МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

В.С. Деева

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ В НЕФТЕГАЗОВОМ ДЕЛЕ

*Рекомендовано в качестве учебного пособия
Редакционно-издательским советом
Томского политехнического университета*

Издательство
Томского политехнического университета
2018

УДК 622.276:004.94(075.8)
ББК 33.36:32.971.3я73
Д26

Деева В.С.

Д26 Компьютерное моделирование в нефтегазовом деле : учебное пособие / В.С. Деева ; Томский политехнический университет. – Томск : Изд-во Томского политехнического университета, 2018. – 86 с.

ISBN 978-5-4387-0806-3

В пособии рассмотрены основные разделы курса «Компьютерное моделирование в нефтегазовом деле». Состоит из восьми теоретических разделов, в которых описаны методы приближенного решения нелинейных алгебраических уравнений, систем линейных алгебраических уравнений, задач аппроксимации и интерполяции, краевых задач и задач Коши для обыкновенных дифференциальных уравнений, а также статистические методы и законы распределения. Теоретическое изложение иллюстрируется примерами расчетов в программном продукте Excel, облегчающими самостоятельное знакомство с материалом

Предназначено для студентов, обучающихся по всем специальностям направления 21.03.01 «Нефтегазовое дело».

УДК 622.276:004.94(075.8)
ББК 33.36:32.971.3я73

Рецензенты

Доктор физико-математических наук, профессор
заведующий кафедрой динамики полета
Томского государственного университета
В.И. Биматов

Доктор технических наук, профессор
заведующий кафедрой химии Томского государственного
архитектурно-строительного университета
Ю.С. Саркисов

ISBN 978-5-4387-0806-3

Предисловие

Специалистам в нефтегазовой отрасли необходимы знания в области математического моделирования для решения задач моделирования и оптимизации, поскольку для многих процессов строгое физическое моделирование невозможно или сложно. В условиях модели наиболее просто и оптимально по времени можно исследовать влияние различных режимных параметров, оптимизировать процесс, а также проверить новые эффективные решения.

Изложение настоящего учебного пособия обусловлено отсутствием подобных сведений в смежных дисциплинах и основано на предположении о знакомстве читателя с курсами высшей математики и основ программирования, физики, термодинамики, гидрогазодинамики, согласованными с государственным стандартом программ обучения в университете технического профиля.

Целью данной работы является систематизация, углубление и закрепление знаний по дисциплине «Компьютерное моделирование в нефтегазовом деле». Для ее достижения автором применен системный подход к использованию современных математических инструментов: кроме изложения теоретического материала в нем есть примеры использования математического аппарата в современном программном табличном процессоре Excel для каждой темы.

Главная задача специалиста, изучающего любой физический процесс, заключается в выявлении закономерностей поведения этого процесса, получении функциональной зависимости между входными параметрами и выходным результатом. Часто это сводится к математически формализованному описанию физических процессов и технических систем.

Разделы в учебнике расположены в порядке возрастания сложности соответственно темам дисциплины «Компьютерное моделирование в нефтегазовом деле». Каждый раздел содержит в себе примеры решения типовых задач, которые встречаются при решении проблем современной нефтегазодобычи и подготовки к ним. Такое построение удобно для самостоятельной работы студентов, а также проведение лабораторных и практических занятий. При подготовке книги был учтен многолетний опыт преподавания дисциплины «Компьютерное моделирование в нефтегазовом деле» студентам нефтяных специальности в Научно-исследовательском Томском Политехническом университете.

Автор выражает благодарность рецензентам – В.И. Биматову, заведующему кафедрой динамики полета Томского государственного университета, профессору, доктору физико-математических наук, а также Ю.С. Саркисову, заведующему кафедрой химии Томского государственного архитектурно-строительного университета, профессору, доктору технических наук, за ценные замечания и помощь при подготовке пособия.

Введение

В настоящее время процесс подготовки специалистов нефтегазовой отрасли не представляется возможным без подготовки в области информационных технологий, которые используются в профессиональной деятельности каждого работника. В формировании профессиональных компетенций активную роль играет обучение определенным профессиональным навыкам работы в компьютерных программах, что позволяет применять их и в дальнейшем в профессиональной деятельности. В теоретической части изучаются сами программные продукты, а на практике решаются задачи с их использованием, выполняются конкретные проекты. При моделировании можно исследовать влияние определенных параметров, провести, оптимизировать, а также проверить новые эффективные решения, которые в реальных условиях провести сложно или даже невозможно.

Практически на всех предприятиях нефтегазового сектора быстрыми темпами накапливается большое количество информации, полученной при исследованиях: результаты геологической документации скважин, выработок, спектральных и химических анализов руд, пород и минералов, данные геофизических и геохимических измерений и др. Исходная информация чаще всего накапливается в бумажном виде. Ее необходимо перенести на машинные носители, преобразуя в цифровую или символьную форму с помощью различных технических средств. Одно из важнейших направлений научно-технического прогресса – внедрение автоматизированных методов накопления, хранения, обработки и передачи информации с целью повышения эффективности.

Можно сказать, что с 90-х гг. наступил новый период, связанный с широким распространением персональных компьютеров, которые постепенно стали доступны каждому, позволяя оперативно обрабатывать поступающую информацию.

В настоящее время современные компьютеры в народном хозяйстве, в том числе в нефтегазовой отрасли, используются:

1) для накопления информации и хранения ее на носителях. При необходимости производится систематизация (сортировка, получение выборок и пр.) по определенному критерию;

2) обработки информации преимущественно на базе различных доступных предприятию методов для описания или сравнения свойств, классификации объектов и прогнозирования их дальнейшего поведения;

3) для моделирования научных и прикладных задач, решения которых находятся методами математического моделирования заданных объектов и явлений;

4) автоматизации технологических операций, распространенных в нефтегазовом деле, таких как построение геологических карт и разрезов, подсчет запасов и ресурсов, проектирование разведочных и эксплуатационных работ и др.

Построение компьютерной модели базируется на отвлечении от конкретной природы изучаемого объекта или явления. Чем больше существенных, важных свойств выявлено и учтено в компьютерной модели, тем более приближенной она окажется к реальной модели. Компьютерное моделирование заключается в проведении серии вычислительных экспериментов на компьютере, целью которых является анализ, интерпретация и сопоставление результатов моделирования с реальным поведением изучаемого объекта и, при необходимости, последующее уточнение модели.

Связь курса математики с профессиональной подготовкой будущих инженеров нефтегазового дела – одно из условий, обеспечивающих целостность и глубину знаний обучающихся, их самостоятельность в решении инженерных задач. После окончания обучения студенты должны уметь корректно ставить задачи с четкой формулировкой ограничений, грамотно выбирать математическую модель и рассчитывать ее параметры, делать выводы и давать рекомендации на основе полученных результатов. Однако для интерпретации результатов студент должен иметь хорошую подготовку в области численного моделирования, теории вероятностей, математической статистики.

Математические методы исследования используются для решения задач с определенной точностью и достоверностью путем математического моделирования, и при этом на выходе получается оптимальное решение. Понятие «математические методы» ассоциируется с определением «математическая модель». Математическая модель – это совокупность алгебраических формул, по которым вычисляются искомые величины. Однако чаще всего поведение параметров описывается сложными уравнениями или системой уравнений, найти решение которых можно только с использованием современных быстродействующих ЭВМ. В данном учебном пособии представлены практически значимые задачи нефтегазовой отрасли, которые решаются математическими методами (линейные и нелинейные уравнения, системы линейных уравнений, аппроксимация, интерполяция, соответствие законам распределения и статистические методы) с помощью современных компьютерных систем – Excel – мощный инструмент, позволяющий сосредоточить внимание на логике методов и алгоритмов, освобождая от необходимости освоения громоздких вычислительных процедур благодаря встроенным в программное обеспечение функциям.

Раздел 1. Понятия модели и моделирования

В научных исследованиях большую роль играют гипотезы, то есть предсказания, основанные на небольшом количестве наблюдений или опытов. При формулировании и проверке правильности гипотез большое значение имеет аналогия – сходство двух объектов. Гипотезы, отражающие реальный мир, должны обладать наглядностью и сводиться к удобным для исследования логическим схемам, их обычно связывают с экспериментом. Такие логические схемы, упрощающие логические построения и позволяющие проводить эксперименты для уточнения природы исследуемых явлений, называют моделями, а процесс построения модели – моделированием. Суть моделирования заключается в переходе от изучения исходного явления, процесса или технической системы непосредственно в реальности к другому явлению, процессу или технической системе, которую называют моделью. Основная цель такого перехода – облегчить исследование, сделать доступным определение интересующих нас величин, искусственно воспроизвести исследуемые явления.

Дадим определение понятию *моделирование* как замещение исследуемого оригинала (объекта или явления) его условным образом, описанием или возможно другим объектом, который называют *моделью*. Под моделью можно понимать как физический, так и абстрактный объект, свойства которого схожи со свойствами исследуемого объекта или явления и поведение которого близко к оригиналу в рамках некоторых допущений и приемлемых погрешностей. Процесс моделирования проводится с целью изучение свойств оригинала путем исследования модели, а не самого объекта, вследствие того, что на реальном объекте проводить исследования либо дорого, либо неудобно, а то и вовсе невозможно в силу ряда причин: длительности эксперимента, отсутствия реального объекта (в случае, когда он еще проектируется) и др.

В процессе моделирования исследователь разделяет свойства исходного объекта на существенные и второстепенные, исходя из заданных требований к модели. Основываясь на целевых задачах, специалисту необходимо найти в исходном объекте только те черты, которые имеют непосредственное отношение к интересующей стороне его функционирования.

Именно этим, то есть решаемой задачей и имеющимися средствами, определяются основные требования к любой модели, и поэтому можно четко выделить следующие критерии:

- адекватность – означает близкое отображение свойств объекта, при этом могут учитываться только существенные стороны объекта;
- полнота – предоставление всей необходимой информации об объекте в рамках гипотез, принятых при построении модели;
- гибкость – возможность воспроизведения различных ситуаций во всем диапазоне изменения условий и параметров;
- трудоемкость разработки должна быть приемлемой для имеющегося времени и программных средств.

Логично, что компьютерная модель сложной системы должна адекватно отображать все основные факторы и взаимосвязи и тем самым обеспечивать отражение характеристик реальных ситуаций, критериев и ограничений. При этом она должна быть достаточно универсальной, чтобы при необходимости можно было описывать близкие по назначению объекты с достаточной степенью достоверности, и в то же время достаточно простой, чтобы пользователь мог выполнить необходимые исследования, не выходя за пределы разумных затрат. Все это говорит о том, что моделирование представляет собой, скорее, искусство, чем сформировавшуюся науку, с самостоятельным набором средств отображения явлений и процессов реального мира.

Любые методы обработки экспериментальных данных содержат в своей основе явную или неявную модель изучаемого объекта или происходящего с ним явления (события).

На практике все модели можно разделить на два больших класса: физические и математические.

Физические модели – это выполненные в определенном масштабе макеты объектов. Например, модель кристаллической решетки минерала, модель идеальных кристаллов с различными наборами граней, морфологические модели рудных тел и др.

Математическая модель представляет собой совокупность представлений, предположений, гипотез и аксиом, отражающих существо изучаемого геологического объекта или явления. Она выражается в математической форме и позволяет описывать, анализировать и прогнозировать свойства объектов или последствия явлений.

В основе математического моделирования лежит принцип системного подхода. То есть для исследования при моделировании возможно выделить объект или явление (или группу), которые рассматриваются как отдельная система, имеющая какие-то физические или условные границы, а также связи как между частями системы, так и между отдельными свойствами. Объекты, расположенные за пределами системы, принимаются в модели в качестве окружающей среды.

Компьютерная модель – это программная реализация математической модели, дополненная различными служебными программами, при этом сочетающая в себе как абстрактные, так и физические черты. Как физическое устройство она может входить в состав испытательных стендов, тренажеров и виртуальных лабораторий. Компьютерная модель обладает уникальным набором полезных свойств: простота создания и модификации модели, высокая точность получаемых результатов, неограниченная функциональная сложность. Поэтому в настоящее время компьютерное моделирование так распространено.

Конечной целью моделирования может быть описание и классификация объектов, понимание геологической природы объектов и явлений, предсказание (прогнозирование) поведения или свойств системы, а в некоторых случаях и управление системой на основе контроля ее состояния. Например, при разведке и эксплуатации месторождения необходимо понять его строение и происхождение, прогнозировать количество и качество минерального сырья, управлять процессом эксплуатации с целью рационального использования недр и решать много других практических задач.

Для того чтобы создать модель, т. е. смоделировать процесс или явление, необходимо провести определенную последовательность действий.

Во-первых, надо определить систему, т. е. задать границы, перечислить входящие в объект подобъекты и их свойства, а иногда и определить взаимосвязи между ними.

Во-вторых, нужно понять, в чем будут (а каким образом) измеряться характеристики свойств объектов, входящих в систему. Другими словами, создать исходные данные для математической обработки. Добавим, что этот этап не обязателен для всех задач, тогда изучению подвергаются предполагаемые значения, заданные автором модели.

Следующий шаг – создание представления о сущности изучаемой системы, какая из существующих гипотез соответствует ее формальной сути. Можно выдвинуть несколько гипотез и, уже основываясь на последующем математическом моделировании, сделать заключение об их соответствии.

На четвертом этапе система представляется в математической форме, то есть в виде формул, правил, уравнений и прочее. Это и есть математическая постановка задачи. После этого этапа часто приходится возвращаться назад ко второму и третьему этапам для уточнения недостающих сведений.

Затем происходит исследование математической модели. Часто пятый этап сводится к решению формул и уравнений четвертого этапа и вычислению прогнозных значений свойств или параметров явлений. Это и есть получение ответа на задачу. Иногда для принятия решения нужно оценить погрешность прогнозирования. В определенных случаях, если исходные данные колеблются в некоторых пределах, можно исследовать зависимость прогнозных значений от них.

И наконец, последний этап, о котором часто забывают или пропускают, но он является неотъемлемой частью процесса моделирования – это проверка соответствия полученных результатов фактическим данным. Важно определить, соответствует ли математическая модель описанию системы, тезисам, положенным в ее основу. Оценить степень совпадения или сходства фактических данных с теоретическими, вычисленными в ходе решения математической модели. Следует отметить, что проверка не всегда возможна, особенно в тех случаях, когда получение фактических данных затруднено или невозможно.

При этом на каждом из этапов решаются разные задачи и используются отличные методы и средства. В результате математического моделирования могут быть получены различные ответы, зависящие от постановки задачи. При необходимости можно определить прогнозные значения свойств, которые трудно измерить, или оценить степень соответствия модели фактическим данным, а также сравнить модели и установить, какая из них лучше соответствует поставленной задаче, и затем выбрать одну модель.

Все объекты и явления в нефтяной отрасли являются сложными структурами, при этом каждая из них находится под влиянием большого числа факторов, которые трудно, а иногда и невозможно корректно учесть. Необходимо понимать, что любая математическая модель является приближенным отражением реальных природных систем, их математическое моделирование не может дать исчерпывающую характеристику их свойств. И для каждой природной системы можно построить несколько моделей различной степени сложности с различной достоверностью прогнозирования и надежностью. Но всегда можно выбрать оптимальный уровень сложности, при котором принятая достоверность не ухудшает работоспособность модели.

Математическое моделирование рассматривается как средство исследования процессов или явлений с помощью их математических моделей при соблюдении двух условий:

- модель обеспечивает корректное (адекватное) отображение свойств оригинала, существенных с точки зрения исследуемой операции;
- модель позволяет устранить проблемы, присущие проведению исследований на реальных объектах.

В математическом моделировании различают аналитическое и имитационное моделирование.

Аналитическое моделирование – это такое моделирование, при котором изучаются численно-математические (количественные) модели реального объекта в виде алгебраических, дифференциальных и других уравнений, а также предусматривающих осуществление однозначной вычислительной процедуры, приводящей к их точному решению.

Имитационное моделирование – такое моделирование, при котором исследуются математические модели в виде алгоритмов, воспроизводящих функционирование исследуемой системы путем последовательного выполнения большого количества элементарных операций логического типа.

Следующая классификация математических моделей основана на свойствах модели и разделяет их на три группы. В первой группе анализируются характеристики в пределах однородных совокупностей свойств объектов вне связи их с пространственным размещением – группа *статистических моделей*. Они бывают одномерные, двухмерные и многомерные.

Во второй группе анализируются пространственные координаты пунктов наблюдений, то есть изучаются *пространственные геологические поля*. Модели этой группы далее делятся на детерминированные и вероятностные. В *детерминированных моделях* состояние объекта или явления определяется только начальными данными, причем однозначно и полностью предсказуемо в пространстве. *Вероятностные модели* характеризуются тем, что их состояние в настоящем и тем более в будущем неоднозначно определяется исходными данными и их параметры могут быть предсказаны с какой-то вероятностью и только в определенном диапазоне значений.

Третья группа охватывает случайные процессы, в которых учитывается фактор времени.

Жизненный цикл любой модели, в том числе математической и компьютерной, можно разбить на следующие этапы:

1. Сбор информации об объекте, выдвижение гипотез, предмодельный анализ.

2. Определение состава и структуры (взаимосвязей) модели.

3. Построение спецификаций модели, разработка и отладка отдельных ее составляющих для сборки в единое целое.

4. Исследование модели – выбор метода исследования и разработка алгоритма (программы) моделирования.

5. Исследование адекватности, устойчивости, чувствительности модели.

6. Оценка затраченных ресурсов (средств).

7. Интерпретация, анализ результатов моделирования и установление некоторых причинно-следственных связей в исследуемой системе.

8. Генерация отчетов и проектных решений.

9. Уточнение, модификация модели, если это необходимо, и возврат к исследуемой системе с новыми знаниями, полученными с помощью моделирования.

Вопросы и задания для самопроверки

1. Что такое моделирование? С какой целью его используют?

2. Дайте определение понятия «модель». Перечислите основные требования к модели.

3. Какие дополнительные требования предъявляются к компьютерной модели?

4. Что представляет собой физическая модель, а что – математическая?

5. Перечислите и охарактеризуйте этапы моделирования.

6. В чем различие между аналитическим и имитационным моделированием?

7. Классифицируйте модели в зависимости от начальных данных и фактора времени. Более подробно опишите характеристики статистических, детерминированных и вероятностных моделей.

Раздел 2. Основы практического применения пакета Excel

2.1. Знакомство с Excel

Широкое внедрение персональных компьютеров во все сферы инженерной деятельности, в том числе в сферу разведки, добычи и переработки нефти и газа, требует от современного инженера умения пользоваться готовыми компьютерными программами для решения практических задач. Подобные программы освобождают инженера от громоздких и рутинных расчетов, связанных с большой затратой физического труда и времени. При наличии современного интерфейса работа с готовыми программами не обременительна, удобна и чем-то даже напоминает игру, хотя получаемые с ее помощью результаты имеют вполне реальное практическое приложение.

Пакеты компьютерных программ, решающих ту или иную технологическую задачу, могут использоваться как тренажеры для проверки правильности выбранных решений, как математические модели для выбора таких решений, а также как эффективные средства для обучения будущих специалистов. Но главное направление заключается в математическом моделировании месторождений, что позволяет решать вопросы, касающиеся подсчета запасов, определения качества минерального сырья, геолого-экономической оценки месторождений.

Компьютерное моделирование в нефтегазовом деле становится одним из важных составлявших современного предприятия. В условиях жесткой конкуренции внедрение новых информационных технологий компьютерного проектирования в производственный процесс предприятия позволяет:

– повысить техническое качество проектов;

– применить новые технические решения;

– сократить сроки проектирования;

– эффективно реагировать на выдвигаемые заказчиком требования;

– оперативно и качественно вносить изменения и выполнять корректировку проектов;

– выдавать заказчику проектно-сметную документацию в современных цифровых форматах;

– насыщать проектную документацию дополнительной, атрибутивной, информацией, используемой в дальнейшем в эксплуатации;

– использовать технологии параллельного проектирования;

– унифицировать проектные решения и процессы проектирования (использовать готовые фрагменты чертежей: конструктивных и геометрических элементов, унифицированных конструкций, стандартных изделий);

– использовать повторно проектные решения, данные и наработки;

– применять стратегическое проектирование;

– заменять натурные испытания и макетирования математическим моделированием;

– повышать качество управления проектированием;

– применять методы вариантного проектирования и оптимизации.

При разведке месторождений накапливается большое количество информации: геологическая документация разведочных выработок, данные опробования, результаты геофизических, геохимических исследований и др. В дальнейшем информация перерабатывается с целью построения геологических карт, разрезов, проекций рудных тел, подсчета запасов и решения других вопросов. Исходная информация накапливается и затем преобразуется в цифровую или символьную форму и хранится в виде файлов баз данных в Excel, dBase, Access, FoxPro, Paradox, Word, MS DOS и другие.

Электронная таблица Excel состоит из 16384 строк и 256 столбцов. Строки пронумерованы целыми числами от 1 до 16384, а столбцы обозначены буквами латинского алфавита A, B, …, Z, AB, …XFD. На пересечении столбца и строки располагается ячейка. Для указания на конкретную ячейку таблицы мы используем адрес, составляемый из номера строки и обозначения столбца, на пересечении которых эта ячейка находится (например, A1, F5, C20, AA3 и т. д.). В любую ячейку можно ввести исходные данные – число или текст, а также формулу для расчета производной информации. Формула начинается со знака равенства (=). Все формулы записываются в одну строку, поэтому иногда необходимо вводить дополнительные скобки, которых нет в исходных формулах. Ширину столбца и высоту строки можно изменять.

При работе в среде Excel необходимо сначала выделить объект, а затем над выделенным объектом выполнять операцию. Для выделения с помощью мышки: столбца – щелкнуть мышкой по верхней адресной строке; нескольких столбцов – щелкнуть мышкой по заголовку первого столбца и, не отпуская левую клавишу мыши, протащить ее по адресной строке по соответствующим заголовкам; строк – аналогично. Для выделения всех ячеек рабочего листа – щелкнуть мышкой по кнопке на пересечении адресных полос. Для снятия выделения достаточно щелкнуть мышкой по любой невыделенной ячейке рабочего листа.

На рис. 1 показан общий вид окна Excel.

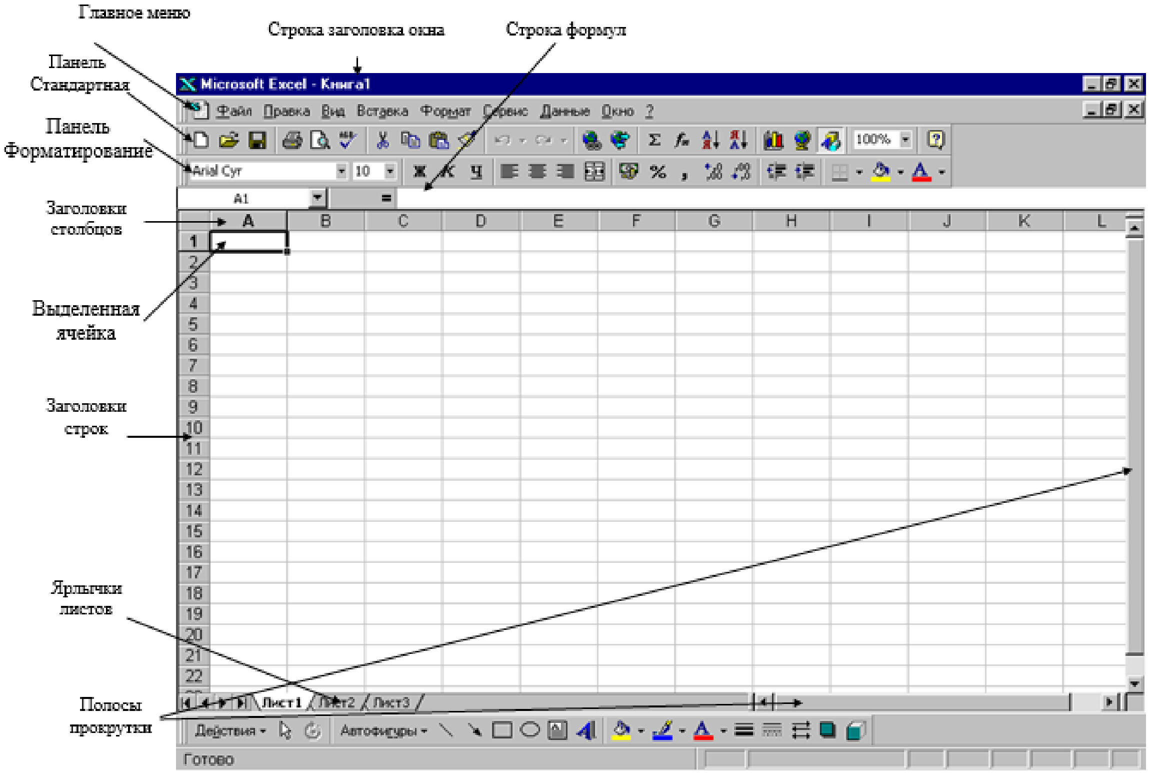


Рис. 1. Вид рабочего листа Excel

С помощью кнопок (Σ fx) выполняются операции, упрощающие ввод *функций в* ячейку электронной таблицы. Если кликнуть на кнопку со знаком суммы, то в текущей ячейке появится функция =СУММ(...). После вызова кнопки ***fx*** на экране появится диалоговое окно *Мастера функций*, после чего появится возможность выбрать функцию, вводимую в текущую ячейку.

В программе Excel можно использовать свыше 400 функций, которые разделены на категории (тематические группы): математические, финансовые, статистические, текстовые, логические, даты и времени (рис. 2).

В общем случае функция – это переменная величина, значение которой зависит от других значений (аргументов). Скобки – обязательная принадлежность функции, даже если у нее нет аргументов. Например, функция ПИ() возвращает число = 3,14... В качестве аргументов функции могут использоваться числа, адреса ячеек, диапазоны ячеек, арифметические выражения и функции. Смысл и порядок следования аргументов однозначно определен описанием функции. Например, если в ячейке G6 записана формула с функцией возведения в степень =СТЕПЕНЬ(A4,2), значением этой ячейки будет значение A4, возведенное в степень 2.

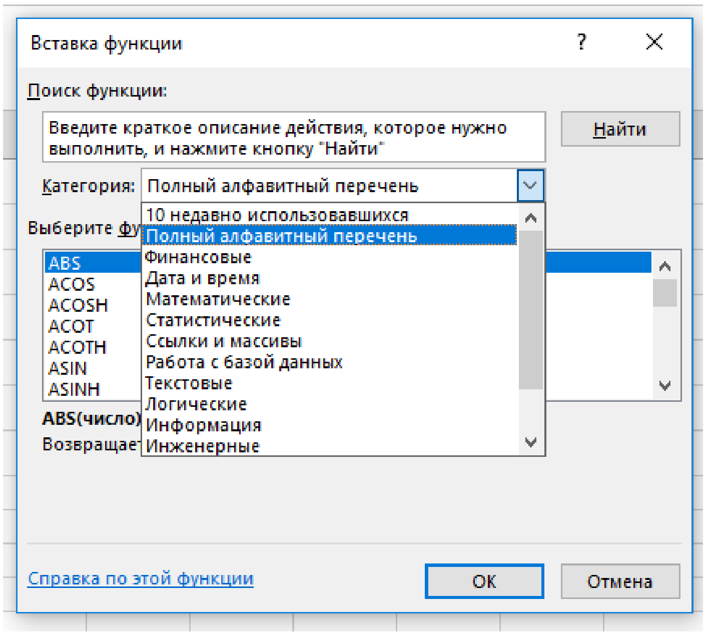


Рис. 2. Мастер вставки функции в Excel

Пусть надо получить 10 элементов арифметической прогрессии с начальным значением 10 и шагом 5.

