МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

ЯРОСЛАВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
им. П. Г. ДЕМИДОВА

ПРЕПОДАВАНИЕ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ НАУК В КЛАССИЧЕСКОМ УНИВЕРСИТЕТЕ

Материалы 2-й научно-методической конференции преподавателей математического факультета и факультета информатики и вычислительной техники Ярославского государственного университета им. П. Г. Демидова

Ярославль 2007

ББК Ч 486.24/29я43
П 71
УДК 51.37

Рекомендовано
Редакционно-издательским советом университета
в качестве научного издания. План 2007 года

Преподавание математики и компьютерных наук в классическом университете: материалы 2-й научно-методической конференции преподавателей математического факультета и факультета информатики и вычислительной техники Ярославского государственного университета им. П. Г. Демидова / отв. редактор М. В. Невский; Яросл. гос. ун-т. – Ярославль: ЯрГУ, 2007. – 175 с.

ISBN 5-8397-0517-9 (978-5-8397-0517-3)

Представлены материалы 2-й научно-методической конференции "Преподавание математики и компьютерных наук в классическом университете", состоявшейся в Ярославском государственном университете им. П. Г. Демидова в феврале 2007 г.

ISBN 5-8397-0517-9 (978-5-8397-0517-3)

Содержание

Предисловие

Подобно другим естественным наукам, математика представляет собой игру, в которую мы играем с окружающим миром, со Вселенной. Самые лучшие математики и самые хорошие преподаватели — это, очевидно, люди, которые прекрасно разбираются в её правилах, а также получают удовольствие от самого процесса игры.

Мартин Гарднер, "Крестики – нолики"

Преподаватель диктует: "окружность — это геометрическое место точек на плоскости, находящихся на одном и том же расстоянии от одной внутренней точки, именуемой центром". Хороший ученик вписывает эту фразу в тетрадь; плохой ученик рисует в ней "человечков", но ни тот, ни другой ничего не поняли. Тогда преподаватель берёт мел и рисует круг на доске. "Ага, — думают ученики, — почему он не сказал сразу: окружность — это кружок, и мы бы сразу поняли".

Анри Пуанкаре, "Наука и метод"

Математика и её приложения отражают все наши представления о мире, причём это отражение происходит в рамках специальных и весьма сложных структур, называемых математическими структурами. Поэтому растить математиков — дело трудное и деликатное.

В Ярославском государственном университете имени П. Г. Демидова подготовка специалистов в области математики и прикладой математики осуществляется на математическом факультете и факультете информатики и вычислительной техники (ИВТ). Эти два факультета располагаются в одном 7–м здании ЯрГУ, на левом берегу Волги. Нынешний 2006 год был для них юбилейным: математическому факультету исполнилось 30 лет, а факультету информатики и вычислительной техники — 20 лет. За это время факультетами подготовлено большое количество высококвалифицированных работников по ряду наукоёмких специальностей и направлений. Их выпускники входят в ведущий кадровый состав многих предприятий и организаций Ярославля и всего региона.

Преподавателями и сотрудниками факультетов накоплен значительный и ценный опыт в научно-педагогической области. Возникла идея зафиксировать этот опыт посредством проведения своей "заволжской" научно-методической конференции и издания сборника её материалов.

Решение провести в феврале 2007 года научно-методическую конференцию "Преподавание математики и компьютерных наук в классическом университете" и опубликовать материалы конференции было принято на заседании Учёного совета математического факультета 15 мая 2006 года. Предложение участвовать в конференции позднее было с одобрением встречено Учёным советом факультета ИВТ. Это уже вторая научно-методическая конференция с таким названием (первая состоялась в 2005 году).

Организаторы конференции предложили преподавателям двух факультетов подготовить доклады следующей примерной тематики:

1) специальные вопросы, возникающие при преподавании математических и компьютерных дисциплин;

2) вопросы методического характера, связанные с организацией преподавания этих дисциплин;

3) актуальные проблемы математического образования и образования в области компьютерных наук.

В оргкомитет поступили 37 докладов преподавателей, представляющих все семь кафедр математического факультета (алгебры и математической логики; дифференциальных уравнений; компьютерной безопасности и математических методов обработки информации; математического анализа; математического моделирования; общей математики; теории функций и функционального анализа), а также все пять кафедр факультета ИВТ (вычислительных и программных систем; дискретного анализа; информационных и сетевых технологий; компьютерных сетей; теоретической информатики). Как видно из содержания, в докладах рассматриваются весьма разнообразные вопросы. связанные с преподаванием математических и компьютерных дисциплин.

В настоящем сборнике читатель найдёт все представленные авторами тексты. Они печатаются в авторской редакции, то есть сохраняют не только суждения, но и особенности стиля авторов. Компьютерный набор в системе LaTeX сделан самими авторами. Ответственным редактором был собран воедино весь материал, внесена некоторая (чаще техническая) правка в отдельные доклады, а также осуществлена необходимая подготовка сборника к печати.

Не секрет, что среди преподавателей есть талантливые учёные, скептически относящиеся к методической работе. Однако представляется бесспорным то, что любой исследователь постоянно занимается методическим осмыслением близкого ему научного материала.

М. Невский,
ответственный редактор,
заместитель декана математического факультета
по научной работе

О преподавании дисциплины "Теория псевдослучайных генераторов"

В. А. Бадоев

Рассматривается один из подходов к преподаванию дисциплины специализации "Теория псевдослучайных генераторов".

В список дисциплин специализации "Математические методы защиты информации" по специальности "Компьютерная безопасность" решением УМО по образованию в области информационной безопасности включена дисциплина "Теория псевдослучайных генераторов". Она требует знания основ теории вероятностей и математической статистики. Теория вероятностей и математическая статистика изучаются одновременно с этой дисциплиной, что способствует, на наш взгляд, лучшему усвоению студентами основных понятий и результатов. По разным причинам при изложении дисциплины "Теория псевдослучайных генераторов" имеет смысл повторить основные определения теории вероятности и упомянуть некоторые свойства. К этим понятиям относятся: случайная величина, ряд распределения случайной величины, функция распределения, плотность, моменты случайной величины, системы случайных величин, коэффициент корреляции, проверка гипотез.

При изучении дисциплины "Теория псевдослучайных генераторов" прослеживается связь и переход от непрерывной теории вероятности к дискретной, а от дискретной — к "компьютерной", основанной на применении персонального компьютера (первичная обработка и анализ исходных данных).

Проблематика курса состоит в следующем. Рассматриваются основные на сегодняшний день методы генерации базисной случайной величины (случайной величины, равномерно распределенной на [0,1]). Для генерации базовой случайной величины используются: мультипликативный конгруэнтный метод; метод линейных смешанных формул; квадратичный метод Ковэю; метод основанный на свойстве воспроизводимости равномерного закона.

В качестве дополнительного материала студентам предлагается самостоятельно изучить методы генерации основных классических случайных величин: нормальной, показательной, биномиальной, пуассоновской. За основной метод генерации случайной последовательности предлагается взять метод обратной функции. В результате к окончанию изучения дисциплины курса у студентов формируется набор программ, генерирующих и анализирующих случайные последовательности.

На заключительном этапе изучения курса студентам предлагается сгенерировать случайную последовательность с заданной плотностью распределения и провести ее анализ с помощью изученных методов. Для анализа точности сгенерированной случайной величины применяются критерии Пирсона, Колмогорова, Смирнова – Крамера – Мизеса, тест "совпадения моментов". Функцию плотности распределения обычно конструируют с использованием элементарных функций.

Приведем пример задания, которое выдается студентам по окончании изучения курса.

Плотность f(x) распределения случайной величины задана графически. Требуется по ней:

1) методом обратной функции сгенерировать случайную последовательность длины $n = 100, 10000, 100000, 1000000$, *имеющей плотность распределения* $f(x)$;

2) сделать группировку полученной последовательности;

3) найти эмпирическое среднее значение и эмпирическую дисперсию и сравнить их с теоретическими значениями;

4) в случае, если различия между эмпирическими и теоретическими результатами (см. предыдущий пункт) несущественны, используя критерии согласия, проверить гипотезу — может ли полученная последовательность иметь исходную плотность распределения; в случае, если различия между эмпирическими и теоретическими результатами существенны, следует проверить правильность алгоритма генерации последовательности;

5) оценить трудоемкость получения одного значения случайной величины, имеющей плотность распределения $f(x)$.

Решение задачи дополнить геометрической интерпретацией — построить гистограмму и график плотности распределения в одной системе координат. Сравнить качество случайных последовательностей (имеющих плотность $f(x)$*), полученных на основе базисной случайной величины, сгенерированной различными методами. Дать обоснованный ответ, какой из методов предпочтительнее.*

В заключение хотелось бы остановиться на некоторых важных методических моментах. Нам представляется, что было бы правильным, если бы все студенты писали программы на одном языке программирования. При первичном изложении дисциплины студентам была дана свобода в выборе языка программирования, что привело к затруднению полной проверки алгоритмов в программе и практически к невозможности помочь студентам находить ошибки в реализации алгоритмов. Особое внимание следует обратить на быстродействие алгоритмов.

Считаю курс "Теории псевдослучайных генераторов" своевременным и нужным. При изложении математического анализа и алгебры на младших курсах эти предметы подкрепляются численными методами и компьютерной реализацией, а теория вероятностей в этом смысле оставалась не более чем теорией. Настоящая дисциплина выполняет аналогичную вспомогательную функцию по отношению к дисциплине "Теория вероятностей".

Классификация тестовых заданий на основе оценки остаточной энтропии (на примере заданий ЕГЭ по математике за 2005 год)

И. В. Баженов

Работа посвящена методу оценки селективной способности тестовых заданий с помощью показателя остаточной энтропии, который характеризует неопределенности системы ответов в случае исключения каждого из заданий из теста. На основе модели проведен анализ результатов тестирования ЕГЭ, приведена классификация заданий по селективной ценности.

Библиография: 2 названия.

В настоящее время задачи управления качеством образования выходят на первый план, поэтому необходим мониторинг состояния знаний учащихся. Традиционно используемая пятибалльная шкала не позволяет полностью объективно отразить уровень знаний каждого учащегося. Использование тестов для контроля уровня обученности позволяет получить объективную и надежную информацию о результатах обучения.

Рассмотрим тест, содержащий варианты ответов. Испытуемый получает отметку о выполнении либо невыполнении задания. Пусть каждый правильный ответ оценивается на один балл, неправильный — на ноль баллов. В результате каждый испытуемый по итогам тестирования получает набор ответов из нулей и единиц. Множество векторов ответов из k таких заданий содержит 2^k элементов.

Задачей данного исследования была оценка селективной способности каждого задания и влияния каждого из заданий на общую неопределенность системы теста.

Селективная способность теста может быть оценена с помощью показателя энтропии, который отражает неопределенность системы.

Согласно Шеннону, показатель энтропии определяется равенством

$$E = -\sum_{k=1}^{N} \tilde{p}_k \log_2 \tilde{p}_k, \tag{1}$$

где N — число состояний системы ($N = 2^{30}$), $\tilde{p}_k$ — вероятность получения одного из состояний системы.

Показатель энтропии (1) для теста позволяет судить об однородности системы векторов из нулей и единиц, отражающих ответы испытуемых. Максимальная энтропия для k заданий теста достигается при абсолютной однородности системы, когда каждый испытуемый равновероятно получает один из векторов системы, т.е. $E_{max} = k$.

Наличие в тесте слишком простых или слишком сложных заданий уменьшают неопределенность системы и, следовательно, энтропия уменьшается.

Пусть при условии независимости ответов каждый испытуемый отвечает правильно на k-й вопрос с вероятностью p_k и не отвечает с вероятностью $1 - p_k$. Тогда показатель энтропии, исходя из (1), представляется в следующем виде:

$$E = -\sum_{k=1}^{30} (p_k \log_2 p_k + (1 - p_k) \log_2 (1 - p_k)). \tag{2}$$

Для того, чтобы оценить селективную способность каждого из заданий теста, введем понятие частичной энтропии, которая характеризует неопределенность системы ответов на одно из заданий теста. Показатель определяется равенством

$$E_k = -\left(p_k \log_2 p_k + (1 - p_k) \log_2 (1 - p_k)\right). \tag{3}$$

Пусть $X = \{x_1,x_2,...,x_k\}$ — набор заданий теста. Требуется упорядочить этот набор согласно степени влияния каждого из этих заданий на селективную способность теста.

Остаточная энтропия для i-го задания теста определяется по формуле

$$\grave{E}_i = -\sum_{k=1,k\neq i}^{N} \hat{p}_k \log_2 \hat{p}_k, \tag{4}$$

где N — число состояний системы, $\hat{p}_k$ - вероятность получения одного из состояний системы ответов на тест, представленный набором $\grave{X} = \{x_1,x_2,...x_i,x_{i-1},x_k\}$.

Задания теста можно упорядочить по абсолютному отклонению остаточной энтропии от общей энтропии теста.

Разница остаточной энтропии i-го задания теста есть

$$\Delta E_i = E - \grave{E}_i, \tag{5}$$

где E — показатель энтропии теста, определяемый (1).

Оценка заданий по отклонению остаточной энтропии позволяет учитывать возможную зависимость ответов испытуемых.

Автор воспользовался данными ЕГЭ по математике, проведенному в 2005 году в г. Ярославле. Тест был рассмотрен в части A и B, которые содержали варианты ответов. Тест содержит 10 заданий типа A и 11 заданий типа B. Задания B9, B10 и B11 предлагали испытуемым самостоятельно написать ответ, что исключало выбор ответа "наугад". Энтропия теста в части A и B была оценена по формуле (1) ($E_{AB} = 11,1384$, для сравнения $E_{max(AB)} = 21 > E_{AB}$).

Приведем статистические данные. Для оценки "трудности" заданий части A и B была вычислена вероятность их выполнения p_i, i — номер задания. Чем ниже вероятность p_i, тем больше трудность.

Классификация заданий по трудности (в скобках указана вероятность p_i) : A1 (0,8286), A8 (0,8166), A3 (0,7928), A2 (0,7841), A4 (0,7386), A5 (0,7024), B1 (0,6891), A7 (0,6871), A9 (0,6060), A6 (0,5794), B2 (0,5792), B3 (0,5462), A10 (0,5460) , B4 (0,4612), B6 (0,3775), B5 (0,3753), B7 (0,2519), B8 (0,2386), B9 (0,1862), B10 (0,1788), B11 (0,0949).

Согласно (4) определена остаточная энтропия каждого из заданий. В качестве параметра для классификации заданий использовалось абсолютное отклонение остаточной энтропии от энтропии теста, определяемое (5). Задания упорядочены в порядке убывания отклонений ΔE_i (в скобках указано значение остаточной энтропии): A6(0,284534), B6(0,269617), B5(0,256569), A10(0,243958), B8(0,243441), A9(0,233683), A7(0,224423), B7(0,222534), A5(0,221957), B9(0,202763), B10(0,193163), B4(0,191258), B2(0,190293), B3(0,171134), A2(0,165318), A4(0,151229), B1(0,148927), B11(0,134968), A8(0,129787), A3(0,126406), A1(0,122159).

Как видно из полученной классификации, максимальным отклонением остаточной энтропии обладает задание A6, что говорит о его наибольшей значимости в селективности теста. Данное задание теста предлагает найти производную функции, что требует знаний правил дифференцирования и представляет определенные трудности для части учащихся, но в то же время задание не представляет сложности для другой части испытуемых. Задание имеет достаточно высокое значение показателя частичной энтропии, вычисленной согласно (3), равное 0,98, что говорит о его высокой селективной способности вне зависимости от ответов учащихся на смежные задания теста.

Задания, обладающие минимальным отклонением остаточной энтропии, такие как A1, A3, A8, B1, являются достаточно простыми для учащихся, их решает большая часть испытуемых. В то время как с заданием B11 по геометрии справляется наименьшее число испытуемых, вероятность того, что это задание решено, минимальна. Эта группа заданий оказывает наименьшее влияние на селективную способность теста.

Как уже отмечалось выше, селективную способность заданий можно оценить с помощью показателя частичной энтропии. Приведем классификацию заданий в соответствии с показателем частичной энтропии (значение показателя для каждого из заданий указано в скобках): B4(0,995649), A10(0,993883), B3(0,993825), B2(0,981846), A6(0,981745), A9(0,967348), B6(0,956223), B5(0,954635), A7(0,896503), B1(0,894263), A5(0,878393), A4(0,828923), B7(0,814203), B8(0,792624), A2(0,752562), A3(0,736004), B9(0,693494), A8(0,687425), B10(0,677456), A1(0,660877), B11(0,452467).

Следует заметить, что классификация заданий данного теста в соответствии с отклонением остаточной энтропии отличается от классификации заданий по частичной энтропии. Это говорит о зависимости ответов испытуемых на одно из заданий теста от ответов на другие задания.

Метод оценки остаточной энтропии заданий теста позволяет проанализировать селективную способность тестовых заданий в рамках всего теста, являясь более корректным, чем метод оценки частичной энтропии, с точки зрения статистики, так как не использует ограничение независимости ответов учащихся. Оценка остаточной энтропии позволяет повысить информационную "стоимость" теста за счет выявления заданий, имеющих наименьшую селективную ценность в общей структуре теста и дальнейшей корректировки теста.

Список литературы

[1] *Майоров В. В., Майорова Н. Л.* К вопросу об оценки уровня селективности тестовых заданий // Математика в Ярославском университете: Сб. обзорных статей. К 30-летию математического факультета. Ярославль, 2006. С. 301–305.

[2] *Хакен Г.* Синергетика. М.: 362 с.

О междисциплинарных связях курса "Сети Петри"

В. А. Башкин

Рассматриваются возможности курса "Сети Петри" по расширению научного кругозора учащихся в области некоторых разделов математики и теоретической информатики.

Библиография: 11 названий.

Дисциплина "Сети Петри" — одна из дисциплин цикла "Общепрофессиональные дисциплины" учебного плана по специальности "Прикладная математика и информатика" для студентов, специализирующихся по кафедре теоретической информатики. Целью изучения дисциплины является ознакомление учащихся с основами теории сетей Петри, а также с примыкающими разделами дискретной математики; обучение основным методам моделирования и анализа параллельных и распределенных систем; ознакомление с историей развития теории сетей Петри.

Важными особенностями данного курса являются, с одной стороны, его насыщенность математическими понятиями и подходами, и, с другой стороны, тесная связь с практикой. В частности, в ходе изучения дисциплины студенты знакомятся с такими областями применения сетей Петри, как моделирование и проектирование систем, верификация программ и протоколов, управление параллельными и распределенными системами (в том числе технологическими и бизнес-процессами).

Сети Петри — один из наиболее популярных классов формальных моделей, используемых для моделирования и анализа систем (в первую очередь параллельных и распределенных) [1, 2]. Это весьма наглядный формализм, обладающий простой графической формой записи и в то же время достаточной выразительностью. В сетях Петри существуют конструкции, позволяющие адекватно моделировать основополагающие структуры параллельных и распределенных вычислений (последовательное соединение, выбор, распараллеливание, синхронизация) [11]. Основы теории сетей Петри были заложены в 60–е годы XX века [10], с тех пор теория активно развивается, причем важные научные результаты появляются в ней и в настоящее время.

Рассмотрим особенности данной дисциплины, позволяющие в ходе её изучения затронуть ряд смежных областей математики и информатики.

Теория алгоритмов и формальных языков. Изучение класса сетей Петри начинается с его позиционирования среди известных студентам классов формальных моделей вычислений. Уже на самых первых лекциях можно достаточно просто и наглядно показать, что обыкновенные сети Петри занимают по выразительности промежуточное положение между конечными системами (конечными автоматами) и универсальными вычислителями (машинами Тьюринга и т. д.). Удобнее всего это сделать при помощи сравнения классов языков, порождаемых соответствующими классами систем [1, 7].

В настоящее время существует большое число разнообразных классов систем на основе обыкновенных сетей Петри, что позволяет в ходе их изучения познакомить студентов с целым спектром классов вычислительной мощности. Рассматриваются автоматные сети Петри, равномощные конечным автоматам, и два класса универсальных вычислителей — сети Петри с ингибиторными дугами и сети Петри с приоритетами [1] (что в очередной раз напоминает студентам о феномене, сформулированном в тезисе Чёрча – Тьюринга). Также изучается ряд промежуточных классов, в частности, обыкновенные сети Петри и сети Петри со свободным выбором.

Одной из целей дисциплины "Сети Петри" можно считать формирование у студентов представления о возможностях и ограничениях формального подхода к моделированию. В связи с этим важное значение приобретают знания о разрешимости и вычислительной сложности различных алгоритмических проблем для различных классов формальных моделей. В ходе односеместрового курса сделать полный обзор невозможно, поэтому конкретной целью является более общее изучение данной темы, результатом которого для студента должен стать тезис о том, что при увеличении сложности используемого при моделировании класса систем растет трудоемкость анализа получаемой модели (вплоть до неразрешимости свойств).

Рассматривается ряд алгоритмических проблем, а именно проблемы останова, достижимости (достижимость конкретного состояния), ограниченности (конечность множества состояний), живости (возможность активации перехода в любом достижимом состоянии), эквивалентности множеств достижимости и равенства языков. При их изучении студентов можно познакомить со следующими интересными фактами:

- проблема эквивалентности множеств достижимости неразрешима для сетей с более чем четырьмя неограниченными позициями и разрешима для остальных [6];

- проблема равенства языков неразрешима для сетей с более чем одной неограниченной позицией и разрешима для остальных [6];

- проблема достижимости NP-полна для обыкновенных сетей Петри, при этом для ограниченных сетей существует полиномиальный алгоритм [3].

Формальные языки обычно изучаются в университете в рамках классов языков из иерархии Хомского. Теория сетей Петри дает хорошую возможность расширить кругозор студентов. Дело в том, что класс языков обыкновенных сетей Петри не удается четко расположить в рамках этой иерархии, так как он не сравним с классом контекстно-свободных языков. Этот результат достаточно легко доказывается на примерах (см. [1, 7]). Интересно, что доказательство существования КС-языков, не распознаваемых сетями Петри (см. [7]), попутно выявляет важное свойство сетей Петри — мощность множества состояний растет в них в полиномиальной зависимости от длины срабатывания, в отличие от магазинных автоматов и других КС-систем, где этот рост экспоненциален. Впрочем, это не означает меньшей выразительности сетей Петри — память в них расходуется по-другому, что позволяет порождать многие не-КС-языки. На этом примере очень удобно объяснить студентам особенности различных способов организации бесконечной памяти: очереди, стека (КС-системы) и вектора (сети Петри).

Факультативно в данной теме можно рассмотреть еще одну иерархию классов формальных моделей (и соответствующих формальных языков) — иерархию Майра систем переписывания процессов (PRS-систем) [9]. В отличие от иерархии Хомского, это не цепочка классов, а решетка, и в ней имеются узлы, соответствующие как последовательным системам (КС-языки), так и параллельным (языки сетей Петри). Иерархия Майра — новый и достаточно красивый способ классификации формальных языков, но изучение его требует предварительного изучения синтаксиса алгебр процессов.

Теория множеств и математическая логика. Первая тема, изучаемая в ходе данного курса — мультимножества. Знание понятия конечного мультимножества и операций над мультимножествами совершенно необходимо, поскольку именно мультимножества используются в течение всего курса для обозначения состояний сети Петри.

При рассмотрении одной из основных областей практического использования сетей Петри — верификации — очевидным образом возникает необходимость изучения способов спецификации "правильного" поведения системы. На практике для такой спецификации чаще всего используются различные темпоральные логики. По мнению автора, в рамках данного курса достаточно познакомить студентов с самыми общими принципами задания свойств системы при помощи формул (удобнее всего использовать синтаксис логики CTL [8]).

Факультативно можно рассмотреть ещё один вид логики, связанный с теорией сетей Петри — интуиционистскую линейную логику Жирара [5]. Она отличается от классической тем, что рассматривает посылки вывода как расходуемые ресурсы. Это можно легко промоделировать при помощи сетей Петри [4].

Алгебра. Одним из способов записи сети Петри является матричный. При этом способе каждое срабатывание сети задается матричным уравнением (т.н. уравнение состояний) с участием матрицы инцидентности и векторов предусловия и постусловия. Изучение этого уравнения методами линейной алгебры, в частности, нахождение его специальных решений — так называемых P-инвариантов и T-инвариантов [2] — дает мощный инструмент для анализа свойств модели.

Еще один пример применения алгебраического подхода в теории сетей Петри — специальные хорошо структурированные классы сетей. В частности, регулярные сети Котова [1] формируются из конечного числа образующих элементов при помощи заданного набора операций.

Теория вероятностей. В ходе данного курса рассматриваются стохастические сети Петри, которые используются для моделирования и анализа систем с вероятностями.

Список литературы

[1] *Котов В. Е.* Сети Петри. М.: Наука. 1984.

[2] *Питерсон Дж.* Сети Петри и моделирование систем. М.: Мир. 1984.

[3] *Esparza J.* Decidability and complexity of Petri net problems — an introduction // LNCS. 1998. V.1491. P.374–428.

[4] *Farwer B.* A Linear Logic view of object Petri nets // Fundamenta Informatica, vol. 37, №3, 1999. P. 225–246.

[5] *Girard J.-Y.* Linear Logic // Theoretical Computer Science, vol. 50, 1987. P. 1–102.

[6] *Jančar P.* Decidability questions for bisimilarity of Petri nets and some related problems // Proc. of STACS'94. LNCS. 1993. V.775. P. 581–592.

[7] *Jantzen M.* Language Theory of Petri Nets // LNCS. 1987. V. 254.

[8] *Manna, Z., Pnueli A.* The Temporal Logic of Reactive and Concurrent Systems. Springer-Verlag, 1992.

[9] *Mayr R.* Process Rewrite Systems // Information and Computation. Vol. 156, No 1. 2000. P. 264-286.

[10] *Petri C. A.* Kommunikation mit Automaten. PhD theses. Bonn: Institute für Instrumentelle Mathematik. 1962.

[11] *Reisig W.* Petri net models of distributed algorithms // LNCS. 1995. V. 1000.

Об одной задаче по теме "Системы линейных уравнений"

М. А. Башкин

На практических занятиях по алгебре со студентами специальности "Компьютерная безопасность" предлагается рассматривать метод решения систем линейных уравнений с целыми коэффициентами в целых числах.

Библиография: 6 названий.

В данной статье речь пойдет о методе решения систем линейных уравнений с целыми коэффициентами в целых числах (см. [1]). Автор предлагает включить этот метод в число задач, рассматриваемых на первом курсе специальности "Компьютерная безопасность" по дисциплине "Алгебра". Метод решения систем линейных уравнений в целых числах согласуется с темой "Системы линейных уравнений" и не требует предварительного освещения на лекции, так как студенты к этому времени уже имеют необходимые знания. Действительно, тема "Системы линейных уравнений" является последней темой второго раздела программы, а необходимые сведения для понимания обоснования метода решения содержатся в предыдущих темах этого раздела. Поэтому проблем с обоснованием у студентов возникнуть не должно. Рассмотрение данного метода имеет еще и междисциплинарное значение. Например, он находит свое применение в теории чисел. В результате мы прослеживаем связь алгебры с теорией чисел на основании темы "Системы линейных уравнений". Поскольку метод решения может быть записан как алгоритм, его легко запрограммировать.

Задачи, где требуется найти все целочисленные решения системы линейных уравнений, содержатся во всех изданиях сборника задач по алгебре под редакцией А. И. Кострикина (см. [4], [5] и [6]). Кроме них можно использовать и другие сборники задач, где содержатся системы линейных уравнений с целыми коэффициентами. Например, любое издание сборника задач по линейной алгебре под редакцией И. В. Проскурякова ([2] или [3]). При этом стоит отметить, что в учебной литературе эта тема практически не встречается.

Рассмотрим задачу решения систем линейных уравнений с целыми коэффициентами в целых числах более подробно. Пусть

$$\left\{ \begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ \multicolumn{3}{c}{\dotfill} \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \end{array} \right. \tag{1}$$

— система линейных уравнений с целыми коэффициентами. Необходимо найти все целочисленные решения системы (1).

Чтобы получить ответ на вопрос задачи, необходимо преобразовать систему (1) так, чтобы множество целых решений усматривалось непосредственно. При этом надо позаботиться о том, чтобы новая система была эквивалентна исходной. Сделать это можно с помощью уже известных студентам к этому времени преобразований метода Гаусса, при которых сохраняется условие целочисленности коэффициентов рассматриваемой системы, т. е. с помощью перестановки уравнений, умножения уравнения на отличное от нуля число, прибавление к одному уравнению другого уравнения, умноженного на любое целое число. Но всегда ли такой подход приведет к желаемому результату? Рассмотрев систему из одного уравнения, например, $2x_1+3x_2+4x_3 = 5$, убеждаемся, что это не так.

Рассмотрим матрицу A системы (1). Не отказываясь от элементарных преобразований строк расширенной матрицы системы уравнений, введем следующие элементарные преобразования для первых n столбцов:

А. перестановка столбцов;

Б. умножение столбцов на ± 1;

В. прибавление к столбцу другого столбца, умноженного на целое число.

Эти преобразования вместе со своими обратными не выводят за пределы целых чисел. Истолковываются они как следствия линейных замен неизвестных, выполняемых последовательно.

Наша цель — привести матрицу A к виду

$$D = diag\{d_1, d_2, \ldots, d_r\}, \tag{2}$$

где $d_1, \ldots d_r$ — целые положительные числа и $d_i | d_{i+1}$.

Приведение матрицы системы A к виду (2) осуществляем с помощью следующего алгоритма:

1. переставляя строки и столбцы, перемещаем в левый верхний угол преобразуемой мотрицы наименьшее по модулю число из элементов матрицы, не равных нулю;

2. прибавляя к строке другую, умноженную на целое число, а к столбцу — другой столбец, максимально уменьшаем модули всех остальных элементов первого столбца и первой строки;

3. выполняем шаги 1 и 2 до тех пор, пока в левом верхнем углу матрицы не окажется наименьший по модулю из не равных нулю элементов, а все остальные элементы первой строки и первого столбца не станут нулями;

4. переходим к матрице меньшего размера, возникающей из полученной матрицы игнорированием первой строки и первого столбца, затем аналогично преобразуем эту матрицу и т.д.

5. умножая строки на ± 1, добиваемся положительности d_i, $i = 1, \ldots, r$.

Для поиска решений системы мы рассматриваем расширенную матрицу $(A\,B)$ системы (1), где

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \multicolumn{3}{c}{\dotfill} \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

и приписываем снизу к ее подматрице коэффициентов при неизвестных единичную матрицу размера $n \times n$:

$$\begin{pmatrix} A & B \\ E & \end{pmatrix}.$$

При этом элементарные преобразования строк матрицы A распространяем на столбец свободных членов, а элементарные преобразования столбцов на единичную матрицу. В результате получаем матрицу

$$\begin{pmatrix} D & C \\ U & \end{pmatrix}.$$

Тогда система совместна, если $d_i | c_i$ при $i = 1, \ldots, r$, $c_k = 0$ при $k > r$, а общее решение задается формулой

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = U \begin{pmatrix} c_1/d_1 \\ \vdots \\ c_r/d_r \\ y_{r+1} \\ \vdots \\ y_n \end{pmatrix},$$

где $y_{r+1}, \dots, y_n$ — любые целые числа. Если хотя бы одно из этих двух условий не выполняется, то система целых решений не имеет.

Обоснование рассмотренного метода можно найти в книге [1] §85.

В заключение приведем два примера оформления решений задач, взятых из [6].

Пример 1 (№8.24(а)). Найти все целочисленные решения системы уравнений

$$2x_1 + 3x_2 + 4x_3 = 5.$$

Решение.

$$\begin{pmatrix} 2 & 3 & 4 & 5 \\ 1 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 4 & 5 \\ 1 & -1 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 4 & 5 \\ -1 & 1 & 0 & \\ 1 & 0 & 0 & \\ 0 & 0 & 1 & \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & 4 & 5 \\ -1 & 3 & 0 & \\ 1 & -2 & 0 & \\ 0 & 0 & 1 & \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 5 \\ -1 & 3 & 4 & \\ 1 & -2 & -4 & \\ 0 & 0 & 1 & \end{pmatrix}.$$

Следовательно $d_1 = 1$, $c_1 = 5$, $U = \begin{pmatrix} -1 & 3 & 4 \\ 1 & -2 & -4 \\ 0 & 0 & 1 \end{pmatrix}$.

Общее решение данной системы —

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 & 3 & 4 \\ 1 & -2 & -4 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -5 + 3y_2 + 4y_3 \\ 5 - 2y_2 - 4y_3 \\ y_3 \end{pmatrix},$$

где y_2 и y_3 — целочисленные параметры.

Полезно провести проверку. (В [6] ответ для данной задачи указан неправильно.)

Ответ: $\{(-5 + 3y_2 + 4y_3, 5 - 2y_2 - 4y_3, y_3) \,|\, y_2, y_3 \in \mathbb{Z}\}$.

Пример 2 (№8.24(б)). Найти все целочисленные решения системы уравнений

$$\begin{cases} 2x_1 + 3x_2 - 11x_3 - 15x_4 = 1, \\ 4x_1 - 6x_2 + 2x_3 + 3x_4 = 2, \\ 2x_1 - 3x_2 + 5x_3 + 7x_4 = 1. \end{cases}$$

Решение.

$$\begin{pmatrix} 2 & 3 & -11 & -15 & 1 \\ 4 & -6 & 2 & 3 & 2 \\ 2 & -3 & 5 & 7 & 1 \\ 1 & 0 & 0 & 0 & \\ 0 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -11 & -15 & 1 \\ -10 & 4 & 2 & 3 & 2 \\ -5 & 2 & 5 & 7 & 1 \\ -1 & 1 & 0 & 0 & \\ 1 & 0 & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & -8 & -11 & 0 \\ 0 & 12 & -50 & -68 & 6 \\ -1 & 3 & -11 & -15 & \\ 1 & -2 & 11 & 15 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & -2 & 0 & 0 & 6 \\ 0 & 0 & -48 & -3 & -24 \\ -1 & -11 & -63 & -4 & \\ 1 & 11 & 64 & 4 & \\ 0 & 1 & 6 & -1 & \\ 0 & 0 & 0 & 1 & \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & -2 & 0 & 0 & 6 \\ 0 & 0 & -3 & 0 & -24 \\ -1 & -11 & -4 & 1 & \\ 1 & 11 & 4 & 0 & \\ 0 & 1 & -1 & 22 & \\ 0 & 0 & 1 & -16 & \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & 8 \\ -1 & -11 & -4 & 1 & \\ 1 & 11 & 4 & 0 & \\ 0 & 1 & -1 & 22 & \\ 0 & 0 & 1 & -16 & \end{pmatrix}.$$

Следовательно $d_1 = d_2 = d_3 = 1$, $c_1 = 1$, $c_2 = -3$, $c_3 = 8$,
$U = \begin{pmatrix} -1 & -11 & -4 & 1 \\ 1 & 11 & 4 & 0 \\ 0 & 1 & -1 & 22 \\ 0 & 0 & 1 & -16 \end{pmatrix}.$

Общее решение данной системы —

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 & -11 & -4 & 1 \\ 1 & 11 & 4 & 0 \\ 0 & 1 & -1 & 22 \\ 0 & 0 & 1 & -16 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -3 \\ 8 \\ y_4 \end{pmatrix} = \begin{pmatrix} y_4 \\ 0 \\ -11 + 22y_4 \\ 8 - 16y_4 \end{pmatrix},$$

где y_4 — целочисленный параметр.

