

ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ В ПЕРЕРАБОТКЕ НЕФТИ И ГАЗА

*Методические указания по выполнению курсовой работы
для студентов специальности 15,03,02*

САНКТ-ПЕТЕРБУРГ
2017

Министерство образования и науки Российской Федерации
Федеральное бюджетное образовательное учреждение высшего образования
Санкт-Петербургский горный институт

Кафедра информатики и компьютерных технологий

ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ В ПЕРЕРАБОТКЕ НЕФТИ И ГАЗА

*Методические указания по выполнению курсовой работы для
студентов специальности 15.03.02*

**САНКТ-ПЕТЕРБУРГ
2017**

УДК 519.86:622.3.012 (075.83)

Информационные процессы в переработке нефти и газа: Методические указания по выполнению курсовой работы для студентов специальности 15,03,02/ Санкт-Петербургский горный университет. Сост.: Т.Р. Косовцева, В.В.Беляев, 60 с.

Методические указания предназначены для студентов специальности 15,03,02 «Оборудование нефтегазопереработки» дневной формы обучения и содержат задания для курсовой работы, необходимые теоретические сведения для решения задач из предметной области и рекомендации по использованию математического пакета MathCad и табличного процессора MS Excel.

Табл. 2. Рис.15. Библиогр.: 10 назв.

Научный редактор - доц. А.Б. Маховиков.

© Санкт-Петербургский горный университет,
2017 г.

ВВЕДЕНИЕ

Целью курсовой работы по дисциплине «Информационные технологии» является углубление знаний по информатике и программированию, полученных студентами при изучении дисциплины на I курсе. Курсовая работа дает возможность студенту овладеть основными принципами построения алгоритмов, методами вычислений и их реализации на персональном компьютере (ПК), приобрести навыки постановки задач, построения математических моделей, получения физических закономерностей при обработке экспериментальных данных и их анализ а также получить представление о применении персонального компьютера и наиболее распространенных пакетов программ при решении задач из предметной области.

Применение математических моделей, их реализация на ПК позволяет проанализировать наиболее существенные взаимосвязи различных показателей, получить оптимальное решение и сравнить его с другими, наметить пути устранения недостатков и показать, к каким качественно новым выводам можно прийти, используя математические модели и ПК.

Из курса информатики известно, что весь процесс получения результатов с применением ПК требует значительных усилий и умения планировать свои действия. Применение ПК позволяет сократить работу, затрачиваемую на вычисления, увеличить количество рассматриваемых вариантов с целью выбора оптимального решения, а также повысить достоверность и точность результатов.

Курсовая работа предполагает решение каждым студентом нескольких задач из предметной области, используя полученные знания из курса информатики, в частности, навыки работы в математическом пакете MathCad и табличном процессоре MS Excel

Методические указания содержат всю необходимую информацию для выполнения курсовой работы:

- сведения об основных этапах работы, начиная от формализации задач и кончая защитой отчета о выполненной работе;
- рекомендации по программированию, отладке программ и вводу исходных данных;
- постановку предлагаемых для решения задач и разработку математической модели;
- указания по вводу расчетных формул и по способу формализации данных с соответствующими примерами;
- варианты задач;
- рекомендательный библиографический список.

1. ПОРЯДОК ВЫПОЛНЕНИЯ КУРСОВОЙ РАБОТЫ

1. Согласование темы с руководителем работы (срок – сентябрь).
2. Изучение литературы по теме:
 - повторение, углубленное изучение разделов учебников (конспектов лекций), относящихся к выбранной теме;
 - ознакомление с литературой, рекомендуемой настоящими методическими указаниями, конспектирование и цитирование необходимых для решения поставленной задачи теоретических положений;
 - самостоятельный подбор дополнительной литературы;
 - подбор справочного материала.
3. Составление текста заданий в соответствии с номером варианта.
4. Формализация исходной информации, выбор методов решения.
5. Выполнение расчетов в табличном процессоре MS Excel.
6. Выполнение расчетов в системе MathCad (или в ИИС TP).
7. Написание пояснительной записки (отчета о работе).
8. Подготовка графического материала, иллюстрирующего содержание работы и выводы автора.
9. Сдача работы на проверку ее руководителю, доработка текста, графики.
10. Защита курсовой работы.

Выдача заданий по курсовой работе производится не позднее, чем через две недели после начала занятий. Во время выдачи заданий объявляются сроки выполнения студентом отдельных этапов, назначается дата сдачи отчета на проверку и дата защиты работы.

При выставлении оценки по курсовой работе учитываются качество отчета, знания студента по существу работы, оригинальность и творческий подход к выбору методов решения, а также своевременность выполнения всей работы и отдельных ее этапов.

Студент обязан не менее одного раза в месяц информировать руководителя курсовой работы о выполненных этапах.

2. ТРЕБОВАНИЯ К ОТЧЕТУ ПО РАБОТЕ

Отчет по курсовой работе (пояснительная записка) должен содержать следующие разделы:

- титульный лист;
- задание по курсовой работе;
- аннотация;
- оглавление;
- введение;
- теоретические сведения;
- текст каждой задачи (постановка задачи);
- исходные данные;
- подробное описание решения задачи при использовании табличного процессора MS Excel;
- выбор метода решения и описание алгоритмов (блок-схемы);
- выполнение расчетов в системе MathCad;
- выполнение расчетов в любой системе программирования;
- результаты расчетов в виде графиков и таблиц;
- анализ решения задачи, выводы;
- библиографический список.

На титульном листе указывается официальное название института, вид работы, наименование кафедры и название дисциплины, тема курсовой работы, фамилия и инициалы студента, шифр группы, дата оформления отчета, должность, фамилия и инициалы руководителя работы, место для выставления оценки.

В аннотации приводятся краткие сведения о содержании работы (на русском и иностранном языках).

Введение должно содержать информацию о наиболее часто используемых программных средствах при решении математических и прикладных задач.

Теоретические сведения по каждой задаче должны содержать информацию, необходимую для ее решения в общем виде. При указании формул следует разъяснить смысл всех величин, входящих в них.

Текст каждой задачи составляется студентом с учетом постановки задачи и конкретных данных, соответствующих номеру студента в списке группы.

Решение задачи с помощью табличного процессора MS Excel должно демонстрировать этапы расчета с необходимыми для их понимания комментариями. В отчете приводятся фрагменты рабочих листов в режиме отображения данных и в режиме отображения формул. Результаты вычислений следует использовать в качестве теста для проверки правильности решения, полученного на персональном компьютере.

Формализация задачи предполагает, что должны быть рассмотрены вопросы: в какой форме представить исходные данные для их ввода в компьютер, какие формулы и в какой последовательности следует применить для получения промежуточных и окончательных результатов, какова точность вычисления всех параметров и правил их округления. Конкретные рекомендации для каждой задачи имеются в настоящих разделах методических указаний.

Описание алгоритмов должно быть структурированным, «сверху вниз». Это означает, что сначала нужно выделить укрупненные этапы решения задачи. Затем каждый этап разбивается на более мелкие шаги и т.д. Процесс детализации завершается тогда, когда все шаги становятся очевидными для программирования, т.е. их можно представить либо одним оператором, либо небольшим количеством очевидных операторов.

Расчеты в системе MathCad выполняются по разработанному алгоритму. Все расчеты необходимо прокомментировать.

Программа составляется на одном из известных студенту языке программирования, и должна содержать достаточное количество комментариев для понимания ее текста. Все используемые в программе переменные сводятся в таблицу идентификаторов. Образец таблицы приводится в табл.1.

Таблица 1

Таблица идентификаторов для задачи 1.

Обозначение в формуле	Обозначение в программе	Комментарий
i	i	Номер элемента массива в строке
$\sum a_i$	Summa_A	Сумма элементов массива A_i

Исходные данные нужно представить в том виде, в каком их вводят в компьютер. Если их много, то следует записать их в файл и вводить в программу из этого файла. Распечатка должна содержать исходную программу, вводимую информацию и результаты выполнения программы.

Анализ результатов расчетов в MathCad предполагает сравнение с результатами вычислений в MS Excel, а также их смысловую оценку и выводы.

В конце отчета о курсовой работе нужно дать список использованной литературы по информатике и геодезии.

Пояснительная записка составляется с использованием текстового процессора MS Word. Для заголовков разделов и основного текста нужно создать стили. Оглавление в документ вставляется автоматически средствами MS Word на основе созданных стилей для разделов и подразделов.

3. РЕКОМЕНДАЦИИ ПО СОСТАВЛЕНИЮ ПРОГРАММ

При разработке программы нужно следовать принципам структурного программирования: поэтапная детализация, использование только базовых структур (следование, ветвление, цикл), повышение наглядности программы.

Уже на стадии разработки программы нужно продумать мероприятия по ее отладке (подготовка тестов, включение в программу операторов вывода промежуточных результатов, учет особых случаев, ошибок вывода).

В общем случае нужно быть готовым к неожиданностям при запуске программы и поэтому иметь твердые копии (распечатки текста) программы и исходной информации для их восстановления в случае необходимости.

Разрабатывая программу, нужно помнить о целесообразности оформления некоторых важных ее частей в виде подпрограмм. Метод подпрограмм облегчает написание и отладку программы.

Другие указания по программированию приводятся при рассмотрении конкретных задач, входящих в курсовую работу.

ЗАДАНИЕ

Для каждой задачи требуется:

1. Построить в MS Excel график таблично заданной функции.
2. Вычислить коэффициент корреляции в MS Excel для случая линейной зависимости между исходными параметрами.
3. В зависимости от вида графика и величины коэффициента корреляции, выбрать несколько классов эмпирических функций из следующих вариантов: линейная функция, экспоненциальная функция, квадратичная (полиномиальная) функция.
4. Определить конкретный вид выбранных эмпирических функций, решив системы линейных уравнений матричным методом в Excel и вычислить значения коэффициентов. Вычислить коэффициенты детерминированности для полученных функций. Построить графики теоретических функций, с наложением фактических данных.
5. Составить алгоритм вычислений эмпирических функций по методу наименьших квадратов в виде блок-схемы.
6. Выбрать один из численных методов решения систем линейных уравнений и описать его.
7. Написать программы для вычисления коэффициентов эмпирических формул по методу наименьших квадратов в математическом пакете Mathcad и среде программирования

(например, VBA)..

8. Оценить статистическая значимость коэффициента детерминированности с помощью критерия Фишера.

9. Вычислить среднюю ошибку аппроксимации для всех полученных уравнений.

Сравнить все результаты вычислений. Определить, какая из полученных эмпирических формул лучшим образом аппроксимирует заданную формулу.

10. Составить Пояснительную записку в текстовом редакторе MS WORD.

ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

Построение эмпирических формул методом наименьших квадратов

Часто при анализе фактических результатов измерений или экспериментов возникает необходимость найти в явном виде функциональную зависимость между этими фактическими величинами.

Для нахождения аналитической взаимосвязи между двумя величинами X и Y производят ряд наблюдений; в результате получается измеренных значений:

x	x_1	x_2	...	x_i	...	x_n
y	y_1	y_2	...	y_i	...	y_n

Поскольку табличные результаты получаются как итог каких-либо экспериментов, эти значения называются эмпирическими или опытными значениями. Таким образом, исходными данными являются два одномерных массива одинаковой длины, содержащие эмпирические данные.