Эта задача решается с использованием верхнего меню. В ячейку С1 заносим начальное значение, выделяем 10 ячеек для элементов прогрессии. Затем открываем в верхнем меню вкладку «редактирование/прогрессия». В открывшемся окошке выбираем тип прогрессии, указываем величину шага. В результате в столбце С1–С10 находятся элементы прогрессии (10, 15, 20, …, 55).

Рассчитать сумму элементов этой прогрессии можно как сумму элементов столбца, выбрав редактирование/автосумма (∑). В результате в ячейке С11 появится результат 325.

Логические операции используются в том случае, когда значение ячейки необходимо вычислять в зависимости от выполнения или невыполнения одного или нескольких условий. В ячейку записывается условие в виде функции логического ветвления (ЕСЛИ) с тремя аргументами:

ЕСЛИ(логич.выраж.; значение 1, если истина; значение 2, если ложь)

Логическое выражение – любое значение или выражение, принимающие значения ИСТИНА или ЛОЖЬ. Например, В2 > 47 – логическое выражение, принимающее значение 1, если в ячейке В2 находится число, больше, чем 47, в противном случае функция принимает значение 2.

Пусть дан объем добычи нефти (баррелей в сутки) по каждой скважине в течение недели. Требуется получить информацию о скважинах, у которых средняя добыча за неделю выше среднего значения. Использовать логические операции.

Для решения задачи сначала найдем средние значения добычи каждой скважины. Для этого используем встроенную функцию СРЗНАЧ в Excel, выбрав ее из списка, выделив в качестве аргументов ячейки B11:G11. Затем, используя эту же функцию, найдем среднее значение добычи нефти в ячейке C15. Появится результат 239,09524.

Далее используем функцию ЕСЛИ, записав в ячейку B13 для первого продавца условие:

=ЕСЛИ(B11>=$B$15;"лучшая";"***").

Затем скопируем эту функцию для других скважин в ячейки C11:G11. В итоге в строке 13 мы увидим информацию о скважинах, у которых средний объем добычи нефти за неделю выше среднего: в столбце такой скважины будет текст «лучшая».

В данном примере перед номером строки и столбца ячейки B15 стоит знак $. Он используется в случаях, когда адрес ячейки при копировании и перемещении не должен меняться. Такая адресация называется абсолютной. При обычном копировании формул из одной ячейки в другую ссылки на ячейки в Excel автоматически меняются. Такая адресация называется относительной.

2.2. Построение графиков функций

Для отражения полученных или собранных числовых данных в Excel имеются средства для создания графиков и диаграмм, которые в наглядной форме представляют зависимости и тенденции. Microsoft Excel позволяет создавать множество различных типов графиков: линейчатая диаграмма, гистограмма, линейный график, точечная, круговая и пузырьковая диаграмма, диаграмма-поверхность и другие. За построение графиков и диаграмм в Excel отвечает *Мастер диаграмм*, который вызывается через функцию *Вставки* → *Диаграмма* главного меню. Эта функция считается одним из основных достоинств Excel. Поэтому важно научиться строить графики для получения максимальной отдачи от Excel. Рассмотрим несколько примеров построения графиков.

Построение графиков функции одной переменной

Пример 1. Для примера построим графики функций $y_1 = \ln(x)–2$ и $y_2 = –2x+1$ в диапазоне [0,2; 3] с шагом 0,2.

Заголовок задания запишем в ячейке F1. Затем укажем заголовки столбцов в ячейках: A3 – «№» – порядковый номер точки; B3 – «x» – значение аргумента; C3 – «$y_1 = \ln(x)–2$», D3 – «$y_2 = –2x+1$» – вычисленное значение функции соответствующего аргумента.

Заполним саму таблицу следующим образом:

1. Заполним диапазон ячеек A4–A18 последовательностью 1, 2, …, 20. При этом возможно использовать автозаполнение, для чего в ячейках A4–A5 необходимо ввести числа 1 и 2 соответственно, а затем, выделив их, потянуть за нижний правый угол до достижения необходимого значения.

2. Заполним диапазон ячеек B4–B18 последовательностью от 0,2 до 3 с шагом 0,2.

3. В ячейках C4–C18 рассчитаем значения функции $y_1 = ln(x)–2$. Для этого находим и выбираем функцию LOG10 (число). В качестве аргумента функции указываем адрес соответствующей ячейки в столбце B, для чего щелкаем мышкой по нужной ячейке. В данном случае также есть возможность использовать автозаполнение (рис. 3).

4. В ячейках D4–D18 рассчитаем значения функции $y_2 = –2x+1$. В качестве аргумента функции указываем адрес соответствующей ячейки в столбце B, для чего щелкаем мышкой по нужной ячейке. В данном случае также есть возможность использовать автозаполнение (рис. 3).

C7 =LOG10(B7)-2

	A	B	C	D
3	№	*x*	y = lg(*x*) - 2	y = - 2*x*+1
4	1	0,2	-2,6990	0,6
5	2	0,4	-2,3979	0,2
6	3	0,6	-2,2218	-0,2
7	4	0,8	-2,0969	-0,6
8	5	1,0	-2,0000	-1,0
9	6	1,2	-1,9208	-1,4
10	7	1,4	-1,8539	-1,8
11	8	1,6	-1,7959	-2,2
12	9	1,8	-1,7447	-2,6
13	10	2,0	-1,6990	-3,0
14	11	2,2	-1,6576	-3,4
15	12	2,4	-1,6198	-3,8
16	13	2,6	-1,5850	-4,2
17	14	2,8	-1,5528	-4,6
18	15	3,0	-1,5229	-5,0

Рис. 3. Создание таблицы функций

5. Строим графики функции, выполняя последовательно следующие шаги:

5.1. Выделим диапазон ячеек C3:D18 (вместе с заголовками).

5.2. В главном меню вызываем *Вставка → График* и выбираем один вариант из предложенных типов графиков. На экране появляется график функции в первоначальном варианте. Теперь необходимо привести его в соответствие требованиям стандарта НИ ТПУ.

В третьей строке меню выбираем подменю *Выбрать данные*, после чего на экране появится окно *Выбор источника данных*. В нашем случае необходимо изменить подписи горизонтальной оси. Для этого надо нажать *Изменить* – в появившемся окне курсор будет в строке *Диапазон подписи оси*, мышкой выделяем диапазон аргументов – ячейки B4:D18 и нажимаем OK. Видим, что на графике по оси *x* появились необходимые значения – снова нажимаем OK.

Видим, что шрифт на графике не Time New Roman, поэтому двойным щелчком мышки поочередно выделяем значения оси *y*, значения оси *x*, название графика и, выходя в главное меню, изменяем шрифт и при необходимости его размер (рис. 4).

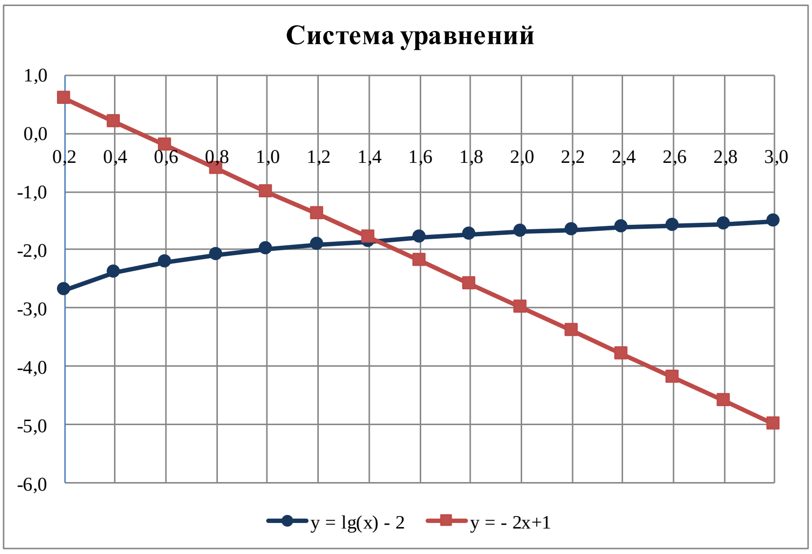


Рис. 4. Построение графиков функций

2.3. Построение плоскости 1-го порядка

Уравнение вида A*x* + B*y* + C*z* + D = 0 называется общим уравнением плоскости. Важные частные случаи уравнения плоскости возникают при равенстве нулю некоторых коэффициентов A, B, C, D. Если D = 0, то уравнение A*x* + B*y* + C*z* = 0 определяет плоскость, проходящую через начало координат. Если A = 0, то уравнение B*y* + C*z* + D = 0 определяет плоскость, параллельную оси *x*, аналогично при B = 0 и C = 0.

Пример 2. Построить плоскость, заданную уравнением

$$2x + 4y - 2z + 2 = 0.$$

Построим часть плоскости, лежащей в 1-м квадранте для *x* от 0 до 6 с шагом 0,5 и для *y* от 0 до 6 с шагом 1,0. Для этого решим уравнение относительно z: $z = x + 2y + 1$. Значения x вводим в столбец А (можно использовать автозаполнение), начиная со строки 2 и до строки 14. Значения y вводим в строку 1, начиная со столбца В и заканчивая столбцом Н. В ячейку В2 записываем формулу для расчета z: =$A2+2*B$1+1 и, нажав указателем мыши на черную точку в нижнем правом углу ячейки В2, заполняем этой формулой весь диапазон ячеек В2:Н14.

Теперь обращаемся к *Мастеру диаграмм*. Выделив весь диапазон ячеек В2:Н14, в главном меню выбираем *Вставка → Другие диаграммы → Проволочная поверхность* и нажимаем ОК. На экране появляется плоскость в первоначальном варианте. В третьей строке меню выбираем подменю *Выбрать данные*, после чего на экране появится окно *Выбор источника данных*. Для изменения подписей по оси x надо нажать *Изменить* в поле *Подписи горизонтальной оси (категории)* и мышкой выделить диапазон аргументов – ячейки А2:А14. Для изменения подписей по оси y в поле *Элементы легенды (ряды)* необходимо последовательно поменять каждый из семи рядов: в рабочее поле *Имя ряда* для ряда 1 вводим ячейку В1, после нажатия ОК первым элементом показано значение 0; для ряда 2 вводим ячейку С1, после нажатия ОК вторым элементом показано значение = 1 и т. д. Таким образом, после ввода последнего значения ряда $y = 6$ и закрытия окна нажатием ОК мы получаем практически готовый график. Остается только изменить шрифт на Time New Roman и в результате получаем плоскость, представленную на рис. 5.

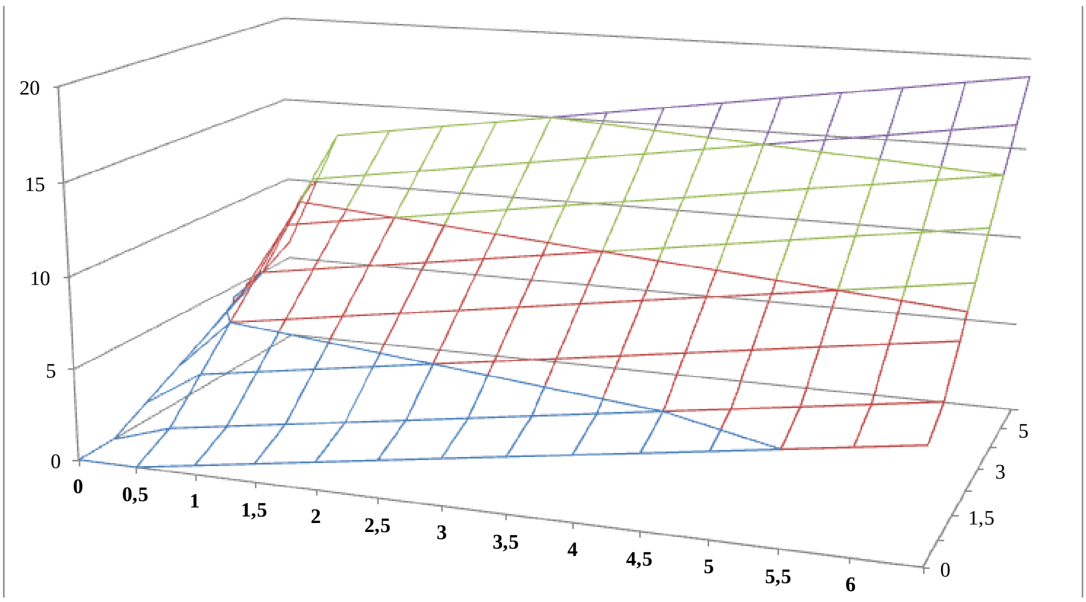


Рис. 5. Вид построенной в программе Excel плоскости z = x+2·y+1

2.4. Построение плоскости 2-го порядка

Пример 3. Построить двухполосный гиперболоид, заданный уравнением $x^2 / 9 + y^2 / 36 = z$.

Переменная x имеет диапазон изменений (–3; +3), y – (–2; +2), шаг 0,5 для обеих переменных. Уравнение относительно z:

$$z = \frac{x^2}{9} + \frac{y^2}{36}.$$

Как и в примере 2, значения x вводим в столбец A, начиная со строки 2 и до строки 14. Значения y – в строку 1, начиная со столбца B, в ячейку B2 вводим уравнение для z:

z = ($A2^2)/9+(B$1^2)/36.

Далее заполняем этой формулой весь диапазон ячеек B2:J14 (рис. 6).

Теперь обращаемся к *Мастеру диаграмм*. Выделив весь диапазон ячеек B2:J14, в главном меню выбираем *Вставка → Другие диаграммы → Проволочная поверхность* и нажимаем OK. По аналогии с предыдущим примером изменяем подписи по осям. По оси x: нажать *Изменить* в поле *Подписи горизонтальной оси (категории)* и мышкой выделить диапазон аргументов – ячейки A2:A14.

Для изменения подписей по оси y в поле *Элементы легенды (ряды)* необходимо последовательно поменять каждый из семи рядов: в рабочее поле *Имя ряда* для ряда 1 вводим ячейку B1, после нажатия OK первым элементом показано значение 0; для ряда 2 вводим ячейку C1, после нажатия OK вторым элементом показано значение = 0,5 и т. д. Теперь для более наглядного представления графика поменяем минимальное значение вертикальной оси. Для этого активируем меню вертикальной оси двойным щелчком мышки и в открывшимся окне вместо минимального значения *авто* набираем 1,0.

Для изменения подписей по оси x надо нажать *Изменить* в поле *Подписи горизонтальной оси (категории)* и мышкой выделить диапазон аргументов – ячейки A2:A14. Для изменения подписей по оси y в поле *Элементы легенды (ряды)* необходимо последовательно поменять каждый из рядов: в рабочее поле *Имя ряда* для ряда 1 вводим ячейку B1, после нажатия OK первым элементом показано значение –3; для ряда 2 вводим ячейку C1, после нажатия OK вторым элементом показано значение = 0,5 и т. д. Таким образом, после ввода последнего значения ряда y = 3 и закрытия окна нажатием OK мы получаем практически готовый график. Остается только изменить шрифт на Time New Roman и в результате получаем плоскость, представленную на рис. 7.

C4 =$A4^2/9+C$1^2/36

	A	B	C	D	E	F	G	H	I	J
1	**x /y**	**-2**	**-1,5**	**-1**	**-0,5**	**0**	**0,5**	**1**	**1,5**	**2**
2	**-3**	1,1111	1,0625	1,0278	1,0069	1	1,0069	1,0278	1,0625	1,1111
3	**-2,5**	0,8056	0,7569	0,7222	0,7014	0,6944	0,7014	0,7222	0,7569	0,8056
4	**-2**	0,5556	0,5069	0,4722	0,4514	0,4444	0,4514	0,4722	0,5069	0,5556
5	**-1,5**	0,3611	0,3125	0,2778	0,2569	0,25	0,2569	0,2778	0,3125	0,3611
6	**-1**	0,2222	0,1736	0,1389	0,1181	0,1111	0,1181	0,1389	0,1736	0,2222
7	**-0,5**	0,1389	0,0903	0,0556	0,0347	0,0278	0,0347	0,0556	0,0903	0,1389
8	**0**	0,1111	0,0625	0,0278	0,0069	0	0,0069	0,0278	0,0625	0,1111
9	**0,5**	0,1389	0,0903	0,0556	0,0347	0,0278	0,0347	0,0556	0,0903	0,1389
10	**1**	0,2222	0,1736	0,1389	0,1181	0,1111	0,1181	0,1389	0,1736	0,2222
11	**1,5**	0,3611	0,3125	0,2778	0,2569	0,25	0,2569	0,2778	0,3125	0,3611
12	**2**	0,5556	0,5069	0,4722	0,4514	0,4444	0,4514	0,4722	0,5069	0,5556
13	**2,5**	0,8056	0,7569	0,7222	0,7014	0,6944	0,7014	0,7222	0,7569	0,8056
14	**3**	1,1111	1,0625	1,0278	1,0069	1	1,0069	1,0278	1,0625	1,1111

Рис. 6. Исходные данные решения уравнения $z = \dfrac{x^2}{9} + \dfrac{y^2}{36}$

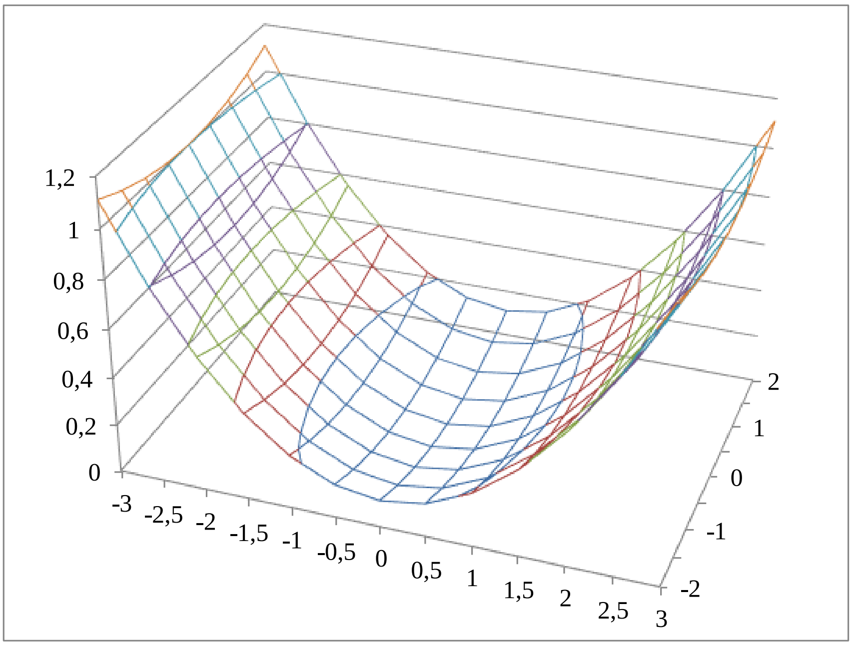


Рис. 7. Вид построенной в программе Excel поверхности $z = \dfrac{x^2}{9} + \dfrac{y^2}{36}$

Вопросы и задания для самопроверки

1. Опишите структуру рабочего листа. Какие основные окна Excel вы можете назвать?

2. Какие средства Excel помогают визуализировать информацию? Какие функции используются в Excel для создания графиков и диаграмм (линейчатая диаграмма, гистограмма и другие)?

3. Что такое автозаполнение? Приведите примеры встроенных функций Excel.

Раздел 3. Решение нелинейных алгебраических уравнений

Одной из математических задач, часто возникающих в различных разделах физики, химии, нанотехнологий и других, является решение нелинейных уравнений с одним неизвестным. В общем случае нелинейное уравнение с одним неизвестным можно записать в виде $F(x^*) = 0$. Понятие нелинейности уравнения означает, что его графиком не является прямая линия. Решить подобное уравнение – значит найти значение x^*, которое называют корнем уравнения, такое, что $x^* \subset R$, при этом $F(x^*) = 0$. Необходимо понимать, что нелинейное уравнение может иметь несколько корней, то есть точек x_1^*, x_2^*, x_3^*, и т. д., в которых функция $F(x)$ пересекает ось x.

Методы решения нелинейного уравнения делят на точные (аналитические) и приближенные (итерационные). Если вы используете точный метод, то корень находится за конечное число действий и представляется некоторой алгебраической формулой. В приближенных методах процесс нахождения решения может быть бесконечен и зависит от требуемой точности. Приближенным решением называется бесконечная последовательность $\{x_n\}$, такая, что $\lim_{n\to\infty} x_n = x^*$. По определению предела, для любого (сколь угодно малого) ε найдется такое N, что при любом $n > N$, $|x_n - x^*| < \varepsilon$. Члены этой последовательности x_n называются последовательными приближениями к решению, или итерациями. Наперед заданное число ε – точность метода, а N – это количество итераций, которое необходимо выполнить, чтобы получить решение с заданной точностью ε.

3.1. Методы нахождения приближенного решения

Существуют различные методы нахождения, т. е. способы построения последовательности итераций, приближенного решения $\{x_n\}$. Для того чтобы принять решение, какой приближенный метод использовать, уравнение $F(x) = 0$ надо исследовать на наличие корней и уточнить, где эти корни находятся, т. е. найти интервалы изоляции корней. Это значит надо найти отрезок, на котором корень уравнения не только существует, но и является единственным. То есть, иначе говоря, каждому корню должен соответствовать свой интервал изоляции. Поэтому, если корней несколько, то и интервалов столько же, и для каждого нужно найти интервал изоляции. Известны несколько способов исследования функции: аналитический, табличный, графический. Аналитический способ состоит в нахождении экстремумов функции $F(x)$, исследовании ее поведения при $x \to \pm\infty$, нахождении участков возрастания и убывания функции. Табличный способ – это построение таблицы, состоящей из столбца аргумента x и столбца значений функции $F(x)$. О наличии корней свидетельствуют перемены знака функции. Чтобы не произошла потеря корней, шаг изменения аргумента должен быть достаточно мелким, а интервал изменения – достаточно широким. И наконец, графический способ – это построение графика функции $F(x)$ и определение числа корней по количеству пересечений графика с осью x.

После того, как интервалы изоляции корней найдены, можно переходить непосредственно к нахождению корней. Ознакомимся с несколькими итерационными методами, которые позволяют найти корень на заданном интервале изоляции $[a, b]$, используя различные формулы.

1. Метод деления отрезка пополам. Помним о том, что мы нашли интервал $[a, b]$, который содержит единственный корень уравнения. Найдем середину данного отрезка: $c = (a + b)/2$. Понятно, что искомый корень остался в одной половине: $[a, c]$ или $[c, b]$. Тогда проверяем следующее условие: если $F(a) \cdot F(c) < 0$, то корень попал в интервал $[a, c]$ и деление отрезка нужно снова повторить, но при этом приняв в качестве нового конца точку c, другими словами, $b = c$. В противоположном случае, когда корень на отрезке $[c, b]$, изменяем значение левого конца отрезка: $a = c$ (рис. 8). Поскольку корень всегда заключен внутри отрезка, итерационный процесс можно останавливать, если длина отрезка станет меньше заданной точности: $|b - a| < \varepsilon$.

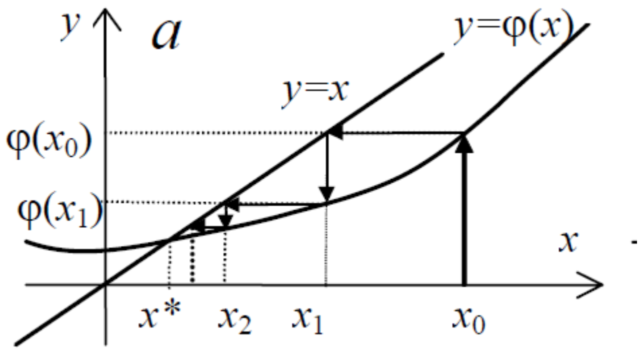


Рис. 8. Метод деления отрезка пополам

2. Метод хорд. В этом методе кривая $F(x)$ заменяется прямой линией – хордой, стягивающей точки $(a, F(a))$ и $(b, F(b))$. В зависимости от знака выражения $F(a) \cdot F''(a)$ метод хорд имеет два варианта. Пусть $F(a) \cdot F''(a) > 0$. Тогда $x_0 = b$, точка a будет оставаться неподвижной. Уравнение хорды, проходящей через точки $(a, F(a))$ и $(x_0, F(x_0))$,

$$y = F(a) + \frac{F(x_0) - F(a)}{x_0 - a}(x - a).$$

Следующее приближение x_1 находим как точку пересечения хорды с осью x:

$$x_1 = a - \frac{F(a)(x_0 - a)}{F(x_0) - F(a)}.$$

Пусть теперь $F(a) \cdot F''(a) < 0$. Тогда $x_0 = a$, точка b неподвижна. Проведем хорду, соединяющую точки $(b, F(b))$ и $(x_0, F(x_0))$:

$$y = F(x_0) + \frac{F(b) - F(x_0)}{b - x_0}(x - x_0).$$

Вычислим точку пересечения хорды с осью x:

$$x_1 = x_0 - \frac{F(x_0)(b - x_0)}{F(b) - F(x_0)}.$$

На следующем шаге мы возьмем вычисленное значение x_1 в качестве x_0. И так далее (рис. 9). Таким образом, формула для метода хорд выглядит следующим образом:

если $F(a) \cdot F''(a) > 0$, то $x_0 = b$, $x_{k+1} = a - \frac{F(a)(x_k - a)}{F(x_k) - F(a)}$;

если $F(a) \cdot F''(a) < 0$, то $x_0 = a$, $x_{k+1} = x_k - \frac{F(x_0)(b - x_k)}{F(b) - F(x_k)}$.