Полезно провести проверку.

Ответ: $\{(y_4, 0, -11 + 22y_4, 8 - 16y_4) \mid y_4 \in \mathbb{Z}\}$.

Список литературы

[1] *Ван дер Варден Б. Л.* Алгебра. М.: Наука. 1976.

[2] *Проскуряков И. В.* Сборник задач по линейной алгебре. М.: Наука, 1984.

[3] *Проскуряков И. В.* Сборник задач по линейной алгебре. М.: Лаборатория базовых знаний, 2002.

[4] Сборник задач по алгебре / Под ред. А. И. Кострикина. М.: Наука, 1987.

[5] Сборник задач по алгебре / Под ред. А. И. Кострикина. М.: Факториал, 1995.

[6] Сборник задач по алгебре / Под ред. А.И.Кострикина. М.: Физматлит, 2001.

Об участии кафедры общей математики ЯрГУ им. П.Г. Демидова в работе Всероссийского семинара преподавателей математики университетов и педагогических вузов

Л. П. Бестужева

Дается краткий обзор основных результатов исследований преподавателей кафедры, представленных на семинаре за период 1990–2006 гг.

Библиография: 3 названия.

Всероссийский семинар преподавателей математики университетов и педагогических вузов проводится ежегодно, начиная с 1987 года, в различных городах России по инициативе А. Г. Мордковича. Имя А. Г. Мордковича знакомо всем, кто интересуется проблемами школьного математического образования и подготовки преподавателей математики в педагогических вузах и классических университетах. В 1987 году им была защищена докторская диссертация на тему «Профессионально-педагогическая направленность математической подготовки будущих учителей в педагогических институтах». Автор этой заметки имела возможность познакомиться с концепцией А. Г. Мордковича на его лекциях в МГПИ им. В. И. Ленина, будучи на повышении квалификации в этом вузе. Концепция базируется на четырех принципах. Согласно принципу рациональной фундаментальности, студент должен получить фундаментальную математическую подготовку, обеспечивающую ему знания, умения и навыки, далеко выходящие за рамки школьного курса математики. Фундаментальность является не целью, а средством подготовки учителя, а потому должна быть согласована с нуждами приобретаемой профессии. Принцип бинарности утверждает, что основу построения общих и специальных курсов должно составлять объединение общенаучной и методической линий. Согласно принципу ведущей идеи, на первый план выдвигаются взаимосвязи конкретного математического курса с соответствующим школьным предметом, что должно способствовать осмыслению будущим учителем структуры как вузовского, так и школьного курсов математики, осознанию системы изучаемых понятий и фактов как научного фундамента школьной математики. И, наконец, принцип непрерывности означает, что одно из условий профессиональной подготовки учителя состоит в непрерывности постижения студентами педагогической деятельности с первых дней обучения в вузе [1].

В конце 80-х и самом начале 90-х годов семинар проводился два раза в год: весной и осенью, а затем — один раз в год, осенью, как правило, в конце сентября. Материалы семинара (тезисы докладов участников) начали выпускаться с 1989 года. В 1990 году очередной семинар проводился в Ярославле. Принимал семинар ЯГПИ им. К. Д. Ушинского. Именно к этому времени относится начало участия практически всех преподавателей кафедры обшей математики в работе семинара. Каждый семинар имеет свою тему, в рамках которой участники семинара выступают с сообщениями в соответствии с госбюджетной темой кафедры «Проблемы эффективности учебного процесса в высшей школе: содержательные и технологические аспекты» в 90-е годы, нынешней темой «Проблемы качества послевузовского и дополнительного профессионального образования в контексте Болонского процесса» и своими научными интересами. С тех пор в сборниках материалов семинара появилось более 70 тезисов докладов преподавателей кафедры. Остановимся подробнее на содержании этих публикаций.

Занимаясь теорией и практикой многоуровневого университетского педагогического образования, В. А. Кузнецова исследовала различные аспекты этой проблемы: особенности подготовки преподавателя в классическом университете в форме дополнительного образования, организационно-методические вопросы и стандарты университетского педагогического образования в виде Государственных требований к минимуму содержания и уровню подготовки выпускников для получения дополнительной квалификации «Преподаватель», многопрофильность и вариативность подготовки педагогов в классических университетах, процессы фундаментализации и гуманизации в математическом образовании, вопросы формирования профессиональной культуры. Касаясь содержания обучения, В. А. Кузнецова указывает на необходимость усиления геометрической подготовки будущих преподавателей в классическом университете и предлагает использовать для этого элективные курсы. В последнее время в поле зрения В. А. Кузнецовой проблемы участия Российской высшей школы в Болонском процессе и вопросы проектирования образовательных систем.

Вопросы геометрической подготовки рассматривались Е. В. Никулиной, Л. Б. Медведевой и И. Р. Овсянниковой, что свидетельствует об актуальности этой проблемы в классическом университете в настоящее время. Многоуровневая система образования характеризуется появлением различных видов аттестационных работ студентов: квалификационных, выпускных, дипломных. Особенности этих работ обсуждали в своих докладах Ю. И. Большаков, Л. Б. Медведева и И. Р. Овсянникова. В условиях практически постоянного реформирования или модернизации школьного математического образования большое значение имеет система переподготовки учителей. На роль региональных центров педагогических исследований в переподготовке учителей указывают В. А. Кузнецова и Л. Б. Медведева. Не осталась без внимания и концепция математического образования в 12-летней школе. О педагогической подготовке учителя математики в русле этой концепции пишет Ю. И. Большаков. Ряд докладов Л. П. Бестужевой, Л. Б. Медведевой и И. Р. Овсянниковой посвящен Практикуму по решению задач — одной из дисциплин образовательной программы для получения дополнительной квалификации «Преподаватель».

Хорошо известно, что проблема обучения решению задач многоаспектна. О ее значимости свидетельствуют многочисленные публикации на эту «вечную тему». На автора этой заметки большое влияние оказали взгляды профессора С.–Петербургского государственного университета О. И. Иванова на систему математической и методической подготовки преподавателей профильных школ. В основе этой системы лежит принцип интегративности, который понимается как единство фундаментальных знаний в области математики, психологии, педагогики и методики преподавания математики, единство знаний элементарной и высшей математики. Выполнение принципа интегративности применительно к обучению решению задач обеспечивается, по мнению О. А. Иванова, двойной направленностью практикума, суть которой состоит в том, что «система учебных задач сливается с методикой обучения» [2]. Отметим некоторое противоречие: двойная направленность практикума не предполагает его использования для психолого-педагогических знаний, заявленных в принципе интегративности. Опираясь на концепцию профессионально-педагогической направленности математической подготовки учителей (А. Г. Мордкович) и учитывая принцип интегративности (О. А. Иванов), автор заметки счел возможным расширить направленность практикума по решению задач и использовать его для взаимообусловленных математических, психолого-педагогических и методических знаний. Некоторые аспекты конкретной реализации расширенной направленности практикума обсуждаются в докладах, посвященных обучению различным способам решения задач, включению эвристик в подготовку преподавателя математики, конструированию задач как учебной деятельности, обеспечивающей развитие творческой активности студентов.

Занимаясь проблемами подготовки преподавателей математики в классических университетах, В. Ф. Чаплыгин затрагивает в своих докладах вопросы преподавания математических дисциплин, в частности математического анализа, и создания учебных пособий, обсуждает обоюдную важность элементарной и высшей математики, обосновывает необходимость фундаментальных знаний и следования дидактическим принципам, особое внимание уделяет роли математики в формировании духовной культуры личности. В. Ф. Чаплыгин указывает на важнейшие моменты в подготовке преподавателя: развитие понятия числа, использование примеров и контрпримеров, обучение логическим рассуждениям.

Отметим, что в работе семинара в разные годы принимали участие не только преподаватели кафедры общей математики, но и преподаватели других кафедр математического факультета: Л. Ю. Белова, Т. Г. Бычкова, М. А. Доброхотова, Н. Л. Майорова, Н. Я. Кругляк, Е. Р. Семко, Н. В. Сенчакова, С. И. Яблокова, а также преподаватель факультета психологии Г. Ф. Третьякова. Тезисы ряда докладов были написаны в соавторстве с преподавателями ЯГПУ им. К. Д. Ушинского.

На XXV (юбилейном) семинаре, который проводился в Кирове, научный руководитель семинара А. Г. Мордкович отметил тот факт, что через семинар многие исследователи «обкатали» свои идеи, разработки, концепции, прошли своеобразную предзащиту кандидатских и докторских диссертаций. Семинарский «ученый совет», пишет А. Г. Мордкович, отличается от обычных диссертационных советов массовостью, полным пониманием существа излагаемых вопросов, внимательным и заинтересованным отношением к докладчику и исключительной благожелательностью [3]. С этим нельзя не согласиться. Доклад В. А. Кузнецовой о многоуровневой системе образования на одном из семинаров во время подготовки ею докторской диссертации вызвал оживленную дискуссию и инициировал тему следующего семинара, специально посвященного проблемам двухступенчатой подготовки учителя математики (Липецк, 1993 г.). На семинаре (Вологда, 2001 г.) во время работы над кандидатской диссертацией выступала и автор этой заметки. Заданные вопросы, замечания, высказанные участниками семинара, оказали неоценимую помощь в дальнейшей работе диссертантов.

Все семинары завершаются интересными экскурсионными программами, связанными с достопримечательностями тех городов России, где проходят семинары. Автору этой заметки особенно памятно посещение Кирилло-Белозерского и Ферапонтовых монастырей с фресками Дионисия на Вологодчине.

В работе семинара участие принимают и наши коллеги из ЯГПУ им. К. Д. Ушинского. Когда знакомишься со сборниками материалов семинара, обращает на себя внимание тот факт, что Ярославль регулярно представлен 8-20 участниками, что свидетельствует о творческой профессиональной активности ярославцев, которым близки проблемы школьного математического образования и подготовки учителей математики.

Отметим, что постоянные участники семинара упоминают его полное название только в официальных документах, а между собой называют его только так — «семинар Мордковича». И это о многом говорит. В этом году «детищу» А. Г. Мордковича исполнилось двадцать пять.

Список литературы

[1] *Мордкович А. Г.* О профессионально-педагогической направленности математической подготовки будущих учителей // Математика в школе. 1984. №6.

[2] *Иванов О. А.* Теоретические основы построения системы специальной математической и методической подготовки преподавателей профильных школ. СПб.: Изд-во С.-Петерб. ун-та, 1997.

[3] *Мордкович А. Г.* О Всероссийском научном семинаре преподавателей математики университетов и педагогических вузов // Проблемы подготовки учителя математики к преподаванию в профильных классах: Материалы XXV Всерос. семинара преподавателей математики ун-тов и педвузов. Киров; М.: ВятГГУ, МГПУ, 2006.

Риманова и псевдориманова метрики в области D евклидова пространства

Ю. И. Большаков

Представлен фрагмент курса лекций "Избранные вопросы геометрии и анализа", читаемого автором студентам третьего курса, обучающихся по специальности "Математика" и проходящих специализацию "Геометрия и топология".

Библиография: 3 названия.

1. Определение римановой и псевдоримановой метрик, длина дуги, угол между кривыми, скалярное произведение касательных векторов

Определение 1.1. *Мы будем говорить, что в области $D \subset \mathbb{R}^n$ задана риманова метрика, если в каждой криволинейной системе координат (z) на D определена $n \times n$–положительно определенная матрица G_z, элементы которой суть гладкие функции координат $(z) = (z_1, z_2, \ldots, z_n)$, причем при смене системы координат $(z) \mapsto (y)$ матрица G_z изменяется по закону:*

$$G_y = \left(\tfrac{\partial z}{\partial y} \right)^t G_z \left(\tfrac{\partial z}{\partial y} \right). \tag{1.1}$$

Здесь $\left(\frac{\partial z}{\partial y} \right) = \left(\frac{\partial(z_1,\, z_2,\, \ldots, z_n)}{\partial(y_1,\, y_2, \ldots,\, y_n)} \right)$ — матрица Якоби замены $(z) \mapsto (y)$. Отметим, что при замене G_z на G_y по формуле (1.1) свойство быть положительно определенной матрицей с гладкими компонентами сохраняется.

О существовании римановой метрики. Если в $\mathbb{R}^n$ скалярное произведение векторов (= столбцов) задано стандартным образом, то длина дуги l определена по формуле:

$$l = \int\limits_{\alpha}^{\beta} \sqrt{\sum_{i=1}^{n} (\dot{x}_i)^2}\, dt = \int\limits_{\alpha}^{\beta} \sqrt{\dot{x}^t \dot{x}}\, dt, \tag{1.2}$$

где $\dot{x}^t = (\dot{x}_1, \dot{x}_2, \ \ldots, \ \dot{x}_n)$, $x_i = x_i(t)$, $i = 1, \ 2, \ \ldots, n$, $\alpha \le t \le \beta$. Соотношение (1.2) равносильно

$$\left(\tfrac{dl}{dt} \right)^2 = \dot{x}^t \dot{x}. \tag{1.3}$$

Выражение, стоящее в правой части (1.3), суть матричное умножение. При смене криволинейной системы координат $(y) \mapsto (x)$ координаты касательных векторов меняются по закону:

$$\dot{x} = \left(\frac{\partial x}{\partial y} \right) \dot{y} \tag{1.4}$$

и, следовательно, инвариантность длины дуги (1.3) приводит к равенству:

$$\dot{x}^t \dot{x} = \dot{y}^t \left(\frac{\partial x}{\partial y} \right)^t \left(\frac{\partial x}{\partial y} \right) \dot{y}. \tag{1.5}$$

Используя ценное правило дифференцирования сложной функции, легко получить формулу (1.1).

Тем самым, мы установили существование римановой метрика в $D \subset \mathbb{R}^n$. Именно, если задать в каждой точке x области D матрицу $G_x = I_n$ — единичную $n \times n$–матрицу, тогда при смене криволинейных координат $(z) \mapsto (y)$ требование сохранения длин всех дуг приводит к формуле (1.1). Метрика $G_x = I_n$ называется евклидовой.

Определение псевдоримановой метрики почти полностью повторяет соответствующее определение римановой метрики, оно лишь заменяет положительную определенность G_z на её невырожденность. Более подробно.

Определение 1.2. *Мы будем говорить, что в области $D \subset \mathbb{R}^n$ задана псевдориманова метрика, если в каждой криволинейной системе координат (z) на D определена симметрическая невырожденная матрица G_z, элементы которой суть гладкие функции координат $(z) = (z_1,\ z_2,\ \ldots,\ z_n)$, причем при смене системы координат $(z) \mapsto (y)$ матрица G_z изменяется по закону (1.1).*

В качестве примера псевдоримановой метрики рассмотрим псевдоевклидову метрику $G_z = I_p \oplus (-I_{n-p})$. В силу теоремы инерции число p является инвариантом. При смене системы координат матрица G_z преобразуется по закону (1.1).

Пусть в области $D \subset \mathbb{R}^n$, отнесенной к криволинейной системе координат (z), задана риманова метрика G_z, тогда мы можем определить длину дуги гладкой кривой $\gamma : [\alpha,\ \beta] \to D$, заданной уравнениями $z_k = z_k(t),\ \ k = 1,\ 2,\ \ldots, n :$

$$l = \int\limits_{\alpha}^{\beta} \sqrt{\dot{z}^t G_z \dot{z}} dt. \tag{1.6}$$

Из соотношения (1.1) легко следует инвариантность длины дуги, определенной формулой (1.6).

Скалярное произведение $< \xi,\ \eta >$ касательных векторов к кривым γ и δ в точке $\gamma(t_0) = \delta(t_0) = p$ в системе координат (x) определяется по формуле:

$$< \xi,\ \eta > = < \dot{\gamma}(t_0),\ \dot{\delta}(t_0) > = (\dot{x}_\gamma(t_0))^t G_{x_0} \dot{x}_\delta(t_0), \tag{1.7}$$

где $x_\gamma(t)$ и $x_\delta(t)$ — суть координатная запись кривых γ и δ соответственно. Соотношение (1.1) гарантирует инвариантнсть скалярного произведения векторов.

Зная скалярное произведение касательных векторов к паре пересекающихся в данной точке p кривых, мы можем легко определить угол φ между этими кривыми в точке $p = x_0$:

$$cos\,\varphi = \frac{(\dot{x}_\gamma(t_0))^t G_{x_0} \dot{x}_\delta(t_0)}{\sqrt{(\dot{x}_\gamma(t_0))^t G_{x_0} \dot{x}_\gamma(t_0)}\sqrt{(\dot{x}_\delta(t_0))^t G_{x_0} \dot{x}_\delta(t_0)}}. \tag{1.8}$$

Ввиду инвариантности скалярного произведения касательных векторов только что сформулированное определение $cos\,\varphi$ также инвариантно.

Пример 1.1. *Показать, что отображение $f:\ \mathbb{R}^2 \to \mathbb{R}^2$, заданное соотношениями*

$$\begin{cases} y_1 = x_1 + sin\,x_2 \\ y_2 = \frac{1}{2} sin\,x_1 + x_2, \end{cases} \tag{1.9}$$

определяет криволинейную систему координат на $\mathbb{R}^2$, т.е. диффеоморфизм $\mathbb{R}^2$ на себя.

В самом деле, функции $y_i = y_i(x_1,\ x_2),\ i = 1, 2$, принадлежат классу C^1, якобиан $\left| \frac{\partial y}{\partial x} \right| \neq 0$ во всех точках $x \in \mathbb{R}^2$. Кроме того, (1.9) — биекция $\mathbb{R}^2$ на себя. Докажем последнее утверждение.

Если $x_1 + sin\ x_2 = x_1' + sin\ x_2'$ и $\frac{1}{2} sin\ x_1 + x_2 = \frac{1}{2} sin\ x_1' + x_2'$, то $|x_1' - x_1| = 2|sin\ \frac{x_2'-x_2}{2}||cos\ \frac{x_2'+x_2}{2}| \leq 2|sin\ \frac{x_2'-x_2}{2}| \leq |x_2' - x_2| = \frac{1}{2}|sin\ x_1' - - sin\ x_1| \leq \frac{2}{2}|sin\ \frac{x_1'-x_1}{2}||cos\ \frac{x_1'+x_1}{2}| \leq \frac{|x_1'-x_1|}{2}|$. Поэтому $x_1' = x_1$ и $x_2 = x_2'$. Пусть, далее, y_1 и y_2 — произвольные фиксированные действительные числа. Покажем, что существует пара чисел $x_1 = x_1^0,\ x_2 = x_2^0$, удовлетворяющая (1.9). С этой целью докажем разрешимость уравнения $y_1 = x_1 + sin\ (y_2 - \frac{1}{2} sin\ x_1)$ относительно x_1. Рассмотрим функцию $\varphi(t) = t + sin\ (y_2 - \frac{1}{2} sin\ t) - y_1$: на концах отрезка $[y_1 - 2,\ y_1 + 2]$ она принимает значения разных знаков. В силу её непрерывности существует точка $t = x_1^0$, в которой $\varphi(x_1^0) = 0$. Тогда из 2–го соотношения (1.9) находим $x_2 = x_2^0$. Это, в конечном счёте, и доказывает диффеоморфность $\mathbb{R}^2$ на себя.

Замечание 1.1. *Гладкость и отличность от нуля якобиана отображения в области D не гарантирует диффеоморфности областей.*

В самом деле, в области $D = \mathbb{R}^2 \backslash 0$ отображение f, заданное функциями

$$\begin{cases} y_1 = x_1^3 - 3x_1x_2^2 \\ y_2 = 3x_1^2x_2 - x_2^3 \end{cases}, \tag{1.10}$$

обладает двумя первыми свойствами. Однако, в точках $(\ cos\ \frac{\pi}{4},\ sin\ \frac{\pi}{4}\ )$ и $(\ -cos\ \frac{\pi}{12},\ sin\ \frac{\pi}{12}\ )$ значения отображения f совпадают.

Заметим, что в (1.10) $y_1 = Re\ z^3,\ y_2 = Im\ z^3$, где $z = x_1 + ix_2$.

Пример 1.2. *Вычислить длину дуги $\gamma:\ x_1 = a,\ x_2 = t,\ 0 \leq t \leq \frac{\pi}{2}$, если $G_y = I_2$, а криволинейные координаты (x) и (y) на $\mathbb{R}^2$ связаны соотношениями (1.9).*

Поскольку

$$G_x = \left(\frac{\partial y}{\partial x} \right)^t G_y \left(\frac{\partial y}{\partial x} \right) = \begin{pmatrix} 1+\frac{1}{4}cos^2\, x_1 & cos\, x_2 + \frac{1}{2}cos\, x_1 \\ cos\, x_2+\frac{1}{2}cos\, x_1 & 1+cos^2\, x_2 \end{pmatrix},$$

то $\dot{x}^t G_x \dot{x} = 1 + cos^2\, t = 2\left(1 - \left(\frac{1}{\sqrt{2}}\, sin^2\, t\right)\right)$. Поэтому $l = \sqrt{2}\int\limits_0^{\frac{\pi}{2}} \sqrt{1-\left(\frac{1}{\sqrt{2}}\right)^2 sin^2\, t}\; dt = \sqrt{2}\cdot\frac{\pi}{2}\left(1 - \frac{1}{2^2}k^2 - \frac{1^2\cdot 3}{2^2\cdot 4^2}k^4 - \frac{1^2\cdot 3^2\cdot 5}{2^2\cdot 4^2\cdot 6^2}k^6 - \ldots\right)$, где $k = \frac{1}{\sqrt{2}}$ (см. [1], с. 154, п. 774.1).

Пример 1.3. *Вычислить длину дуги параболы* $x_1 = t,\ x_2 = t^2,\ 0 \le t \le \frac{1}{\sqrt{2}}$ *на псевдоевклидовой плоскости* (x) *с* $G_x = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

Решение. $l = \int\limits_0^{\frac{1}{\sqrt{2}}} \sqrt{-\dot{x}_1^2 + \dot{x}_2^2}\, dt = \int\limits_0^{\frac{1}{\sqrt{2}}} \sqrt{-1+4t^2}\, dt = I_1 + I_2.$

$I_1 = i\int\limits_0^{\frac{1}{2}} \sqrt{1-4t^2}\; dt = \frac{i}{2}\int\limits_0^{1}\sqrt{1-z^2}\; dz = \frac{i}{2\cdot 2}\left(z\sqrt{1-z^2} + arcsin\, z\right)|_0^1 = i\frac{\pi}{8}$. (см. [1], п. 350.01). $I_2 = \int\limits_{\frac{1}{2}}^{\frac{1}{\sqrt{2}}} \sqrt{-1+4t^2}\; dt = \frac{1}{2}\int\limits_1^{\sqrt{2}}\sqrt{z^2-1}\; dz = \frac{1}{2\cdot 2}(z\sqrt{z^2-1} - ln\, |z + {}$
${} + \sqrt{z^2-1}|)|_1^{\sqrt{2}} = \frac{1}{4}(\sqrt{2} - ln\,(1+\sqrt{2}))$ (см. [1], п. 290.01). Т.о, $l = \frac{\pi}{8}i + \frac{1}{4}(\sqrt{2} -$
$- ln\,(1+\sqrt{2}))$.

Замечание 1.2. *Первое слагаемое в выражении для* l *в примере 1.3 является чисто мнимым. Оно отвечает так называемой "мировой линии". Геометрически мировая линия обладает следующим определяющим свойством: в любой её точке касательный вектор* ξ *лежит внутри изотропного конуса. В нашем случае* $\xi = (1,\ 2t)^t$ *лежит внутри угла* $\left(-\frac{\pi}{4},\ \frac{\pi}{4}\right)$, *ибо изотропные прямые имеют уравнения:* $x_1 = t,\ x_2 = \pm t,\ t \in \mathbb{R}$.

Упражнение 1.1. *Показать, что отображение* $f : D_1 \to D_1'$, *заданное соотношениями* $x_1 = \rho\ cos\ \varphi,\ x_2 = \rho\ sin\ \varphi$, *определяет криволинейную систему координат на* $D_1 = \{(\rho,\ \varphi)|\rho > 0,\ 0 < \varphi < 2\pi\}$ *со значениями в* $D_1' = \mathbb{R}^2\backslash\{(x_1,\ 0)|x_1 \ge 0\}$. *Здесь* $(x) = (x_1,\ x_2),\ y = (y_1,\ y_2) = (\rho,\ \varphi)$. *Вычислить* G_y, *если* $G_x = I_2$. *Написать формулу для вычисления длины дуги в полярной системе координат* $y = (\rho,\ \varphi)$.

Упражнение 1.2. *Решить задачи, сформулированные в упражнении 1.1, для цилиндрической и сферической систем координат в* $D_i \subset \mathbb{R}^3$ *и* D_i', *описав соответствуюшие области с помощью систем неравенств при* $i = 1, 2$.

2. Геометрия на сфере S^2

2.1. Метрика на сфере S^2

Пусть сфера $S^2 \subset \mathbb{R}^3$ имеет уравнение $x_1^2 + x_2^2 + x_3^2 = R^2$, а кривая $\gamma \subset S^2$ определена соотношениями $x_i = x_i(t),\ i = 1, 2, 3$, тогда $\left(\frac{dl}{dt}\right)^2 =$

$= \dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2 = \dot{x}^t \dot{x}$. В сферических координатах $y = (r,\ \theta, \psi)$:

$$\begin{cases} x_1 = r\ sin\ \theta\ cos\ \psi \\ x_2 = r\ sin\theta\ cos\ \psi \\ x_3 = r\ cos\ \theta \end{cases} \tag{2.11}$$

$\left(\frac{dl}{dt}\right)^2 = \dot{x}^t \dot{x} = \dot{y}^t G_y \dot{y} = (\dot{r},\ \dot{\theta},\ \dot{\psi}) \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\ sin\ \theta \end{bmatrix} \begin{bmatrix} \dot{r} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = R^2(\dot{\theta}^2 + sin^2\ \theta \dot{\psi}^2)$, ибо $r = R = const$. Таким образом, на сфере радиуса $r = R$ объемлющим евклидовым пространством индуцируется метрика $G_y = \begin{pmatrix} R^2 & 0 \\ 0 & R^2\ sin\ \theta \end{pmatrix}$ в системе координат $y = (\theta,\ \psi,)$ где θ — широта, отсчитываемая от (северного) полюса N, а ψ — долгота, отсчитываемая вдоль экватора от нулевого (гринвического) меридиана. Если рассмотреть стереографическую проекцию точки $P(\theta,\ \psi)$ на плоскость $(x_1,\ x_2)$, отнесенную к полярной системе координат $(z) = (\rho,\ \varphi)$, то, как нетрудно понять, $tg\frac{\theta}{2} = \frac{R}{\rho},\ \varphi = \psi$ и $G_z = \left(\frac{\partial y}{\partial z}\right)^t G_y \left(\frac{\partial y}{\partial z}\right) = \frac{4R^4}{(R^2+\rho^2)^2} \begin{pmatrix} 1 & 0 \\ 0 & \rho^2 \end{pmatrix}$, поэтому $\left(\frac{dl}{dt}\right)^2 = \frac{4R^4}{(R^2+\rho^2)^2}(\dot{\rho}^2 + \rho^2\dot{\varphi}^2)$.

Иначе говоря, метрика сферы G_z пропорциональна метрике евклидовой плоскости (см. упражнение 1.1). Говорят, что метрика сферы конформно евклидова. Смысл этого термина ясен — углы между двумя кривыми в подобных метриках равны.

2.2. Длина дуги локсодромы

Определение 2.2.1. *Локсодромой называют кривую $\gamma \subset S^2$, которая пересекает все меридианы под одним и тем же углом α.*

Пусть $\gamma : \theta = \theta(t),\ \psi = \psi(t)$ — локсодрома, а меридиан $\mu :\ \theta = t,\ \psi = \psi_0$, тогда $\left(\frac{dl}{dt}\right)^2 = R^2(\dot{\theta}^2 + sin\ \theta\ \dot{\psi}^2)$ и, кроме того,

$$cos\ \alpha = \frac{\begin{pmatrix} 1 \\ 0 \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ 0 & sin^2\theta \end{pmatrix} \begin{pmatrix} \dot{\theta} \\ \dot{\psi} \end{pmatrix}}{\sqrt{\begin{pmatrix} 1 \\ 0 \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ 0 & sin^2\theta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}} \sqrt{\begin{pmatrix} \dot{\theta} \\ \dot{\psi} \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ 0 & sin^2\theta \end{pmatrix} \begin{pmatrix} \dot{\theta} \\ \dot{\psi} \end{pmatrix}}} =$$

$\frac{\dot{\theta}}{\sqrt{\dot{\theta}^2 + sin^2\ \theta\ \dot{\psi}^2}}$. Поэтому $l = \int\limits_{t_1}^{t_2} R\sqrt{\dot{\theta}^2 + sin^2\ \theta\ \dot{\psi}^2}\ dt = R\int\limits_{t_1}^{t_2} \frac{d\theta}{cos\ \alpha} = \frac{R}{cos\ \alpha}(\theta(t_2) - \theta(t_1))$.

Если $\theta(t_1) = 0,\ \theta(t_2) = \pi$, то $l = \frac{\pi R}{cos\ \alpha}$. В частности, если $\alpha = 0$, то $l = \pi R$ — половина длины меридиана.

С целью более глубокого усвоения курса рекомендуем ознакомиться с книгами [2] и [3].

Список литературы

[1] *Двайт Г. Б.* Таблицы интегралов и другие математические формулы. М.: Наука, 1969. 228 с.

[2] *Мищенко А. С., Фоменко А. Т.* Курс дифференциальной геометрии и топологии. М.: МГУ, 1980. 440 с.

[3] *Дубровин Б. А., Новиков С. П., Фоменко А. Т.* Современная геометрия. М.: Наука, 1979. 760 с.

О некоторых подходах к решению задач по векторной алгебре

Г. М. Бродский

Сопоставляются творческие и систематические методы решения задач по векторной алгебре и предлагается правило, помогающее организовать рассуждения в случае систематического подхода.

Библиография: 7 названий.

Мысль о целесообразности показывать студентам на практических занятиях по математике весьма общий характер многих рассуждений, используемых при решении конкретных задач, не является новой. Здесь пойдет речь не об алгоритмах, для которых явные формулировки на лекциях и применение на практических занятиях стали общепринятыми, а о соображениях эвристического характера, облегчающих решение самых разнообразных задач. Автор обращается к этой теме не впервые. В адресованных молодым преподавателям методических указаниях [1] предложены некоторые правила, дающие единый метод решения задач на прямые и плоскости. В статье [2] сформулированы шесть правил организации рассуждений при решении задач на составление уравнений множеств точек. Идеи статьи реализованы в главе 1 учебного пособия [3], где упомянутые шесть правил преподносятся студентам (не как утверждения, которые обязательно следует запомнить, а как способы рассуждений, которым рекомендуется научиться). Одно из этих правил ([3], правило 4) используется автором в преподавании всех разделов курса „Геометрия и алгебра“ для студентов специальности „Прикладная математика и информатика“.

Переходя к обсуждению методов решения задач по векторной алгебре, разобьем их на два класса. К первому отнесем *творческие* методы, идеи которых связаны прежде всего со спецификой рассматриваемой задачи, а не с применением какого-либо весьма общего способа рассуждений. Второй класс составляют *систематические* методы, основанные на общих подходах, заслуживающих стандартизации. У методов указанных двух типов может быть и общее. Например, и творческие, и систематические методы могут использовать какие-либо соображения симметрии. Но у них есть и существенные различия. Творческие методы могут быть остроумными и приводить к очень короткому решению задачи. Однако многие из них становятся неприменимыми при незначительном изменении условия задачи. Студенту не всегда легко найти творческое решение задачи, в то время как правильное применение известного систематического метода гарантирует получение результата. Прежде чем перейти к сравнению творческих и систематических методов на примерах, сформулируем поддерживающее систематический подход

Правило. *При решении задач по векторной алгебре целесообразно все векторы и точки, участвующие в решении задачи, выражать через какой-либо минимальный набор векторов и точек, необходимый для задания рассматриваемого в задаче объекта (т.е. такой, что при удалении любого вектора или любой точки из него рассматриваемый объект оказывается определенным неоднозначно). Например, если речь идет о доказательстве тождества, то имеет смысл выразить через выбранный минимальный набор векторов и точек левую и правую части тождества и затем сравнить результаты. Если речь идет о нахождении вектора или точки, то искомый вектор или точку надо попытаться выразить через выбранный минимальный набор векторов и точек, и т.п. Минимальный набор векторов и точек следует выбирать так, чтобы ограничения на эти векторы и точки (в случае их появления из условия задачи) записывались как можно проще.*

Сформулированное правило основано на свойствах *геометрических векторных пространств* $\mathbb{L}_n$ $(n = 1, 2, 3)$, образованных всеми векторами в $\mathbb{A}_n$, где $\mathbb{A}_1$, $\mathbb{A}_2$ и $\mathbb{A}_3$ — *геометрические точечные пространства*, образованные всеми точками прямой, плоскости и трехмерного пространства соответственно. Одно из этих свойств заключается в следующем: в $\mathbb{L}_n$ всякая минимальная система образующих является базисом. Поскольку этим свойством, вообще говоря, не обладают свободные модули над кольцами, использование правила теряет смысл в геометрии над кольцом. Впрочем, изучение геометрии над кольцом выходит за рамки дисциплины „Геометрия и алгебра“. В нижеследующих примерах через $\vec{r}_T$ обозначается радиус-вектор точки T относительно выбранного полюса.