Если между величинами x и y существует некоторая функциональная зависимость, но ее аналитический вид неизвестен, то возникает практическая задача – найти эмпирическую формулу

$$y^T = F(x, a_1, a_2, \dots, a_m), \quad (1)$$

где a_1, a_2, \dots, a_m – коэффициенты. Вид функции и значения коэффициентов a_1, a_2, \dots, a_m подбираются таким образом, чтобы значения $y_i^T = F(x_i, a_1, a_2, \dots, a_m)$, вычисленные по эмпирической

формуле при различных значениях x_i , как можно меньше отличались бы от опытных значений y_i .

Нахождение аналитической зависимости между эмпирическими величинами называется *аппроксимацией функции*, заданной таблично.

Удачный выбор эмпирической формулы в значительной мере зависит от опыта и знаний исследования в предметной области, используя которые он может правильно указать класс функций.

При аппроксимации определяют класс функций, из которых выбирается теоретическая функция $F(x, a_1, a_2, \dots, a_m)$, и далее отыскивают наилучшие значения коэффициентов.

Чаще всего для аппроксимации используют *метод наименьших квадратов* (МНК).

Поясним геометрический смысл этого метода.

Каждая пара чисел (x_i, y_i) из исходной таблицы определяет точку M_i на плоскости $ХОУ$. Используя формулу (1) с различными значениями коэффициентов a_1, a_2, \dots, a_m , можно построить множество кривых, которые будут являться графиками теоретических функций $F(x, a_1, a_2, \dots, a_m)$. Величина $y_i^T = F(x_i, a_1, a_2, \dots, a_m)$ называется теоретическим значением функции в точке x_i . Разность $(y_i^T - y_i)$ называется *отклонением* или *остатком* и представляет собой расстояние по вертикали от точки M_i до графика эмпирической функции.

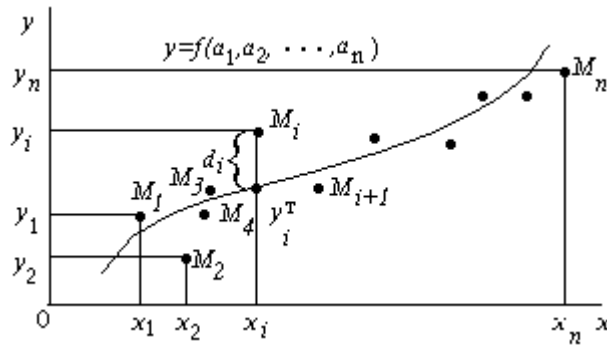


Рис.1. Геометрический смысл МНК

Согласно методу наименьших квадратов наилучшими коэффициентами a_1, a_2, \dots, a_m считаются те, для которых сумма квадратов отклонений найденных теоретических значений функции от заданных эмпирических значений будет минимальной. Следовательно, задача состоит в определении коэффициентов a_1, a_2, \dots, a_m , таким образом, чтобы сумма квадратов отклонений была наименьшей.

$$S(a_1, a_2, \dots, a_m) = \sum_{i=1}^n [f(x_i; a_1, a_2, \dots, a_m) - y_i]^2 \quad (2)$$

Построение эмпирических формул состоит из двух этапов:

1. Выяснение общего вида этой формулы
2. Определение ее наилучших параметров.

Если из теоретических соображений характер зависимости между величинами x и y неизвестен, то вид эмпирической зависимости может быть произвольным. Предпочтение отдается простым формулам, обладающим хорошей точностью.

Большое значение имеет изображение полученных экспериментальных данных в декартовых или в специальных системах координат. По положению точек можно примерно угадать вид зависимости путем установления сходства между построенным графиком и образцами известных кривых.

Для того, чтобы найти набор коэффициентов a_1, a_2, \dots, a_m ,

при которых достигается минимум функции $S(a_1, a_2, \dots, a_m)$, определяемой формулой (2), используем необходимое условие экстремума функции нескольких переменных - равенство нулю частных производных. В результате получим нормальную систему для определения коэффициентов a_1, a_2, \dots, a_m :

$$\frac{\partial S}{\partial a_1} = 0; \frac{\partial S}{\partial a_2} = 0; \dots; \frac{\partial S}{\partial a_m} = 0. \quad (3)$$

Таким образом, нахождение коэффициентов a_1, a_2, \dots, a_m сводится к решению системы (3). Эта система упрощается, если эмпирическая формула (1) линейна относительно параметров a_1, a_2, \dots, a_m , тогда система (3) будет линейной.

Конкретный вид системы (3) зависит от того, из какого класса эмпирических формул мы ищем зависимость (1). В случае линейной зависимости $y = a_1 + a_2 x$ система (3) примет вид:

$$\begin{cases} a_1 n + a_2 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \\ a_1 \sum_{i=1}^n x_i + a_2 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \end{cases} \quad (4)$$

где a_1 и a_2 - неизвестные, а суммы $(\sum_{i=1}^n x_i)$; $(\sum_{i=1}^n y_i)$ и т.д. дают конкретные значения коэффициентов и свободных членов в системе линейных уравнений (4). Эта линейная система может быть решена любым известным методом (с помощью обратной матрицы, методом Гаусса, простых итераций, по формулам Крамера и т.д.).

В случае квадратичной зависимости $y = a_1 + a_2 x + a_3 x^2$ система (3) примет вид

$$\left\{ \begin{array}{l} a_1 n + a_2 \sum_{i=1}^n x_i + a_3 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i \\ a_1 \sum_{i=1}^n x_i + a_2 \sum_{i=1}^n x_i^2 + a_3 \sum_{i=1}^n x_i^3 = \sum_{i=1}^n x_i y_i \\ a_1 \sum_{i=1}^n x_i^2 + a_2 \sum_{i=1}^n x_i^3 + a_3 \sum_{i=1}^n x_i^4 = \sum_{i=1}^n x_i^2 y_i \end{array} \right. \quad (5)$$

Линеаризация нелинейных зависимостей

В ряде случаев в качестве эмпирической зависимости берут функцию, в которую неопределенные коэффициенты входят нелинейно. При этом функцию выбирают, как правило, такого вида, чтобы можно было ее линеаризовать, т.е. свести к линейной. К числу таких зависимостей относятся, например,

- степенная

$$y = a_1 x^{a_2} \quad (6)$$

- экспоненциальная

$$y = a_1 \cdot e^{a_2 \cdot x} \quad (7)$$

- показательная

$$y = a_1 \cdot a_2^x \quad (8)$$

В приведенных выше зависимостях a_1 и a_2 являются коэффициентами, которые необходимо численно определить.

Для указанных выше зависимостей это достигается путем логарифмирования.

В случае *степенной* зависимости линеаризация достигается путем логарифмирования уравнения (6). В результате чего получаем соотношение:

$$\ln y = \ln a_1 + a_2 \ln x \quad (9)$$

Обозначим $\ln y$, $\ln x$ и $\ln a_1$ соответственно через z , t и b , тогда зависимость (8) может быть записана в виде

$$z = b + a_2 t,$$

что позволяет применить формулы (4) с заменой a_1 на b и

пересчетом исходных данных $z_i = \ln y_i$, а $t_i = \ln x_i$. После вычисления b определяем значение коэффициента a_1 исходной зависимости по формуле $a_1 = e^b$.

В случае экспоненциальной зависимости линеаризация достигается путем логарифмирования уравнения (7), после чего получаем соотношение

$$\ln y = \ln a_1 + a_2 \cdot x \quad (10)$$

Обозначим $\ln y$ и $\ln a_1$ соответственно через Z и c , тогда зависимость (7) может быть записана в виде

$$z = c + a_2 \cdot x,$$

что позволяет применить формулы для вычисления коэффициентов линейной зависимости (с заменой a на c и y_i на z_i).

Линеаризующие преобразования для различных видов функций приведены в табл. 2.

Таблица 2

<i>Исходная функция</i>	<i>Замена</i>	<i>Линейное уравнение</i>
$Y = a_1 \cdot e^{a_2 \cdot X}$, экспоненциальная	$\ln Y = Z$ $c = \ln a_1$	$Z = a_1 + a_2 \cdot X$
$Y = a_1 \cdot X^{a_2}$ показательная	$\ln Y = Z$ $c = \ln a_1$ $\ln X = T$	$Z = a_1 + a_2 \cdot T$
$Y = a_1 \cdot a_2^X$ степенная	$\ln Y = Z$ $c = \ln a_1$ $d = \ln a_2$	$Z = c + d \cdot X$
$Y = a_1 + \frac{a_2}{X}$ равносторонняя гипербола	$\frac{1}{X} = T$	$Y = a_1 + a_2 T$

Специальный вид линейной зависимости.

В традиции некоторых разделов науки зависимость содержания различных ионов (минерализации) от плотности пластовой воды принято представлять в виде

$$Y = c \cdot (X - 1000) \quad , \quad (11)$$

где X - плотности пластовой воды ,

Y – степень минерализации,

c - некоторый параметр, зависящий от типа растворенных ионов (минералов).

Такой вид зависимости ряд преимуществ по сравнению с традиционным. Во-первых, он обеспечивает нулевое значение степени минерализации для дистиллированной воды, которая имеет плотность равную 1000, и при этом значении X , очевидно Y равен нулю.

Во-вторых, параметр c имеет простой смысл - он показывает на сколько повышается степень минерализации для пластовой воды если её плотность возрастает на единицу по сравнению с плотностью дистиллированной воды.

В качестве примера приведем эмпирическую зависимость содержания ионов от плотности воды, поступающей в скважину вместе с нефтью. Для этого был проанализирован состав вод по скважинам горизонта D_1 Ромашкинского месторождения, эксплуатирующимся в течение длительного времени. Для основных ионов пластовой воды – Na^+ , K^+ , Ca^{2+} , Mg^{2+} , Cl^- – указанная зависимость в пределах изменения плотности пластовых вод 1030 – 1185 kg/m^3 типичная зависимость приведена рис. 2. Очевидно, что она носит линейный характер и хорошо аппроксимируется уравнением

$$c(\rho) = c \cdot (\rho - 1000) \quad , \quad (12)$$

где ρ – плотность воды, поступающей в скважину, c – постоянная величина для данного вида иона, характеризующая концентрацию.

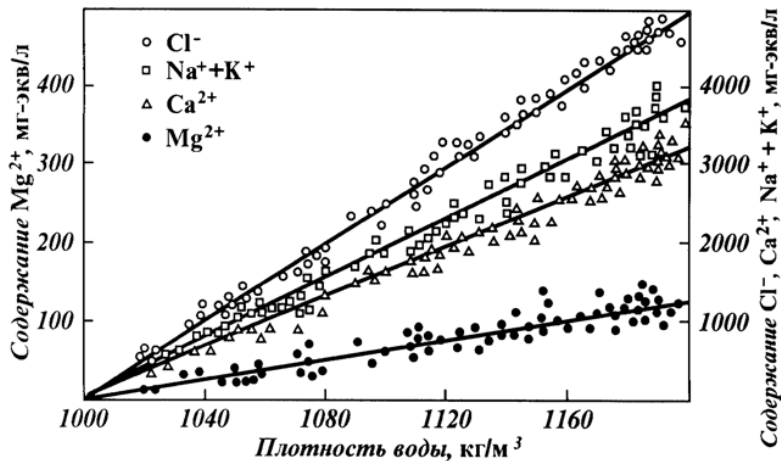


Рис.2. Зависимости содержания ионов от плотности воды, поступающей в скважину вместе с нефтью

Полученная закономерность подтверждена результатами обработки данных об изменении состава вод при заводнении продуктивных пластов девонского горизонта и верхнего карбона на 121-м месторождении Урало-Поволжья и Западной Сибири.

