

Лабораторная работа №1

Тема: хранилища данных и аналитическая обработка данных

Цель работы: получение практических навыков проектирования, разработки и использования хранилищ данных.

Задание: спроектируйте БД в многомерной модели представления данных используя модель звезды или снежинки (в реляционной базе) согласно полученному варианту (используя программу ERWin или MS Visio или их аналоги, модель должна включать не менее 5 сущностей), реализуйте спроектированную базу в СУБД MS SQL Server.

Внесите в базу тестовые данные (не менее 10 строк у каждую таблицу).

Реализуйте аналитические запросы к базе, используя следующие конструкции секционирование (partitioning), упорядочивание (order by), кадрирование (с использованием rows и range), аналитических функций сведения (pivot, unpivot), ранжирования функций (row_number, rank, dense_rank), получения значения строк (first_value, last_value, lead, lag), статистические (var, varp, stdevp, stdev).

Для справки по синтаксису используйте ресурс <http://technet.microsoft.com>.

Отчет по лабораторной работе должен содержать:

1. Фамилию и номер группы учащегося, задание
2. Описание многомерной модели (схема)
3. Физическую модель БД (sql-код)
4. Перечень тестовых данных (в виде таблиц)
5. Код запросов, задача (вопрос) для решения которых можно использовать полученные наборы данных (для каждого запроса), и результаты их выполнения (принтскрин с базы).

Варианты заданий

- 1) Погодные условия в регионе
- 2) Продажа комплектующих изделий
- 3) Демографическая ситуация в регионе
- 4) Продажа земельных участков
- 5) Рынок труда
- 6) Больница
- 7) Железнодорожный транспорт
- 8) Авиационные перевозки
- 9) Олимпиада
- 10) Футбол
- 11) Туристический бизнес
- 12) Социальные сети
- 13) Интернет-провайдер
- 14) Здравоохранение
- 15) Автострахование
- 16) Кредитование
- 17) Экология
- 18) Правонарушения
- 19) Литература
- 20) Компьютеры

Лабораторная работа №2

Тема: анализ данных.

Цель работы: Получение практических навыков анализа данных.

Задание: Используя программное средство Weka, выполните анализ данных согласно полученному варианту. Работа состоит из нескольких этапов:

- 1) Подготовка данных для анализа в полученной согласно варианту предметной области (атрибутов должно быть не менее 10, строк с данными не менее 100, строки должны быть уникальными)
- 2) Загрузка данных в систему, рассмотрение описания данных (максимальных, минимальных значений и т.д.)
- 3) Построение моделей различными методами:
 - Регрессионной,
 - Классификации
 - Кластеризации
 - Ассоциативной
- 4) Исследование моделей, их интерпретация и выводы о возможности их применения

Отчет по лабораторной работе должен содержать:

1. Фамилию и номер группы, задание
2. Описание данных
3. Описание процесса построения моделей
4. Описание полученного результата (с визуализацией)
5. Интерпретация полученных результатов и выводы
6. Листинги данных и моделей.

Основные теоретические сведения

Weka (Waikato Environment for Knowledge Analysis) — свободное программное обеспечение для анализа данных, написанное на Java в университете Уайкато (Новая Зеландия), распространяющееся по лицензии GNU GPL. Система представляет собой систему библиотек функции обработки данных, плюс несколько графических интерфейсов к этим библиотекам. Основной интерфейс системы - Explorer. Он позволяет выполнять практически все действия, которые предусмотрены в системе.

Также в системе Weka предусмотрены другие интерфейсы - Knowledge Flow для работы с большими массивами данных (Explorer загружает все данные в память сразу, и потому работа с большими массивами затруднена) и Experimenter для экспериментального подбора наилучшего метода анализа данных.

Weka предоставляет доступ к SQL-базам через Java Database Connectivity (JDBC) и в качестве исходных данных может принимать результат SQL-запроса. Возможность обработки множества связанных таблиц не поддерживается, но существуют утилиты для преобразования таких данных в одну таблицу, которую можно загрузить в *Weka*.

Explorer имеет несколько панелей.

- 1) Панель предобработки Preprocess panel позволяет импортировать данные из базы, CSV файла и т. д., и применять к ним алгоритмы фильтрации, например, переводить количественные признаки в дискретные, удалять объекты и признаки по заданному критерию.
- 2) Панель классификации Classify panel позволяет применять алгоритмы классификации и регрессии (в *Weka* они не различаются и называются classifiers) к выборке данных, оценивать предсказательную способность алгоритмов, визуализировать ошибочные предсказания, ROC-кривые, и сам алгоритм, если это возможно (в частности, решающие деревья).
- 3) Панель поиска ассоциативных правил Associate panel решает задачу выявления всех значимых взаимосвязей между признаками.
- 4) Панель кластеризации Cluster panel даёт доступ к алгоритму k-средних, EM-алгоритму для смеси гауссианов и другим.
- 5) Панель отбора признаков Select attributes panel даёт доступ к методам отбора признаков.
- 6) Панель визуализации Visualize строит матрицу графиков разброса (scatter plot matrix), позволяет выбирать и увеличивать графики, и т. д.

WEKA использует Java, так что если на вашем компьютере нет JRE, выберите для установки версию WEKA, включающую в себя JRE.



Рисунок 1 - Стартовое окно WEKA

При запуске WEKA, пакет предлагает вам на выбор 4 графических интерфейса для работы с WEKA и вашими данными. Выберите Explorer.

Анализ данных подразумевает наличие самих данных в системе. Для того чтобы загрузить данные в WEKA, их следует преобразовать в формат, понятный этому программному пакету. Наиболее подходящим форматом для загрузки данных в WEKA является формат Attribute-Relation File Format (ARFF), который сначала определяет тип загружаемых данных, а потом указывает собственно данные.

В файле формата ARFF вы указываете название и тип данных для каждого столбца таблицы, а затем собственно данные по строкам.

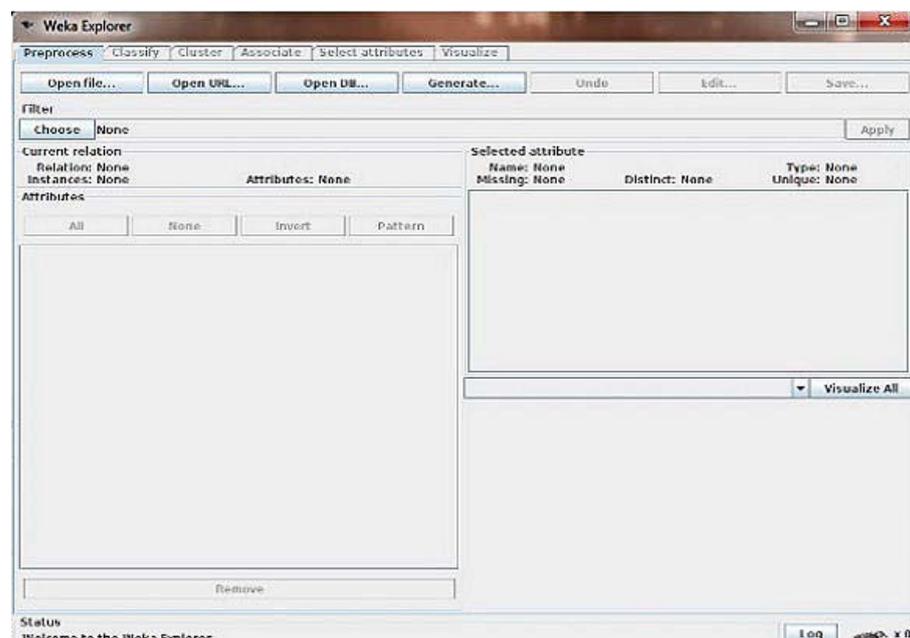


Рисунок 2 - Окно WEKA Explorer

Подготовка данных состоит в добавлении полей метаданных в начало файла. На отдельных строчках добавляется следующая информация:

-названия зависимости @relation имя,

- описания атрибутов @attribute: имя, тип
- @data перед началом самих данных.

Различают следующие типы данных:

- численные (numeric, real, integer),
- перечислимые(nominal) (задаются перечислением вида {i1, ..., in}),
- строковые (string),
- дата (date [date format]).
- составной тип (relational),

Пример использования различных моделей анализа данных в Weka

Регрессия

В моделях регрессионного анализа используются всего два типа данных: NUMERIC и DATE.

Метод регрессионного анализа является самым простым и, пожалуй, наименее эффективным методом интеллектуального анализа данных (удивительно, как часто эти качества сопутствуют друг другу). Самая простая модель анализа использует один входной (независимый) параметр и один результирующий (зависимый) параметр, модель можно усложнить, добавив несколько десятков входных параметров, но в любом случае общий подход будет один и тот же: на основании нескольких независимых переменных определяется один зависимый результат. Таким образом, модель регрессионного анализа используется для прогнозирования значения одной зависимой переменной, исходя из известных значений нескольких независимых параметров.

Наверняка, каждый из нас хотя бы раз сталкивался с регрессионной моделью, а может быть, и проводил в уме самостоятельный регрессионный анализ. Наиболее очевидный пример – определение стоимости дома. Цена на дом (зависимая переменная) определяется несколькими независимыми параметрами: какова площадь дома и размер участка, используется ли в оформлении кухни гранитные плиты, каково качество и срок службы сантехники и так далее. Так что, если вам случалось когда-либо продавать или покупать дом, то, скорее всего, вы использовали регрессионный анализ для определения его цены. Вы оценивали параметры похожих домов в этом же районе и цену, по которой эти дома были проданы (т.е. создавали модель), а затем подставляли параметры вашего дома в полученную зависимость и рассчитывали предполагаемую стоимость вашего дома.

Давайте воспользуемся моделью регрессионного анализа для определения цены дома и разберем конкретный пример. В таблице внизу указаны фактические параметры домов, выставленных на продажу в моем районе. На основании этих данных я попробую оценить стоимость моего дома (и воспользуюсь этими результатами, чтобы опротестовать

предъявленную мне сумму налога на недвижимость).

Таблица 1. Регрессионная модель оценки стоимости дома

Площадь дома (кв.футы)	Размер участка	Количество спален	Гранитная отделка на кухне	Современное сантехническое оборудование?	Продажная цена
3529	9191	6	0	0	\$205,000
3247	10061	5	1	1	\$224,900
4032	10150	5	0	1	\$197,900
2397	14156	4	1	0	\$189,900
2200	9600	4	0	1	\$195,000
3536	19994	6	1	1	\$325,000
2983	9365	5	0	1	\$230,000
3198	9669	5	1	1	????

Файл с данными (отношение \ таблица Дом, атрибуты размер дома, ванная и т.д. и собственно данные через запятую):

```
@RELATION house
@ATTRIBUTE houseSize NUMERIC
@ATTRIBUTE lotSize NUMERIC
@ATTRIBUTE bedrooms NUMERIC
@ATTRIBUTE granite NUMERIC
@ATTRIBUTE bathroom NUMERIC
@ATTRIBUTE sellingPrice NUMERIC
@DATA
3529,9191,6,0,0,205000
3247,10061,5,1,1,224900
4032,10150,5,0,1,197900
2397,14156,4,1,0,189900
2200,9600,4,0,1,195000
3536,19994,6,1,1,325000
2983,9365,5,0,1,230000
```

Запустите WEKA и выберите опцию Explorer. В результате откроется закладка Preprocess окна Explorer. Щелкните на кнопке Open File и выберите созданный вами ARFF-файл. Окно WEKA Explorer с загруженными данными о домах показано на рисунке 3.

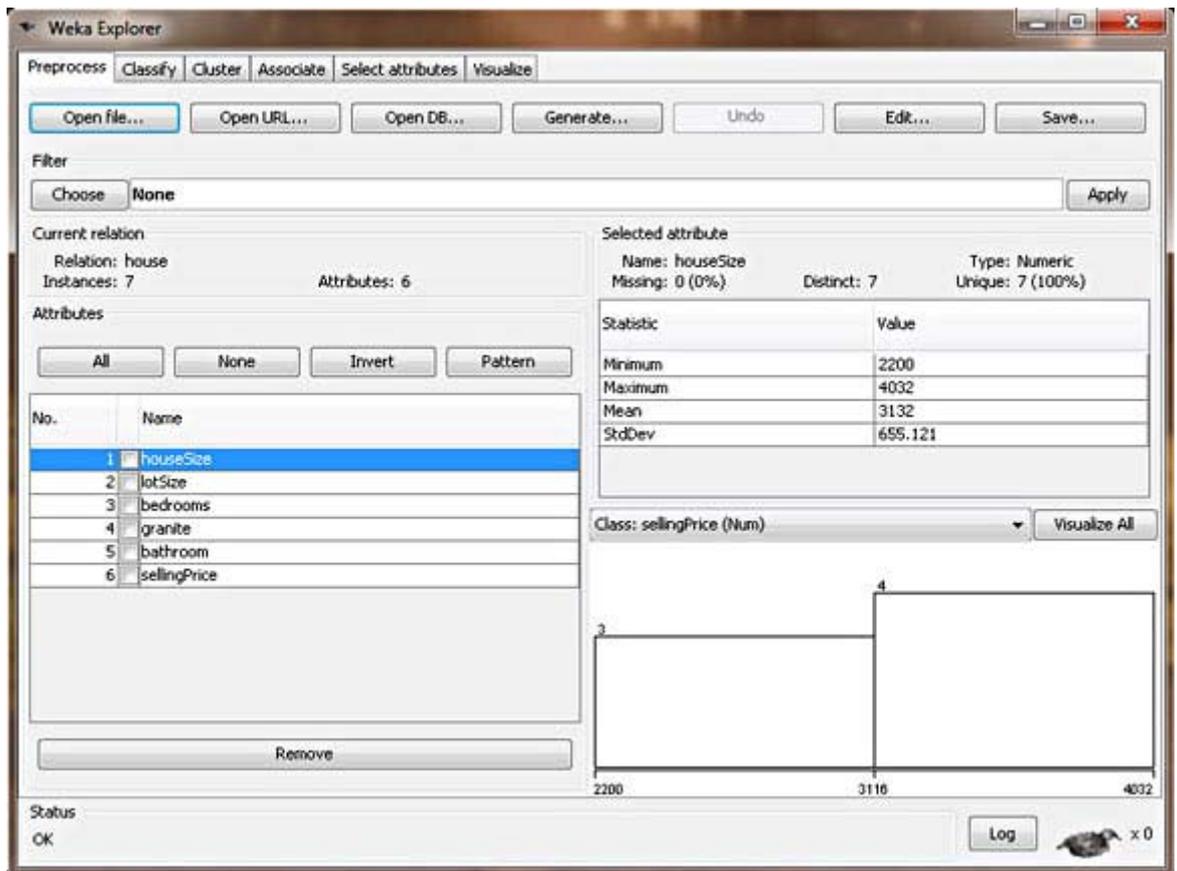


Рисунок 3- Окно WEKA Explorer с загруженными данными о домах

В этом окне вы можете проверить данные, на основании которых вы собираетесь строить модель. В левой части окна Explorer показаны параметры объектов (Attributes), которые соответствуют заголовкам столбцов нашей исходной таблицы, а также указано количество объектов (Instances), т.е. строк таблицы. Если вы щелкните мышкой на одном из заголовков столбцов, то в правой панели будет выведена полная информация о наборе данных в данном столбце. Например, если мы выберем столбец houseSize в левой панели (он выбран по умолчанию), то в правой панели отобразится дополнительная статистическая информация по этому столбцу. Будет показано максимальное значение в столбце (4032 кв.фута) и минимальное значение (2200 кв.футов). Кроме того, будет подсчитано среднее значение (3131 кв.фут) и стандартное отклонение (655 кв.футов) (стандартное отклонение – статистический показатель рассеивания значений случайной величины). Наконец, здесь же вам предлагается возможность визуального анализа данных (кнопка Visualize All). Поскольку в нашей таблице данных не так много, то их визуальное отображение не дает такой наглядной аналитической картины, как в случае использования сотен или тысяч показателей.

Для того чтобы создать модель, откройте закладку **Classify**. В качестве первого шага, нам надо выбрать тип модели для анализа, чтобы указать WEKA, каким образом мы хотим анализировать наши данные, и какую модель построить:

- 1) Щелкните на кнопке **Choose** и разверните меню **functions**.
- 2) Выберите опцию **LinearRegression**.

Таким образом, мы указали WEKA, что мы хотим создать модель регрессионного анализа. Как вы заметили, меню включает целое множество моделей. Множество! Это еще раз подчеркивает факт нашего весьма поверхностного знакомства с областью интеллектуального анализа данных. Обратите внимание: в меню включена опция **SimpleLinearRegression**, однако мы не используем ее, поскольку этот тип модели определяет значение зависимой переменной по значениям одного независимого параметра, а у нас их целых шесть. Если вы выбрали правильную модель, то окно WEKA Explorer должно выглядеть так, как показано на рисунке 4.

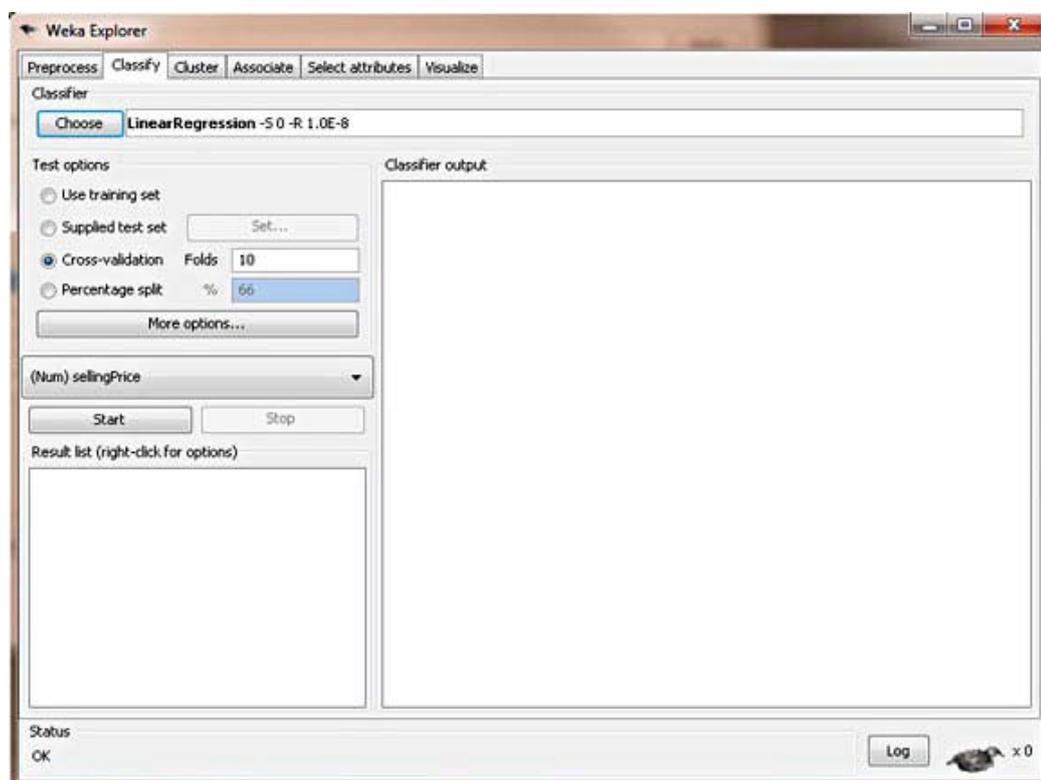


Рисунок 4 - Модель линейного регрессионного анализа WEKA

После того, как мы выбрали тип модели, нужно указать WEKA, какие данные должны использоваться для ее создания. Несмотря на то, что ответ на этот вопрос для нас вполне очевиден – нужно взять данные из созданного нами ARFF-файла – существует несколько других, более сложных, возможностей предоставления данных для анализа. Опция **Supplied test set** позволяет указать дополнительный набор тестовых данных для модели, опция **Cross-validation** использует несколько наборов данных, усредняет их и строит модель на основе средних значений, а опция **Percentage split** использует в качестве базы для модели процентиля набора данных. Эти способы применяются для создания аналитических моделей, которые мы рассмотрим в следующих статьях этой серии. В случае регрессионного анализа нам нужна опция **Use training set**. В этом случае WEKA

создаст модель на базе данных из загруженного ARFF-файла.

Завершающий этап создания модели – выбор зависимой переменной (столбца, в котором находится неизвестное нам значение, которое требуется рассчитать). В нашем примере – это цена дома, так как именно это значение мы и хотим узнать. Сразу после секции Test options находится раскрывающийся список, в котором вам нужно выбрать зависимый параметр. По умолчанию должен быть выбран атрибут **sellingPrice**. Если это не так, выберите сами этот параметр.

Мы определили все параметры и можем приступить к созданию модели. Нажмите кнопку **Start**. В результате окно WEKA должно выглядеть так, как показано на рисунке 5.