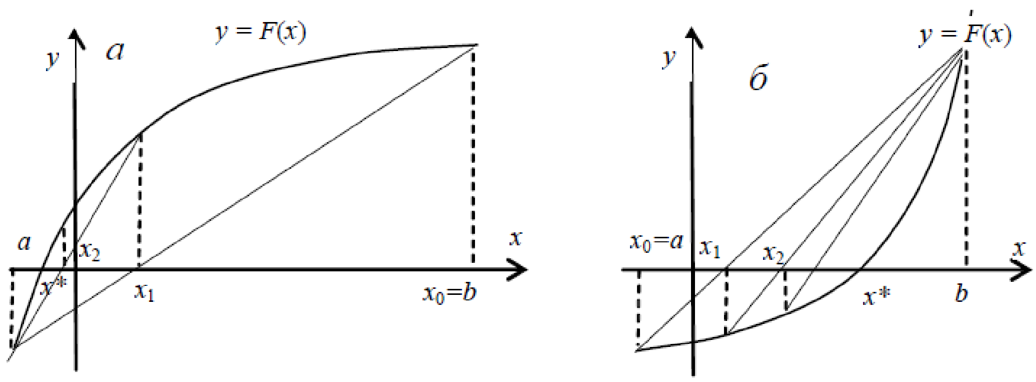


Рис. 9. Метод хорд

Окончание итерационного цикла в этом методе происходит по условию малости невязки уравнения: $|F(x_k)| < \varepsilon$.

3. Метод Ньютона. Он также часто называется метод касательных. Этот метод основан на замене исходного нелинейного уравнения $F(x) = 0$ линейным уравнением, которое можно легко решить.

Пусть x_0 – начальное приближение (любая точка интервала изоляции). Построим касательную к функции $y = F(x)$, проходящую через точку $(x_0, F(x_0))$: $y = F(x_0) + F'(x_0) \cdot F(x_0 - x)$.

Найдем пересечение касательной с осью x:

$$x_1 = x_0 - \frac{F(x_0)}{F'(x_0)}.$$

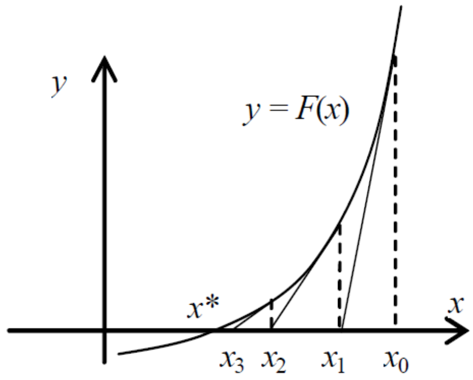


Рис. 10. Метод Ньютона

На следующей итерации в качестве x_0 надо взять полученное в ходе предыдущего вычисления значение x_1 (рис. 10). И так продолжать итерационный процесс до тех пор, пока, как и в методе хорд, не выполнится условие, что невязка уравнения становится меньше заданной точности ε: $|F(x_1)| < \varepsilon$.

3.2. Решение одного уравнения с одним неизвестным в Excel

Рассмотрим использование метода деления отрезка пополам на простом примере.

Пример 4. Найти корни уравнения $x^2 \cdot \cos(2x) + 1 = 0$ на отрезке $[0; 2\pi]$.

Для отделения корней построим таблицу значений функции $y = x^2 \cdot \cos(2x) + 1$ на интервале $[0; 2\pi]$ с шагом $\pi/12$. Для этого заполним диапазон ячеек A4–A28 последовательностью 1, 2, …, 20, диапазон ячеек B4–B28 последовательностью от 0 до $+2\pi$ с шагом $\pi/6$. В ячейках C4–C23 рассчитаем значения функции $y = x^2 \cdot \cos(2x) + 1$, в качестве аргумента функции указываем адрес соответствующей ячейки в столбце B, для чего щелкаем мышкой по нужной ячейке.

Выделим отрезки, на которых функция меняет знак с «+» на «–» или наоборот. Для нахождения таких интервалов используем функцию ЕСЛИ. Сначала запишем в ячейку D4 для $x = 0$ условие, как показано на рис. 11.

Затем скопируем функцию ЕСЛИ в ячейки D5:D28 для других значений x. Тогда в итоге в столбце D появится информация об интервалах, в которых находится корень уравнения. Для интервала, в котором находится корень, в строке будет текст «корень». В вашем файле выделите цветом ячейки, в которых появился текст «корень», и убедитесь, что у вас это интервалы [1,047; 1,309], [2,094; 2,356], [3,927;4,189] и [5,236; 5,498].

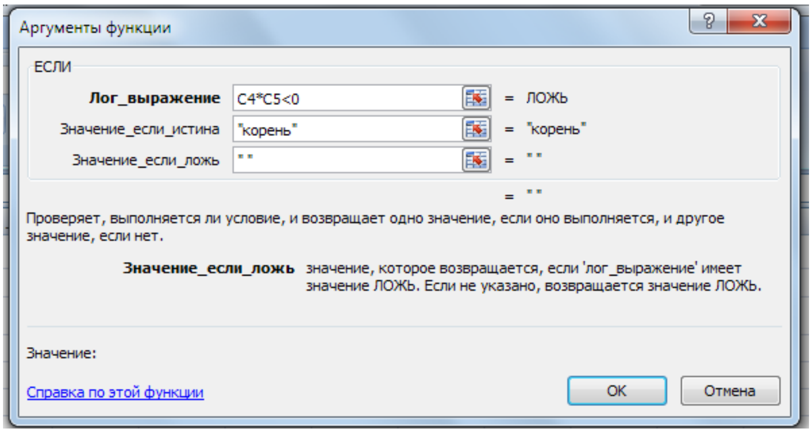


Рис. 11. Окно мастера функции ЕСЛИ

Построим график функции $y = x^2 \cdot \cos(2x) + 1$. Для этого выделим диапазон ячеек C3:C28 (вместе с заголовком); в главном меню вызываем *Вставка → График* и выбираем один вариант из предложенных типов графиков. На экране появляется график функции в первоначальном варианте. Теперь необходимо привести его в соответствие требованиям стандарта НИ ТПУ.

В третьей строке меню выбираем подменю *Выбрать данные*, после чего на экране появится окно *Выбор источника данных*. В нашем случае необходимо изменить подписи горизонтальной оси. Для этого надо нажать *Изменить* – в появившемся окне курсор будет в строке *Диапазон подписи оси*, мышкой выделяем диапазон аргументов – ячейки B4–B28 – и нажимаем ОК. Видим, что на графике по оси x появились необходимые значения – снова нажимаем ОК.

Уточнение корней продемонстрируем на двух методах: методе подбора параметров (А) и методе половинного деления (Б).

А) При использовании встроенной возможности Excel *Подбор параметров* корень находится путем последовательных приближений, но это делают внутренние возможности системы. Для пользователей необходимо только вручную ввести значение, очень близкое к искомому корню. Например, для первого отрезка локализации корня [1,047; 1,309] возьмем за начальное приближение точку этого отрезка 1,25 и запишем его в ячейку E1. В ячейку F1 введем формулу функции, для которой мы ищем корни. В главном меню находим *Данные → Анализ «что-если» → Подбор параметра*, на экране появляется диалоговое окно. В поле *Установить в ячейке* вводим ссылку на ячейку, содержащую формулу; в поле *Значение* записываем значение правой части уравнения; в поле *Изменяя значение* в ячейке записываем адрес ячейки, содержащей переменную (рис. 12). Нажимаем ОК.

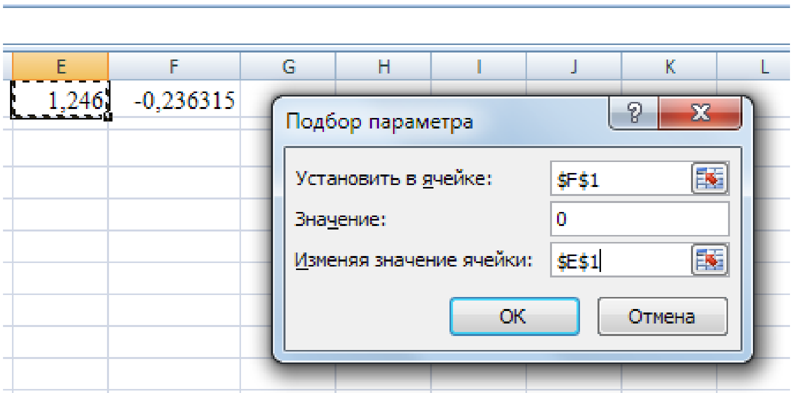


Рис. 12. Окно подбора параметра

На экране появляется *Результат подбора параметра*, в котором показано найденное решение. Значение корня и функции появляются в ячейках E1 и F1 соответственно (рис. 13).

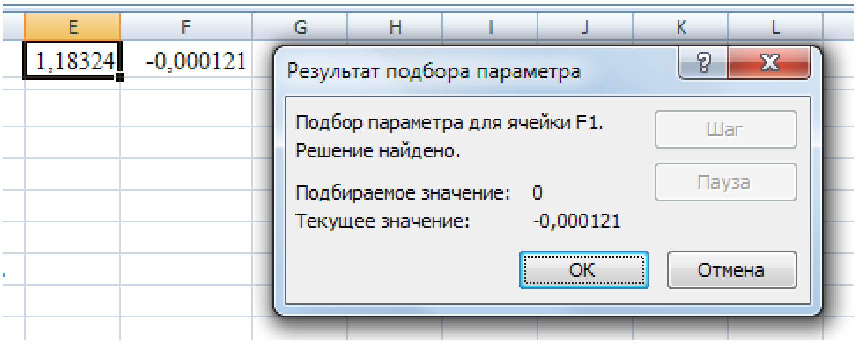


Рис. 13. Окно результата подбора параметра

Аналогично необходимо найти другие три корня. Для этого также вручную вводятся значения, близкие к искомому корню, из найденных на первом этапе интервалов, например 2,2 для [2,094; 2,356], 4 для [3,927; 4,189] и 5,3 для [5,236; 5,498]. Каждый раз копируем формулу функции, для которой мы ищем корни. И в итоге для нашего примера решениями уравнения должны быть значения: 1,183; 2,257; 3,959; 5,481.

Б) Алгоритм метода половинного деления сводится к повторяющемуся на каждом шаге делению отрезка локализации корня пополам, причем каждый последующий отрезок является половиной предыдущего. Процесс деления отрезка продолжается до тех пор, пока его длина не станет меньше 2ε, где ε – точность нахождения корня. В этом случае любая точка отрезка локализации отличается от корня не более чем на ε.

В качестве первого приближения к корню возьмем любую точку приблизительно в середине отрезка $[a, b]$, т. е. $x = (a + b)/2$.

Если $F(a) \cdot F(x) \le 0$, то корень находится в интервале $[a, x]$, который берем за новый отрезок локализации корня. Если $F(a) \cdot F(b) > 0$, то за новый отрезок локализации берем интервал $[x, b]$.

Новый отрезок локализации в два раза меньше первоначального. Процесс деления отрезка продолжаем до тех пор, пока его длина не станет меньше 2ε.

Рассмотрим этот алгоритм для второго отрезка локализации, найденного на первом этапе, – [2,094; 2,356], с точностью 0,001. Создадим новую таблицу следующим образом:

1. Занесем в ячейку A2 начальную точку интервала, в ячейку B2 – конечную точку интервала, значение функции $y = x^2 \cdot \cos(2x) + 1$ для нее – в ячейке E2. Рассчитаем среднюю точку в ячейке C2, используя встроенную функцию СРЕДЗН(), значение функции для нее – в ячейке D2. В ячейке F2 проведем вычисление $F(a) \cdot F(x)$ для определения знака результата умножения (сама расчетная величина для нас не имеет никакого значения!!) для установления нового отрезка локализации.

2. В соответствующей строке столбца G мы будем сравнивать длину отрезка с двойной точностью нахождения корня и решать, следует ли продолжать процесс деления отрезка или корень найден.

3. В ячейке A3 мы запишем формулу ЕСЛИ(D2<=0;A2;C2), а в ячейке B3 – ЕСЛИ(D2<=0;C2;B2). Таким образом, мы проверим условие для определения нового отрезка локализации. В ячейки C3–G3 мы скопируем ячейки C2–G2.

4. Далее мы копируем ячейки A3–G3 в ячейки A4–G4, затем A4–G4 в ячейки A5–G5 и т. д. до тех пор, пока в столбце F не появится надпись: «Корень равен…» (рис. 14).

G4 =ЕСЛИ((B4-A4)<2*$I$1;"Корень равен "&C4;"Корень не найден")

	A	B	C	D	E	F	G	H	I
1	a	b	x	$F(x)$	$F(a)$	$F(a) \cdot F(x)$		Точность нахождения корня	0,001
2	2,094	2,356	2,2250	-0,2841352	-1,1954180	0,3397	Корень не найден		
3	2,225	2,356	2,2905	0,3126637	-0,2841352	-0,0888	Корень не найден		
4	2,2250	2,2905	2,2578	0,0028430	-0,2841352	-0,0008	Корень не найден		
5	2,2250	2,2578	2,2414	-0,1435388	-0,2841352	0,0408	Корень не найден		
6	2,2414	2,2578	2,2496	-0,0710668	-0,1435388	0,0102	Корень не найден		
7	2,2496	2,2578	2,2537	-0,0342911	-0,0710668	0,0024	Корень не найден		
8	2,2537	2,2578	2,2557	-0,0157688	-0,0342911	0,0005	Корень не найден		
9	2,2557	2,2578	2,2567	-0,0064741	-0,0157688	0,0001	Корень не найден		
10	2,2567	2,2578	2,2572	-0,0018183	-0,0064741	0,0000	Корень равен 2,25723828125		

Рис. 14. Результаты уточнения корня методом половинного деления

Вопросы и задания для самопроверки

1. Какие уравнения называются нелинейными алгебраическими уравнениями? Какие методы существуют для решения таких уравнений?

2. Перечислите этапы нахождения решения. Какое необходимое условие существования корня на отрезке?

3. Дайте определение понятию «итерация». Какие вы знаете критерии остановки итерационного процесса?

4. В чем заключается суть метода деления отрезка пополам?

5. В чем заключается суть метода хорд?

6. В чем заключается суть метода Ньютона?

Раздел 4. Решение систем линейных алгебраических уравнений

Многие прикладные задачи в технике, экономике и других областях сводятся к решению матриц, которые, по сути, представляют собой систему линейных алгебраических уравнений (СЛАУ), в которой элементы матрицы являются соответствующими коэффициентами. Около 75 % всех вычислительных задач приводит к решению таких систем. Существует много методов и современных пакетов прикладных программ для решения СЛАУ, но для того чтобы их успешно использовать, необходимо разбираться в основах построения методов и алгоритмов, иметь представление о недостатках и преимуществах каждого из методов.

Мы будем решать лишь системы, у которых число уравнений совпадает с числом неизвестных, и при этом решение будет единственным.

Пусть дана линейная система n уравнений с n неизвестными, где a_{ij}, b_i ($i, j = 1,\ldots, n$) – соответственно коэффициенты при неизвестных и свободные члены уравнений:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2; \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases}$$

4.1. Решение СЛАУ методом Крамера

Метод Крамера относится к прямым методам решения линейных систем. Он основан на замене i-го столбца на столбец свободных членов и нахождении неизвестных с помощью определителей матриц $x_i = \det A_i / \det \mathrm{A}$.

Пусть дана линейная система n уравнений с n неизвестными, где a_{ij} – коэффициенты при неизвестных; b_i – свободные члены уравнений, $i, j = 1, \ldots, n$, $n \leq 5$.

Составляем основную матрицу $\mathrm{A} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$, находим ее определитель D. На базе матрицы А составляем вспомогательные матрицы $\mathrm{A}_1, \ldots, \mathrm{A}_n$ следующим образом: в матрице A_1 вместо первого столбца подставляем столбец свободных членов $\bar{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, в матрице A_2

второй столбец заменяем столбцом свободных членов и т. д. Вычисляем их определители $D_1, \ldots, D_n$. Каждое неизвестное x_i находится делением D_i на D.

Определитель матрицы – это многочлен, комбинирующий элементы квадратной матрицы таким образом, что его значение сохраняется при транспонировании и линейных комбинациях строк или столбцов.

Пример 5. Пусть дана линейная система из пяти уравнений с пятью неизвестными:

$$\begin{cases} 3x_1 - x_2 + x_3 + 2x_5 = 18; \\ 2x_1 - 5x_2 + x_4 + x_5 = -7; \\ x_1 - x_4 + 2x_5 = 8; \\ 2x_2 + x_3 + x_4 - x_5 = 10; \\ x_1 + x_2 - 3x_3 + x_4 = 1. \end{cases}$$

Запишем систему в виде матрицы A, а также вспомогательные матрицы $A_1, \ldots, A_5$, в каждой из которых заменим соответствующий i-й столбец столбцом свободных членов (рис. 15).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		3	-1	1	0	2		18	-1	1	0	2		**18**
2		2	-5	0	1	1		-7	-5	0	1	1		**-7**
3	**A=**	1	0	0	-1	2	**A_1=**	8	0	0	-1	2		**8**
4		0	2	1	1	-1		10	2	1	1	-1		**10**
5		1	1	-3	1	0		1	1	-3	1	0		**1**
6														
7		3	18	1	0	2		3	-1	18	0	2		
8		2	-7	0	1	1		2	-5	-7	1	1		
9	**A_2=**	1	8	0	-1	2	**A_3=**	1	0	8	-1	2		
10		0	10	1	1	-1		0	2	10	1	-1		
11		1	1	-3	1	0		1	1	1	1	0		
12														
13		3	-1	1	18	2		3	-1	1	0	18		
14		2	-5	0	-7	1		2	-5	0	1	-7		
15	**A_4=**	1	0	0	8	2	**A_5=**	1	0	0	-1	8		
16		0	2	1	10	-1		0	2	1	1	10		
17		1	1	-3	1	0		1	1	-3	1	1		

Рис. 15. Основная и вспомогательные матрицы примера 5

Вычислим определители матриц, используя функцию МОПРЕД. Для этого выделим ячейку, в которой мы хотим получить значение определителя матрицы, на панели инструментов нажмем кнопку f_x, найдем функцию МОПРЕД, в появившемся окне мастера функций в качестве аргумента функции мышкой выделим нужную нам матрицу (рис. 16). В ячейках A19:A24 расположим результаты вычислений определителей матриц A, A_1, …, A_5.

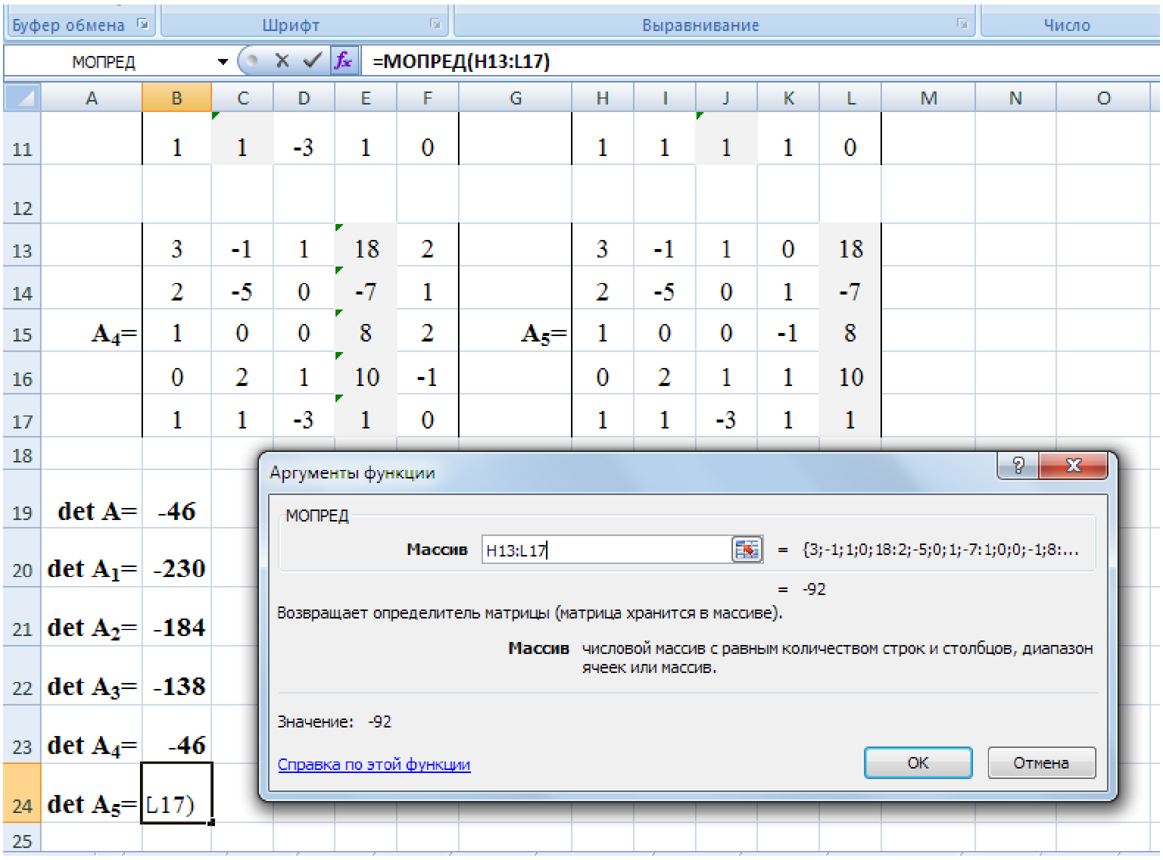


Рис. 16. Расчет определителей матриц A, A_1, ..., A_5

Далее вычисляем:

$$x_1 = \det A_1/\det A = 5;$$

$$x_2 = \det A_2/\det A = 4;$$

$$x_3 = \det A_3/\det A = 3;$$

$$x_4 = \det A_4/\det A = 1;$$

$$x_5 = \det A_5/\det A = 2.$$

Сделаем проверку (рис. 17).

fx =B1*$G$19+C1*$G$20+D1*$G$21+E1*$G$22+F1*$G$23

E	F	G	H	I	J	K	L	M	N	O
0	2		18	-1	1	0	2		**18**	18
1	1		-7	-5	0	1	1		-7	-7
-1	2	**A_1=**	8	0	0	-1	2		**8**	8
1	-1		10	2	1	1	-1		**10**	10
1	0		1	1	-3	1	0		**1**	1

Рис. 17. Проверка решения системы линейных уравнений примера 5

4.2. Решение СЛАУ методом Гаусса

Наиболее известным и популярным точным способом решения линейных систем является метод Гаусса. Этот метод относится к *прямым методам* решения линейных систем. Он основан на приведении матрицы системы к треугольному виду путем последовательного исключения неизвестных из уравнений системы (прямой ход метода Гаусса) и последующем решении этой треугольной системы, начиная с последнего уравнения (обратный ход метода Гаусса).

Пусть дана линейная система n уравнений с n неизвестными, где a_{ij}, b_i $(i, j = 1,\dots, n)$ – соответственно коэффициенты при неизвестных и свободные члены уравнений:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2; \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n. \end{cases}$$

На первом шаге из второго и всех последующих уравнений системы исключается x_1 с помощью первого уравнения. На втором шаге исключается x_2 из третьего и всех последующих уравнений с помощью второго уравнения и т. д. При этом, если в уравнении с номером k отсутствует неизвестная x_k ($a_{kk} = 0$), то производится перестановка этого уравнения с любым нижестоящим уравнением, содержащим эту переменную.

Такой процесс исключения называется прямым ходом Гаусса и необходимо продолжать его до тех пор, пока в левой части последнего уравнения не останутся нули и только один член с неизвестным x_n.

Обратный ход метода Гаусса состоит в последовательном вычислении неизвестных. Решая последнее уравнение, находят единственное неизвестное x_n. Далее, используя это значение, из предыдущего уравнения вычисляют x_{n-1} и т. д. Последним находят x_1 из первого уравнения.

Пример 6. Пусть дана линейная система из трех уравнений с тремя неизвестными:

$$\begin{cases} 2x_1 + 6x_2 - 2x_3 = 10; \\ 4x_1 + x_2 - x_3 = 20; \\ 3x_1 + 4x_2 = 10. \end{cases}$$

Запишем систему в виде матрицы (рис. 18).

	A	B	C	D	E	F
1						
2		2	6	-2	10	
3		4	1	-1	20	
4		3	4	0	10	

Рис. 18. Исходные данные

Сначала сделаем коэффициенты перед переменной x_1 во всех уравнениях равными единице, поделив каждое уравнение на коэффициент, стоящий перед этой переменной, т. е. поделив первое уравнение на a_{11}, второе на a_{21}, а третье на a_{31} (рис. 19).

2	Запишем систему в виде:					
3		2	6	-2	10	(1).
4		4	1	-1	20	(2) .
5		3	4	0	10	(3).
6	1 шаг:					
7	Разделим на					
8	2	1	3	-1	5,0	(1).
9	4	1	0,25	-0,25	5,0	(2) .
10	3	1	1,33	0	3,33	(3).

Рис. 19. Первый этап решения СЛАУ методом Гаусса

Теперь избавимся от переменной x_1 во втором и третьем уравнениях, вычтя из них первое (рис. 20).

11	2 шаг:					
12	Вычтем первое уравнение					
13		1	3	-1	5,0	(1).
14	(2) - (1)	0	-2,75	0,75	0,0	(2) .
15	(3) - (1)	0	-1,67	1	-1,67	(3).

Рис. 20. Второй этап решения СЛАУ методом Гаусса

Теперь нам нужно с помощью второго уравнения избавиться от переменной x_2 в третьем уравнении. Сделаем это тоже в два этапа. Сначала поделим второе уравнение на a_{22}, а третье уравнение на a_{23}. Далее преобразуем третье уравнение, вычтя из него второе (рис. 21).

17	3 шаг:					
18	Разделим на					
19	↓	1	3	-1	5,0	(1).
20	-2,75	0	1	-0,27	0,0	(2) .
21	-1,67	0	1	-0,60	1,00	(3).
22	4 шаг:					
23	Вычтем уравнение (2) из (3)					
24	↓	1	3	-1	5,0	(1).
25		0	1	-0,27	0,0	(2) .
26	(3) - (2)	0	0	-0,33	1,00	(3).
27						

Рис. 21. Результат решения системы линейных уравнений

Получим систему, приведенную к треугольному виду:

$$\begin{cases} x_1+3x_2-x_3=5; \\ x_2-0{,}27x_3=0; \\ -0{,}33x_3=1. \end{cases}$$

Найдем значения неизвестных, начиная с третьего (рис. 22).

x_3=	E26/D26=	-3,05556	
x_2=	E25 – D25*C28=		-0,83333
x_1=	E24 – D24*C28 – C24*D29 =		4,44444

Рис. 22. Результат решения системы линейных уравнений примера 9

Сделаем проверку, подставив полученные значения неизвестных в левые части уравнений системы, чтобы убедиться в выполнении условий:

$$\begin{cases} 2*4{,}444-6*0{,}833+2*3{,}056=10; \\ 4*4{,}444-4*0{,}833+3{,}056=20; \\ 3*4{,}444-4*0{,}833=10. \end{cases}$$

В файле Excel расчет производится следующим образом:

$$\begin{cases} =B2*E30+C2*D29+D2*C28; \\ =B3*E30+C3*D29+D3*C28; \\ =B4*E30+C4*D29. \end{cases}$$

4.3. Матричный метод решения СЛАУ

Пусть дана линейная система n уравнений с n неизвестными, где a_{ij}, b_i $(i, j = 1,\dots, n)$ – соответственно коэффициенты при неизвестных и свободные члены уравнений:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2; \\ \qquad\qquad\qquad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases}$$

Данную систему можно записать в виде матричного уравнения: $\bar{\bar{\mathrm{A}}} * \bar{\mathrm{X}} = \bar{\mathrm{B}}$, где $\bar{\bar{\mathrm{A}}}$ – матрица коэффициентов; $\bar{\mathrm{X}}$ – вектор неизвестных; $\bar{\mathrm{B}}$ – столбец свободных членов. Вектор неизвестных $\bar{\mathrm{X}}$ можно найти, решая систему линейных алгебраических уравнений с использованием обратной матрицы, для чего умножаем левую и правую часть матричного уравнения на обратную матрицу $\bar{\bar{\mathrm{A}}}^{-1}$ слева. Получим следующее соотношение: $\bar{\mathrm{X}} = \bar{\bar{\mathrm{A}}}^{-1} \cdot \bar{\mathrm{B}}$. Матрица (A^{-1}) называется обратной по отношению к квадратной матрице А, если при умножении этой матрицы на исходную и справа и слева получается единичная матрица: $\mathrm{A} \cdot \mathrm{A}^{-1} = \mathrm{A}^{-1} \cdot \mathrm{A} = \mathrm{E}$. Необходимым и достаточным условием существования обратной матрицы является невырожденность исходной (т. е. ее определитель отличен от 0).