Чтобы определить коэффициент c в зависимости (11), достаточно сделать замену

$$Z = (X - 1000). \quad (13)$$

Тогда зависимость (11) примет вид

$$Y = c \cdot Z. \quad (14)$$

Чтобы найти коэффициент c с помощью МНК воспользуемся следующими соотношениями. Цель МНК найти такое значение c чтобы сумма квадратов отклонений $S(c)$ была минимальной. Сумма квадратов отклонений в данном случае равна

$$S(c) = \sum_{i=1}^n (c \cdot Z_i - Y_i)^2$$

Необходимым условием экстремума является соотношение

$$\frac{d(S(c))}{dc} = 0 \quad (15)$$

Соотношение (15) может быть преобразовано к виду:

$$c \cdot \sum_{i=1}^n Z_i^2 = \sum_{i=1}^n Y_i Z_i$$

Откуда можно определить c :

$$c = \frac{\sum_{i=1}^n Y_i Z_i}{\sum_{i=1}^n Z_i^2} \quad (16)$$

Элементы теории корреляции

График теоретической зависимости $Y^T(x)$, полученный по найденной эмпирической формуле, называется *кривой регрессии*.

Для проверки согласия (справедливости) построенной кривой регрессии с результатами эксперимента, как правило, используют следующие числовые характеристики: *коэффициент корреляции* и *коэффициент детерминированности*.

Коэффициент корреляции является мерой линейной связи между зависимыми величинами. Он показывает, насколько хорошо, в среднем, может быть представлена (вычислена) одна из величин в виде линейной функции от другой.

Коэффициент корреляции вычисляется по формуле:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (17)$$

где $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$, $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$ - среднеарифметические значения по x и y соответственно.

Коэффициент корреляции по абсолютной величине не превосходит 1. Чем ближе $|r|$ к 1, тем теснее линейная связь между x и y , тем более целесообразна аппроксимация таблично заданной

функции линейной зависимостью.

Особо подчеркнем, что если коэффициент корреляции существенно меньше 1, это не означает отсутствие зависимости между x и y . Это означает только, что не применима линейная аппроксимация, но можно искать аппроксимирующую зависимость среди степенных, экспоненциальных, квадратичных и других видов функций.

Чтобы определить насколько хорошо построенная зависимость отображает эмпирические данные, водится еще одна характеристика – *коэффициент детерминированности* R^2 .

Пусть $S_{ост}$ – сумма квадратов отклонений теоретических значений функции от эмпирических данных:

$$S_{ост} = \sum_{i=1}^n (y_i - y_i^T)^2 . \quad (18)$$

Полученная величина характеризует отклонение теоретических результатов от экспериментальных данных. Чем больше $S_{ост}$, тем хуже выбранная теоретическая функция описывает экспериментальные данные и, наоборот, чем меньше $S_{ост}$, тем лучше выбранная теоретическая функция описывает экспериментальные данные.

Введем понятие регрессионной суммы квадратов:

$$S_{рег} = \sum_{i=1}^n (y_i^T - \bar{y})^2 . \quad (19)$$

Эта величина характеризует разброс теоретических данных относительно среднего значения.

Для линейной зависимости справедливо следующее соотношение:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i^T - \bar{y})^2 + \sum_{i=1}^n (y_i^T - y_i)^2 \quad (20)$$

Обозначим: $S_{полн} = \sum_{i=1}^n (y_i - \bar{y})^2$, тогда, для линейной зависимости очевидно, справедливо следующее равенство:

$$S_{полн} = S_{ост} + S_{рег} \quad (21)$$

Коэффициент детерминированности R^2 определяется по формуле:

$$R^2 = 1 - \frac{S_{ост}}{S_{полн}} \quad (22)$$

Поскольку $S_{полн} \geq 0$, $S_{ост} \geq 0$ и из формулы (22) следует что $S_{полн} \geq S_{ост}$, то из формулы (14) следует

$$0 \leq R^2 \leq 1 \quad (23)$$

Чем меньше остаточная сумма квадратов $S_{ост}$ по сравнению с общей суммой квадратов $S_{полн}$, тем больше значение коэффициента детерминированности R^2 .

Коэффициент детерминированности R^2 показывает, насколько хорошо полученная теоретическая функция описывает взаимосвязь между эмпирическими данными. Если этот коэффициент равен 1, то имеет место полное совпадение выбранной теоретической модели с фактическими данными. В противоположном случае, если коэффициент детерминированности близок к нулю, то выбранная эмпирическая формула неудачна, и она не может использоваться для вычисления значений функции.

Коэффициент детерминированности служит показателем тесноты связи между фактором x и откликом y , описываемой данным уравнением.

Иногда показателям тесноты связи можно дать качественную оценку (*шкала Чеддока*) (табл.3).

Таблица 3

<i>Количественная мера тесноты связи</i>	<i>Качественная характеристика силы связи</i>
0,1-0,3	Слабая
0,3-0,5	Умеренная
0,5-0,7	Заметная
0,7-0,9	Высокая

0,9-0,99	Весьма высокая
----------	----------------

Коэффициент детерминированности R^2 может быть преобразован к следующему виду:

$$R^2 = 1 - \frac{S_{ост}}{S_{общ}} = \frac{S_{общ} - S_{ост}}{S_{общ}} = \frac{S_{факт}}{S_{общ}} = \frac{S_{факт}/n}{S_{общ}/n} = \frac{D_{факт}}{D_{общ}}, \quad (23)$$

где $D_{факт}$ – факторная дисперсия;
 $D_{общ}$ – общая дисперсия.

Таким образом, коэффициент детерминированности R^2 равен доле вариации Y объясняемой вариацией фактора X .

В случае линейной зависимости двух переменных коэффициент детерминированности равен квадрату коэффициента корреляции ($R^2 = r_{xy}^2$).

Далее оценивается *статистическая значимость* коэффициента детерминированности и параметров полученного уравнения, то есть оценка вероятности того, что данные величины не примут нулевые значения.

Проверка значимости уравнения в целом, то есть гипотезы о наличии линейной зависимости между X и Y , проводится с помощью *критерия Фишера*.

Проверка значимости уравнения в целом предполагает проверку нулевой гипотезы об отсутствии линейной связи между X и Y , то есть $H_0: R^2 = 0$, альтернативная гипотеза $H_1: R^2 \neq 0$, то есть R^2 существенно отличен от нуля и уравнение значимо. Если нулевая гипотеза справедлива, то $S_{ост}$ мало отличается от $S_{факт}$.

Для отклонения H_0 необходимо, чтобы регрессионная (факторная) дисперсия превышала остаточную в несколько раз. Схема проверки гипотезы совпадает с общей схемой проведения дисперсионного анализа (табл. 3).

Для линейного уравнения регрессии справедливо выражение

$$R^2 = 1 - \frac{S_{ост}}{S_{общ}} = 1 - \frac{S_{ост}}{S_{ост} + S_{факт}} = 1 - \frac{1}{1 + \frac{S_{факт}}{S_{ост}}}. \quad (24)$$

Отсюда следует, что чем больше отношение $\frac{S_{факт}}{S_{ост}}$, тем ближе значение коэффициента детерминированности к единице. Это утверждение справедливо и для нелинейной регрессии.

Для сравнения качества различных моделей используется скорректированный индекс детерминации $-R_{adj}^2$, содержащий поправку на число степеней свободы:

$$R_{adj}^2 = 1 - (1 - R^2) \cdot \frac{(n-1)}{(n-m-1)}. \quad (25)$$

Другой оценкой качества уравнения регрессии является средняя ошибка аппроксимации - среднее отклонение теоретических значений от фактических, которая определяется по формуле:

$$A = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \cdot 100 \%. \quad (26)$$

Модель считается пригодной для прогноза, если величина A не превышает 8%-10%.

2. ПРИМЕРЫ РЕШЕНИЯ ЗАДАЧ

ЗАДАЧА 1

На рис.3 приведены экспериментальные данные о содержании ионов Cl (c , мг-экв/л) и плотности воды (ρ , кг/м³), поступающей в скважину вместе с нефтью.

Плотность, кг/м ³	ρ ₀	X	1015,0	1033,0	1034,0	1044,0	1054,0	1063,0	1074,0	1091,0	1095,0	1095,0	1107,0	1111,0	1119,0
Содержание ионов Cl ⁻ , мг-экв/л	c	Y	511,76	700,00	767,64	895,58	1023,52	1279,40	1535,28	1791,16	2302,92	2814,68	3070,56	2558,80	2814,68
Плотность, кг/м ³	ρ ₀	X	1123,0	1123,0	1128,0	1133,0	1140,0	1149,0	1151,0	1155,0	1161,0	1164,0	1169,0	1175,0	
Содержание ионов Cl ⁻ , мг-экв/л	c	Y	3070,56	3198,50	3326,44	3454,38	3582,32	3710,26	3838,20	3966,14	4094,08	4222,02	4349,96	4477,90	

Рис.3 Исходные данные для построения эмпирической зависимости.

Задание.

Исследовать эмпирическую зависимость содержания ионов Cl⁻ от плотности воды, поступающей в скважину вместе с нефтью, по следующему плану:

1. Построить, используя МНК, и исследовать «специализированное» уравнение (12).
2. Вычислить R_{12}^2 - коэффициент детерминированности, характеризующий качество этого «специализированного» уравнения.
3. Используя МНК, построить и исследовать линейное уравнение «общего» вида

$$c(\rho) = a_2 \cdot \rho + a_1 \quad (27)$$

4. Определить a_1 и a_2 в уравнении (24), решив систему (4).
5. Вычислить R_{24}^2 - коэффициент детерминированности, характеризующий качество линейного уравнения общего вида
6. Сравнить качество уравнений (12) и (24), сопоставив величины R_{12}^2 и R_{24}^2 .
7. Вычислить по этим уравнениям прогнозные значения величины содержания ионов Cl⁻, если плотность пластовой воды равна $\rho_{\text{прогнози}}$. В качестве $\rho_{\text{прогнози}}$ взять величину равную $\rho_{\text{max}} - 0.1 \cdot (\rho_{\text{max}} - \rho_{\text{min}})$, где ρ_{max} и ρ_{min} максимальное и минимальное значения плотности ρ в таблице исходных данных. Вычислить абсолютную и относительную разницу прогнозных

значений. Сделать вывод о целесообразности использовании этих уравнений.

8. Используя МНК, построить и исследовать квадратичную ($y = a_1 + a_2x + a_3x^2$) и экспоненциальную ($Y = a_1 \cdot e^{a_2 \cdot X}$) зависимости. Вычислить коэффициенты детерминированности по формуле (24) для полученных уравнений квадратичной и экспоненциальной зависимостей, сравнить их с коэффициентом детерминированности, характеризующий качество линейного уравнения общего вида.

9. Вычислить среднюю ошибку аппроксимации A и скорректированный индекс детерминации $-R_{adj}^2$ для полученных моделей. Сравнить полученные коэффициенты, сделать выводы.

Указания к решению

Все необходимые вычисления выполнить в электронных таблицах MS Excel и системе MathCAD.

Расчеты провести в электронных таблицах MS Excel тремя способами:

- 1) численно, используя расчетные формулы приведенные выше;
- 2) используя средство «Диаграмма-Тренд»;
- 3) используя встроенную функцию ЛИНЕЙН для получения параметров линейной зависимости.