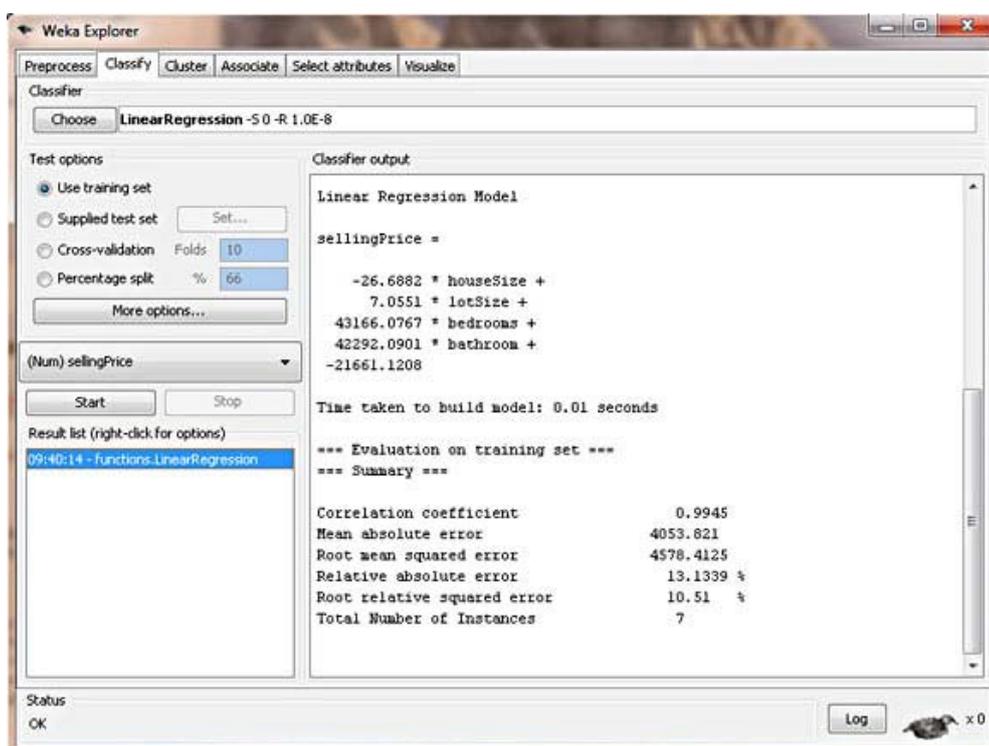


Рисунок 5 - Регрессионная модель WEKA для расчета стоимости дома

Готовая модель регрессионного анализа

$$\begin{aligned} \text{sellingPrice} = & (-26.6882 * \text{houseSize}) + \\ & (7.0551 * \text{lotSize}) + \\ & (43166.0767 * \text{bedrooms}) + \\ & (42292.0901 * \text{bathroom}) \\ & - 21661.1208 \end{aligned}$$

Рассчитаем цену конкретного дома:

$$\begin{aligned} \text{sellingPrice} = & (-26.6882 * 3198) + \\ & (7.0551 * 9669) + \\ & (43166.0767 * 5) + \\ & (42292.0901 * 1) \\ & - 21661.1208 \end{aligned}$$

$$\text{sellingPrice} = 219,328$$

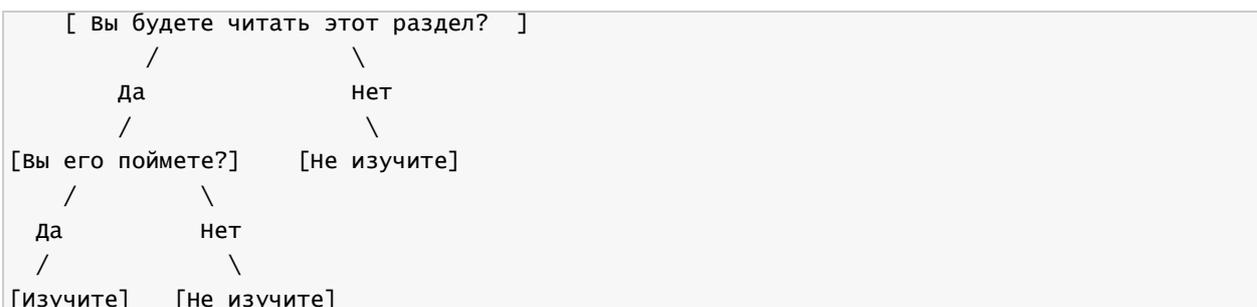
Рассмотрим зависимости между данными нашей модели и постараемся сделать определенные выводы относительно правил формирования цен на недвижимость.

- Гранитные элементы в оформлении кухни не влияют на цену дома — WEKA использует только те данные, которые, согласно статистике, влияют на точность модели (влияние каждого независимого параметра на зависимую переменную определяется с помощью коэффициента детерминации R-квадрат, обсуждение которого выходит за рамки этой статьи). Таким образом, параметры, не имеющие достаточного влияния на зависимую переменную, в модели не учитываются. Наша регрессионная модель свидетельствует о том, что использование гранита на кухне не влияет на цену дома.
- Состояние ванных комнат и сантехники влияет на цену дома — поскольку мы используем значения 0 или 1 в качестве показателя модернизации ванных комнат, то соответствующий коэффициент регрессионной модели демонстрирует нам, как современное сантехническое оборудование влияет на цену дома, а именно добавляет 42292\$ к его цене.
- Большая площадь дома снижает его цену — Согласно модели WEKA, по мере роста площади домов, цена снижается. Это следует из того, что модель включает переменную houseSize с отрицательным коэффициентом. Что же получается? Увеличение площади дома на 1 кв.фут снижает его стоимость на 26\$? Подобное утверждение кажется очевидной бессмыслицей. Мы же рассматриваем дома в Америке: чем больше, тем лучше, особенно в Техасе, где я живу. Как же это понимать? Это классический пример случая «каков вопрос, таков и ответ». На самом деле, размер дома не является независимой величиной. Этот параметр связан, например, с количеством спален - очевидно, что в больших домах и количество спален больше. Так что наша модель, увы, не идеальна, но мы можем ее поправить. Запомните: закладка Preprocess позволяет удалить столбцы из набора данных. В качестве самостоятельного упражнения, удалите столбец houseSize и создайте новую модель. Проверьте, как изменение набора данных отразится на цене дома, и какая из двух моделей больше соответствует реальности (уточненная цена моего дома \$217,894).

Классификация

Метод классификации (также известный как метод классификационных деревьев или деревьев принятия решений) - это алгоритм анализа данных, который определяет

пошаговый способ принятия решения в зависимости от значений конкретных параметров. Дерево этого метода имеет следующий вид: каждый узел представляет собой точку принятия решения на основании входных параметров. В зависимости от конкретного значения параметра вы переходите к следующему узлу, от него – к следующему узлу, и так далее, пока не дойдете до листа, который и дает вам окончательное решение. Звучит довольно запутанно, но на самом деле метод достаточно прямолинеен. Давайте обратимся к конкретному примеру.



Это простое классификационное дерево определяет ответ на вопрос «Изучите ли вы принцип построения классификационного дерева?» В каждом узле вы отвечаете на соответствующий вопрос и переходите по соответствующей ветке к следующему узлу, до тех пор, пока не дойдете до листа с ответом «да» или «нет». Эта модель применима к любым сущностям, и вы сможете ответить, в состоянии ли эти сущности изучить классификационные деревья, с помощью двух простых вопросов. В этом и состоит основное преимущество классификационных деревьев – они не требуют чрезмерного количества информации для создания достаточно точного и информативного дерева решений.

4.8 (модификация C4.5)

Набор данных, который мы будем использовать для примера классификационного анализа, содержит информацию, собранную дилерским центром BMW. Центр начинает рекламную кампанию, предлагая расширенную двухгодичную гарантию своим постоянным клиентам. Подобные кампании уже проводились, так что дилерский центр располагает 4500 показателями относительно предыдущих продаж с расширенной гарантией. Этот набор данных обладает следующими атрибутами:

- Распределение по доходам [0=\$0-\$30k, 1=\$31k-\$40k, 2=\$41k-\$60k, 3=\$61k-\$75k, 4=\$76k-\$100k, 5=\$101k-\$150k, 6=\$151k-\$500k, 7=\$501k+]
- Год/месяц покупки первого автомобиля BMW
- Год/месяц покупки последнего автомобиля BMW
- Воспользовался ли клиент расширенной гарантией

Загрузите файл `bmw-training.arff` (см.раздел Загрузка) в программный пакет WEKA,

используя те же шаги, которые мы проделали для загрузки данных в случае регрессионного анализа. Замечание: в предлагаемом файле содержатся 3000 из имеющихся 4500 записей. Мы разделили набор данных так, чтобы часть их использовалась для создания модели, а часть – для проверки ее точности, чтобы убедиться, что модель не является подогнанной под конкретный набор данных.

Откройте закладку Classify, выберите опцию trees, а затем опцию J48. Убедитесь, что выбрана опция Use training set, чтобы пакет WEKA при создании модели использовал именно те данные, которые мы только что загрузили в виде файла. Нажмите кнопку Start и предоставьте WEKA возможность поработать с нашими данными. Результирующая модель должна выглядеть так:

```

Number of Leaves :      28

Size of the tree :      43

Time taken to build model: 0.18 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances      1774           59.1333 %
Incorrectly Classified Instances    1226           40.8667 %
Kappa statistic                     0.1807
Mean absolute error                  0.4773
Root mean squared error              0.4885
Relative absolute error              95.4768 %
Root relative squared error          97.7122 %
Total Number of Instances           3000

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
                0.662   0.481   0.587     0.662   0.622     0.616    1
                0.519   0.338   0.597     0.519   0.555     0.616    0
weighted Avg.   0.591   0.411   0.592     0.591   0.589     0.616

=== Confusion Matrix ===

  a    b  <-- classified as
1009  516 |   a = 1
710  65  |   b = 0

```

Наиболее существенные данные – это показатели классификации "Correctly Classified Instances" (59.1%) и "Incorrectly Classified Instances" (40.9%). Кроме того, следует обратить внимание на число в первой строке столбца ROC Area (0.616). Чуть позже мы подробно обсудим эти значения, пока же просто запомните их. Наконец, таблица Confusion Matrix показывает количество ложноположительных (516) и

ложноотрицательных (710) распознаваний. Поскольку показатель точности нашей модели – 59,1%, то в первоначальном рассмотрении ее нельзя назвать достаточно хорошей.

Вы сможете увидеть дерево, если щелкнете правой кнопкой мышки в панели результирующей модели. В контекстном меню выберите опцию Visualize tree. На экране отобразится визуальное представление классификационного дерева нашей модели (рисунок 3), однако в данном случае картинка мало чем нам поможет. Еще один способ увидеть дерево модели – прокрутить вверх вывод в окне Classifier Output, там вы найдете текстовое описание дерева с узлами и листьями.

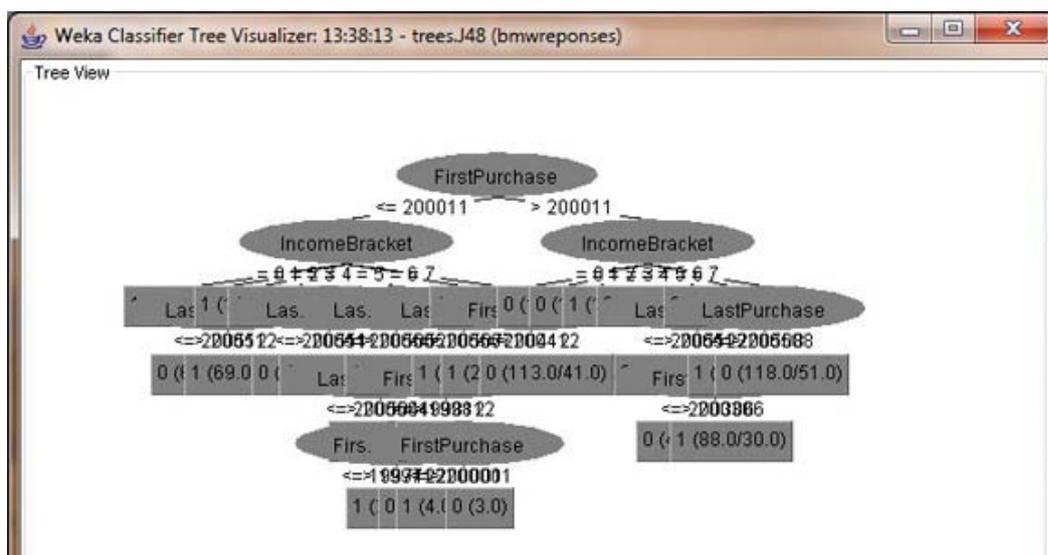


Рисунок 6 -Визуальное представление дерева классификации

Остался последний этап проверки классификационного дерева: нам надо пропустить оставшийся набор данных через полученную модель и проверить, насколько результаты классификации будут отличаться от реальных данных. Для этого в секции Test options выберите опцию Supplied test set и нажмите на кнопку Set. Укажите файл bmw-test.arff, содержащий оставшиеся 1500 данных, которые не были включены в обучающий набор. При нажатии на кнопку Start WEKA пропустит тестовые данные через модель и покажет результат работы модели. Давайте нажмем на Start и проверим, что у нас получилось.

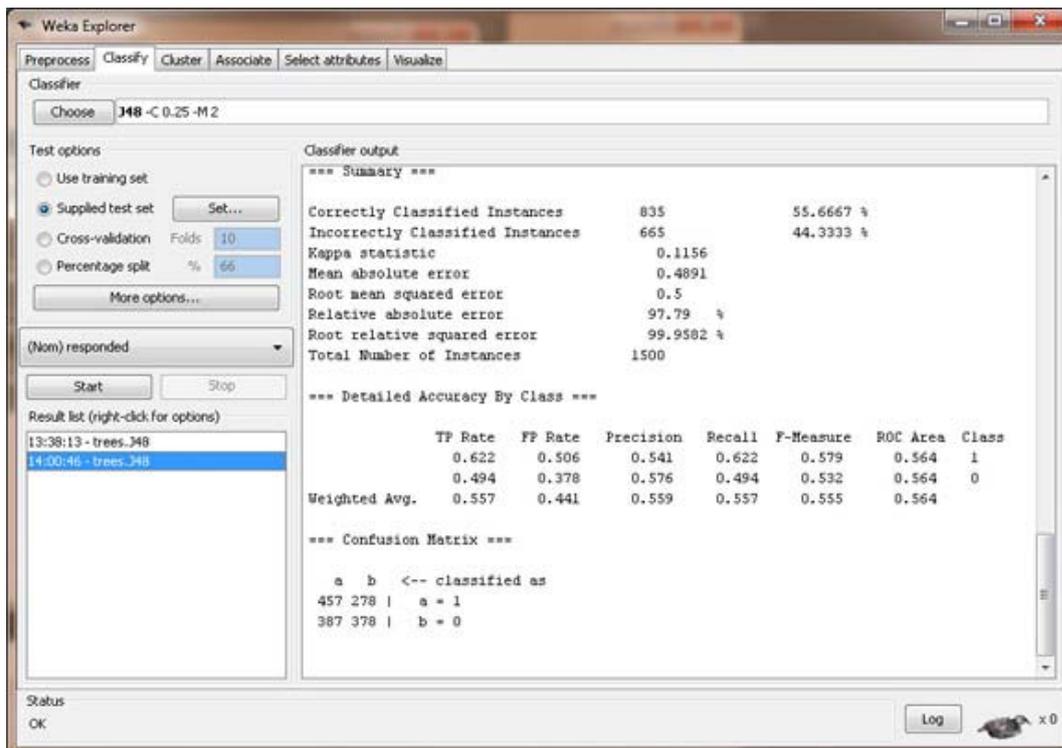


Рисунок 7 - Проверка классификационного дерева

Сравнивая показатель Correctly Classified Instances для тестового набора (55,7%) с этим же показателем для обучающего набора (59,1%), мы видим, что точность модели для двух разных наборов данных примерно одинакова. Это значит, что новые данные, которые будут использоваться в этой модели в будущем, не снизят точность ее работы.

Однако, поскольку собственно точность модели довольно низка (всего лишь 60% данных классифицировано верно), мы имеем полное право остановиться и сказать: «Она работает с точностью чуть выше 50%, с таким же успехом мы можем просто пытаться угадать значение случайным образом». Существуют случаи, когда использование алгоритмов интеллектуального анализа данных приводит к созданию неудачной аналитической модели.

Классификационная модель не подходит для анализа имеющихся у нас данных.

Метод ближайших соседей

Алгоритм метода ближайших соседей во многом схож с алгоритмом, используемым в методе кластеризации. Метод определяет расстояние между неизвестной точкой и всеми известными точками данных. Самый простой и наиболее распространенный способ определения расстояния – это нормализованное евклидово расстояние.

Загрузим файл `bmw-training.arff` в WEKA, выполнив в закладке Preprocess.

Точно так же, как мы проделали это для методов регрессионного анализа и классификации в предыдущих статьях, мы должны открыть закладку Classify. В панели

Classify нужно выбрать опцию lazy, а затем Ibk (здесь IB означает Instance-Based – обучение на примерах, а k указывает на количество соседей, поведение которых мы хотим исследовать). Убедитесь, что вы выбрали опцию Use training set, чтобы использовать набор данных, который мы только что загрузили в WEKA. Нажмите кнопку Start.

Результат вычислений Ibk

```

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances      2663           88.7667 %
Incorrectly Classified Instances    337           11.2333 %
Kappa statistic                     0.7748
Mean absolute error                 0.1326
Root mean squared error             0.2573
Relative absolute error             26.522 %
Root relative squared error         51.462 %
Total Number of Instances          3000

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
                0.95    0.177    0.847     0.95    0.896     0.972    1
                0.823    0.05    0.941     0.823   0.878     0.972    0
Weighted Avg.   0.888    0.114    0.893     0.888   0.887     0.972

=== Confusion Matrix ===

  a    b  <-- classified as
1449  76 |    a = 1
261 1214 |    b = 0

```

У модели, использующей метод ближайших соседей, показатель точности равен 89% - совсем неплохо для начала, учитывая то, что точность предыдущей модели составляла всего 59%. Практически 90% точности – это вполне приемлемый уровень. Давайте рассмотрим результаты работы метода в терминах ложных определений, чтобы вы смогли на конкретном примере увидеть, как именно WEKA может использоваться для решения реальных вопросов бизнеса.

Результаты использования модели на нашем наборе данных показывают, что у нас есть 76 ложноположительных распознаваний (2.5%) и 261 ложноотрицательных распознаваний (8.7%). В нашем случае ложноположительное распознавание означает, что модель считает, что данный покупатель приобретет расширенную гарантию, хотя на самом деле он отказался от покупки. Ложноотрицательное распознавание, в свою очередь, означает, что согласно результатам анализа данный покупатель откажется от расширенной гарантии, а на самом деле он ее купил. Предположим, что стоимость каждой рекламной листовки, рассылаемой дилером, составляет \$3, а покупка одной расширенной гарантии приносит ему 400\$ дохода. Таким образом, ошибки ложного распознавание в

терминах расходов и доходов нашего дилера будут выглядеть следующим образом: $400\$ - (2.5\% * \$3) - (8.7\% * 400) = \$365$. Следовательно, ложное распознавание ошибается в пользу дилера. Сравним этот показатель с данными модели классификации: $\$400 - (17.2\% * \$3) - (23.7\% * \$400) = \304 . Как вы видите, использование более точной модели повышает потенциальный доход дилера на 20%.

Naive Bayes (наивный байесовский метод)

"Наивная" классификация - достаточно прозрачный и понятный метод классификации. "Наивной" она называется потому, что исходит из предположения о взаимной независимости признаков.

Свойства наивной классификации:

1. Использование всех переменных и определение всех зависимостей между ними.
2. Наличие двух предположений относительно переменных:
 - все переменные являются одинаково важными;
 - все переменные являются статистически независимыми, т.е. значение одной переменной ничего не говорит о значении другой.

Для использования этого метода в системе Weka панели Classify нужно выбрать опцию NativeBayes.

Результаты интерпретируются также как и при использовании других методов классификации.

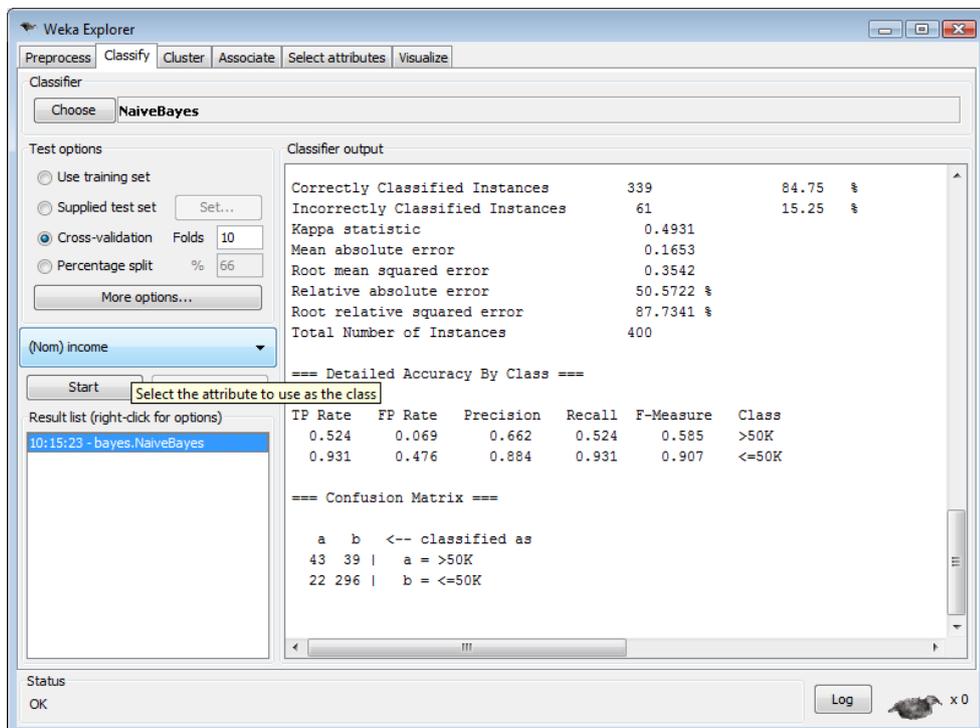


Рисунок 8 – Окно выбора метода

1R

Метод классификации 1R – один из самых простых и понятных методов

классификации. Применяется как к числовым данным, которые разбиваются на промежутки, так и к данным типа nominal.