Пример 1 ([6], задача 10). Дан пространственный или плоский четырехугольник $ABCD$. Найти такую точку M, чтобы

$$\overrightarrow{MA}+\overrightarrow{MB}+\overrightarrow{MC}+\overrightarrow{MD}=\vec{0}. \tag{1}$$

Решение. Способ 1 (творческий). Пусть K, L, P, Q, R, S — середины отрезков AB, CD, BC, DA, AC, BD. Поскольку $\overrightarrow{MA}+\overrightarrow{MB}=2\overrightarrow{MK}$ и $\overrightarrow{MC}+\overrightarrow{MD}=2\overrightarrow{ML}$, то (1) равносильно равенству $\overrightarrow{MK}+\overrightarrow{ML}=\vec{0}$. Видим, что искомая точка M существует, единственна и является серединой отрезка, соединяющего середины противоположных сторон AB и CD . Ясно, что M можно охарактеризовать также как середину отрезка, соединяющего середины противоположных сторон BC и DA , как середину отрезка, соединяющего середины диагоналей AC и BD , и, следовательно, как точку пересечения любых двух из (или всех трех) отрезков KL, PQ, RS.

Способ 2 (систематический). Приняв за полюс произвольную точку O, используем для задания четырехугольника $ABCD$ минимальную информацию: O, $\vec{r}_A$, $\vec{r}_B$, $\vec{r}_C$, $\vec{r}_D$. Переписав (1) в виде

$$\vec{r}_A-\vec{r}_M+\vec{r}_B-\vec{r}_M+\vec{r}_C-\vec{r}_M+\vec{r}_D-\vec{r}_M=\vec{0},$$

находим, что

$$\vec{r}_M=\frac{1}{4}(\vec{r}_A+\vec{r}_B+\vec{r}_C+\vec{r}_D).$$

Существование и единственность искомой точки M доказаны. Для выяснения ее геометрического смысла достаточно учесть, что

$$\vec{r}_M=\frac{1}{2}\left(\frac{\vec{r}_A+\vec{r}_B}{2}+\frac{\vec{r}_C+\vec{r}_D}{2}\right)=$$
$$=\frac{1}{2}\left(\frac{\vec{r}_A+\vec{r}_D}{2}+\frac{\vec{r}_B+\vec{r}_C}{2}\right)=\frac{1}{2}\left(\frac{\vec{r}_A+\vec{r}_C}{2}+\frac{\vec{r}_B+\vec{r}_D}{2}\right).$$

Заметим, что в отличие от способа 1, способ 2 остается применимым при замене условия (1), скажем, на такое:

$$3\overrightarrow{MA}+7\overrightarrow{MB}-9\overrightarrow{MC}+\overrightarrow{MD}=\vec{0}.$$

Пример 2([6], задача 133). В треугольнике ABC проведены медианы AD, BE и CF. Вычислить

$$u=(\overrightarrow{BC},\overrightarrow{AD})+(\overrightarrow{CA},\overrightarrow{BE})+(\overrightarrow{AB},\overrightarrow{CF}).$$

Решение. Способ 1 (творческий). Имеем:

$$u=\left(\overrightarrow{BC},\frac{\overrightarrow{AB}+\overrightarrow{AC}}{2}\right)+\left(\overrightarrow{CA},\frac{\overrightarrow{BA}+\overrightarrow{BC}}{2}\right)+\left(\overrightarrow{AB},\frac{\overrightarrow{CA}+\overrightarrow{CB}}{2}\right),$$

что после использования линейности скалярного произведения по второму аргументу и приведения подобных членов приводит к ответу $u = 0$.

Способ 2 (систематический). Приняв за полюс точку A, используем для задания треугольника ABC следующую минимальную информацию: A, $\vec{r}_B$, $\vec{r}_C$. Тогда

$$u = \left(\vec{r}_C - \vec{r}_B, \frac{\vec{r}_B + \vec{r}_C}{2}\right) + \left(-\vec{r}_C, \frac{1}{2}\vec{r}_C - \vec{r}_B\right) + \left(\vec{r}_B, \frac{1}{2}\vec{r}_B - \vec{r}_C\right),$$

что после использования линейности скалярного произведения по каждому аргументу и приведения подобных членов приводит к ответу $u = 0$.

Способ 3 (творческий). Приняв за полюс точку M пересечения медиан AD, BE и CF, используем для задания треугольника ABC избыточную информацию: M, $\vec{r}_A$, $\vec{r}_B$, $\vec{r}_C$ (напомним, что $\vec{r}_A + \vec{r}_B + \vec{r}_C = \vec{0}$). Тогда

$$u = \left(\vec{r}_C - \vec{r}_B, -\frac{3}{2}\vec{r}_A\right) + \left(\vec{r}_A - \vec{r}_C, -\frac{3}{2}\vec{r}_B\right) + \left(\vec{r}_B - \vec{r}_A, -\frac{3}{2}\vec{r}_C\right),$$

что после использования линейности скалярного произведения по каждому аргументу и приведения подобных членов приводит к ответу $u = 0$.

Способ 4 (систематический). Приняв за полюс ту же точку M, что и в способе 3, используем для задания треугольника ABC минимальную информацию: M, $\vec{r}_A$, $\vec{r}_B$. Тогда

$$u = \left(-\vec{r}_A - 2\vec{r}_B, -\frac{3}{2}\vec{r}_A\right) + \left(2\vec{r}_A + \vec{r}_B, -\frac{3}{2}\vec{r}_B\right) + \left(\vec{r}_B - \vec{r}_A, -\frac{3}{2}(-\vec{r}_A - \vec{r}_B)\right),$$

что после использования линейности скалярного произведения по каждому аргументу и приведения подобных членов приводит к ответу $u = 0$.

Пример 3 ([6], задача 134). Дан прямоугольник $ABCD$ и точка M (которая может лежать как в плоскости прямоугольника, так и вне ее). Показать, что: 1) $(\overrightarrow{MA}, \overrightarrow{MC}) = (\overrightarrow{MB}, \overrightarrow{MD})$; 2) $MA^2 + MC^2 = MB^2 + MD^2$.

Решение. Способ 1 (творческий). Заметим, что для произвольных $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ имеет место эквивалентность

$$\begin{cases} \alpha = \beta, \\ \gamma = \delta \end{cases} \Longleftrightarrow \begin{cases} \alpha + \gamma = \beta + \delta, \\ \alpha - \gamma = \beta - \delta. \end{cases}$$

Если положить $\alpha = MA^2 + MC^2$, $\beta = MB^2 + MD^2$, $\gamma = 2(\overrightarrow{MA}, \overrightarrow{MC})$, $\delta = 2(\overrightarrow{MB}, \overrightarrow{MD})$, то останется доказать, что $\alpha + \gamma = \beta + \delta$ и $\alpha - \gamma = \beta - \delta$. Действительно,

$$\alpha + \gamma = (\overrightarrow{MA} + \overrightarrow{MC}, \overrightarrow{MA} + \overrightarrow{MC}) = (2\overrightarrow{MP}, 2\overrightarrow{MP}) = \\ = (\overrightarrow{MB} + \overrightarrow{MD}, \overrightarrow{MB} + \overrightarrow{MD}) = \beta + \delta,$$

где P — точка пересечения диагоналей прямоугольника $ABCD$, и

$$\alpha - \gamma = (\overrightarrow{MA} - \overrightarrow{MC}, \overrightarrow{MA} - \overrightarrow{MC}) = (\overrightarrow{CA}, \overrightarrow{CA}) = CA^2 = \\ = DB^2 = (\overrightarrow{MB} - \overrightarrow{MD}, \overrightarrow{MB} - \overrightarrow{MD}) = \beta - \delta.$$

Способ 2 (систематический). Если принять за полюс точку M и использовать для задания прямоугольника $ABCD$ и точки M минимальную информацию M, $\vec{r}_A$, $\vec{r}_B$, $\vec{r}_D$, то через нее условие „параллелограмм $ABCD$ является прямоугольником“ запишется в виде $(\vec{r}_B - \vec{r}_A, \vec{r}_D - \vec{r}_A) = 0$, неудобном для применения. Поэтому примем за полюс точку A и используем для задания прямоугольника $ABCD$ и точки M минимальную информацию A, $\vec{b}$, $\vec{d}$, $\vec{m}$, где $\vec{b} = \vec{r}_B$, $\vec{d} = \vec{r}_D$, $\vec{m} = \vec{r}_M$. Тогда условие „параллелограмм $ABCD$ является прямоугольником“ можно записать в виде $(\vec{b}, \vec{d}) = 0$. Имеем:

$$(\overrightarrow{MA}, \overrightarrow{MC}) = (-\vec{m}, \vec{b} + \vec{d} - \vec{m}), \qquad (\overrightarrow{MB}, \overrightarrow{MD}) = (\vec{b} - \vec{m}, \vec{d} - \vec{m}),$$
$$MA^2 + MC^2 = (\vec{m}, \vec{m}) + (\vec{b} + \vec{d} - \vec{m}, \vec{b} + \vec{d} - \vec{m}),$$
$$MB^2 + MD^2 = (\vec{b} - \vec{m}, \vec{b} - \vec{m}) + (\vec{d} - \vec{m}, \vec{d} - \vec{m}).$$

Преобразуя правые части полученных равенств с использованием линейности скалярного произведения по каждому аргументу, приводя подобные члены и учитывая условие $(\vec{b}, \vec{d}) = 0$, убеждаемся в справедливости доказываемых равенств.

Отметим, что показ студентам нескольких способов решения одной и той же задачи, кажется, становится традицией (см., например, [3–7]).

Список литературы

[1] *Бродский Г. М.* Аналитическая геометрия: Методические указания к проведению практических занятий. Ярославль: ЯрГУ, 1978. 12 с.

[2] *Бродский Г. М.* О задачах на составление уравнений множеств точек // Проблемы повышения эффективности образовательного процесса в высших учебных заведениях. Ярославль: ЯрГУ, 2004. С. 34–46.

[3] *Бродский Г. М.* Избранные вопросы аналитической геометрии и линейной алгебры: учеб. пособие. Ярославль: ЯрГУ, 2005. 136 с.

[4] *Ким Г. Д., Крицков Л. В.* Алгебра и аналитическая геометрия: Теоремы и задачи. Т.I. М.: ИКД „Зерцало-М“, 2003. 430 с.

[5] *Ким Г. Д., Крицков Л. В.* Алгебра и аналитическая геометрия: Теоремы и задачи. Т.II, ч.1. М.: ИКД „Зерцало-М“, 2003. 170 с.

[6] *Моденов П. С., Пархоменко А. С.* Сборник задач по аналитической геометрии. М.: Наука, 1976. 384 с.

[7] *Резниченко С. В.* Аналитическая геометрия в примерах и задачах (Алгебраические главы). М.: Изд-во МФТИ, 2001. 576 с.

О системном характере объектов, изучаемых в арифметике целых чисел и многочленов

Г. М. Бродский

При изучении арифметики целых чисел и многочленов в курсе „Геометрия и алгебра“ для студентов специальности „Прикладная математика и информатика“ предлагается подчеркивать системный в той или иной степени характер ряда понятий.

Библиография: 1 название.

В книге [1] Ю.И. Манин назвал „социологическим“ подход к математическому объекту, при котором он рассматривается не с точки зрения его внутренней структуры, а как член сообщества себе подобных, т.е. некоторой *системы*. Этот термин, употребленный при описании языка категорий, на наш взгляд, заслуживает, во-первых, более широкой трактовки, во-вторых, замены на более точный термин *системный*, и, в-третьих, использования в преподавании курса „Геометрия и алгебра“. Внутреннее и системное описания объекта дополняют друг друга, причем для многих объектов, изучаемых в геометрии и алгебре, системное описание играет особенно важную роль. Например, для понятия комплексного числа, допускающего различные внутренние описания (пара действительных чисел (a, b), матрица $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$ с действительными элементами, вычет $a + bx$ кольца многочленов $\mathbb{R}[x]$ по модулю $x^2 + 1$, и т.д.) на переднем плане оказывается вопрос об определении операций над комплексными числами. Не случайно принято давать определение поля комплексных чисел $\mathbb{C}$, а под комплексными числами понимать его элементы. Иногда внутреннее описание объекта вообще сводится к выбору обозначений. Например, элементы поля Галуа $\mathbb{F}_2$ обычно обозначают как 0 и 1. Однако с системной точки зрения они отличны от действительных чисел 0 и 1: ведь $1+1=0$ для $0, 1 \in \mathbb{F}_2$. Другими словами, поле $\mathbb{F}_2$ не является подполем поля $\mathbb{R}$ несмотря на справедливость включения множеств $\mathbb{F}_2 \subseteq \mathbb{R}$. Так что поле $\mathbb{F}_2$ не относят к числовым полям, а его элементы вряд ли стоит называть числами. Аналогично обстоит дело с полем (классов) вычетов $K[x]/p(x)$, где $p(x)$ — неприводимый многочлен над полем K. Его элементы можно записывать не как классы вычетов по модулю $p(x)$, а как сами вычеты по этому модулю. При этом вновь сталкиваемся с ситуацией, когда имеет место включение множеств $K[x]/p(x) \subseteq K[x]$, однако $K[x]/p(x)$ не является подполем кольца $K[x]$. На наш взгляд, нецелесообразно поля $K[x]/p(x)$ называть полями многочленов: ведь элементы поля $K[x]/p(x)$ и остатки от деления на $p(x)$ в кольце $K[x]$ (рассматриваемые как многочлены из $K[x]$) различны с системной точки зрения.

Представляется особенно полезным обращать внимание студентов на системный характер ряда понятий, изучаемых в разделе „Арифметика целых чисел и многочленов“ курса „Геометрия и алгебра“. Показу возможного использования этого приема в виде списка понятий, подчеркивание системного в той или иной степени характера которых может уберечь студентов от ошибок и помочь при решении задач, предпошлем несколько замечаний. Для краткости условимся коммутативные области целостности называть просто областями. В связи с тем, что области $\mathbb{Z}$ и $K[x]$, где K — поле, являются евклидовыми, арифметику целых чисел и арифметику многочленов в курсе „Геометрия и алгебра“ естественно изучать параллельно. Это может вызывать у некоторых студентов путаницу двоякого рода. Во-первых, как будет продемонстрировано ниже, целые числа как элементы области $\mathbb{Z}$ и целые числа как элементы области $\mathbb{R}[x] \supseteq \mathbb{R} \supseteq \mathbb{Z}$ могут вести себя различным образом (про их свойства, проявляющие себя таким образом, будем говорить, что они имеют *системный* характер). Во-вторых, в ситуации, когда K — подполе поля P (так что $K[x]$ — подкольцо кольца $P[x]$) определения объекта или свойства (в терминах, имеющих смысл как для целых чисел, так и для многочленов), примененные к многочленам из $K[x]$ и к тем же многочленам, рассматриваемым как элементы из $P[x]$, могут оказываться равносильными или нет (в последнем случае условимся говорить, что объект или свойство носит *сильно системный* характер). Существующие терминологические традиции (вместо термина „неприводимый многочлен“ используется более точный: „неприводимый многочлен над полем K“) решают возникающие при этом проблемы лишь частично.

Приведем упомянутый выше список.

1. Деление с остатком носит системный, но не сильно системный характер. Действительно, результаты деления с остатком 7 на 3 раз-

личны в $\mathbb{Z}$ и $\mathbb{R}[x]$: $7 = 3 \cdot 2 + 1$ в $\mathbb{Z}$ и $7 = 3(7/3) + 0$ в $\mathbb{R}[x]$. Но если $f, g \in K[x]$ и $g \neq 0$, то, как видно из соответствующего алгоритма, результаты деления с остатком f на g в $K[x]$ и $P[x]$ одинаковы (здесь и всюду ниже K — подполе поля P).

2. Отношение делимости носит системный, но не сильно системный характер. Действительно, $3 \nmid 7$ в $\mathbb{Z}$, но $3 \mid 7$ в $\mathbb{R}[x]$. Однако из пункта 1 легко вывести, что для любых $f, g \in K[x]$ утверждения „$f \mid g$ в $K[x]$“ и „$f \mid g$ в $P[x]$“ равносильны.

3. Свойство быть делителем единицы (обратимым элементом области) носит системный, но не сильно системный характер. Действительно, $3 \nmid 1$ в $\mathbb{Z}$, но $3 \mid 1$ в $\mathbb{R}[x]$. В то же время для любого $f \in K[x]$ утверждения „$f \mid 1$ в $K[x]$“ и „$f \mid 1$ в $P[x]$“ равносильны в силу пункта 2.

4. Отношение ассоциированности носит системный, но не сильно системный характер. Действительно, $3 \nsim 7$ в $\mathbb{Z}$, но $3 \sim 7$ в $\mathbb{R}[x]$. Далее, для любых $f, g \in K[x]$ утверждения „$f \sim g$ в $K[x]$“ и „$f \sim g$ в $P[x]$“ равносильны в силу пункта 2 (ведь два многочлена ассоциированы тогда и только тогда, когда они делят друг друга).

5. Для $a, b \in \mathbb{Z}$ (соответственно $a, b \in K[x]$) через ННОД(a, b) обозначим неотрицательный (соответственно нормированный) наибольший общий делитель a и b. Для $a, b \in \mathbb{Z}$ (соответственно $a, b \in K[x]$) нам будет удобно использовать наряду с ННОД(a, b) обозначение ННОД$_{\mathbb{Z}}(a, b)$ (соответственно ННОД$_{K[x]}(a, b)$). Понятие ННОД носит системный, но не сильно системный характер. Действительно, ННОД$_{\mathbb{Z}}(6, 9) = 3$, но ННОД$_{\mathbb{R}[x]}(6, 9) = 1$. Кроме того, ННОД$_{K[x]}(f, g) =$ ННОД$_{P[x]}(f, g)$ для любых $f, g \in K[x]$, так как применение алгоритма Евклида к $f, g \in K[x]$ и нормирование полученного результата не выводят за пределы $K[x]$.

6. Отношение взаимной простоты носит системный, но не сильно системный характер. Действительно, 3 и 6 не являются взаимно простыми целыми числами, но являются взаимно простыми многочленами из $\mathbb{R}[x]$. Однако для любых $f, g \in K[x]$ их взаимная простота в $K[x]$ равносильна взаимной простоте в $P[x]$ в силу пункта 5 (ведь $f, g \in K[x]$ взаимно просты в $K[x]$ тогда и только тогда, когда ННОД$_{K[x]}(f, g) = 1$).

7. Понятие простого (неприводимого) элемента области носит системный и даже сильно системный характер. Действительно, 3 — простой элемент области $\mathbb{Z}$, но 3 не является неприводимым многочленом над полем $\mathbb{R}$. Кроме того, многочлен $x^2 + 1$ неприводим над полем $\mathbb{R}$, но приводим над полем $\mathbb{C}$.

Напомним, что использование факта, отмеченного в пункте 5, необходимо при решении ряда задач на нахождение наибольшего общего делителя двух многочленов.

Список литературы

[1] *Манин Ю. И.* Лекции по алгебраической геометрии. Ч.1. Аффинные схемы. М.: Изд-во МГУ, 1970. 134 с.

Вариации на тему "Предел функции нескольких переменных"

Т. Г. Бычкова

Описывается фрагмент учебного процесса по математическому анализу.

Признаюсь, что приступаю к указанной в названии теме на практических занятиях всегда с тайным трепетом и сомнениями в душе. Ведь это, пожалуй, самая трудная часть курса. Изучая её и решая задачи, мы будем одновременно учиться доказывать разные утверждения. Получится ли? Я не знаю. Но я знаю, что мы должны пройти этот путь ...

Итак, начинаем, как обычно, с теории.

По аналогии с одномерным анализом есть два эквивалентных определения предела функции нескольких переменных в точке: по Коши (на $\varepsilon - \delta$ языке) и по Гейне (на языке последовательностей). Мы рассматриваем оба определения, делая акцент на втором. Приведем его.

Определение. *Пусть $E \subset \mathbb{R}^k$ $(k > 1)$ и a — предельная точка множества E. Число b называют пределом функции $f : E \to \mathbb{R}$ в точке a, если для любой последовательности z_n, сходящейся к a при $n \to \infty$ и удовлетворяющей условиям*

$$z_n \in E, \qquad z_n \neq a \qquad (n \in \mathbb{N}), \tag{1}$$

соответствующая последовательность значений функции $f(z_n)$ сходится к b.

Сформулировав определение, обращаю внимание своих учеников на то, что z_n и a — элементы k-мерного пространства $\mathbb{R}^k$ и условие "$z_n \to a$

при $n \to \infty$" $\Longleftrightarrow \sqrt{\sum\limits_{m=1}^{k} (z_n^{(m)} - a^{(m)})^2} \to 0$ при $n \to \infty$, что в свою очередь равносильно покоординатной сходимости. Далее предлагаю подумать над тем, как используя данное определение, можно доказать, что $\lim\limits_{z \to a} f(z)$ не существует. Совместными усилиями приходим к пониманию того, что для этого достаточно указать две последовательности z'_n, z''_n, сходящиеся к точке a при $n \to \infty$ и обладающие свойством (1), для которых

$$\lim_{n\to\infty} f(z'_n) \neq \lim_{n\to\infty} f(z''_n). \tag{2}$$

Замечаю,что подобную технику мы уже использовали на первом курсе, когда доказывали, что не существует $\lim\limits_{x\to 0} \sin\frac{1}{x}$, $\lim\limits_{x\to\infty} \cos x$, $\lim\limits_{x\to 0} e^{\frac{1}{x}}$ и т. д.

Теперь уже можно переходить к задачам. При этом, как правило, мы ограничиваемся функциями двух переменных. Итак, предлагаю доказать, что не существует $\lim\limits_{\substack{x\to 0\\ y\to 0}} f(x,y)$, где

$$f(x,y) = \frac{xy}{xy + (x-y)^2}. \tag{3}$$

После недолгих раздумий находим две последовательности

$$z'_n = (x'_n, y'_n) = \left(\frac{1}{n}, 0\right), \; z''_n = (x''_n, y''_n) = \left(\frac{1}{n}, \frac{1}{n}\right), \tag{4}$$

для которых

$$\lim_{n\to\infty} f(x'_n, y'_n) = 0, \; \lim_{n\to\infty} f(x''_n, y''_n) = 1. \tag{5}$$

Вроде бы все доказано. Но всем ли понятно? На всякий случай еще раз проговариваю, что завершить доказательство можно, рассуждая от противного: предположим, что $\lim\limits_{\substack{x\to 0\\ y\to 0}} f(x,y)$ существует и равен b. Тогда для любой последовательности $z_n = (x_n, y_n)$, сходящейся к точке $(0,0)$ и обладающей свойством (1),

$$\lim_{n\to\infty} f(x_n, y_n) = b$$

и, стало быть,

$$\lim_{n\to\infty} f(x'_n, y'_n) = \lim_{n\to\infty} f(x''_n, y''_n),$$

что противоречит (5). Следовательно, предел функции (3) в точке $(0,0)$ не существует.

В ряде случаев и при вычислении предела целесообразно переходить к последовательностям, например, при решении такой задачи: найти

$$\lim_{\substack{x\to 0\\ y\to 0}} (x^2+y^2)^{x^2y^2}. \tag{6}$$

Замечаем, что
$$(x^2+y^2)^{x^2y^2} = e^{x^2y^2\ln(x^2+y^2)}$$
и достаточно вычислить предел показателя, то есть
$$\lim_{\substack{x\to 0\\ y\to 0}} x^2y^2\ln(x^2+y^2). \tag{7}$$
Предлагаю самостоятельно дома доказать, что предел (7) равен нулю. Но на следующем занятии, как правило, рассказываю решение сама.

Итак, пусть $z_n = (x_n, y_n)$ — произвольная последовательность типа (1), сходящаяся к точке $(0,0)$. Используя неравенство Коши
$$x^2y^2 \leqslant \frac{(x^2+y^2)^2}{4}$$
и замену $t_n = x_n^2 + y_n^2$ (заметим, что $t_n \to 0$ при $n \to \infty$), выписываем цепочку неравенств:
$$x_n^2y_n^2|\ln(x_n^2+y_n^2)| \leqslant \frac{(x_n^2+y_n^2)^2}{4}|\ln(x_n^2+y_n^2)| = \frac{t_n^2}{4}|\ln t_n|. \tag{8}$$
Выражение, стоящее в (8) справа, стремится к нулю при $n \to \infty$. Это следует из того, что
$$\lim_{t\to +0} t^2\ln t = 0. \tag{9}$$
Прошу аудиторию объяснить данный факт. Задача — вычислительная и трудностей не вызывает. Кто-нибудь обязательно предложит переписать функцию $t^2\ln t$ в виде $\dfrac{\ln t}{t^{-2}}$ и найти предел (9) по правилу Лопиталя. Итак, $0 \leqslant x_n^2y_n^2|\ln(x_n^2+y_n^2)| \leqslant \frac{t_n^2}{4}|\ln t_n| \to 0$ при $n \to \infty$. Теперь всем уже понятно, что настал момент воспользоваться леммой о встречных последовательностях, более известной под названием леммы "о двух миллиционерах", и сделать вывод, что предел (7) равен нулю. Стало быть, исходный предел (6) равен 1.

Рассматриваемая тема получает развитие (а методика находит применение) и далее, при изучении таких понятий, как непрерывность и дифференцируемость. Пусть, к примеру, требуется доказать, что функция
$$f(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2}, & \text{если } x^2+y^2 \neq 0;\\ 0, & \text{если } x^2+y^2 = 0 \end{cases}$$
не является непрерывной в точке $(0,0)$.

Выясняем, что в данном случае можно и не доказывать, что $\lim\limits_{\substack{x\to 0\\ y\to 0}} f(x,y)$ не существует, а достаточно указать одну последовательность (x_n, y_n) типа (1), сходящуюся к точке $(0,0)$ при $n \to \infty$, для которой
$$\lim_{n\to\infty} f(x_n, y_n) \neq 0. \tag{10}$$

По аналогии с ранее решенными задачами пробуем уже знакомые варианты (4). Не проходят. После более или менее длительного перебора находим подходящую последовательность

$$(x_n, y_n) = \left(\frac{1}{n}, \frac{1}{n^2}\right).$$

Для нее условие (10) выполнено. Ведь

$$\lim_{n\to\infty} f(x_n, y_n) = \frac{1}{2}.$$

Значит, рассматриваемая функция не является непрерывной в точке $(0,0)$.

Понятие дифференцируемости функции нескольких переменных — одно из самых трудных в математическом анализе. Оно вызывает у студентов непонимание и много вопросов. Тем не менее, сформулировав определение, я перехожу к задачам, рассчитывая на приобретенный нами опыт.

Пусть надо доказать, что функция $f(x,y) = \sqrt{|xy|}$ не дифференцируема в точке $(0,0)$, хотя обе частные производные $f'_x(0,0)$ и $f'_y(0,0)$ существуют. Без труда вычисляем:

$$f'_x(0,0) = 0, \;\; f'_y(0,0) = 0. \tag{11}$$

А дальше? Жду какое-то время, а затем предлагаю воспользоваться методом от противного. Итак, предположим, что функция дифференцируема в точке $(0,0)$. Тогда

$$f(x,y) - f(0,0) = f'_x(0,0)x + f'_y(0,0)y + o(\sqrt{x^2+y^2}),$$

то есть, в силу (11),

$$\sqrt{|xy|} = o(\sqrt{x^2+y^2}). \tag{12}$$

И что это значит? Оказывается, знают не все. Еще раз повторяю, что (12) просто по определению означает, что

$$\lim_{\substack{x\to 0\\ y\to 0}} \frac{\sqrt{|xy|}}{\sqrt{x^2+y^2}} = 0. \tag{13}$$

Без труда доказываем, используя уже знакомые варианты последовательностей (4), что предел (13) не существует. Следовательно, функция не дифференцируема в $(0,0)$.

Следующие задачи: “являются ли дифференцируемыми в точке $(0,0)$ функции а) $f(x,y) = \sqrt[3]{xy}$; б) $f(x,y) = \sqrt[3]{x^3+y^3}$?” решаем легко, по аналогии.

Теперь пора предложить задачу потруднее: исследовать на дифференцируемость в точке $(0,0)$ функцию

$$f(x,y) = \begin{cases} e^{-\frac{1}{x^2+y^2}}, & \text{если } x^2+y^2 \neq 0; \\ 0, & \text{если } x^2+y^2 = 0. \end{cases}$$

Приходится подсказать. Предлагаю попробовать доказать, что

$$e^{-\frac{1}{x^2+y^2}} = o(\sqrt{x^2+y^2}). \tag{14}$$

И далее не вмешиваюсь, жду. Выписываем предел, находим замену, применяем правило Лопиталя. Получается цепочка:

$$\lim_{\substack{x\to 0\\ y\to 0}} \frac{e^{-\frac{1}{x^2+y^2}}}{\sqrt{x^2+y^2}} = \lim_{z\to+\infty} ze^{-z^2} = \lim_{z\to+\infty} \frac{z}{e^{z^2}} = \lim_{z\to+\infty} \frac{1}{2ze^{z^2}} = 0.$$

Следовательно, (14) верно, и функция дифференцируема в точке $(0,0)$, причем частные производные $f'_x(0,0) = f'_y(0,0) = 0$. “Почему?” — непременно спросит кто-нибудь. “По определению,” — отвечаю я.

Вот и все. Мы переходим к вычислительным задачам: дифференциалы, экстремумы, замена переменных и т. д. Учебный процесс продолжается.

Преподавание компьютерных дисциплин на кафедре вычислительных и программных систем

В. В. Васильчиков

Рассмотрены вопросы преподавания учебных дисциплин компьютерной направленности. Описана работа кафедры вычислительных и программных систем ЯрГУ, а также отмечены основные проблемы и пути их решения.

Учебная нагрузка кафедры вычислительных и программных систем почти целиком состоит из курсов компьютерного блока, которые включают и дисциплины общего профессионального ядра и специальные дисциплины. И если первые достаточно устоялись как в содержательном, так и методическом аспекте, то вторая группа дисциплин требует постоянного внимания и переработки. Это в основном дисциплины, посвященные современным информационным технологиям и научным направлениям, появившимся в этой области. За последние три года преподавателями кафедры был разработан целый ряд таких учебных курсов.

- "Программирование в MS Windows". Курс посвящен разработке программного обеспечения в среде программирования MS Visual C++ 6 с использованием библиотеки MFC. Читается для направления 510200 и специальности 010500 "Прикладная математика и информатика" в 7 семестре и специальности 010503 "Математическое обеспечение и администрирование ИС" в 6 семестре.

- "Программирование в сетях Windows". Курс посвящен разработке сетевых приложений для платформы Win32 с использованием интерфейсов программирования NetBIOS и Winsock и наиболее распространенных сетевых протоколов. Читается для направления 510200 и специальности 010500 "Прикладная математика и информатика" в 8 семестре и специальности 010503 "Математическое обеспечение и администрирование ИС" в 7 семестре.

- "Разработка приложений для .NET Framework" . Объем курса 140 часов. Курс состоит из четырех частей: "Язык программирования C#"; "Программирование в .NET Framework"; "Разработка Web-приложений с использованием ASP.NET"; "Разработка Windows-приложений в .NET Framework". Курс читается для группы специализации студентов математического факультета и как факультатив для студентов 4 курса факультета ИВТ.

- Два учебных курса, посвященных операционным системам семейства UNIX и их администрированию: "Сетевые ОС" для специальности 010502 "Прикладная информатика (в экономике)" в 7 семестре и "ОС семейства UNIX и их администрирование" для специальности 010500 "Прикладная математика и информатика"в 7–8 семестрах.

- Четыре учебных курса для студентов специальности 010503 "Математическое обеспечение и администрирование ИС", читающиеся в 7–8 семестрах: "Компьютерное моделирование", "Параллельное программирование", "Алгоритмы оптимизации и гибридные системы", "Рекурсивно-логическое программирование".

- Для студентов 4 курса специальности 010500 "Прикладная математика и информатика" была введена учебная дисциплина "Программирование в СУБД Oracle".

Кроме того, содержание целого ряда читавшихся ранее учебных дисциплин было в значительно мере переработано и актуализировано.

Вместе с тем существует множество компьютерных дисциплин, остающихся за пределами изучающихся на факультете учебных курсов, и в то же время весьма востребованных на рынке труда. Здесь можно назвать и администрирование Windows и сетей Windows, и владение современными технологиями автоматизации разработки программных проектов и моделей бизнес-систем, CASE-средствами (типа IBM Rational Rose), языком визуального моделирования UML и многое другое.

О причинах такого положения будет сказано ниже, а пока автору хотелось бы рассмотреть, как обстоят дела с реализацией минимального, с его точки зрения, набора дисциплин программистской направленности на факультете ИВТ.

На первом курсе студенты должны овладеть базовыми навыками алгоритмизации вычислительных процессов для решения экономических и расчетных задач, научиться пользоваться основными структурами данных, познакомиться с основными алгоритмами решения типовых задач, изучить один из языков высокого уровня (на факультете ИВТ это язык C), научиться основным приемам отладки и тестирования программ, изучить принципы разработки программ и отдельных программных модулей. При этом в качестве инструментального средства вполне подойдет достаточно простая среда разработки — turboC или BorlandC 3.1, так что на изучение самого инструмента потребуется не слишком много времени. Здесь учебные дисциплины и методика преподавания достаточно устоялись, количество часов представляется достаточным.

На втором курсе предполагается овладение концепцией объектно-ориентированного программирования и соответственно переход к языку C++. При этом важно четкое понимание, что объектно-ориентированный подход отнюдь не исключает подхода структурного, и умение сделать правильный выбор в зависимости от характера задачи. Здесь очень важно сформировать хорошую методику для овладения этим принципиально иным подходом к организации решения задачи, и в настоящий момент преподавателями кафедры ведутся такие разработки. В качестве инструмента на начальном этапе может быть использован тот же BorlandC 3.1 (и даже лучше, если он), в дальнейшем возможен переход к средам быстрой разработки приложений для Windows, например, Borland C++ Builder. После перехода к данному инструменту возможно и формирование начальных навыков работы с базами данных. Надо сказать, что степень понимания основных идей ООП и навыки объектно-ориентированного проектирования систем в настоящий момент у студентов недостаточны. Преподавателями кафедры проводятся методические разработки по улучшению преподавания данного компонента.

