегодовая стоимость ОПФ;
 X13 - среднегодовой фонд заработной платы ППП;
 X14- фондовооруженность труда;
 X15- оборачиваемость нормируемых оборотных средств;
 X16 - оборачиваемость ненормируемых оборотных средств;
 X17 - непроизводственные расходы.

Таблица 1 – **Варианты заданий 1-15 по кластерному анализу**

№ варианта	Результативный признак, Y	Номера факторных признаков, X
1	1	6, 8, 12, 17
2	1	8, 11, 13, 17
3	1	6, 8, 14, 17
4	1	8, 11, 14, 17
5	1	7, 11, 12, 13
6	1	8, 9, 13, 14
7	1	5, 6, 7, 9
8	1	5, 9, 11, 17
9	3	8, 10, 15, 16
10	3	5, 10, 15, 17
11	3	5, 7, 11, 12
12	3	8, 9, 10, 17
13	2	4, 5, 6, 9
14	2	4, 6, 8, 9
15	2	4, 5, 8, 17

Таблица 2 – **Таблица исходных данных**

№ пред-при-ятия	Y1	Y2	Y3	X4	X5	X6	X7	X8	X9	X10
1	9,26	204,2	13,26	0,23	0,78	0,40	1,37	1,23	0,23	1,45
2	9,38	209,6	10,16	0,24	0,75	0,26	1,49	1,04	0,39	1,30
3	12,11	222,6	13,72	0,19	0,68	0,40	1,44	1,80	0,43	1,37
4	10,81	236,7	12,85	0,17	0,70	0,50	1,42	0,43	0,18	1,65
5	9,35	62,0	10,63	0,23	0,62	0,40	1,35	0,88	0,15	1,91
6	9,87	53,1	9,12	0,43	0,76	0,19	1,39	0,57	0,34	1,68
7	8,17	172,1	25,83	0,31	0,73	0,25	1,16	1,72	0,38	1,94
8	9,12	56,5	23,39	0,26	0,71	0,44	1,27	1,70	0,09	1,89
9	5,88	52,6	14,68	0,49	0,69	0,17	1,16	0,84	0,14	1,94
10	6,30	46,6	10,05	0,36	0,73	0,39	1,25	0,60	0,21	2,06
11	6,22	53,2	13,99	0,37	0,68	0,33	1,13	0,82	0,42	1,96
12	5,49	30,1	9,68	0,43	0,74	0,25	1,10	0,84	0,05	1,02
13	6,50	146,4	10,03	0,35	0,66	0,32	1,15	0,67	0,29	1,85
14	6,61	18,1	9,13	0,38	0,72	0,02	1,23	1,04	0,48	0,88
15	4,32	13,6	5,37	0,42	0,68	0,06	1,39	0,66	0,41	0,62

Продолжение таблицы 2.2

№ пред- приятия	X11	X12	X13	X14	X15	X16	X17
1	26006	167,69	47750	6,40	166,32	10,08	17,72
2	23935	186,10	50391	7,80	92,88	14,76	18,39
3	22589	220,45	43149	9,76	158,04	6,48	26,46
4	21220	169,30	41089	7,90	93,96	21,96	22,37
5	7394	39,53	14257	5,35	173,88	11,88	28,13
6	11586	40,41	22661	9,90	162,30	12,60	17,55
7	26609	102,96	52509	4,50	88,56	11,52	21,92
8	7801	37,02	14903	4,88	101,16	8,28	19,52
9	11587	45,74	25587	3,46	166,32	11,52	23,99
10	9475	40,07	16821	3,60	140,76	32,40	21,76
11	10811	45,44	19459	3,56	128,52	11,52	25,68
12	6371	41,08	12973	5,65	177,84	17,28	18,13
13	26761	136,14	50907	4,28	114,48	16,20	25,74
14	4210	42,39	6920	8,85	93,24	13,32	21,21
15	3557	37,39	5736	8,52	126,72	17,28	22,97

Содержание отчета

- 1 Тема и цель работы.
- 2 Диаграмма классов и текст программы.
- 3 Результаты выполнения программы.

Контрольные вопросы

- 1 Для чего предназначена кластеризация? Характеристики кластера.
- 2 Методы нормирования признаков объектов.
- 3 Какие бывают меры различия между объектами?
- 4 Сущность метода К средних и метода максимина.
- 5 Основные изученные понятия языка Java (приложение А, параграф 19).

Список рекомендуемой литературы

- 1 **Антонов, А. В.** Системный анализ. Учеб. для вузов/А.В. Антонов. - М.: Высш. шк., 2004. - 454 с: ил.
- 2 **Орлов, А. И.** Теория принятия решений : Учебное пособие / А.И.Орлов. – М. : Издательство «Экзамен», 2005. – 656 с.
- 3 П.Ноутон, Г.Шилдт. JAVA 2 – 1050 с. [электронный вариант]
- 4 **Пантелеев, А. В.** Методы оптимизации в примерах и задачах : Учеб. пособие / А. В. Пантелеев, Т. А. Летова. – 2-е изд. – М.: Высш. шк., 2005. – 544 с: ил.
- 5 **Реклейтис, Г.** Оптимизация в технике / Реклейтис Г., Рейвиндран А., Рэгсдел К. – Т. 1 и 2, в 2-х т. – М. : Мир, 1986.
- 6 **Смородинский, С. С.** Методы анализа и принятия решений в слабоструктурированных задачах : Учеб. пособие по курсу “Методы и системы принятия решений”. Для студентов специальности “Автоматизированные системы обработки информации” / Смородинский С.С., Батин Н.В. – Минск : БГУИР, 2002. – 116 с.
- 7 **Тарасенко, Ф. П.** Прикладной системный анализ (Наука и искусство решения проблем): Учебник. - Томск: Изд-во Том. ун-та» 2004. - 186 с.
- 8 Теоретические основы системного анализа / Новосельцев В.И. [и др.]; под ред. В. И. Новосельцева. - М.: Майор, 2006. - 592 с. : ил.
- 9 **Черноруцкий, И. Г.** Методы оптимизации в теории управления: Учебное пособие / И. Г. Черноруцкий. – СПб.: Питер, 2004. — 256 с: ил.
- 10 **Черноруцкий, И. Г.** Методы принятия решений. – СПб.: БХВ-Петербург. 2005. – 416 с: ил.