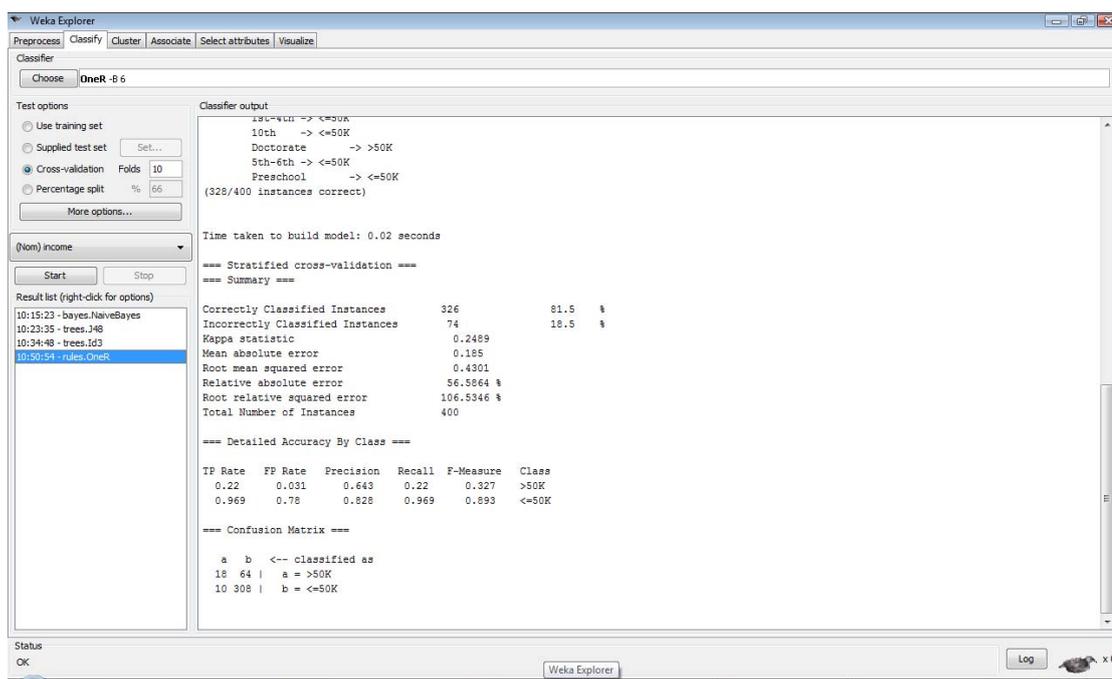


Рисунок 9

Для использования этого метода в системе Weka панели Classify нужно выбрать опцию OneR.

Результаты интерпретируются также как и при использовании других методов классификации.

SVM

Для этого метода не требуется каких-либо преобразований исходной выборки.

Данный метод является алгоритмом классификации с использованием математических функций. Метод использует нелинейные математические функции. Номинальные данные преобразуются в числовые. Основная идея метода опорных векторов – перевод исходных векторов в пространство более высокой размерности и поиск максимальной разделяющей гиперплоскости в этом пространстве.

Для использования этого метода в системе Weka панели Classify нужно выбрать опцию SMO.

Результаты интерпретируются также как и при использовании других методов классификации.

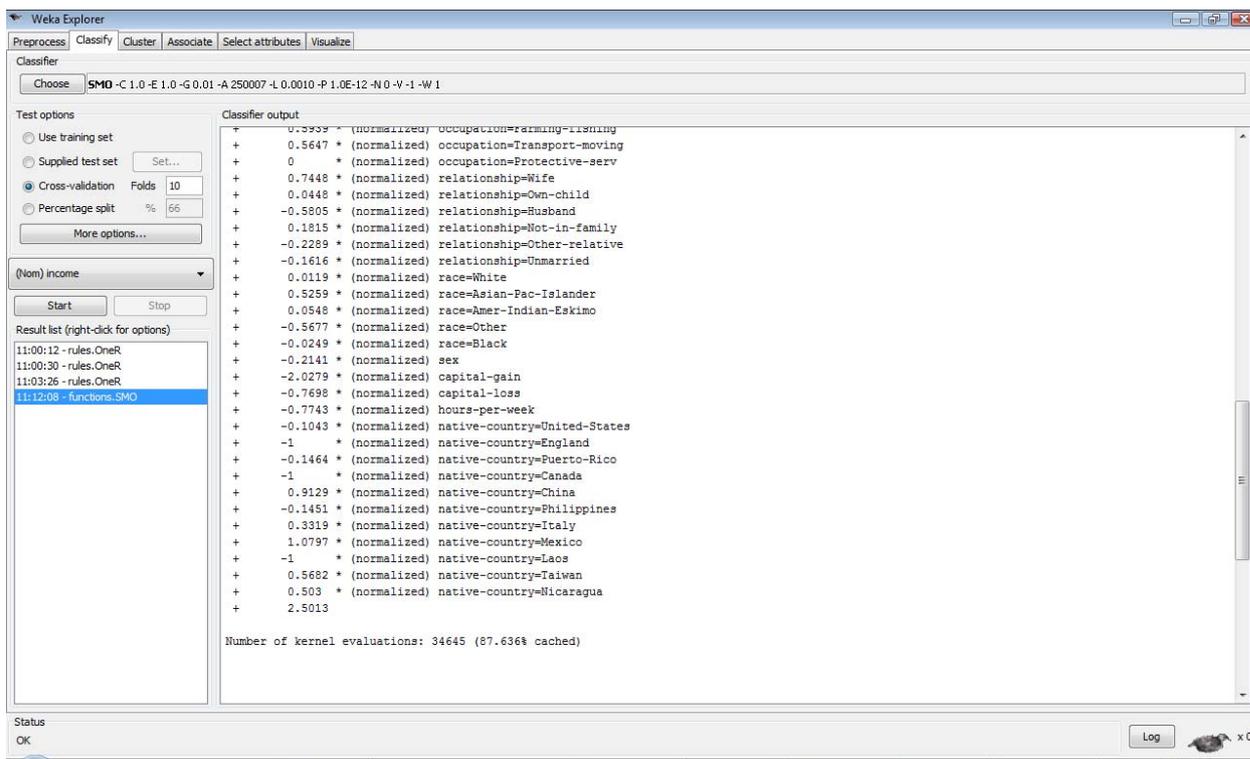


Рисунок 10

Кластеризация

Кластеризация позволяет разбить данные на группы, каждая из которых имеет определенные признаки. Метод кластерного анализа используется в тех случаях, когда необходимо выделить некоторые правила, взаимосвязи или тенденции в больших наборах данных. В зависимости от потребностей бизнеса, вы можете выделить несколько различных групп данных. Одно из явных преимуществ кластеризации по сравнению с классификацией состоит в том, что для разбиения множества на группы может использоваться любой атрибут (если вы помните, метод классификации использует только определенное подмножество атрибутов). В качестве основного недостатка метода кластеризации следует упомянуть тот факт, что составитель модели должен заранее решить, на сколько групп следует разбить данные. Для человека, который не имеет никакого представления о конкретном наборе данных, принять такое решение достаточно затруднительно. Следует ли нам создать три группы или пять групп? А может, нам нужно определить десять групп? Может потребоваться несколько итераций проб и ошибок, для того чтобы определить оптимальное количество кластеров.

Тем не менее, для среднестатистического пользователя кластеризация может оказаться наиболее полезным методом интеллектуального анализа данных. Этот метод позволит вам быстро разбить ваши данные на отдельные группы и сделать конкретные выводы и предположения относительно каждой группы. Математические методы, реализующие кластерный анализ, довольно сложны и запутаны, так что в случае кластеризации мы будем целиком полагаться на вычислительные возможности WEKA.

Загрузите файл `bmw-browsers.arff` в WEKA, выполнив те же шаги, которые мы проделали ранее для открытия данных в закладке **Preprocess**.

Поскольку мы хотим разбить имеющиеся у нас данные на кластеры, вместо закладки **Classify** нам потребуется закладка **Cluster**. Нажмите на кнопку **Choose** и в предлагаемом меню выберите опцию **SimpleKMeans**.

Щелкните на опции **SimpleKMeans** (дизайн пользовательского интерфейса оставляет желать лучшего, но работать с ним можно). Единственный атрибут алгоритма, который нас интересует – это поле `numClusters`, которое указывает на количество кластеров для разбиения (напоминаем, что это значение вам нужно выбрать еще до создания модели). Изменим значение по умолчанию (2) на 5.

Результаты кластеризации

Attribute	Cluster#					
	Full Data (100)	0 (26)	1 (27)	2 (5)	3 (14)	4 (28)
Dealership	0.6	0.9615	0.6667	1	0.8571	0
Showroom	0.72	0.6923	0.6667	0	0.5714	1
ComputerSearch	0.43	0.6538	0	1	0.8571	0.3214
M5	0.53	0.4615	0.963	1	0.7143	0
3Series	0.55	0.3846	0.4444	0.8	0.0714	1
Z4	0.45	0.5385	0	0.8	0.5714	0.6786
Financing	0.61	0.4615	0.6296	0.8	1	0.5
Purchase	0.39	0	0.5185	0.4	1	0.3214
Clustered Instances						
0	26 (26%)					
1	27 (27%)					
2	5 (5%)					
3	14 (14%)					
4	28 (28%)					

Данные кластеризации показывают, каким образом сформирован каждый кластер: значение «1» означает, что у всех данных в этом кластере соответствующий атрибут равен 1, а значение «0» означает, что у всех данных в этом кластере соответствующий атрибут равен 0. Данные соответствуют среднему значению атрибута на кластере. Каждый кластер характеризует определенный тип поведения клиентов, таким образом, на основании нашего разбиения мы можем сделать некоторые полезные выводы:

- Кластер 0— эту группу посетителей можно было бы назвать «мечтатели». Они бродят вокруг дилерского центра, рассматривая машины, выставленные на внешней парковке, но никогда не заходят внутрь, и, хуже того, никогда ничего не покупают.
- Кластер 1— эту группу следовало бы назвать «поклонники M5», поскольку

они сразу же подходят к выставленным автомобилям этой модели, полностью игнорируя BMW серии 3 или Z4. Тем не менее, эта группа не отличается высокими показателями покупки машин – всего 52%. Это потенциально может свидетельствовать о недостаточно продуманной стратегии продаж и о необходимости улучшить работу дилерского центра, например, за счет увеличения количества продавцов в секции M5.

- Кластер 2— эта группа настолько мала, что мы могли бы назвать ее выбраковкой. Дело в том, что данные этой группы статистически довольно разбросаны, и мы не можем сделать каких-либо определенных заключений относительно поведения посетителей, попавших в этот кластер (подобная ситуация может указывать на то, что вам следует сократить количество кластеров в модели)
- Кластер 3— эту группу следовало бы назвать «любимцы BMW», потому что посетители, попавшие в это кластер, всегда покупают машину и получают необходимое финансирование. Обратите внимание, данные этого кластера демонстрируют интересную модель поведения этих покупателей: сначала они осматривают выставленные на парковке машины, а затем обращаются к поисковой системе дилерского центра. Как правило, они покупают модели M5 или Z4, но никогда не берут модели третьей серии. Данные этого кластера указывают на то, что дилерскому центру следует активнее привлекать внимание к поисковым компьютерам (может быть, вынести их на внешнюю парковку), и кроме того, следует найти какой-нибудь способ выделить модели M5 и Z4 в результатах поиска, чтобы гарантированно обратить на них внимание посетителей. После того, как посетитель, попавший в этот кластер, выбрал определенную модель автомобиля, он гарантированно получает необходимый кредит и совершает покупку.
- Кластер 4— эту группу можно назвать «начинающие владельцы BMW», поскольку они всегда ищут модели 3 серии и никогда не интересуются более дорогими M5. Они сразу же проходят в демонстрационный зал, не тратя время на осмотр машин на внешней стоянке. Кроме того, они не пользуются поисковой системой центра. Примерно 50% этой группы получают одобрение по кредиту, тем не менее, покупку совершают всего 32% участников. Анализируя данные этого кластера, можно сделать следующий вывод: посетители этой группы хотели бы купить свой первый BMW и точно знают, какая машина им нужна (модель 3 серии с минимальной

конфигурацией). Однако, для того чтобы купить машину, им нужно получить положительное решение по кредиту. Чтобы повысить уровень продаж среди посетителей 4 кластера, дилерскому центру следовало бы понизить уровень требований для получения кредита или снизить цены на модели 3 серии.

Еще один интересный способ изучения результатов кластеризации – это визуальное представление данных. Щелкните правой кнопкой мышки в секции Result List закладки Cluster (и вновь элементы пользовательского интерфейса оставляют желать лучшего). В контекстном меню выберите опцию Visualize Cluster Assignments. В результате откроется окно с графическим представлением результатов кластеризации, настройки которого вы можете выбрать наиболее удобным для вас образом. Для нашего примера, измените настройку оси X так, чтобы она соответствовала количеству автомобилей M5 (M5 (Num)), а настройку оси Y – так, чтобы она показывала количество купленных автомобилей (Purchase (Num)), и укажите выделение каждого кластера отдельным цветом (для этого установите значение поля Color в Cluster (Nom)). Такие настройки помогут нам оценить распределение по кластерам в зависимости от того, сколько человек интересовалось BMW M5, и сколько человек купило эту модель. Кроме того, сдвиньте указатель Jitter примерно на три четверти в сторону максимума, это искусственным образом увеличит разброс между группами точек, чтобы вам было удобнее их просматривать.

Соответствует ли визуальное отображение кластеризации тем заключениям, которые мы сделали на основании данных в листинге 5? Как мы видим, в окрестности точки $X=1$, $Y=1$ (посетители, которые интересовались автомобилями модели M5 и купили их) расположены только два кластера: 1 и 3. Аналогично, в окрестности точки $X=0$, $Y=0$ расположены только два кластера: 4 и 0. Соответствует ли это нашим выводам? Да, соответствует. Кластеры 1 и 3 покупают BMW M5, в то время как кластер 0 не покупает ничего, а кластер 4 ищет BMW серии 3. На рисунке 8 показано визуальное отображение кластеров нашей модели. Мы предлагаем вам самостоятельно попрактиковаться в обнаружении других трендов и течений, меняя настройки осей X и Y.

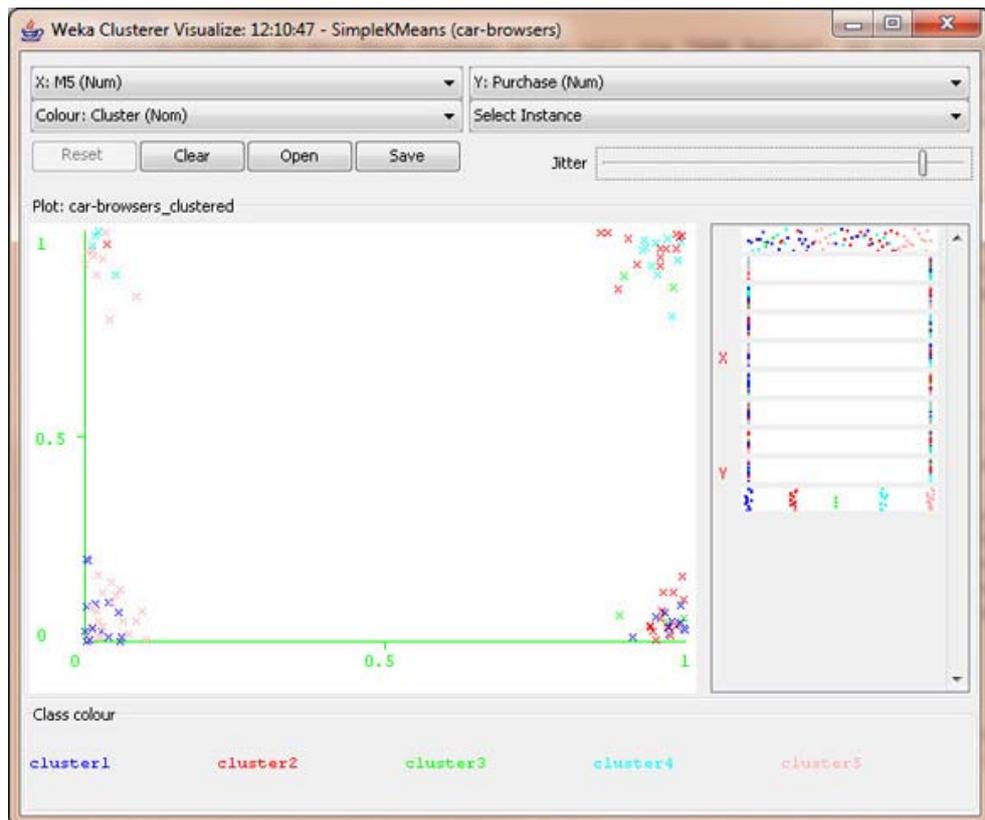


Рисунок 11 - Визуальное отображение кластеризации

Построение ассоциативных правил

Нахождение ассоциативных правил происходит почти так же, как и классификация. На вкладке Associate выбирается метод нахождения, для него выставляются параметры кликом на его названии, после чего нажимается кнопка Start и анализируется вывод (перед началом использования метода Априори необходимо применить фильтр RemoveType и удалить numeric-атрибуты). В нашем случае ассоциативные правила строятся по методу Априори.

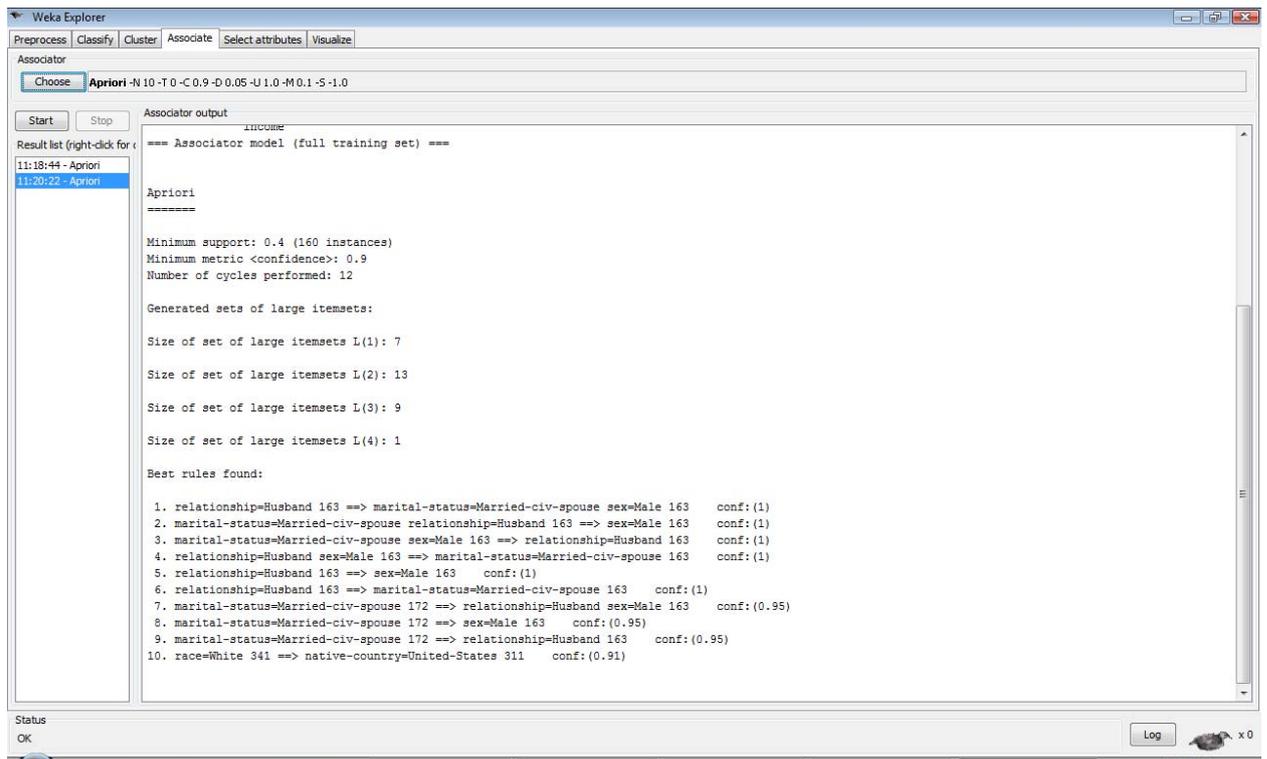


Рисунок 12

Варианты задания

- 1) Погодные условия в регионе
- 2) Продажа комплектующих изделий
- 3) Демографическая ситуация в регионе
- 4) Продажа земельных участков
- 5) Рынок труда
- 6) Больница
- 7) Железнодорожный транспорт
- 8) Авиационные перевозки
- 9) Олимпиада
- 10) Футбол
- 11) Туристический бизнес
- 12) Социальные сети
- 13) Интернет-провайдер
- 14) Здравоохранение
- 15) Автострахование
- 16) Кредитование
- 17) Экология
- 18) Правонарушения
- 19) Литература
- 20) Компьютеры

Лабораторная работа №3

Тема: оптимизация.