Расчеты провести в системе MathCAD двумя способами:

- 1) численно, используя расчетные формулы приведенные выше;
- 2) используя встроенную функцию line().

Ввод исходных данных обеспечить из текстового файла.

Решение

Расчеты для определения коэффициентов **линейной** зависимости (12) и (24) с использованием MS Excel приведены на рис.4-16.

На рис.4-5 в ячейке C35 содержится вычисленный коэффициент корреляции r , он равен 0,988. Это значение близко к единице, что позволяет сделать вывод о том, что содержании ионов

Cl⁻ (с, мг-экв/л) и плотности воды (ρ , кг/м³), поступающей в скважину вместе с нефтью связаны линейной зависимостью.

На рис.4-5 также проведены расчеты по определению коэффициента c для «специализированного» уравнения (12). Коэффициент c для иона Cl⁻, характеризующий концентрацию, находится в ячейке G34 и в данном случае равен 25,20. Таким образом, искомое «специализированное» уравнение имеет вид

$$c(\rho) = 25.20 \cdot (\rho - 1000) \quad (28)$$

Для определения качества уравнения на рис.6-7 показаны вычисления коэффициент детерминированности R^2 . Значение R_{12}^2 приведено в ячейке J34, оно равно 0.9710. Такая величина R^2 (близкая к единице), позволяет сделать вывод, что полученное «специализированное» уравнение, хорошо описывает эмпирические данные.

	A	B	C	D	E	F	G
1	Зависимость содержания ионов Cl⁻ от плотности пластовой воды						
2	Исходные данные			Расчеты для "специализированного" уравнения			
3							
4		Плотность	Содержание ионов Cl⁻	"Приведенная плотность"			
5		г0	с	г0'			
6		X	Y	Z=X-1000	Z*Z	Z*Y	Y-Y _{ср.}
7		1015,0	511,76	15,0	225	7676,4	-2182,5
8		1033,0	700,00	33,0	1089	23100	-1994,3
30		1169,0	4349,96	169,0	28561	735143,2	1655,7
31		1175,0	4477,90	175,0	30625	783632,5	1783,6
32	Сумма	27706,0	67356,7		345990,0	8717774,5	0,0
33	Ср.знач.	1108,2	2694,3				
34						c=	25,20
35		г=	0,9877205				

3

Рис.4 Решение задачи 1. Расчеты для «специализированного» уравнения на листе MS Excel в режиме отображения данных (начало).

	B	C	D	E	F	G
6	X	Y	Z=X-1000	Z*Z	Z*Y	Y-Уср.
7	1015	511,76	=B7-1000	=D7*D7	=D7*C7	=C7-\$C\$33
8	1033	700	=B8-1000	=D8*D8	=D8*C8	=C8-\$C\$33
30	1169	4349,96	=B30-1000	=D30*D30	=D30*C30	=C30-\$C\$33
31	1175	4477,9	=B31-1000	=D31*D31	=D31*C31	=C31-\$C\$33
32	=СУММ(B7:B31)	=СУММ(C7:C31)		=СУММ(E7:E31)	=СУММ(F7:F31)	=СУММ(G7:G31)
33	=СРЗНАЧ(B7:B31)	=СРЗНАЧ(C7:C31)				
34					c=	=F32/E32
35	r=	=КОРРЕЛ(B7:B31;C7:C31)				

Рис.5 Решение задачи 1. Расчеты для «специализированного» уравнения на листе MS Excel в режиме отображения формул (начало).

	H	I	J	K	L	M
6	(Y-Уср.)^2	Утеор	Утеор-Y	(Утеор-Y)^2	Утеор-Уср	(Утеор-Уср)^2
7	4763348,2	376,3	-135,5	18357,9	-2318,0	5373128,2
8	3977111,2	827,8	127,8	16330,6	-1866,5	3483741,8
30	2741310,7	4239,3	-110,7	12247,0	1545,0	2387100,6
31	3181337,4	4389,8	-88,1	7761,3	1695,5	2874828,6
32	39318357,3	67878,9	522,1	1141569,3	522,1	33418663,8
33	Сполн.			Сост.		Sperp.
34		R^2=	0,9710			

Рис.6 Решение задачи 1. Расчеты для «специализированного» уравнения на листе MS Excel в режиме отображения данных (окончание).

	G	H	I	J	K	L	M
6	Y-Уср.	(Y-Уср.)^2	Утеор	Утеор-Y	(Утеор-Y)^2	Утеор-Уср	(Утеор-Уср)^2
7	=C7-\$C\$33	=G7*G7	=\$G\$34*(B7-1000)	=I7-C7	=J7*J7	=I7-\$C\$33	=L7*L7
8	=C8-\$C\$33	=G8*G8	=\$G\$34*(B8-1000)	=I8-C8	=J8*J8	=I8-\$C\$33	=L8*L8
30	=C30-\$C\$33	=G30*G30	=\$G\$34*(B30-1000)	=I30-C30	=J30*J30	=I30-\$C\$33	=L30*L30
31	=C31-\$C\$33	=G31*G31	=\$G\$34*(B31-1000)	=I31-C31	=J31*J31	=I31-\$C\$33	=L31*L31
32	=СУММ(G7:G31)	=СУММ(H7:H31)	=СУММ(I7:I31)	=СУММ(J7:J31)	=СУММ(K7:K31)	=СУММ(L7:L31)	=СУММ(M7:M31)
33		Сполн.			Сост.		Sperp.
34	=F32/E32		R^2=	=1-K32/H32			

Рис.7 Решение задачи 1. Расчеты для «специализированного» уравнения на листе MS Excel в режиме отображения формул

(окончание).

Расчеты для определения коэффициентов зависимости (24), т.е. «общего» уравнения с использованием MS Excel приведены на рис.8-11. При этом коэффициенты системы (4) определены в интервале ячеек B66:D66 (рис.8-9).

Система нормальных уравнений примет вид

$$\begin{cases} 25 \cdot a_1 + 27706 \cdot a_2 = 67356.7 \\ 27706a_1 + 30757990 \cdot a_2 = 76074514.46 \end{cases}$$

Коэффициенты матрицы системы содержатся в интервале ячеек B70:C71; вектор правых частей - в интервале ячеек E70:E71.

Для нахождения решения используем обратную матрицу, которая содержится в интервале ячеек B74:C75. Вектор решения содержится в интервале ячеек E74:E75.

Таким образом, a_1 и a_2 равны -27094,23 и 29,879 соответственно, и уравнение «общего» вида может быть записано как

$$c(\rho) = 26.879 \cdot \rho - 27094.23 \quad (29)$$

	A	B	C	D	E	F	G
39	Расчеты для "общего" уравнения						
40		X	Y	X*X	X*Y	Y-Усп.	(Y-Усп.)^2
41	1	1015	511,76	1030225	519436,40	-2182,5	4763348,2
42	2	1033	700,00	1067089	723100,00	-1994,3	3977111,2
43	3	1034	767,64	1069156	793739,76	-1926,6	3711901,6
63	23	1164,00	4222,02	1354896	4914431,28	1527,8	2334021,3
64	24	1169,00	4349,96	1366561	5085103,24	1655,7	2741310,7
65	25	1175,00	4477,90	1380625	5261532,50	1783,6	3181337,4
66	Сумма	27706,00	67356,74	30757990	76074514,46	0,0	39318357,3
67	Ср.знач.	1108,2	2694,3				Сполн.
68							
69	A				B		
70		25	27706		67356,7		
71		27706	30757990		76074514,5		
73		A^(-1)					
74		23,173107	-0,020874		-27094,23865	=a1	
75		-0,020874	1,884E-05		26,87911305	=a2	

Рис.8 Решение задачи 1. Расчеты для «общего» уравнения на листе MS Excel в режиме отображения данных (начало).

	B	C	D	E	F
40	X	Y	X*X	X*Y	Y-Усп.
41	=B7	=C7	=B41^2	=B41*C41	=C41-\$C\$67
42	=B8	=C8	=B42^2	=B42*C42	=C42-\$C\$67
43	=B9	=C9	=B43^2	=B43*C43	=C43-\$C\$67
63	=B29	=C29	=B63^2	=B63*C63	=C63-\$C\$67
64	=B30	=C30	=B64^2	=B64*C64	=C64-\$C\$67
65	=B31	=C31	=B65^2	=B65*C65	=C65-\$C\$67
66	=СУММ(B41:B65)	=СУММ(C41:C65)	=СУММ(D41:D65)	=СУММ(E41:E65)	=СУММ(F41:F65)
67	=СРЗНАЧ(B41:B65)	=СРЗНАЧ(C41:C65)			
69	A			B	
70	=A65	=B66		=C66	
71	=B66	=D66		=E66	
73	A^(-1)				
74	=МОБР(B70:C71)	=МОБР(B70:C71)		=МУМНОЖ(B74:C75;E70:E71)	=a1
75	=МОБР(B70:C71)	=МОБР(B70:C71)		=МУМНОЖ(B74:C75;E70:E71)	=a2

Рис.9 Решение задачи 1. Расчеты для «общего» уравнения на листе MS Excel в режиме отображения формул (начало).

Для определения качества «общего» уравнения на рис.10-11 показаны вычисления коэффициент детерминированности R^2 . Значение R_{24}^2 приведено в ячейке I68, оно равно 0.9756. Такая величина R_{24}^2 (близкая к единице), позволяет сделать выводы, что полученное «общее», хорошо описывает эмпирические данные и делает это лучше, чем «специализированное» уравнение у которого $R_{12}^2=0.9710$. Это улучшение не велико, поскольку разница в коэффициентах мала и равна 0,0046, т.е. имеет порядок 10^{-3} .

	Н	И	Ж	К	Л
40	Утеор	Утеор-У	(Утеор-У)^2	Утеор-Усп	(Утеор-Усп)^2
41	188,1	-323,7	104781,0	-2506,2	6281081,0
42	671,9	-28,1	790,4	-2022,4	4090038,9
43	698,8	-68,9	4743,9	-1995,5	3982041,6
63	4193,0	-29,0	839,3	1498,8	2246339,5
64	4327,4	-22,5	506,9	1633,2	2667260,3
65	4488,7	10,8	117,1	1794,4	3220049,3
66		0,0	959687,8	0,0	38358669,4
67			Сост.		Сперп.
68	R^2=	0,9756			

Рис.10 Решение задачи 1. Расчеты для «общего» уравнения на листе MS Excel в режиме отображения данных(окончание).

	F	G	Н	И	Ж	К	Л
40	У-Усп.	(У-Усп.)^2	Утеор	Утеор-У	(Утеор-У)^2	Утеор-Усп	(Утеор-Усп)^2
41	=C41-\$C\$67	=F41*F41	=\$E\$75*B41+\$E\$74	=H41-C41	=I41*I41	=H41-\$C\$67	=K41*K41
42	=C42-\$C\$67	=F42*F42	=\$E\$75*B42+\$E\$74	=H42-C42	=I42*I42	=H42-\$C\$67	=K42*K42
43	=C43-\$C\$67	=F43*F43	=\$E\$75*B43+\$E\$74	=H43-C43	=I43*I43	=H43-\$C\$67	=K43*K43
63	=C63-\$C\$67	=F63*F63	=\$E\$75*B63+\$E\$74	=H63-C63	=I63*I63	=H63-\$C\$67	=K63*K63
64	=C64-\$C\$67	=F64*F64	=\$E\$75*B64+\$E\$74	=H64-C64	=I64*I64	=H64-\$C\$67	=K64*K64
65	=C65-\$C\$67	=F65*F65	=\$E\$75*B65+\$E\$74	=H65-C65	=I65*I65	=H65-\$C\$67	=K65*K65
66	=СУММ(F41:F65)	=СУММ(G41:G65)		=СУММ(I41:I65)	=СУММ(J41:J65)	=СУММ(K41:K65)	=СУММ(L41:L65)
67		Сполн.			Сост.		Сперп.
68			R^2=	=1-J66/G66			

Рис.11 Решение задачи 1. Расчеты для «общего» уравнения на листе MS Excel в режиме отображения формул (окончание).