На третьем курсе автору видится уместным переход к более серьезным инструментам разработки Windows-приложений, например, MS VisualC++, детальное знакомство с механизмами работы ОС Windows и управления ими, освоение навыков разработки сетевых программ, изучение конкретных систем управления базами данных.

Далее, на четвертом и в начале пятого курса представляется целесообразным изучение платформы .NET, технологий автоматизации разработки программных проектов, Web-технологий, администрирования ОС Windows и Linux.

Но практическая реализация даже этой (минимальной!) программы упирается в ряд серьезных проблем.

1. Недостаток часов, отведенных в РУП под компьютерные дисциплины.

   На первых двух курсах часов, отведенных под основные дисциплины ("Информатика" и "Основы программирования" на первом курсе и "Системное и прикладное программное обеспечение" на втором) в принципе хватает на обозначенные цели. Начиная с третьего курса, основных дисциплин компьютерной направленности в РУП не предусмотрено вовсе, а объем спецдисциплин и курсов по выбору недостаточен, тем более что есть еще и ряд необходимых математических дисциплин. Поэтому, например, для того, чтобы прочитать на четвертом курсе специальности 010500 курсы "Программирование в MS Windows" и "Программирование в сетях Windows" пришлось убрать из РУП две ранее читавшихся дисциплины, а курс "Разработка приложений для .NET Framework" и вовсе читается факультативно. Кроме того, в последние годы обозначилась тенденция к уменьшению лабораторных часов под спецдисциплины (практически они сведены к нулю). Автор слабо себе представляет, как можно изучать программистские дисциплины, не подходя к компьютеру.

2. Отсутствие больших компьютерных аудиторий, оснащенных необходимыми техническими средствами.

   Проведение занятий по большинству из перечисленных предметов целесообразно проводить в компьютерных классах с проектором. Тем самым в несколько раз можно увеличить скорость изложения материала, а студенты должны иметь возможность сразу же опробовать тот или иной элемент на практике. Имеющиеся в настоящий момент компьютерные классы на 10–12 компьютеров недостаточны для полноценного проведения таких занятий. Оптимальным вариантом представляется класс с 25 компьютерами — в большем классе трудно держать под контролем всех учеников и помогать им в решении возникающим проблем. В настоящее время по вопросу о создании такого учебного класса принято положительное решение и уже начались необходимые работы.

Ясно, что в существующую систему планирования нагрузки (лекции и лабораторные занятия) проведение таких занятий в компьютерном классе не очень хорошо вписывается. Да и вообще, оптимальным было бы обучение каждому такому курсу проводить в течение полного рабочего дня за одну-две недели. Впрочем, это, по-видимому, пока нереально, поскольку потребует реорганизации всего учебного процесса.

3. Отсутствие квалифицированных преподавателей по многим из необходимых дисциплин.

   Изучить предмет по книгам (которых, как правило, нет) и потом проводить занятия со студентами нереально. Для того чтобы преподавать, скажем, администрирование сетей, надо как минимум пару лет отработать администратором. Чтобы со знанием дела говорить о конкретной технологии или среде разработки, надо иметь опыт практической работы с ней.

   Полученные однажды знания и навыки очень быстро устаревают, нужно приобретать новые, а потом еще методически подготовить курс и разработать необходимые материалы: слайды для лекций, лабораторные и экзаменационные задания, учебные и методические пособия (это при нагрузке 850–900 часов на ставку). А потом через 4–5 лет все повторить сначала.

   Сейчас отчасти проблема кадров решается за счет энтузиастов из числа наших выпускников, а ныне специалистов, которые соглашаются проводить занятия по актуальным курсам из любви к процессу, для души (не за тысячу же рублей зарплаты!). Часть занятий проводится также аспирантами и магистрантами.

   А что касается штатных преподавателей, необходимо предоставить им возможность пройти переподготовку (повышение квалификации). К сожалению, среди предлагаемых министерством вариантов повышения квалификации очень трудно найти курсы по современным компьютерным технологиям.

   Проблема нехватки квалифицированных преподавателей в области Computer Science, как можно понять из различных публикаций, не менее остро стоит во всем мире. И, вероятно, наиболее реальный способ ее преодоления заключается в использовании возможностей фирм и организаций, заинтересованных в трудоустройстве наших выпускников. Они могли бы участвовать в разработке учебных программ по спецдисциплинам, подбирать и оплачивать курсы по переподготовке и повышению квалификации для преподавателей университета, наконец, проводить занятия по некоторым учебным курсам. К сожалению, таких примеров пока очень мало.

Опыт использования технических средств для проведения занятий по компьютерным дисциплинам

В. В. Васильчиков

На основе опыта подготовки и проведения ряда учебных дисциплин рассматриваются преимущества использования современных технических средств для проведения занятий, формулируются основные требования к необходимому учебному и методическому обеспечению.

В преподавании дисциплин компьютерного блока, по-видимому, особое место занимают учебные курсы, посвященные изучению современных информационных технологий и соответствующих научных направлений. Это обусловлено целым рядом причин.

- Во-первых, это очень высокая востребованность знаний и практических навыков в данных областях. Причем требования потенциальных работодателей к уровню владения материалом и широте его охвата весьма и весьма высок. Автору не раз доводилось выслушивать от представителей фирм и организаций, где работают наши выпускники, сетования на то, что они (выпускники) не обладают знанием той или иной конкретной технологии или конкретной среды разработки. К сожалению, разнообразие технологий и инструментальных средств настолько велико, что удовлетворить всем требованиям (иногда довольно специфическим), по-видимому, невозможно. Однако очевидно, что кроме общей компьютерной и программистской культуры выпускники должны владеть и наиболее распространенными и востребованными на рынке технологиями и средствами разработки.

- В области информационных технологий, как ни в какой другой, очень быстро происходят изменения как по естественным причинам, так и вследствие рыночной конъюнктуры. Постоянно появляются новые подходы к хранению и обработке информации, соответствующие инструментальные средства и программные комплексы. Вместе с тем научная и методическая литература не успевает за стремительными изменениями, происходящими в этой области. Зачастую к моменту выхода книг по той или иной тематике (а для их издания на русском языке требуется еще не меньше года), появляются новые версии, а то и вовсе новые технологии. А ведь нужно время для их освоения преподавателями и подготовки учебных курсов и всех сопутствующих материалов. Впрочем, об этом мы чуть подробнее поговорим ниже.

- Наконец, учебной работой в данной области готовы заниматься далеко не все преподаватели как математического факультета, так и факультета ИВТ. В результате значительная часть этих курсов поручается молодым преподавателям и сотрудникам, аспирантам. А именно среди них (при условии высокой квалификации в области информационных и компьютерных технологий) текучесть кадров особенно высока. Причины этого положения дел достаточно ясны, и с ними сталкиваются во всем мире. Однако нам хотелось бы, чтобы результатами учебно-методических разработок преподавателей можно было бы воспользоваться и после прекращения их педагогической деятельности.

В данном докладе автору хотелось бы сказать о том, как можно было бы улучшить преподавание упомянутых учебных курсов за счет применения различных технических и программных средств, какие при этом проблемы решаются и какие возникают.

Этот доклад является неким обобщением опыта, полученного при подготовке и проведении занятий по ряду учебных курсов за последние два года. За это время автором были подготовлены и прочитаны курсы "Программирование в MS Windows" и "Программирование в сетях Windows". Первый посвящен разработке программного обеспечения в среде программирования MS Visual C++ 6 с использованием библиотеки MFC, второй — разработке сетевых приложений для платформы Win32. Кроме того, был подготовлен и прочитан 140–часовой учебный курс "Разработка приложений для .NET Framework". При его подготовке использовались материалы учебных курсов Microsoft, что, в частности, позволило сделать некоторые выводы о методике преподавания.

Для чтения всех курсов использовались презентации в формате MS Power Point и проектор. С одной стороны, это дает существенные преимущества, с другой, порождает некоторые дополнительные проблемы.

Остановимся на них чуть подробнее. К преимуществам следует отнести следующие моменты.

- Резко возрастает скорость изложения материала, минимум в 3–5 раз. Это очень существенный момент, поскольку объем материала, как правило, весьма значителен, а количество учебных часов невелико.

  Хотя, по-видимому, преимуществом это является при работе с достаточно подготовленными студентами. Вряд ли такая скорость требуется при работе с первокурсниками. Кроме того, несколько снижается обратная связь с аудиторией, а при преподавании на младших курсах необходимо оперативно реагировать на вопросы, недостаточное понимание, и т. п. Тут уж без доски обойтись трудно.

- Появляется возможность продемонстрировать и обсудить объемные и достаточно сложные куски программного кода, для чего обычная доска совершенно не пригодна. Кроме того, с помощью проектора можно сразу продемонстрировать приемы работы в соответствующей инструментальной среде.

- Поскольку материалы лекции готовятся в электронной форме, их удобно в дальнейшем использовать для подготовки учебных пособий и раздаточного материала для самостоятельной работы студентов. Это также очень важный момент, поскольку без самостоятельной работы, тщательно спланированной преподавателем, такой большой объем информации усвоить почти невозможно.

- Представленные проекты и программный код немедленно могут быть протестированы студентами (предполагается, что занятия проводятся в компьютерном классе).

К возможным проблемам и неудобствам можно отнести следующее.

- Большой объем подготовительной работы. К примеру, при подготовке курса "Программирование в MS Windows" у автора на подготовку всех материалов для двухчасового занятия уходило не менее 8-10 часов. Кроме собственно, презентаций требуется подготовить заготовки учебных проектов и подробное описание работы с ним (добавление тех или иных возможностей и функциональности) по каждой рассматриваемой теме. Это особенно важно, так как не все студенты успевают выполнить необходимые упражнения во время занятий, в этом случае они могут проделать это дома.

- Для полноценной работы необходим достаточно высокий уровень квалификации студентов. В частности, для упомянутого курса предполагалось, что студенты владеют навыками самостоятельной установки программных комплексов, настройки необходимых параметров операционной системы, имеют навыки работы с командной строкой, поскольку часть утилит предполагают именно такой интерфейс.

  Здесь можно упомянуть еще одну проблему, не имеющую прямого отношения к теме доклада, но связанную с преподаванием современных информационных технологий. При изучении Visual C++, например, невозможно выполнить ряд учебных проектов, связанных с использованием технологий OLE и ActiveX, поскольку они предполагают доступ к системному реестру, а у студентов, естественно, таких привилегий нет. При изучении среды программирования .NET эта проблема (недостаток привилегий) возникает еще чаще. Частично ее удалось решить совместными усилиями системных администраторов и автора за счет некоторых обходных маневров, однако полностью снять можно только, если Microsoft задумается о необходимости проведения занятий в учебных классах, где невозможно предоставить всем студентам права администраторов. Кстати, судя по публикациям в Internet, у американских университетов проблемы такие же.

- Большой объем материала воспринимается на слух тяжеловато, а качественное конспектирование затруднительно. Поэтому преподавателю следует сразу позаботиться о подготовке и предоставлении студентам учебных материалов в электронной форме. Особенно это важно в первые годы проведения занятий по данному курсу, пока не вышло соответствующее учебное пособие. А подготовить его следует не позднее, чем курс будет прочитан в первый раз.

- Маленькие компьютерные аудитории плохо приспособлены для проведения занятий в такой форме. Необходимо иметь классы примерно на 25–30 компьютеров. Больше, по-видимому, тоже нецелесообразно, поскольку трудно контролировать процесс выполнения учебных упражнений.

В настоящий момент для перечисленных учебных курсов разработаны почти все необходимые учебные, демонстрационные и методические материалы. Подготовлены два учебных пособия, одно из которых уже издано. Начата работа по разработке некоторых вспомогательных программных средств, которые преподаватель может использовать при проведении занятий в компьютерных классах: программ для автоматической регистрации студентов и ведения журнала посещаемости, а также для проведения тестирования и зачетных мероприятий.

Методические вопросы изложения связи программы и подпрограммы в компьютерных курсах

О. В. Власова, Н. Б. Чаплыгина

Одной из сложных для понимания студентами тем в курсах "Информатика" и "Языки программирования" является организация связи между программами. Предлагается вариант изложения этой темы, использующий программирование на низком уровне. Приводится пример программы на Ассемблере, позволяющий понять все стороны управления и обмена информацией, сопровождающих процесс такой связи.

Опыт преподавания дисциплин, связанных с приобретением навыков программирования на языке высокого уровня для студентов младших курсов свидетельствует о том, что одним из наиболее трудных моментов для понимания является организация связи между вызывающей и вызываемой функциями. Главной составляющей в организации такой связи является передача информации в вызываемую функцию и возвращение информации в вызывающую функцию. Значительная часть студентов допускает ошибки, связанные с организацией связи программы с подпрограммой.

Типичными являются такие

а) синтаксические ошибки, как обозначение идентификаторов формальных параметров в заголовке функции без соответствующих спецификаций, несоответствие списков формальных и фактических параметров при вызове функции;

б) семантические ошибки, связанные с преобразованием данных при несоответствии типа фактического параметра формальной спецификации в описании функции;

в) логические ошибки, порожденные неверно выбранной формой передачи параметров: значением или ссылкой, или неумением разделить данные на две категории: параметры и рабочие (локальные) переменные.

Данные определяются с помощью операторов-описаний, например:

int count, k;   float fl;   long mas[10];

При выполнении этих операторов-описаний переменным *count, k, fl* и *mas* выделяется оперативная память, после чего имена этих переменных могут быть использованы в конструировании других операторов или выражений. Важно дать понять студентам, что с каждым именем переменной связаны две характеристики: значение переменной, т. е. содержимое оперативной памяти, и ее адрес. При использовании имени переменной в выражении в одном случае используется ее значение, в другом — ее адрес, а в третьем — и то, и другое. Это зависит от контекста, в котором используется данный идентификатор. При передаче параметров это различие в использовании переменных становится существенным.

Приведем типичный пример ошибочного описания функции, которая должна обменять значениями две целочисленные переменные:

```
void change(int a, int b)
{ int r=a; a=b; b=r;
  return;
}
```

После выполнении оператора-вызова функции *change(a, b)* студент с удивлением обнаруживает, что переменные *a* и *b* сохранили свои исходные значения. Приведенное описание формальных параметров означает передачу параметров по значению, т. е. при входе в функцию change для каждого из описанных таким образом параметров появляется его "двойник" — локальная переменная, в которую копируется значение фактического параметра, и все операции производятся с этой локальной переменной, что никак не отражается на передаваемом фактическом параметре. Для того, чтобы добиться цели, необходимо объявить параметры, от которых ожидается изменение значения (выходные параметры), параметрами, передаваемыми адресом или, иначе, по ссылке. Заголовок функции в этом случае должен быть таким

void change(int& a, int& b)

Постигнуть разницу между двумя описанными формами передачи параметров студенту помогает освоение организации связи программы и подпрограммы на Ассемблере и затем связи программы на языке высокого уровня и подпрограммы на Ассемблере. Приведем пример организации такой связи.

Постановка задачи.

1) Программа на языке C++ передает двумерный целочисленный массив ассемблерной функции с тем, чтобы получить от нее линейный

массив, составленный из максимальных по столбцам элементов исходного массива. 2) Вызываемая функция формирует требуемый выходной линейный массив и его размер и, в учебных целях, передает размер выходного массива как возвращаемое значение целого типа. Заголовок вызываемой функции max_as :

extern "C"int max_as(int n, int m, int a[N][N], int& k, int p).*
Входные целочисленные параметры n и m — размеры исходной матрицы, параметр массивового типа — исходная матрица. Выходные параметры: целое k — число элементов в полученном линейном массиве, p — его адрес.

Описания переменных и обращение к функции:

```
#define N 20
   int n,m;       // размеры исходной матрицы
   int a[N][N];   // исходная матрица
   int k;         // размер выходного массива
   int p[N];      // выходной массив
int k_ax = max_as (n, m, a, k, p); // обращение к функции
```

При обращении к функции информация о параметрах и точке возврата помещается в стек. Порядок сохранения параметров является обратным порядку перечисления в списке заголовка функции, т. е. первый по списку параметр оказывается на вершине стека. Благодаря такому механизму в C++ можно создавать функции с переменным числом параметров. Передавая параметр, нужно точно представлять, какую информацию помещаем в стек: адрес или значение. По адресу параметра можно получить и его значение, но зная значение параметра, нельзя установить его адрес. Для входных параметров значение адреса является излишним, кроме того, передача по значению защищает параметр от несанкционированных изменений. Во многих случаях достаточно передать лишь значение параметра, что упрощает обращение к нему в вызываемой программе. Для выходных же параметров для их формирования необходимы адреса. Адреса передаются и для входного параметра массивового типа, поскольку его значение занимает большой объем памяти, в этом случае передача значения потребует дополнительных ресурсов времени и памяти. Скалярный параметр, имеющий один из основных предопределенных типов, помещается в стек своим значением, (например, $int\ n,\ int\ m$). Если параметр является указателем или ссылкой, то в стек также помещается его значение, но в этих случаях значением является адрес памяти (например, $int\&\ k,\ int*\ p$). Для функций дальнего вызова адреса или указатели помещаются в стек двойным словом: сегментным адресом и смещением параметра. Возвращаемое функцией значение может быть скалярного типа. В этом случае оно возвращается через регистры: $al,\ ax,\ dx : ax,$ — в зависимости от их объема памяти. В нашем примере возвращаемое значение имеет тип int, поэтому вызывающая функция будет ожидать его в регистре ax. Затем в стек помещается адрес точки возврата — адрес следующей за вызовом выполняемой команды: для близкого вызова — только смещение, для дальнего вызова — сегментный адрес и смещение и производится передача управления в функцию. После получения управления из функции необходимо очистить стек.

Обязанности вызываемой функции таковы: получив управление, позаботиться о сохранении информации о регистрах; получить через стек доступ к входным и выходным параметрам; выполнить необходимые действия; восстановить регистры из стека; вернуть управление на точку возврата.

Приведем исходный текст вызываемой функции $\_max\_as$, которая формирует линейный массив, составленный из максимальных по столбцам элементов переданного ей массива, и находит его длину.

```
nconst equ 20
_max_as proc far ; описание процедуры дальнего вызова
  push bp   ; сохранение регистров в стеке
  mov bp,sp ;  фиксируем точку в стеке
  push bx
  push ds
  push di
  push si
  push es
; выбор информации о параметрах из стека
    mov dx,6[bp]   ; параметр n
    mov cx,8[bp]   ; параметр m
    lds bx,10[bp]   ; вектор адреса a - в регистрах ds:bx, адрес  a[0][0]
    les si,18[bp]    ; вектор адреса p -в регистрах es:si, адрес p[0]
  cikl: push cx
  ; обработка очередного столбца
    mov cx,6[bp]    ; n
    dec cx             ; число сравнений
    mov di,bx        ; адрес начала столбца
    mov ax,word ptr[di]  ; в регистре ax "отсеиваем"максимум
    c1: add di,nconst*2  ; адрес следующего элемента столбца
      cmp ax,word ptr [di]
      jge con
      mov ax,word ptr [di]   ; найден больший элемент
    con: loop c1           ; цикл по столбцу
    ; сейчас в ах - максимальный элемент просмотренного столбца
    mov word ptr es:[si],ax    ;  адрес p[si]
    add si,2              ; адрес следующего элемента вектора
    add bx,2             ; адрес следующего столбца массива
    pop cx
  loop cikl              ; цикл по столбцам
```

```
sub si,word ptr 18[bp]    ; вычисление длины массива p
shr si,1                  ;значение для параметра k
mov ax,si
les di,14[bp]             ; адрес параметра k
mov word ptr es:[di],ax   ; помещаем k по своему адресу
pop es   ; восстанавливаем регистры из стека
pop si
pop di
pop ds
pop bx
pop bp
ret; корректировать стек - обязанность вызывающей программы
_max_as endp
```

Атрибут far в заголовке процедуры передает команде run информацию о дальнем типе перехода. После получения управления функция сохраняет все используемые регистры в стеке. Необходимо оставить содержимое всех регистров в прежнем состоянии, если действия функции не направлены специально на формирование их содержимого. Этому правилу должны следовать все вызываемые функции и процедуры. В большинстве случаев для формирования выходного значения используется регистр ax. В нашем примере предполагается его изменение: в конце работы функции он должен содержать возвращаемое значение. Поэтому его нет среди сохраняемых регистров.

Содержимое стека при обращении к параметрам

Смещение в стеке	Содержимое стека
	здесь можно разместить локальные переменные
+0	bp
	сегмент и смещение точки возврата
+6	значение n
+8	значение m
+10	сегмент и смещение a
+14	сегмент и смещение k
+18	сегмент и смещение p

Параметры, передаваемые значением, копируются из стека. Например, параметр n копируется в регистр dx командой $mov\ dx, 6[bp]$. Параметр, передаваемый адресом, выбирается из стека так, чтобы можно было использовать его адрес для дальнейшей работы. Например, команда $les\ si, 18[bp]$ позволяет в дальнейшем обращение к параметру с помощью адреса $es : si$. По окончании работы функция восстанавливает содержимое всех регистров, кроме ax, и передает управление командой ret по адресу возврата, который извлекается из стека. Удалить параметры из стека по правилам языка C++ предстоит вызывающей программе.

После отладки программы на низком уровне студент лучше усваивает организацию передачи информации между программными функциями.

Какие задачи допустимы на олимпиадах по математике и информатике

С. Г. Волчёнков

На примере двух задач, предлагавшихся на олимпиадах по программированию, сделаны некоторые выводы об особенностях олимпиад по математике и информатике.

Данная статья не претендует на полноту освещения вопроса, вынесенного в название. В ней просто приведены две задачи, предлагавшиеся автором на олимпиадах по информатике. Первая — на заключительном этапе Всероссийской олимпиады школьников, вторая — в полуфинале студенческого командного чемпионата мира по программированию ACM. Обе эти задачи объединяет не только их принадлежность к классической теории чисел и общая проблематика, связанная с соотношением *чистая математика — информатика*, но также и то, что такие задачи совершенно недопустимы в олимпиадах по математике. В первые годы становления информатики в ней в основном решались математические задачи, так как первыми информатиками, естественно, стали математики. Затем в этой области появилось много задач и из других разделов науки, техники и просто здравого смысла. Сейчас сложилась ситуация, когда на различных соревнованиях по информатике стараются даже не давать задач, опирающихся на слишком специальные знания из других областей, в частности, математики. Возникает угроза потери для информатики ряда красивых идей и методов, разработанных в смежных науках. Приводимые ниже задачи и комментарии к ним должны показать, что далеко не всегда специальные математические знания должны рассматриваться как препятствие дать данную задачу на обычных школьных или студенческих соревнованиях по информатике.

Задача 1. Разбиение на пары с простой суммой.

Для заданного натурального числа N требуется разбить числа $1, 2, 3, \ldots, N$ на максимально возможное число пар с простой суммой.

$N < 10000000$.

Пример такого разбиения при $N = 8$: (1, 2), (3, 4), (5, 8), (6, 7).

Сначала небольшой математический комментарий. В теории чисел известен факт, называемый постулатом Бертрана, согласно которому для любого натурального N между N и $2N$ всегда существует простое число. Указанный факт какое-то время оставался недоказанным, отсюда и закрепившееся за ним название постулата. Благодаря работам русского математика П. Л. Чебышёва постулат Бертрана был доказан. Сейчас он известен не только профессиональным математикам, но и многим школьникам, однако считается достаточно сложным, и даже на математических олимпиадах не принято давать задач, опирающихся в своих решениях на него. С помощью постулата Бертрана наша задача решается легко. Отметим сначала, что если N нечётно, то одному числу пары заведомо не будет и одно число можно отбросить. Так вот. отбросить можно само N, так как для чётного N (это будет показано ниже) все числа от 1 до N разбиваются на нужные пары. В самом деле, рассмотрим то самое простое чисто , которое должно быть между N и $2N$. Образуем первую пару, дающую простую сумму: N и $p - N$. Заметим, что такую же сумму дают и другие пары: $N - 1$ и $(p - N) + l$. $N - 2$ и $(p - N) + 2$, и т. д. Множество чисел от $p - N$ до N оказалось разбитым на нужные пары. Поэтому далее процесс можно рекурсивно продолжить для оставшихся чисел от 1 до $p - N - 1$. Отметим, что число $p - N - 1$ опять-таки чётно, что требуется для продолжения процесса.

Описанный алгоритм при $N = 8$ сработает следующим образом.

Находим первое простое число, большее 8, но меньшее 16. Это 11. Составляем пары чисел, дающие в сумме 11: (8, 3), (7, 4), (6, 5).

Продолжаем рекурсивно решать задачу для оставшихся четырёх чисел.

Первым простым числом большим 4 является 5. Пары чисел, дающие в сумме 5: (4, 1), (3, 2).

На этом все числа исчерпаны.

Казалось бы, предложенное решение непосредственно опирается на постулат Бертрана, и давать данную задачу школьникам нельзя. Но так ли это? А что изменилось бы, если бы постулат Бертрана не был бы верен? Ответ прост. Число N не сможет ни с одним из имеющихся чисел дать простую сумму и, значит, заведомо должно быть исключено из рассмотрения, а процесс надо продолжать для оставшихся чисел. Алгоритм же поиска нужного простого числа и алгоритм установления, что такого числа нет, очевидно, один и тот же. Итак, единственно зачем оказался нужным сам постулат — это приятное удивление от того, что задача всегда решается до последнего числа!

Задача 2. Разбиение на группы с простой суммой.

Для заданного натурального числа $N > 1$ требуется разбить множество чисел $1, 2, 3, \ldots, N$ на минимально возможное число групп с простой суммой.

$N < 1000000$.

Примеры.

$N = 4$. Искомое разбиение на две группы: (1, 2) и (3, 4).

$N = 6$. Тоже две группы: (2) и (1, 3, 4, 5, 6).

Математический комментарий к этой задаче, по форме очень похожей на предыдущую, опирается на гораздо более серьезный математический факт. Более того, на недоказанную гипотезу Х. Гольдбаха, дополненную Л. Эйлером, утверждающую, что любое чётное число представимо в виде суммы двух простых, а также на теорему И. М. Виноградова о представлении достаточно большого нечётного числа в виде суммы трех простых.

Небольшой анализ задачи. Сумма первых N натуральных чисел равна $S = \frac{N(N+1)}{2}$, что хорошо известно любому школьнику, начавшему изучать арифметические прогрессии. Из этой формулы легко заметить, что для всех $N > 2$ сумма всех чисел до N включительно не является простым числом и, значит, количество искомых групп более одной. А вот далее вступают в действие вышеупомянутые математические факты. Если S — чётное чисто, то, по гипотезе Гольдбаха, количество групп равно двум. Если же S — нечётно, то, по теореме Виноградова, групп должно быть не более трех. Остается заметить, что их может быть и две, если $S - 2$ оказывается простым. В другую группу при этом попадает одно простое число 2.

Опять же зададим себе вопрос, как решать задачу, если мы не знаем ни гипотезы Гольдбаха (которая, может быть, еще и не верна), ни теоремы Виноградова. Оказывается, что решать задачу всё равно надо будет так же.

Сначала находим S по формуле. Кстати, если школьник забыл эту формулу или не знал её, он может успешно вычислить S, просуммировав N чисел на имеющемся в его распоряжении компьютере. Далее, определяем чётность S. В случае нечётного S вычитаем из S двойку и проверяем $S-2$ на простоту. Заметим, что это единственный шанс получить результат, состоящий из двух групп, так как при любом разбиении нечётного числа на два слагаемых одно из них будет чётным, т. е. может оказаться простым, только если равно 2. Если результат является простым числом, то задача решена. Если нет, то вычитаем 3 (или любое другое нечётное число, не превышающее N, но 3 — естественнее) и получаем заведомо чётное число $S - 3$, которое, как и в случае чётного S. надо разделять на две группы. Почему на две, а не больше? Да потому, что требуется найти наименьшее число групп и поиск надо начинать с меньших значений. То, что поиск на этом и закончится (а он закончится, так как гипотеза Гольдбаха проверена для всех чисел, которые могут быть заданы в тестах, и даже в миллионы раз больших!), явится для школьника приятным подарком. Остается порекомендовать не забыть в программе отдельно рассмотреть случай $N = 2$ — единственный случай простоты числа S.

Итак, рассмотрены две задачи, которые, безусловно, придуманы на основании сложных математических фактов. Для решения же этих задач, по большому счету, не нужны ни знания этих фактов, ни сами эти факты. Может быть, приведённые задачи являются уникальными в этом смысле? Отнюдь нет.

Мы уже давно привыкли искать совершенные паросочетания в двудольном графе, не заботясь о выполнении условий теоремы Холла (по ходу дела проверится!) или строить максимальные потоки, не оценивая минимальные разрезы, производить перебор вариантов, не только не оценив заранее объем этого перебора, но даже зная о теоретической необозримости этого перебора (игра в шахматы, и не только). И здесь следует заявить — информатика должна решать массу задач. Хорошо, если эти задачи решаются с помощью математических или иных точных методов. Но если это не так, то задачи всё равно надо решать, и решения задач далеко не всегда могут опираться на исследованные другими науками кочки.

Ни одна наука не может в абсолютной мере разрешать или запрещать те или иные действия. Говорят, что майский жук, согласно уравнениям аэродинамики, не может летать. К счастью, *жук этих уравнений не знает — и летает!*

Некоторые особенности преподавания темы "Уравнения с частными производными первого порядка"

С. Д. Глызин, Е. П. Кубышкин

Предлагается план изложения темы, указанной в названии.

Знание уравнений с частными производными первого порядка требуется студентам третьего курса при изучении курса “Уравнения с частными производными“ или курса “Уравнения математической физики“, каждый из которых начинается, как правило, с классификации уравнений второго порядка, изложение которой предполагает знание теории уравнений с частными производными первого порядка. В силу того, что построение решений уравнений с частными производными первого порядка, изучаемых в курсе, сводится к интегрированию систем обыкновенных дифференциальных уравнений, указанную тему целесообразно излагать в курсе обыкновенных дифференциальных уравнений сразу после темы “Первые интегралы“. Авторам представляется полезным следующий план изложения обсуждаемой темы.

Общие понятия. Задача Коши. Рассматривается случай, когда искомая функция зависит от двух независимых переменных $(x, y) \in \Omega$, где Ω – область на плоскости. Общий вид уравнения первого порядка

$$F(x, y, u, u_x, u_y) = 0, \tag{1}$$

где $F(\cdot)$ — гладкая функция своих переменных. Здесь и ниже под гладкой функцией понимается функция класса C^1 своих переменных. Под решением уравнения (1) в области Ω понимается функция $u(x, y) \in C^1(\Omega)$, обращающая уравнение (1) в тождество. Обозначим через γ гладкую кривую, лежащую в Ω. Пусть на γ задана гладкая функция $g(x, y)$. Задача Коши для уравнения (1) ставится следующим образом: требуется построить решение уравнения (1), удовлетворяющее условию $u(x, y)\big|_\gamma = g(x, y)$.

Почти линейные уравнения. Рассмотрение начнем с уравнения

$$a(x, y)u_x + b(x, y)u_y = 0, \tag{2}$$

где $a(x,y), b(x,y)$ — гладкие в Ω функции. Это уравнение используется при классификации уравнений второго порядка. Рассмотрим в Ω систему обыкновенных дифференциальных уравнений

$$\frac{dx}{dt} = a(x,y),\ \frac{dy}{dt} = b(x,y), \tag{3}$$

которую назовем характеристической системой уравнения (2). Здесь t — некоторый параметр. Интегральные кривые системы (3), расположенные в Ω, называются характеристиками уравнения (2).

Гладкую функцию $\varphi(x,y)$, определенную в Ω, будем называть первым интегралом системы уравнений (3), если она постоянна на любом ее решении.

Утверждение 1. *Для того, чтобы функция $\varphi(x,y)$ была решением уравнения (2), необходимо и достаточно, чтобы она являлась первым интегралом системы (3).*

Отметим, что $\mathfrak{F}(\varphi(x,y))$, где $\mathfrak{F}(\cdot)$ некоторая гладкая функция, также является первым интегралом системы (3).

Построим решение задачи Коши для уравнения (2). Пусть $\gamma = \{x(\alpha), y(\alpha)\}$ — гладкая кривая, лежащая в Ω, на которой задана гладкая функция $g(\alpha)$. Пусть для некоторого α_0

$$\det\begin{pmatrix} x'(\alpha_0) & a(x_0,y_0) \\ y'(\alpha_0) & b(x_0,y_0) \end{pmatrix} \neq 0, \quad x_0 = x(\alpha_0),\ y_0 = y(\alpha_0). \tag{4}$$

Условие (4) означает, что в точке (x_0, y_0) (и некоторой ее окрестности) кривая γ не касается характеристик уравнения (2). Функция $u = g(\varphi_*^{-1}(\varphi(x,y))$ определяет единственное решение уравнения (2) в окрестности (x_0,y_0), проходящее через $g(\alpha)$. Здесь $\varphi_*(\alpha) = \varphi(x(\alpha), y(\alpha))$. Для дальнейшего построения решения вдоль γ необходимо воспользоваться методом продолжения.

Рассмотрим общее почти линейное уравнение

$$a(x,y)u_x + b(x,y)u_y = c(x,y,u), \tag{5}$$

Здесь $c(x,y,u)$ — гладкая функция своих переменных, определенная при $(x,y) \in \Omega$, $|u| \leqslant M$.

Утверждение 2. *Пусть на гладкой кривой $\gamma = \{x(\alpha), y(\alpha)\}$, принадлежащей Ω, задана гладкая функция $g(\alpha)$ и для некоторого α_0 выполнено условие (4). Тогда в некоторой окрестности точки (x_0, y_0) существует единственное решение уравнения (5), обращающееся на γ в $g(\alpha)$.*

Квазилинейные уравнения. Уравнения вида

$$a(x,y,u)u_x + b(x,y,u)u_y = c(x,y,u), \tag{6}$$

где $a(\cdot), b(\cdot), c(\cdot)$ — гладкие функции своих переменных при $(x,y) \in \Omega$, $|u| \leqslant M$, называется квазилинейным уравнением первого порядка.