Цель работы: Получение практических навыков в области оптимизации данных.

Задание: Используя программное средство MS Excel, решите задачу оптимизации, согласно полученному варианту.

Отчет по лабораторной работе должен содержать:

1. Фамилию и номер группы, задание
2. Описание задач
3. Описание процесса решения
4. Решение

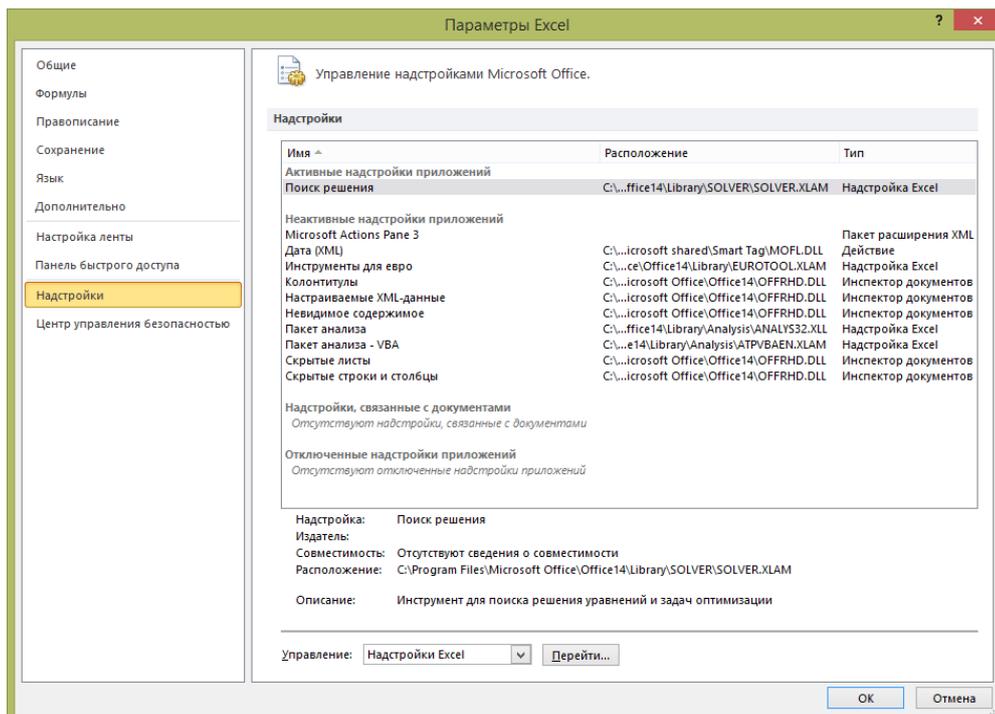
Основные теоретические сведения

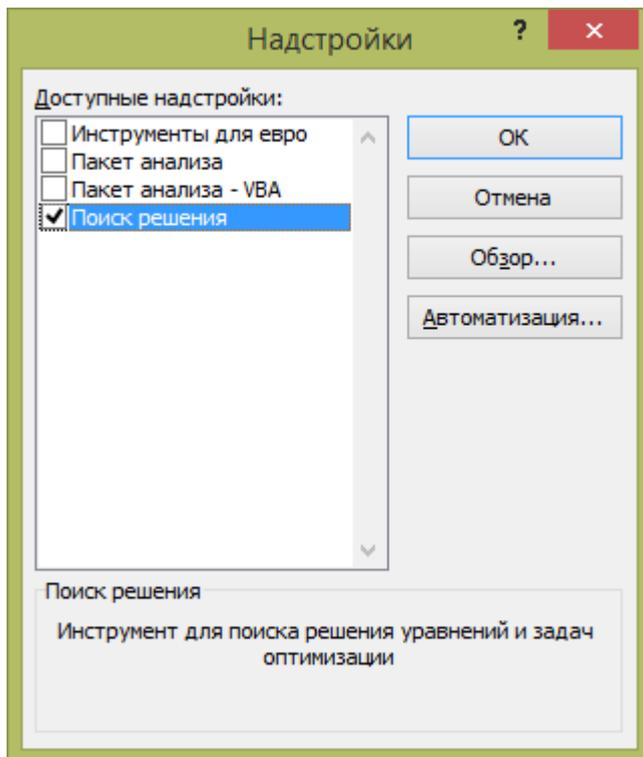
В MS Excel для решения задач оптимизации используется надстройка **ПОИСК РЕШЕНИЯ**

Сначала надстройку **Поиск решения** необходимо подключить (до первого использования)

В MS Excel 2003: **Сервис / Надстройки / Поиск решения / ОК**

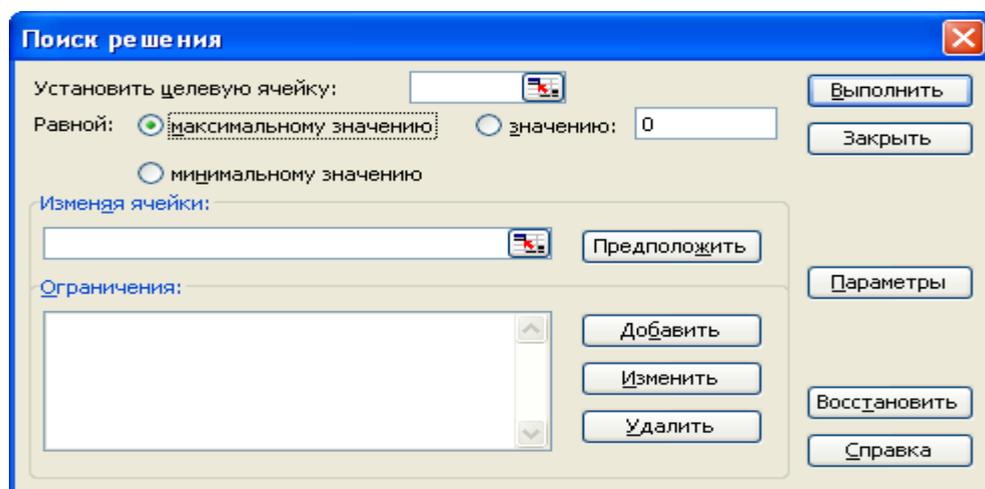
В MS Excel 2010: **Файл / Параметры / Надстройки / Управление надстройками/ Поиск решения / ОК**



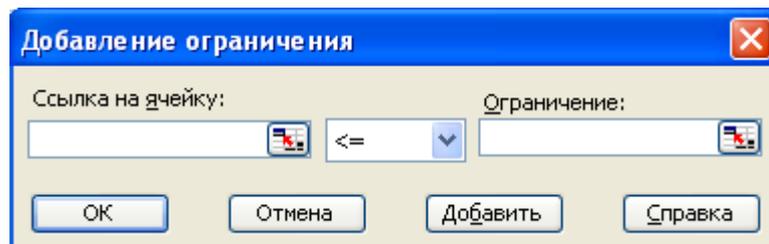


В MS Excel кнопка **Поиск решения** появится во вкладке **Данные**.

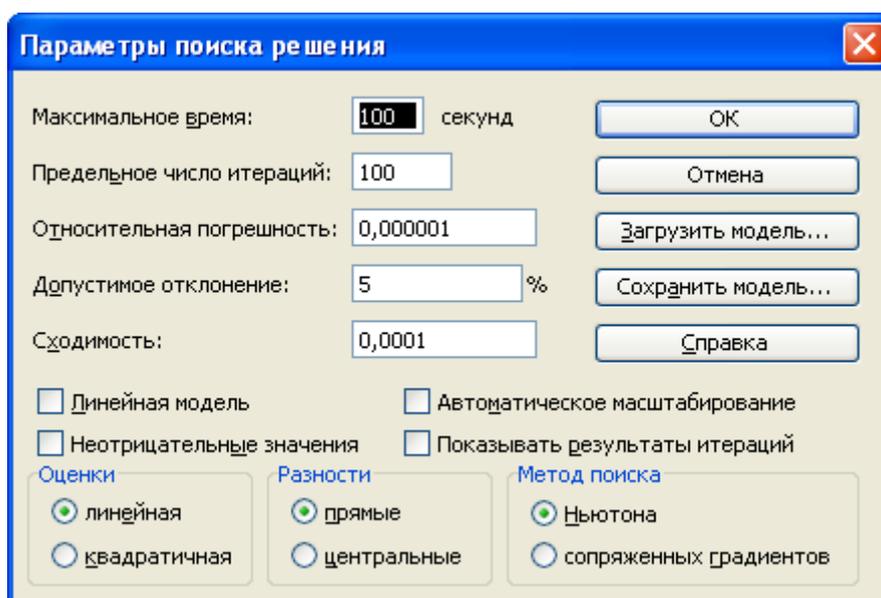
В программе Excel в меню Сервис применяя команду Поиск решения, откроется диалоговое окно где устанавливается адрес целевой ячейка, диапазон переменных.



С помощью кнопки **Добавить** вводятся необходимые ограничения.



Кнопка **Параметры** открывает диалоговое окно **Параметры поиска решения**, где по умолчанию стоит определенный набор команд.



По умолчанию значение допустимого отклонения стоит 5%. Это значит, что процедура оптимизации продолжается только до тех пор, пока значение целевой функции будет отличаться от оптимального не более чем на 5%. Более высокие значения допустимого отклонения ускоряют работу средства Поиск решения при оптимизации моделей, однако существует риск, что найденное значение будет значительно отличаться от истинного оптимума соответствующей задачи. Устанавливая значение допустимого отклонения, например, равным 0 %, мы заставляем Поиск решения находить истинный оптимум задачи за счет, возможно, более длительного времени решения.

Для улучшения работы средства Поиска решения настройка диалогового окна Параметры поиска решения часто применяется при решении задач нелинейного программирования.

Значение в поле Сходимость используется для завершения процесса поиска решения, когда изменение целевой функции происходит очень медленно. Если установить меньшее значение сходимости, чем предусмотрено по умолчанию (0,0001), программа продолжит процесс оптимизации даже при малых изменениях целевой функции.

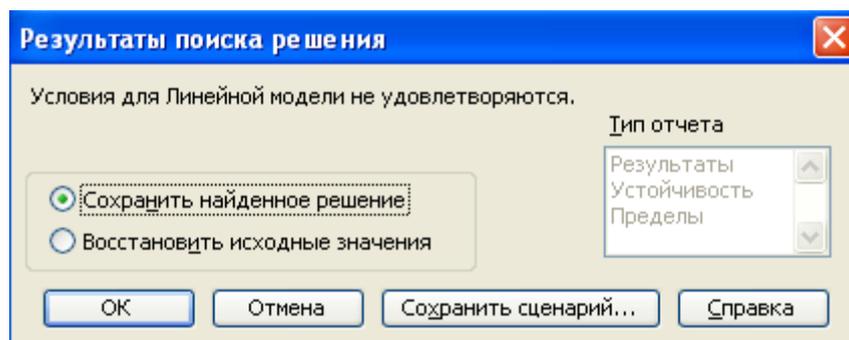
Если установить в области Оценки переключатель квадратичная, Поиск решения будет применять для вычисления различных оценок более точную квадратичную аппроксимацию, а не линейную (по умолчанию). Кроме того, установка в области Разности переключателя центральные вместо переключателя прямые приведет к тому, что Поиск решения для вычисления частных производных будет применять более точную аппроксимацию, используя большее количество точек.

Обе эти установки улучшают вычисляемые числовые оценки функций нелинейной модели, однако могут увеличить время решения, поскольку на каждой итерации следует производить дополнительные вычисления.

В диалоговом окне Параметры поиска решения можно также задать метод поиска решения. Метод сопряженных градиентов в процессе оптимизации использует меньше памяти, но требует большего количества вычислений, при заданном уровне точности, чем заданный по умолчанию метод Ньютона.

Значение в поле Относительная погрешность, определяет, на сколько точно должно совпадать вычисленное значение левой части ограничения со значением правой части, чтобы данное ограничение было выполнено.

Команда Выполнить запускает решение задачи. Поиск решения просит уточнить: сохранить ли найденное решение или нет.



При задании в диалоговом окне Поиска решения правых частей ограничений всегда следует указывать ссылки на ячейки в табличной модели.

Ячейки в правых частях неравенств в табличной модели должны содержать константы, а не формулы.

Процедура **Поиск решения** представляет собой мощный инструмент для выполнения сложных вычислений. Она позволяет по находить значения переменных, удовлетворяющих указанным критериям оптимальности, при условии выполнения заданных ограничений.

Наилучшие результаты она позволяет получить для задач выпуклого (в том числе линейного) программирования при условии отсутствия ограничений типа «равно».

Поиск решения можно использовать и для решения задач математического программирования других типов, но в этом случае процедура поиска часто заканчивается неудачей, а при благоприятном исходе находит лишь один из локальных оптимумов. Поэтому решение таких задач с помощью данной процедуры следует предварять их аналитическим исследованием на предмет свойств области допустимых решений, чтобы выбрать подходящие начальные значения и сделать правильное заключение о качестве и практической применимости полученного решения.

Результаты оптимизации оформляются в виде отчетов трёх типов:

- **Результаты.** Отражаются исходное (до оптимизации) и оптимальное значения целевой функции, значения переменных до и после оптимизации, а также формулы ограничений и дополнительные сведения об ограничениях.
- **Устойчивость.** Содержит сведения о чувствительности решения к малым изменениям в формуле целевой функции или в формулах ограничений. Отчет не создается для моделей, значения переменных в которых ограничены множеством целых чисел.
- **Пределы (Ограничения).** Состоит из верхнего и нижнего значения целевой функции и списка переменных, влияющих на нее, их нижних и верхних границ. Отчет не создается для моделей, значения переменных в которых ограничены множеством целых чисел. Нижней границей является наименьшее значение, которое может принимать переменная (влияющая ячейка) при условии, что значения других переменных (влияющих ячеек) фиксированы и удовлетворяют заданным ограничениям.

Пример решения задачи линейного программирования

Решим в MS Excel задачу линейного программирования

$$F = x_1 + 2x_2 + 2x_3 + x_4 + 6x_5 \rightarrow \min$$

$$\begin{cases} 2x_2 + x_4 + 2x_5 = 4 \\ x_1 + x_2 + 4x_5 = 5 \\ x_3 + x_5 = 3 \\ x_j \geq 0, j = \overline{1,5} \end{cases}$$

1. Создадим область переменных. Ячейки **B2:B6** будут играть роль переменных

	A	B	C
1	Переменные		
2	x1=		
3	x2=		
4	x3=		
5	x4=		
6	x5=		
7			

2. Введем формулу вычисления значений целевой функции

	A	B	C
1	Переменные		
2	x1=		
3	x2=		
4	x3=		
5	x4=		
6	x5=		
7	Целевая функция		
8	=B2+2*B3+2*B4+B5+6*B6		
9			

3. Создадим область ограничений

В ячейках A11:A13 будем вычислять **левые части** ограничений в системе

В ячейках B11:B13 введем **правые части** ограничений системы

	A	B	C
1	Переменные		
2	x1=		
3	x2=		
4	x3=		
5	x4=		
6	x5=		
7	Целевая функция		
8	0		
9	Ограничения		
10	Левая часть	Правая часть	
11		4	
12		5	
13		3	

В ячейках A11:A13 будем вычислять **левые части** ограничений в системе

	A	B	C
1	Переменные		
2	x1=		
3	x2=		
4	x3=		
5	x4=		
6	x5=		
7	Целевая функция		
8	0		
9	Ограничения		
10	Левая часть	Правая часть	
11	=2*B3+B5+2*B6		
12		5	
13		3	

В ячейках A11:A13 будем вычислять **левые части** ограничений в системе

	A	B
1	Переменные	
2	x1=	
3	x2=	
4	x3=	
5	x4=	
6	x5=	
7	Целевая функция	
8	0	
9	Ограничения	
10	Левая часть	Правая часть
11	0	4
12	=B2+B3+4*B6	5
13		3

4. Вызовем окно диалога **Поиск решения**

	A	B	C	D	E	F	G	H	I	J
1	Переменные									
2	x1=									
3	x2=									
4	x3=									
5	x4=									
6	x5=									
7	Целевая функция									
8	0									
9	Ограничения									
10	Левая часть	Правая часть								
11	0	4								
12	0	5								
13	0	3								

Поиск решения

Установить целевую ячейку:

Равной: максимальному значению значению:

минимальному значению

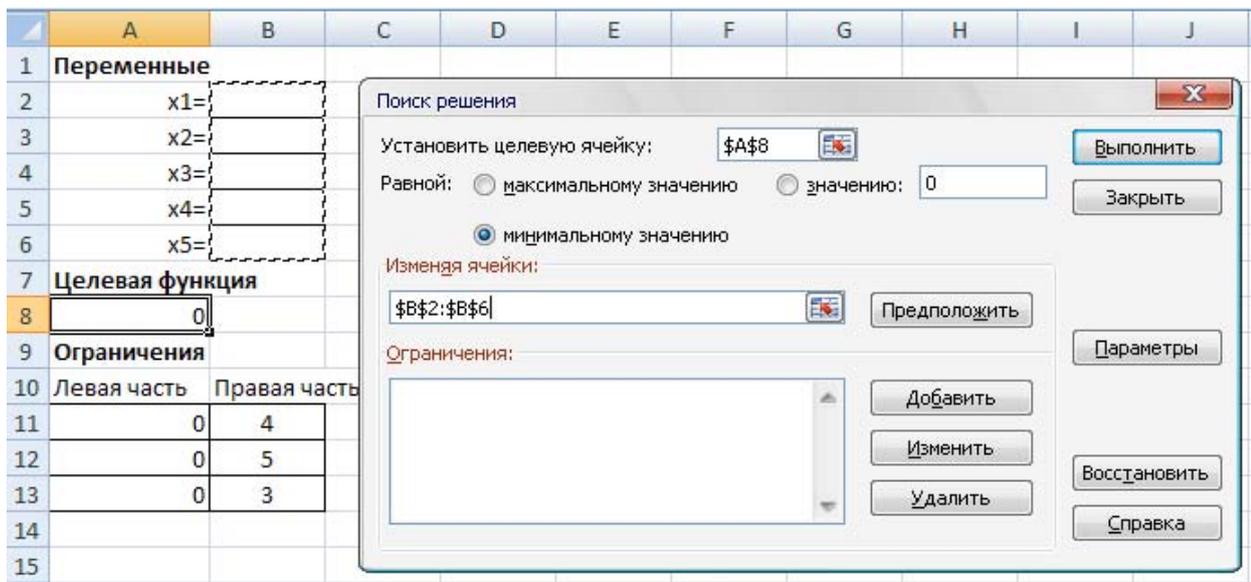
Изменяя ячейки:

Ограничения:

Устанавливаем целевую ячейку **A8** (там где вычисляется значение целевой функции)

Указываем направление оптимизации – минимизация (по условию)

В поле **Изменяя ячейки** указываем ячейки переменных **B2:B6**



Осталось нажать кнопку **Выполнить**

	A	B
1	Переменные	
2	x1=	0
3	x2=	1
4	x3=	2
5	x4=	0
6	x5=	1
7	Целевая функция	
8		12
9	Ограничения	
10	Левая часть	Правая часть
11	4	4
12	5	5
13	3	3

Ответ:

$$F_{min} = F(0,1,2,0,1) = 12$$

Пример решения транспортной задачи

Фирма имеет 4 фабрики и 5 центров распределения ее товаров. Фабрики фирмы располагаются в Денвере, Бостоне, Новом Орлеане и Далласе с производственными возможностями 200, 150, 225 и 175 единиц продукции ежедневно, соответственно. Центры распределения товаров фирмы располагаются в Лос-Анджелесе, Далласе, Сент-Луисе, Вашингтоне и Атланте с потребностями в 100, 200, 50, 250 и 150 единиц продукции ежедневно, соответственно. Хранение на фабрике единицы продукции, не поставленной в центр распределения, обходится в \$0,75 в день, а штраф за просроченную поставку единицы продукции, заказанной потребителем в центре распределения, но там не

находящейся, равен \$2,5 в день Стоимость перевозки единицы продукции с фабрик в пункты распределения приведена в таблице "Транспортные расходы":

Таблица "Транспортные расходы"

		1	2	3	4	5
		Лос-Анджелес	Даллас	Сен-Луис	Вашингтон	Атланта
1	Денвер	1,50	2,00	1,75	2,25	2,25
2	Бостон	2,50	2,00	1,75	1,00	1,50
3	Новый Орлеан	2,00	1,50	1,50	1,75	1,75
4	Даллас	2,00	0,50	1,75	1,75	1,75

Необходимо так спланировать перевозки, чтобы минимизировать суммарные транспортные расходы.