Такой метод решения СЛАУ называется матричным. Рассмотрим пример.

Пример 7. Пусть матрица коэффициентов $\bar{\bar{\mathrm{A}}}$ введена в диапазон ячеек A1:C3, а вектор-столбец правой части $\bar{\mathrm{B}}$ – в диапазон E1:E3. Найдем обратную матрицу коэффициентов $\bar{\bar{\mathrm{A}}}^{-1}$ с помощью функции МОБР. Результирующую матрицу расположим в диапазоне A5:C7 (рис. 23).

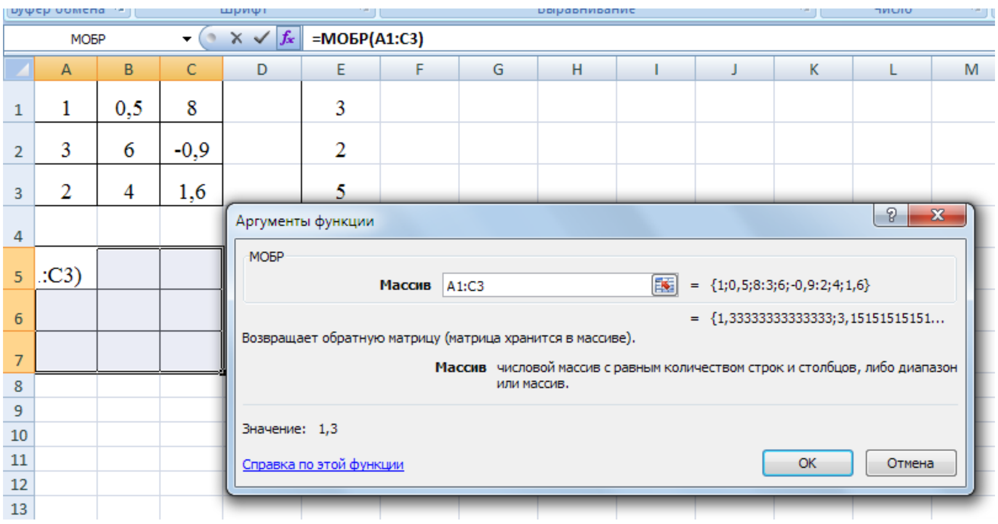
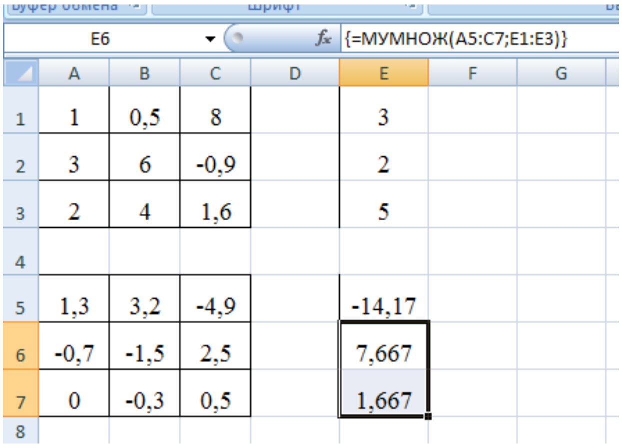


Рис. 23. Нахождение обратной матрицы

Для нахождения вектора неизвестных $\overline{X}$ выделим диапазон для его расположения – ячейки E5:E7, а затем по формуле $\overline{X} = \overline{\overline{A}}^{-1} \cdot \overline{B}$ с помощью функции МУМНОЖ получим результат (рис. 24).



E6 {=МУМНОЖ(A5:C7;E1:E3)}

	A	B	C	D	E	F	G
1	1	0,5	8		3		
2	3	6	-0,9		2		
3	2	4	1,6		5		
4							
5	1,3	3,2	-4,9		-14,17		
6	-0,7	-1,5	2,5		7,667		
7	0	-0,3	0,5		1,667		
8							

Рис. 24. Результат решения системы линейных уравнений

Пример 8. Имеются три вертикальные скважины, в которых определены абсолютные отметки кровли пласта (табл. 1).

Таблица 1

Номер скважины	Координаты скважины, м			Абсолют. отметка
n	x	y	z	пласта z, $м$
1	355	142	1	125,6
2	210	163	1	148,3
3	224	281	1	105,2

Необходимо рассчитать абсолютную отметку кровли пласта с координатами x = 240 м, y = 200 м.

Составим систему уравнений:

$$355a + 142b + c = 125{,}6;$$

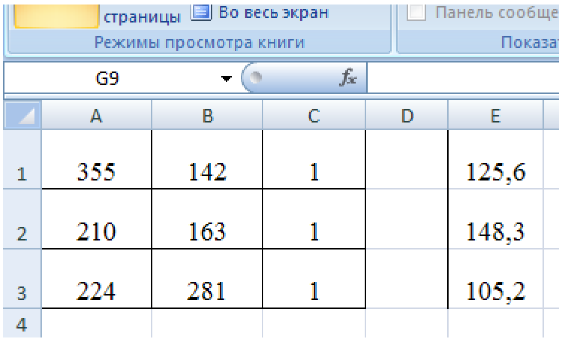
$$210a + 163b + c = 148{,}3;$$

$$224a + 201b + c = 105{,}2.$$

Для решения данной системы в программе Excel в диапазон ячеек A1:C3 введем матрицу исходных данных, а в диапазон E1:E3 – вектор-столбец правой части (рис. 25).

Первым шагом в диапазоне A5:C7 получим обратную матрицу с помощью функции МОБР. Для этого выделяем ячейки A5:C7, на панели инструментов *Стандартная* нажимаем кнопку f_x, выбираем категорию

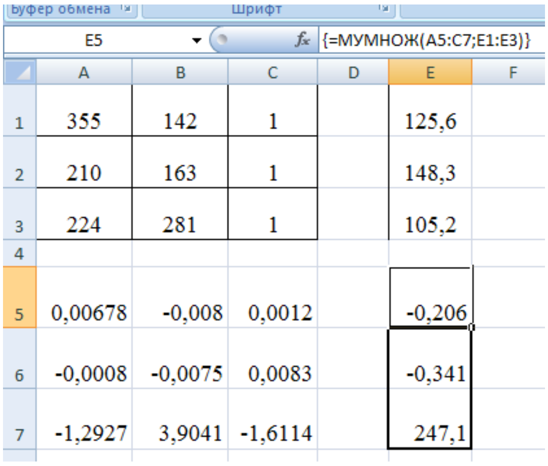
Математические, находим функцию МОБР, нажимаем ОК, появляется окно мастера функций. Мышкой выделяем ячейки A1:C3 в качестве массива, после чего переходим с помощью мышки в строку формул, нажимаем клавишу F2, затем – сочетание клавиш Ctrl + Shift + Enter. В результате в диапазоне A5:C7 появляется обратная матрица.



	A	B	C	D	E
1	355	142	1		125,6
2	210	163	1		148,3
3	224	281	1		105,2
4					

Рис. 25. Исходные данные для решения СЛАУ

Теперь выделим диапазон в ячейках E5:E7 и с помощью функции МУМНОЖ найдем коэффициенты *a*, *b* и *c*. Для этого на панели инструментов *Стандартная* нажимаем кнопку f_x, выбираем категорию *Математические*, находим функцию МУМНОЖ и нажимаем Enter. Мышкой выделяем ячейки A5:C7 в качестве Массива1, после чего переходим мышкой в строку формулы, нажимаем клавишу F2, в качестве Массива2 выделяем ячейки E1:E3, нажимаем клавишу F2, а затем – сочетание клавиш Ctrl + Shift + Enter. В результате в диапазоне E5:E7 появляется матрица, являющаяся решением системы уравнений (рис. 26).



	A	B	C	D	E	F
1	355	142	1		125,6	
2	210	163	1		148,3	
3	224	281	1		105,2	
4						
5	0,00678	-0,008	0,0012		-0,206	
6	-0,0008	-0,0075	0,0083		-0,341	
7	-1,2927	3,9041	-1,6114		247,1	

Рис. 26. Результат решения системы линейных уравнений

То есть значения коэффициентов следующие: *a* = – 0,206;
b = – 0,341;
c = 247,1.

Следовательно, уравнение имеет вид

$$z = -\,0{,}206x - 0{,}341y + 247{,}1.$$

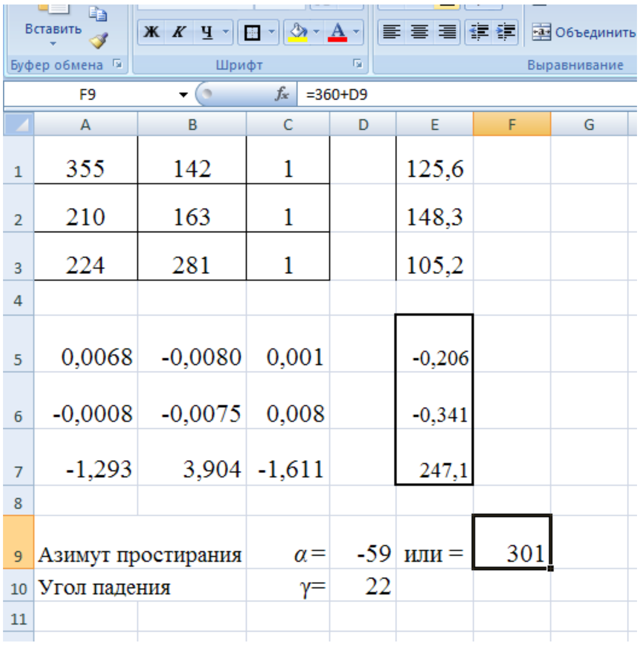
Подставляя в уравнение заданные координаты, найдем абсолютную отметку кровли пласта в точке x = 240 м, y = 200 м:

$$z = -\,0{,}206 \cdot 240 - 0{,}341 \cdot 200 + 247{,}1 = 129{,}5 \text{ м}.$$

Зная коэффициенты уравнения, можно извлечь дополнительную геологическую информацию об элементах залегания кровли пласта (рис. 27): азимут простирания $\alpha = \mathrm{arctg}(-b/a)$, угол падения $\gamma = \mathrm{arctg}\sqrt{a^2 + b^2}$.

Для данного примера рассчитаем в Excel:

- азимут простирания α =ATAN(–E6/E5)*180/ПИ() = –59°= 301°;
- угол падения γ =ATAN(КОРЕНЬ(E5^2+E6^2))*180/ПИ() = 22°.



	A	B	C	D	E	F	G
1	355	142	1		125,6		
2	210	163	1		148,3		
3	224	281	1		105,2		
4							
5	0,0068	-0,0080	0,001		-0,206		
6	-0,0008	-0,0075	0,008		-0,341		
7	-1,293	3,904	-1,611		247,1		
8							
9	Азимут простирания		α =	-59	или =	301	
10	Угол падения		γ=	22			
11							

Рис. 27. Результат выполнения примера 8

Вопросы и задания для самопроверки

1. Что такое система линейных алгебраических уравнений (СЛАУ)? Запишите СЛАУ в скалярном, векторном и матричном виде. При каком условии существует единственное решение?
2. Какой алгоритм решения методом обратной матрицы?
3. Какой алгоритм решения методом Крамера?
4. Какой алгоритм решения методом Гаусса?
5. Какой из трех изученных методов наиболее трудоемкий, а какой более точный?

Раздел 5. Решение обыкновенных дифференциальных уравнений

Обыкновенные *дифференциальные уравнения* (ОДУ) представляют собой зависимость между искомой функцией одного аргумента, самим аргументом и производными функции различного порядка. Такие уравнения в общем виде можно записать так: $F(x, y, y', ..., y^{(n)})$. При этом старшая степень производной n называется порядком дифференциального уравнения. Общее решение дифференциального уравнени n-го порядка содержит n произвольных постоянных $C_1, \ldots, C_n$, т. е. общее решение имеет вид $y = \varphi(C_1, ..., C_n)$.

Для выделения единственного решения необходимо задать n дополнительных условий. В зависимости от способа задания этих условий существуют два различных типа задач: *задача Коши* и *краевая задача*, которые и применяются для решения практических задач.

Если дополнительные условия задаются в одной точке, то такая задача называется *задачей Коши.* Дополнительные условия в задаче Коши называются начальными условиями, а точка x_0, в которой они задаются, – начальной точкой. Решения задачи Коши находятся для $x > x_0$.

Если дополнительные условия задаются в более чем одной точке, т. е. при различных значениях независимой переменной, то такая задача называется *краевой.* Сами дополнительные условия называются краевыми, или граничными. На практике, как правило, краевые условия задаются в двух точках ($x = a$ и $x = b$), которые являются границами решения дифференциальных уравнений.

5.1. Решение задачи Коши

Как ни странно, но в Excel нет встроенных методов решения дифференциальных уравнений.

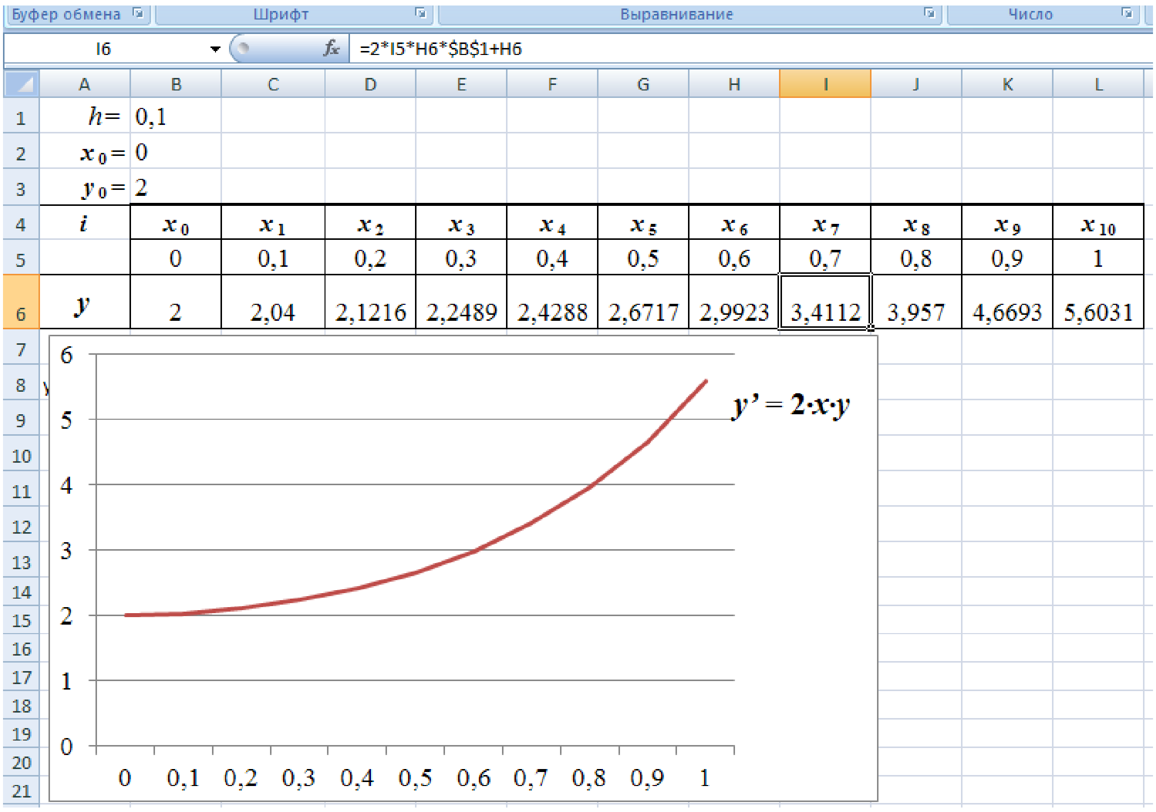
Для решения дифференциального уравнения в Excel применяется метод конечных разностей, суть которого состоит в том, что область непрерывного значения аргумента заменяется дискретным множеством точек – узлами. Узлы составляют разностную сетку. Функция, рассчитанная на заданной сетке, называется сеточной функцией. Таким образом, решение дифференциального уравнения сводится к отысканию значений сеточной функции в узлах, то есть целью является построение таблицы

$\boldsymbol{x_i}$	x_0	x_1	…	x_n
$\boldsymbol{y_i}$	y_0	y_1	…	y_n

При нахождении приближенного решения вычисления проводятся с расчетным шагом $h = \frac{x_n - x_0}{n}$, расчетными узлами служат точки $x_i = x_0 + ih$, где $i = 0, 1, \ldots, n$, промежутка $[x_0, \; x_n]$. Значения сеточной функции рассчитываются по формуле $y_{i+1} = y_i + hf(x_i, y_i, y_{i-k+1})$. Такая замена называется аппроксимацией.

***Пример** 9*. Необходимо решить дифференциальное уравнение $y' = 2xy$, $y(0) = 2$ на отрезке [0, 1].

Решение. Разобьем заданный отрезок [0, 1] на 10 частей с шагом 0,1 и заменим заданное дифференциальное уравнение сеточной функцией $y_{i+1} = y_i + h2x_iy_i$ (рис. 28).



$h=$	0,1										
$x_0=$	0										
$y_0=$	2										
i	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
y	2	2,04	2,1216	2,2489	2,4288	2,6717	2,9923	3,4112	3,957	4,6693	5,6031

Рис. 28. Решение $y' = 2xy$, $y(0) = 2$ на отрезке [0, 1]

Если в правой части уравнения $y_{i+1} = y_i + hf(x_i, y_i, y_{i-k+1})$ отсутствует y_{i+1}, то есть функции рассчитывается только по предыдущим значениям, то метод считается *явным k-шаговым*.

Если в правой части уравнения $y_{i+1} = y_i + hf(x_i, y_i, y_{i-k+1})$ присутствует y_{i+1}, то его решение усложняется и применяются итерационные методы, называемые *неявными*.

Мы рассмотрим два наиболее простых явных метода решения задачи Коши.

5.2. Решение задачи Коши методом Эйлера

Метод Эйлера основан на разложении искомой функции в ряд Тейлора в окрестностях узлов сетки. При этом отбрасываются все члены, содержащие производные выше первого порядка. Алгоритм решения задачи Коши методом Эйлера достаточно простой:

1. Задание начальных данных.
2. $i = 0$, $y_0 = y0$.
3. $y_{i+1} = y_i + hf(x_i, y_i)$.
4. Если $i < N$, перейти на пункт 3.
5. Вывод решения задачи Коши в узловых точках и построение графика искомой функции по точкам $\{x_i , y_i\}$, $i = 0, 1, 2, \ldots, N$.

N – число разбиений отрезка $[a, b]$ на подынтервалы постоянной длины h. При этом шаг разбиения $h = \dfrac{b-a}{N}$. Координаты узловых точек x_i вычисляются по формуле $x_i = a + hi$, где $i = 0, 1, 2, \ldots, N$. Значения сеточной функции $y_1 = y_0 + hf(x_0, y_0)$, $y_2 = y_1 + hf(x_1, y_1)$ и т. д.

Погрешность метода на каждом шаге вычислений имеет порядок h^2.

5.3. Решение задачи Коши методом Рунге–Кутта второго порядка

Метод Рунге–Кутта второго порядка является самым популярным, т. к. дает хороший порядок точности. Описывается формулой

$$y_0 = y0,\ \ y_{i+1} = y_i + ((1 - \alpha)\, k_1 + \alpha \cdot k_2) \cdot h,$$

где $k_1 = f\,(x_i, y_i\,)$, $k_2 = f\,(x_i + \dfrac{h}{2\alpha}, y_i + \dfrac{h}{2\alpha}k_1)$; α – параметр на интервале [0, 1], который задает скорость сходимости численного решения, подбирается для конкретной задачи. При малых значениях числа разбиений N наиболее оптимальным значением α является 0,26.

5.4. Решение задачи Коши методом Рунге–Кутта четвертого порядка

На практике для решения задачи Коши чаще всего применяют метод Рунге–Кутта четвертого порядка точности, который описывается формулой

$$y_0 = y0,\ \ y_{i+1} = y_i + (k_1 + 2 \cdot k_2 + 2 \cdot k_3\ + k_4)\,\frac{h}{6},$$

где

$$k_1 = f(x_i, y_i), \qquad k_2 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2}k_1),$$

$$k_3 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2}k_2), \qquad k_4 = f(x_i + h, y_i + h \cdot k_3).$$

Погрешность метода на каждом шаге вычислений имеет порядок h^5.

Пример 10. Решим уравнение $dy/dx = 2(x^2 + y)$ на интервале [0 , 1], начальное значение $y(0) = 1$, $0 \le x \le 1$, $h = 0{,}1$. Для оценки точности задано точное решение в виде функции $u(x) = 1{,}5\ell^{2x} - x^2 - x - 0{,}5$.

1. Расчеты по формуле Эйлера.

В столбец A занесем номер итерации начиная с 0, в столбец B – значения x с шагом $h = 0{,}1$ начиная с 0. В столбце C будем проводить расчеты y_{i+1} по формуле Эйлера, в столбце D – точное решение, в столбце E – погрешность вычисления по формуле Эйлера с точным расчетом (рис. 29).

C14 =C13+2*(B13^2+C13)*$C$8

	A	B	C	D	E	F
7						
8		*h*=	0,1			
9		*x*₀=	0			
10		*y*₀=	1			
11		***X***	***Эйлер***	***Точное***	***Ошибка вычисления***	***k1***
12	0	0	1	1	0	
13	1	0,1	1,2	1,2221	0,02210	
14	2	0,2	1,442	1,49774	0,05574	
15	3	0,3	1,7384	1,84318	0,10478	
16	4	0,4	2,10408	2,27831	0,17423	
17	5	0,5	2,5569	2,82742	0,27053	
18	6	0,6	3,11828	3,52018	0,40190	
19	7	0,7	3,81393	4,3928	0,57887	
20	8	0,8	4,67472	5,48955	0,81483	
21	9	0,9	5,73766	6,86447	1,12681	
22	10	1	7,04719	8,58358	1,53639	

Рис. 29. Таблица расчета функции y' = 2(x2 + y) методом Эйлера и ее погрешности по сравнению с точным значением

2. Расчеты по формуле Рунге–Кутта второго порядка точности.

В ячейке F13 рассчитаем коэффициент k_1 по формуле =2*(B12^2+H12), где значение $y_0 = 1$ в ячейке H12 из начального условия.

В ячейке F12 рассчитаем k_2 по формуле =2*(B12+$C$8/(2*$H$10))^2+(H12+F13*($C$8/(2*$H$10))), где в ячейке H10 записан параметр α (в данном случае 0,26).

В ячейке H13 рассчитано значение y_1 по формуле = H12 – ((1– $H$10)*F13 +$H$10*G13)*$C$8, а в ячейке I13 – погрешность вычисления по формуле Рунге–Кутта второго порядка с точным значением по формуле = H13 – E13. Далее копируем все расчеты вниз на 9 строк до y_{10} (рис. 30).

=H14+((1-$H$10)*F15+$H$10*G15)*$C$8

	E	F	G	H	I
			$y'=2\cdot(x^2+y)$		
			α=	0,26	
а ни	***Точное***	***k1***	***k2***	***II Рунге-Кутта***	*Ошибка вычислени*
	1			1	
1	1,222104	2	2,8432	1,2219	-0,0002
7	1,4977	2,46385	3,56237	1,4969	-0,0009
8	1,8432	3,07374	4,48376	1,8409	-0,0023
2	2,2783	3,86181	5,65185	2,2736	-0,0047
5	2,8274	4,86725	7,12093	2,8189	-0,0085
9	3,5202	6,13789	8,9572	3,5060	-0,0141
9	4,3928	7,73207	11,2415	4,3705	-0,0223
8	5,4895	9,72098	14,0722	5,4557	-0,0338
8	6,8645	12,1914	17,5698	6,8147	-0,0498
4	8,5836	15,2494	21,8808	8,5121	-0,0715

Рис. 30. Таблица расчета функции y'=2(x2 + y) методом Рунге–Кутта II порядка и погрешность

3. Расчеты по формуле Рунге–Кутта четвертого порядка точности.

В ячейке J13 рассчитаем коэффициент k_1 по формуле = 2*(B12^2 + N12), где значение $y_0 = 1$ в ячейке N12 из начального условия.

В ячейке K13 рассчитаем k_2 по формуле = 2*((B12 + $C$8/2)^2 + (N12 + J13*($C$8/2))), где в ячейке C8 шаг $h = 0{,}1$. В ячейке L13 рассчитаем k_3 по формуле = 2*((B12 + $C$8/2)^2 + (N12 + K13*($C$8/2))), а в ячейке M13 – коэффициент k_4 по формуле = 2*((B12 + $C$8)^2 + (N12 + L13*$C$8)).

В ячейке N13 рассчитано значение y_1 по формуле = N12 + (J13 + 2*K13 + 2*L13 + M13)*$C$8 / 6, а в ячейке O13 – погрешность вычисления y_1 по формуле Рунге–Кутта четвертого порядка с точным значением по формуле =N13 – E13. Далее копируем все расчеты вниз на 9 строк до y_{10} (рис. 31).

=N12+(J13+2*K13+2*L13+M13)*$C$8/6

J	K	L	M	N	O
k1	***k2***	***k3***	***k4***	***IV Рунге-Кутта***	*Ошибка вычислен*
				1	
2,0000	2,2050	2,2255	2,4651	1,2221	0,0000
2,4642	2,7356	2,76277	3,07676	1,4977	0,0000
3,0755	3,4280	3,46326	3,86811	1,8432	0,0000
3,8663	4,3180	4,36313	4,87896	2,2783	0,0000
4,8766	5,4492	5,5065	6,15788	2,8274	0,0000
6,1548	6,8753	6,94731	7,76424	3,5201	-0,0001
7,7603	8,6613	8,75138	9,77053	4,3927	-0,0001
9,7655	10,8870	10,9992	12,2653	5,4894	-0,0001
12,2589	13,6498	13,7889	15,3567	6,8643	-0,0001
15,3486	17,0685	17,2405	19,1767	8,5834	-0,0002

Рис. 31. Таблица расчета функции y'=2(x2 + y) методом Рунге–Кутта VI порядка и погрешность

Отобразим полученные результаты на графике (рис. 32).