Будем искать решение уравнения (6) в виде неявной функции

$$\varphi(x, y, u) = C,$$

где C — произвольная постоянная. С учетом равенств $u_x = -\varphi_x/\varphi_u$, $u_y = -\varphi_y/\varphi_u$, имеем для определения $\varphi(x, y, u)$ уравнение

$$a(x, y, u)\varphi_x + b(x, y, u)\varphi_y + c(x, y, u)\varphi_u = 0. \qquad (7)$$

Уравнение (7) отличается от (2) лишь тем, что искомая функция и коэффициенты уравнения зависят от трех переменных. Введем характеристическую систему

$$\frac{dx}{dt} = a(x, y, u),\ \frac{dy}{dt} = b(x, y, u),\ \frac{du}{dt} = c(x, y, u) \qquad (8)$$

и будем рассматривать ее в области $(x, y, u) \in \Omega \times [-M, M]$. Интегральные кривые системы (8), принадлежащие указанной области, будем называть характеристиками уравнения (7). Обозначим через $\varphi_j(x, y, u)$ $(j = 1, 2)$ два независимых первых интеграла системы (8). Отметим, что функция $\mathfrak{F}(\varphi_1(x, y, u), \varphi_2(x, y, u))$, где $\mathfrak{F}(\cdot)$ — произвольная гладкая функция двух переменных, также будет первым интегралом системы (8).

Выберем кривую $\gamma = \{x(\alpha), y(\alpha)\}$, лежащую в Ω, на которой зададим гладкую функцию $g(\alpha)$ $|g(\alpha)| \leqslant M$. Предположим, что кривая $\Gamma = \{x(\alpha), y(\alpha), g(\alpha)\}$ в точке $x_0 = x(\alpha_0)$, $y_0 = y(\alpha_0)$, $u_0 = g(\alpha_0)$ не касается характеристик уравнения (7). Это означает, что

$$\operatorname{rang} \begin{pmatrix} a(x_0, y_0, u_0) & b(x_0, y_0, u_0) & c(x_0, y_0, u_0) \\ x'(\alpha_0) & y'(\alpha_0) & g'(\alpha_0) \end{pmatrix} = 2. \qquad (9)$$

Утверждение 3. *Пусть кривая Γ в точке x_0, y_0, u_0 удовлетворяет перечисленным выше условиям. Тогда уравнение (7) имеет в окрестности этой точки единственное решение, совпадающее с функцией $g(\alpha)$.*

Для дальнейшего построения решения вдоль кривой необходимо воспользоваться методом продолжения.

Положим $\varphi_j(x(\alpha), y(\alpha), g(\alpha)) = \varphi_{j*}(\alpha)$ $(j = 1, 2)$. Тогда решение задачи Коши в окрестности точки x_0, y_0, u_0 может быть представлено в неявном виде как

$$g\big(\varphi_{1*}^{-1}(\varphi_1(x, y, u))\big) - g\big(\varphi_{1*}^{-2}(\varphi_2(x, y, u))\big) = 0.$$

Остановимся на геометрической интерпретации понятия решения. Если имеется некоторая поверхность $u = u(x, y)$, то вектор $(u_x, u_y, -1)$ определяет направление нормали к этой поверхности. Таким образом, уравнение (6) сводится к требованию, чтобы в каждой точке поверхности $u = u(x, y)$ направление, определяемое полем $(a(\cdot), b(\cdot), c(\cdot))$ находилось в касательной плоскости к поверхности. Отсюда следует, что если поверхность $u = u(x, y)$ является решением уравнения (6), то она состоит из характеристик этого уравнения.

О преподавании спецкурса "Алгоритмы на графах" студентам специальности "Прикладная математика и информатика"

В. Л. Дольников, О. П. Якимова

Рассматриваются некоторые вопросы преподавания спецкурса "Алгоритмы на графах".

Теория графов — важный математический инструмент, широко используемый в химии, генетике, исследовании операций, лингвистике, проектировании. Непосредственное и детальное представление практических систем, таких, как распределительные сети и системы связи, приводит к графам большого размера, успешный анализ которых зависит от существования "хороших" алгоритмов. Поэтому в рамках курса уделяется большое внимание разработке и описанию эффективных алгоритмов, решающих практические задачи. Работа по этой проблематике значительно обогащает алгоритмическую культуру студентов, совершенствует технику написания программ, так как алгоритмы являются неиссякаемым источником задач любого уровня сложности.

В настоящее время учебный план предусматривает 36 часов занятий. Семестровый курс заканчивается экзаменом. В связи с обширностью предмета перед преподавателем встает нетривиальная задача об отборе наиболее важных тем и наилучшей организации подачи материала. Предлагается следующая структура курса.

1. Представление графа в памяти компьютера. Выбор соответствующей структуры данных для представления графа имеет принципиальное значение при разработке эффективных алгоритмов. При решении задач используются четыре основных способа описания графа: матрица инциденций; матрица смежности; списки смежных вершин и перечни ребер. В дальнейшем студентам предлагается оценить трудоемкость алгоритма для того или иного представления графа.

2. Алгоритм поиска в глубину. На его основе разбираются вопросы достижимости и связности графов, построения сильных компонент и решается задача отыскания всех двусвязных компонент.

3. Алгоритм поиска в ширину. Задача о кратчайшем пути, соединяющем заданные вершины (алгоритм Дейкстры). Рассматривается задача нахождения критического пути в графе проектирования работ и для ее решения разбирается алгоритм нахождения пути в бесконтурном графе. Алгоритм Флойда отыскания кратчайших путей между всеми парами вершин.

4. Деревья, их свойства. Построение остовного дерева минимальной стоимости с помощью алгоритмов Краскала и Прима. Их сравнительный анализ. Приложение этих задач к проектированию газовых и электрических распределительных сетей.

5. Алгоритмы поиска в деревьях. Деревья двоичного поиска. Сбалансированные двоичные деревья. АВЛ-деревья, красно-черные деревья: вставка, удаление, поиск элемента. Трудоемкость. Оценка максимальной и средней высоты сбалансированного дерева с n узлами. Сферы применения сбалансированных двоичных деревьев.

6. Методы поиска во внешней памяти на основе деревьев. В–деревья: вставка, удаление, поиск элементов. В+деревья, как основа для построения индексов в базах данных.

7. Обходы. Эйлеровы циклы и эйлеровы графы. Критерий существования эйлеровых путей и циклов в неориентированных и ориентированных графах. Алгоритм Флери построения эйлерова цикла. Задача китайского почтальона и ее приложения к сборке мусора, к доставке почты. Гамильтоновы циклы и гамильтоновы графы, достаточные условия существования гамильтонова цикла. NP-полные задачи. Задача коммивояжера. Эвристические алгоритмы для ее решения.

8. Независимость и покрытия. Рассматриваются две задачи, связанные с выбором экстремальных множеств вершин с заданными свойствами. Способы решения задач об определении максимальных независимых и минимальных доминирующих множеств, последняя обобщается до задачи о покрывающих множествах. Приложения задачи о покрывающих множествах к составлению расписаний вылета самолетов и работы транспорта, к задаче об информационном поиске.

9. Наибольшие паросочетания и задача о назначениях. Алгоритм построения наибольшего паросочетания в двудольном графе. Транспортная задача. Алгоритм построения совершенного паросочетания минимального веса в двудольном взвешенном графе.

10. Раскраски. Задача составления расписаний, распределения оборудования. Алгоритм последовательной раскраски. Раскраска планарных графов.

11. Задача о максимальном потоке. Алгоритм Форда – Фалкерсона.

Следует отметить, что спецкурс "Алгоритмы на графах" во многом является обобщающим, так как некоторые знания по теории графов и алгоритмам на графах студенты получили в рамках других учебных курсов, прежде всего дискретной математики и информатики.

В ходе изучения предмета студенты выполняют индивидуальные задания по решению прикладных задач, требующих модификации известных алгоритмов. Хорошим источником нетривиальных задач являются сайты олимпиад по информатике, включая широко известный чемпионат мира по программированию, проводимый под эгидой международной организации ACM (Association for Computing Machinery), например, acm.timus.ru.

О теоретико-множественном и логическом введении

В. Г. Дурнев

Обосновывается целесообразность включения в Государственные образовательные стандарты по математическим специальностям дисциплины "Элементы теории множеств и математической логики".

Библиография: 2 названия.

Достаточно трудно оспорить утверждение, что теоретико-множественные методы к середине XX века проникли во многие разделы математики, стали составной частью их фундамента. Это вызвало необходимость включения в учебные планы математических специальностей университетов дисциплин, дающих некоторое первоначальное введение в теорию множеств и математическую логику. По сложившейся традиции элементы теории множеств и математической логики включаются в отдельные математические дисциплины, изучаемые студентами первого курса. Однако, на наш взгляд, это приводит к некоторым нежелательным последствиям, хотя и не катострофическим, но и небезобидным. При разработке стандартов второго поколения автор этих строк предлагал включить в стандарт по специальности "Математика" в блок естественно научных дисциплин семестровый курс под условным названием "Элементы теории множеств и математической логики", который, кстати, уже много лет включается в учебные планы по этой специальности на математическом факультете Ярославского университета имени П. Г. Демидова. Однако это предложение не нашло поддержки среди представителей ряда российских университетов. Одним из аргументов при этом была ссылка на то, что по сложившейся традиции теоретико-множественные вопросы рассматриваются на начальном этапе изучения дисциплины "Математический анализ". Поэтому представляет интерес хотя бы бегло посмотреть, как это делается. За основу мы взяли замечательную, на наш взгляд, книгу В. А. Зорича "Математический анализ", допущенную Министерством высшего и среднего специального образования СССР в качестве учебника для студентов университетов, обучающихся по специальностям "Математика" и "Механика" [1].

Первая глава начинается с введения логической символики. В рассмотрение вводятся, как говорится на стр. 13, «символы математической логики $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$ для обозначения соответственно отрицания "не" и связок "и", "или", "влечет", "равносильно"». Однако из приводимых вслед за этим примеров видно, что автор имеет ввиду скорее *союзы*, чем *связки*. А что такое *логическая связка,* без привлечения истинностных таблиц объяснить затруднительно. На стр. 15 вскользь упоминается один из основных законов классической логики — *закон исключенного третьего*, который почему-то назван *принципом исключенного третьего*, и из текста в начале стр. 15 совершенно не ясно, почему можно, как утверждается ниже, «использовать также принцип исключенного третьего, в силу которого высказывание $A \vee \neg A$ (A или $\neg A$) считается истинным независимо от конкретного содержания высказывания A». И уж совсем загадочным представляется следующее за этим предложение: «Следовательно, мы одновременно принимаем, что $\neg(\neg A) \Leftrightarrow A$, т. е. повторное отрицание равносильно исходному высказыванию». Мы не призываем уже на первых шагах изучения математики уделять особое внимание таким вопросам, как существование весьма влиятельного направления математической мысли — интуиционизма, представители которого не готовы принять ни закон исключенного третьего, ни снятие двойного отрицания. Нас интересует лишь статус рассмотренного текста. Если это "беглое" напоминание вопросов, уже рассмотренных во вводном курсе математической логики, то такое достаточно поверхностное рассмотрение столь фундаментальных логических вопросов может быть в какой-то мере и оправдано. Если же этот текст претендует на ознакомление слушателей с новыми для них понятиями, то статус текста становится небесспорным.

Квантор существования $\exists$ вопреки логической традиции называется логическим оператором "существует", или "найдется". Однако, эти два слова имеют несколько различный смысл, поэтому в математической логике принято говорить "существует", хотя, как потом выясняется, и смысл этого слова может пониматься весьма различными способами.

Немало вопросов вызывает и следующий текст на стр. 19.

«Например, запись $\forall x((x \in A) \Leftrightarrow (x \in B))$ означает, что для любого объекта x соотношения $x \in A$ и $x \in B$ равносильны. Поскольку множество вполне определяется своими элементами, указанное высказывание принято обозначать короткой записью

$$A = B,$$

читаемой "A равно B", обозначающей совпадение множеств A и B».

Однако из приведенного текста не удается вывести, что если $A = B$ и $A \in C$, то $B \in C$. А ведь в канторовском пояснении понятия множества ключевыми служат слова *объединение в единое целое*, что позволяет, построив по некоторым исходным элементам множество A, строить затем множество C, одним из элементов которого является A, и т. д. И как установить невозможность ситуации, при которой $A = B$ и $A \in C$, но $B \notin C$?

Все эти вопросы, конечно, следует обсуждать не на лекциях по математическому анализу, а в специально для этого предназначенном курсе.

Теорема об антирефлексивности отношения $cardA \leq cardB$ $((cardA \leq cardB) \wedge (cardB \leq cardA) \Rightarrow (cardA = cardB))$ почему-то названа теоремой Шредера – Бернштейна, хотя традиционно ее называют теоремой Кантора – Бернштейна, иногда теоремой Кантора – Бернштейна – Шредера. Теорема же Цермело о сравнимости любых двух множеств по мощности (либо $cardA \leq cardB$, либо $cardB \leq cardA$) почему-то названа теоремой Кантора, хотя именно в связи с доказательством этой теоремы Цермело первым явно сформулировал *аксиому выбора.* Конечно, эта теорема следует из предположения Кантора о возможности вполне упорядочить произвольное множество. Но доказау эту возможность именно Цермело.

На стр. 38 – 39 приводится список некоторых аксиом теории множеств, однако уже первая из аксиом списка — аксиома объемности — сформулирована, на наш взгляд, не бесспорным образом — в ней на самом деле дается определение равенства множеств. Цитируем текст на стр. 38.

«Аксиома объемности. *Множества A и B равны тогда и только тогда, когда они имеют одни и те же элементы*».

А звучать она могла бы, например, так.

«Аксиома объемности. *Если множества A и B состоят из одних и тех же элементов, то они равны*».

При этом обратная импликация: *если множества A и B равны, то они состоят из одних и тех же элементов* — относится к логическим аксиомам, а не к аксиомам теории множеств.

Завершает список аксиом текст в конце стр. 39.

«Аксиомы 1 – 7 составляют аксиоматику теории множеств, известную как аксиоматика Цермело – Френкеля.

К ней обычно добавляется еще одна независимая от аксиом 1 – 7 и часто используемая в анализе

8. Аксиома выбора. Для любого семейства непустых множеств существует множество C такое, что, каково бы ни было множество X данного семейства, множество $X \cap C$ состоит из одного элемента».

По поводу этого текста можно сделать два замечания.

Во-первых, в такой формулировке аксиома просто не верна, как показывает следующий простой пример: для семейства, состоящего из множеств $\{1\}$, $\{2\}$ и $\{1, 2\}$ соответствующее множество C не существует.

Во-вторых, вскользь утверждается *независимость аксиомы выбора от остальных аксиом* 1 – 7. Однако подобное можно утверждать лишь *при условии, что исходная система аксиом непротиворечива.*

Конечно, все эти достаточно тонкие вещи можно обсуждать лишь в отдельном курсе, но тогда зачем их вскользь упоминать? Это, на наш взгляд, может создать у студента *иллюзию понимания*, что в свою очередь приводит к недоуменным вопросам типа: почему студенты успешно сдают отдельные фундаментальные математические дисциплины, но испытывают трудности с таким "простейшим предметом", как аналитическая геометрия? Просто, на наш взгляд, в аналитической геометрии существенно меньше возможностей для "создания иллюзии понимания" в силу весьма конкретного характера материала.

Рассмотрим текст на стр. 47.

«Относительно любой абстрактной системы аксиом сразу же возникают по крайней мере два вопроса.

Во-первых, совместимы ли эти аксиомы, т. е. существует ли множество, удовлетворяющее всем перечисленным условиям. Это вопрос о *непротиворечивости аксиоматики*».

Однако *вопрос о непротиворечивости аксиоматики* традиционно формулируется как вопрос о невозможности вывести из рассматриваемой системы аксиом некоторое утверждение вместе с его отрицанием. Вопрос же: существует ли множество (структура), удовлетворяющее данной системе аксиом — это вопрос о *совместности аксиоматики.* Впрочем, в силу теоремы Геделя о полноте с точки зрения классической математики эти два вопроса равносильны.

Требует уточнения и вопрос, касающийся гипотезы континуума. В цитируемой книге говорится:

«Уже на заре теории множеств возник вопрос о том, существуют ли множества промежуточной мощности между счетными множествами и множествами мощности континуума, и было высказано предположение, называемое гипотезой континуума, что промежуточные мощности отсутствуют. Вопрос оказался глубоко затрагивающим основания математики. Он был решен в 1963 г. современным американским математиком П. Коэном. Коэн доказал неразрешимость гипотезы континуума, показав, что она сама, как и ее отрицание, не противоречит принятой в теории множеств аксиоматике, а потому гипотеза континуума не может быть ни доказана, ни опровергнута в рамках этой аксиоматики.

Ситуация вполне аналогичная независимости пятого постулата Евклида о параллельных от остальных аксиом геометрии.»

На наш взгляд, этот текст нуждается в принципиальных уточнениях.

Континуум-гипотеза была сформулирована Г. Кантором и получила название *континуум-гипотеза Кантора*. В знаменитые проблемы Д. Гильберта она входит под номером 1 и называется "Проблема Кантора о мощности континуума".

Проблема, сформулированная Кантором не относилась к какой-либо аксиоматике теории множеств (их во времена Кантора просто еще не было), а относилась к теории множеств в том виде, как ее создал сам Кантор. Поэтому и аналогия с геометрией не вполне правомерна.

Первое продвижение в исследованиях о месте континуум-гипотезы в аксиоматической теории множеств сделал К. Гедель в 1940 г.

Результаты К. Геделя и П. Коэна имеют ценность лишь *при условии непротиворечивости той системы аксиом, независимость от которой устанавливается*. В случае же ее противоречивости из нее выводима, в частности, как сама континуум-гипотеза, так и ее отрицание!

Хотя аксиома выбора и сформулирована в рассматриваемой книге в явном виде, однако не показано, как она используется. Конечно, это можно было бы сделать уже при доказательстве счетности объединения счетного семейства счетных множеств, но это трудновато для восприятия. Однако необходимость аксиомы выбора явно видна при доказательстве эквивалентности определений предела функции в точке по Коши и по Гейне: в каждой $1/n$-окрестности $V_n(a)$ точки a существует точка x_n такая, что $f(x_n) \notin V(A)$, но как «одновремено выбрать счетное множество $\{x_n\}$ таких точек»?

Можно было бы привести и еще ряд примеров определенной "недостаточной строгости изложения с точки зрения современной логики". Однако возникает более важный вопрос, возможно ли в таких идейно и технически сложных дисциплинах как математический анализ одновременно достичь высокого уровня строгости и сохранить доступность? В то же время, если в "более простых" дисциплинах не стремиться к соответствующему уровню строгости, то что в них останется?

Именно стремление к изложению математической дисциплины на "современном уровне строгости" часто делает ее трудной для восприятия студентами, а не ее конкретное содержание. Но если к этому не стремиться, то не превращаем ли мы математические дисциплины в "учения, которым студенты должны верить и следовать, потому что их к этому принуждают преподаватели"?

В другом достаточно популярном учебнике — книге Л. Д. Кудрявцева "Курс математического анализа"[2], также допущенной Министерством высшего и среднего специального образования СССР в качестве учебника для студентов физико-математических и инженерно-физических специальностей вузов, теоретико-множественным вопросам уделяется еще меньше внимания — они рассматриваются лишь в первом параграфе первой главы, занимающем всего 11 страниц. И здесь вопросы возникают уже во втором абзаце, который звучит так:

«Множества A и B называются *равными*, если они состоят из одних и тех же элементов. Таким образом, равенство $A = B$ означает, применительно к множествам, что одно и то же множество обозначено разными буквами A и B».

Этот текст, по всей видимости, предлагается в качестве *определения равенства множеств*, но не сказано, какие свойства равенства $A = B$ при этом выполнены, в частности, из того, что $A = B$ и $A \in C$ следует ли, что $B \in C$.

При этом интересно заметить, что создатель теории множеств Г. Кантор в этом параграфе не упоминается, хотя о Д. Пеано говорится в связи с аксиоматизацией для систем натуральных чисел. Однако совершенно неясной остается роль рассматриваемых аксиом Д. Пеано — не показано, как, отправляясь от них, можно ввести сложение и умножение натуральных чисел, и т. д.

Конечно, можно в какой-то мере считать, что при изучении функций одной действительной переменной теоретико-множественный язык не играет достаточно важной роли, но в случае функций двух и более переменных подобное утверждать вряд ли обосновано.

Заметим, что аксиома выбора используется уже при доказательстве счетности объединения счетного множества счетных множеств, при доказательстве эквивалентности определений предела функции в точке по Гейне и по Коши, эквивалентности различных определений предельной точки множества, в доказательстве теоремы Кантора о равномерной непрерывности функции, непрерывной на отрезке (на компакте), и т. д. Эти логические вопросы весьма не просты, однако и совершенно их игнорировать, как это практикуется много лет, вряд ли оправдано после того, как во второй половине XX века был проведен глубокий логический анализ оснований ряда разделов математики. В частности, в значительной мере была выявлена роль одного из наиболее знаменитых (дискуссионных) утверждений теоретико-множественной математики — аксиомы выбора.

Подобные проблемы возникают и при преподавании других математических дисциплин — изучение различных уточнений понятия "алгоритм" (машины Тьюринга, рекурсивные функции и др.) без достаточно глубокого понимания необходимости этих уточнений, написание все новых и новых программ на все новых и новых языках программирования без попытки уяснения, что же такое синтаксически правильная программа на том или ином языке.

Конечно, трудно понять, что важнее: выработка навыков решения задач из тех или иных сборников или уяснение логических основ дисциплины. Для решения этих двух тесно связанных между собой проблем от студентов требуются усилия различного характера. На наш взгляд, по ряду объективных и субъективных причин традиционно большее внимание уделяется решению первой задачи. Однако нам кажется, что именно успешное решение второй задачи в большей мере способствует формированию "математически образованного" специалиста. Со временем навыки забудутся, но "математическая образованность" должна остаться. На решение этих вопросов в определенной мере и нацелена дисциплина "Элементы теории множеств и математической логики", которую, на наш взгляд, целесообразно включить в разрабатываемые в настоящее время Государственные образовательные стандарты третьего поколения по математическим специальностям.

Список литературы

[1] *Зорич В. А.* Математический анализ. Часть 1. М.: Наука, 1981. 544 с.

[2] *Кудрявцев Л. Д.* Курс математического анализа (в 2-х томах): Учебник для студентов университетов и вузов. Т. 1. М.: Высшая школа, 1981. 681 с.

О продолжении линейного функционала в конечномерном пространстве

И. П. Иродова

Разобраны примеры продолжения функционала с сохранением нормы.

Библиография: 1 название.

Тема "Линейные непрерывные функционалы" является одной из основных тем в курсе "Функциональный анализ", который читается студентам-математикам в 5–6 семестрах в объеме 85 часов лекций и 53 часов лабораторных занятий. Центральное место в этой теме занимает теорема Хана – Банаха о продолжении линейного функционала с сохранением нормы. Важность теоремы Хана – Банаха объясняется тем, что она имеет многочисленные приложения.

В этой статье будет рассмотрен вопрос о построении продолжения линейного функционала на плоскости и в трехмерном пространстве. На конкретном примере будет показано, что в неевклидовых пространствах единственности продолжения может не быть.

Цель, которая преследовалась при решении задач этого цикла, заключается в том, чтобы студенты, с одной стороны, усвоили формулировку теоремы Хана – Банаха и смежные с ней вопросы, а с другой стороны, поняли простую геометрическую иллюстрацию этой теоремы в конечномерном пространстве.

Перейдем к формулировке основного результата. Для этого напомним основные определения (см., например, [1]).

Определение 1. *Пусть L — линейное нормированное пространство и f — линейный непрерывный функционал, заданный на L. Число*

$$\|f\|_L = \sup_{x\neq 0} \frac{|f(x)|}{\|x\|}$$

мы назовем ***нормой*** *функционала f.*

Отметим, что в силу непрерывности f норма функционала конечна. Приведем формулировку теоремы, которая дает геометрическую интерпретацию нормы функционала.

Теорема 1. *Пусть f — линейный непрерывный функционал, заданный на L. Тогда*

$$\|f\|_L = \frac{1}{\inf_{x \in A_f} \|x\|};$$

здесь гиперплоскость A_f определяется формулой

$$A_f = \{x \in L : \quad f(x) = 1\}.$$

Таким образом, норма функционала — это величина, обратная к расстоянию от гиперплоскости A_f до нуля.

Определение 2. *Пусть L_0 — подпространство L и f_0 — непрерывный линейный функционал, заданный на L_0. Непрерывный функционал f называется* ***продолжением*** *f_0, если*

$$f(x) = f_0(x), \quad x \in L_0.$$

Сформулируем теперь теорему Хана – Банаха.

Теорема 2. *Пусть L — линейное нормированное пространство, f_0 — линейный непрерывный функционал, заданный на L_0. Тогда функционал f_0 может быть продолжен до некоторого линейного функционала f без увеличения нормы, то есть так, чтобы выполнялось равенство*

$$\|f_0\|_{L_0} = \|f\|_L.$$

Можно дать геометрическую интерпретацию теоремы Хана – Банаха. Как следует из теоремы 1, уравнение $f_0(x) = 1$ определяет в L_0 гиперплоскость, лежащую на расстоянии $\frac{1}{\|f_0\|}$ до нуля. Продолжая f_0 без увеличения нормы до функционала на всем L, мы проводим через эту частичную гиперплоскость "большую" гиперплоскость на всем L, причем расстояние до нуля остается прежним.

Выберем теперь в качестве L пространство l_p^n. Напомним, что норма в l_p^n, $1 \leqslant p \leqslant \infty$, задается формулой

$$\|x\|_{l_p^n} := \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}};$$

как обычно, если $p = \infty$, то сумма заменяется на максимум.

Далее мы будем рассматривать случаи $n = 2$ и $n = 3$.

Теперь сформулируем теорему об общем виде функционала, заданного в l_p^n.

Теорема 3. *Пусть f — линейный непрерывный функционал, заданный на l_p^n, $1 \leqslant p \leqslant \infty$. Тогда существует вектор $a = (a_1, ..., a_n)$ такой, что*

$$f(x) = \sum_{i=1}^{n} a_i x_i$$

и, кроме того,

$$\|f\| = \|a\|_{l_q^n};$$

здесь $\frac{1}{p} + \frac{1}{q} = 1$.

Все необходимые формулировки приведены, перейдем к решению задач.

Задача 1. *Пусть L_0 — подпространство l_1^2, определяемое равенством*

$$L_0 = \left\{ x \in l_1^2 : \quad x_1 + x_2 = 0 \right\}.$$

На L_0 задан функционал $f_0(x) = x_1 - 2x_2$. Найти продолжение функционала f_0, чтобы выполнялись условия теоремы Хана – Банаха.

Решение. Вычислим норму функционала f_0. По определению имеем:

$$\|f_0\|_{L_0} = \sup_{x_1+x_2=0} \frac{|x_1 - 2x_2|}{|x_1| + |x_2|} = \frac{3}{2}.$$

Из теоремы 3 следует, что искомый функционал f имеет вид

$$f(x) = a_1 x_1 + a_2 x_2$$

и

$$\|f\|_{l_1^2} = \max(|a_1|, |a_2|).$$

Так как функционал f является продолжением f_0, то

$$\begin{cases} a_1 x_1 + a_2 x_2 = x_1 - 2x_2 \\ x_1 + x_2 = 0. \end{cases}$$

Отсюда следует, что $a_1 - a_2 = 3$. Учитывая, что $\|f\|_L = \|f_0\|_{L_0}$, получаем второе условие на числа $a_1, a_2 :\quad \max(|a_1|, |a_2|) = \frac{3}{2}$. Таким образом, нужно решить уравнение $\max(|a_1|, |a_1 - 3|) = \frac{3}{2}$. Построив график $y = \max(|a_1|, |a_1 - 3|)$, находим, что $a_1 = \frac{3}{2}$. Итак, искомый функционал имеет вид

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

Приведем пример, показывающий что функционал f_0 может иметь не единственное продолжение.

Задача 2. *Пусть подпространство L_0 пространства l^2_∞ определяется равенством*

$$L_0 = \left\{x \in l^2_\infty : \quad x_1 = x_2\right\}.$$

а

$$f_0(x) = x_1 + 3x_2.$$

Найти продолжение f_0.

Решение. Используем геометрический подход. Вычислим норму f_0 по теореме 1. Сейчас гиперплоскость $A_{f_0} = \{x \in L_0 : \quad x_1 + 3x_2 = 1\}$ состоит из одной точки $M = \left(\frac{1}{4}, \frac{1}{4}\right)$. Так как расстояние от A_{f_0} до 0 равно $\frac{1}{4}$, то $\|f_0\| = 4$. Построим теперь гиперплоскость (а сейчас это прямая) $A_f = \{x \in L : \quad f(x) = 1\}$ так, чтобы она прошла через точку M, и расстояние до нуля было равно $\frac{1}{4}$. Напомним, как находить расстояние от точки до прямой. Берем маленький шар с центром в заданной точке и "раздуваем" его до первого касания с прямой. В нашем случае мы должны нарисовать шар с центром в нуле и радиусом $\frac{1}{4}$. Выясним, как выглядит шар в пространстве l^2_∞. Вспоминая, как задается норма в l^2_∞, получаем формулу

$$B[0, \frac{1}{4}] = \left\{x \in l^2_\infty : \quad \max(|x_1|, |x_2|) = \frac{1}{4}\right\}.$$

Таким образом, шар в пространстве l^2_∞ — это квадрат со сторонами, параллельными осям координат. Заметим, что точка M попала в "вершину" шара $B[0, \frac{1}{4}]$. Это означает, что можно проводить любую прямую, которая проходит через M и заключена между прямыми $x_1 = \frac{1}{4}$ и $x_2 = \frac{1}{4}$ (прямая не должна пересекать шар). Уравнение любой такой прямой имеет вид

$$a_1\left(x_1 - \frac{1}{4}\right) + a_2\left(x_2 - \frac{1}{4}\right) = 0, \quad a_1, a_2 \geqslant 0.$$

Это уравнение можно записать по-другому:

$$\frac{4}{a_1 + a_2}(a_1 x_1 + a_2 x_2) = 1.$$

Итак, функционал

$$f(x) = \frac{4}{a_1 + a_2}(a_1 x_1 + a_2 x_2),$$

где $a_1, a_2 \geqslant 0$ и $a_1^2 + a_2^2 \neq 0$, является искомым.

Заметим, что если точка попадает на грань шара (как в задаче 1), то продолжение единственное, а если в вершину (как в задаче 2), то единственности нет. В частности, отсюда следует, что в евклидовом пространстве l^2_2 продолжение единственно, так как шар в l^2_2 "круглый".

Задача 3. *Пусть подпространство L_0 пространства l_1^3 есть*

$$L_0 = \left\{x \in l_\infty^2 : \quad x_1 = t, x_2 = 2t, x_3 = 3t\right\}.$$

а

$$f_0(x) = 2x_1 + 3x_2 - x_3.$$

Найти продолжение f_0.

Решение. Вычислим норму функционала f_0. Имеем:

$$\|f_0\| = \sup_t \frac{|2t + 6t - 3t|}{|t| + |2t| + 3t} = \frac{5}{6}.$$

Функционал f имеет вид

$$f(x) = a_1x_1 + a_2x_2 + a_3x_3.$$

Так как f является продолжением f_0 и их нормы совпадают, то получаем следующую систему:

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 = 2x_1 + 3x_2 - x_3 \\ x_1 = t \\ x_2 = 2t \\ x_3 = 3t \\ \max\left(|a_1|, |a_2|, |a_3|\right) = \frac{5}{6} \end{cases}.$$

Первые четыре условия дают равенство $a_1+2a_2+3a_3 = 5$. Итак, остается решить такую задачу:

$$\max\left(|5 - 2a_2 - 3a_3|, |a_2|, |a_3|\right) = \frac{5}{6}.$$

Можно просто перебрать варианты, когда одно из чисел равно $\frac{5}{6}$, а два других не превосходят $\frac{5}{6}$. Мы решим задачу по-другому. Так как $|5-$ $- 2a_2 - 3a_3| \leqslant \frac{5}{6}$, то $\frac{25}{6} \leqslant 2a_2 + 3a_3 \leqslant \frac{35}{6}$. С другой стороны, $a_2 \leqslant \frac{5}{6}$ и $a_3 \leqslant \frac{5}{6}$, а значит, $2a_2 + 3a_3 \leqslant \frac{25}{6}$. Следовательно, существует единственное решение задачи $a_1 = a_2 = a_3 = \frac{5}{6}$.

Список литературы

[1] *Колмогоров А. Н., Фомин С. В.* Элементы теории функций и функционального анализа. М: Наука, 1976. 544 с.

О преподавании теории чисел

Л. С. Казарин, В. К. Шалашов

Плато сказал: "Бог есть геометрия", Якоби произнес: "Бог это арифметика". Явился Кронекер и изрек: "Бог создал натуральные числа, а все остальное — работа человека".

Феликс Клейн

Курс "Теория чисел", как мы его представляем, изложен в нашем учебном пособии (Казарин Л.С., Шалашов В.К. " Теория чисел", Ярославль, 2003–2004). Его цель — дать элементарное изложение классических теорем теории чисел с вкраплениями современных результатов. Теория чисел всегда занимала уникальное место в математике. Почти каждый век со времен античности был свидетелем новых блестящих результатов, относящихся к свойствам чисел. Большинство выдающихся математиков внесло вклад в теорию чисел.

Почему теория чисел непреодолимо влечет как ведущих математиков, так и многих любителей? В то время как многие задачи в этой области знания чрезвычайно трудны, их можно сформулировать в терминах, достаточно простых, чтобы вызвать интерес у лиц без достаточной математической подготовки. Некоторые из просто звучащих вопросов выдержали интеллектуальный штурм в течение веков и остаются среди наиболее трудных нерешенных проблем математики.

Следует отметить, что теория чисел — один из наиболее подходящих предметов для начального изучения математики. Она не требует длительной предварительной подготовки, содержание ее понятно и доступно многим, а ее методы позволяют студентам познакомиться с научным подходом к решению задач. Студенты, работающие в этой области, имеют возможность использовать метод проб и ошибок в сочетании с применением любознательности, интуиции и изобретательности.