- Поскольку данная модель сбалансирована (суммарный объем произведенной продукции равен суммарному объему потребностей в ней), то в этой модели не надо учитывать издержки, связанные как со складированием, так и с недопоставками продукции.

Для решения данной задачи построим ее математическую модель.

$$Z = \sum_{i=1}^4 \sum_{j=1}^5 c_{ij} x_{ij}$$

Неизвестными в данной задаче являются объемы перевозок. Пусть x_{ij} - объем перевозок с i -ой фабрики в j -й центр распределения. Функция цели - это суммарные транспортные расходы, т. е. где c_{ij} – стоимость перевозки единицы продукции с i -и фабрики j -й центр распределения.

Неизвестные в данной задаче должны удовлетворять следующим ограничениям:

- Объемы перевозок не могут быть отрицательными.
- Так как модель сбалансирована, то вся продукция должна быть вывезена с фабрик, а потребности всех центров распределения должны быть полностью удовлетворены.

В результате имеем следующую модель: Минимизировать:

$$Z = \sum_{i=1}^4 \sum_{j=1}^5 c_{ij} x_{ij}$$

при ограничениях:

$$\sum_{i=1}^4 x_{ij} = b_{ij}, j \in [1,5]$$

$$x_{ij} \geq 0, i \in [1,4], j \in [1,5]$$

$$\sum_{i=1}^4 x_{ij} = a_i, i \in [1,4],$$

где a_i - объем производства на i -й фабрике, b_j — спрос в j -м центре распределения.

Решение задачи с помощью MS Excel.

1. Ввести данные, как показано на рис. 6.

В ячейки A1:E4 введены стоимости перевозок. Ячейки A6:E9 отведены под значения неизвестных (объемы перевозок). В ячейки G6:G9 введены объемы производства на фабриках, а в ячейки A11:E11 введена потребность в продукции в пунктах распределения. В ячейку F10 введена целевая функция =СУММПРОИЗВ(A1:E4;A6:E9).

	A	B	C	D	E	F	G	H
1	1,5	2	1,75	2,25	2,25			
2	2,5	2	1,75	1	1,5			
3	2	1,5	1,5	1,75	1,75			
4	2	0,5	1,75	1,75	1,75			
5								
6						0	200	
7						0	150	
8						0	225	
9						0	175	
10	0	0	0	0	0	0		
11	100	200	50	250	150			
12								

В ячейки A10:E10 введены формулы

=СУММ(A6:A9)

=СУММ(B6:B9)

=СУММ(C6:C9)

=СУММ(D6:D9)

=СУММ(E6:E9) определяющие объем продукции, ввозимой в центры распределения.

В ячейки F6:F9 введены формулы

=СУММ(A6:E6)

=СУММ(A7:E7)

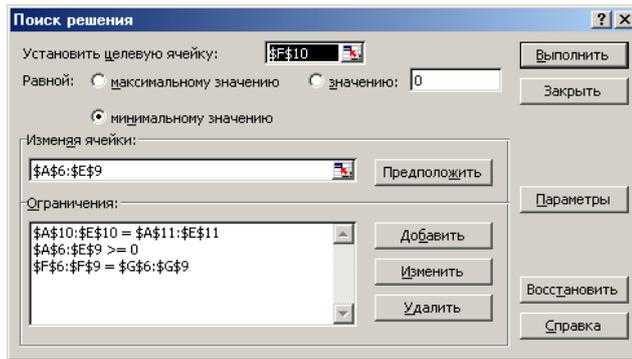
=СУММ(A8:E8)

=СУММ(A9:E9) вычисляющие объем продукции, вывозимой с фабрик.

2. Выбрать команду **Сервис/Поиск решения** и заполнить открывшееся диалоговое окно **Поиск решения**, как показано на рис. 7.

Внимание! В диалоговом окне **Параметры поиска решения** необходимо

установить флажок **Линейная модель**.



3. После нажатия кнопки **Выполнить** средство поиска решений находит оптимальный план поставок продукции и соответствующие ему транспортные расходы.

	A	B	C	D	E	F	G	H
1	1,5	2	1,75	2,25	2,25			
2	2,5	2	1,75	1	1,5			
3	2	1,5	1,5	1,75	1,75			
4	2	0,5	1,75	1,75	1,75			
5								
6	100	0	50	50	0	200	200	
7	0	0	0	150	0	150	150	
8	0	25	0	50	150	225	225	
9	0	175	0	0	0	175	175	
10	100	200	50	250	150	975		
11	100	200	50	250	150			
12								

Пример решения задачи нелинейного программирования

Задача называется задачей нелинейного программирования, если её математическая модель имеет вид

$$\begin{cases} g_1(x_1, x_2, \dots, x_n) = b_1, \\ g_2(x_1, x_2, \dots, x_n) = b_2, \\ \dots \\ g_m(x_1, x_2, \dots, x_n) = b_m, \end{cases}$$

$$f(x_1, x_2, \dots, x_n) \rightarrow \max(\min),$$

в которой среди g_i или f есть нелинейные функции.

В отличие от задач линейного программирования не существует единого метода для решения задач нелинейного программирования.

Задачи нелинейного программирования в Microsoft Excel решаются так же как и задачи линейного программирования (см. 1.2), с той лишь разницей, что в окне "Параметры поиска решения" необходимо сбросить флаги "Линейная модель" и, если это необходимо, "Неотрицательные значения".

Пример. Решить в Microsoft Excel следующую задачу нелинейного

программирования:

$$\text{найти } f = 4x_1 + x_1^2 + 8x_2 + x_2^2 \rightarrow \min \text{ при условии } x_1 + x_2 = 180.$$

В данной модели система ограничений состоит из одного линейного уравнения и нелинейной целевой функции.

Решение.

1. Заполняем ячейки на рабочем листе необходимыми переменными, целевой функцией и ограничениями:

	A	B	C
1	x1	x2	f
2			=4*A2+A2^2+8*B2+B2^2
3	Система ограничений		
4	=A2+B2		180

2. В окне "Параметры поиска решения" сбрасываем флаги "Линейная модель" (так как решаемая задача есть задача нелинейного программирования) и "Неотрицательные значения" (в условии задачи нет ограничений на знаки переменных).

3. После нажатия кнопки "Выполнить" получаем ответ:

	A	B	C
1	x1	x2	f
2	91	89	17278
3	Система ограничений		
4	180		180

из которого следует, что минимальное значение целевой функции равно 17278 и достигается при $x_1 = 91$ и $x_2 = 89$.

Пример использования метода градиентного спуска

Алгоритм

1. Заданную квадратичную функцию

$$f(x_1, x_2) = a_{11}x_1^2 + a_{22}x_2^2 + 2a_{12}x_1x_2 + b_1x_1 + b_2x_2 + c$$

представить в виде:

$$f(x_1, x_2) = \frac{1}{2} X^T A X + X^T B + C,$$

$$\text{где } X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad A = \begin{pmatrix} 2a_{11} & 2a_{12} \\ 2a_{12} & 2a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad C = c.$$

Градиент функции

$$\nabla f(x_1, x_2) = AX + B.$$

1. Следующее приближение к экстремуму рассчитывается по формуле:

$$X^{(k+1)} = X^{(k)} - h\nabla f(X^{(k)}) \quad (\text{с равным шагом})$$

или

$$X^{(k+1)} = X^{(k)} - h_k \nabla f(X^{(k)}) \quad (\text{с дробным шагом: если } f(X^{(k+1)}) > f(X^{(k)}), \text{ то}$$

шаг уменьшается, например, $h_{k+1} = \frac{h_k}{3}$).

2. Итерационный процесс заканчивается, как только $R^i < \varepsilon$.

Реализация в MS Excel

1. Минимизировать квадратичную функцию

$$f(x_1, x_2) = 11x_1^2 + 3x_2^2 + 6x_1x_2 - 2\sqrt{10}x_1 + 6\sqrt{10}x_2 - 22,$$

начиная движение от точки $x_1^{(0)} = 0, x_2^{(0)} = 0$.

Подготовить исходные данные на листе:

Указание: Элементы матрицы В рассчитываются с использованием функции КОРЕНЬ().

	А	В	С	Д	Е	Ф	Г
1	МЕТОД ГРАДИЕНТНОГО СПУСКА						
2				точность	$\varepsilon =$	0,0001	
3							
4	А =	22	6	В =	-6,3246	С =	-22
5		6	6		18,9737		
6							
7	$X^{(0)} =$	0					
8		0					
9							
10	шаг итераций	1					

2. Приблизиться к точке минимума с помощью таблицы вычислений:

Вид рабочего листа с результатом расчета

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	МЕТОД ГРАДИЕНТНОГО СПУСКА													
2				точность	$\varepsilon =$	0,0001								
3														
4	A =	22	6	B =	-6,3246	C =	-22							
5		6	6		18,9737									
6														
7	$X^{(0)} =$	0												
8		0												
9														
10	шаг итераций	1												
11														
12	номер итерации	0	1	2	3	4	5	6	7	8	9			
13	X	0	6,32456	0	2,10819	1,40546	1,6397	1,56162	1,58765	1,57897	1,58186			
14		0	-18,974	0	-6,3246	-4,2164	-4,9191	-4,6849	-4,7629	-4,7369	-4,7456			
15	$f(X)$	-22	378	-22	-66,444	-71,383	-71,931	-71,992	-71,999	-72	-72			
16	grad $f(X)$	-6,3246	18,9737	-6,3246	2,10819	-0,7027	0,23424	-0,0781	0,02603	-0,0087	0,00289			
17		18,9737	-56,921	18,9737	-6,3246	2,10819	-0,7027	0,23424	-0,0781	0,02603	-0,0087			
18	$\ grad f(X)\ $	20	60	20	6,66667	2,22222	0,74074	0,24691	0,0823	0,02743	0,00914			
19	шаг итераций	1	0,33333	0,33333	0,33333	0,33333	0,33333	0,33333	0,33333	0,33333	0,33333			
20	Погрешность		400	400	44,4444	4,93827	0,5487	0,06097	0,00677	0,00075	8,4E-05			
21														итерационный процесс закончен

Вид рабочего листа с формулами

	A	B	C	D	E	F	G	H	I	J
12	номер итерации	0	=B12+1							
13	X	{=B7:B8}	{=B13:B14-B19*B16:B17}							
14										
15	$f(X)$	{=0,5*МУМНОЖ(МУМНОЖ(ТРАНСП(B13:B14);\$B\$4:\$C\$5);B13:B14)+МУМНОЖ(ТРАНСП(B13:B14);\$E\$4:\$E\$5)+\$G\$4}								
16	grad $f(X)$	{=МУМНОЖ(\$B\$4:\$C\$5;B13:B14)+\$E\$4:\$E\$5}								
17										
18	$\ grad f(X)\ $	=КОРЕНЬ(СУММКВ(B16:B17))								
19	шаг итераций	=B10	=ЕСЛИ(C15<B15;B19;B19/3)							
20	Погрешность		=ABS(C15-B15)							
21			=ЕСЛИ(C20>=\$F\$2;"", "итерационный процесс закончен")							

Пример использования метода наискорейшего градиентного спуска

Алгоритм

1. Заданную квадратичную функцию представить в виде:

$$f(x_1, x_2) = \frac{1}{2} X^T A X + X^T B + C.$$

Градиент функции

$$\nabla f(x_1, x_2) = A X + B.$$

2. Шаг метода в направлении спуска (определяемый из условия минимума функции $f(X)$):

$$h = \frac{\|\nabla f(X)\|^2}{[\nabla f(X)]^T \cdot A \nabla f(X)}.$$

3. Следующее приближение к экстремуму рассчитывается по формуле:

$$X^{(k+1)} = X^{(k)} - h \nabla f(X^{(k)}).$$

4. Итерационный процесс заканчивается, как только $R^i < \varepsilon$.

Реализация в MS Excel

Вид рабочего листа с результатом расчета

	A	B	C	D	E	F	G
1	МЕТОД НАИСКОРЕЙШЕГО ГРАДИЕНТНОГО СПУСКА						
2				точность	$\varepsilon =$	0,0001	
3							
4	A =	22	6	B =	-6,3246	C =	-22
5		6	6		18,9737		
6							
7	$X^{(0)} =$	0					
8		0					
9							
10	номер итерации	0	1	2			
11	X	0	1,58114	1,58114			
12		0	-4,7434	-4,7434			
13	$f(X)$	-22	-72	-72			
14	$\text{grad} f(X)$	-6,3246	8,9E-16	8,9E-16			
15		18,9737	0	0			
16	$\ \text{grad} f(X)\ $	20	8,9E-16	8,9E-16			
17	шаг итераций	0,25	0,04545	0,04545			
18	Погрешность		50	0			
19				итерационный процесс закончен			

Вид рабочего листа с формулами

	A	B	C	D	E	F	G	H	I	J	K
10	номер итерации	0	=B10+1								
11	X	(=B7:B8)	(=B11:B12-B17*B14:B15)								
12											
13	$f(X)$	(=0,5*МУМНОЖ(МУМНОЖ(ТРАНСП(B11:B12);\$B\$4:\$C\$5);B11:B12)+МУМНОЖ(ТРАНСП(B11:B12);\$E\$4:\$E\$5)+\$G\$4)									
14	$\text{grad} f(X)$	(=МУМНОЖ(\$B\$4:\$C\$5;B11:B12)+\$E\$4:\$E\$5)									
15											
16	$\ \text{grad} f(X)\ $	=КОРЕНЬ(СУММКВ(B14:B15))									
17	шаг итераций	(=B16^2/МУМНОЖ(ТРАНСП(B14:B15);МУМНОЖ(\$B\$4:\$C\$5;B14:B15)))									
18	Погрешность		=ABS(C13-B13)								
19			=ЕСЛИ(C18>\$F\$2;"", "итерационный процесс закончен")								

Метод Гаусса–Зейделя (наискорейшего покоординатного спуска)

Алгоритм

1. Заданную квадратичную функцию представить в виде:

$$f(x_1, x_2) = \frac{1}{2} X^T A X + X^T B + C.$$

Градиент функции

$$\nabla f(x_1, x_2) = A X + B.$$

1. Направление спуска – i -ый орт пространства $e^i = (0 \dots 0 1 0 \dots 0)^T$. Шаг метода в направлении спуска (определяемый из условия минимума функции $f(X)$):

$$h = \frac{[e^i]^T e^i}{[e^i]^T \cdot A e^i}.$$

2. Следующее приближение к экстремуму рассчитывается по формуле:

$$X^{(k+1)} = X^{(k)} - h \nabla f(X^{(k)}).$$
3. Итерационный процесс заканчивается, как только $R^i < \varepsilon$.

Реализация в MS Excel

Вид рабочего листа с результатом расчета

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	МЕТОД ГАУССА-ЗЕЙДЕЛЯ															
2				точность	$\varepsilon =$	0,0001										
4	A =	22	6	B =	-6,3246	C =	-22									
5		6	6		18,9737											
7	$X^{(0)} =$	0														
8		0														
10	номер итерации	0	1	2	3	4	5	6	7	8	9	10	11			
11	X	0	0,28748	1,14992	1,22832	1,46353	1,48492	1,54906	1,5549	1,57239	1,57398	1,57875	1,57919			
12		0	-0,8624	-3,4498	-3,685	-4,3906	-4,4547	-4,6472	-4,6647	-4,7172	-4,7219	-4,7363	-4,7376			
13	$f(X)$	-22	-38,529	-68,281	-69,51	-71,723	-71,815	-71,979	-71,986	-71,998	-71,999	-72	-72			
14	$\text{grad} f(X)$	-6,3246	-5,1746	-1,7249	-1,4113	-0,4704	-0,3849	-0,1283	-0,105	-0,035	-0,0286	-0,0095	-0,0078			
15		18,9737	15,5239	5,17464	4,23379	1,41126	1,15467	0,38489	0,31491	0,10497	0,08588	0,02863	0,02342			
16	$\ \text{grad} f(X)\ $	20	16,3636	5,45455	4,46281	1,4876	1,21713	0,40571	0,33194	0,11065	0,09053	0,03018	0,02469			
17	направление	1	0	1	0	1	0	1	0	1	0	1	0			
18	спуска	0	1	0	1	0	1	0	1	0	1	0	1			
19	шаг итераций	0,04545	0,16667	0,04545	0,16667	0,04545	0,16667	0,04545	0,16667	0,04545	0,16667	0,04545	0,16667			
20	Погрешность		16,5289	29,7521	1,22942	2,21296	0,09144	0,1646	0,0068	0,01224	0,00051	0,00091	3,8E-05			
21														итерационный процесс закончен		

Вид рабочего листа с формулами

	A	B	C	D	E	F	G	H	I	J	K
10	номер итерации	0	=B10+1								
11	X	(=B7:B8)	(=B11:B12-B19*B14:B15)								
13	$f(X)$	(=0,5*МУМНОЖ(МУМНОЖ(ТРАНСП(B11:B12);\$B\$4:\$C\$5);B11:B12)+МУМНОЖ(ТРАНСП(B11:B12);\$E\$4:\$E\$5)+\$G\$4)									
14	$\text{grad} f(X)$	(=МУМНОЖ(\$B\$4:\$C\$5;B11:B12)+\$E\$4:\$E\$5)									
15											
16	$\ \text{grad} f(X)\ $	=КОРЕНЬ(СУММКВ(B14:B15))									
17	направление	1	=ЕСЛИ(B17=1;0;1)								
18	спуска	=ЕСЛИ(B17=1;0;1)	=ЕСЛИ(C17=1;0;1)								
19	шаг итераций	(=СУММПРОИЗВ(B17:B18;B17:B18)/СУММПРОИЗВ(B17:B18;МУМНОЖ(\$B\$4:\$C\$5;B17:B18)))									
20	Погрешность		=ABS(C13-B13)								
21			=ЕСЛИ(C20>\$F\$2;"", "итерационный процесс закончен")								

Использование встроенных инструментов MS EXCEL

1. Безусловная оптимизация с использованием надстройки Поиск решения MS Excel: **Сервис – Поиск решения:**

	A	B	C	D	E	F	G	H
1				точность	$\varepsilon =$	0,0001		
2								
3	A =	11	3	B =	-6,3246	C =	-22	
4		3	3		18,9737			
5								
6	X =	0						
7		0						
8								
9	$f(X)$	-22						

Поиск решения X

Установить целевую ячейку: Выполнить

Равной: максимальному значению значению: Закреть

минимальному значению

Изменяя ячейки:

Предположить

Ограничения:

Добавить
Изменить
Удалить

Параметры
Восстановить
Справка

LP5 ПоискРешения

Варианты задания

Часть 1. Линейное программирование

Решить задачу линейного программирования

1	$F(x) = 2x_1 - 5x_2 \rightarrow \max$ $\begin{cases} 2x_1 + 3x_2 \geq 6 \\ 3x_1 - x_2 \geq -1 \\ 2x_1 - 4x_2 \geq 4 \end{cases}$	2	$F(x) = -x_1 + 4x_2 \rightarrow \max$ $\begin{cases} x_2 \leq 2 \\ 2x_1 - x_2 \geq -1 \\ 2x_1 - 4x_2 \leq 4 \end{cases}$
3	$F(x) = 3x_1 + 4x_2 \rightarrow \max$ $\begin{cases} x_1 + x_2 \leq 2 \\ -x_1 + 5x_2 \geq 10 \\ x_1 \geq 2 \end{cases}$	4	$F(x) = 21x_1 + 4x_2 \rightarrow \max$ $\begin{cases} x_1 + x_2 \geq 12 \\ 2x_1 - x_2 \leq -1 \\ 2x_1 - 4x_2 \geq 4 \end{cases}$
5	$F(x) = 2x_1 + 3x_2 \rightarrow \max$ $\begin{cases} x_1 + x_2 \geq 10 \\ 8x_1 - x_2 \geq -8 \\ x_2 \leq 2 \end{cases}$	6	$F(x) = 7x_1 + 3x_2 \rightarrow \max$ $\begin{cases} 3x_1 + 4x_2 \geq 12 \\ x_1 - x_2 \geq -1 \\ 2x_1 - 4x_2 \geq 4 \end{cases}$
7	$F(x) = 3x_1 + 2x_2 \rightarrow \max$ $\begin{cases} x_1 \geq 1 \\ 8x_1 - 2x_2 \geq -8 \\ x_2 \geq 2 \end{cases}$	8	$F(x) = 3x_1 + x_2 \rightarrow \max$ $\begin{cases} x_1 \geq 1 \\ 2x_1 + x_2 \geq 8 \\ 2x_1 - x_2 \leq 4 \end{cases}$
9	$F(x) = 4x_1 + 5x_2 \rightarrow \max$ $\begin{cases} x_2 \geq 5 \\ x_1 - 2x_2 \geq -2 \\ 6x_1 - 2x_2 \geq 6 \end{cases}$	10	$F(x) = 3x_1 - 7x_2 \rightarrow \min$ $\begin{cases} x_1 + x_2 \geq 1 \\ x_1 - 2x_2 \geq -2 \\ 3x_1 - x_2 \geq 6 \end{cases}$
11	$F(x) = -x_1 - x_2 \rightarrow \min$ $\begin{cases} x_2 \leq 1 \\ x_1 - 2x_2 \geq 2 \\ 2x_1 + 3x_2 \geq 12 \end{cases}$	12	$F(x) = 6x_1 + 7x_2 \rightarrow \max$ $\begin{cases} 3x_1 + 5x_2 \geq 14 \\ 2x_1 - 4x_2 \geq -6 \\ 3x_1 - 8x_2 \leq 6 \end{cases}$

13	$F(x) = 6x_1 + 7x_2 \rightarrow \max$ $\begin{cases} 3x_1 + 5x_2 \geq 15 \\ x_1 - 4x_2 \leq -4 \\ 3x_1 - 2x_2 \geq 7 \end{cases}$	14	$F(x) = -x_1 - x_2 \rightarrow \min$ $\begin{cases} 2x_1 + 6x_2 \geq 12 \\ x_1 + x_2 \geq 2 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$
15	$F(x) = 2x_1 + 5x_2 \rightarrow \max$ $\begin{cases} 3x_1 + 5x_2 \geq 0 \\ 2x_1 - 4x_2 \leq -10 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$	16	$F(x) = x_1 + 3x_2 \rightarrow \max$ $\begin{cases} 3x_1 + 5x_2 \geq 15 \\ 2x_1 - 4x_2 \leq 10 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$
17	$F(x) = x_1 + 3x_2 \rightarrow \max$ $\begin{cases} 3x_1 - 8x_2 \leq 9 \\ x_1 - 2x_2 \geq -4 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$	18	$F(x) = 4x_1 + 3x_2 \rightarrow \max$ $\begin{cases} 3x_1 + 5x_2 \geq 16 \\ x_1 - 2x_2 \geq 7 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$
19	$F(x) = x_1 - 9x_2 \rightarrow \min$ $\begin{cases} 3x_1 + 5x_2 \geq 17 \\ 2x_1 - 4x_2 \geq -6 \\ 3x_1 - 8x_2 \leq 6 \end{cases}$	20	$F(x) = x_1 + 14x_2 \rightarrow \min$ $\begin{cases} x_2 \leq 1 \\ 2x_1 + 3x_2 \geq 13 \\ x_1 - 2x_2 \geq 2 \end{cases}$

Часть 2. Транспортная задача

Решить транспортную задачу. Заданы мощности поставщиков $a_i (i = 1, 2, 3)$, емкости потребителей $b_j (j = 1, 2, 3)$ и матрица стоимостей перевозок единицы продукции от каждого поставщика каждому потребителю. Требуется найти план перевозок, при котором суммарные транспортные затраты будут наименьшими.