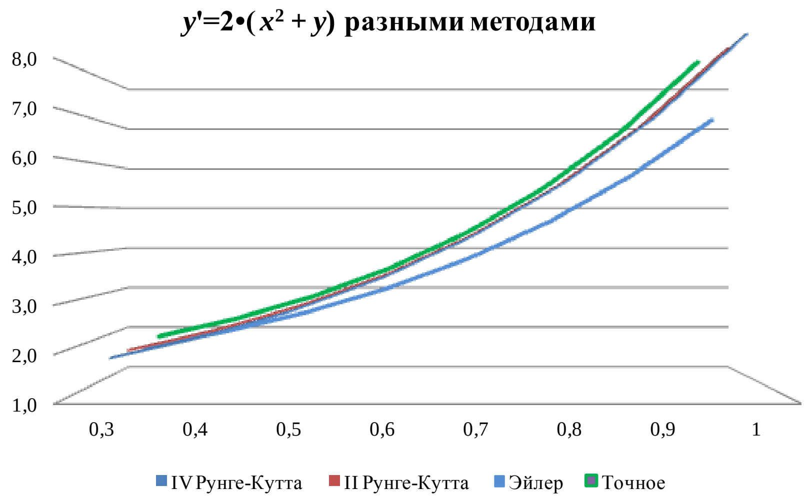


Рис. 32. Графики функции y' = 2(x2 + y) на интервале [0; 1], рассчитанной методами Эйлера, Рунге–Кутта II, VI порядка, и точного решения u = 1,5ℓ2x – x2 – x – 0,5

Если кривые на графике сливаются и трудноразличимы, то нужно изменить формат рядов данных, вызвав контекстное меню и проделав необходимую корректировку (рис. 32).

Вопросы и задания для самопроверки

1. Дайте определение понятию обыкновенные дифференциальные уравнения.

2. Какие условия налагают краевая задача, задача Коши?

3. Раскройте алгоритм решения задачи Коши методом Эйлера.

4. Чем отличается решение задачи Коши методом Рунге–Кутта второго порядка от решения методом Рунге–Кутта четвертого порядка?

Раздел 6. Аппроксимация экспериментальных данных

В вычислительной практике часто возникает задача восстановления функции $f(x)$ на отрезке $[a, b]$, если известны ее значения $y_0 = f(x_0)$, $y_1 = f(x_1),\ldots, y_n = f(x)$ в отдельных фиксированных точках x_0 , $x_1,\ldots, x_n$ отрезка. Такая задача имеет место при табличном задании функции. Значения $y_0, y_1,\ldots, y_n$ в этом случае – продукт измерения физической величины на наборе x_0, $x_1, \ldots,$ x_n аргумента x. Чтобы приближенно восстановить функцию $f(x)$ на всем отрезке $[a, b]$, строят аппроксимирующую функцию $\varphi(x)$, расчеты по которой в определенном смысле приближаются к экспериментально полученным значениям.

К табличному заданию функции прибегают также в том случае, когда аналитический вид функции $f(x)$ известен, но сложен и требует большого объема вычислений для определения ее отдельных значений. Над такой функцией, кроме того, трудно выполнить математические операции дифференцирования и интегрирования. Замена $f(x)$ приближенной функцией $j(x)$ позволяет упростить вычисления.

Пусть величина y является функцией аргумента x, но вид аналитической зависимости $y = f(x)$ неизвестен, функция задана в виде таблицы в точках x_0, x_1, ..., x_n, где значения $y_0 = f(x_0), \ldots, y_n = f(x_n)$. Решается задача замены функции $f(x)$ более простой функцией $\varphi(x)$, чтобы эти функции совпадали в заданной области наилучшим образом. Одним из методов решения данной проблемы является аппроксимация: построение функции $\varphi(x) = P_k(x)$, *аппроксимирующей* данную функцию $f(x)$ и являющейся многочленом степени k вида $P_k(x) = \Sigma p_j x^j$ $(j = 0, \ldots, k)$, коэффициенты которого p_j определяются из условия минимизации суммы квадратов отклонений (*метод наименьших квадратов*) в точках $x_0, x_1, \ldots, x_n$ значений многочлена $P_k(x)$ от $y_0, y_1, \ldots, y_n$ соответственно, т. е. путем решения задачи: $\min S = \Sigma(P_k(x_i) - y_i)^2$.

Для аппроксимации функции сначала необходимо выбрать степень аппроксимирующего многочлена, т. е. среди многочленов разной степени выбрать тот, который лучше отражает поведение заданной таблично функции. Возможны следующие варианты функций:

1. Линейная: $\varphi(x) = a \cdot x + b$. Обычно применяется в простейших случаях, когда экспериментальные данные возрастают или убывают с постоянной скоростью (рис. 33).

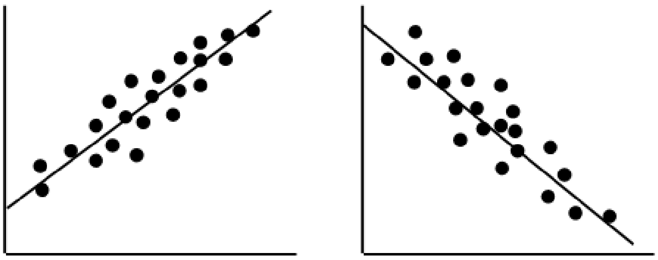


Рис. 33. Линейная аппроксимация исходных данных

2. Полиноминальная: $\varphi(x) = a_0 + a_1x + a_2x^2 + a_nx^n$, где a_i – константы. Степень полинома определяется количеством экстремумов (рис. 34).

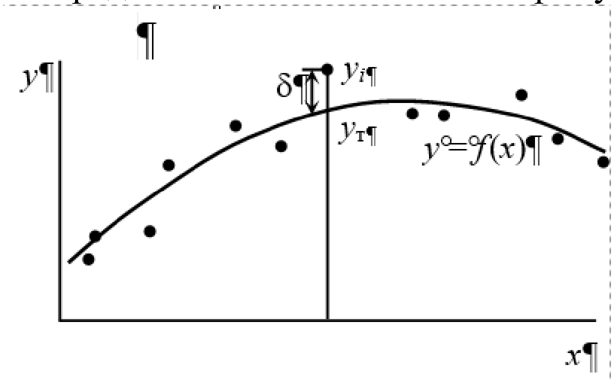


Рис. 34. Полиноминальная аппроксимация исходных данных

3. Логарифмическая: $\varphi(x) = a \cdot \ln x + b$. Применяют для описания данных, которые вначале быстро растут или убывают, а затем стабилизируются.

4. Степенная: $\varphi(x) = ax^n$. Данные не должны иметь нулевых или отрицательных значений.

5. Экспоненциальная: $\varphi(x) = be^{ax}$. Данные не должны иметь нулевых или отрицательных значений (рис. 35).

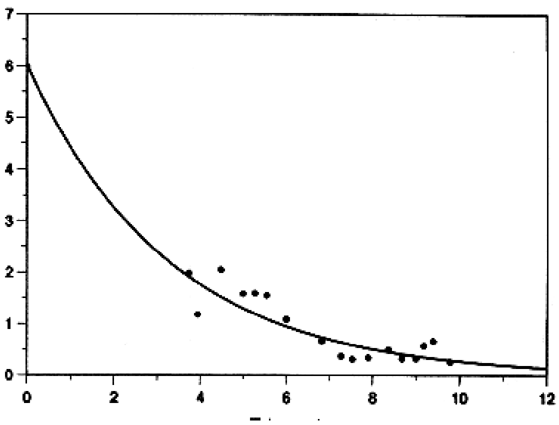


Рис. 35. Экспоненциальная аппроксимация исходных данных

Затем с помощью метода наименьших квадратов определить параметры p_j аппроксимирующего многочлена, чтобы $S = \Sigma(P_k(x_i) - y_i)^2 \to \min$.

Для их поиска дифференцируют функционал S по p_j и получают систему линейных уравнений, содержащую $(k + 1)$ линейное уравнение, зависящее от $(k + 1)$ неизвестного параметра $p_0, p_1, ..., p_k$:

$$S = \Sigma(p_0 + p_1x + p_2x^2 + p_kx^k - y_i)^2.$$

После определенных математических операций получим:

$$(n+1)\,p_0 + p_1\sum x_i + p_2\sum x_i^2 + \ldots + p_k\sum x_i^k = \sum y_i,\ \ i = 0, \ldots n;$$

$$p_0\sum x_i + p_1\sum x_i^2 + p_2\sum x_i^3 + \ldots + p_k\sum x_i^{k+1} = \sum y_i\,x_i,\ \ i = 0, \ldots n;$$

$$\cdots\cdots\cdots$$

$$p_0\sum x_i^k + p_1\sum x_i^{k+1} + p_2\sum x_i^{k+2} + \ldots + p_k\sum x_i^{2k} = \sum y_i\,x_i^k,\ \ i = 0, \ldots n.$$

Решая эту систему линейных уравнений, находим значения параметров.

Пример 11. Найти с помощью метода наименьших квадратов аппроксимирующий многочлен для таблично заданной функции $y = f(x)$, приняв предположение, что $f(x)$ является линейной.

x	0,75	1,5	2,25	3,0	3,75
y	2,5	1,2	1,12	2,25	4,28

Решение. Так как исходная функция предполагается линейной, то в качестве аппроксимирующего многочлена выберем многочлен первой степени вида

$$P_1(x) = p_0 + p_1x.$$

Тогда $S = \Sigma(p_0 + p_1x_i - y_i)^2$, $i = 0, \ldots, 4$. Система линейных уравнений для поиска параметров p_0 и p_1 будет иметь следующий вид:

$$(n + 1)p_0 + p_1\Sigma x_i = \Sigma y_i, \quad i = 0, \ldots, 4, \text{ т. е. } n = 4;$$

$$p_0\Sigma x_i + p_1\Sigma x_i^2 = \Sigma y_ix_i, \quad i = 0, \ldots, 4.$$

Проведем расчеты значений коэффициентов при неизвестных параметрах (табл. 2).

Таблица 2

$n + 1 = 5$

i	x	y	x^2	$x\,y$
0	0,75	2,5	0,56	1,88
1	1,5	1,2	2,25	1,80
2	2,25	1,12	5,06	2,52
3	3	2,25	9,00	6,75
4	3,75	4,28	14,06	16,05
Σ	11,25	11,35	30,94	29,00

Получаем систему

$$\begin{cases} 5p_0 + 11{,}25 \cdot p_1 = 11{,}35; \\ 11{,}25p_0 + 30{,}94p_1 = 29, \end{cases}$$

решая которую любым известным методом (например, матричным), получим: $p_0 \approx 0{,}89$, $p_1 \approx 0{,}62$.

Тогда аппроксимирующая функция имеет вид

$$y = 0{,}89 + 0{,}62x.$$

Рассчитаем для нее таблицу и построим график, отобразив исходные данные (рис. 36).

i	x	$y_{линей}$
0	0,75	1,348
1	1,5	1,809
2	2,25	2,27
3	3	2,731
4	3,75	3,192

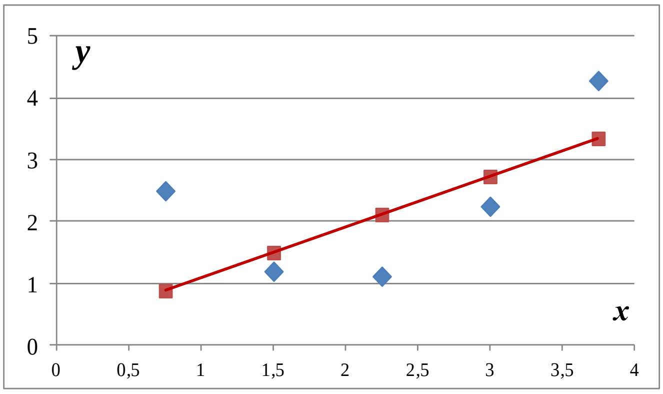


Рис. 36. Таблица и график линейной аппроксимации и исходных данных

Желание исследователей описывать результаты опытов стандартными статистическими распределениями понятно. Их основные характеристики хорошо изучены. В ряде случаев статистические функции, согласующиеся с экспериментальными данными, позволяют выявлять динамические процессы, обусловливающие флуктуации физических величин. Но в тех случаях, когда экспериментальное распределение имеет сложную форму и не описывается стандартной функцией распределения, все-таки удобно иметь аппроксимационную формулу, удовлетворительно описывающую данные. Нередко в таких случаях в качестве аппроксимационной функции выбирают полиномы различных степеней.

Пример 12. Найти с помощью метода наименьших квадратов аппроксимирующий многочлен для функции $y = f(x)$ примера 1, приняв предположение, что $f(x)$ является квадратичной.

x	0,75	1,5	2,25	3,0	3,75
y	2,5	1,2	1,12	2,25	4,28

Решение. Так как исходная функция предполагается квадратичной, то в качестве аппроксимирующего многочлена выберем многочлен второй степени вида

$$P_2(x) = p_0 + p_1x + p_2x^2.$$

Тогда $S = \Sigma(p_0 + p_1x_i + p_2x_i^{\,2} - y_i)^2$, $i = 0, \ldots, 4$. Система линейных уравнений для поиска параметров p_0, p_1 и p_2 будет иметь следующий вид:

$$(n + 1)p_0 + p_1\sum x_i + p_2\sum x_i^2 = \sum y_i, \quad i = 0, \ldots, 4;$$

$$p_0\sum x_i + p_1\sum x_i^2 + p_2\sum x_i^3 = \sum y_i x_i, \quad i = 0, \ldots, 4;$$

$$p_0\sum x_i^2 + p_1\sum x_i^3 + p_2\sum x_i^4 = \sum y_i x_i^2, \quad i = 0, \ldots, 4.$$

Проведем расчеты значений коэффициентов при неизвестных параметрах (табл. 3).

Таблица 3

$n + 1 = 5$

i	**x**	**y**	x^2	x^3	x^4	xy	x^2y
0	0,75	2,5	0,56	0,42	0,32	1,88	1,41
1	1,5	1,2	2,25	3,38	5,06	1,80	2,70
2	2,25	1,12	5,06	11,39	25,63	2,52	5,67
3	3	2,25	9,00	27,00	81,00	6,75	20,25
4	3,75	4,28	14,06	52,73	197,75	16,05	60,19
Σ	11,25	11,35	30,94	94,92	309,76	29,00	90,21

Получаем систему

$$\begin{cases} 5\cdot p_0 + 11{,}25\cdot p_1 + 30{,}94\cdot p_2 = 11{,}35; \\ 11{,}25\cdot p_0 + 30{,}94\cdot p_1 + 94{,}92\cdot p_2 = 29; \\ 30{,}94\cdot p_0 + 94{,}92\cdot p_1 + 309{,}76\cdot p_2 = 90{,}21, \end{cases}$$

решая которую любым известным методом, получим: $p_0 \approx 4{,}82$, $p_1 \approx -3{,}88$, $p_2 \approx 1{,}00$.

Аппроксимирующая функция имеет вид

$$y = 4{,}82 - 3{,}88x + x^2.$$

Рассчитаем для нее таблицу и построим график, отобразив исходные данные (рис. 37).

i	*x*	$y_{кв.}$
0	0,75	2,472
1	1,5	1,247
2	2,25	1,146
3	3	2,169
4	3,75	4,316

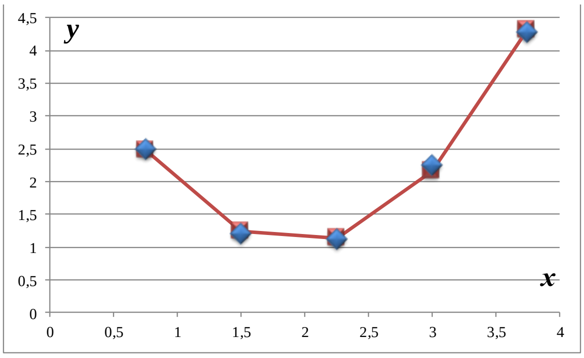


Рис. 37. Таблица и график квадратичной аппроксимации и начальных данных

Пример 13. Найти с помощью метода наименьших квадратов аппроксимирующий многочлен для таблично заданной функции $y = f(x)$, приняв предположение, что $f(x)$ является экспоненциальной:

x	0,75	1,5	2,25	3,0	3,75
y	2,5	1,2	1,12	2,25	4,28

Решение. Экспоненциальная функция имеет вид

$$y = a_0 \cdot e^{a1x}.$$

Приведем ее к линейному виду, прологарифмируя:

$$\ln y = \ln a_0 + a_1 x,$$

и обозначив:

$$\varphi = \ln y; \quad p_0 = \ln a_0; \quad p_1 = a_1.$$

Тогда система линейных уравнений для поиска параметров p_0 и p_1 будет иметь следующий вид:

$$(n + 1)p_0 + p_1\Sigma x_i = \Sigma\varphi_i, \qquad i = 0, \ldots, 4;$$
$$p_0\Sigma x_i + p_1\Sigma x_i^2 = \Sigma\varphi_i x_i, \qquad i = 0, \ldots, 4.$$

Проведем расчеты значений коэффициентов при неизвестных параметрах (табл. 4).

Таблица 4

i	*x*	*y*	x^2	*φ*	*φx*
0	0,75	2,5	0,56	0,916	0,687
1	1,5	1,2	2,25	0,182	0,273
2	2,25	1,12	5,06	0,113	0,255
3	3	2,25	9,00	0,811	2,433
4	3,75	4,28	14,06	1,454	5,452
Σ	11,25	11,35	30,94	3,48	9,10

Получаем систему

$$\begin{cases} 5p_0+11{,}25p_1=3{,}48; \\ 11{,}25p_0+30{,}94p_1=9{,}1, \end{cases}$$

решая которую любым методом, получим: $p_0 \approx 0{,}18$, $p_1 \approx 0{,}23$.

После этого возвратимся к показательной функции и найдем параметры $a_0 = e^{p0} = e^{0,18} = 1{,}2$ и $a_1 = 0{,}23$.

Следовательно, аппроксимирующая функция имеет вид $y = 1{,}2e^{0,23x}$. Рассчитаем для нее таблицу и построим график, отобразив исходные данные (рис. 38).

i	*x*	*y*	*y экс.*
0	0,75	2,5	1,426
1	1,5	1,2	1,690
2	2,25	1,12	2,004
3	3	2,25	2,377
4	3,75	4,28	2,818

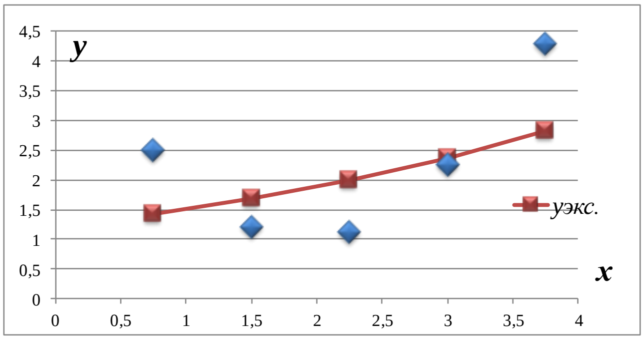


Рис. 38. Таблица и график экспоненциальной аппроксимации и начальных данных

В Excel имеются специальные возможности, помогающие найти и построить аппроксимирующие функции.

1. Функция для вычисления коэффициентов регрессии:

ЛИНЕЙН (массив_y; массив_x; Конст; статистика). Конст – необязательное значение, которое указывает, требуется ли, чтобы выполнялось соотношение $y = ax$. Статистика – логическое значение, указывающее, требуется ли вернуть дополнительную статистику по регрессии.

2. Функция для расчета значения по методу наименьших квадратов:

ТЕНДЕНЦИЯ (массив_y; массив_x; Новый_массив_x; Конст),
где массив_y – значения функции, которые уже известны для соотношения $y = ax + b$. Новый_массив_x – новые значения x, для которых функция ТЕНДЕНЦИЯ возвращает ожидаемые значения y. Если этот параметр пропущен, то предполагается, что он совпадает с массивом x. Конст – аналогично ЛИНЕЙН().

Пример 14. Для таблично заданной функции $y = f(x)$ построить аппроксимирующую функцию, вычислить ожидаемое значение в точках 0; 0,75; 1,75; 2,8; 4,5.

0	0,5	1	1,5	2	2,5	3	3,5	4
5,75	3,39	2,81	3,25	3,75	4,11	4,85	4,45	5,25

Решение. Для того, чтобы рассчитать значения коэффициентов p_0 и p_1, выделим ячейки K72:L72, обратимся к мастеру функций и в категории *Статистические* выберем функцию ЛИНЕЙН. Заполним появившееся диалоговое окно соответствующим образом (рис. 39), установив курсор в окне редактирования. Нажмем клавишу F2, а затем комбинацию клавиш Ctrl + Shift + Enter. В результате в ячейках K72:L72 появятся значения $p_0 = 0{,}204$, $p_1 = 3{,}77$.

В случае линейно заданной функции аппроксимирующая функция будет иметь вид

$$y = 3{,}77 + 0{,}204\,x.$$

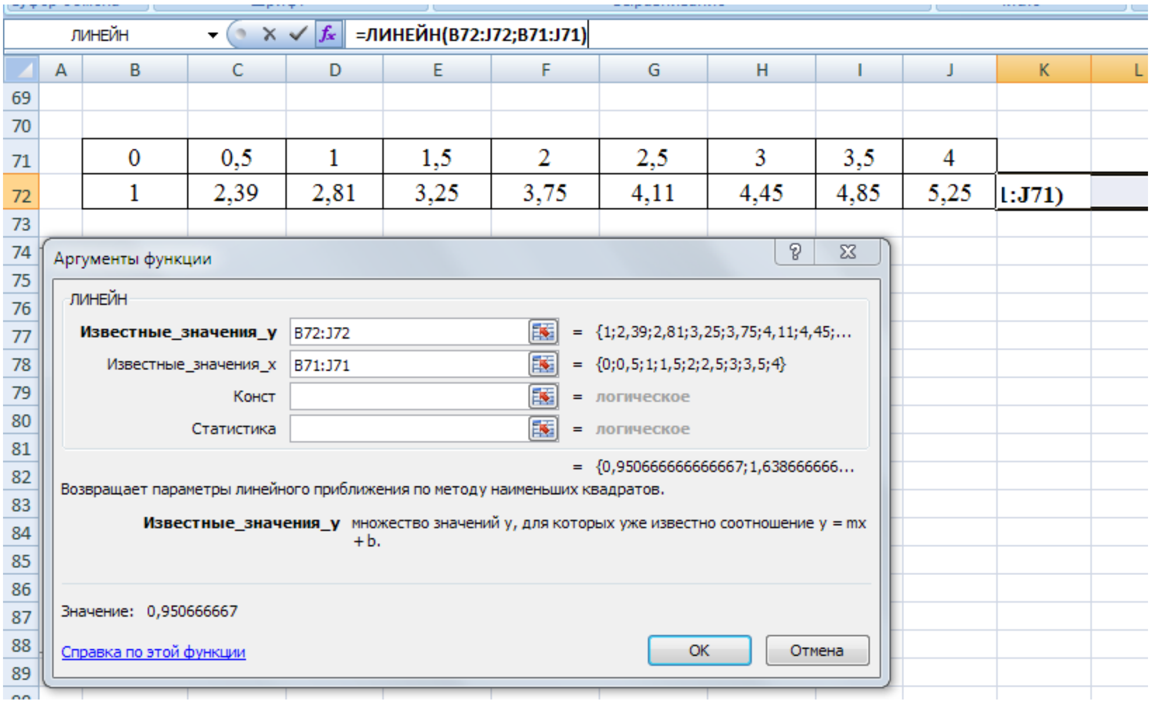


Рис. 39. Окно мастера функции ЛИНЕЙН()

Для вычисления ожидаемого значения в точках 0; 0,75; 1,75; 2,8; 4,5 занесем их в ячейки B74:F74. Затем выделим диапазон ячеек B75:F75, обратимся к мастеру функций и в категории *Статистические* выберем функцию ТЕНДЕНЦИЯ. Заполним появившееся диалоговое окно соответствующим образом (рис. 40), установив курсор в окне редактирования. Нажмем клавишу F2, а затем комбинацию клавиш Ctrl + Shift + Enter. В результате в ячейках B75:F75 появятся значения функции y:

3,771	3,924	4,128	4,342	4,609

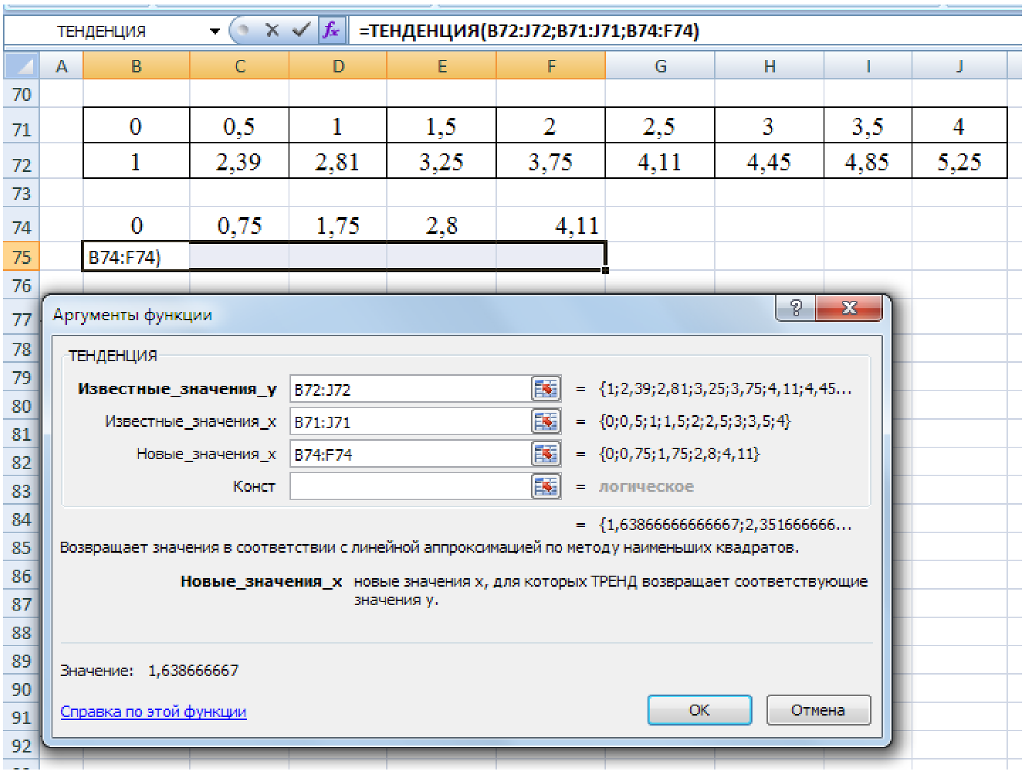


Рис. 40. Окно мастера функции ТЕНДЕНЦИЯ()

Теперь подберем аппроксимирующую функцию графическим методом. Для этого по исходным данным строим график. Щелкаем правой кнопкой мыши по одной из точек графика. В появившемся диалоговом окне выбираем команду *Добавить линию тренда* (рис. 41).