Удивительным образом, несмотря на кажущуюся отвлеченность этой науки, оказалось, что она имеет массу применений в обыденной жизни от быстрых алгоритмов вычисления (свертки, числового преобразования Фурье, работающих в современных специализированных процессорах) до интеллектуальных карточек, алгоритмов цифровой подписи и защиты информации. Это обстоятельство придает дополнительный стимулирующий импульс к изучению самой теории чисел и ее приложений, перекидывает мостики с сопредельными математическими дисциплинами. В то же время существование мощного подспорья в экспериментальной работе, каковым являются современные компьютеры, заставляет задумываться над алгоритмической стороной дела. Показательно, что усвоение принципа математической индукции легче дается, если понять, что большинство обычных рассуждений по индукции легко преобразуется в вычислительный алгоритм, использующий оператор цикла. По-видимому, каждое время требует своего подхода к преподаванию теории чисел и, стало быть, невозможно написать учебник по теории чисел на все времена.

Несколько слов о задачах. Они в достаточном количестве завершают большинство параграфов. Трудность их колеблется от чисто механического применения теории до необходимости проведения довольно сложного логического рассуждения. Они требуют активного участия студентов, ибо никто не сможет изучить теорию чисел без решения задач. При этом задачи вычислительного характера развивают базовую технику и проверяют понимание основных мотивов и понятий. Задачи теоретического характера дают практику в конструировании доказательств. Следует отметить небольшое упущение, состоящее в малом количестве задач, имеющих тип "верно ли...?" или "что будет, если...?", и т. д. Иначе говоря, мало задач поискового плана. Здесь авторы, признавая свой невольный промах, надеются на искусство преподавателя, ведущего практические занятия. Очень сомнительно, что можно сформулировать поисковые задачи учебного плана, которые можно использовать в течение длительного времени. Благодаря Интернету решения таких задач очень скоро становятся известны большинству студентов и, тем самым, утрачивают свою привлекательность.

Правило множителей Лагранжа в курсе математического анализа

В. С. Климов, А. Ю. Ухалов

Обсуждается методика изложения темы "Правило множителей Лагранжа" в курсе математического анализа. Приводится доказательство теоремы Лагранжа, не опирающееся на теорему о неявной функции.

Библиография: 1 название.

В курсе математического анализа теория функций одной переменной, как правило, излагается с достаточной полнотой. При переходе же к функциям многих переменных изложение существенно осложняется громоздкими выкладками. Даже когда основная идея утверждения проста, полное изложение доказательства часто оказывается технически сложным и его трудно привести целиком на лекции.

Лекторы вынуждены идти на упрощение курса — опускать важные теоремы или приводить их без доказательства. Одно такое упрощение влечет за собой другие, и в результате страдает целостность курса.

В частности, часто сокращению подвергается изложение круга вопросов, связанных с теоремой о неявной функции. Сама теорема о неявной функции для многомерного случая, как правило, приводится без доказательства. В результате оказываются плохо обоснованными и другие важные темы, например, теорема о правиле множителей Лагранжа. Доказательство этой теоремы, приводимое обычно в учебниках и лекционных курсах, существенно опирается на теорему о неявной функции. Логичным становится опустить и доказательство теоремы Лагранжа об условном экстремуме, так как из под фундамента ее доказательства уже вынут главный камень.

Оказывается, однако, что для многих утверждений многомерного анализа существуют доказательства, использующие минимальное количество фактов из предшествующего материала.

В настоящей заметке приводится элементарное доказательство теоремы о правиле множителей Лагранжа, опирающееся только на теорему Вейерштрасса.

Пусть на открытом множестве U пространства R^n определены два конечных семейства функций $f_i : U \to R$ $(i \in I)$ и $f_i : U \to R$ $(i \in I_0)$ $(I,\ I_0$ — конечные непересекающиеся множества). Пусть множество

$$E = \{x \in U | f_i(x) \leqslant 0 \quad (i \in I), \quad f_i(x) = 0 \quad (i \in I_0)\}$$

непусто. Рассматривается задача

$$f_0(x) \to \min, \tag{1}$$

$$f_i(x) \leqslant 0 \quad (i \in I), \qquad f_i(x) = 0 \quad (i \in I_0), \qquad x \in U. \tag{2}$$

Здесь $f_0 : U \to R$ — действительная функция на множестве U. Элемент $\hat{x}$ из E называют локальным решением задачи (1), (2), если найдется такое $\delta > 0$, что $f_0(x) \geqslant f_0(\hat{x})$ $\forall x \in E \cap B(\hat{x}, \delta)$. Здесь и далее $B(\hat{x}, R) = \{x \in R^n, \|x - \hat{x}\| \leqslant R\}$.

Введем в рассмотрение функцию

$$g(t) = \begin{cases} \frac{t^2}{2}, & t \geqslant 0, \\ 0, & t < 0. \end{cases}$$

Функция $g(t)$ дифференцируема и

$$g'(t) = \begin{cases} t, & t \geqslant 0, \\ 0, & t < 0. \end{cases}$$

Производная $g'(t)$ непрерывна на всей оси. Положим

$$\Phi(x) = \frac{1}{2} \sum_{i \in I_0} f_i^2(x) + \sum_{i \in I} g[f_i(x)].$$

Как нетрудно видеть, равенство $\Phi(x) = 0$ эквивалентно включению $x \in E$. Действительно, если $x \in E$, то $f_i(x) = 0 (i \in I_0)$, $g[f_i(x)] = 0 (i \in I)$, поэтому $\Phi(x) = 0$. Верно и обратное. Если функции f_i $(i \in I_1 = I \cup I_0)$ непрерывно дифференцируемы, то функция $\Phi(x)$ также непрерывно дифференцируема.

Теорема 1. *Пусть $\hat{x}$ — локальное решение задачи (1), (2) и функции $f_i (i \in \{0\} \cup I_1)$ непрерывно дифференцируемы в окрестности точки $\hat{x}$. Тогда найдутся такие числа $\hat{\lambda}_i$, не все равные нулю, что*

$$\hat{\lambda}_0 f_0'(\hat{x}) + \sum_{i \in I_1} \hat{\lambda}_i f_i'(\hat{x}) = 0, \tag{3}$$

$$\hat{\lambda}_i f_i'(\hat{x}) = 0 \quad (i \in I_1), \tag{4}$$

$$\hat{\lambda}_0 \geqslant 0, \hat{\lambda}_i \geqslant 0 \quad (i \in I). \tag{5}$$

Числа $\hat{\lambda}_i$ называют множителями Лагранжа, а саму теорему 1 — правилом множителей Лагранжа. Ясно, что вместе с набором $\hat{\lambda}_i$ соотношениям (3) – (5) удовлетворяет набор $t\hat{\lambda}_i$ $(t > 0)$ — набор множителей Лагранжа определяется неоднозначно. Наиболее важным необходимым условием оптимальности является равенство (3), называемое условием стационарности. Соотношение (4) называют условием дополнительной нежесткости. Если $f_i(\hat{x}) < 0$ для некоторого индекса i из I, то ограничение $f_i(x) \leqslant 0$ называют мягким в точке $\hat{x}$; в этом случае из (4) следует равенство $\hat{\lambda}_i = 0$.

Доказательство.

Каждому натуральному числу N поставим в соответствие функцию

$$\Phi^N(x) = f_0(x) + N\Phi(x) + \frac{|x - \hat{x}|^2}{2}$$

и экстремальную задачу

$$\Phi^N(x) \to \min, \qquad x \in B(\hat{x}, R). \tag{6}$$

Число $R > 0$ считаем настолько малым, что $B(\hat{x}, R) \subset U$, $f_0(x) \geqslant f_0(\hat{x})$ $\forall x \in E \cap B(\hat{x}, R)$, функции f_i непрерывно дифференцируемы на шаре $B(\hat{x}, R)$.

Существование решения x^N задачи (6) следует из теоремы Вейерштрасса. Справедливо неравенство $\Phi^N(x^N) \leqslant \Phi^N(\hat{x})$ или, более подробно,

$$f_0(x^N) + N\Phi(x^N) + \frac{|x^N - \hat{x}|^2}{2} \leqslant f_0(\hat{x}). \tag{7}$$

В частности, из (7) вытекает оценка $N\Phi(x^N) \leqslant f_0(\hat{x}) - f_0(x^N)$. Последовательность $x^N \in B(\hat{x}, R)$. Если $\bar{x}$ — предельная точка этой последовательности, то $\Phi(\bar{x}) = 0$. Из (7) следует неравенство

$$f_0(\bar{x}) + \frac{|\bar{x} - \hat{x}|^2}{2} \leqslant f_0(\hat{x}).$$

С другой стороны, $\bar{x} \in E \cap B(\hat{x}, R)$, следовательно $f_0(\bar{x}) \geqslant f_0(\hat{x})$, поэтому $\bar{x} = \hat{x}$. Всякая предельная точка последовательности x^N совпадает с $\hat{x}$. Это означает, что $x^N \to \hat{x}$. При больших N элемент x^N есть внутренняя точка шара $B(\hat{x}, R)$.

Так как x^N — внутренняя точка шара $B(\hat{x}, R)$ и x^N минимизирует функцию Φ^N на этом шаре, то x^N — стационарная точка функции Φ^N: $(\Phi^N)'(x^N) = 0$, т. е.

$$f_0'(x^N) + N\sum_{i \in I_0} f_i(x^N) f_i'(x^N) + N\sum_{i \in I} g'(f_i(x^N)) f_i'(x^N) + (x^N - \hat{x})^T = 0$$

Разделим это равенство на положительное число K^N, так чтобы получилось соотношение

$$\lambda_0^N f_0'(x^N) + \sum_{i \in I_1} \lambda_i^N f_i'(x^N) + \frac{(x^N - \hat{x})^T}{K^N} = 0, \tag{8}$$

в котором

$$(\lambda_0^N)^2 + \sum_{i \in I_1} (\lambda_i^N)^2 = 1, \quad \forall N. \tag{9}$$

Таким образом,

$$\lambda_0^N = \frac{1}{K^N}, \quad \lambda_i^N = \frac{N}{K^N} f_i(x^N) \quad (i \in I_0), \quad \lambda_i^N = \frac{N}{K^N} g'(f_i(x^N)) \quad (i \in I), \tag{10}$$

$$K^N = \left(1 + N^2 \sum_{i \in I_0} f_i^2(x^N) + N^2 \sum_{i \in I} |g'(f_i(x^N))|^2\right)^{\frac{1}{2}} \geqslant 1.$$

Каждая из последовательностей λ_i^N $(N = 1, 2, \dots)$ ограничена, поэтому, не уменьшая общности, можно считать, что $\lambda_i^N \to \hat{\lambda}_i$, $\forall i \in \{0\} \cup I_1$. Переходя в (9) к пределу, получаем равенство

$$\hat{\lambda}_0^2 + \sum_{i \in I_1} \hat{\lambda}_i^2 = 1.$$

Из этого равенства следует, что хотя бы одно из чисел $\hat{\lambda}_i$ отлично от нуля.

Полученный набор $\hat{\lambda}_i$ является искомым. Действительно, устремляя в (8) N к ∞, получаем условие стационарности (3) (используются соотношения $x^N \to \hat{x}$, $\lambda_i^N \to \hat{\lambda}_i$).

Если $f_i(\hat{x}) = 0$, то $\hat{\lambda}_i f_i(\hat{x}) = 0$, и условие дополнительной нежесткости выполнено. Если $f_i(\hat{x}) < 0$, то $f_i(x^N) < 0$ $(N >> 1)$, следовательно,

$$\lambda_i^N = \frac{N}{K^N} g'(f(x^N)) = 0,$$

поэтому $\hat{\lambda}_i = 0$, т. е. (4) справедливо и в этом случае. В силу (10) $\lambda_0^N \geqslant 0$, $\lambda_i^N \geqslant 0$ $(i \in I)$, поэтому $\hat{\lambda}_0 \geqslant 0$, $\hat{\lambda}_i \geqslant 0$ $\forall i \in I$. Теорема доказана. $\square$

Варианты теоремы 1 сохраняются, если одно из множеств I_0, I пусто.

Теорема 2. *Пусть $\hat{x}$ — локальное решение задачи*

$$f_0(x) \to \min,$$

$$f_i(x) \leqslant 0 \quad (i \in I), \quad x \in U.$$

и функции f_i непрерывно дифференцируемы в окрестности точки $\hat{x}$. Тогда найдутся такие числа $\hat{\lambda}_i \geqslant 0$, не все равные 0, что выполнены условия (3), (4) с $I_1 = I$.

Теорема 3. *Пусть $\hat{x}$ — локальное решение задачи*

$$f_0(x) \to \min,$$

$$f_i(x) = 0 \quad (i \in I_0), \quad x \in U.$$

и функции f_i непрерывно дифференцируемы в окрестности точки $\hat{x}$. Тогда найдутся такие числа $\hat{\lambda}_i \geqslant 0$, не все равные 0 одновременно, что выполнено условие (3) с $I_1 = I_0$.

Теоремы 1 – 3 выражают правило множителей Лагранжа. Во всех теоремах $\hat{\lambda}_0 \geqslant 0$. Если $\hat{\lambda}_0 > 0$, то разделив на $\hat{\lambda}_0$, получаем снова набор множителей Лагранжа, для которого $\hat{\lambda}_0 = 1$.

Лагранж считал, что всегда можно взять $\hat{\lambda}_0 = 1$. Однако это законно, лишь если $\hat{\lambda}_0 > 0$. Любое предположение, гарантирующее неравенство $\hat{\lambda}_0 > 0$, называют условием регулярности. В экстремальной задаче

$$f_0(x) = x_2 \to \min, \quad f_1(x_1, x_2) = x_2^3 - x_1^2 = 0$$

решение $\hat{x} = (0,0)^T$, условие регулярности не выполнено. Действительно,

$$f_0'(\hat{x}) = (0,1), \quad f_1'(\hat{x}) = (0,0),$$

и равенство $\hat{\lambda}_0 f_0'(\hat{x}) + \hat{\lambda}_1 f_1'(\hat{x}) = 0$ возможно лишь при $\hat{\lambda}_0 = 0$.

Основу доказательства теоремы 1 составляет широко используемый в оптимизации метод штрафных функций, см. [1], гл. 8. Там же приведены достаточно обозримые условия регулярности задачи (1), (2).

Список литературы

[1] *Поляк Б. Т.* Введение в оптимизацию. М.: Наука, 1983.

О некоторых резервах для повышения уровня математической подготовки студентов

А. Ю. Колесов, А. Н. Куликов

На основе данных современной психологии и педагогики авторы предлагают, а часто просто напоминают ряд методических приемов, способствующих улучшению уровня математической подготовки выпускников математического факультета.

Процесс формирования специалиста в любой отрасли знаний предполагает различные этапы. В частности, он включает в себя разбиение на дисциплины. Если говорить далее лишь о двух специальностях: "Математика" и "Прикладная математика и информатика", то обычно выделяют три цикла дисциплин. Это цикл математических дисциплин, информатика и цикл общеобразовательных дисциплин. Далее речь пойдет о первом цикле — цикле математических дисциплин, хотя сказанное частично относится и к двум оставшимся циклам.

Когда студент заканчивает свое образование в университете и ему выдается соответствующий диплом, то предполагается, что он приобрел достаточную для дальнейшей трудовой деятельности сумму знаний и умений. Принято считать, что первое и второе описано в документах, которые называются "Государственные образовательные стандарты".

Но трудно себе представить преподавателя, который в процессе своей деятельности еженедельно сверялся бы с ними. На самом деле, как правило, преподаватель ориентируется на свой личный опыт, представления, на коллективную точку зрения коллег по факультету. Итак представления о необходимом минимуме знаний и умений в значительной мере у любого преподавателя все-таки присутствуют. Так что обычно речь идет о проблеме достижения этого уровня.

Известна систематизация учебных целей по Блуму, который предложил шесть степеней усвоения содержания.

1. Знание — это способность воспроизводить факты, понятия, законы. (Дать определение производной, интеграла и т. д. Привести правила дифференцирования, интегрирования.)

2. Понимание — это способность переводить информацию из одной формы в другую, объяснить, интерпретировать, сравнить. (Например, сравнить два определения непрерывности — по Коши и по Гейне.)

3. Применение — это способность применять знания, опыт, навыки в относительно новых ситуациях. (Здесь иллюстрации излишни — это умение решать задачи средней сложности из стандартных задачников и руководств. Речь не идет, конечно, о простых задачах, имеющих характер упражнений, которые обычно сопутствуют определениям).

Уровни 4, 5, 6 — это умение разложить информацию на составляющие, способность создавать новое информационное поле на основе синтеза предшествующих информаций, давать качественные и количественные суждения, относящиеся к изучаемому предмету, теории. Далее мы не будем обсуждать эти три ступени иерархии, так как они затрагиваются на таких этапах обучения, как выполнение курсовых и дипломных работ.

В идеале, конечно, от преподавателя требуется вывести студента на третий уровень предлагаемой систематизации, но этот уровень на практике, как правило, соответствует отличной оценке, что вряд ли достижимо для всех студентов. На практике большинство преподавателей хорошо понимает, что третьего уровня можно достичь лишь для той части курса, который входит в так называемый обязательный минимум: под ним обычно понимают тот объем знаний, без которого трудно или даже невозможно усваивать смежные дисциплины.

Из психологических исследований известно, что человек запоминает лишь 10% от прочитанного, 20% от услышанного, 30% от увиденного, 50% от услышанного и увиденного, 70% от сделанного им самим. Эти данные частично объясняют, почему лекционнные формы обучения продолжают доминировать в процессе обучения, хотя призывы "реформаторов" направлены на замену их какими-то "новыми, более прогрессивными" формами работы со студентами. Второй вывод из этих исследований состоит в том, что 50% -ная эффективность при усвоении нового близка к предельной, как бы большего мы не желали. Больший процент эффективности можно достичь лишь повторениями. Понятно, что объемы часов не могут позволить сделать это многократно. Поэтому на практике остается не так много возможностей.

Первая состоит в том, чтобы в рамках данного курса повторять некоторые основные моменты предшествующих. Так, в курсе обыкновенных дифференциальных уравнений практикуется повторение некоторых понятий, определений и навыков из курса "Линейная алгебра и геометрия". В курсе "Уравнения математической физики" обычно предполагается краткий экскурс в теорию рядов Фурье. Практикуется повтор некоторых методов решения обыкновенных дифференциальных уравнений. Заметим, что это не обязательно делать в форме прямолинейного повторения. Часто можно найти завуалированную форму, форму сравнительного повторения. Поясним это на примере.

При изучении в рамках курса "Уравнения математической физики" систем уравнений с частными производными гиперболического типа можно отдельно выделить частный случай таких систем:

$$\frac{\partial u}{\partial t} = A\frac{\partial u}{\partial x}, \tag{1}$$

где $u = u(t,x) \in \mathbb{R}^n$, $x \in \mathbb{R}$, $t \geq 0$, а A — квадратная матрица, элементы которой $a_{jk} \in \mathbb{R}$ $(j,k = 1,2,\ldots,n)$. Напомним, что эта система называется гиперболической, если выполнено одно из двух условий:

1) все собственные числа матрицы действительны и просты;

2) A – симметричная матрица.

В этом случае рассмотрение системы (1) вместе с начальными условиями

$$u(0,x) = f(x) \in \mathbb{R}^n \tag{2}$$

приводит к возможности найти решение задачи Коши (1), (2) в явном виде. Это предусматривает нахождение инвариантов Римана, с помощью которых и строится решение этой задачи. Нахождение базисных решений, соответствующих каждому инварианту Римана, идентично нахождению линейно независимых решений системы обыкновенных дифференциальных уравнений

$$\dot{y} = Ay.$$

Вторая возможность состоит в использовании резервов, заложенных в такие курсы как "История математики" и "Концепции современного естествознания". Поясним это на первом примере. Курс "История математики" обычно строится как гуманитарный предмет, где просто перечисляются достижения математики (скорее математиков) на том или ином историческом отрезке. По-видимому, для студентов принесли бы большую пользу уменьшение чисто исторической информации и замена ее на рассказ о сути некоторых достижений. Например, вместо простого перечисления имен Петербургской математической школы, среди которых упоминается, конечно, и А. М. Ляпунов, следует привести небольшой экскурс в основное его достижение — создание теории устойчивости движения. Последнее позволит заодно повторить целый фрагмент курса "Обыкновенные дифференциальные уравнения", а также частично повторить и некоторые другие разделы этого курса. Такой подход, естественно, позволит улучшить подготовку студентов перед сдачей Государственных экзаменов. Еще больше возможности дает КСЕ с объемом часов более годового лекционного курса.

Об одном подходе к изложению теории физического маятника

Е. П. Кубышкин

Излагается подход к описанию динамики физического маятника.

Библиография: 1 название.

Физическим маятником принято называть твердое тело, которое может вращаться вокруг неподвижной горизонтальной оси в однородном поле тяжести. На этом физическом объекте демонстрируются законы механики, свойства решений дифференциальных уравнений, некоторые положения теории колебаний. Ниже излагается подход к описанию динамики физического маятника, основанный на идеях монографии [1].

Введем инерциальную систему координат $OXYZ$ таким образом, чтобы ось OZ совпадала с осью вращения маятника, а OY ось была направлена вертикально вниз. Свяжем с маятником систему координат $Oxyz$ таким образом, чтобы центр масс лежал на оси Oy, а оси Oz и OZ совпадали. Введем обобщенную координату φ — угол между осями Oy и OY. Система является консервативной. Кинетическая и потенциальная энергии определяются выражениями: $T(\dot\varphi) = J_z\dot\varphi^2/2$, $\Pi(\varphi) = -mga\cos\varphi$, где J_z — момент инерции твердого тела относительно оси OZ, m — масса тела, g — ускорение свободного падения, а интеграл энергии есть

$$\frac{1}{2}\dot\varphi^2 - \omega_0^2\cos\varphi = h, \quad \omega_0^2 = \frac{mga}{J_z}. \tag{1}$$

Уравнение движения (уравнение Лагранжа второго рода) имеет вид

$$\ddot{\varphi} + \omega_0^2 \sin\varphi = 0. \tag{2}$$

Анализируя выражение (1), можно качественно определить поведение траекторий уравнения (2) на фазовой плоскости $(\varphi, \dot{\varphi})$. Так, при $h < -\omega_0^2$ движение невозможно. При $h = -\omega_0^2$ маятник находится в положении равновесия $\varphi = 2k\pi$ $(k = 0, \pm 1, \dots)$, $\dot{\varphi} = 0$. Эти точки типа центр. Они окружены замкнутыми кривыми, соответствующими колебаниям системы, которые имеют место при $-\omega_0^2 < h < \omega_0^2$. Картина фазовых траекторий периодична с периодом 2π. При $h = \omega_0^2$ возможны два типа движений. Один тип движения соответствует положению равновесия маятника $\varphi = \pi + 2k\pi$ $(k = 0, \pm 1, \dots)$. Это точки типа седло. Второй тип движений — это сепаратрисы, идущие из седла в седло. При $h > \omega_0^2$ абсолютная величина угла φ монотонно возрастает. Движение маятника будет вращательным. Сепаратрисы разделяют области колебательных и вращательных движений.

Рассмотрим схему интегрирования уравнения (2).

1. $h = -\omega_0^2$. В этом случае возможны согласно (1) лишь состояния равновесия, отмеченные ранее.

2. $-\omega_0^2 < h < \omega_0^2$. Маятник в этом случае совершает колебательные движения. Обозначим через α угол, на который отклоняется маятник из положения равновесия $\varphi = 0$. Тогда $h = -\omega_0^2 \cos\alpha$ и (1) запишется в виде

$$\dot{\varphi}^2 = 2\omega_0^2(\cos\varphi - \cos\alpha). \tag{3}$$

Положим $k = \sin(\alpha/2)$ и сделаем замену переменных $\sin(\varphi/2) = k\sin(\psi)$. В результате (3) примет вид:

$$\dot{\psi}^2 = \omega_0^2(1 - k^2\sin^2\psi).$$

Если при $t = 0$ $\varphi = 0$, то

$$\omega_0 t = \int_0^{\psi} \frac{dx}{\sqrt{1 - k^2\sin^2 x}} = F(\psi, k). \tag{4}$$

Функция $u = F(\psi, k)$ называется эллиптическим интегралом первого рода. Величина $0 \leqslant k < 1$ называется модулем эллиптического интеграла. $K(k) = F(\pi/2, k)$ называется полным эллиптическим интегралом первого рода. Справедливо следующее представление в виде сходящегося ряда:

$$K(k) = \frac{\pi}{2}\Big(1 + \frac{1}{4}k^2 + \dots\Big).$$

Функция, являющаяся обращением эллиптического интеграла (4), называется амплитудой и обозначается $\psi = \mathrm{am}\,(u)$. Введем эллиптический синус $z = \mathrm{sn}\,(u) = \sin(\mathrm{am}\,(u))$ и эллиптический косинус $z = \mathrm{cn}\,(u) = \cos(\mathrm{am}\,(u))$, которые являются периодическими функциями с периодом $4K(k)$. Между названными функциями существуют определенные соотношения [1], аналогичные соотношениям между тригонометрическими функциями.

Из равенства (4) имеем $\psi = \mathrm{am}\,(\omega_0 t)$. Отсюда $\varphi = 2\arcsin(k\mathrm{sn}\,(\omega_0 t))$. Функция φ периодична по t с периодом $T = 4K(k)/\omega_0$. Если α мало, то $T \cong 2\pi\sqrt{l/g}$ — период колебаний математического маятника. Для математического маятника $a = l$, $J_z = ml^2$.

3. $h = -\omega_0^2$. Из (1) имеем $\dot{\varphi}^2 = 4\omega_0^2\cos^2(\varphi/2)$. Если при $t = 0$ $\varphi = 0$ и $\varphi > 0$, то, интегрируя уравнение, имеем $\varphi = \pi + 4\,\mathrm{arctg}(\exp(\omega_0 t))$.

4. $h > \omega_0^2$. Пусть при $t = 0$ $\varphi = 0$, $\dot{\varphi} = \dot{\varphi}_0$. Тогда $h = \dot{\varphi}^2/2 - \omega_0^2$ и (1) запишется в виде

$$\dot{\varphi}^2 = \dot{\varphi}_0^2(1 - k_1^2\sin^2(\varphi/2)), \quad k_1^2 = 4\omega_0^2/\dot{\varphi}_0^2.$$

Так как $h > \omega_0^2$, то $\dot{\varphi}_0^2 > 4\omega_0^2$ и, следовательно, $k_1^2 < 1$. Тогда $\dot{\varphi}t/2 = F(\varphi/2, k_1)$ и $\varphi = 2\mathrm{am}\,(\dot{\varphi}t/2)$ — возрастающая функция.

Список литературы

[1] *Журавский А. М.* Справочник по эллиптическим функциям. М.: Из-во АН СССР. 1941. 235 с.

О дисциплине специализации "Верификация программ"

Е. В. Кузьмин, В. А. Соколов

Обсуждаются распределение нагрузки на аудиторную и внеаудиторную работу студентов по дисциплине специализации «Верификация программ» и затрагиваемые в лекциях и заданиях научные направления по тематике предмета.

Библиография: 7 названий.

В Ярославском государственном университете им. П. Г. Демидова на факультете информатики и вычислительной техники для студентов 5-го курса специальности «Прикладная математика и информатика» преподаётся дисциплина специализации «Верификация программ».

В рамках этого учебного курса отводится значительное число внеаудиторных часов на самостоятельную работу студентов. В этой статье обсуждаются распределение нагрузки на аудиторную и внеаудиторную работу студентов и затрагиваемые в лекциях и заданиях научные направления по тематике предмета.

Верификация представляет собой анализ корректности программных систем относительно спецификации, в которой задаются исследуемые программные свойства. Современные методы анализа корректности включают в себя тестирование, метод доказательства теорем (theorem proving) [1] и метод проверки модели (model checking) [2, 3].

Тестирование применяется после окончательного написания программы. Но, как известно (Э. Дейкстра), если при тестировании ошибки найдены не были, это ещё не означает, что их нет вовсе. Отсюда следует необходимость рассмотрения формальных методов доказательства корректности программ. Поскольку тестирование – это способ проверки правильности программ, с которым студенты так или иначе (обычно в бессистемной форме) сталкивались при выполнении различных лабораторных работ, связанных с программированием, это направление предлагается студентам на самостоятельное изучение. Аудитории ставится задача подготовить ряд докладов и рефератов, отражающих современное положение дел в науке по тестированию. Основная цель – обратить внимание студентов на то, что тестирование представляет собой серьёзную проблематику, по которой в настоящее время проводится активная научно-исследовательская работа, а также определить круг проблем и ограничений, существующих в тестировании.

Метод доказательства теорем является очень трудоёмким методом с сильной привязкой к семантике языка программирования. Поэтому в рамках курса (на лекциях) вводится «простой» язык программирования [1], включающий привычные конструкции условного перехода и цикла, но имеющий ограничения на типы данных: допускаются переменные только целочисленного типа (или булевого типа). После определения семантики операторов введённого языка программирования, разъяснения основных принципов доказательства программ, написанных на этом языке, и рассмотрения нескольких примеров каждому студенту предлагается (по вариантам) доказать корректность небольшой программы. При этом допускаются «мягкая» и «жёсткая» формы задания. В первом случае формальная спецификация даётся вместе с текстом программы, во втором случае студенту необходимо самостоятельно произвести спецификацию формальным образом по сопровождению на естественном (русском) языке, т.е. построить ряд предикатов, отражающих начальные условия и ожидаемый результат выполнения программы. Пример задания приведён ниже.

Докажите формально, что следующий алгоритм записывает в а приближённое значение квадратного корня числа n:

$a, b := 0, n{+}1;$

do $a{+}1 \neq b \;\rightarrow\;\; d := (a{+}b){\div}2;$

    if $d{*}d \leq n \;\rightarrow\; a := d$

    [] $d{*}d > n \;\rightarrow\; b := d$

    fi

od

Спецификация: $\{Q\colon\ 0 \leq n\}$ – *предусловие*, $\{R\colon\ 0 \leq a^2 \leq n < (a{+}1)^2\}$ – *постусловие*, $\{P\colon\ a < b \leq n{+}1 \wedge a^2 \leq n < b^2\}$ – *инвариант цикла*, $\{t\colon b - a + 1\}$ – *ограничивающая функция.*

Важно отметить, что доказательство корректности такого рода программы проходит в терминах предикатов (что позволяет на практике применять полученные ранее знания по теории исчисления предикатов), занимает от трёх до пяти страниц и требует от студента предельной точности рассуждений. Более того, особенностью метода доказательства теорем является то, что в ходе рассуждений можно установить корректность программы относительно спецификации, однако если этого не удаётся сделать, то это, вообще говоря, не означает наличия ошибок в программе, а требует более тонкого разбора. Перефразируя известное изречение Э. Дейкстры (о тестировании) можно сказать, что метод доказательства теорем способен доказать отсутствие о шибок в программе (относительно спецификации), но не всегда доказывает их присутствие. С позиции технологии программирования метод доказательства теорем следует рассматривать как дополнение тестирования, позволяющее получать более совершенную стратегию отладки программ.

Метод проверки модели (model checking) – это метод автоматической верификации программных систем. При этом подходе для программы строится модель с конечным числом состояний, а свойства модели (спецификация) выражаются на языке темпоральной логики. Конечная модель и формулы темпоральной логики подаются на вход программе-верификатору, и далее проверка истинности формул для модели осуществляется автоматически (в основном с применением переборных алгоритмов). Поэтому в рамках обсуждаемой дисциплины по методу проверки модели важно рассмотреть следующие разделы: построение конечной модели программы, спецификация свойств на языках темпоральных логик и программы-верификаторы с описанием методов и алгоритмов, применяемых в этих программных средствах.

Необходимо заметить, что построение *конечной адекватной* (исходной программе) модели является весьма серьёзной задачей, поскольку модель может не учитывать ряд программных свойств или порождать несуществующие. Несмотря на то, что проверяются свойства модели, объектом исследования всё же по-прежнему остаётся программа.

Для спецификации свойств программ используются наиболее распространённые и широко применяемые темпоральные логики: логика линейного времени LTL (linear-time temporal logic) и логика ветвящегося времени CTL (branching-time temporal logic или computation tree logic). Логики имеют довольно простые синтаксис и семантику и не встречают трудностей в понимании у аудитории. В то же время прикладной аспект, касающихся применения этих логик, студентам не кажется столь прозрачным и полезным без достаточно простых, но семантически нагруженных примеров. Важно отметить, что в имеющейся литературе (статьи, учебники и монографии) таких примеров либо нет, либо они имеют вырожденный характер (применение логик в этом случае становится бессмысленным). В публикациях основной упор сделан на алгоритмы и методики. В работах, касающихся практического применения метода проверки модели, объектами рассмотрения являются или (довольно нетривиальные) параллельные алгоритмы, или протоколы передачи данных, перегруженные деталями реализации и требующие специальной подготовки и дополнительных знаний.

По мнению авторов, выходом из сложившейся ситуации может служить демонстрация метода проверки модели на программах, построенных с применением автоматного подхода к программированию [5–7]. Технология автоматного программирования является достаточно эффективной при построении программного обеспечения для «реактивных» систем и систем логического управления. Эта технология, не исключая других методов построения программного обеспечения «без ошибок», существенно более конструктивна, так как позволяет начинать «борьбу с ошибками» еще на стадии алгоритмизации. Автоматный подход к программированию с точки зрения моделирования и анализа программных систем имеет ряд преимуществ по сравнению с традиционным подходом. При автоматном программировании исключается проблема адекватности модели программе, поскольку набор взаимодействующих автоматов, описывающий логику программы, уже является адекватной конечной моделью, по которой формально и изоморфно строится программный модуль (что является бесспорным плюсом автоматной технологии). К тому же свойства программной системы в виде автоматов формулируются и специфицируются естественным и понятным образом (в том числе и на языках темпоральных логик). Проверка свойств осуществляется в терминах, которые естественно вытекают из автоматной модели программы.

При автоматном подходе к проектированию и построению программ выделяются две части: системно независимая и системно зависимая. Первая часть реализует логику программы и задаётся системой вза-

Рис. 1: Пользовательский интерфейс банкомата

имодействующих автоматов Мура – Мили. Проектирование каждого автомата состоит в создании по словесному описанию (декларации о намерениях) схемы связей, описывающей его интерфейс, и графа переходов, определяющего его поведение. По этим двум документам формально и изоморфно может быть построен модуль программы (и затем реализована системно зависимая часть), соответствующий автомату.