1	<table border="1"> <thead> <tr> <th>b_j</th> <th>40</th> <th>120</th> <th>170</th> </tr> </thead> <tbody> <tr> <th>a_i</th> <td></td> <td></td> <td></td> </tr> <tr> <td>90</td> <td>5</td> <td>6</td> <td>8</td> </tr> <tr> <td>65</td> <td>6</td> <td>9</td> <td>10</td> </tr> <tr> <td>75</td> <td>4</td> <td>7</td> <td>5</td> </tr> </tbody> </table>	b_j	40	120	170	a_i				90	5	6	8	65	6	9	10	75	4	7	5	2	<table border="1"> <thead> <tr> <th>b_j</th> <th>25</th> <th>40</th> <th>35</th> </tr> </thead> <tbody> <tr> <th>a_i</th> <td></td> <td></td> <td></td> </tr> <tr> <td>20</td> <td>3</td> <td>6</td> <td>4</td> </tr> <tr> <td>90</td> <td>5</td> <td>9</td> <td>3</td> </tr> <tr> <td>60</td> <td>4</td> <td>8</td> <td>6</td> </tr> </tbody> </table>	b_j	25	40	35	a_i				20	3	6	4	90	5	9	3	60	4	8	6
b_j	40	120	170																																								
a_i																																											
90	5	6	8																																								
65	6	9	10																																								
75	4	7	5																																								
b_j	25	40	35																																								
a_i																																											
20	3	6	4																																								
90	5	9	3																																								
60	4	8	6																																								

3	<table border="1"> <tr> <td>b_j</td> <td>16</td> <td>20</td> <td>35</td> </tr> <tr> <td>a_i</td> <td></td> <td></td> <td></td> </tr> <tr> <td>15</td> <td>6</td> <td>7</td> <td>5</td> </tr> <tr> <td>8</td> <td>5</td> <td>6</td> <td>4</td> </tr> <tr> <td>20</td> <td>9</td> <td>10</td> <td>6</td> </tr> </table>	b_j	16	20	35	a_i				15	6	7	5	8	5	6	4	20	9	10	6	4	<table border="1"> <tr> <td>b_j</td> <td>20</td> <td>12</td> <td>8</td> </tr> <tr> <td>a_i</td> <td></td> <td></td> <td></td> </tr> <tr> <td>22</td> <td>7</td> <td>6</td> <td>3</td> </tr> <tr> <td>12</td> <td>8</td> <td>4</td> <td>2</td> </tr> <tr> <td>16</td> <td>2</td> <td>3</td> <td>1</td> </tr> </table>	b_j	20	12	8	a_i				22	7	6	3	12	8	4	2	16	2	3	1
b_j	16	20	35																																								
a_i																																											
15	6	7	5																																								
8	5	6	4																																								
20	9	10	6																																								
b_j	20	12	8																																								
a_i																																											
22	7	6	3																																								
12	8	4	2																																								
16	2	3	1																																								
5	<table border="1"> <tr> <td>b_j</td> <td>19</td> <td>31</td> <td>10</td> </tr> <tr> <td>a_i</td> <td></td> <td></td> <td></td> </tr> <tr> <td>20</td> <td>5</td> <td>8</td> <td>3</td> </tr> <tr> <td>10</td> <td>2</td> <td>4</td> <td>2</td> </tr> <tr> <td>12</td> <td>7</td> <td>6</td> <td>3</td> </tr> </table>	b_j	19	31	10	a_i				20	5	8	3	10	2	4	2	12	7	6	3	6	<table border="1"> <tr> <td>b_j</td> <td></td> <td></td> <td></td> </tr> <tr> <td>a_i</td> <td>14</td> <td>20</td> <td>22</td> </tr> <tr> <td>50</td> <td>3</td> <td>8</td> <td>9</td> </tr> <tr> <td>18</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td>12</td> <td>2</td> <td>7</td> <td>6</td> </tr> </table>	b_j				a_i	14	20	22	50	3	8	9	18	3	4	5	12	2	7	6
b_j	19	31	10																																								
a_i																																											
20	5	8	3																																								
10	2	4	2																																								
12	7	6	3																																								
b_j																																											
a_i	14	20	22																																								
50	3	8	9																																								
18	3	4	5																																								
12	2	7	6																																								
7	<table border="1"> <tr> <td>b_j</td> <td>20</td> <td>18</td> <td>17</td> </tr> <tr> <td>a_i</td> <td></td> <td></td> <td></td> </tr> <tr> <td>30</td> <td>9</td> <td>7</td> <td>4</td> </tr> <tr> <td>15</td> <td>5</td> <td>3</td> <td>2</td> </tr> <tr> <td>45</td> <td>10</td> <td>8</td> <td>5</td> </tr> </table>	b_j	20	18	17	a_i				30	9	7	4	15	5	3	2	45	10	8	5	8	<table border="1"> <tr> <td>b_j</td> <td>18</td> <td>40</td> <td>12</td> </tr> <tr> <td>a_i</td> <td></td> <td></td> <td></td> </tr> <tr> <td>32</td> <td>9</td> <td>8</td> <td>4</td> </tr> <tr> <td>15</td> <td>8</td> <td>7</td> <td>3</td> </tr> <tr> <td>7</td> <td>4</td> <td>3</td> <td>2</td> </tr> </table>	b_j	18	40	12	a_i				32	9	8	4	15	8	7	3	7	4	3	2
b_j	20	18	17																																								
a_i																																											
30	9	7	4																																								
15	5	3	2																																								
45	10	8	5																																								
b_j	18	40	12																																								
a_i																																											
32	9	8	4																																								
15	8	7	3																																								
7	4	3	2																																								
9	<table border="1"> <tr> <td>b_j</td> <td>17</td> <td>13</td> <td>25</td> </tr> <tr> <td>a_i</td> <td></td> <td></td> <td></td> </tr> <tr> <td>20</td> <td>8</td> <td>3</td> <td>6</td> </tr> <tr> <td>15</td> <td>4</td> <td>2</td> <td>5</td> </tr> <tr> <td>30</td> <td>9</td> <td>4</td> <td>7</td> </tr> </table>	b_j	17	13	25	a_i				20	8	3	6	15	4	2	5	30	9	4	7	10	<table border="1"> <tr> <td>b_j</td> <td>12</td> <td>19</td> <td>9</td> </tr> <tr> <td>a_i</td> <td></td> <td></td> <td></td> </tr> <tr> <td>18</td> <td>5</td> <td>8</td> <td>2</td> </tr> <tr> <td>22</td> <td>8</td> <td>9</td> <td>4</td> </tr> <tr> <td>15</td> <td>6</td> <td>7</td> <td>3</td> </tr> </table>	b_j	12	19	9	a_i				18	5	8	2	22	8	9	4	15	6	7	3
b_j	17	13	25																																								
a_i																																											
20	8	3	6																																								
15	4	2	5																																								
30	9	4	7																																								
b_j	12	19	9																																								
a_i																																											
18	5	8	2																																								
22	8	9	4																																								
15	6	7	3																																								
11	<table border="1"> <tr> <td>b_j</td> <td>14</td> <td>20</td> <td>30</td> </tr> <tr> <td>a_i</td> <td></td> <td></td> <td></td> </tr> <tr> <td>25</td> <td>4</td> <td>5</td> <td>9</td> </tr> <tr> <td>10</td> <td>2</td> <td>3</td> <td>3</td> </tr> <tr> <td>12</td> <td>4</td> <td>6</td> <td>8</td> </tr> </table>	b_j	14	20	30	a_i				25	4	5	9	10	2	3	3	12	4	6	8	12	<table border="1"> <tr> <td>b_j</td> <td>40</td> <td>12</td> <td>20</td> </tr> <tr> <td>a_i</td> <td></td> <td></td> <td></td> </tr> <tr> <td>17</td> <td>8</td> <td>4</td> <td>9</td> </tr> <tr> <td>30</td> <td>6</td> <td>3</td> <td>7</td> </tr> <tr> <td>15</td> <td>5</td> <td>2</td> <td>4</td> </tr> </table>	b_j	40	12	20	a_i				17	8	4	9	30	6	3	7	15	5	2	4
b_j	14	20	30																																								
a_i																																											
25	4	5	9																																								
10	2	3	3																																								
12	4	6	8																																								
b_j	40	12	20																																								
a_i																																											
17	8	4	9																																								
30	6	3	7																																								
15	5	2	4																																								
13	<table border="1"> <tr> <td>b_j</td> <td>17</td> <td>21</td> <td>8</td> </tr> <tr> <td>a_i</td> <td></td> <td></td> <td></td> </tr> <tr> <td>24</td> <td>5</td> <td>7</td> <td>4</td> </tr> <tr> <td>16</td> <td>4</td> <td>8</td> <td>3</td> </tr> <tr> <td>20</td> <td>6</td> <td>9</td> <td>4</td> </tr> </table>	b_j	17	21	8	a_i				24	5	7	4	16	4	8	3	20	6	9	4	14	<table border="1"> <tr> <td>b_j</td> <td>20</td> <td>12</td> <td>37</td> </tr> <tr> <td>a_i</td> <td></td> <td></td> <td></td> </tr> <tr> <td>15</td> <td>5</td> <td>3</td> <td>7</td> </tr> <tr> <td>10</td> <td>3</td> <td>2</td> <td>3</td> </tr> <tr> <td>24</td> <td>6</td> <td>4</td> <td>8</td> </tr> </table>	b_j	20	12	37	a_i				15	5	3	7	10	3	2	3	24	6	4	8
b_j	17	21	8																																								
a_i																																											
24	5	7	4																																								
16	4	8	3																																								
20	6	9	4																																								
b_j	20	12	37																																								
a_i																																											
15	5	3	7																																								
10	3	2	3																																								
24	6	4	8																																								

15	<table border="1" data-bbox="432 203 719 472"> <tbody> <tr> <td>$a_i \backslash b_j$</td> <td>10</td> <td>7</td> <td>18</td> </tr> <tr> <td>15</td> <td>6</td> <td>3</td> <td>7</td> </tr> <tr> <td>18</td> <td>4</td> <td>2</td> <td>9</td> </tr> <tr> <td>12</td> <td>5</td> <td>3</td> <td>8</td> </tr> </tbody> </table>	$a_i \backslash b_j$	10	7	18	15	6	3	7	18	4	2	9	12	5	3	8	16	<table border="1" data-bbox="1075 203 1350 472"> <tbody> <tr> <td>$a_i \backslash b_j$</td> <td>9</td> <td>31</td> <td>20</td> </tr> <tr> <td>20</td> <td>3</td> <td>9</td> <td>8</td> </tr> <tr> <td>14</td> <td>4</td> <td>6</td> <td>7</td> </tr> <tr> <td>12</td> <td>2</td> <td>4</td> <td>5</td> </tr> </tbody> </table>	$a_i \backslash b_j$	9	31	20	20	3	9	8	14	4	6	7	12	2	4	5
$a_i \backslash b_j$	10	7	18																																
15	6	3	7																																
18	4	2	9																																
12	5	3	8																																
$a_i \backslash b_j$	9	31	20																																
20	3	9	8																																
14	4	6	7																																
12	2	4	5																																
17	<table border="1" data-bbox="437 524 713 792"> <tbody> <tr> <td>$a_i \backslash b_j$</td> <td>20</td> <td>14</td> <td>16</td> </tr> <tr> <td>30</td> <td>5</td> <td>2</td> <td>6</td> </tr> <tr> <td>15</td> <td>2</td> <td>1</td> <td>3</td> </tr> <tr> <td>25</td> <td>4</td> <td>2</td> <td>8</td> </tr> </tbody> </table>	$a_i \backslash b_j$	20	14	16	30	5	2	6	15	2	1	3	25	4	2	8	18	<table border="1" data-bbox="1075 524 1350 792"> <tbody> <tr> <td>$a_i \backslash b_j$</td> <td>20</td> <td>10</td> <td>30</td> </tr> <tr> <td>35</td> <td>6</td> <td>3</td> <td>7</td> </tr> <tr> <td>15</td> <td>3</td> <td>2</td> <td>4</td> </tr> <tr> <td>20</td> <td>5</td> <td>4</td> <td>8</td> </tr> </tbody> </table>	$a_i \backslash b_j$	20	10	30	35	6	3	7	15	3	2	4	20	5	4	8
$a_i \backslash b_j$	20	14	16																																
30	5	2	6																																
15	2	1	3																																
25	4	2	8																																
$a_i \backslash b_j$	20	10	30																																
35	6	3	7																																
15	3	2	4																																
20	5	4	8																																
19	<table border="1" data-bbox="437 844 713 1113"> <tbody> <tr> <td>$a_i \backslash b_j$</td> <td>21</td> <td>30</td> <td>32</td> </tr> <tr> <td>16</td> <td>5</td> <td>9</td> <td>7</td> </tr> <tr> <td>32</td> <td>4</td> <td>6</td> <td>5</td> </tr> <tr> <td>20</td> <td>3</td> <td>5</td> <td>4</td> </tr> </tbody> </table>	$a_i \backslash b_j$	21	30	32	16	5	9	7	32	4	6	5	20	3	5	4	20	<table border="1" data-bbox="1075 844 1350 1113"> <tbody> <tr> <td>$a_i \backslash b_j$</td> <td>25</td> <td>19</td> <td>21</td> </tr> <tr> <td>40</td> <td>5</td> <td>3</td> <td>6</td> </tr> <tr> <td>17</td> <td>2</td> <td>1</td> <td>2</td> </tr> <tr> <td>23</td> <td>7</td> <td>4</td> <td>8</td> </tr> </tbody> </table>	$a_i \backslash b_j$	25	19	21	40	5	3	6	17	2	1	2	23	7	4	8
$a_i \backslash b_j$	21	30	32																																
16	5	9	7																																
32	4	6	5																																
20	3	5	4																																
$a_i \backslash b_j$	25	19	21																																
40	5	3	6																																
17	2	1	2																																
23	7	4	8																																

Часть 3. Нелинейное программирование

Решите задачу нелинейного программирования

$$1. f(x) = -3x_1^2 - x_2^2 + 4x_1 + 4x_2 - 2x_1x_2 \rightarrow \max$$

$$\begin{cases} 3x_1 + 4x_2 \leq 12 \\ x_1 - 2x_2 \leq 2 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$2. f(x) = -3x_1^2 - x_2^2 + 12x_1 + 4x_2 \rightarrow \max$$

$$\begin{cases} x_1 + x_2 \leq 0 \\ \frac{1}{2}x_1 - \frac{1}{2}x_2 \leq -1 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$3. f(x) = -\frac{1}{2}x_1^2 - \frac{1}{2}x_2^2 + x_1 + 2x_2 \rightarrow \max$$

$$\begin{cases} 2x_1 + 3x_2 \leq 2 \\ x_1 + 4x_2 \leq 5 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$4. f(x) = 3x_1 - 2x_2 - \frac{1}{2}x_1^2 - x_2^2 + x_1x_2 \rightarrow \max$$

$$\begin{cases} 2x_1 + 3x_2 \leq 2 \\ x_1 + 4x_2 \leq 5 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$5. f(x) = 3x_1 - 2x_2 - \frac{1}{2}x_1^2 - x_2^2 + x_1x_2 \rightarrow \max$$

$$\begin{cases} x_1 \leq 3 \\ x_2 \leq 6 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$6. f(x) = -4x_1 + 8x_2 - x_1^2 - \frac{3}{2}x_2^2 + 2x_1x_2 \rightarrow \max$$

$$\begin{cases} x_1 + x_2 \leq 3 \\ x_1 - x_2 \leq 1 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$7. f(x) = -4x_1 + 8x_2 - x_1^2 - \frac{3}{2}x_2^2 + 2x_1x_2 \rightarrow \max$$

$$\begin{cases} x_1 + x_2 \leq 1 \\ x_1 \leq 4 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$8. f(x) = -4x_1 + 8x_2 - x_1^2 - \frac{3}{2}x_2^2 + 2x_1x_2 \rightarrow \max$$

$$\begin{cases} 3x_1 + 5x_2 \leq 15 \\ x_1 - x_2 \leq 1 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$9. f(x) = 3x_1 - 2x_2 - \frac{1}{2}x_1^2 - x_2^2 + x_1x_2 \rightarrow \max$$

$$\begin{cases} -x_1 + 2x_2 \leq 2 \\ 2x_1 - x_2 \leq 2 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$10. f(x) = -x_1 + 6x_2 - x_1^2 - 3x_2^2 + 3x_1x_2 \rightarrow \max$$

$$\begin{cases} 4x_1 + 3x_2 \leq 12 \\ -x_1 + x_2 \leq 1 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$11. f(x) = -x_1 + 6x_2 - x_1^2 - 3x_2^2 + 3x_1x_2 \rightarrow \max$$

$$\begin{cases} x_1 + x_2 \leq 3 \\ -2x_1 + x_2 \leq 2 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$12. f(x) = -x_1 + 6x_2 - x_1^2 - 3x_2^2 + 3x_1x_2 \rightarrow \max$$

$$\begin{cases} x_1 - x_2 \leq 0 \\ x_2 \leq 5 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$13. f(x) = 6x_2 - x_1^2 - \frac{3}{2}x_2^2 + 2x_1x_2 \rightarrow \max$$

$$\begin{cases} 3x_1 + 4x_2 \leq 12 \\ -x_1 + x_2 \leq 12 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$14. f(x) = 6x_2 - x_1^2 - \frac{3}{2}x_2^2 + 2x_1x_2 \rightarrow \max$$

$$\begin{cases} -x_1 + 2x_2 \leq 2 \\ x_1 \leq 4 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0$$

$$15. f(x) = 6x_2 - x_1^2 - \frac{3}{2}x_2^2 + 2x_1x_2 \rightarrow \max$$

$$\begin{cases} 3x_1 + 4x_2 \leq 12 \\ -x_1 - 2x_2 \leq -2 \end{cases}$$
$$x_1 \geq 0, x_2 \geq 0$$

$$16. f(x) = 4x_1 + 4x_2 - 3x_1^2 - x_2^2 + 2x_1x_2 \rightarrow \max$$

$$\begin{cases} 3x_1 + 6x_2 \leq 18 \\ -x_1 + x_2 \leq 12 \end{cases}$$
$$x_1 \geq 0, x_2 \geq 0$$

$$17. f(x) = 8x_1 + 12x_2 - x_1^2 - \frac{3}{2}x_2^2 \rightarrow \max$$

$$\begin{cases} -x_1 + 2x_2 \leq 2 \\ x_1 \leq 6 \end{cases}$$
$$x_1 \geq 0, x_2 \geq 0$$

$$18. f(x) = 8x_1 + 12x_2 - x_1^2 - \frac{3}{2}x_2^2 \rightarrow \max$$

$$\begin{cases} -3x_1 + 2x_2 \leq 0 \\ 4x_1 + 3x_2 \leq 12 \end{cases}$$
$$x_1 \geq 0, x_2 \geq 0$$

$$19. f(x) = 3x_1 - 2x_2 - \frac{1}{2}x_1^2 - x_2^2 + x_1x_2 \rightarrow \max$$

$$\begin{cases} -2x_1 - x_2 \leq -2 \\ 2x_1 + 3x_2 \leq 6 \end{cases}$$
$$x_1 \geq 0, x_2 \geq 0$$

$$20. f(x) = 6x_1 + 4x_2 - x_1^2 - \frac{1}{2}x_2^2 - x_1x_2 \rightarrow \max$$

$$\begin{cases} x_1 + 2x_2 \leq 2 \\ -2x_1 + x_2 \leq 0 \end{cases}$$
$$x_1 \geq 0, x_2 \geq 0$$