Расчеты по этим уравнениям прогнозных значений величины содержания ионов Cl^- , если плотность пластовой воды равна заданному значению $\rho_{\text{прогноз}}$, приведено на рис.12-13.

	A	B	C	D	E	F
78	Вычисление прогнозного значения					
79		x_прогноз	1159,0			
80						
81		Кoeffициенты			Отклонение	
82	Уравнение	a1	a2 или c	y_прогнозн	абсолютное	относительное, %%
83	специализированное		25,20	4006,26	52,39	1,31
84	общее	-27094,24	26,879113	4058,65		

Рис.12 Решение задачи 1. Расчеты для определения прогнозного значения на листе MS Excel в режиме отображения данных

	D	E	F	G	H	I
78	Вычисление прогнозн					
79		x_прог	=МАКС(Е7:Е31)-0,1*(МАКС(Е7:Е31)-МИН(Е7:Е31))			
80						
81		Кoeff			Разница	
82	Уравнение	a1	a2 или c	y_прогнозн	абсолютная	относительная, %%
83	специализированное		=J34	=J34*(F79-1000)	=G84-G83	=H83/G83*100
84	общее	=H74	=H75	=H75*F79+H74		

Рис.13 Решение задачи 1. Расчеты для определения прогнозного значения на листе MS Excel в режиме отображения формул.

Для определения величины $\rho_{\text{прогноз}}$ в ячейку F79 вводим формулу, соответствующую выражению $\rho_{\text{max}} - 0.1 \cdot (\rho_{\text{max}} - \rho_{\text{min}})$, где ρ_{max} и ρ_{min} максимальное и минимальное значения плотности ρ в таблице исходных данных (интервал ячеек E7:E31).

В результате получим значение $\rho_{\text{прогноз}}$ равное 1159,0.

Для вычисления прогнозных значений величины содержания ионов Cl^- подставим это значение в найденные уравнения. Полученные прогнозные значения величины содержания ионов для «специализированного» и «общего» уравнений находятся в ячейках

D83 и D84 соответственно. Абсолютная и относительная разности прогнозных значений приведены в ячейках E83 и F83.

Коэффициент c зависимости (12) также может быть найден с помощью средств MS Excel, если воспользоваться специальной возможностью средства «Диаграмма-Тренд». Для этого на вкладке «Параметры» устанавливаем отметку (\checkmark) в элементе управления с надписью «пересечение кривой с осью Y в точке 0» (рис.14).

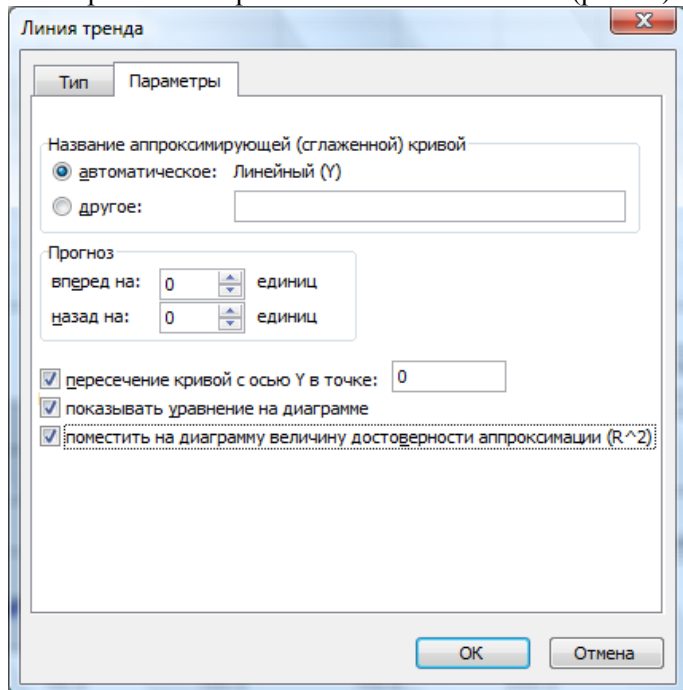


Рис.14 Вкладка «Параметры» средства «Диаграмма-Тренд»

Результат применения этого средства приведен на рис.15. Заметим, что по оси абсцисс откладываются не истинные значения плотности, а их отклонения от плотности дистиллированной воды, равной 1000. Коэффициент c и коэффициент детерминированности R_{12}^2 получились равными 25,197 и 0,9711, что полностью совпало со значениям, вычисленным по расчетным формулам.

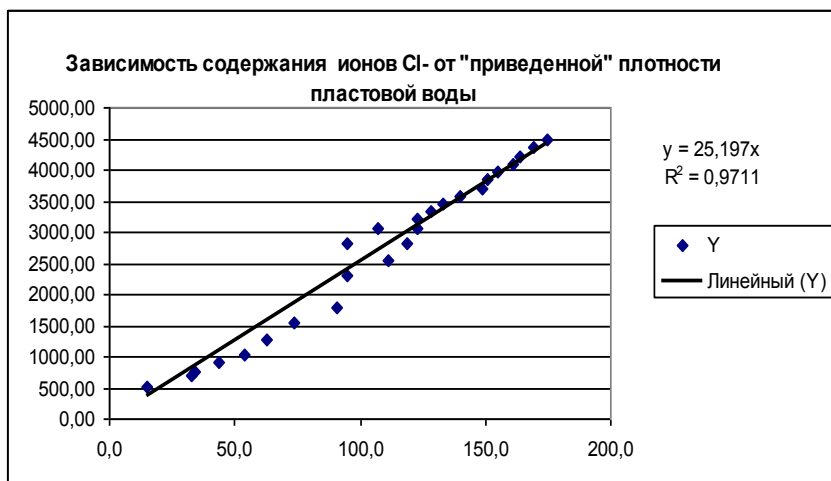


Рис.15 Решение задачи 1. Построение средством «Диаграмма-Тренд» зависимости для «специализированного» уравнения на листе MS Excel.

Коэффициенты a_1 и a_2 уравнения «общего» вида (24) также определены с помощью средства MS Excel «Диаграмма-Тренд» на рис.16. Все полученные величины совпали со значениями, вычисленными по расчетным формулам.

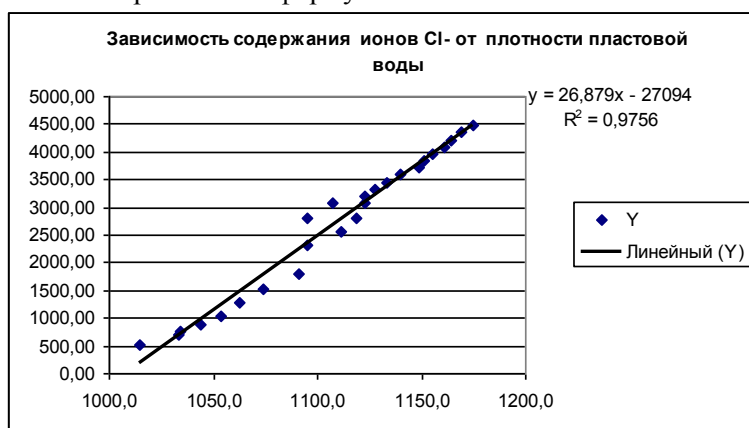


Рис.16 Решение задачи 1. Построение средством «Диаграмма-Тренд» зависимости для «общего» уравнения на листе MS Excel .

Расчеты для определения коэффициентов зависимости (12) и (24) с использованием **MathCAD** приведены на рис.17-21.

Исходные данные запишем в текстовый файл. Для этого можно скопировать исходные данные из интервала ячеек В7:С31 (рис.4) в буфер обмена и перенести в текстовый файл созданный с помощью блокнота. Далее (при необходимости) установить символ «точка» в качестве разделителя между целой и дробной частью чисел, сохраним полученный файл с именем mnk_oil.txt. Вид текстового файла приведен на рис.17. Этот текстовый файл будет использован для ввода данных как при расчетах в MathCAD.

Все вычисления в MathCAD достаточно прозрачны. Для ввода данных используется команда ► **Вставка** ► **Компонента** из главного меню MathCAD.

Будет запущен Мастер (специальная программа), который обеспечит выполнение всей процедуры ввода, результат ввода будет находиться в двумерном массиве W .

Далее эти данные (по столбцам) присваиваем одномерным массивам x и y , с которыми выполняем все дальнейшие операции.

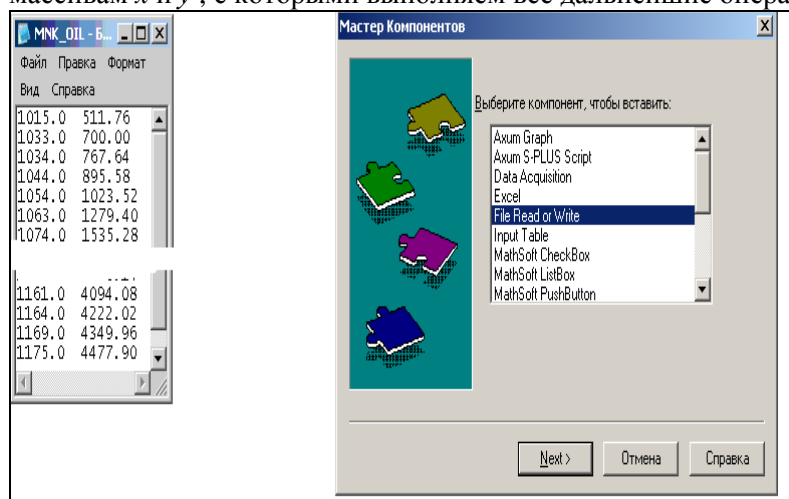


Рис.17 Решение задачи 1. Подготовка данных и запуск *Мастера Компонентов* для ввода данных в системе MathCAD .

Зависимость содержания ионов Cl⁻ от плотности пластовой воды. Построение эмпирической формулы МНК.

Ввод исходных данных в общий массив W из текстового файла с помощью команды Insert Component (вставить компонент)

W :=



W =

	1	2
1	$1.015 \cdot 10^3$	511.76
2	$1.033 \cdot 10^3$	700
3	$1.034 \cdot 10^3$	767.64
4	$1.044 \cdot 10^3$	895.58
5	$1.054 \cdot 10^3$...