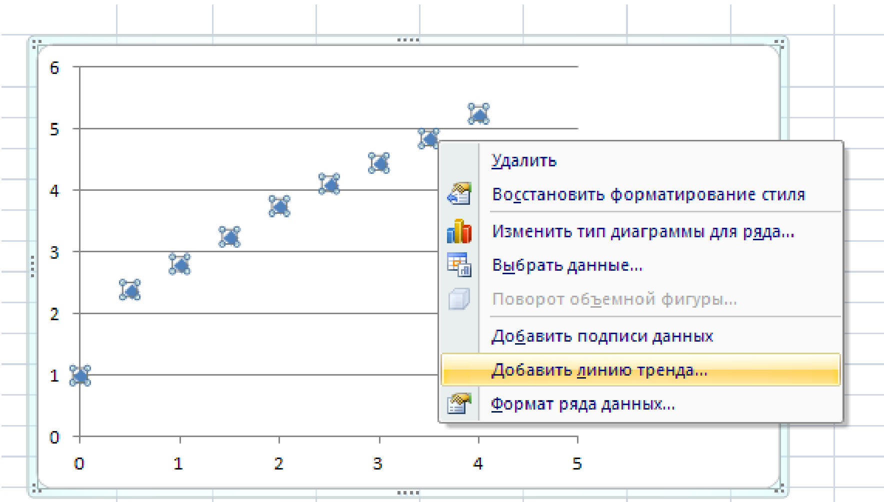


Рис. 41. Контекстное меню графика

Появляется диалоговое окно *Линия тренда* (рис. 42). На первой вкладке мы можем выбрать тип линии тренда – линейный, полиноминальный (любой степени) и т. д., дать название тренду. Также можно показать уравнение на диаграмме или величину достоверности R^2, установить флажок, если есть необходимость, на закладке. Величина достоверности показывает зависимость аппроксимирующей функции от исходных данных. Чем ближе она к 1, тем выше степень приближения. На других вкладках есть возможность форматирования линий аппроксимации (толщина, цвет и т. д.).

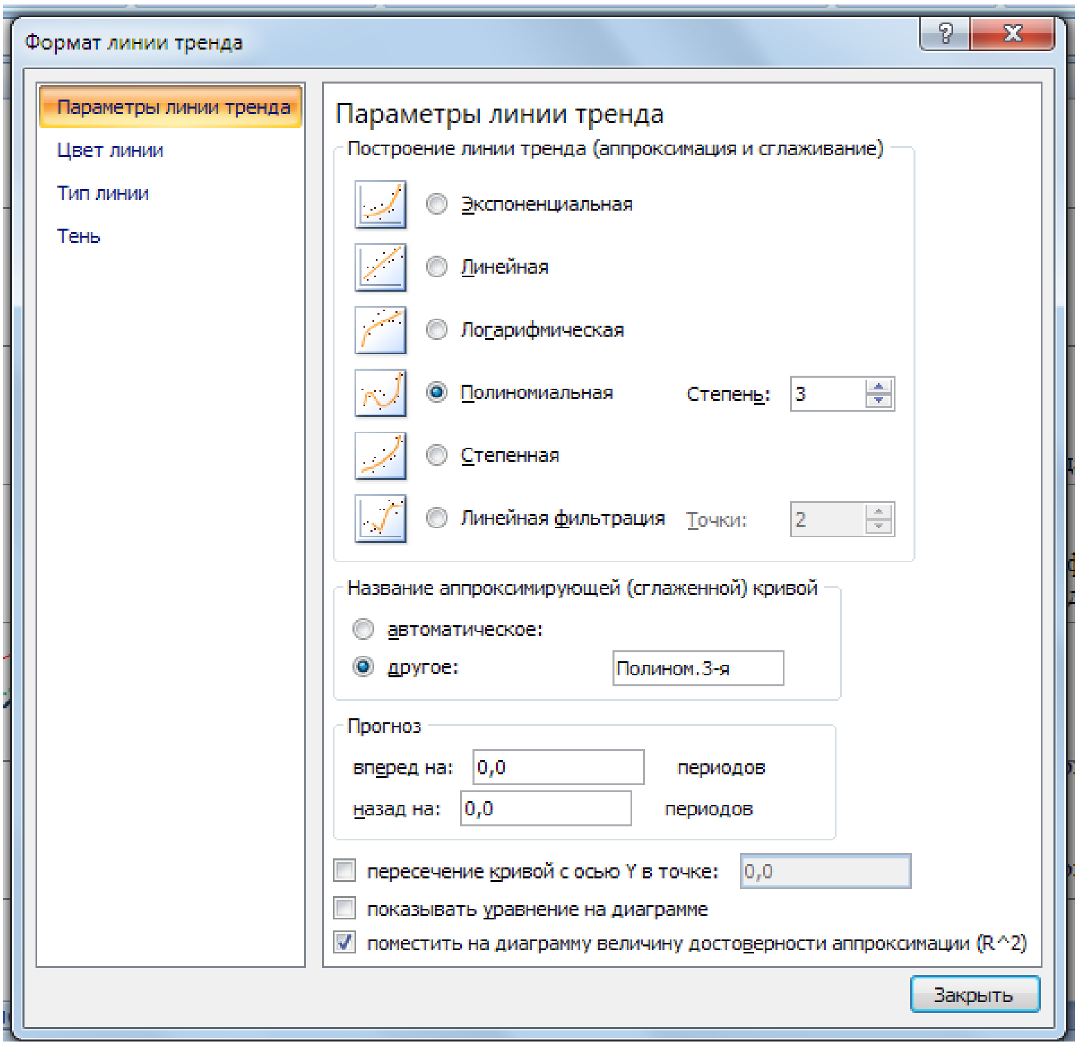


Рис. 42. Диалоговое окно Линии тренда

Проанализировав полученные на графике данные, можно сделать вывод, что лучше всего аппроксимирует исходные данные полиномиальная функция 3-й степени ($R^2 = 0{,}8759$) по сравнению с линейной ($R^2 = 0{,}0812$) и квадратичной ($R^2 = 0{,}5645$) (рис. 43).

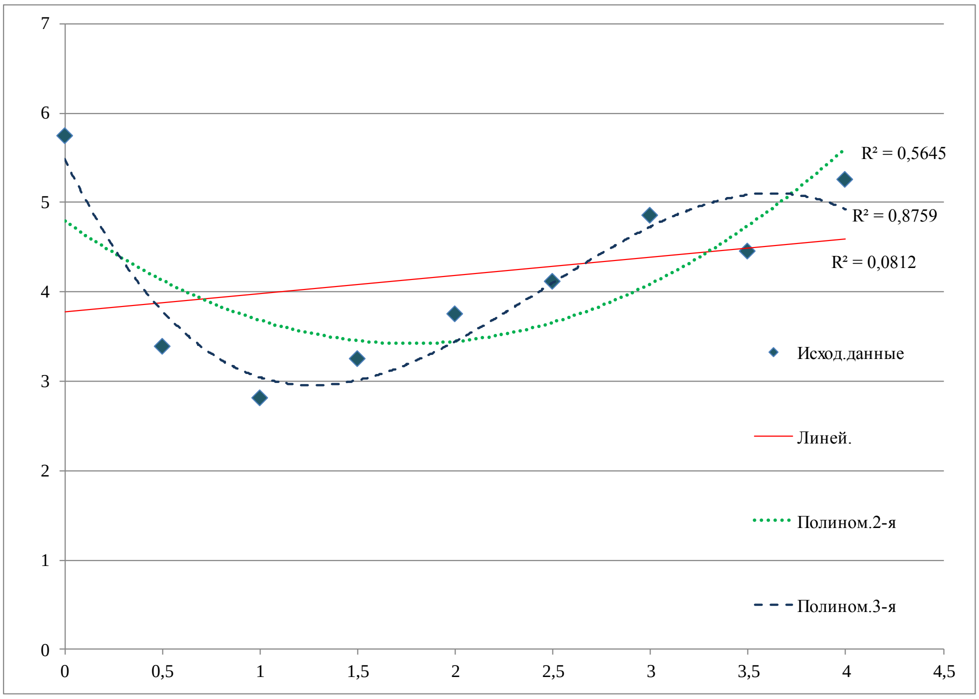


Рис. 43. Линии аппроксимирующих функций

Вопросы и задания для самопроверки

1. Что такое аппроксимация? Какая функция называется аппроксимирующей?

2. В чем заключается метод наименьших квадратов?

3. В чем суть линейной аппроксимации? Какой ее график?

4. В чем суть полиноминальной аппроксимации? Напишите формулу в общем виде. Какой ее график?

5. В чем суть логарифмической аппроксимации? Какой ее график?

6. Какие встроенные функции Excel позволяют упростить процесс аппроксимации?

6. Как можно построить на графике в Excel аппроксимирующие функции?

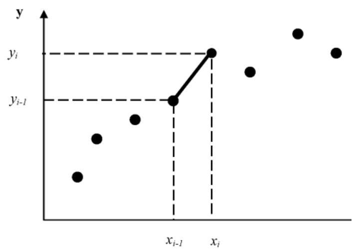
Раздел 7. Интерполяция экспериментальных данных

Пусть величина y является функцией аргумента x, но вид аналитической зависимости $y = f(x)$ неизвестен, функция задана таблично в точках x_0, x_1, ..., x_n, где значения $y_0 = f(x_0)$, …, $y_n = f(x_n)$. Решается задача замены функции $f(x)$ более простой функцией $\varphi(x)$, чтобы эти функции совпадали в заданной области наилучшим образом. Одним из методов решения данной проблемы является интерполяция: построение функции $\varphi(x) = P_n(x)$, являющейся *интерполяционным многочленом* степени n и обладающей тем свойством, что ее значения в точках x_0, x_1, ..., x_n дополнительным условием является обязательное совпадение значений функций $f(x)$ и $\varphi(x)$ в заданных точках. Интерполяция функции производится на некотором отрезке $[x_0, x_n]$.

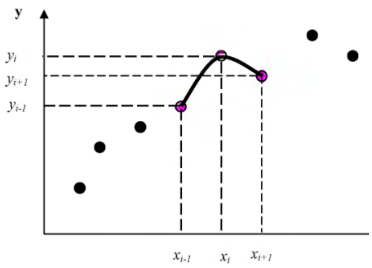
Различают *глобальную* интерполяцию, когда на весь интервал строится один многочлен, и *локальную* – когда для каждой части интервала строится свой многочлен.

Простейшая локальная интерполяция – линейная, при которой заданные точки соединяются прямыми отрезками, функция получается в виде ломаной с вершинами в данных точках. Для каждого участка $[x_{i-1}, x_i]$ в качестве интерполяционного многочлена используется уравнение прямой, проходящей через концы этого отрезка.

Для i-го интервала: $y = a_i \cdot x + b_i$, $x_{i-1} \le x \le x_i$, $a_i = \dfrac{y_i - y_{i-1}}{x_i - x_{i-1}}$, $b_i = y_{i-1} - a_i \cdot x_{i-1}$. Поэтому для нахождения приближенного значения функции в точке x сначала определяем, в какой интервал попадает значение x, а затем подставляем его в формулу.



При локальной квадратичной интерполяции в качестве интерполяционной функции на участке $[x_{i-1}, x_{i+1}]$ используется квадратный трехчлен: $y = a_i x^2 + b_i x + c_i$, $x_{i-1} \le x \le x_{i+1}$, графиком которого является парабола. Функция строится по трем точкам: x_{i-1}, x_i, x_{i+1}.



Для нахождения коэффициентов a, b и c необходимо составить систему из трех уравнений:

$$\begin{cases} a x_{i-1}^2 + b x_{i-1} + c = y_{i-1}; \\ a x_i^2 + b x_i + c = y_i; \\ a x_{i+1}^2 + b x_{i+1} + c = y_{i+1}, \end{cases}$$

решить ее на интервале $[x_{i-1}, x_{i+1}]$, а затем подставить найденные коэффициенты в уравнение $y = ax^2 + bx + c$.

Пример 15. Найти значение функции $y = f(x)$, заданной таблично в точке $x = 2$:

1) при линейной интерполяции;

2) квадратичной интерполяции.

i	***0***	***1***	***2***	***3***	***4***
x	0,75	1,5	2,25	3,0	3,75
y	2,5	1,2	1,12	2,25	4,28

Решение

1. Видно, что $x = 2$ лежит во втором интервале [1,5; 2,25], то есть $i = 2$, а коэффициенты a_2 и b_2 рассчитываются следующим образом:

$$a_2 = \frac{y_2 - y_1}{x_2 - x_1} = \frac{1{,}12 - 1{,}2}{2{,}25 - 1{,}5} = -0{,}1067;$$

$$b_2 = y_1 - a_2 \cdot x_1 = 1{,}2 - (-0{,}1067) \cdot 1{,}5 = 1{,}36.$$

Значение функции в точке $x = 2$:

$$y = a_2 \cdot x + b_2 = -0{,}1067 \cdot 2 + 1{,}36 = 1{,}147 \text{ (рис. 44).}$$

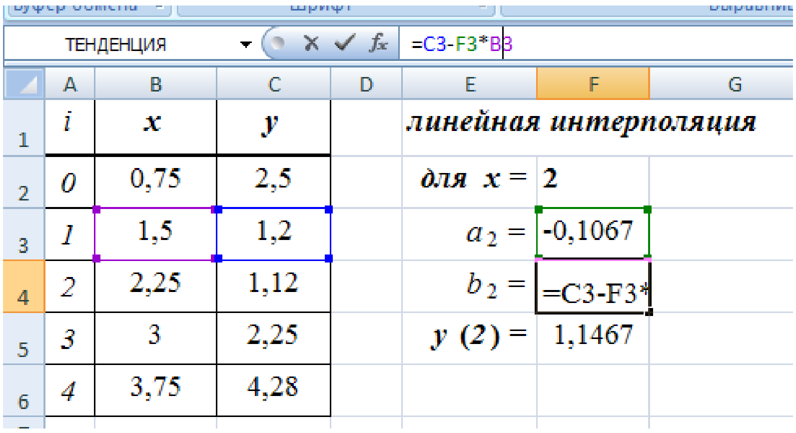
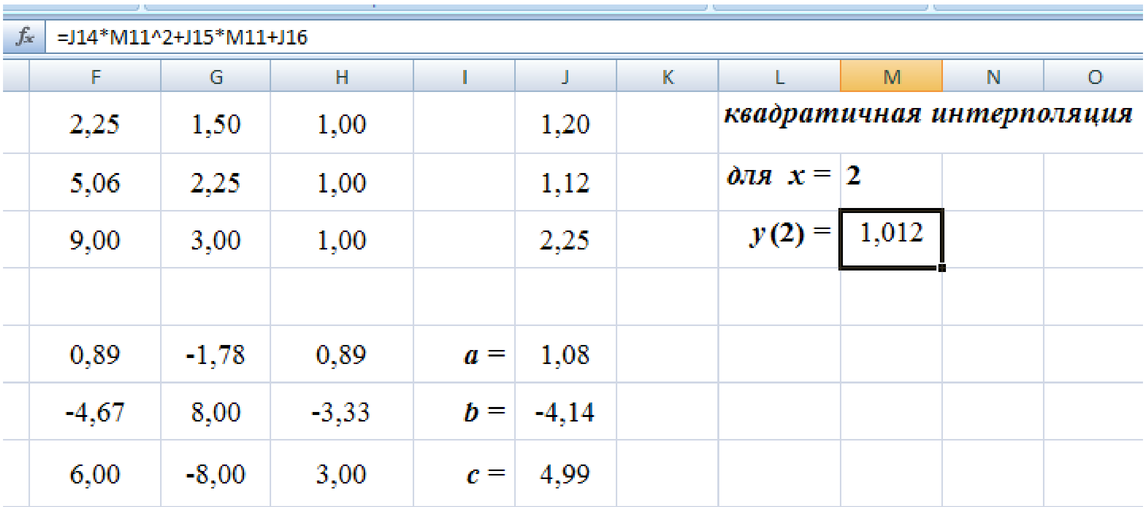


Рис. 44. Расчет линейной интерполяции для x = 2

2. Так как $x = 2$ лежит в интервале [1,5; 2,25], то коэффициенты a, b и c рассчитываются из условия:

$$\begin{cases} x_1^2 a + x_1 b + c = y_1; \\ x_2^2 a + x_2 b + c = y_2; \\ x_3^2 a + x_3 b + c = y_3; \end{cases} \qquad \begin{cases} 1{,}5^2 a + 1{,}5b + c = 1{,}2; \\ 2{,}25^2 a + 2{,}25b + c = 1{,}12; \\ 3^2 a + 3b + c = 2{,}25, \end{cases}$$

а $y(2) = a \cdot 2^2 + b \cdot 2 + c$ рассчитывают следующим образом (рис. 45):



=J14*M11^2+J15*M11+J16

F	G	H	I	J	K	L	M	N	O
2,25	1,50	1,00		1,20		***квадратичная интерполяция***			
5,06	2,25	1,00		1,12		***для x =***	**2**		
9,00	3,00	1,00		2,25		***y* (2) =**	1,012		
0,89	-1,78	0,89	***a* =**	1,08					
-4,67	8,00	-3,33	***b* =**	-4,14					
6,00	-8,00	3,00	***c* =**	4,99					

Рис. 45. Расчет квадратичной интерполяции для $x = 2$

В результате получаем: $a = 1{,}08$; $b = 4{,}14$ и $c = 4{,}99$, а значение функции $y(2) = 1{,}08 \cdot 2^2 + 4{,}14 \cdot 2 + 4{,}99$.

Заметим, что значения, полученные при локальных линейной и квадратичной интерполяциях различаются.

Имеются различные формулы для построения интерполяционного многочлена: формулы Лагранжа, Ньютона, Гаусса.

Рассмотрим интерполяционный многочлен Лагранжа. Он имеет вид

$$P_n(x) = \Sigma y_j p_j(x)\ , \quad \text{где} \ \ p_j(x) = \prod_{k=0,\ k\neq j}^{n} \frac{(x - x_k)}{(x_j - x_k)}.$$

При этом легко заметить, что каждый многочлен $p_j(x)$ обладает тем свойством, что он обращается в ноль при всех табличных значениях аргумента x, кроме $x = x_j$. А при $x = x_j$ его значение равно единице. При этом полученный интерполяционный многочлен $P_n(x)$ обладает необходимым свойством: его значения в точках $x_0, x_1, \ldots, x_n$ совпадают со значениями $y_0, y_1, \ldots, y_n$ функции $f(x)$ в этих же точках. Можно также заметить, что его степень на единицу меньше количества табличных значений функции.

Пример 16. Построить интерполяционный многочлен Лагранжа для заданной таблично функции $y = f(x)$ и его график.

i	***0***	***1***	***2***	***3***	***4***
x	188	188,1	188,2	188,3	188,4
y	62,1	64	61,5	63,5	62,1

Решение. Если имеется пять значений функции в заданных точках, то интерполяционный многочлен будет четвертой степени и иметь следующий вид:

$$P_4(x) = \Sigma y_i \cdot p_i(x),\ i = 0, \ldots, 4,$$

где $p_i(x)$ являются так же многочленами четвертой степени от x и получаются следующим образом:

$$p_0(x) = \frac{(x-188{,}1)(x-188{,}2)(x-188{,}3)(x-188{,}4)}{(188-188{,}1)(188-188{,}2)(188-188{,}3)(188-188{,}4)};$$

$$p_1(x) = \frac{(x-188)(x-188{,}2)(x-188{,}3)(x-188{,}4)}{(188{,}1-188)(188{,}1-188{,}2)(188{,}1-188{,}3)(188{,}1-188{,}4)};$$

$$p_2(x) = \frac{(x-188)(x-188{,}1)(x-188{,}3)(x-188{,}4)}{(188{,}2-188)(188{,}2-188{,}1)(188{,}2-188{,}3)(188{,}2-188{,}4)};$$

$$p_3(x) = \frac{(x-188)(x-188{,}1)(x-188{,}2)(x-188{,}4)}{(188{,}3-0188)(188{,}3-188{,}1)(188{,}3-188{,}2)(188{,}3-188{,}4)};$$

$$p_4(x) = \frac{(x-188)(x-188{,}1)(x-188{,}2)(x-188{,}3)}{(188{,}4-188)(188{,}4-188{,}1)(188{,}4-188{,}2)(188{,}4-188{,}3)}.$$

Сам интерполяционный многочлен имеет вид

$$P_4(x) = 62{,}1p_0(x) + 64p_1(x) + 61{,}5p_2(x) + 63{,}5p_3(x) + 62{,}1p_4(x).$$

Рассчитаем знаменатели коэффициентов $p_i(x)$ (рис. 46). Тогда многочлены $p_i(x)$ будут иметь вид:

$$p_0(x) = \frac{(x-188{,}1)(x-188{,}2)(x-188{,}3)(x-188{,}4)}{0{,}002};$$

$$p_1(x) = \frac{(x-188)(x-188{,}2)(x-188{,}3)(x-188{,}4)}{(-0{,}001)};$$

$$p_2(x) = \frac{(x-188)(x-188{,}1)(x-188{,}3)(x-188{,}4)}{0{,}0004};$$

$$p_3(x) = \frac{(x-188)(x-188{,}1)(x-188{,}2)(x-188{,}4)}{(-0{,}001)};$$

$$p_4(x) = \frac{(x-188)(x-188{,}1)(x-188{,}2)(x-188{,}3)}{0{,}002}.$$

D19 =ПРОИЗВЕД(D14:D18)

	A	B	C	D	E	F
13	*i*	*для* p_0	*для* p_1	*для* p_2	*для* p_3	*для* p_4
14	*0*		0,10	0,20	0,30	0,40
15	*1*	-0,1		0,1	0,2	0,30
16	*2*	-0,2	-0,10		0,10	0,20
17	*3*	-0,3	-0,20	-0,10		0,10
18	*4*	-0,4	-0,30	-0,20	-0,10	
19	***П***	0,002	-0,001	0,000	-0,001	0,002

Рис. 46. Расчет знаменателей $p_i(x)$

Далее построим график интерполяционного многочлена Лагранжа. Для этого создадим таблицу: в первой колонке – значения x на интервале [188; 188,4] с шагом h = 0,02.

Для каждого значения x рассчитаем $p_0(x)$, $p_1(x)$, $p_2(x)$, $p_3(x)$ и $p_4(x)$ (рис. 47). Например, p_1(0,85) в ячейке С37 рассчитывается по формуле = ((A26 – $B$21)*(A26 – $D$21)*(A26 – $E$21)*(A26 – $F$21))/ $C$19,

C26 =((A26-$B$21)*(A26-$D$21)*(A26-$E$21)*(A26-$F$21))/$C$19

	A	B	C	D	E	F	G	H
23	***h*=**	**0,02**						
24		p_0	p_1	p_2	p_3	p_4	***P4(x)***	
25	188	1,00	0,00	0,00	0,00	0,00	62,10	188
26	188,02	0,64	0,64	-0,43	0,18	-0,03	63,82	
27	188,04	0,37	1,00	-0,56	0,23	-0,04	64,66	
28	188,06	0,19	1,14	-0,49	0,19	-0,03	64,83	
29	188,08	0,07	1,13	-0,28	0,10	-0,02	64,55	
30	188,1	0,00	1,00	0,00	0,00	0,00	64,00	188,1
31	188,12	-0,03	0,81	0,30	-0,09	0,01	63,33	
32	188,14	-0,04	0,58	0,58	-0,15	0,02	62,65	

Рис. 47. Расчет $p_i(x)$ для интервала [188; 188,4]

Теперь мы можем рассчитать значения интерполяционного многочлена $P_4(x)$. Шаг выбираем самостоятельно, исходя из необходимости и достаточности количества точек для построения графика. При этом убедимся, что полученный интерполяционный многочлен Лагранжа $P_4(x)$ обладает необходимым свойством: в точках x_0, x_1, …, x_n, а именно 188; 188,1; 188,2; 188,3; 188,4, его значения совпадают со значениями y_0, y_1, …, y_n, а именно 62,1; 64; 61,5; 63,5; 62,1, функции $f(x)$, заданной таблично (рис. 48).

G26 =$B$22*B26+$C$22*C26+$D$22*D26+$E$22

	A	B	C	D	E	F	G	H	I
23	*h*=	0,02							
24		p_0	p_1	p_2	p_3	p_4	$P_4(x)$		
25	188	1,00	0,00	0,00	0,00	0,00	62,10	188	62,1
26	188,02	0,64	0,64	-0,43	0,18	-0,03	63,82		
27	188,04	0,37	1,00	-0,56	0,23	-0,04	64,66		
28	188,06	0,19	1,14	-0,49	0,19	-0,03	64,83		
29	188,08	0,07	1,13	-0,28	0,10	-0,02	64,55		
30	188,1	0,00	1,00	0,00	0,00	0,00	64,00	188,1	64
31	188,12	-0,03	0,81	0,30	-0,09	0,01	63,33		
32	188,14	-0,04	0,58	0,58	-0,15	0,02	62,65		
33	188,16	-0,03	0,36	0,81	-0,15	0,02	62,08		
34	188,18	-0,02	0,16	0,95	-0,11	0,01	61,68		
35	188,2	0,00	0,00	1,00	0,00	0,00	61,50	188,2	61,5
36	188,22	0,01	-0,11	0,95	0,16	-0,02	61,55		
37	188,24	0,02	-0,15	0,81	0,36	-0,03	61,83		
38	188,26	0,02	-0,15	0,58	0,58	-0,04	62,29		
39	188,28	0,01	-0,09	0,30	0,81	-0,03	62,88		
40	188,3	0,00	0,00	0,00	1,00	0,00	63,50	188,3	63,5
41	188,32	-0,02	0,10	-0,28	1,13	0,07	64,04		
42	188,34	-0,03	0,19	-0,49	1,14	0,19	64,35		
43	188,36	-0,04	0,23	-0,56	1,00	0,37	64,27		
44	188,38	-0,03	0,18	-0,43	0,64	0,64	63,60		
45	188,4	0,00	0,00	0,00	0,00	1,00	62,10	188,4	62,1

Рис. 48. Расчет $P_4(x)$ для интервала [188; 188,4]

Построим график полученного интерполяционного многочлена Лагранжа: $P_4(x) = 62{,}1p_0(x) + 64p_1(x) + 61{,}5p_2(x) + 63{,}5p_3(x) + 62{,}1p_4(x)$ (рис. 49).