Часть 4. Градиентные методы

Найти экстремум функции, используя метод градиентного, наискорейшего градиентного и наискорейшего покоординатного спуска, сравнить результаты

1. $f(\mathbf{x}) = 7x_1^2 + 2x_1x_2 + 5x_2^2 + x_1 - 10x_2 \rightarrow \min$

2. $f(\mathbf{x}) = 3x_1^2 - 3x_1x_2 + 4x_2^2 - 2x_1 + x_2 \rightarrow \min$

3. $f(\mathbf{x}) = x_1^2 + 2x_2^2 + e^{x_1^2 + x_2^2} - x_1 + x_2 \rightarrow \min$

4. $f(\mathbf{x}) = \sqrt{x_1^2 + x_2^2 + 1} + \frac{1}{2}x_1 - \frac{1}{2}x_2 \rightarrow \min$

5. $f(\mathbf{x}) = x_1^2 + 4x_1x_2 + 13x_1^2 + 5x_2 \rightarrow \min$

6. $f(\mathbf{x}) = 5x_1^2 - 4x_1x_2 + 5x_2^2 - x_1 - x_3 \rightarrow \min$

7. $f(\mathbf{x}) = x_1^4 + 2x_2^4 + x_1^2x_2^2 + 2x_1 + x_2 \rightarrow \min$

8. $f(\mathbf{x}) = x_1^2 + 3x_2^2 + \cos(x_1 + x_2) \rightarrow \min$

9. $f(\mathbf{x}) = \sqrt{1 + 2x_1^2 + x_2^2} + e^{x_1^2 + 2x_2^2} - x_1 - x_2 \rightarrow \min$

10. $f(\mathbf{x}) = x_1 + 5x_2 + e^{x_1^2 + x_2^2} \rightarrow \min$

11. $f(\mathbf{x}) = 4x_1^2 + 4x_1x_2 + 6x_2^2 - 17x_1 \rightarrow \min$

12. $f(\mathbf{x}) = 2x_1^2 - 2x_1x_2 + 3x_2^2 + x_1 - 3x_2 \rightarrow \min$

13. $f(\mathbf{x}) = 10x_1^2 + 3x_1x_2 + x_2^2 + 10x_2 \rightarrow \min$

14. $f(\mathbf{x}) = x_1^2 - 2x_1x_2 + 6x_1^2 + x_1 - x_2 \rightarrow \min$

15. $f(\mathbf{x}) = x_1^4 + x_2^4 + \sqrt{2 + x_1^2 + x_2^2} - 2x_1 + 3x_2 \rightarrow \min$

16. $f(\mathbf{x}) = 2x_1^2 + 3x_2^2 - 2\sin\left(\frac{x_1 - x_2}{2}\right) + x_2 \rightarrow \min$

17. $f(\mathbf{x}) = \ln(1 + 3x_1^2 + 5x_2^2) + \cos(x_1 - x_2) \rightarrow \min$

18. $f(\mathbf{x}) = x_1^2 + e^{x_1^2 + x_2^2} + 4x_1 + 3x_2 \rightarrow \min$

19. $f(\mathbf{x}) = x_1 + 2x_2 + 4\sqrt{1 + x_1^2 + x_2^2} \rightarrow \min$

$$20. f(\mathbf{x}) = 2x_1 - 5x_2 + e^{x_1^2 + \frac{1}{2}x_2^2} \rightarrow \min$$

Часть 5. Решение задачи оптимизации

Постройте математическую модель задачи и решите ее с помощью MS Excel.

Задача 1

На швейной фабрике для изготовления четырёх видов изделий может быть использована ткань трёх артикулов. Нормы расхода тканей всех артикулов на пошив одного изделия приведены в таблице. В ней так же указаны имеющиеся в распоряжении фабрики общее количество тканей каждого артикула и цена изделия данного вида. Определить, сколько изделий каждого вида должна произвести фабрика, чтобы стоимость изготовленной продукции была максимальной. Сколько ткани каждого из артикулов может сэкономить фабрика не теряя прибыли? Насколько минимально нужно поднять цену на четвертое изделие, чтобы это увеличило прибыль? Что произойдет с прибылью, если фабрике будет необходимо выпускать изделие 3 в количестве не меньше 5 штук?

Артикул ткани	Норма расхода ткани (м) на одно изделие вида				Общее количество ткани
	1	2	3	4	
I	1	-	2	1	180
II	-	1	3	2	210
III	4	2	-	4	800
Цена одного изделия (руб.)	9±2	6	4±3	7	

Задача 2

Предприятие выпускает четыре вида продукции и использует три типа основного оборудования: токарное, фрезерное и шлифовальное. Затраты времени на изготовление единицы продукции для каждого из типов оборудования приведены в таблице. В ней же указаны общий фонд рабочего времени каждого из типов оборудования, а также прибыль от реализации одного изделия данного вида. Определить такой объем выпуска каждого из изделий, при котором общая прибыль от их реализации является максимальной. Что произойдет с общей прибылью, если прибыль от продажи продукции вида 4 вырастет втрое? Как изменится общая прибыль, если предприятию будет необходимо выпускать не меньше 1 единиц продукции 3? Если увеличить время использования фрезерных станков на 80 станко-часов, то можно ли будет уменьшить время использования других станков?

Тип оборудования	Затраты времени (станко-ч) на единицу продукции вида				Общий фонд рабочего времени (станко-ч)
	1	2	3	4	
Токарное	2	1	1	3	300
Фрезерное	1	-	2	1	70
Шлифовальное	1	2	1	-	340
Прибыль от реализации единицы продукции (руб.)	8	3	2±1	1	

Задача 3

Для перевозок груза на трёх линиях могут быть использованы суда трёх типов. Производительность судов при использовании их на различных линиях характеризуется данными, приведёнными в таблице. В ней же указаны общее время, в течение которого суда каждого типа находятся в эксплуатации, и минимально необходимые объёмы перевозок на каждой линии. Определить, какие суда, на какой линии и в течение какого времени следует использовать, чтобы обеспечить максимальную загрузку судов с учётом возможного времени их эксплуатации. Как изменится общий объем перевозок, если производительность судов вида III на третьей линии возрастет вдвое, а на второй — уменьшится до 8 млн. тонномиль в сутки? Возможно ли в этом случае выполнить заданный объем перевозок? На какой линии выгоднее всего использовать суда вида I?

Тип судна	Производительность судов (млн. тонномиль в сутки) на линии			Общее время эксплуатации и судов
	1	2	3	
I	8	14	11	300
II	6	15	13	300
III	12	12	4	300
Заданный объём перевозок (млн. Тонно-миль)	3000	5400	3300	

Задача 4

Компания "Bermuda Paint" специализируется на производстве технических лаков. Представленная ниже таблица содержит информацию о ценах продажи и соответствующих издержках производства единицы полировочного и матового лаков.

Лак	Цена продажи 1 галлона, ф. ст.	Издержки производства 1 галлона, ф. ст
Матовый	13,0	9,0
Полировочный	16,0	10,0

Для производства 1 галлона матового лака необходимо затратить 6 мин. трудозатрат, а для производства одного галлона полировочного лака — 12 мин. Резерв фонда рабочего времени составляет 400 чел.-ч. в день. Размер ежедневного запаса необходимой химической смеси равен 100 унциям, тогда как ее расход на один галлон матового и полировочного лаков составляет 0,05 и 0,02 унции соответственно. Технологические возможности завода позволяют выпускать не более 3000 галлонов лака в день.

В соответствии с соглашением с основным оптовым покупателем компания должна поставлять ему 5000 галлонов матового лака и 2500 галлонов полировочного лака за каждую рабочую неделю (состоящую из 5 дней). Кроме того, существует профсоюзное соглашение, в котором оговаривается минимальный объем производства в день, равный 2000 галлонов. Администрации данной компании необходимо определить ежедневные объемы производства каждого вида лаков, которые позволяют получать максимальный общий доход.

Требуется:

а) Определить ежедневный оптимальный план производства и соответствующую ему величину дохода.

б) Для исходной задачи (не учитывающей сверхурочные работы) определить промежуток изменений показателя единичного дохода за 1 галлон полировочного лака, в котором исходное оптимальное решение остается прежним.

Задача 5

Найти решение, состоящее в определении плана изготовления изделий А, В и С, обеспечивающего максимальный их выпуск, в стоимости выраженной с учётом ограничений на возможное использование сырья трёх видов. Нормы расхода сырья каждого вида на одно изделие, цена одного изделия соответствующего вида, а также имеющегося сырья, приведены в таблице. Можно ли сэкономить сырье не уменьшая общей прибыли? Что произойдет с прибылью, если перед предприятием поставлена задача выпустить не менее пяти изделий вида А?

Вид сырья	Нормы затрат (кг) на одно изделие			Общее количество сырья (кг)
	A	B	C	
I	18	15	12	360
II	6	4	8	192
III	5	3	3	180
Цена одного изделия (руб.)	9±1	10±2	16	-

Задача 6

Полиграфическая компания выпускает рекламные издания LinksLetter и Ragged Edge, которые покупатели могут брать в местных магазинах и ресторанах. Компания получает доход, продавая место для размещения рекламы в своих изданиях. Стоимость LinksLetter составляет \$50 за тысячу экземпляров, а стоимость Ragged Edge — \$100 за тысячу экземпляров. Чтобы напечатать тысячу экземпляров LinksLetter требуется один час, а печать тысячи экземпляров Ragged Edge занимает всего полчаса. На следующей неделе ресурс времени печати составит 120 ч. Обе рекламные газеты складываются фальцевальной машиной, ресурс рабочего времени которой составляет 200 ч в неделю, причем она складывает обе газеты с одинаковой скоростью 1000 экземпляров в час. Компания хочет полностью использовать время печатного станка, минимизировав при этом затраты на производство печатной продукции. Определите оптимальный производственный план и его минимальную стоимость.

Предположим, что цели менеджера полиграфической компании изменились. Теперь он решил максимизировать получаемую от публикаций прибыль. Он определил, что прибыль от тысячи экземпляров LinksLetter составляет \$25, а от тысячи экземпляров Ragged Edge — \$45. Необходимо напечатать не менее 60000 экземпляров LinksLetter и не менее 30000 экземпляров Ragged Edge. Ограничения на ресурс рабочего времени печатного станка и фальцевальной машины остаются прежними. Каким будет оптимальный производственный план? Какие ограничения являются связывающими?

Задача 7

Завод может производить пять различных продуктов в произвольном соотношении. В выпуске каждого продукта принимают участие три станка, как показано в таблице. Все цифры даны в минутах на фунт продукта.

Продукт	Время работы станка, мин/фунт		
	1	2	3
A	12	8	5
B	7	9	10
C	8	4	7
D	10	0	3
E	7	11	2

Ресурс рабочего времени каждого станка составляет 128 ч в неделю. Все продукты конкурентоспособны и все их произведенное количество может быть продано по цене \$5, \$4, \$5, \$4 и \$4 за фунт продукта A, B, C, D и E соответственно. Переменные затраты на зарплату составляют \$4 в час для станков 1 и 2 и \$3 в час для станка 3. Стоимость материалов, затраченных на выпуск каждого фунта продуктов A и C, составляет \$2, а продуктов B, D и E — \$1. Руководство хочет максимизировать прибыль компании. Сколько часов отработает каждый станок, и в каких единицах измеряются теневые цены для ограничений, задающих ресурс рабочего времени для станков? Какую цену фирма может позволить себе заплатить за получение дополнительного часа рабочего времени станка 2? На сколько может увеличиться цена продажи продукта A, прежде чем изменится оптимальный производственный план?

Задача 8

На ткацкой фабрике для изготовления трёх артикулов ткани используются станки двух типов, пряжи и красители. В таблице указаны производительность станка каждого типа, нормы расхода пряжи и красителей, цена 1 метра ткани данного артикула, а также общий фонд рабочего времени станков каждого типа, имеющихся в распоряжении фабрики, фонды пряжи и красителей и ограничения на возможный выпуск тканей данного артикула.

Ресурсы	Нормы затрат на 1 м ткани артикула			Общее количество ресурсов
	1	2	3	

Производительность станков (станко-ч):				
I типа	0,02	-	0,04	200
II типа	0,04	0,03	0,01	500
Пряжа (кг)	1,0	1,5	2,0	15000
Красители (кг)	0,03	0,02	0,025	450
Цена 1м ткани (руб.)	5	8	8	-
Выпуск ткани (м):				
Минимальный	1000	2000	2500	-
Максимальный	2000	9000	4000	-

Составить такой план изготовления тканей, согласно которому будет произведено возможное количество тканей данного артикула, а общая стоимость всех тканей максимальна. Можно ли будет при этом сэкономить ресурсы пряжи и красителей? Будут ли полностью загружены станки?

Задача 9

Машиностроительное предприятие для изготовления четырёх видов продукции использует токарное, фрезерное, сверлильное, расточное и шлифовальное оборудование, а также комплектующие изделия.

Кроме того, сборка изделий требует выполнения определённых сборочно-наладочных работ. Нормы затрат всех видов на изготовление каждого из изделий приведены в таблице. В этой же таблице указаны наличный фонд каждого из ресурсов, прибыль от реализации единицы продукции данного вида, а также ограничения на возможный выпуск продукции 2-го и 3-го вида.

Найти план выпуска продукции, при котором прибыль от её реализации является максимальной.

Ресурсы	Нормы затрат на изготовление одного изделия				Общий объём ресурсов
	1	2	3	4	

Производительность оборудования (человек-ч):					
Токарного	550	-	620	-	64270
Фрезерного	40	30	20	20	4800
Сверлильного	86	110	150	52	22360
Расточного	160	92	158	128	26240
Шлифовального	-	158	30	50	7900
Комплекующие изделия (шт)	3	4	3	3	520
Сборочно-наладочные работы (человек-ч)	4,5	4,5	4,5	4,5	720
Прибыль от реализации одного изделия (руб.)	315	278	573	370	-
Выпуск (шт.):					
Минимальный	-	40	-	-	-
Максимальный	-	-	120	-	-

Найти план выпуска продукции, при котором прибыль от ее реализации является максимальной. Можно ли будет при этом сэкономить комплекующие изделия? Будет ли полностью загружено оборудование?

Задача 10

Для поддержания нормальной жизнедеятельности человеку необходимо потреблять не менее 118 г белков, 56 г жиров, 500 г углеводов, 8 г минеральных солей. Количество питательных веществ, содержащихся в 1 кг каждого вида потребляемых продуктов, а также цена 1 кг каждого из этих продуктов приведены в следующей таблице:

Питательные вещества	Содержание (г) питательных веществ в 1 кг продуктов						
	Мясо	рыба	молоко	Масло	сыр	крупа	картофель
Белки	180	190	30	10	260	130	21
Жиры	20	3	40	865	310	30	2
Углеводы	-	-	50	6	20	650	200
Минеральные соли	9	10	7	12	60	20	10
Цена 1 кг продуктов (руб.)	1,8	1,0	0,28	3,4	2,9	0,5	0,1

Составить дневной рацион, содержащий не менее минимальной суточной нормы потребности человека в необходимых питательных веществах при минимальной общей стоимости потребляемых продуктов.

Задача 11

Для перевозок трёх видов продукции предприятие использует два типа технологического оборудования и два вида сырья. Нормы затрат сырья и времени на изготовление одного изделия каждого вида приведены в таблице. В ней же указаны общий фонд рабочего времени каждой из групп технологического оборудования, объёмы имеющегося сырья каждого вида, а также цена одного изделия данного вида и ограничения на возможный выпуск каждого из изделий.

Ресурсы	Нормы затрат на одно изделие вида			Общее количество ресурсов
	1	2	3	
Производительность оборудования (норм-ч):				
I типа	2	-	4	200
II типа	4	3	1	500
Сырьё (кг):				
1-го вида	10	15	20	1495
2-го вида	30	20	25	4500
Цена одного изделия (руб.)	10	15	20	-
Выпуск (шт.):				
Минимальный	10	20	25	-
Максимальный	20	40	100	-

Составить такой план производства продукции, согласно которому будет изготовлено необходимое количество изделий каждого вида, а общая стоимость всей изготавливаемой продукции максимальна. Можно ли будет при этом получить экономию сырья? Будет ли полностью загружено оборудование? Что произойдет с величиной прибыли, если цену на изделие 1 увеличить на 20%?

Задача 12

При производстве четырёх видов кабеля выполняется пять групп технологических операций. Нормы затрат на 1 км кабеля данного вида на каждой из групп операций,

прибыль от реализации 1 км каждого вида кабеля, а также общий фонд рабочего времени, в течение которого могут выполняться эти операции, указаны в таблице.

Технологическая операция	Нормы затрат времени (ч) на обработку 1 км кабеля вида				Общий фонд рабочего времени (ч)
	1	2	3	4	
Волочение	1,2	1,8	1,6	2,4	7200
Наложение изоляции	1,0	0,4	0,8	0,7	5600
Скручивание элементов в кабель	6,4	5,6	6,0	8,0	11176
Освинцевание	3,0	-	1,8	2,4	3600
Испытание и контроль	2,1	1,5	0,8	3,0	4200
Прибыль от реализации 1 км кабеля	1,2	0,8	1,0	1,3	-

Определить такой план выпуска кабеля, при котором общая прибыль от реализации изготавливаемой продукции является максимальной. Кабель какого вида производить выгоднее всего?

Задача 13

На мебельной фабрике изготавливается пять видов продукции: столы, шкафы, диваны-кровати, кресла-кровати и тахты. Нормы затрат труда, а также древесины и ткани на производство единицы продукции данного вида приведены в таблице.

Ресурсы	Норма расхода ресурса на единицу продукции					Общее количество ресурсов
	стол	шкаф	диван-кровать	кресло-кровать	тахта	
Трудозатраты (человека-ч)	4	8	12	9	10	3456
Древесина (м ³)	0,4	0,6	0,3	0,2	0,3	432
Ткань (м)	-	-	6	4	5	2400
Прибыль от реализации одного изделия (руб.)	8	10	16	14	12	-

Выпуск (шт.):						
Минимальный	120	90	20	40	30	-
Максимальный	480	560	180	160	120	-

В этой же таблице указана прибыль от реализации одного изделия каждого вида, приведено общее количество ресурсов данного вида, имеющееся в распоряжении фабрики, а также указано (на основе изучения спроса), в пределах каких объёмов может изготавливаться каждый вид продукции.

Определить план производства продукции мебельной фабрикой, согласно которому прибыль от её реализации является максимальной. Можно ли при этом будет сэкономить древесину и ткань?

Задача 14

Из четырех видов сырья необходимо составить смесь, в состав которой должно входить не менее 26 ед. химического вещества А, 30 ед. – вещества В и 24 ед. – вещества С. Количество единиц химического вещества, содержащегося в 1 кг сырья каждого вида, указано в таблице. В ней же приведена цена 1 кг сырья каждого вида.

Составить смесь, содержащую не менее необходимого количества данного вида и имеющую минимальную стоимость.

Вещество	Количество единиц вещества, содержащегося в 1 кг сырья вида			
	1	2	3	4
А	1	1	-	4
В	2	-	3	5
С	1	2	4	6
Цена 1 кг сырья (руб.)	5	6	7	8

Задача 15

Для производства двух видов изделий А и В используется токарное, фрезерное и шлифовальное оборудование. Нормы затрат времени для каждого из типов оборудования на одно изделие данного вида приведены в таблице. В ней же указан общий фонд рабочего времени каждого из типов оборудования, а также прибыль от реализации одного изделия.

Тип оборудования	Затраты времени (станко-часов) на обработку одного изделия	Общий фонд полезного рабочего времени
------------------	--	---------------------------------------

	A	B	
Фрезерное	10	8	168
Токарное	5	10	180
Шлифовальное	6	12	144
Прибыль от реализации одного изделия (руб.)	14	18	

Найти план выпуска изделий вида А и В, обеспечивающий максимальную прибыль от их реализации.