Вывод исходных данных

	1		1
1	$1.015 \cdot 10^3$	1	511.76
2	$1.033 \cdot 10^3$	2	700
3	$1.034 \cdot 10^3$	3	767.64
4	$1.044 \cdot 10^3$	4	895.58
5	$1.054 \cdot 10^3$	5	$1.024 \cdot 10^3$
6	$1.063 \cdot 10^3$	6	$1.279 \cdot 10^3$
7	$1.074 \cdot 10^3$	7	$1.535 \cdot 10^3$
8	$1.091 \cdot 10^3$	8	$1.791 \cdot 10^3$
9	$1.095 \cdot 10^3$	9	$2.303 \cdot 10^3$
10	$1.095 \cdot 10^3$	10	$2.815 \cdot 10^3$
11	$1.107 \cdot 10^3$	11	$3.071 \cdot 10^3$
12	$1.111 \cdot 10^3$	12	$2.559 \cdot 10^3$
13	$1.119 \cdot 10^3$	13	$2.815 \cdot 10^3$
14	$1.123 \cdot 10^3$	14	$3.071 \cdot 10^3$
15	$1.123 \cdot 10^3$	15	$3.199 \cdot 10^3$
16	$1.128 \cdot 10^3$	16	$3.326 \cdot 10^3$
17	$1.133 \cdot 10^3$	17	$3.454 \cdot 10^3$
18	$1.14 \cdot 10^3$	18	$3.582 \cdot 10^3$
19	$1.149 \cdot 10^3$	19	$3.71 \cdot 10^3$
20	$1.151 \cdot 10^3$	20	$3.838 \cdot 10^3$
21	...	21	...

Коэффициент корреляции между величинами x и y
 $\text{corr}(x, y) = 0.988$

Переопределение начального индекса в массивах

$\text{ORIGIN} := 1$

Количество наблюдений

$N := \text{rows}(W)$

$N = 25$

Пределы изменения индекса векторов

$i := 1..N$

Эмпирические данные раскладываем по отдельным "векторам"

$x_i := W_{i,1}$ $y_i := W_{i,2}$

Вычисление некоторых характеристик эмпирических данных

$y_{cp} := \text{mean}(y)$

$y_{cp} = 2.694 \times 10^3$

$$S_{total} := \sum_{i=1}^N (y_i - y_{cp})^2$$

$S_{total} = 3.932 \times 10^7$

$x_{min} := \text{min}(x)$

$x_{min} = 1.015 \times 10^3$

$x_{max} := \text{max}(x)$

$x_{max} = 1.175 \times 10^3$

Размах величины X

$R_x := x_{max} - x_{min}$

$R_x = 160$

Прогнозное значение x (при котором будем вычислить теоретический y)

$x_{prognoz} := x_{max} - 0.1 \cdot R_x$

$x_{prognoz} = 1.159 \times 10^3$

Рис.18 Решение задачи 1 в системе MathCAD (начало).

Расчеты для "общего" уравнения

1 способ. Вычисление коэффициентов "общего" уравнения, с помощью решения нормальной системы уравнений

$$S_x := \sum_{i=1}^N (x_i) \quad S_y := \sum_{i=1}^N [(y)_i] \quad S_{xy} := \sum_{i=1}^N (x_i \cdot y_i) \quad S_{xx} := \sum_{i=1}^N ((x_i \cdot x_i))$$

$$S_x = 2.771 \times 10^4 \quad S_y = 6.736 \times 10^4 \quad S_{xy} = 7.607 \times 10^7 \quad S_{xx} = 3.076 \times 10^7$$

$$A := \begin{pmatrix} N & S_x \\ S_x & S_{xx} \end{pmatrix} \quad B := \begin{pmatrix} S_y \\ S_{xy} \end{pmatrix} \quad \begin{pmatrix} a1 \\ a2 \end{pmatrix} := A^{-1} \cdot B$$

$$A = \begin{pmatrix} 25 & 2.771 \times 10^4 \\ 2.771 \times 10^4 & 3.076 \times 10^7 \end{pmatrix} \quad B = \begin{pmatrix} 6.736 \times 10^4 \\ 7.607 \times 10^7 \end{pmatrix} \quad \text{Коэффициенты "общего" уравнения}$$

$$\begin{pmatrix} a1 \\ a2 \end{pmatrix} = \begin{pmatrix} -2.709 \times 10^4 \\ 26.879 \end{pmatrix}$$

2 способ. Вычисление коэффициентов "общего" уравнения, с помощью функции line

$$fit := line(x, y)$$

Коэффициенты "общего" уравнения

$$fit_1 = -2.709 \times 10^4 \quad fit_2 = 26.879$$

"Общее" уравнение в явном виде

$$c(ro) := a1 + a2 \cdot ro$$

Вычисление теоретических значения y для "общего" уравнения

$$yt_i := fit_2 \cdot x_i + fit_1$$

Вычисление коэффициента детерминации R^2 для "общего" уравнения

$$S_{resid} = \sum_{i=1}^N (y_i - yt_i)^2$$

Коэффициент детерминации для "общего" уравнения

$$R_{square} = 1 - \frac{S_{resid}}{S_{total}} \quad R_{square} = 0.976$$

Рис.19 Решение задачи 1 в системе MathCAD (продолжение).

Расчеты для "специализированного" уравнения

z	1	10
1	15	
2	33	
3	34	
4	44	
5	54	
6	63	
7	74	
z =	8	91
	9	95
	10	95
	11	107
	12	111
	13	119
	14	123
	15	123
	16	128

$$S_{zz} := \sum_{i=1}^N (z_i)^2 \quad S_{zy} := \sum_{i=1}^N (z_i \cdot y_i) \quad c := \frac{S_{zy}}{S_{zz}}$$

$S_{zz} = 3.46 \times 10^5 \quad S_{zy} = 8.718 \times 10^6$ Коэффициент "с"
"специализированного"
уравнение
 $c = 25.197$

"Специализированное" уравнение
 $c_spec(ro) := c \cdot (ro - 1000)$

Теоретические значения y для "специализированного" уравнения
 $yt_spec_i := c_spec(x_i)$

Вычисление коэффициента детерминации R^2 для "специализированного" уравнения

$$S_{resid_spec} := \sum_{i=1}^N (y_i - yt_spec_i)^2 \quad S_{resid_spec} = 1.137 \times 10^6$$

Коэффициент детерминации R^2 для "специализированного" уравнения
 $Rsquare_spec := 1 - \frac{S_{resid_spec}}{S_{total}} \quad Rsquare_spec = 0.971$

Рис.20 Решение задачи 1 в системе MathCAD (продолжение).

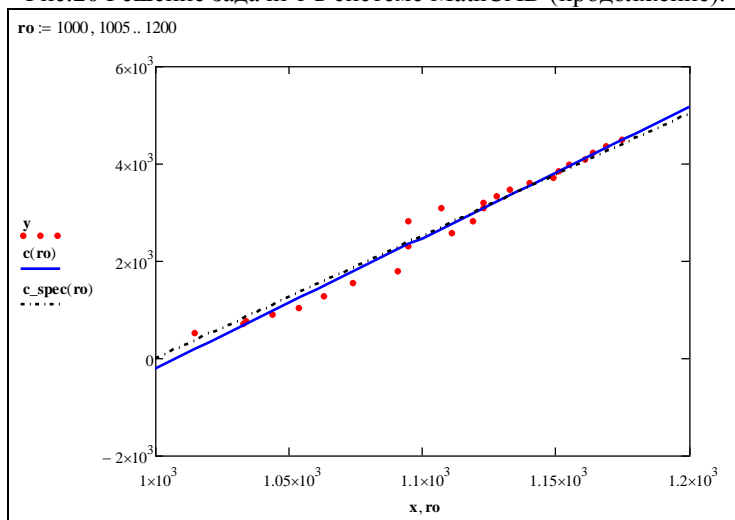


Рис.21 Решение задачи 1 в системе MathCAD (продолжение).

Вычисление прогнозного значения для	$x_prognoz = 1.159 \times 10^3$
"Специализированное" уравнение	$c_spec(x_prognoz) = 4.006 \times 10^3$
"Общее" уравнение	$c(x_prognoz) = 4.059 \times 10^3$
Разности	
Абсолютная	$delta := c(x_prognoz) - c_spec(x_prognoz)$ $delta = 52.393$
Относительная,%%	$\frac{delta}{c(x_prognoz)} \cdot 100 = 1.291$

Рис.22 Решение задачи 1 в системе MathCAD (окончание).

Все вычисленные величины и зависимости в системе MathCAD полностью совпали с соответствующими значениями, вычисленными в Excel.

Основные выводы по решению задачи. Эмпирические зависимости содержания ионов Cl^- (c , мг-экв/л) от плотности воды (ρ , кг/м³), могут быть записаны в виде:

$$c(\rho) = 25.20 \cdot (\rho - 1000) \quad \text{— «специализированное» уравнение;}$$

$$c(\rho) = 26.879 \cdot \rho - 27094.23 \quad \text{— уравнение «общего» вида.}$$

Сопоставим величины коэффициентов детерминации $R_{12}^2 = 0.9710$ и $R_{24}^2 = 0.9756$. Как и следовало ожидать, $R_{12}^2 < R_{24}^2$, т.е. уравнение «общего» вида формально лучше описывает экспериментальные данные, чем «специализированное» уравнение. Заметим, что это отличие незначительно, поскольку объясненная доля дисперсии величины « c » «общим» уравнением всего на 0,45% ($R_{12}^2 - R_{24}^2 = 0.9756 - 0.9710 = 0,0045$) больше объясненной доли дисперсии величины « c » «специализированным» уравнением. Таким образом, принимая во внимание, что R_{12}^2 весьма близко к единице и это уравнение имеет простой и ясный физический смысл, приходим к выводу, что для описания зависимости содержания

ионов Cl^- от плотности воды, поступающей в скважину вместе с нефтью содержания можно использовать «специализированное» уравнение. Другим доводом в пользу использования «специализированного» уравнения является тот факт, что графики этих зависимостей, приведенные на рис.21 отличаются весьма незначительно.

Прогнозные значения величины содержания ионов Cl^- вычисленные по этим уравнениям при $\rho_{\text{прогноз}} 1159,0$ равны 4006,26 и 4058,65 соответственно. Отклонение этих величин незначительно и равно 1,31 проценту.

ВАРИАНТЫ ЗАДАНИЙ

Задача 1

Варианты задачи 1 приведены на рис.28-30.

	E	F	G	H	I	J	K	L
1	Зависимость содержания ионов Cl⁻ от плотности пластовой воды							
2	Вариант 1		Вариант 2		Вариант 3			
3	Содержание ионов Cl ⁻ , мг.экв-л	Плотность пластовой воды, кг/ м ³ куб.	Содержание ионов Cl ⁻ , мг.экв-л	Плотность пластовой воды, кг/ м ³ куб.	Содержание ионов Cl ⁻ , мг.экв-л	Плотность пластовой воды, кг/ м ³ куб.		
4	X	Y	X	Y	X	Y		
5	1015,0	380,76	1024,0	704,76	1018,0	316,76		
6	1033,0	723,58	1032,0	984,58	1032,0	752,64		
7	1034,0	901,64	1034,0	952,64	1037,0	1061,58		
8	1044,0	881,52	1040,0	903,52	1040,0	1117,52		
9	1054,0	1196,40	1046,0	1170,40	1045,0	1175,40		
10	1063,0	1378,28	1062,0	1636,28	1061,0	1385,28		
11	1074,0	1751,16	1065,0	1793,16	1066,0	1606,16		
12	1091,0	2312,92	1088,0	2482,92	1085,0	2483,92		
13	1095,0	2389,92	1093,0	2239,92	1086,0	2372,92		
14	1095,0	2464,80	1097,0	2630,80	1098,0	2508,80		
15	1107,0	2778,68	1108,0	2640,68	1108,0	2636,68		
16	1111,0	2825,68	1114,0	2923,68	1111,0	2915,68		
17	1119,0	3025,56	1117,0	2982,56	1119,0	3008,56		
18	1123,0	3264,56	1123,0	2870,56	1121,0	2942,56		
19	1123,0	3320,50	1129,0	3197,50	1122,0	3369,50		
20	1128,0	3370,44	1134,0	3510,44	1133,0	3365,38		
21	1133,0	3374,38	1135,0	3585,38	1134,0	3424,44		
22	1140,0	3743,32	1140,0	3529,26	1140,0	3764,32		
23	1149,0	3644,26	1143,0	3403,32	1144,0	3647,26		
24	1151,0	3779,20	1154,0	3727,20	1146,0	3847,20		
25	1155,0	4148,14	1155,0	4005,14	1154,0	3812,14		
26	1161,0	4010,08	1156,0	4116,08	1160,0	4183,08		
27	1164,0	4125,02	1161,0	4127,02	1165,0	4285,02		
28	1169,0	4477,96	1172,0	4464,96	1174,0	4180,96		
29	1175,0	4536,90	1174,0	4428,90	1175,0	4393,90		

Рис.28. Варианты задачи 1 (начало).