Таким образом, в качестве задания студентам может быть предложено произвести построение поведенческой модели системы автоматов и спецификации для «автоматной» программы, построенной, например, по иерархической модели [4]. Под поведенческой моделью понимается структура Крипке (или конечная система переходов), которая непосредственно описывает поведение (последовательность действий и состояний при исполнении) программы. При этом описание структуры Крипке должно преследовать следующую цель. Спецификация в виде формул темпоральной логики с учётом структуры Крипке должна охватывать максимально возможное число свойств, которые были (и могут быть) выражены (для программы) на естественном языке.

Далее следует пример модели «автоматной» программы, предлагаемой для верификации методом проверки модели.

Рассмотрим системно независимую часть автоматной программы, которая отвечает за логику управления основными операциями банкомата (пользовательский интерфейс, работа с картой, выдача денег).

Процесс работы банкомата заключается в следующем. После того, как карта вставлена в банкомат, на дисплее появляется надпись «Введите Ваш персональный PIN-код». С помощью оцифрованных кнопок вводится PIN-код. Если PIN-код верен, предлагается выбрать одно из двух действий: «Снятие наличных» (обезличенная кнопка 6) и «Остаток на счете» (обезличенная кнопка 7). Если нажата кнопка «Остаток на счёте», банкомат выдает чек и возвращает карту. Для получения денег нажимается кнопка «Снятие наличных», а на дисплее появляется

Рис. 2: Схема взаимодействия автоматов управления

надпись «Печатать чек?». После выбора нужного варианта появляется надпись «Выберите сумму», а обезличенным кнопкам присваиваются значения: 3000, 2000, 1500, 1000, 500, 200, 100 и «другая». Далее происходит проверка введенной денежной суммы на предмет превышения лимита. Если лимит превышен, предлагается ввести другую сумму. В случае корректности указанной суммы появляется карта и чек, если он был затребован. После взятия карты банкомат выдает деньги. Кнопка «Сброс» прерывает текущую операцию и возвращает карту, а кнопка «Отмена» позволяет ввести заново неверно набранную информацию. Если деньги или карта не были взяты в течение 20 секунд, они захватываются банкоматом.

Банкомат разбивается на чётко выраженные подобъекты управления (пользовательский интерфейс, механизмы захвата карты и выдачи денег), которые имеют свои собственные подсистемы управления. Подсистемы управления представлены в виде конечных автоматов Мура–Мили, выстроенных в иерархию. В результате логическая часть системы управления банкоматом имеет вид иерархической модели системы взаимодействующих автоматов [4]. Автомат находящийся выше по иерархии, управляет своими вызываемыми автоматами путем генерации событий и передачи им управления с этими событиями. Кроме того, автомат следит за состояниями вызываемых автоматов, так как от них могут зависеть его собственные переходы по состояниям. На рис. 2 представлена схема взаимодействия автоматов. События обозначаются символом **e**, запросы к объекту управления (входные переменные) — **x**, текущее состояние некоторого автомата **Ai** хранится в переменной **yi**, а выходные воздействия обозначаются **z**.

Основной автомат $A0$ отвечает за логику взаимодействия с пользователем через панель управления. Автомат $A0$ реагирует на нажатие кнопок на панели инструментов, производит процедуру чтения/записи служебной информации в память банкомата, запрашивает необходимую информацию в банковской сети, и, кроме того, постоянно (по системному таймеру, генерирующему событие $e0$) проверяет возможность (условия) перехода в следующее состояние. Автомат $A0$ имеет вызываемые автоматы $A1$ и $A2$, взаимодействие с которыми осуществляется передачей управления посредством отправки им определённых событий и отслеживанием их текущих состояний. Схема связей и граф переходов автомата управления банкоматом $A0$ представлена на рис. 4.

Автомат $A1$ взаимодействует с автоматом управления $A0$, передавая последнему своё текущее состояние и получая от него события «извлечь деньги» и «проверить переходы». Автомат $A1$ обращается с запросами параметров к датчику денег и таймеру паузы и осуществляет выходные воздействия на механизм выдачи денег и таймер.

Автомат $A2$ отвечает за управление механизмом захвата/возврата карты, обращаясь с запросами к датчику карты и таймеру паузы и воздействуя на таймер и механизм управления захвата/выдачи карты.

Схемы связей и графы переходов автоматов управления механизмами выдачи денег $A1$ и захвата/возврата карты $A2$ представлены на рис. 3.

Для всей системы взаимодействующих автоматов правило перехода в новое состояние в общем случае можно описать следующим образом. После получения события автомат в зависимости от своего текущего состояния реагирует (или никак не реагирует) на событие, опрашивает параметры объекта управления (входные переменные), учитывает состояния вложенных автоматов, затем производит последовательность выходных воздействий, включая и те выходные воздействия, которые необходимо совершить при попадании в новое состояние, и только после этого переводится в новое состояние (которое в случае петли может быть тем же самым).

Выходное воздействие первого типа, направленное на объект управления, считается сразу же осуществлённым после его применения. Выходное воздействие второго типа, представляющее собой передачу управления с событием вызываемому автомату, считается выполненным только лишь после реакции вложенного автомата на это событие, которая заключается в том, что либо автомат переходит в новое состояние (срабатывает один из переходов вызываемого автомата), либо событие игнорируется вызываемым автоматом (ни один из переходов сработать не может). До тех пор пока выходное воздействие второго типа не выполнится, работа (процесс перехода в новое состояние) главного автомата приостанавливается.

Важно отметить, что правила переходов, а точнее условия срабаты-

Рис. 3: Схемы связей и графы переходов автоматов $A1$ и $A2$

вания переходов, должны удовлетворять условию детерминированности (или ортогональности), т. е. при поступлении некоторого события может быть готов к срабатыванию не более чем один переход. Если ни один переход в текущем состоянии при поступлении некоторого события сработать не может (условия переходов не выполняются), то событие игнорируется.

Рассмотрим ряд примеров **темпоральных свойств** автоматной модели системы управления банкоматом, заданных на естественном (русском) языке.

«Сброс в начальное состояние». Если нажата кнопка «Сброс» до перехода банкомата в состояние снятия денег, то в любом случае до того, как банкомат начнёт новый цикл работы (т. е. попросит вставить карту), все автоматы управления вернутся в свои начальные состояния.

«Корректное завершение». Если банкомат перешёл в состояние снятия денег, то он обязательно вернётся в начальное состояние, и до того,

Рис. 4: Схема связей и граф переходов автомата управления банкоматом $A0$

как он запросит вставить карту, автомат выдачи денег и автомат захвата карты вернутся в свои начальные состояния.

«Корректный сброс». После перехода банкомата в состояние «Снятие денег» вплоть до начального состояния «Ожидание карты» нельзя произвести операцию сброса (т. е. завершить транзакцию нажатием кнопки «Сброс»).

«Обнуление поля суммы». После того, как банкомат вышел из состояния «Выбор суммы» без применения кнопки «Сброс», поле суммы обязательно очищается прежде, чем банкомат опять вернётся в это же состояние.

«Возврат карты». Если карта была вставлена, то рано или поздно она будет возвращена (т. е. автомат карты $A2$ обязательно перейдёт в состояние «Карта на выдаче»), если в процессе работы использовались кнопки «Ввод» и «Сброс».

Важно заметить, что построение структуры Крипке, описывающей поведение автоматной модели (системы взаимодействующих автоматов), представляется довольно сложной задачей. В то же время, при предоставлении вместе с автоматной моделью описания структуры Крипке подавляющее большинство студентов успешно (и не без интереса) справляются с задачей формальной спецификации указанных выше темпоральных свойств на языках логик LTL и CTL. При этом ознакомление с технологией автоматного программирования требует небольшого количества времени, отведённого на самоподготовку. Всё это в большой мере способствует пониманию мотивации применения темпоральных логик для спецификации и верификации программных систем.

После того, как поведенческая модель, т. е. структура Крипке, построена для системы взаимодействующих автоматов, и произведена спецификация свойств автоматной модели на языке темпоральной логики, следующими разделами, которые могли бы быть затронуты в рамках курса «Верификация программ», являются способы представления множества состояний структуры Крипке, алгоритмы проверки истинности формулы темпоральной логики для структуры Крипке, а также программы-верификаторы, разрабатываемые и поддерживаемые ведущими научно-исследовательскими лабораториями мира. Студентам предлагается самостоятельно изучить первые две тематики, отсылая их, например, к [2], а также применить для задачи рассмотренного только что вида один из верификаторов по выбору. При этом важным является понимание принципов, на которых работает выбранная программа-верификатор, а также осознание ограничений, накладываемых в верификаторе на размер множества состояний поведенческой модели и вид формул используемой темпоральной логики. Кроме того, необходимо детально изучить интерфейсный формализм (взаимодействующие процессы, сети Петри, системы переходов специального вида и т. д.), с помощью которого в верификаторе задаётся входная проверяемая модель (структура Крипке), т. к. само представление исследуемой модели в рамках этого формализма может оказаться далеко не тривиальной задачей. Всё это требует от исследователя определённых творческих усилий при выполнении анализа корректности программной системы.

Заключение. Таким образом, нами были рассмотрены тематики и задания, предлагаемые в рамках курса «Верификация программ». Выделены направления для самостоятельного изучения с учётом большого объёма внеаудиторных часов, отведённых на самоподготовку. Предложены и обоснованы задачи для самостоятельной и лабораторной работы, направленные на лучшее понимание материалов лекций и предмета дисциплины.

Список литературы

[1] *Грис Д.* Наука программирования /Д. Грис; пер. с англ. М.: Мир, 1984. 416 с.

[2] *Кларк Э. М., Грамберг О., Пелед Д.* Верификация моделей программ: Model Checking. М.: МЦНМО, 2002. 416 с.

[3] *Кузьмин Е. В., Соколов В. А.* Структурированные системы переходов. М.: ФИЗМАТЛИТ, 2006. 178 с.

[4] *Кузьмин Е. В.* Иерархическая модель автоматных программ // Моделирование и анализ информационных систем. Т. 13, 1. ЯрГУ, 2006. С. 27–34.

[5] *Шалыто А. А.* Switch-технология. Алгоритмизация и программирование задач логического управления. СПб.: Наука, 1998. 628 с. http://is.ifmo.ru/books/switch/1/

[6] *Шалыто А. А.* Автоматное проектирование программ. Алгоритмизация и программирование задач логического управления // Известия академии наук. Теория и системы управления. 2000. № 6. С. 63–81. (http://is.ifmo.ru, «Статьи»)

[7] *Шалыто А. А., Туккель Н. И.* SWITCH-технология – автоматный подход к созданию программного обеспечения «реактивных» систем // Программирование. 2001. № 5. С. 45–62. http://is.ifmo.ru/works/switch/1/

Вопросы построения систем дистанционного контроля знаний с применением онтологий

В. А. Курчидис, Д. Ю. Кашалкин, А. С. Назанский

Рассматриваются вопросы организации систем дистанционного контроля знаний, основанные на выделении в их архитектуре онтологических сервисов. Решение семантического аспекта проблемы взаимодействия пользователей (приложений) и систем контроля знаний основывается на использовании модели семантической сервис — ориентированной архитектуры. Предлагаемая архитектура позволяет проводить семантический анализ знаний испытуемых, а также осуществлять эффективный поиск, вызов и публикацию сервисов контроля знаний обучающих систем во Всемирной Сети.

Библиография: 3 названия.

Появление и развитие WWW во многом определили изменение облика архитектур современных информационных систем в разных областях. Эти изменения оказывают влияние на такую важную область современного образования, как дистанционное обучение и, в частности, дистанционный контроль знаний. При построении архитектур, лежащих в основе распределенных информационных систем контроля знаний, важными вопросами являются проблемы публикации и обнаружения во Всемирной Сети таких систем, а также их взаимодействие с другими приложениями и пользователями (проблема интероперабельности). Причем проблема интероперабельности все чаще рассматривается не только на синтаксическом, но и на семантическом уровне, что позволяет передавать как данные, так и их семантику от пользователя к системе и осуществлять последующий смысловой анализ полученных знаний. Для решения данных вопросов предлагается подход к организации систем дистанционного контроля знаний, основанный на выделении в архитектуре таких систем онтологических сервисов.

В качестве средства достижения необходимой гибкости ИТ-инфраструктуры системы рассматривается сервис-ориентированная архитектура (СОА) [1]. В ее основе лежит разделение бизнес-логики приложения и логики представления. При этом бизнес-логика может предлагаться (например, через Интернет) в виде отдельной службы с прикладным интерфейсом, который на программном уровне реализуется в виде соответствующего API (Application Program Interface). Примером реализации СОА является архитектура web-служб-сервисов, доступных в World Wide Web (WWW) [2]. Эта архитектура является важным шагом на пути к "бесшовной" интеграции распределенных программных компонентов с использованием Web-стандартов.

Однако существующие технологии Web-служб, такие как SOAP (Simple Object Access Protocol), WSDL (Web Services Description Language) и UDDI (Universal Description, Discovery and Integration), основанные на использовании общих стандартов, обеспечивают интероперабельность только на синтаксическом уровне. В таких условиях невозможна обработка и передача семантики данных, хотя именно смысловая нагрузка информации и ее анализ являются основополагающими факторами эффективной работы информационной системы дистанционного контроля знаний.

Для решения семантического аспекта проблемы взаимодействия пользователей (приложений) и систем контроля знаний предлагается соответствующая модель семантической сервис-ориентированной архитектуры. Ее использование направлено на автоматизацию решения таких задач, как поиск, композиция, селекция сервисов, что позволяет создавать гибкие механизмы для работы с обучающей системой в распределенной среде.

В предлагаемой архитектуре сервис контроля знаний рассматривается как некоторая функция $F : I \to O$, где I — конечное множество входных параметров (например, ответ студента), а O — конечное множество выходных параметров (с помощью которых происходит оценивание ответов). Помимо этого, функция характеризуется предусловиями (P) и эффектами (E) (в случае, если сервис предполагает изменение состояния предметной области). Множество P характеризует условия, при выполнении которых возможно выполнение сервиса. Множество описывает эффекты от применения сервиса, оказываемые на предметную область.

Для представления семантики множеств I, O, P, E предлагается использование онтологий. Онтология может быть определена как формальная, эксплицитная, совместно разделяемая спецификация концептуализации. Формализм описания онтологий достигается путем использования языков дескриптивной логики (в частности SHOIN(D))[3], которые позволяют создавать модель предметной области (домена) и работать с ней. Модель домена представляется в виде базы знаний $K = (T, R, A)$, где K — база знаний, T — TBox (терминология), R — RBox (отношения), A — ABox (утверждения относительно объектов). Подобное описание позволяет осуществлять вывод неявных знаний (относительно предметной области) с использованием специальных алгоритмов (алгоритм категоризации, проверки на непротиворечивость, и т. д.).

Для автоматизации задач поиска и селекции сервиса контроля знаний разработан алгоритм сравнения запроса пользователя и описания сервиса, применяющий метод категоризации дескриптивной логики к дескриптивным логическим программам, лежащим в основе описания запросов и сервисов. Для проверки выполнения предусловий сервиса и непротиворечивости его эффектов применяется алгоритм непротиворечивости дескриптивной логики.

Предложенная архитектура позволяет проводить семантический анализ знаний испытуемых, а также осуществлять эффективный поиск, вызов и публикацию сервисов контроля знаний обучающих систем во Всемирной Сети. Независимость семантической сервис-ориентированной архитектуры от конкретной предметной области и технологии позволяет применять ее для самых разных предметных областей дистанционного обучения.

Список литературы

[1] Курчидис В. А., Кашалкин Д. Ю., Назанский А. С. Принципы организации Web-служб в системах с семантической сервис-ориентированной архитектурой // Яросл. гос. ун-т им. П.Г. Демидова. Ярославль, 2006. Деп. в ВИНИТИ 20.04.2006, № 526–В2006.

[2] W3C Working Group. Web Services Architecture. http://www.w3.org/TR/2004/NOTE–ws–arch–20040211/

[3] Baader F. et al. The description logic handbook — Theory, implementation and applications // Cambridge University Press, 2003. 583 с.

Проблемы преподавания объектно-ориентированного программирования

Н. С. Лагутина, Ю. А. Ларина

Рассматриваются проблемы преподавания объектно-ориентированного подхода к программированию, возникающие при переходе от обучения методам структурного программирования к изучению новой технологии.

Библиография: 2 названия.

Наука и высшее образование взаимосвязаны в обществе, претендующем на развитие и совершенствование. Эффективность процессов высшего образования и научной деятельности определяется уровнем технологичности, характеризующим каждый из этих двух связанных процессов. Разработка новой научно-методической концепции преподавания современных информационных технологий включает в себя совокупность организационных мероприятий, методов, системных средств, технологических установок, направленных на формирование новых знаний, необходимых для высокоэффективной профессиональной деятельности в условиях чрезвычайно быстро развивающейся сферы компьютерных систем.

Одним из главных препятствий развития информационных технологий в нашей стране является острый недостаток специалистов, умеющих не только правильно распорядиться имеющимся багажом научных знаний, но и способных постоянно осуществлять поиск и быть готовыми оценивать актуальные нововведения в данной области. Современный вуз должен осознавать необходимость активного взаимодействия с внешним миром. Это задача как содержания образования, так и используемых технологий обучения. Важно, применяя ту или иную методологию обучения студентов вносить коррективы в зависимости от образовательной подготовленности студентов, соответствования выпускаемых специалистов запросам современного рынка труда.

Современные информационные технологии — быстро развивающаяся система, в которой постоянно появляются новые программные среды, приложения, подходы к обработке информации, направленные на решение задач совершенствования техники программирования.

В настоящий момент программирование как сфера профессиональной деятельности включает в себя знание различных платформ

(Windows, Unix, Linux, и т. д.), языков программирования (C, C++, Java и др.), интернет-приложений, сетевых технологий, систем разработки и управления базами данных, и т. д. Необходимость подготовки специалистов, владеющих основами современного программирования порождает ряд проблем, связанных с преподаванием соответствующих курсов и дисциплин специализации.

Во-первых, объективная инерционность образовательной системы не соответствует стремительному изменению старых и появлению новых средств и методов программирования. Во-вторых, отсутствует системный подход к повышению квалификации преподавателей в сфере информационных технологий. В-третьих, отсутствуют соответствующие учебно-методические средства (техническая и особенно учебная литература, современные программные продукты, и т. д.).

Концепция объектно-ориентированного программирования появилась в начале 70-х годов прошлого века как систематизированный подход к алгоритмической формализации сложных предметных областей. Двадцать и более лет назад программисты реализовывали свои проекты путем непосредственного написания кода. С возрастанием размера и сложности проектов становилось все яснее, что такой подход неэффективен. Проблема заключалась в непропорциональном возрастании сложности процесса создания самих программ. Как бы ни применялся структурный подход, он не позволяет в достаточной степени упростить большие сложные программы.

Основополагающей идеей объектно-ориентированного подхода является объединение данных и действий, производимых над этими данными, в единое целое, которое называется объектом. Типичная программа на объектно-ориентированном языке состоит из совокупности объектов, взаимодействующих между собой посредством вызова методов друг друга [1]. Разработка программы превращается в процесс эволюционного программирования, который основывается на сохранении целостности объектов программы, при этом внесение изменений не затрагивает внутреннюю организацию существующих в программе объектов.

Одна из задач методики преподавания программирования вообще связана с переходом от концепции структурного программирования к объектно-ориентированному. Главной проблемой является практическое отсутствие методики преподавания объектно-ориентированного программирования. Необходима разработка учебно-методических пособий, объясняющих концепцию объектно-ориентированного программирования и технологию создания соответствующих программ. Следует обратить особое внимание на подбор задач для лабораторных занятий со студентами, которые, с одной стороны позволяли бы усвоить основные понятия и принципы объектно-ориентированного программирования, а с другой стороны, не были бы слишком сложны и велики по объему.

Для изучения объектно-ориентированного подхода на факультете ИВТ выбран язык C++, так как в ряду объектно-ориентированных языков программирования C++ занимает место наиболее концептуально строгого универсального языка программирования, область применения которого легко расширяется от системных задач до прикладных систем. Некоторые другие объектно-ориентированые языки также успешно развиваются, однако пока их распространение в значительной степени уступает C++ [2].

Следует отметить, что язык программирования C++ существенно отличается по своей концепции от базового языка C. Перед большинством студентов, освоивших структурный подход, при котором программа рассматривается как набор последовательно выполняемых инструкций, возникает проблема перехода к представлению программы в виде совокупности объектов, обладающих сходными свойствами, и набором действий, которые можно с ними производить. Проблема усугубляется тем, что переход к новой концепции программирования сопровождается отказом от старых хорошо изученных и успешно применяемых технологий.

Эта трудность порождает проблему составления заданий для лабораторных работ, которые в силу объективных причин не могут быть слишком объемными и сложными, чтобы отразить преимущества объектно-ориентированного подхода. Как правило, предлагаемые задания могут быть успешно решены с использованием структурного подхода, что провоцирует учащихся пойти уже изученным путем и создает проблему методики преподавания объектно-ориентированного программирования.

Основное внимание при обучении объектно-ориентированному программированию следует уделять не вопросам написания кода, а способам организации программы. Чтобы студенты могли лучше понять и оценить значение объектно-ориентированного подхода, необходимо подчеркнуть, в чем проявляется ограниченность структурного стиля построения программ. Мышление в терминах объектов оказывается более простым и наглядным, чем в терминах функций, поскольку программные объекты схожи с объектами реального мира.

Решение обозначенных проблем силами преподавательского коллектива при поддержке со стороны администрации вуза позволит поднять качество обучения специалистов в сфере новых информационных технологий на новую ступень и сделать выпускников университета еще более конкурентноспособными на рынке труда.

Список литературы

[1] *Лафоре Р.* Объектно-ориентированное программирование в С++. СПб.: Питер, 2004.

[2] *Шилдт Г.* Полный справочник по С++. М.: Вильямс, 2004.

Задача адаптации нейросистем в курсе "Концепции современного естествознания"

В. В. Майоров, Е. В. Коновалов, Г. В. Шабаршина

Рассматривается решение задачи по исследованию нейронной сети, составленной из нейроноподобных элементов — нейронных клеточных автоматов. Исследование может быть использовано в качестве дополнительного материала в курсе лекций "Концепции современного естествознания" для студентов факультета ИВТ специальности "Прикладная математика и информатика".

Библиография: 4 названия.

Как нужно учить студента, чтобы учащийся запомнил изучаемый материал и смог применить современные математические методы и программное обеспечение для решения задач науки, экономики, управления и т. д.? Отчасти эту задачу могут решать спецкурсы и спецсеминары. Именно здесь можно проследить ход развития научной мысли, приложить общие методы к конкретным задачам; здесь предоставляется возможность расширения кругозора студента.

Семинар по теории нейронных сетей на факультете информатики и вычислительной техники активно работает более десяти лет. Результаты, полученные участниками семинара, используются при чтении спецкурсов, на спецсеминарах, служат исходным материалом для НИР студентов (курсовые и дипломные работы), магистров (магистерские диссертации), используются аспирантами для работы над кандидатскими диссертациями. Нейронные сети применяются во многих сферах деятельности человека: системы обработки сигналов, системы обработки и распознавания изображений, задачах информационной безопасности, прогнозирования, диагностики и т. д. Поэтому работа студента на спецкурсе. семинаре, несомненно, расширяет научный кругозор учащегося. Одновременно задачи требуют хорошей математической подготовки студента. Например, для исследования моделей, базирующихся на системах уравнений с запаздыванием, используется качественная теория дифференциальных уравнений, в частности, асимптотические методы исследования сингулярно возмущенных уравнений [1]. В других случаях применяются методы анализа разностных уравнений [2]. Далее, компьютерные эксперименты, отражающие теоретические построения, требуют хорошей программистской подготовки студента.

§1. Нейронные клеточные автоматы как формальные нейроны

Понятие нейронного клеточного автомата (НКА) введено в работе [3]. Каждый из рассматриваемых ниже НКА описывается вектором состояний $(u^{(t)}, p^{(t)})$ и параметрами (T_R, q). Здесь $u^{(t)} > 0$ — мембранный потенциал, где t — текущий момент времени. Величина $p^{(t)}$ $(p^0 \geq p^{(t)} > 0)$ называется пороговым значением мембранного потенциала. Если для всех $s \in [t, t+\tau]$ $(\tau > 0)$ выполнено неравенство $u^{(s)} < p^{(s)}$, то для мембранного потенциала и его порогового значения справедливы соотношения: $u^{(t+\tau)} = u^{(t)} \exp(\Theta\tau)$, $p^{(t+\tau)} = p^{(t)} \exp(-p\Theta\tau)$, где $\Theta > 0$, $p > 0$. Если в некоторый момент времени $u^{(t)} = p^{(t)}$, то НКА генерирует мгновенный импульс (спайк) и посылает его всем связанным с ним автоматам. Существенно, что после генерации импульса автомат на время T_R переходит в рефрактерное состояние (состояние абсолютной невосприимчивости НКА к воздействию со стороны других автоматов). В этом состоянии мембранный потенциал имеет значение u^0. В момент генерации импульса значение порога устанавливается равным p^0 (в дальнейшем, как уже говорилось, спадает по экспоненте). Если автомат-приемник не находится в рефрактерном состоянии, то под действием пришедших спайков в прилегающих к нему синапсах выделяются медиаторы. В рефрактерном состоянии медиаторы не выделяются. Медиаторы присутствуют в течение времени T_1 (время жизни медиатора). Считаем, что $T_1 < T_R$. Однако если автомат-приемник под действием медиаторов сгенерировал импульс, то медиаторы на прилежащих к нему синапсах разрушается. В присутствии медиатора на одном синапсе динамика мембранного потенциала подчинена закону $u^{(t+\tau)} = u^{(t)} \exp(\Theta\tau(1 + q))$. Величину q в модели НКА назовем синаптическим весом. Она характеризует эффективность воздействия медиатора. В случае, если на промежутке времени $[t, t + \tau]$ медиатор присутствует на n синапсах, то в последней формуле $q = \sum_{i=1}^{n} q_i$, где q_i — синаптические веса единичных синапсов. Будем называть синапс возбудительным для $q > 0$, для $q < 0$ — тормозным. В изолированном состоянии каждый

НКА периодически генерирует спайки через время T_2, которое определяется из условия $u^0 \exp(\Theta(T_2 - T_R)) = p^0 \exp(-p\Theta T_2)$.

Формально определенная динамика мембранного потенциала НКА соответствует развитию потенциала биологического нейрона. Модель НКА наиболее близка к модели биологического нейрона, построенной на основе дифференциальных уравнений с запаздыванием [4]. НКА сохраняет главные черты этой модели и позволяет избежать технических трудностей, связанных с интегрированием систем дифференциальных уравнений.

§2. Задача о воздействии пачки спайков на автомат

Предположим, что спайк НКА произошел в нулевой момент времени. На вышедший из рефрактерного состояния автомат действует группа из m импульсов, поступающая через синапсы возбудительного типа. Пронумеруем синапсы в порядке начала выделения медиатора в них. Пусть синаптические коэффициенты одинаковы и равны q. Будем считать, что моменты $t_i, i = 1, 2, \ldots, m$, начала выделения медиатора в синапсах упорядочены следующим образом: $T_R < t_1 \leq t_2 \leq, \ldots, \leq t_m$.

Определение 1. *Воздействие назовем пачечным, если медиатор в m-м синапсе появился раньше, чем распался медиатор в первом синапсе, т.е. если $t_m < t_1 + T_1$.*

Определение 2. *Пусть на автомат действует пачка импульсов. Будем говорить, что он непосредственно реагирует на эту пачку, если его спайк происходит после начала выделения медиатора в m-ом синапсе, но раньше, чем распался медиатор в первом синапсе, т.е. если $t_m < t^{sp} < t_1 + T_1$, где t^{sp} — момент спайка данного НКА.*

Введем обозначения $\xi_i = t_i - t_{i-1}, (i = 2, \ldots, m)$ для временных рассогласований между моментами i-го и $(i-1)$-го спайков пачки. Запаздывание спайка НКА относительно последнего спайка пачки обозначим через $\xi = t_{sp} - t_m$.

Лемма 1. *Пусть НКА непосредственно реагирует на пачку спайков, поступающих через синапсы возбудительного типа. Тогда для временного рассогласования ξ справедлива формула*

$$\xi = \left((T_2 - t_m)(p+1) - q \sum_{i=2}^{m} (i-1)\xi_i \right) / (1 + p + mq). \qquad (1)$$

Доказательство. Динамика мембранного потенциала u для $t < t^{sp}$ (пока не произошел спайк) описывается формулой:

$$u^{(t)} = u^0 \exp \left[\Theta \left(t - T_R + q \left(\sum_{i=2}^{m} (i-1)(t_i - t_{i-1}) + m(t - t_m) \right) \right) \right].$$

Момент времени, в который мембранный потенциал равен пороговому значению, т.е.

$$u^{(t^{sp})} = p^{(t^{sp})},$$

соответствует генерации спайка автоматом. Отсюда получаем

$$(t^{sp} - t_m)(1 + p + mq) = (T_2 - t_m)(p + 1) - q\sum_{i=2}^{m}(i-1)(t_i - t_{i-1}).$$

Переходя к обозначениям через временные рассогласования, получим (1).

Здесь следует обсудить условия, при которых реакция автомата на пачку будет непосредственной.

Далее рассматривается задача о периодическом пачечном воздействии на автомат. Пусть одна и та же пачка спайков поступает на синапсы данного автомата с периодом T. Предположим, что автомат непосредственно реагирует на каждую такую пачку. Обозначим через t_i^1, где $i = 1, 2, \ldots, m$, моменты начала выделения медиатора в синапсах под влиянием первой пачки. В ответ на воздействие первоначальной пачки спайков НКА генерирует импульс в момент времени $t_{sp}^1 = t_m^1 + \xi_a^1$ (ξ_a^1 определяется по формуле (1)). Относительно этого момента следующее выделение медиатора в m-ом синапсе начинается спустя время $T - \xi_a^1$. Второй спайк автомата запаздывает относительно момента начала второго выделения медиатора на m-ом синапсе на величину

$$\xi_a^2 = ((T_2 - T + \xi_a^1)(p + 1) - q\sum_{i=2}^{m}(i-1)\xi_i)/(1 + p + mq). \qquad (2)$$

Можно продолжить обсуждение и показать, что перидическое пачечное воздействие навязывает автомату свою частоту.

§3. Организация колебаний в кольцевой структуре из N нейронных клеточных автоматов

Рассмотрим кольцевую структуру из N одинаковых автоматов возбудительного типа, в которой i-й НКА ($i = 1, ..., N$) имеет синапсы, оканчивающиеся на $(i-1)$-м и $(i+1)$-м автоматах. При этом $(N+1)$-й автомат отождествляется с первым, а предшественником первого автомата будем считать N-й автомат. Опишем процесс последовательной генерации импульсов автоматами кольца. Пусть в нулевой момент времени $t_1^1 = 0$ состоялся спайк первого элемента, а спустя некоторое время ξ_2^1 под воздействием спайка первого автомата генерирует спайк второй НКА. Генерацию спайка вторым автоматом отнесем к моменту времени $t_2^1 = \xi_2^1$. В момент времени $t_3^1 = \xi_2^1 + \xi_3^1$ состоялся спайк третьего элемента. Далее, в порядке возрастания номеров, следуют моменты спайков остальных элементов кольца: $t_4^1 < \ldots < t_N^1$. Временные промежутки или рассогласования между моментами спайков $(i-1)$-го и i-го НКА обозначим через $\xi_i^1 = t_i^1 - t_{i-1}^1$, $(i = 2, ..., N)$. Спайк последнего N–го элемента состоится в момент времени $t_N^1 = \sum_{i=2}^{N} \xi_i^1$.

Описанный на промежутке $[t_1^1; t_N^1]$ процесс назовем однократным прохождением волны возбуждения по кольцу или первым тактом прохождения волны.

Если первый автомат к моменту времени t_N^1 уже вышел из рефрактерного состояния, то он попадает под влияние спайка N-го элемента кольца. Мы можем вычислить момент времени t_1^2 генерации второго спайка первым автоматом и указать рассогласование ξ_1^2 между спайками первого и N автоматов. Далее, рассуждения проводятся последовательно для второго, третьего и так далее автоматов, и определяются новые рассогласования ξ_i^2, $i = 2, 3, \ldots, N$, между спайками НКА на втором такте прохождения волны возбуждения по кольцу. Эти величины выражаются через ξ_i^1, $i = 2, 3, \ldots, N$. Применяя эти рассуждения для второго и третьего тактов, мы получим, что временные рассогласования на третьем такте прохождения волны возбуждения по кольцу выражаются через соответствующие рассогласования на втором такте. Возникает итерационный процесс. Можно доказать, что он сходится.

Параметры волны зависят от синаптических коэффициентов воздействия автоматов. Ими можно распорядиться так, чтобы промежутки времени между спайками соседних НКА принимали заранее заданные значения.

Обозначим через ξ_i^0 ожидаемые в стационарном режиме значения промежутков времени между спайками $(i-1)$-го и i-го автоматов $(i = 1, 2, \ldots, N)$.

Теорема 1. *Пусть числа ξ_i^0 $(i = 1, 2, \ldots, N)$ удовлетворяют условиям*

$$0 < \xi_i^0 < T_1; \sum_{i=1}^{N} \xi_i^0 < T_2; \sum_{j=1}^{N} \xi_j^0 - \xi_i^0 > T_R \ (i = 1, 2, \ldots, N), \qquad (3)$$

и синаптические веса определены формулами

$$q_{k-1,k} = (1+p)(T_2 - \sum_{j=1}^{N} \xi_j^0)/\xi_k^0, \qquad (4)$$

$$k = 1, 2, 3, \ldots, N.$$

Тогда существует устойчивый режим работы кольца, при котором спайки автоматов следуют в порядке возрастания номеров и временные рассогласования между спайками $(i-1)$-го и i-го автоматов $(i = 1, 2, \ldots, N)$ принимают значения ξ_i^0.

§4. Модель адаптации нейронного клеточного автомата

Задачу о выборе синаптических весов в нейронных системах называют проблемой обучения. Существуют два принципиально различных подхода к ее решению. Согласно одному из них, веса вычисляются вне нейронной сети, а затем импортируются в синапсы. Нахождение весов обычно сводится к решению некоторой оптимизационной задачи. В случае классических нейронных сетей часто минимизируют суммарное квадратичное рассогласование желаемого и реального состояния выходов нейронов.