Задача 16

На звероферме могут выращиваться черно-бурые лисицы и песцы. Для обеспечения нормальных условий их выращивания используется три вида кормов. Количество корма каждого вида, которое должны ежедневно получать лисицы и песцы, приведено в таблице. В ней же указаны общее количество корма каждого вида, которое может быть использовано зверофермой, и прибыль от реализации одной шкурки лисицы и песца.

Найти оптимальное соотношение количества кормов и численности поголовья лис и песцов.

Вид корма	Количество единиц корма, которое ежедневно должны получать		Общее количество корма
	A	B	
Вид 1	2	3	180
Вид 2	4	1	240
Вид 3	6	7	426
Прибыль от реализации одной шкурки (руб.)	16	12	

Задача 17

Для изготовления различных изделий А, В и С предприятие использует три различных вида сырья. Нормы расхода сырья на производство одного изделия каждого вида, цена одного изделия А, В и С, а также общее количество сырья каждого вида, которое может быть использовано предприятием, приведены в таблице:

Вид сырья	Норма затрат сырья (кг) на одно изделие			Общее количество сырья (кг)
	A	B	C	
Вид 1	18	15	12	360
Вид 2	6	4	8	192
Вид 3	5	3	3	180
Цена одного изделия (руб.)	9	10	16	

Изделия А, В и С могут производиться в любых соотношениях (сбыт обеспечен), но производство ограничено выделенным предприятию сырьем каждого вида.

Составить план производства изделий, при котором общая стоимость всей произведенной предприятием продукции является максимальной.

Задача 18

На швейной фабрике для изготовления четырех видов изделий может быть использована ткань трех артикулов. Нормы расхода тканей всех артикулов на пошив одного изделия приведены в таблице. В ней же указаны имеющиеся в распоряжении фабрики общее количество тканей каждого артикула и цена одного

изделия данного вида. Определить, сколько изделий каждого вида должна произвести фабрика, чтобы стоимость изготовленной продукции была максимальной.

Артикул ткани	Норма расхода ткани (м) на одно изделие вида				Общее количество ткани (м)
	Вид 1	Вид 2	Вид 3	Вид 4	
Артикул 1	1	-	2	1	180
Артикул 2	-	1	3	2	210
Артикул 3	4	2	-	4	800
Цена одного изделия (руб.)	9	6	4	7	

Задача 19

Фабрика "GRM pie" выпускает два вида каш для завтрака — "Crunchy" и "Chewy". Используемые для производства обоих продуктов ингредиенты в основном одинаковы и, как правило, не являются дефицитными. Основным ограничением, накладываемым на объем выпуска, является наличие фонда рабочего времени в каждом из трех цехов фабрики.

Управляющему производством Джою Дисону необходимо разработать план производства на месяц. В приведенной ниже таблице указаны общий фонд рабочего времени и число человеко-часов, требуемое для производства 1 т продукта.

Цех	Необходимый фонд рабочего времени, чел.-ч/г		Общий фонд рабочего времени, чел.-ч. в месяц
	"Crunchy"	"Chewy"	
А. Производство	10	4	1000
В. Добавка приправ	3	2	360
С. Упаковка	2	5	600

Доход от производства 1 т "Crunchy" составляет 150 ф. ст., а от производства "Chewy" — 75 ф. ст. На настоящий момент нет никаких ограничений на возможные объемы продаж. Имеется возможность продать всю произведенную продукцию.

Требуется: сформулировать модель линейного программирования, максимизирующую общий доход фабрики за месяц и реализовать решение этой модели.

Задача 20

Нефтяная компания "РТ" для улучшения эксплуатационных качеств и снижения точки замораживания дизельного топлива, которое она производит, добавляет в него определенные химикаты. В каждом бензобаке объемом 1000 л должно содержаться не менее 40 мг химической добавки X, не менее 14 мг химической добавки Y и не менее 18 мг химической добавки Z. Необходимые химические добавки в форме готовых смесей поставляют "РТ" две химические компании А и В. В нижеследующей таблице приведено содержание химических добавок в каждом продукте, поставляемом указанными компаниями.

Продукт	Химические добавки, мг/л		
	X	Y	Z
А	4	2	3
В	5	1	1

Стоимость продукта А — 1,50 ф. ст. за 1 л, а продукта В — 3,00 ф. ст. за 1 л.

Требуется: найти ассортиментный набор продуктов А и В, минимизирующий общую стоимость добавленных в топливо химикатов.

Лабораторная работа №4

Тема: одномерная оптимизация

Цель работы: получение практических навыков оптимизации.

Задание: постройте согласно варианту программу, которая позволяет решать задачу одномерной оптимизации (для написания программы можно использовать любой язык программирования высокого уровня). Интервал неопределенности найти с помощью эвристического алгоритма Свенна (на этом интервале функция должна быть унимодальной). Программа должна выводить график функции на выбранном интервале.

Отчет по лабораторной работе должен содержать:

1. Фамилию и номер группы учащегося, задание
2. Описание поиска интервала
3. Результат выполнения программы.
4. Код программы.

Основные теоретические сведения

Оптимизация функции одной переменной - наиболее простой тип оптимизационных задач. В методах одномерной оптимизации вместо $X=R$ рассматривается отрезок $X=[a,b]$, содержащей искомое решение X_{\min} . Такой отрезок называется отрезком неопределенности, или отрезком локализации. Целевая функция $f(x)$ является унимодальной.

Функция $f(x)$ называется унимодальной на $X=[a,b]$, если существует единственная точка X_{\min} , в которой функция достигает минимума, причем слева от точки функция строго убывает, а справа строго возрастает.

Метод установления границ начального отрезка локализации минимума (алгоритм Свенна).

Шаг 1. Выбрать произвольную начальную точку $x^{(0)}$ и Δ – начальный положительный шаг.

Шаг 2. Вычислить $f(x^{(0)}), f(x^{(0)} + \Delta)$

Шаг 3. Сравнить $f(x^{(0)}), f(x^{(0)} + \Delta)$:

а) если $f(x^{(0)}) > f(x^{(0)} + \Delta)$ то, согласно предположению об унимодальности функции, точка минимума должна лежать правее, чем точка $x^{(0)}$. Положить $a = x^{(0)}, x^{(1)} = x^{(0)} + \Delta, f(x^{(1)}) = f(x^{(0)} + \Delta)$, $k=2$ и перейти на шаг 5.

б) если $f(x^{(0)}) \leq f(x^{(0)} + \Delta)$, то вычислить $f(x^{(0)} - \Delta)$.

Шаг 4. Сравнить $f(x^{(0)} - \Delta), f(x^{(0)})$:

а) если $f(x^{(0)} - \Delta) \geq f(x^{(0)})$, то точка минимума лежит между точками $x^{(0)} - \Delta$ и $x^{(0)} + \Delta$, которые и

образуют границы начального отрезка локализации минимума. Положить $a = x^{(0)} - \Delta$, $b = x^{(0)} + \Delta$ и завершить поиск.

б) если $f(x^{(0)} - \Delta) < f(x^{(0)})$ то, согласно предположению об унимодальности функции, точка минимума должна лежать левее, чем точка $x^{(0)}$. Положить $b = x^{(0)}$, $x^{(1)} = x^{(0)} - \Delta$, $f(x^{(1)}) = f(x^{(0)} - \Delta)$, $\Delta = -\Delta$, $k=2$ и перейти на шаг 5.

Шаг 5. Вычислить $x^{(k)} = x^{(0)} + 2^{k-1} \cdot \Delta$, $f(x^{(k)})$.

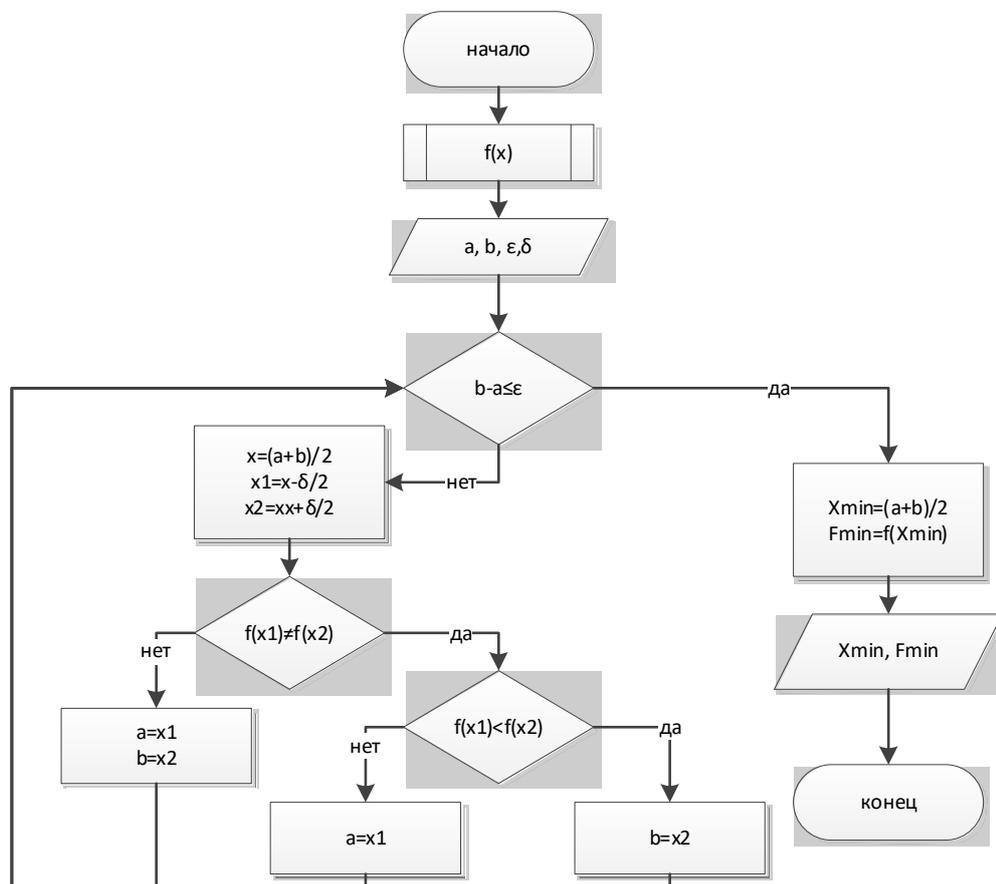
Шаг 6. Сравнить $f(x^{(k)})$, $f(x^{(k-1)})$:

а) если $f(x^{(k-1)}) \leq f(x^{(k)})$, то при $\Delta > 0$ положить $b = x^{(k)}$ при $\Delta < 0$ положить $a = x^{(k)}$ и завершить поиск.

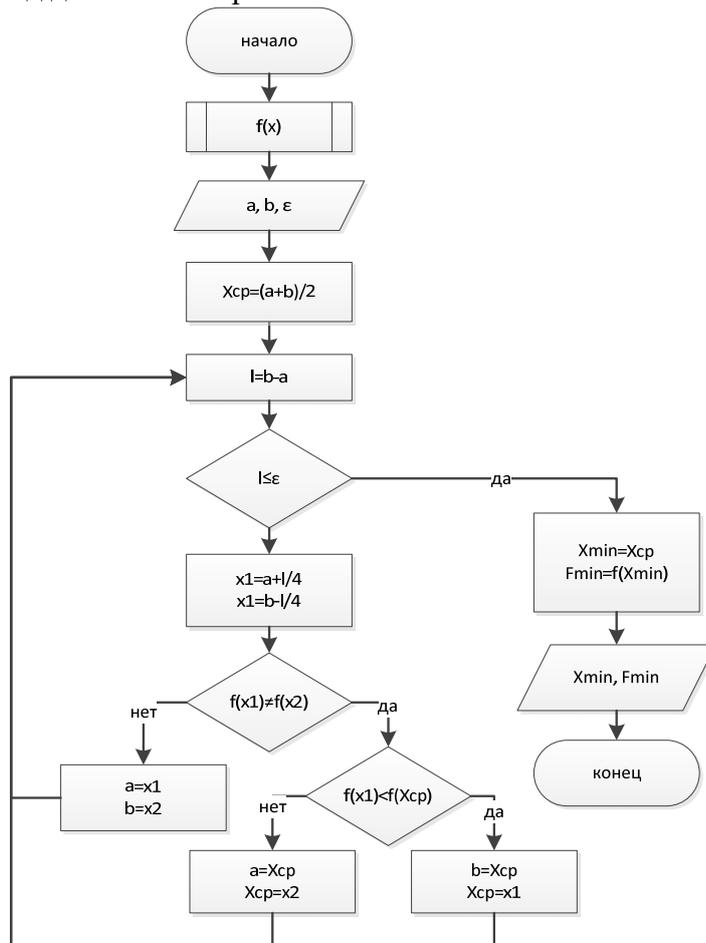
б) если $f(x^{(k-1)}) > f(x^{(k)})$, то при $\Delta > 0$ положить $a = x^{(k-1)}$ при $\Delta < 0$ положить $b = x^{(k-1)}$ и положить $k=k+1$ и перейти на шаг 5.

Для поиска точки X_{\min} с заданной точностью ϵ используют различные методы:

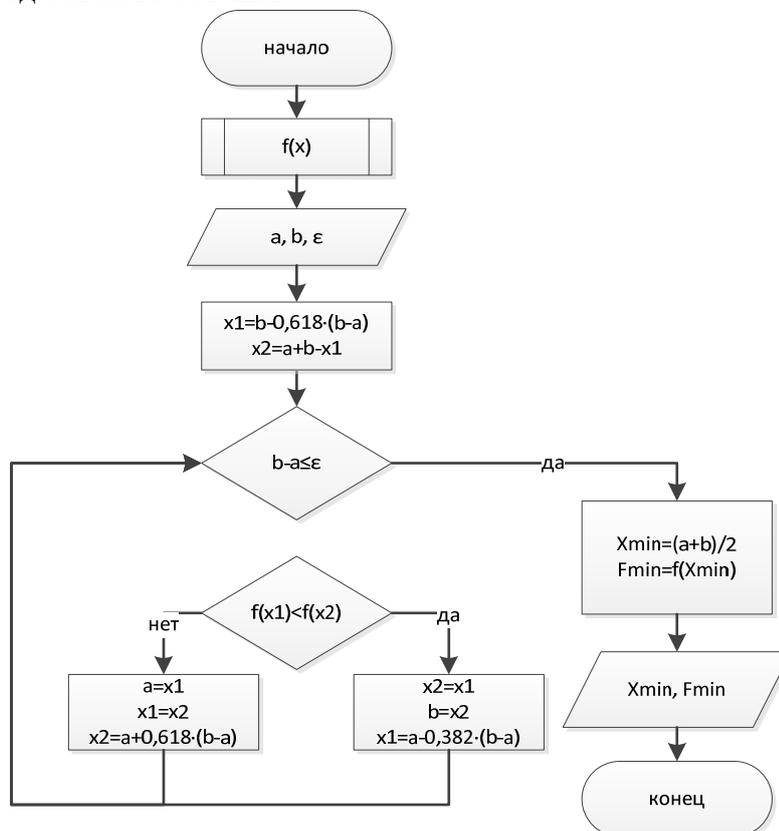
1) метод дихотомии



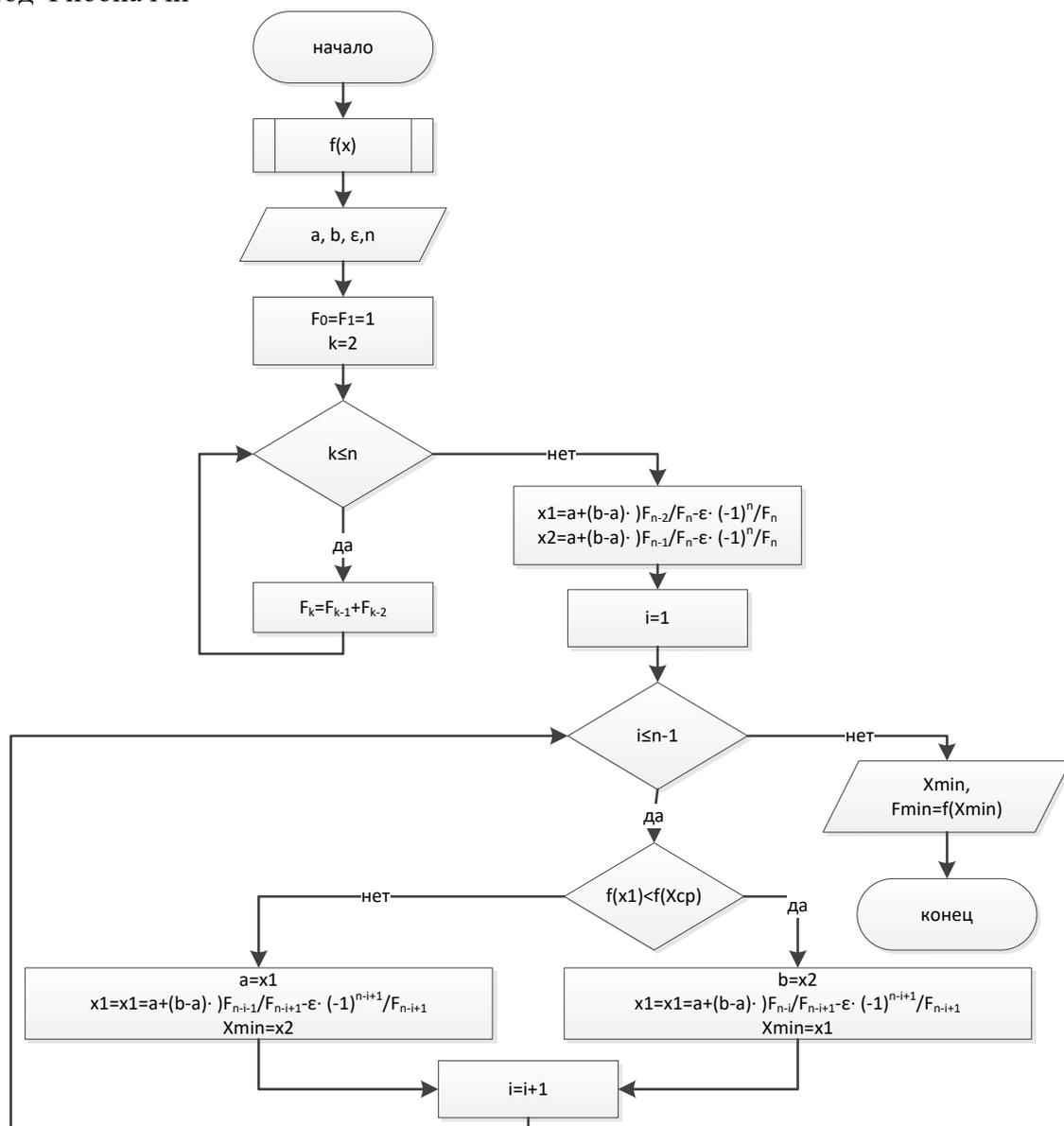
2) метод деления интервала пополам



3) метод золотого сечения



4) метод Фибоначчи



Варианты заданий

Вариант	Метод	Функция
1	Метод дихотомии	$f(x) = 2x^2 + \frac{16}{x}$
2	Метод деления интервала пополам	$f(x) = \frac{127}{4}x^2 - \frac{61}{4}x + 2$
3	Метод золотого сечения	$f(x) = 2x^2 - 12x$
4	Метод Фибоначчи	$f(x) = (500 - 2x)^2$
5	Метод дихотомии	$f(x) = 2x^3 - (x-2) \cdot x$
6	Метод деления интервала пополам	$f(x) = \frac{215}{4}x^3 - \frac{1}{2}x + 4 \cdot 3 + x - 2^x $
7	Метод золотого сечения	$f(x) = 2x^3 - (x-2) \cdot x$
8	Метод Фибоначчи	$f(x) = \frac{127}{4} x-1 ^2 - \frac{34}{4}\lceil x-0,4 \rceil + 2$
9	Метод дихотомии	$f(x) = 2x^2 - 12x$
10	Метод деления интервала пополам	$f(x) = 67x^3 - (x-4) \cdot x$
11	Метод золотого сечения	$f(x) = \frac{127}{4}x^2 - \frac{61}{4}\lfloor x-0,3 \rfloor + 2$
12	Метод Фибоначчи	$f(x) = 2\lfloor x-0,7 \rfloor^2 + \frac{16}{x}$
13	Метод дихотомии	$f(x) = (50 - 13x)^2$
14	Метод деления интервала пополам	$f(x) = \frac{25}{8}x^2 - \frac{1}{3}x + 4 \cdot x-110 $
15	Метод золотого сечения	$f(x) = 6x^3 - (x-2) \cdot x$
16	Метод Фибоначчи	$f(x) = 45x^3 - (x-2) \cdot x$
17	Метод дихотомии	$f(x) = \frac{215}{4}x^3 - \frac{1}{2}x + 4 \cdot x $
18	Метод деления интервала пополам	$f(x) = (50 + x^2 - 2x)^2$
19	Метод золотого сечения	$f(x) = 2\lceil x-0,3 \rceil^2 + \frac{16}{x}$
20	Метод Фибоначчи	$f(x) = (500 - 2x)^2$