	N	O	P	Q	R	S	T	U
1	Зависимость содержания ионов Cl⁻ от плотности пластовой воды							
2	Вариант 4		Вариант 5		Вариант 6			
3	Содержание ионов Cl ⁻ , мг.экв-л	Плотность пластовой воды, кг/ м ³ куб.	Содержание ионов Cl ⁻ , мг.экв-л	Плотность пластовой воды, кг/ м ³ куб.	Содержание ионов Cl ⁻ , мг.экв-л	Плотность пластовой воды, кг/ м ³ куб.		
4	X	Y	X	Y	X	Y		
5	1018,0	471,76	1022,0	564,76	1021,0	457,76		
6	1032,0	802,64	1030,0	935,64	1029,0	777,64		
7	1033,0	938,58	1033,0	1091,58	1037,0	809,58		
8	1036,0	894,52	1040,0	1045,52	1039,0	905,52		
9	1052,0	1373,40	1050,0	1287,40	1054,0	1141,40		
10	1064,0	1378,28	1058,0	1586,28	1065,0	1540,28		
11	1071,0	1644,16	1066,0	1860,16	1072,0	1773,16		
12	1089,0	2386,92	1086,0	2162,92	1089,0	2367,92		
13	1095,0	2496,92	1090,0	2137,92	1093,0	2292,92		
14	1096,0	2399,80	1103,0	2597,80	1099,0	2641,80		
15	1108,0	2782,68	1107,0	2774,68	1110,0	2803,68		
16	1109,0	2916,68	1109,0	2730,68	1113,0	2919,68		
17	1117,0	3020,56	1116,0	2881,56	1116,0	3256,56		
18	1123,0	3004,56	1124,0	3070,56	1117,0	2984,56		
19	1126,0	3228,50	1128,0	3122,50	1124,0	3276,50		
20	1128,0	3291,44	1134,0	3493,44	1127,0	3232,44		
21	1135,0	3531,38	1137,0	3466,38	1132,0	3285,38		
22	1141,0	3448,32	1141,0	3755,26	1142,0	3770,32		
23	1148,0	3809,20	1144,0	3585,32	1148,0	3993,20		
24	1150,0	3786,26	1151,0	3764,20	1149,0	3862,26		
25	1158,0	3878,14	1160,0	4084,14	1156,0	4150,14		
26	1160,0	4271,08	1161,0	4158,08	1160,0	4262,08		
27	1167,0	4038,02	1161,0	4096,02	1163,0	4023,02		
28	1170,0	4314,96	1167,0	4326,96	1170,0	4516,96		
29	1178,0	4595,90	1178,0	4602,90	1175,0	4662,90		

Рис.29. Варианты задачи 1 (продолжение).

	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG
1	Зависимость содержания ионов Cl- от плотности пластовой воды										
2	Вариант 7		Вариант 8		Вариант 9		Вариант 10				
3	Содержание ионов Cl-	Плотность пластовой воды, кг/м куб.	Содержание ионов Cl-	Плотность пластовой	Содержание ионов Cl-	Плотность пластовой	Содержание ионов Cl-	Плотность пластовой	Содержание ионов Cl-	Плотность пластовой	
4	X	Y	X	Y	X	Y	X	Y	X	Y	
5	1020,0	366,76	1018,0	563,76	1021,0	703,76	1022,0	674,76			
6	1032,0	843,58	1031,0	593,64	1030,0	836,64	1034,0	724,64			
7	1034,0	753,64	1037,0	1181,52	1036,0	1133,52	1036,0	993,52			
8	1037,0	838,52	1038,0	760,58	1039,0	702,58	1038,0	723,58			
9	1049,0	1477,40	1048,0	1338,40	1049,0	1336,40	1053,0	1260,40			
10	1055,0	1563,28	1058,0	1487,28	1060,0	1622,28	1056,0	1524,28			
11	1074,0	1761,16	1069,0	1926,16	1068,0	1905,16	1071,0	1905,16			
12	1090,0	2492,92	1089,0	2158,92	1090,0	2110,92	1089,0	2320,92			
13	1095,0	2144,92	1090,0	2269,92	1091,0	2323,92	1090,0	2201,92			
14	1101,0	2501,80	1095,0	2622,80	1099,0	2606,80	1097,0	2614,80			
15	1109,0	2946,68	1113,0	2842,68	1106,0	2867,68	1111,0	2625,68			
16	1109,0	2705,68	1115,0	2821,68	1110,0	2899,68	1111,0	2772,68			
17	1116,0	3002,56	1115,0	3227,56	1117,0	3246,56	1117,0	2875,56			
18	1118,0	2897,56	1125,0	3084,56	1122,0	3223,56	1118,0	3149,56			
19	1128,0	3355,44	1126,0	3128,50	1125,0	3210,50	1121,0	3340,50			
20	1129,0	3194,50	1129,0	3212,44	1126,0	3335,44	1131,0	3303,44			
21	1133,0	3637,38	1132,0	3319,38	1131,0	3566,38	1135,0	3541,38			
22	1138,0	3447,32	1143,0	3712,32	1139,0	3497,32	1135,0	3566,32			
23	1141,0	3597,26	1143,0	3528,26	1148,0	3751,20	1148,0	3532,26			
24	1146,0	3955,20	1152,0	3952,14	1149,0	3535,26	1154,0	3951,20			
25	1151,0	3950,14	1153,0	3904,20	1151,0	3910,14	1155,0	3963,08			
26	1161,0	4015,08	1163,0	4098,08	1164,0	4276,08	1158,0	3934,14			
27	1161,0	4391,02	1165,0	4491,96	1164,0	4289,02	1162,0	4190,02			
28	1173,0	4411,96	1169,0	4250,02	1165,0	4474,96	1165,0	4535,96			
29	1174,0	4423,90	1178,0	4502,90	1177,0	4540,90	1178,0	4537,90			

Рис.30. Варианты задачи 1 (окончание).

Задача 2

Исследовать зависимость насыщения подземных вод газом (Y) от общей минерализации (X , г/л). Для этого, используя МНК построить уравнение $Y = a + b \cdot X$ и исследовать его качество. Экспериментальные данные приведены на рис.31-34.

ЗАВИСИМОСТЬ НАСЫЩЕНИЯ ПОДЗЕМНЫХ ВОД ГАЗОМ ОТ ОБЩЕЙ МИНЕРАЛИЗАЦИИ						
		Вариант 1		Вариант 2		
	Общая минерализация, г/л	Коэффициент насыщения	Общая минерализация, г/л	Коэффициент насыщения	Общая минерализация, г/л	Коэффициент насыщения
	X	Y	X	Y	X	Y
5	21,0	0,41	20,6	0,40	21,2	0,49
6	21,0	0,50	21,3	0,54	21,4	0,38
7	22,0	0,42	22,4	0,59	21,6	0,40
8	22,0	0,55	22,5	0,42	22,3	0,55
9	24,0	0,46	23,6	0,41	24,5	0,42
10	25,0	0,70	25,0	0,70	24,5	0,75
11	28,0	0,80	27,6	0,82	28,0	0,81
12	29,0	0,81	29,0	0,83	29,2	0,86
13	30,1	0,60	30,4	0,64	29,9	0,62
14	31,0	0,83	31,1	0,76	31,1	0,81
15	31,1	0,71	31,3	0,86	31,2	0,71
16	32,6	0,75	32,3	0,73	32,4	0,79
17	33,0	0,85	33,2	0,86	32,7	0,88
18	34,0	0,85	34,2	0,84	34,2	0,81
19	35,0	0,93	35,2	0,96	35,4	0,93
20	36,0	0,82	36,0	0,89	35,6	0,83
21	36,0	0,85	36,4	0,86	36,5	0,80
22	37,0	0,90	37,5	0,90	37,0	0,88
23	38,0	0,90	37,5	0,86	37,7	0,93
24	38,4	0,84	38,3	0,99	38,0	0,86
25	38,4	1,00	38,8	0,88	38,7	1,02
26	39,0	0,90	39,2	0,95	39,0	0,90
27	41,0	1,00	41,0	0,99	41,4	1,00
28	43,0	1,05	43,1	1,07	43,3	1,04
29	43,8	0,95	43,7	1,02	43,6	0,98
30	44,0	1,05	43,8	0,99	43,8	1,07

Рис.31. Варианты исходных данных к задаче 2(начало).

	S	T	U	V	W	X	Y	Z	AA	AB
1										
2		Вариант 6			Вариант 7			Вариант 8		
3		Общая минерализация, г/л	Коэффициент насыщения		Общая минерализация, г/л	Коэффициент насыщения		Общая минерализация, г/л	Коэффициент насыщения	
4		X	Y		X	Y		X	Y	
5		20,8	0,48		20,6	0,52		21,0	0,54	
6		21,4	0,39		21,0	0,38		21,3	0,40	
7		21,5	0,39		21,7	0,38		21,8	0,50	
8		21,6	0,52		22,4	0,56		22,5	0,44	
9		24,1	0,45		24,2	0,43		23,6	0,47	
10		24,8	0,68		25,4	0,72		25,3	0,71	
11		26,2	0,78		27,6	0,79		27,9	0,76	
12		28,7	0,84		28,9	0,83		28,7	0,81	
13		30,5	0,65		29,8	0,63		29,9	0,56	
14		30,9	0,82		30,6	0,82		31,0	0,70	
15		31,3	0,76		31,0	0,73		31,2	0,86	
16		32,7	0,70		32,6	0,86		32,4	0,80	
17		33,3	0,88		33,1	0,74		32,7	0,84	
18		34,4	0,86		34,1	0,83		33,8	0,85	
19		34,9	0,95		35,2	0,96		35,3	0,92	
20		36,3	0,83		35,7	0,82		35,7	0,84	
21		36,5	0,84		36,1	0,90		36,2	0,81	
22		37,4	0,94		37,1	0,93		37,3	0,93	
23		38,0	0,94		37,8	0,87		37,9	0,85	
24		38,1	1,01		38,1	0,99		37,9	0,97	
25		38,6	0,88		38,6	0,82		38,1	0,82	
26		38,9	0,89		39,5	0,88		39,0	0,91	
27		41,1	0,96		41,1	1,04		41,1	1,02	
28		43,5	1,08		43,4	1,05		43,2	1,06	
29		43,8	0,97		43,9	0,98		43,5	0,97	
30		44,0	1,04		43,9	1,09		44,0	1,03	
31										

Рис.32. Варианты исходных данных к задаче 2. (начало).