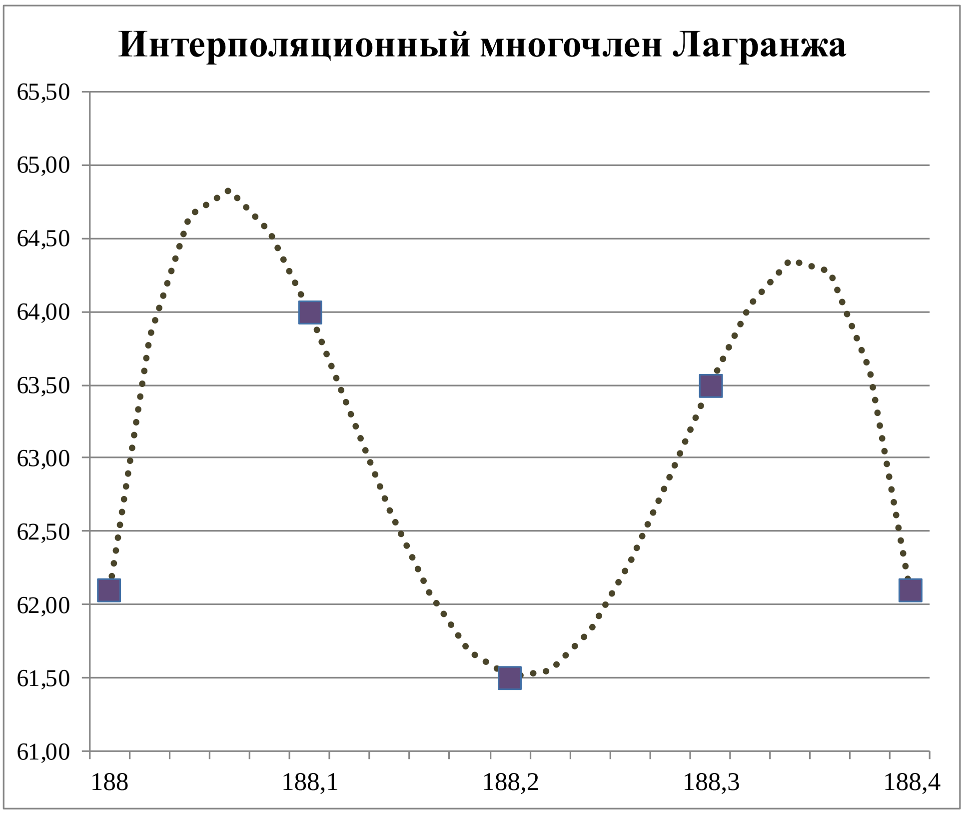


Рис. 49. График $P_4(x)=62{,}1p_0(x)+64p_1(x)+61{,}5p_2(x)+63{,}5p_3(x)+62{,}1p_4(x)$
на интервале [188; 188,4]

Вопросы и задания для самопроверки

1. Что такое интерполяция? Какая функция называется интерполирующей?

2. В чем разница между глобальной и локальной интерполяцией?

3. Что значит «интерполяционный многочлен Лагранжа»? Какое его свойство вы можете указать?

4. В чем заключается отличие между интерполяцией и аппроксимацией?

Раздел 8. Статистическая обработка экспериментальных данных

8.1. Выборка. Точечные оценки результатов измерений

Большое значение в нефтегазовой отрасли имеют экспериментальные исследования. Все полученные результаты наблюдений называются генеральной совокупностью. Все элементы генеральной совокупности характеризуются рядом признаков и могут быть объединены в выборку в зависимости от значений какого-либо признака. Уточним, что и вся генеральная совокупность значений может быть принята в качестве выборки. Оценки истинного значения величины могут быть получены с помощью аппарата математической статистики, когда ее значение возможно получить только в процессе многократного повторения эксперимента.

Пакет Excel богато оснащен набором средств статистической обработки данных. В него включены наиболее часто используемые статистические процедуры: средства описательной статистики, вычисления критериев различия, корреляционные и другие методы, позволяющие проводить необходимый статистический анализ различных типов данных.

В Мастере функции Excel (категория статистические) имеется ряд специальных функций, предназначенных для вычисления выборочных характеристик: медиана, дисперсия, дисперсия генеральной совокупности, математическое ожидание случайной величины и другие.

Пример 17. Пусть дана выборка из 10 экспериментальных значений загрязненности воздуха (концентрации угарного газа), полученных в течение месяца. Используя мастер функций, рассчитаем статистические характеристики для этой выборки (рис. 50).

При некоторых условиях расчет статистических характеристик с помощью таблиц становится громоздким (это происходит когда число исходных данных превышает 50), тогда принято применять компактный метод расчета, включающий в себя предварительную группировку данных. Для этого весь диапазон исходных данных (от минимального до максимального) разбивается на равные интервалы. Удобнее работать с округленными границами, поэтому их часто округляют. Для выборочного метода число исходных данных влияет на количество классов. Выборочный метод используется, когда нет необходимости или возможности исследовать всю совокупность данных. Тогда исследуют выборку и полученные результаты обобщают для всего собрания объектов.

E3 f_x =СРЗНАЧ(B2:B10)

	A	B	C	D	E
1	День месяца	С (мг/куб.м)			
2	1	3		Показатель	Значение
3	4	3,7		среднее значение	3,67
4	9	3		среднее гармоническое	3,55121
5	14	3,2		среднее геометричес.	3,60653
6	15	3,6		медиана	3,60
7	19	3,9		мода	3,0
8	24	3		дисперсия	0,5325
9	28	4,6		стандарт. отклонение	0,72973
10	30	5			
11					

Рис. 50. Результат расчета статистических характеристик выборки

Выборочной (эмпирической) функцией распределения случайной величины X, построенной по выборке $x_1, x_2, \ldots, x_n$, называется функция $F(x)$, равная доле значений x_i, удовлетворяющих условию $x_i < x$, $i = 1, 2, \ldots, n$. Для того чтобы построить выборочную функцию распределения, весь диапазон исходных данных X разбивают на ряд интервалов одинаковой ширины. Что касается числа интервалов, то их чаще всего выбирают не менее 5 и не более 15. Затем рассчитывают число значений X, попавших в каждый интервал. Для нахождения относительной частоты попадания случайной величины X в заданные интервалы нужно разделить эти числа на общее количество наблюдений. По найденным относительным частотам строят гистограмму выборочных функций распределения. Соединив ломаной линией соответствующие точки относительных частот, получим диаграмму, которая называется полигоном частот. При размере выборки (интервала) больше 30 (иногда говорят до бесконечности) выборочные функции распределения превращаются в теоретические. Гистограмма превращается в график плотности распределения, а кумулятивная прямая, в которой по оси абсцисс откладываются интервалы, а по оси ординат – число или доли элементов совокупности, имеющих значение, меньшее или равное заданному, – в график функции распределения.

В Excel для построения выборочной функции распределения используются специальная функция ЧАСТОТА и процедура пакета анализа ГИСТОГРАММА.

Пример 18. Построить эмпирическое распределение профиля высот в метрах для следующей выборки:
64, 57, 63, 62, 58, 61, 63, 60, 60, 61, 65, 62, 62, 60, 64, 61, 59, 59, 63, 61, 62, 58, 58, 63, 61, 59, 62, 60, 60, 58, 61, 60, 63, 63, 58, 60, 59, 60, 59, 61, 62, 62, 63, 57, 61, 58, 60, 64, 60, 59, 61, 64, 62, 59, 65.

В диапазон A2:E12 вводим значения профиля высот. Ширину интервала выбираем равной 1 м. Используя функцию НАИМЕНЬШИЙ, найдем минимальное значение выборки – 57 м, с помощью функции НАИБОЛЬШИЙ – максимальное значение выборки – 65 м. В результате получаем 9 интервалов. В диапазон G2:G10 введем значения интервалов.

Заполним столбец абсолютных частот следующим образом. Выделим диапазон H2:H10. На панели инструментов *Стандартная* нажимаем кнопку f_x, выбираем категорию *Статистические*, находим функцию ЧАСТОТА, нажимаем ОК. Появляется окно мастера функций. Мышкой выделяем ячейки A2:E12 в качестве Массива_данных, после чего переходим мышкой в строку формулы, нажимаем клавишу F2. В качестве Массива_интервалов выделяем ячейки G2:G10. Переходим в строку формулы, нажимаем F2, а затем – сочетание клавиш Ctrl + Shift + Enter. В результате в столбце H2:H10 появляется массив абсолютных частот.

В ячейке H11 найдем общее количество наблюдений. Для этого установим курсор на ячейку H11, на панели инструментов нажимаем кнопку ∑, выбираем диапазон H2:H10, нажимаем Enter.

Теперь рассчитаем относительные частоты. Результаты расположим в ячейках I2:I10. Сначала в ячейке I2 введем формулу H2/$H$11, а затем скопируем ее в остальные ячейки диапазона. Получаем массив относительных частот (рис. 51).

H3 | {=ЧАСТОТА(A2:E12;G2:G10)}

	A	B	C	D	E	F	G	H	I	J
1	Наблюдения						Высота, м	Абсолютные частоты	Относительные частоты	Накопленные частоты
2	64	62	58	63	61		57	2	0,036	0,036
3	57	62	63	58	58		58	6	0,109	0,145
4	63	60	61	60	60		59	7	0,127	0,273
5	62	64	59	59	64		60	10	0,182	0,455
6	58	61	62	60	60		61	9	0,164	0,618
7	61	59	60	59	59		62	8	0,145	0,764
8	63	59	60	61	61		63	7	0,127	0,891
9	60	63	58	62	64		64	4	0,073	0,964
10	60	61	61	62	62		65	2	0,036	1,0
11	61	62	60	63	59			55		
12	65	58	63	57	65		среднее значение		**60,8545**	

Рис. 51. Результат обработки выборки

Заполним столбец накопленных частот. Для этого в ячейку J2 скопируем значение I2, в ячейку J3 введем формулу = J2 + I3 и скопируем ее в остальные ячейки диапазона. Получаем массив накопленных частот. Результат вычислений представлен на рис. 51.

Построим диаграмму относительных и накопленных частот. Сложно поверить, но в Excel 2010 нет специальной гистограммы с графиком с накоплением. Поэтому будем делать следующим образом. Обращаемся к *Мастеру диаграмм*. Выделив диапазон ячеек I2:J10, в главном меню выбираем *Вставка* → *Гистограмма* → *Гистограмма с накоплением*. На экране появляется график в первоначальном варианте. В нем по горизонтальной оси расположены деления 1, 2 и т. д. В подменю выбираем *Выбрать данные*, после чего на экране появится окно *Выбор источника данных*. В поле *Подписи горизонтальной оси* нажимаем *Изменить* и мышкой выделяем диапазон подписей оси – ячейки G2–G10, нажимаем OK и в следующем окне снова OK. На графике по горизонтальной оси появились значения профиля высот от 57 до 65 м.

Далее мышкой щелкаем в диаграмме по одному из столбцов, отражающих накопленные частоты, помеченным становится весь ряд накопленных частот. Нажимаем правую клавишу мышки и в появившемся окне выбираем *Изменить тип диаграммы для ряда* (рис. 52). На экране появляется основное меню выбора (изменения) типа диаграммы. Щелкаем мышкой по иконке *График*, нажимаем OK.

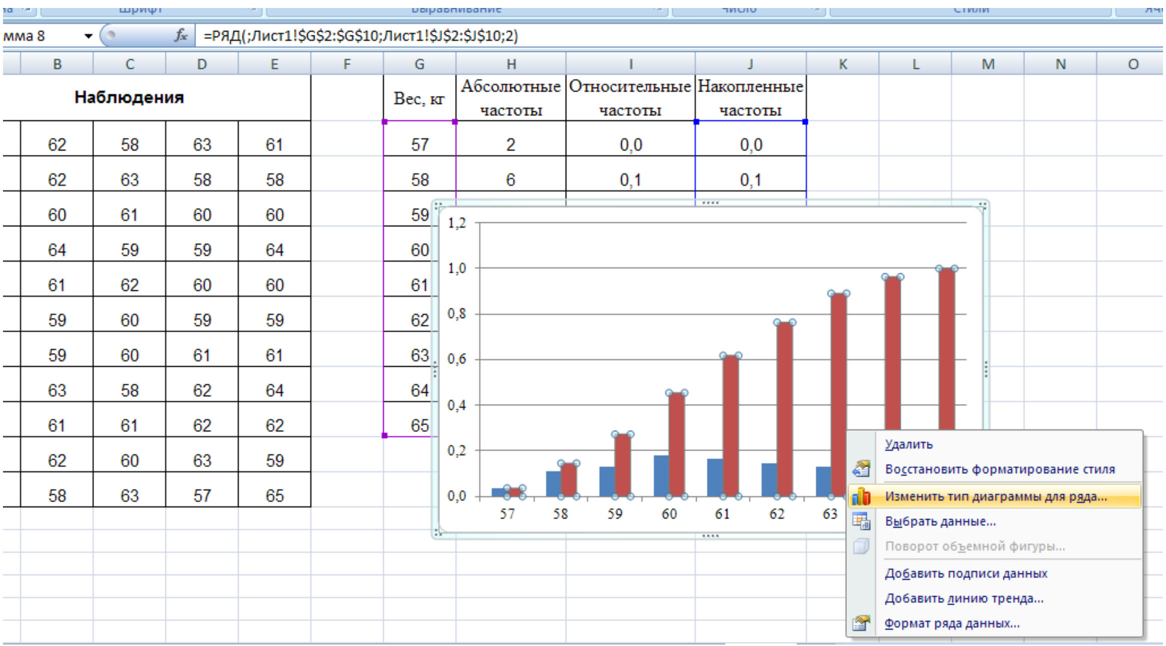


Рис. 52. Меню изменения типа диаграммы для ряда

График на экране принимает вид, где масштабы относительных и накопленных частот отличаются почти в полтора раза. Для наглядности таких диаграмм есть возможность сделать график с двумя вертикальными осями, то есть для гистограммы относительных частот своя шкала, для графика накопленных частот – своя. Для этого выделяем в диаграмме линейный график, щелкая мышкой по нему, затем нажимаем правую клавишу мышки и в появившемся окне *Формат ряда данных* в поле параметры ряда выбираем пункт *по вспомогательной оси*. На экране появляется график с двумя осями. Теперь нужно привести их в соответствие требованиям. Кликаем мышкой на численные значения оси – на экране появляется окно *Формат оси*. Подбираем все необходимые параметры для каждой из осей, редактируем. В результате получаем диаграмму относительных и накопленных частот на одном графике (рис. 53).

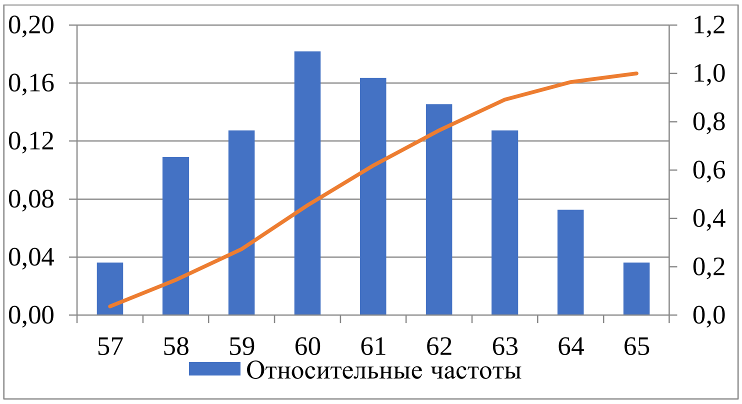


Рис. 53. Диаграмма относительных и накопленных частот

8.2. Погрешности

Все вычисления, проводимые при решении какой-либо задачи, страдают приближенностью. Это происходит в связи с тем, что используемые в этих вычислениях величины несут в себе неточность измерения, определяемую единицей измерения прибора, а также потому, что в процессе вычисления могут производиться округления величин, приводящие также к накоплению погрешности вычисляемых величин. Это надо обязательно учитывать, чтобы всегда иметь представление о величине ошибки полученного результата.

Можно выделить два типа погрешностей. Первый тип появляется в процессе измерений и относится к техническим погрешностям, или погрешностям измерений. Такие погрешности подразделяются на случайные и систематические. Случайные погрешности присутствуют во всех измерениях, они неустранимы и их стараются снизить до разумных пределов путем соответствующей организации работ. Систематические погрешности возникают в результате неправильной методики или технологии измерений, их значения направлены в одну сторону (завышения или занижения результатов измерений). Если такие погрешности появляются, необходимо устранять их либо изменением методики и технологии измерений, либо введением поправок.

Второй тип погрешностей возникает при распространении выборочных данных на генеральную совокупность. Их называют погрешностями аналогии, или погрешностями распространения. Значения их зависят от способа или методики распространения данных выборки на генеральную совокупность, поэтому иногда такие ошибки называют методическими погрешностями.

Для относительных случайных погрешностей существуют допустимые значения, которые приводят в инструкциях по подсчету запасов для каждого вида минерального сырья. Если относительная случайная погрешность окажется больше допустимой, то подсчет запасов будет ненадежным.

Различают два вида погрешностей – абсолютную и относительную. Абсолютная погрешность некоторой величины равна разности между ее истинным значением и приближенным значением, полученным в результате измерения или вычисления. Относительная погрешность – это отношение абсолютной погрешности к модулю приближенного значения величины.

Таким образом, если a – приближенное значение величины x, то выражения для абсолютной и относительной погрешностей запишутся соответственно в виде: $\Delta x = x - a$, $\delta x = \Delta x/|a|$.

Пример 19. Пусть необходимо вычислить значение величины y, а также оценить погрешность этого вычисления. При этом известны оценки абсолютных погрешностей используемых в вычислении величин.

Дано: $y = \sqrt{\frac{a+b}{x^2(1-x)}}$;

$a = 25{,}2$; $\Delta a = 0{,}1$;

$b = 10{,}24$; $\Delta b = 0{,}02$;

$x = 0{,}5$; $\Delta x = 0{,}1$.

Проделаем следующее: разместим в ячейках A1, A2, A3 обозначения ‘a =’, ‘b =’, ‘x =’, а в ячейках B1, B2, B3 их значения соответственно. Аналогично заполним ячейки C1:D3 (рис. 54).

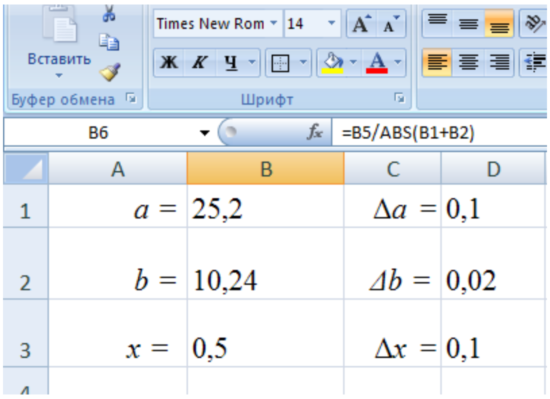


Рис. 54. Исходные данные

Проведем расчеты, используя следующий алгоритм:

$\Delta(a + b) = 0{,}1 + 0{,}02 = 0{,}12;$

$\delta(a + b) = \Delta(a + b)/\ |a + b| = 0{,}12/35{,}44 = 0{,}003386;$

$\delta(x^2) = 2\delta x = 2\ \Delta x/\ |x| = 2\cdot 0{,}1/0{,}5 = 0{,}4;$

$\Delta(1 - x) = \Delta(1) + \Delta(x) = 0 + 0{,}1 = 0{,}1;$

$\delta(1 - x) = \Delta(1 - x)\ /\ |1 - x| = 0{,}1/0{,}5 = 0{,}2;$

$\delta(\sqrt{y}) = \delta(a + b) + \delta(x^2) + \delta(1 - x) = 0{,}003386 + 0{,}4 + 0{,}2 = 0{,}603386;$

$\delta(y) = \frac{1}{2}\delta(\sqrt{y}) = \frac{1}{2}\cdot 0{,}003386 = 0{,}301693;$

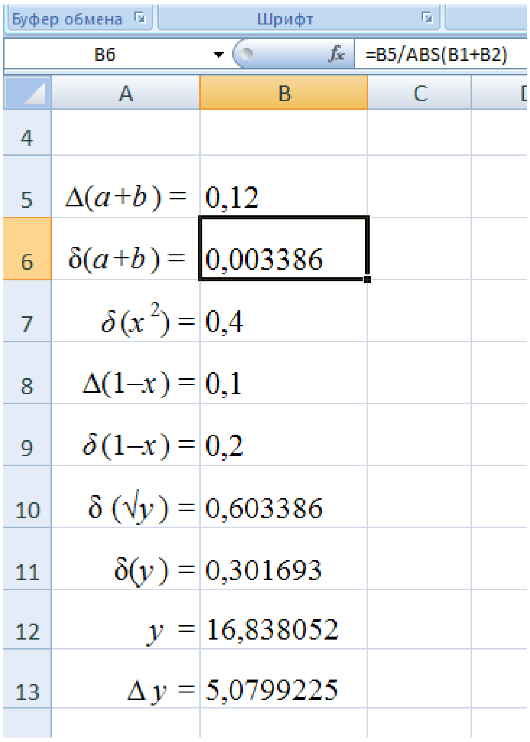
$y = \sqrt{\frac{25{,}2+10{,}24}{0{,}5^2(1-0{,}5)}} = \sqrt{283{,}52} = 16{,}838052;$

$\Delta\, y = \delta\ (y)\cdot|y| = 0{,}301693\cdot 16{,}838052 = 5{,}0799224.$

Так, например, для вычисления относительной погрешности числителя $\delta(a + b)$ в ячейке В6 набираем =, ссылка на В5 (абсолютная погрешность), знак деления (/). Затем обращаемся к мастеру функций, находим функцию ABS, вводим аргументы функции, получаем в строке ввода =B5/ABS(B1+B2), нажимаем Enter. В текущей ячейке появится результат 0,003386.

Для вычисления функции извлечения квадратного корня есть два варианта: использовать функцию КОРЕНЬ(…) или функцию СТЕПЕНЬ (…). Так, например, вычислить значение функции *y* можно или =КОРЕНЬ((B1+B2)/(B3^2*(1-B3)) или =СТЕПЕНЬ ((B1+B2)/(B3^2*(1-B3));1/2). Для вычисления корней с другими показателями используется функция СТЕПЕНЬ (число, степень).

Результаты расчетов представлены на рис. 55.


Буфер обмена | Шрифт

B6 | =B5/ABS(B1+B2)

	A	B	C
4			
5	$\Delta(a+b) =$	0,12	
6	$\delta(a+b) =$	0,003386	
7	$\delta(x^2) =$	0,4	
8	$\Delta(1-x) =$	0,1	
9	$\delta(1-x) =$	0,2	
10	$\delta(\sqrt{y}) =$	0,603386	
11	$\delta(y) =$	0,301693	
12	$y =$	16,838052	
13	$\Delta y =$	5,0799225	

Рис. 55. Результаты расчетов в Excel для примера 19

8.3. Законы распределения случайных величин

При проверке соответствия полученных экспериментальным путем данных теоретическому распределению при обработке статистическими методами часто требуется знание априори определенного закона распределения. В то же время в большинстве реальных случаев при решении задач закон распределения эмпирических данных и его параметры неизвестны. Что же делать?

Закон распределения случайной величины выражается в виде интеграла вероятности: $F(x) = \int\limits_{-\infty}^{x} f(x)dx$, где $F(x)$ – вероятность p того, что значение случайной величины не превысит значения x, т. е. $p = F(x)$; функция под интегралом $f(x)$ – плотность вероятности случайной величины; к кривой, описываемой функцией $f(x)$, стремится гистограмма частостей при увеличении числа наблюдений.

Из рис. 56 видно, что интеграл вероятности $F(x)$ при увеличении значения x монотонно растет от нуля до единицы. Его можно рассматривать как площадь (заштрихована на рис. 56, *б*), ограниченную осью абсцисс, кривой $f(x)$ и отрезком перпендикуляра, проведенного из точки a. Вся площадь под кривой $f(x)$ равна единице, поэтому заштрихованная площадь меньше единицы и соответствует вероятности p.

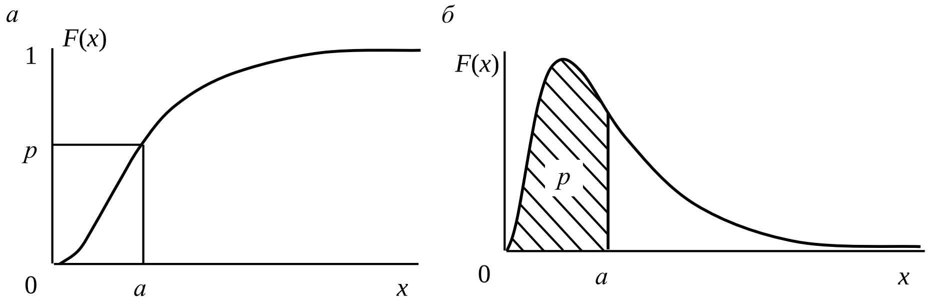


Рис. 56. Графики интеграла вероятности (а) и плотности вероятности (б)

Можно разбить законы распределения случайных величин на две большие группы: дискретные и непрерывные. Случайная величина принимает лишь определенные значения (например, число зерен минералов в пробе), а график плотности вероятности имеет ступенчатый вид для дискретных законов. Это биномиальный, Пуассона, и гипергеометрический законы. Случайная величина для законов с непрерывным распределением может принимать любые значения в области своего существования (например, содержание компонента в руде), а графики плотности вероятности непрерывного распределения имеют плавный вид. К таким законам относятся: нормальный, логнормальный, Стьюдента, χ^2, Фишера и некоторые другие.

Рассмотрим некоторые из часто употребляемых на практике законов распределения.

Чаще других используют ***нормальный***, потому что он носит предельный характер и при определенных условиях к нему приближаются многие другие законы. Нормальный закон описывается интегралом вероятности $F(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{x} \mathrm{e}^{-\frac{(x-\bar{x})^2}{2\sigma^2}} dx$, плотность вероятности имеет вид

$f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$ (рис. 57). Каковы же особенности этого закона и его показатели. Во-первых, его кривая имеет симметричную форму относительно абсциссы, а площадь между кривой и осью абсцисс равна единице. Ветви кривой не ограничены и уходят в плюс и минус бесконечность, сливаясь в удалении с осью абсцисс.

Нормальный закон полностью определяется двумя статистическими характеристиками: средним значением $\bar{x}$, которое определяет положение графика на оси абсцисс, и дисперсией σ^2, которая определяет крутизну ветвей. Кривая плотности вероятности симметричная, асимметрия и эксцесс равны нулю.