Microsoft Excel - МИК_012

Файл Правка Вид Вставка Формат Сервис Данные Окно Справка

10 Ж К Ч

С16 0,75

	J	K	L	M	N	O	P	Q	R	S
1	ЦИИ									
2		Вариант 3			Вариант 4			Вариант 5		
3		Общая минерализация, г/л	Коэффициент насыщения		Общая минерализация, г/л	Коэффициент насыщения		Общая минерализация, г/л	Коэффициент насыщения	
4		X	Y		X	Y		X	Y	
5		20,5	0,52		21,0	0,43		20,9	0,53	
6		21,3	0,36		21,4	0,50		21,0	0,44	
7		22,0	0,58		22,0	0,43		21,7	0,60	
8		22,5	0,43		22,4	0,57		22,3	0,37	
9		23,7	0,48		24,2	0,50		24,2	0,50	
10		25,1	0,73		25,2	0,68		25,0	0,72	
11		28,3	0,82		27,6	0,82		28,3	0,83	
12		28,8	0,78		29,4	0,82		29,1	0,78	
13		29,8	0,61		30,1	0,60		30,3	0,62	
14		30,9	0,81		30,8	0,80		31,4	0,87	
15		31,5	0,67		31,4	0,67		31,5	0,71	
16		32,4	0,76		32,1	0,73		32,6	0,80	
17		32,7	0,86		33,4	0,85		33,0	0,71	
18		34,2	0,84		34,0	0,83		33,5	0,88	
19		35,2	0,95		35,1	0,92		35,0	0,90	
20		35,6	0,86		35,7	0,85		36,3	0,86	
21		36,2	0,78		36,1	0,90		36,5	0,87	
22		36,7	0,87		36,5	0,91		36,5	0,92	
23		38,2	0,90		37,9	1,02		37,6	0,93	
24		38,5	1,02		38,3	0,92		37,9	0,88	
25		38,8	0,79		38,4	0,81		38,3	1,01	
26		39,4	0,95		38,7	0,87		39,3	0,89	
27		40,6	1,01		40,6	1,02		41,4	0,99	
28		42,6	1,03		43,3	1,06		42,5	1,10	
29		43,9	1,05		43,8	1,08		43,8	1,00	
30		44,0	0,92		44,2	0,93		43,8	1,06	
31										

Готово SCRL

Рис.33.Варианты исходных данных к задаче 2 (продолжение).

	AC	AD	AE	AF	AG
1					
2	Вариант 9			Вариант 10	
3	Общая минерализация, г/л	Коэффициент насыщения		Общая минерализация, г/л	Коэффициент насыщения
4	X	Y		X	Y
5	20,5	0,38		20,8	0,51
6	20,6	0,45		21,0	0,42
7	22,0	0,38		21,5	0,51
8	22,3	0,52		22,2	0,41
9	23,8	0,49		23,9	0,45
10	25,1	0,72		25,3	0,68
11	28,2	0,84		28,4	0,84
12	28,5	0,85		28,8	0,80
13	30,3	0,65		30,0	0,60
14	31,4	0,79		31,4	0,79
15	31,4	0,69		31,5	0,68
16	32,5	0,71		32,8	0,78
17	32,7	0,82		33,1	0,81
18	34,3	0,84		34,1	0,85
19	34,8	0,90		35,1	0,97
20	36,3	0,83		35,9	0,77
21	36,4	0,87		36,4	0,87
22	36,5	0,92		37,4	0,94
23	38,1	0,86		38,2	0,90
24	38,3	0,83		38,4	0,86
25	38,5	1,03		38,7	0,95
26	38,9	0,90		38,9	0,98
27	41,0	0,99		41,2	0,97
28	43,1	1,05		43,2	1,09
29	43,9	1,03		43,7	0,94
30	44,1	0,94		43,7	1,00
31					
32					

Рис.34. Варианты исходных данных к задаче 2. (окончание)

Задача 3.

Исследовать зависимость насыщения подземных вод газом (Y) от глубины залегания водоносных горизонтов ($X, м$). Для этого, используя МНК построить уравнение $Y = a + b \cdot X$ и исследовать его качество. Экспериментальные данные для расчетов приведены на рис.35-37.

ЗАВИСИМОСТЬ НАСЫЩЕНИЯ ПОДЗЕМНЫХ ВОД ГАЗОМ ОТ ГЛУБИНЫ ЗАЛЕГАНИЯ ВОДОНОСНЫХ ГОРИЗОНТОВ									
Вариант 1		Вариант 2		Вариант 3		Вариант 4			
Глубина залегания	Коэффициент насыщения	Глубина залегания	Коэффициент насыщения	Глубина залегания	Коэффициент насыщения	Глубина залегания	Коэффициент насыщения	Глубина залегания	Коэффициент насыщения
X	Y	X	Y	X	Y	X	Y	X	Y
2840	0,38	2810	0,36	2860	0,42	2830	0,55		
2850	0,54	2890	0,47	2860	0,5	2850	0,46		
2850	0,51	2900	0,47	2870	0,4	2860	0,64		
2860	0,44	2920	0,55	2880	0,41	2870	0,43		
2900	0,38	2910	0,61	2890	0,55	2880	0,54		
2930	0,59	2890	0,41	2910	0,62	2880	0,44		
2950	0,72	2940	0,72	2930	0,78	2920	0,7		
2970	0,88	2950	0,78	2960	0,73	2950	0,75		
2980	0,74	2960	0,77	2970	0,85	2970	0,79		
2990	0,72	3030	0,76	2990	0,68	2980	0,84		
3010	0,86	2970	0,84	3010	0,94	2980	0,85		
3030	0,75	2990	0,89	3020	0,82	2990	0,87		
3030	0,84	3010	0,81	3030	0,82	3010	0,82		
3030	0,92	3030	0,88	3030	0,86	3030	0,86		
3040	0,89	3080	0,91	3060	0,85	3040	0,8		
3040	0,95	3000	0,95	3070	0,85	3040	0,94		
3040	1,01	3010	0,87	3070	0,85	3050	0,83		
3040	1,14	3020	0,94	3070	0,92	3050	0,96		
3050	0,93	3090	0,98	3070	0,99	3070	0,92		
3080	0,86	3090	1,02	3090	1,06	3080	1,04		
3090	1,06	3120	1,09	3100	1,03	3080	1,06		
3100	1,05	3120	1,11	3110	1,07	3120	0,98		

Рис. 35. Исходные данные к задаче 3(начало).

Microsoft Excel - MNK_OIL2ot belaeva

Файл Правка Вид Вставка Формат Сервис Данные Окно Справка Введите вопрос

10 Ж К Ч

Ф4

	L	M	N	O	P	Q	R	S	T	U	V	W
1												
2		Вариант 5			Вариант 6			Вариант 7			Вариант 8	
3		Глубина залега- ния	Кэффи- циент насыще- ния		Глубина залега- ния	Кэффи- циент насыще- ния		Глубина залега- ния	Кэффи- циент насыще- ния		Глубина залега- ния	Кэффи- циент насыще- ния
4		X	Y		X	Y		X	Y		X	Y
5		2860	0,52		2850	0,45		2820	0,47		2840	0,4
6		2880	0,37		2860	0,52		2820	0,4		2860	0,48
7		2900	0,45		2870	0,42		2870	0,4		2860	0,55
8		2900	0,6		2910	0,53		2870	0,53		2890	0,45
9		2920	0,73		2930	0,39		2870	0,63		2910	0,57
10		2930	0,62		2930	0,78		2920	0,41		2930	0,37
11		2930	0,39		2950	0,57		2950	0,7		2950	0,73
12		2980	0,77		2980	0,69		3000	0,76		2970	0,74
13		2980	0,87		3010	0,78		3000	0,93		2970	0,79
14		3000	0,9		3030	0,81		3010	0,76		3010	0,88
15		3010	0,7		3030	0,83		3030	0,94		3010	0,88
16		3020	0,72		3040	0,82		3040	0,74		3010	0,95
17		3040	0,81		3040	0,89		3050	0,82		3020	0,78
18		3040	0,88		3040	1,01		3060	0,88		3020	0,94
19		3040	0,95		3050	0,89		3060	0,93		3030	0,86
20		3050	1,02		3050	0,85		3070	0,87		3050	0,87
21		3060	0,81		3050	0,92		3070	1,07		3060	1,09
22		3060	0,87		3050	0,9		3080	0,87		3070	0,94
23		3070	0,99		3050	1,02		3080	1,02		3080	0,89
24		3080	0,9		3060	1,05		3090	0,93		3100	1
25		3090	1,03		3070	0,98		3090	1		3110	1
26		3120	1,09		3070	1,1		3120	1,08		3120	1,08
27												

Лист1 Решение Зад_1 данные Решение Зад_1 формулы Зад_1 F4

Действия Автофикуры

Готово SCRL

Рис. 36. Исходные данные к задаче 3(продолжение)..

Microsoft Excel - MNK_OIL2ot belaeva

Файл Правка Вид Вставка Формат Сервис Данные Окно Справка

10 Ж К Ч

Ф4

	X	Y	Z	AA	AB	AC	AD
1							
2		Вариант 9			Вариант 10		
3		Глубина залега ния	Кэффи циент насыще ния		Глубина залега ния	Кэффи циент насыще ния	
4		X	Y		X	Y	
5		2850	0,47		2820	0,41	
6		2850	0,44		2830	0,47	
7		2860	0,44		2830	0,51	
8		2860	0,44		2840	0,39	
9		2890	0,58		2870	0,41	
10		2920	0,51		2900	0,58	
11		2960	0,71		2960	0,7	
12		2960	0,73		2990	0,83	
13		2980	0,75		2990	0,9	
14		2990	0,89		3000	0,83	
15		3000	0,86		3000	0,92	
16		3000	0,89		3010	0,72	
17		3010	0,86		3010	0,95	
18		3030	1		3030	0,93	
19		3040	0,86		3040	0,79	
20		3050	0,89		3040	0,88	
21		3050	0,97		3050	0,86	
22		3060	0,82		3060	1,01	
23		3060	1,03		3060	1,03	
24		3080	0,88		3060	1,07	
25		3080	1,01		3090	0,94	
26		3140	1,1		3110	1,07	
27							
28							

Лист1 Решение Зад_1 данные Ре

Действия Автофигуры

Готово SCRL

Рис. 37. Исходные данные к задаче 3(окончание).

СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

1. *Мановян А.К.* Технология первичной переработки нефти и природного газа. М: Химия, 2001.- 568 с.
2. *Очков В.Ф.* Mathcad 14 для студентов и инженеров: русская версия. ВHV-Петербург, 2009 г., 512 с.
3. *Дмитриев В.В., Ярг Л.А.* Методы и качество лабораторного изучения грунтов: учебное пособие. – М.: КДУ, 2008. – 542 с.
4. *Ломтадзе В.Д.* Словарь по инженерной геологии. – СПб.: СПГГИ (ТУ), 1999, 360 с.
5. *Бондарик Г.К., Ярг Л.А.* Инженерно-геологические изыскания. - М.: КДУ, 2007 г., 424 с.
6. *Иванов И.П., Трижвинский Ю.Б.* Инженерная геодинамика. – СПб.: Наука, 2001.-416 с.
7. Материалы сайта «Степень газонасыщения подземных вод продуктивной толщи юрского гидрогеологического комплекса Харампурского мегавала.»Л.А. Новиков (Томский филиал ИГНиГ СО РАН).