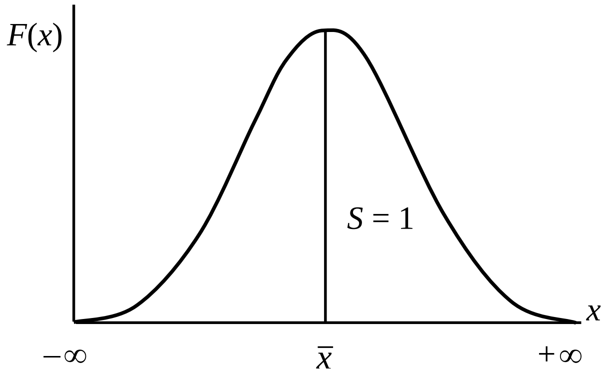


Рис. 57. График плотности вероятности нормального закона

В геохимии широко применяется ***логарифмически-нормальный*** (логнормальный) закон распределения, так как этим законом описывается частота появления низких содержаний химических элементов. Область существования случайной величина в логнормальном законе – от нуля до +∞. Если присутствуют нулевые значения (или следы), что нередко бывает при спектральном и химическом анализах, то это вызывает трудности, так как логарифм нуля равен –∞. Обычно нулевые содержания заменяют какими-то минимальными значениями, например пределом чувствительности. Плотность вероятности логарифмов описывается формулой нормального закона: $f(z) = \frac{1}{\sigma_z\sqrt{2\pi}} e^{-\frac{(z-\bar{z})^2}{2\sigma_z^2}}$, (рис. 58), где $\bar{z}$ – среднее значение логарифмов; σ – среднеквадратичное отклонение логарифмов.

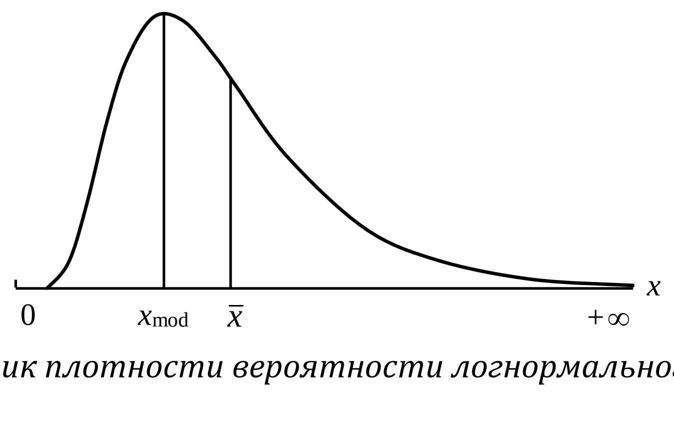


Рис. 58. График плотности вероятности логнормального закона

Для оценки соответствия имеющихся экспериментальных данных нормальному закону распределения используют графический метод, выборочные параметры формы распределения и критерии согласия.

Критериями согласия – это статистические критерии, предназначенные для проверки согласия экспериментальных и рассчитанных по теоретической модели данных. Нулевая гипотеза – это утверждение о том, что распределение генеральной совокупности, из которой получена выборка, не отличается от нормального. Наиболее распространен критерий хи-квадрат (χ^2), который основан на сравнении эмпирических частот интервалов группировки с теоретическими (ожидаемыми) частотами, рассчитанными по формулам нормального распределения. Данный критерий применяется для объема выборки больше 40 ($n \geq 40$), выборочные данные группируются в интервальный ряд с числом интервалов не менее 7, а в каждом интервале не должно находится менее 5 наблюдений (частот).

Распределение χ^2 служит преимущественно для проверки гипотез о соответствии наблюдаемых частот теоретическим законам распределения. Плотность вероятности распределения описывается формулой $f(\chi^2) = f(x) = \dfrac{1}{2^{k/2}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} \mathrm{e}^{-\frac{x}{2}}$ (рис. 59), где k – число степеней свободы, зависящее от числа классов гистограммы n_k (обычно $k = n_k - 3$).

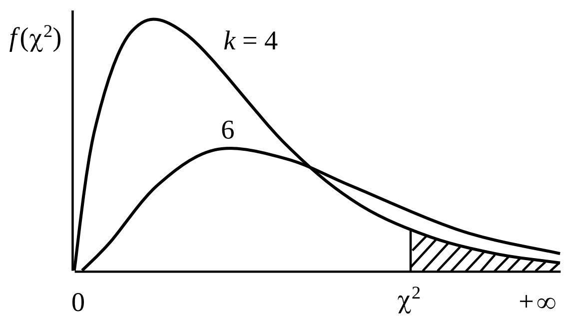


Рис. 59. График распределения χ^2

Видим, что случайная величина χ^2 имеет область существования от нуля до $+\infty$. График плотности вероятности асимметричен, модальное значение $\chi^2_{\mathrm{mod}} = k - 2$. При увеличении числа степеней свободы распределение χ^2 приближается к нормальному с математическим ожиданием k и дисперсией $2k$. Как только число степеней свободы превышает тридцать ($k > 30$), можно переходить к таблицам нормального распределения, заменив величину χ^2 нормированной случайной величиной t: $t = (\chi^2 - k)/\sqrt{2k}$.

Для практических целей требуется иметь таблицу коэффициентов вероятности (табл. 5), играющих роль критериев. В зависимости от вероятности α при заданной степени свободы k критерию χ^2 соответствует заштрихованная площадь на рис. 60.

В Excel критерий χ^2 реализован в функции ХИ2ТЕСТ, которая вычисляет вероятность совпадения наблюдаемых (фактических) значений и теоретических (гипотетических) значений. Если вычисленная вероятность ниже уровня значимости (0,05), то нулевая гипотеза отвергается и утверждается, что наблюдаемые значения не соответствуют нормальному закону распределения. Если вычисленная вероятность близка к 1, то можно говорить о высокой степени соответствия экспериментальных данных нормальному закону распределения.

Параметры функции ХИ2ТЕСТ («фактический_интервал, ожидаемый_интервал»): «фактический_интервал» – интервал данных, которые содержат наблюдения, подлежащие сравнению с ожидаемыми значениями; «ожидаемый_интервал» – интервал данных, которые содержат теоретические (ожидаемые) значения для соответствующих наблюдений.

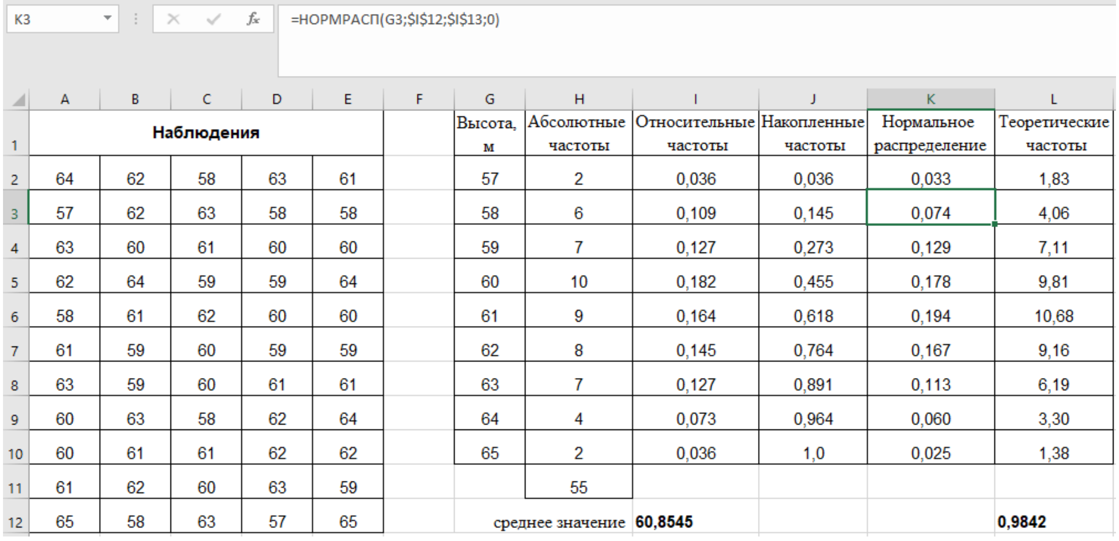
Пример 20. Для выборки распределения профиля высот (в метрах): 64, 57, 63, 62, 58, 61, 63, 60, 60, 61, 65, 62, 62, 60, 64, 61, 59, 59, 63, 61, 62, 58, 58, 63, 61, 59, 62, 60, 60, 58, 61, 60, 63, 63, 58, 60, 59, 60, 59, 61, 62, 62, 63, 57, 61, 58, 60, 64, 60, 59, 61, 64, 62, 59, 65.
проверить соответствие выборочных данных нормальному закону распределения.

Вычисления будут проходить в два этапа:

1. Найдем теоретические частоты нормального распределения, предварительно определив среднее значение и стандартное отклонение выборки. Для этого в ячейке I12 с помощью функции СРЗНАЧ находим среднее значение для диапазона A2:E12. Затем в ячейке I13 с помощью функции СТАНДОТКЛОН находим стандартное отклонение для этих данных. Теоретические частоты определяем с помощью функции НОРМРАСП: устанавливаем курсор в ячейку K2, вызываем указанную функцию, заполняем ее параметры следующим образом: x – ячейка G2; среднее – $I$12; стандартное отклонение – $I$13; интегральный – 0. Нажимаем ОК. Далее копируем формулу в ячейки K3– K10. После чего заполняем столбец L, введя формулу =K2*$H$11 в ячейку L2 и растянув ее в ячейки L3–L10. Результаты вычислений показаны на рис. 60.

2. С помощью функции ХИ2ТЕСТ определим соответствие данных нормальному закону распределения. Для этого устанавливаем курсор в ячейку L12, на панели инструментов *Стандартная* нажимаем кнопку f_x, в появившемся окне выбираем категорию *Статистические* и находим функцию ХИ2ТЕСТ(), нажимаем ОК, вводим требуемые параметры (рис. 61), нажимаем ОК. В ячейке L12 появляется значение 0,9842. Это

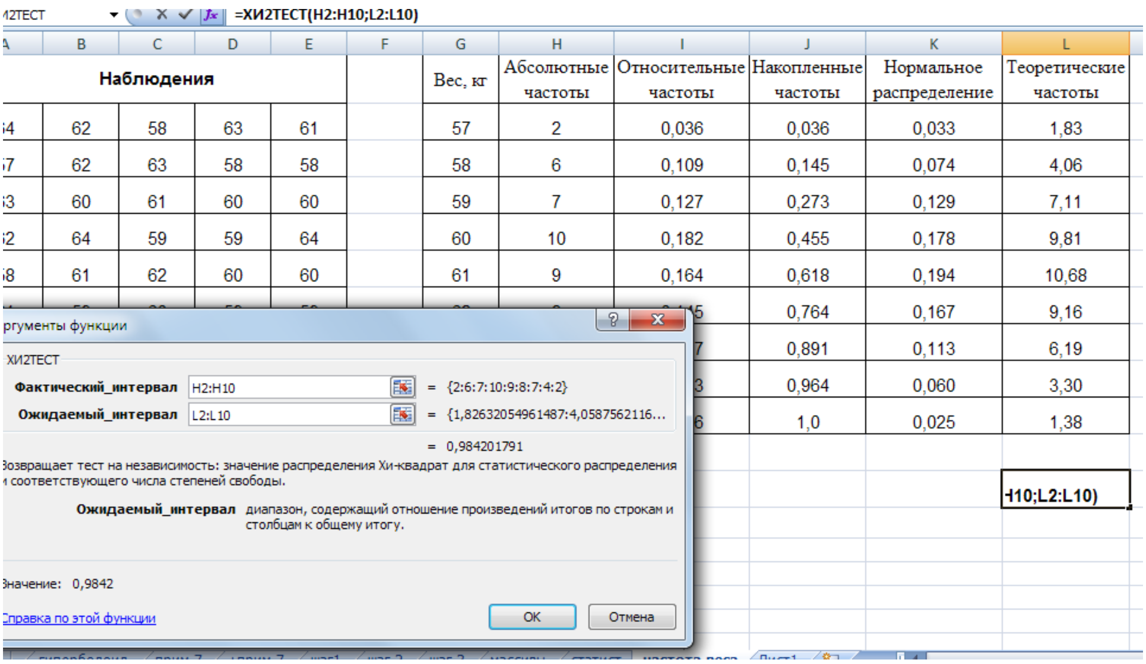
значит, что вероятность соответствия выборочных данных нормальному закону распределения составляет 98,42 %.



К3 =НОРМРАСП(G3;$I$12;$I$13;0)

	A	B	C	D	E	F	G	H	I	J	K	L
1	Наблюдения						Высота, м	Абсолютные частоты	Относительные частоты	Накопленные частоты	Нормальное распределение	Теоретические частоты
2	64	62	58	63	61		57	2	0,036	0,036	0,033	1,83
3	57	62	63	58	58		58	6	0,109	0,145	0,074	4,06
4	63	60	61	60	60		59	7	0,127	0,273	0,129	7,11
5	62	64	59	59	64		60	10	0,182	0,455	0,178	9,81
6	58	61	62	60	60		61	9	0,164	0,618	0,194	10,68
7	61	59	60	59	59		62	8	0,145	0,764	0,167	9,16
8	63	59	60	61	61		63	7	0,127	0,891	0,113	6,19
9	60	63	58	62	64		64	4	0,073	0,964	0,060	3,30
10	60	61	61	62	62		65	2	0,036	1,0	0,025	1,38
11	61	62	60	63	59			55				
12	65	58	63	57	65		среднее значение		**60,8545**			**0,9842**

Рис. 60. Результаты вычислений теоретических частот



=ХИ2ТЕСТ(H2:H10;L2:L10)

B	C	D	E	F	G	H	I	J	K	L
Наблюдения					Вес, кг	Абсолютные частоты	Относительные частоты	Накопленные частоты	Нормальное распределение	Теоретические частоты
62	58	63	61		57	2	0,036	0,036	0,033	1,83
62	63	58	58		58	6	0,109	0,145	0,074	4,06
60	61	60	60		59	7	0,127	0,273	0,129	7,11
64	59	59	64		60	10	0,182	0,455	0,178	9,81
61	62	60	60		61	9	0,164	0,618	0,194	10,68
								0,764	0,167	9,16
								0,891	0,113	6,19
								0,964	0,060	3,30
								1,0	0,025	1,38

Рис. 61. Результаты вычислений χ^2-критерия

8.4. Расчет доверительного интервала

Среднее значение $\bar{x}$ из n независимых значений случайной величины x также является случайной величиной. Обычно среднее значение случайной величины находят по выборке из генеральной совокупности. Математическое ожидание случайной величины в генеральной совокупности $M(x)$ обычно неизвестно. Его можно приближенно оценить с помощью выборочного среднего значения, которое также является случайной величиной. Чаще всего с достаточным основанием предполагается, что случайная величина, как представляющая собою сумму многих случайных величин, имеет распределение, близкое к нормальному. Размах значений нормально распределенной величины составляет приближенно $\pm 3\delta$. Где-то в этом интервале и заключено математическое ожидание $M(x)$. Наиболее вероятно, что оно совпадает со средним значением $\overline{x}$, которое является *точечной оценкой* математического ожидания. Менее вероятно, что математическое ожидание смещено в ту или иную сторону от среднего значения. Интервал возможных значений математического ожидания зависит от вероятности β и выражается через коэффициент вероятности. Данный интервал называется *доверительным интервалом*, или *интервальной оценкой*, математического ожидания. Каждому значению вероятности β соответствует определенный коэффициент вероятности и размер доверительного интервала. То есть для заданной вероятности β требуется определить максимальную погрешность ε_β, которая имеется в эксперименте, если за истинное значение случайной величины принята ее оценка, полученная из опыта, – среднее значение ($X^{ср}$). Полученный интервал называется доверительным, его границы ($X^{ср} \pm \varepsilon_\beta$) – доверительными, вероятности β – доверительной. Чем больше величина доверительной вероятности β, тем шире доверительный интервал (больше величина ε_β), и наоборот, чем больше ширина доверительного интервала, тем с большей вероятностью истинное значение находится в этом интервале. Иногда вместо доверительной вероятности задают уровень значимости $p = 1 - \beta$, который часто выражают в процентах. Уровень значимости показывает вероятность того, что истинное значение измеряемой величины будет лежать за пределами доверительного интервала. На практике величину доверительной вероятности берут равной 0,9; 0,95; 0,99.

Расчет доверительного интервала зависит от объема выборки – n. Если объем выборки больше 30 ($n > 30$), то предполагается, что случайная величина подчиняется нормальному закону распределения и доверительный интервал можно рассчитать с использованием стандарта среднего: $S_{x^{ср}} = S_x / \sqrt{n}$.

В Excel для вычисления границ доверительного интервала и при числе элементов выборки <30 используют функцию ДОВЕРИТ или процедуру *Описательная статистика*.

Функция ДОВЕРИТ(альфа; станд_откл; размер) определяет полуширину доверительного интервала, где альфа – уровень значимости (p), доверительная вероятность равна 100·(1 – альфа) процентам, т. е. альфа = 0,05 означает 95%-й уровень доверительной вероятности; станд_откл – стандартное отклонение; размер – объем выборки (n).

Пример 21. Найти с доверительной вероятностью β = 95 % границы доверительного интервала для времени функционирования скважины, если экспериментально установлено, что у 25 скважин среднее время функционирования составило 140 месяцев, а стандартное отклонение – 2,5 месяца.

Устанавливаем курсор в ячейку A1, на панели инструментов *Стандартная* нажимаем кнопку f_x, в появившемся окне выбираем категорию *Статистические* и находим функцию ДОВЕРИТ, нажимаем ОК, вводим требуемые параметры (рис. 62), нажимаем ОК.

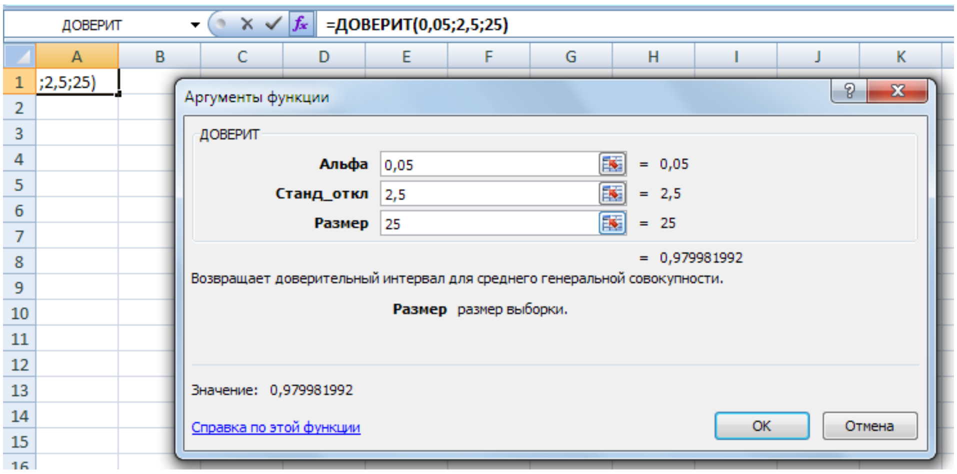


Рис. 62. Окно функции ДОВЕРИТ

В ячейке A1 появляется 0,97998 – значение полуширины 95%-го доверительного интервала для времени функционирования скважины. То есть с 95%-м уровнем надежности можно утверждать, что средняя продолжительность функционирования скважины составляет 140 ± 0,97998 месяца или от 139,02 до 140,98 месяца.

8.5. Отбраковка грубых измерений

Иногда при проведении многократных измерений одной величины оказывается, что одно или несколько значений существенно отличаются от остальных. Такие измерения, как правило, содержат грубую ошибку и называются *грубыми*, или *выбросами*. Перед построением доверительного интервала при заданной доверительной вероятности β необходимо серию измерений проверить на наличие выбросов. Проверка на выбросы проводится по U-критерию, рассчитывается по формуле

$U^p = \frac{|X^{\text{под}} - X^{\text{ср}}|}{\sqrt{S_{x\,n}^{2\,n-1}}}$, где $X^{\text{под}}$ – измерение, подозреваемое на наличие грубой ошибки.

Сравниваем U^p с табличным значением $U_{p,f}$ (табл. 5), которое определяется уровнем значимости p и количеством степеней свободы $f = n - 2$. Если $U^p > U_{p,f}$, то значение $X^{\text{под}}$ содержит грубую ошибку и это измерение необходимо исключить из серии. После исключения всех измерений, содержащих грубые ошибки, можно переходить к построению доверительного интервала.

Таблица 5

Зависимость параметра $U_{p,f}$ от количества измерений (n) и вероятности (p)

n	***Up,f при p***			***n***	***Up,f при p***		
	0,9	***0,95***	***0,99***		***0,9***	***0,95***	***0,99***
3	1,41	1,41	1,41	**15**	2,33	2,49	2,8
4	1,64	1,69	1,72	**16**	2,35	2,52	2,84
5	1,79	1,87	1,96	**17**	2,38	2,55	2,87
6	1,89	2	2,13	**18**	2,4	2,58	2,9
7	1,97	2,09	2,26	**19**	2,43	2,6	2,93
8	2,04	2,17	2,37	**20**	2,45	2,62	2,96
9	2,1	2,24	2,46	**25**	2,54	2,72	3,07
10	2,15	2,29	2,54	**30**	2,61	2,79	3,16
11	2,19	2,34	2,61	**35**	2,67	2,85	3,22
12	2,23	2,39	2,66	**40**	2,72	2,9	3,27
13	2,26	2,43	2,71	**45**	2,76	2,95	3,32
14	2,3	2,46	2,76	**50**	2,8	2,99	3,37

Пример 22. Пусть имеется выборка объемом 4 измерения: $x_1 = 4{,}13$; $x_2 = 4{,}65$; $x_3 = 4{,}62$; $x_4 = 4{,}54$. Уровень значимости $\alpha = 0{,}05$. Провести отбраковку измерений, т. е. определить измерения, содержащие грубую ошибку и исключить их из серии.

Для выполнения задания сделаем следующие шаги:

1. Введем исходные данные, рассчитаем среднее значение выборки с помощью функции СРЗНАЧ, $x_{\text{ср}} = 4{,}485$, рассчитаем дисперсию случайной величины с помощью функции ДИСП ($S_x^2 = 0{,}058$).

2. Исследуем выборку на наличие грубых ошибок. Для этого рассчитаем значение критерия U^p для всех измерений, подозреваемых на ее

наличие, сравним с табличным значением $U_{p,f}$ и, используя функции ЕСЛИ, заполним столбец E, исключив из выборки грубые ошибки (рис. 63).

3. После исключения из выборки грубых ошибок объем выборки уменьшается. Тогда можно переходить к построению доверительного интервала.

Сравним результаты расчетов до и после отбраковки (рис. 64).

=ABS(C6-$C$8)/КОРЕНЬ($C$9*($A$7-1)/$A$7)

	A	B	C	D	E
1			$U_{p,f}$ =	1,69	
2					
3	№		Значения	U_p	
4	1		4,13	1,700	0,00
5	2		4,65	0,790	4,65
6	3		4,62	0,646	4,62
7	4		4,54	0,263	4,54
8	среднее значение		4,485		
9	Дисперсия выборки		0,058		

Рис. 63. Отбраковка измерений

а)

Уровень надежности(95,0%)	0,383767	
Доверительные границы	3,7175	5,2525

б)

Уровень надежности(95,0%)	0,141254	
Доверительные границы	4,3587	4,8480

Рис. 64. Значение доверительного интервала и границ:
а) до исключения грубых ошибок; б) после исключения из выборки грубых ошибок

В особо ответственных случаях контроль проводят несколько раз, чтобы убедиться в обоснованности введения поправок на систематическую погрешность.

Вопросы и задания для самопроверки

1. Сравнить понятия генеральной совокупности и выборки.

2. Что такое выборочная функция? Описать алгоритм и требования для ее построения.

3. Перечислить и кратко охарактеризовать законы распределения случайных величин.

4. Что такое выброс? Для чего проводят отбраковку грубых измерений?

5. Дать определение доверительного интервала. Как он изменяется после отбраковки грубых измерений?

6. Дать определение погрешности. Перечислить виды погрешностей.

Заключение

Научно-технический прогресс в области информатики и вычислительной техники обусловил широкое внедрение компьютеров и современных методов обработки информации во все отрасли народного хозяйства, в том числе в нефтегазовую. Успешное использование математических методов и компьютеров невозможно без повышения уровня математического образования. Предлагаемое учебное пособие помогает читателю получить представление о принципах и особенностях математического моделирования нефтегазовых объектов и явлений, овладеть основными методами математической, в том числе статистической, обработки информации и научиться применять их для решения задач нефтегазовой отрасли. В данном учебном пособии рассмотрены средства визуализации данных (построение диаграмм и графиков) в MS Office Excel. Представленные в нем сведения могут быть применены и при выполнении практических работ в вузе и в дальнейшем, при создании иллюстрационных материалов для отчетов, слайдов презентаций при защите курсовых работ, докладах на конференциях и т. д. В этом пособии изложены основы построения, редактирования и форматирования диаграмм. Дополнительные возможности программы для визуализации данных можно изучить самостоятельно по другим источникам.

Рассмотренные математические методы не исчерпывают всего их многообразия. Приведены лишь наиболее важные и распространенные в нефтегазовых объектах и явлениях, по мнению автора, методы, проверенные на практике и которые могут быть полезны при математической обработке исходных данных. Так, исключены из рассмотрения дисперсионный анализ, частные коэффициенты корреляции, некоторые редкие законы распределения случайных величин.

Новые математические методы, которые появились совсем недавно и находятся в стадии становления, такие как фрактальный анализ, явления самоорганизации вещества и др., в учебное пособие пока не вошли.

Список литературы

1. Борздова Т.В. Основы статистического анализа и обработка данных с применением Microsoft Excel : учебное пособие / Т.В. Борздова. – Минск : ГИУСТ БГУ, 2011. – 75 с.
2. Поротов Г.С. Математические методы моделирования в геологии : учебник / Г.С. Поротов. – Санкт-Петербург : Санкт-Петербургский государственный горный институт (технический университет), 2006. – 223 с.
3. Чен-Син Э.П. Методические указания к лабораторным работам по курсу «Компьютерное моделирование» : учебное пособие / Э.П. Чен-Син, Л.Н. Панюшева. – Москва : РГУ нефти и газа, 2004. – 92 с.
4. Ильченко М.А. Excel в математических и статистических расчетах : методические указания. В 2 частях. Часть 2 / М.А. Ильченко, Л.С. Струкова. – Мичуринск : Наукоград, 2007. – 136 с.
5. Бедарев И.А. Методы вычислений : учебное пособие / И.А. Бедарев, Ю.В. Кратова, Н.Н. Федорова. – 2-е изд., перераб. и доп. – Новосибирск : НГАСУ, 2009. – 116 с.
6. Молоков К.А. Компьютерные технологии в машиностроении : методические указания к выполнению практических работ для студентов / К.А. Молоков, А.В. Славгородская, М.В. Китаев. – Владивосток : Издательский дом Дальневосточного федерального ун-та, 2013. – 40 с.

Оглавление

Учебное издание

ДЕЕВА Вера Степановна

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ В НЕФТЕГАЗОВОМ ДЕЛЕ

Учебное пособие

Корректура *Д.В. Заремба*
Компьютерная верстка *О.Ю. Аршинова*
Дизайн обложки *Т.В. Буланова*

Подписано к печати 29.03.2018. Формат 60×84/16. Бумага «Снегурочка».
Печать CANON. Усл. печ. л. 5,00. Уч.-изд. л. 4,52.
Заказ 65-18. Тираж 100 экз.

Издательство
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