В предыдущем разделе для кольцевой нейронной структуры получены значения синаптических весов, гарантирующие существование аттрактора наперед заданной структуры. Таким образом, задача о внешнем обучении уже решена.

Направление, связанное с импортом весов, широко распространено, но при моделировании биологических систем не представляется естественным. Согласно второму подходу к системе обучения, синаптические веса подстраиваются в процессе функционирования нейронной сети.

Вернемся к рассмотрению ориентированной кольцевой структуры, состоящей из N одинаковых возбудительных нейронов. Пусть синаптические веса выбраны в соответствии с формулой (4) и спайк i-го нейрона запаздывает по отношению к спайку $(i-1)$-го нейрона на величину ξ_i^0, где $0 < \xi_i^0 < T_1$. Данное кольцо назовем эталонным или стандартным. Аналогично будем называть принадлежащие ему нейроны.

Наряду с нейронами эталонного кольца рассмотрим еще N нейронов, которые назовем адаптивными. Каждый i-й адаптивный нейрон является постсинаптическим для i-го и $(i-1)$-го нейронов. Синаптический вес воздействия i-го эталонного нейрона на i-ый адаптивный одинаково фиксирован. В свою очередь, синаптический вес воздействия $(i-1)$-го эталонного нейрона на i-ый адаптивный будет меняться в процессе функционирования сети. Цель адаптации — добиться синхронизации функционирования i-го эталонного и i-го адаптивного нейронов.

Задаче можно дать биологическую интерпретацию. *Спайк отдельного эталонного нейрона является относительно слабым сигналом, который может маскироваться в нейронном шуме. Захват эталонным адаптивного нейрона приводит к усилению сигнала. Ясно, что к i-му эталонному могут подстроиться несколько (много) адаптивных нейронов. Это приведет к дальнейшему усилению сигналов. Волновая структура проявится и выделится из нейронного шума.*

Обозначим через q общее, фиксированное значение синаптического веса воздействия i-го эталонного на i-й адаптируемый нейрон, $Q_{i-1,i}$ —

синаптические веса воздействия $(i-1)$-го эталонного на i-й адаптивный нейрон. Эти веса будут меняться в процессе функционирования нейронной сети.

Выбор синаптических весов $q_{i-1,i}$ по формулам (4) гарантирует, что в стандартном кольце устанавливается колебательный режим, в котором спайк i-го нейрона запаздывает по отношению к моменту генерации импульса $(i-1)$-м нейроном на величину ξ_i^0. На каждый i-й адаптивный нейрон действуют образующие пачку спайки $(i-1)$-го и i-го стандартных нейронов. Период воздействия: $T_v = \sum_{i=1}^{N} \xi_i^0 < T_2$. Если синаптический вес q не мал, а вес $Q_{i-1,i}$ не слишком велик, то генерация импульса i-ым адаптивным нейроном будет запаздывать относительно спайка i-го стандартного нейрона. Реакция будет носить непосредственный характер: спайк i-го адаптивного нейрона произойдет раньше, чем завершится спайк $(i-1)$-го стандартного. При увеличениии синаптического коэффициента $G_{i,i-1}$ уменьшается величина запаздывания спайка i-го адаптивного нейрона относительно спайка i-го стандартного автомата. На данном факте основана идея подстройки.

Модификацию синаптического веса $Q_{i-1,i}$ удобно проводить на интервале времени продолжительностью $T_{\text{Ad}} < T_R$ после спайка i-го адаптивного нейрона. На этом промежутке соответствующий синапс не используется, так как нейрон рефрактерен. Одновременно модификацию веса $Q_{i-1,i}$ естественно проводить, только если на данном промежутке наблюдается спайк i-ого стандартного НКА.

Выберем и зафиксируем произвольно i-й адаптивный нейрон. Пусть в нулевой и в момент времени ξ_i^0 начинаются спайки соответственно $(i-1)$-го и i-го стандартных нейронов. Предположим, что спайк i-го адаптивного нейрона запаздывает по отношению к спайку i-го стандартного на величину $\eta_i > 0$. На промежутке времени $t \in [\xi_i^0 + \eta_i, \xi_i^0 + \eta_i + T_{\text{Ad}}]$ произойдет изменение значения синаптического веса $Q_{i-1,i}$. Будем считать, что $\eta_i + T_{\text{Ad}} < T_1$. Пусть до начала адаптации $(t < \xi_i^0 + \eta_i)$ синаптический коэффициент имел значение $Q_{i-1,i}^0$. Легко видеть, что для $t > \xi_i^0 + \eta_i + T_{\text{Ad}}$, пока не начался новый спайк адаптивного нейрона, синаптический вес принимает значение:

$$Q_{i-1,i} = Q_{i-1,i}^0 + r\eta_i, \tag{5}$$

где $r = q^* T_{\text{Ad}}$.

Независимо от адаптивных нейронов по эталонной кольцевой структуре распространяется волна спайков. Пусть на первоначальном такте прохождения волны возбуждения число $\eta_i > 0$ — рассогласование спайков i-го адаптивного и i-го стандартного нейронов, а $Q_{i-1,i}$ — значение синаптического веса после первоначальной подстройки, протекающей после спайка i-го адаптивного нейрона. Из леммы 1 определяется η_i' — новое запаздывание спайка адаптивного нейрона на следующем такте прохождения волны. Пусть $Q'_{i,i-1}$ — значение синаптического веса после модификации, произошедшей во время спайка адаптивного нейрона на этом такте прохождения волны. Оно может быть найдено на основе формулы (5). В результате для величин η_i и $G'_{i,i-1}$ получим соотношения:

$$\eta'_i = ((T_2 - \sum_{j=1}^{N} \xi_j^0 + \eta_i)(p+1) - Q_{i-1,i}\xi_i^0))/(1 + p + Q_{i,i-1} + q),$$

$$Q'_{i-1,i} = Q_{i-1,i} + r\eta'_i. \qquad (6)$$

На формулы (6) будем смотреть как на итерационный процесс, отражающий последовательные такты процесса адаптации. В этом качестве формулы применимы, если выполнены условия, при которых они выписываются.

Очевидно, что отображение (6) имеет неподвижную точку $\eta_i = 0$, $Q_{i-1,i} = q_{i-1,i}$, где число $q_{i-1,i}$ — соответствующий синаптический вес в стандартном кольце. Далее предлагается исследовать вопрос о сходимости итерационного процесса. Эта задача представляет собой серьезное научное исследование.

Список литературы

[1] *Кащенко С. А., Майоров В. В.* Исследование дифференциально- разностных уравнений, моделирующих импульсную активность нейрона // Математическое моделирование, 1993. Т. 5, № 12. С. 13 – 25.

[2] *Бахвалов Н. С.* Численные методы. М.: Наука, 1973. 631 с.

[3] *Шабаршина Г. В.* Проведение возбуждения по кольцевой структуре нейронных клеточных автоматов // Моделирование и анализ информационных систем. Ярославль, 1994, № 2. С. 116 – 121.

[4] *Майоров В. В., Мышкин И. Ю.* Математическое моделирование нейронов сети на основе уравнений с запаздыванием // Математическое моделирование, 1990. Т. 2, № 11. С. 64 – 76.

О преподавании дисциплины "Методы оптимизации" студентам специальности "Прикладная математика и информатика"

Н. Л. Майорова

Излагаются содержание и методика преподавания дисциплины "Методы оптимизации".

Библиография: 11 названий.

С задачами оптимизации приходится встречаться в различных сферах деятельности человека. Каждое разумное действие является в определенном смысле и оптимальным, так как выбирается после сравнения с другими вариантами. Исторически с задачами оптимизации человечество столкнулось уже в древние века. Так, уже давно были решены разнообразные задачи геометрического типа, связанные со свойствами элементарных фигур.

Наиболее простая задача безусловной оптимизации функций многих переменных привлекла внимание математиков во времена, когда закладывались основы математического анализа. Она во многом стимулировала создание дифференциального исчисления. С появлением дифференциального исчисления появилась возможность исследования более сложных задач. Первый общий аналитический прием решения экстремальных задач был разработан Ферма и является одним из первых крупных результатов анализа. Открыт он был, по-видимому, в 1629 году, но впервые достаточно полно изложен в письме к Робервалю в 1638 году. На современном языке прием Ферма сводится к тому, что в точке экстремума $\tilde{x}$ в задаче без ограничений $f(x) \to \text{extr}$ должно иметь место равенство $\nabla f(\tilde{x}) = 0$. Первый намек на этот результат содержится в словах Кеплера из "Стереометрии винных бочек". Точный смысл рассуждения Ферма приобрели через 46 лет, когда в 1684 году появилась работа Лейбница, в которой закладывались основы математического анализа. В своей статье Лейбниц не только получает в качестве необходимого условия соотношение $\nabla f(\tilde{x}) = 0$ (сейчас этот результат называют теоремой Ферма), но и употребляет второй дифференциал для различения максимума и минимума, то есть, по существу, формулирует условия экстремума второго порядка.

Большинство излагаемых Лейбницем фактов было к тому времени известно также и Ньютону, но его работа, завершенная к 1671 году, была опубликована лишь в 1736 году. Важнейшие результаты по минимизации функций были в дальнейшем получены Эйлером и Лагранжем.

Дисциплина "Методы оптимизации" на специальности "Прикладная математика и информатика" до последнего времени изучалась в шестом семестре и естественным образом продолжала изучаемый в пятом семестре курс линейного программирования. Затем по некоторым "техническим" причинам было предложено параллельное изучение обоих курсов в шестом семестре, что создавало некоторые трудности при изучении нелинейной оптимизации без соответствующих знаний вопросов линейной оптимизации. В настоящее время в связи с загруженностью студентов оба эти курса были передвинуты на седьмой и восьмой семестры. Учебный план по дисциплине "Методы оптимизации" предусматривает 54 лекционных часа (по три часа в неделю) и 18 часов для практических занятий (один час в неделю для каждой группы).

Курс является синтетическим, состоящим из трех последовательно изучаемых частей: нелинейного (или математического) программирования, вариационной задачи и задачи оптимального управления. На первых лекциях излагаются задачи минимизации скалярных функций конечного числа переменных. До некоторой степени этот материал является повторением и обобщением знаний, полученных студентами в курсе математического анализа. Приходится также для более полного понимания использовать геометрическую интерпретацию излагаемых фактов, для чего бывает необходимым вспомнить классификацию поверхностей второго порядка, изучаемую в курсе алгебры, понятие линии и поверхности уровня функций многих переменных, а также некоторые сведения из теории квадратичных форм и критерий Сильвестра знакопостоянства квадратичной формы.

Затем изучается переход от необходимых и достаточных условий оптимальности решений оптимизационных задач без ограничений к соответствующим критериям условных задач, то есть задач с различными типами ограничений на независимые переменные. Студентам рассказывается о том, что условная задача является наиболее общим типом оптимизационных задач нелинейного программирования (как, впрочем, и линейного), к которому приводит большой ряд инженерных задач (с ограничениями на управляемые переменные). Такие ограничения существенно уменьшают размеры области, в которой проводится поиск оптимума. На первый взгляд может показаться, что уменьшение размеров допустимой области должно упростить процедуру поиска оптимума. Между тем, напротив, процесс оптимизации становится более сложным, поскольку установленные критерии безусловной оптимизации нельзя использовать при наличии ограничений. При этом может нарушаться даже основное условие, в соответствии с которым оптимум достигается в стационарной точке, характеризующейся нулевым градиентом. Например, безусловный минимум функции $f(x) = (x-2)^2$ имеет место в стационарной точке $x = 2$. Но если задача оптимизации решается с учетом ограничения $x \geqslant 4$, то будет найден условный экстремум, которому соответствует точка $x = 4$. Эта точка не является стационарной точкой функции $f(x)$, так как $\nabla f(4) = 4$. Поэтому для данного типа задач необходимо сформулировать новые условия оптимальности решений задач с ограничениями.

Опыт многолетнего преподавания данной дисциплины свидетельствует о недостаточности учебного времени, отведенного на практические занятия (примерно 8–9 занятий в группе за семестр на все три большие темы данной дисциплины), и невозможности полностью отработать решение каждого типа задач. Во время экзамена "нерешаемой" для многих студентов (в особенности нерадивых) может являться любая простейшая задача на условный экстремум $xy \to \mathrm{extr}, \quad x + y = 1$. Это связано с тем, что учащиеся пользуются критерием Сильвестра для идентификации стационарной точки (является ли она точкой минимума, максимума или в ней вообще нет экстремума), который "не работает" для задач с ограничениями, а справедлива соответствующая теорема – достаточное условие оптимума для данного типа задач. Большие трудности вызывает также проверка условий регулярности при применении общего метода решений задач с ограничениями – метода множителей Лагранжа, так как все условия регулярности формулируются в терминах точки оптимума, поиск которой и ведется. Общеизвестно, что со времен Лагранжа почти на протяжении целого века правило множителей формулировалось с $\lambda_0 = 1$, хотя без дополнительных предположений, например, линейной независимости векторов-градиентов функций-ограничений в точке оптимума это правило неверно.

Отдельной темой рассматривается частный случай общей задачи условной оптимизации со смешанной системой ограничений – выпуклая задача программирования. Студентам поясняется, что доказываемые свойства выпуклых задач имеют важное значение не только в теории, но и в численных методах оптимизации, поскольку большинство существующих численных методов позволяет, вообще говоря, находить лишь локальные решения, а точнее, лишь стационарные точки задачи. Поэтому для выпуклой задачи отыскание стационарной точки означает отыскание решения, причем глобального.

Другой класс изучаемых в данном курсе задач на экстремум – это задачи вариационного исчисления. Следы интереса к ним можно найти и в античной математике, однако подлинное рождение вариационного исчисления произошло в конце 19 века, когда Иоганн Бернулли в 1696 году сформулировал знаменитую задачу о линии наискорейшего ската (брахистохроне). На современном языке она представляет собой бесконечномерную задачу безусловной оптимизации с минимизируемым функционалом специального интегрального вида. В решении задачи о брахистохроне приняли участие лучшие математики того времени: Лейбниц, Ньютон, Якоб Бернулли, Лопиталь. Первое решение задачи принадлежало Я. Бернулли, второе – Лопиталю, третье – Ньютону. После этого в 18 веке Эйлером и Лагранжем были даны общие методы решения задач вариационного исчисления. Условия экстремума первого порядка были получены Эйлером, а второго – Лагранжем и Якоби. Их работу в 19 веке продолжили Коши, Гаусс, Пуассон, Остроградский и другие. Однако решения задач были неполны до недавнего времени, и лишь в конце 19 века работами Вейерштрасса и Гильберта было дано полное решение основных задач вариационного исчисления.

За отведенное учебным планом время на лекциях и практических задачах удается рассмотреть лишь простейшую вариационную задачу и некоторые ее обобщения для случая зависимости интегранта функционала от нескольких неизвестных функций и его зависимости от старших производных искомой функции. Кроме того изучается задача об экстремуме определенного интеграла, когда искомые функции кроме граничных условий и условий непрерывности должны удовлетворять еще дополнительным требованиям, относящимся к поведению функции во всем промежутке интегрирования. Такой тип связанного экстремума возник из частной задачи о нахождении замкнутой кривой заданной длины, ограничивающей наибольшую площадь, и получил благодаря ей название изопериметрической задачи, Интересным примером такого типа задач является задача Дидоны, которая подробно обсуждается со студентами.

Следует отметить, что как задачи линейного и нелинейного программирования, так и простейшая вариационная задача в течение многих лет выносились на государственную аттестацию выпускников математического факультета. В связи с этим от студентов требуется умение решать типовые задачи нахождения экстремалей интегральных функционалов, имеющих интегранты вида $y^{-\frac{1}{2}}\left(1+\dot{y}^2\right)^{\frac{1}{2}}$, $x\left(1+\dot{y}^2\right)^{\frac{1}{2}}$, $\dot{y}^{-2}\left(y^2+1\right)$, $\left(\dot{y}-\sin x\right)^2$, $-y^2+2y\dot{y}+\dot{y}^2$, $\dot{y}^3 y$ и т. п.

В конце сороковых годов 20 века начался новый этап развития методов оптимизации. Как всегда, толчком послужили задачи, поставленные практикой. Возникли динамическое программирование, теория игр, теория оптимального управления, теория дифференциальных игр и другие. В курсе методов оптимизации студенты очень кратко знакомятся с задачей оптимального управления в смысле быстродействия. При этом рассматривается лишь линейный случай. Дается понятие об управляемых объектах, допустимых управлениях, о задаче управления и уравнениях движения объекта. Выводится уравнения Беллмана, определяющее суть метода динамического программирования, которое, в свою очередь, представляет ценное эвристическое средство, на котором базируется метод решения задачи на быстродействие – принцип максимума Понтрягина. На практических занятиях от студентов требуется, кроме аналитического решения задачи, наглядно представить на фазовой плоскости фазовые состояния объекта – так называемый процесс управления. По окончании семестра студенты выполняют расчетно-графическую работу, состоящую из семи задач по всем разделам курса, и сдают курсовой экзамен. Курс "Методы оптимизации" может играть важную роль в подготовке специалиста, развивать его общую культуру.

Список литературы

[1] *Базара М., Шетти К.* Нелинейное программирование. Теория и алгоритмы. М.: Мир, 1982.

[2] *Поляк Б. Г.* Введение в оптимизацию. М.: Наука, 1983.

[3] *Карманов В. Г.* Математическое программирование. М.: Наука, 1986.

[4] *Габасов Р., Кириллова Ф. М.* Методы оптимизации. Минск, БГУ, 1975.

[5] *Моисеев Н. Н., Иванилов Ю. П., Столярова Ю. М.* Методы оптимизации. М.: Наука, 1978.

[6] *Смирнов В. И., Крылов В. И., Канторович Л. В.* Вариационное исчисление. КуБуч, 1933.

[7] *Сухарев А. Г., Тимохов А. В., Федоров В. В.* Курс методов оптимизации. М.: Наука, 1986.

[8] *Рейклейтис Г., Рейвиндран А., Рэгсдел К.* Оптимизация в технике. М.: Мир, 1986, 1-2 тт.

[9] *Алексеев В. М., Тихомиров В. М., Фомин С. В.* Оптимальное управление. М.: Наука, 1979.

[10] *Горстко А. Б., Домбровский Ю. А., Жак С. В.* Методы оптимизации. Метод. указания. МГУ, 1981.

[11] *Дегтярев Ю. И.* Методы оптимизации. М.: Сов. Радио, 1980.

Константы эквивалентности и собственные значения

М. В. Невский

Излагаются различные подходы к получению оценок для констант, стоящих в неравенствах для эквивалентных норм алгебраических многочленов.

Библиография: 4 названия.

Математика едина. Осознание этого важного мировоззренческого факта у студента если и складывается, то постепенно. Поэтому при преподавании следует весьма ценить возможность иллюстрации связи разделов данной математической дисциплины и разделов других дисциплин.

В настоящей работе приводится пример, когда задача, относящаяся к анализу (оценка констант из неравенств для эквивалентных норм на пространствах алгебраических многочленов), решается с помощью методов линейной алгебры (а именно метода собственных значений приведения квадратичной формы к каноническому виду). Этот материал может быть полезен как студентам младших курсов, изучающим линейную алгебру, так и старшекурсникам, изучающим спецдисциплины по теории функций.

В целях доступности для студентов мы ограничиваемся одномерной ситуацией, хотя аналоги всех результатов справедливы и для функций n переменных.

Для непрерывной функции $f : [a, b] \to \mathbf{R}$ мы используем обозначения

$$\|f\|_{C[a,b]} := \max_{x \in [a,b]} |f(x)|, \quad \|f\|_{L_2[a,b]} := \left(\int\limits_a^b |f(x)|^2 dx \right)^{1/2}.$$

Пусть $k \geqslant 0$ — целое. Через Π_k обозначается совокупность многочленов степени $\leqslant k$. Если $g \in \Pi_k$ имеет вид $g(x) = a_0 + a_1 x + \ldots + a_k x^k$, положим

$$\|g\|_1 := \sum_{j=0}^{k} |a_j|.$$

Известно, что любые две нормы, заданные на конечномерном линейном пространстве, являются эквивалентными, см. [2, с. 71]. В частности, это означает, что существуют константы $C_1, C_2 > 0$, зависящие только от k, такие, что для любого $g \in \Pi_k$ имеют место неравенства

$$\|g\|_{C[0,1]} \leqslant C_1\|g\|_1, \quad \|g\|_1 \leqslant C_2\|g\|_{C[0,1]}.$$

Левое из них получается тривиальным образом вместе с точным значением C_1, равным 1. Обозначим через δ_k точное (то есть минимальное возможное) значение константы C_2. В этом тексте отмечаются различные подходы для определения или оценки чисел δ_k.

Теорема 1. *Пусть $k \in \mathrm{N}$, $\lambda_{\min}$ — минимальное собственное значение ганкелевой матрицы порядка $k+1$*

$$\mathbf{C} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \ldots & \frac{1}{k} & \frac{1}{k+1} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \ldots & \frac{1}{k+1} & \frac{1}{k+2} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \ldots & \frac{1}{k+2} & \frac{1}{k+3} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \frac{1}{k+1} & \frac{1}{k+2} & \frac{1}{k+3} & \ldots & \frac{1}{2k} & \frac{1}{2k+1} \end{pmatrix}. \tag{1}$$

Тогда $\lambda_{\min} > 0$, и для любого многочлена $g \in \Pi_k$ имеет место неравенство

$$\|g\|_1 \leqslant \sqrt{\frac{k+1}{\lambda_{\min}}} \cdot \|g\|_{C[0,1]}. \tag{2}$$

Доказательство. Пусть $g(x) = a_0 + a_1x + \ldots + a_kx^k$. Тогда

$$\|g\|^2_{L_2[0,1]} = \int\limits_0^1 \left(a_0 + a_1x + \ldots + a_kx^k\right)^2 dx = \sum_{i,j=0}^{k} c_{ij}a_ia_j, \tag{3}$$

где $c_{ij} = 1/(i+j+1)$ — элементы матрицы $\mathbf{C}$. Очевидно, $Q(a_0, \ldots, a_k) := \|g\|^2_{L_2[0,1]}$ — положительно определённая квадратичная форма на $\mathbf{R}^{k+1}$. По известным результатам линейной алгебры минимальное собственное значение $\lambda_{\min}$ матрицы $\mathbf{C}$ этой квадратичной формы положительно и имеет место неравенство

$$Q(a_0, \ldots, a_k) \geqslant \lambda_{\min} \sum_{j=0}^{k} a_j^2. \tag{4}$$

Из (3) и (4) следует, что

$$\left(\sum_{j=0}^{k} a_j^2\right)^{1/2} \leqslant \sqrt{\frac{1}{\lambda_{\min}}} \cdot \|g\|_{L_2[0,1]} \leqslant \sqrt{\frac{1}{\lambda_{\min}}} \cdot \|g\|_{C[0,1]}.$$

Осталось применить неравенство Коши:

$$\|g\|_1 = \sum_{j=0}^{k} |a_j| \leqslant \sqrt{k+1} \cdot \left(\sum_{j=0}^{k} a_j^2 \right)^{1/2} \leqslant \sqrt{\frac{k+1}{\lambda_{\min}}} \cdot \|g\|_{C[0,1]}.$$

□

Точное значение δ_k может быть получено уже средствами анализа, а именно с помощью следующего неравенства В. А. Маркова, установленного им в 1892 г., см. [1, с. 220], [3, с. 239 – 241], [4, с. 123]. Пусть $T_k(x) = \cos(k \arccos x)$, $x \in [-1, 1]$, — многочлен Чебышёва 1 рода степени k.

Лемма. *Для $g \in \Pi_k(\mathrm{R})$ и $0 \leqslant j \leqslant k$ имеет место неравенство*

$$\|g^{(j)}\|_{C[-1,1]} \leqslant T_k^{(j)}(1) \cdot \|g\|_{C[-1,1]}, \tag{5}$$

в котором при $g = T_k$ достигается равенство.

Теорема 2. *Для $k > 0$ справедливо равенство*

$$\delta_k = 1 + \sum_{j=1}^{k} \frac{2^j}{j!} \cdot \frac{k^2(k^2-1)\dots(k^2-(j-1)^2)}{1 \cdot 3 \cdot \ldots \cdot (2j-1)}. \tag{6}$$

Доказательство. С помощью замены переменных из (5) получается неравенство

$$\|g^{(j)}\|_{C[0,1]} \leqslant 2^j T_k^{(j)}(1) \cdot \|g\|_{C[0,1]}. \tag{7}$$

Пусть $g(x) = a_0 + a_1 x + \ldots + a_k x^k$. Используем соотношения

$$a_j = \frac{g^{(j)}(0)}{j!}, \quad T_k(1) = 1,$$

$$T_k^{(j)}(1) = \sum_{j=1}^{k} \frac{2^j}{j!} \cdot \frac{k^2(k^2-1)\dots(k^2-(j-1)^2)}{1 \cdot 3 \cdot \ldots \cdot (2j-1)}, \ 1 \leqslant j \leqslant k,$$

(последнее приводится в [3, с. 241]). Из них и (7) получаем:

$$\|g\|_1 = \sum_{j=0}^{k} |a_j| \leqslant \left(1 + \sum_{j=1}^{k} \frac{2^j}{j!} \cdot \frac{k^2(k^2-1)\dots(k^2-(j-1)^2)}{1 \cdot 3 \cdot \ldots \cdot (2j-1)} \right) \|g\|_{C[0,1]}. \tag{8}$$

Оценка (8) является точной. Действительно, пусть $p(x) = T_k(1-2x)$. Тогда $p^{(j)}(0) = (-2)^j T_k^{(j)}(1)$, поэтому

$$p(x) = \sum_{j=0}^{k} (-1)^j \frac{2^j}{j!} T_k^{(j)}(1) x^j.$$

Для этого многочлена $\|p\|_{C[0,1]} = 1$,

$$\|p\|_1 = \sum_{j=0}^{k} \frac{2^j}{j!} T_k^{(j)}(1) = 1 + \sum_{j=1}^{k} \frac{2^j}{j!} \cdot \frac{k^2(k^2-1)\dots(k^2-(j-1)^2)}{1 \cdot 3 \cdot \ldots \cdot (2j-1)},$$

поэтому в случае $g = p$ (8) становится равенством. □

Из теоремы 2 получается, что $\delta_1 = 3$, $\delta_2 = 17$, $\delta_3 = 99$, $\delta_4 = 577$, $\delta_5 = 3363$, $\delta_6 = 20369$, и т. д. В частности, *99 есть наименьшая константа, с которой для любых действительных* a_0, a_1, a_2 *и* a_3 *выполняется*

$$|a_0| + |a_1| + |a_2| + |a_3| \leqslant 99 \max_{0 \leqslant x \leqslant 1} |a_0 + a_1 x + a_2 x^2 + a_3 x^3|.$$

Равенство достигается, например, для чисел $a_0 = 1$, $a_1 = -18$, $a_2 = 48$, $a_3 = -32$, которые являются коэффициентами многочлена

$$g(x) = T_3(1-2x) = 4(1-2x)^3 - 3(1-2x) = 1 - 18x + 48x^2 - 32x^3.$$

Точные значения двух первых констант δ_1 и δ_2 могут быть получены и непосредственно, с помощью оценивания значений в равноотстоящих точках отрезка [0,1] и применения неравенства треугольника.

Действительно, пусть $g \in \Pi_1$, $\|g\|_{C[0,1]} \leqslant 1$. Тогда

$$|g(0)| = |a_0| \leqslant 1, \quad |g(1)| = |a_0 + a_1| \leqslant 1,$$

откуда

$$|a_1| \leqslant 2, \quad \|g\|_1 = |a_0| + |a_1| \leqslant 3.$$

Таким образом, $\delta_1 \leqslant 3$. Для многочлена $g_1(x) = T_1(1-2x) = 1-2x$ выполнено $\|g_1\|_{C[0,1]} = 1$, $\|g\|_1 = 3$. Значит, $\delta_1 = 3$.

Если же $g \in \Pi_2$, $\|g\|_{C[0,1]} \leqslant 1$, то

$$|g(0)| = |a_0| \leqslant 1, \ \left|g\left(\frac{1}{2}\right)\right| = \left|a_0 + \frac{a_1}{2} + \frac{a_2}{4}\right| \leqslant 1, \quad |g(1)| = |a_0 + a_1 + a_2| \leqslant 1.$$

Из этих неравенств последовательно получаем:

$$|4a_0 + 2a_1 + a_2| \leqslant 4, \ |3a_0 + a_1| \leqslant 5, \ |a_1| \leqslant 8;$$

$$\left|\frac{a_1}{2} + \frac{3a_2}{4}\right| \leqslant 2, \quad |2a_1 + 3a_2| \leqslant 8;$$

$$\left|\frac{a_1}{2} + \frac{a_2}{4}\right| \leqslant 2, \quad |2a_1 + a_2| \leqslant 8;$$

$$|2a_2| \leqslant 16, \quad |a_2| \leqslant 8.$$

Поэтому

$$\|g\|_1 = |a_0| + |a_1| + |a_2| \leqslant 1 + 8 + 8 = 17,$$

то есть $\delta_2 \leqslant 17$. В этом случае экстремальным является многочлен $g_2(x) = T_2(1-2x) = 1 - 8x + 8x^2$. Для него все предыдущие неравенства обращаются в равенства, в частности $\|g_2\|_{C[0,1]} = 1$, $\|g_2\|_1 = 17$. Значит, $\delta_2 = 17$.

Точное значение δ_3 получить на этом пути уже не удалось.

В заключение отметим, что с помощью двух отмеченных подходов к оценке δ_k получается следующий результат.

Следствие. *Минимальное собственное значение* $\lambda_{\min}$ *ганкелевой матрицы (1) удовлетворяет неравенству*

$$0 < \lambda_{\min} \leqslant \frac{k+1}{\delta_k^2} =$$

$$= (k+1) \cdot \left(1 + \sum_{j=1}^{k} \frac{2^j}{j!} \cdot \frac{k^2(k^2-1)\ldots(k^2-(j-1)^2)}{1 \cdot 3 \cdot \ldots \cdot (2j-1)}\right)^{-2}.$$

Доказательство. Константа, стоящая в (2), не меньше, чем δ_k. □

Представляется, что затронутые вопросы могут быть исходным пунктом самостоятельной работы студентов в рамках различных дисциплин.

Список литературы

[1] *Дзядык В. К.* Введение в теорию равномерного приближения функций полиномами. М.: Наука, 1977. 512 с.

[2] *Кириллов А. А., Гвишиани А. Д.* Теоремы и задачи функционального анализа. М.: Наука, 1979. 384 с.

[3] *Тиман А. Ф.* Теория приближения функций действительного переменного. М.: Гос. изд-во физ.-мат. лит., 1960. 624 с.

[4] *Rivlin T. J.* Chebyshev polynomials. New York etc. John Wiley & Sons, inc. 1990. 249 p.

Минимальные проекторы при линейной интерполяции на $[0,1]^n$, $n=1,2,3$

М. В. Невский

Даётся полное описание минимальных проекторов при линейной интерполяции на отрезке, на квадрате и на кубе. Этот материал может быть использован в спецкурсах по теории приближения функций.

Библиография: 4 названия.

Одним из апробированных разделов специальной дисциплины "Теория приближения" является полиномиальная интерполяция функций. Некоторый материал на эту тему имеется в учебном пособии автора и И. П. Иродовой [1]. Однако с тех пор, как было написано это пособие, ряд результатов был дополнен. В этом тексте приводятся сведения об оптимальной линейной интерполяции, не содержащиеся в учебном пособии [1]; они были получены в работе автора [3]. Соответствующий двумерному случаю проектор "золотого сечения" рассматривался и в [1], но его минимальность там не отмечалась.

Пусть $Q_n = [0,1]^n$. Ниже $n = 1, 2$ или 3. Рассмотрим задачу об оптимальном выборе узлов при интерполяции функции $f \in C(Q_n)$ с помощью линейных функций, то есть многочленов степени $\leqslant 1$ от n переменных. Совокупность этих многочленов обозначается $\Pi_1(\mathrm{R}^n)$. Нас интересуют такие наборы из $n+1$ узлов, принадлежащих Q_n, при которых величина нормы соответствующего проектора $\|P\| := \|P\|_{C(Q_n)\to C(Q_n)}$ является минимальной. Решение этого вопроса приводит к весьма интересным и разнообразным задачам, анализ которых может быть полезен студентам.

Различные формулы для нормы интерполяционного проектора в n-мерной ситуации даны в работе автора [2], а для $n=1$ и $n=2$ приводятся в [1].

1. Случай $n = 1$

Одномерный случай составляет совсем простое упражнение и приводится для полноты. На отрезке $[0,1]$ требуется выбрать два узла интерполяции, минимизирующих $\|P\|$. Из интерполяционной формулы Лагранжа следует, что норма проектора по узлам $x_1, x_2 \in [0,1], x_1 < x_2$, равна

$$\|P\| = \frac{1}{x_2 - x_1} \max\left(x_1 + x_2, 2 - x_1 - x_2\right).$$

Нетрудно видеть, что всегда $\|P\| \geqslant 1$, причём $\|P\| = 1$ тогда и только тогда, когда одновременно $x_1 = 0, x_2 = 1$. Это означает, что в одномерной ситуации имеется ровно один минимальный проектор. Его узлы совпадают с концами отрезка $[0,1]$, а норма равна 1.

2. Случай $n = 2$

Обозначим через τ наименьший корень уравнения $t^2 - 3t + 1 = 0$. Тогда $\tau = (3 - \sqrt{5})/2$. Это число связано с магическим "золотым сечением", так как $\tau/(1-\tau) = 1 - \tau$.

Лемма 1. *Имеет место равенство:*

$$\nu := \min_{0 \leqslant t \leqslant s < 1} \max \left(\frac{(1-s)(1-t)}{1-st}, \frac{s}{1-st} \right) = \frac{1-\tau}{1+\tau} = \frac{\sqrt{5}}{5}. \tag{1}$$

Минимум в (1) достигается лишь при $s = t = \tau$.

Лемма 2. *Пусть $0 \leqslant t \leqslant s < 1$, $P : C(Q_2) \to \Pi_1(\mathrm{R}^2)$ — интерполяционный проектор по узлам $A_1 = (0,0)$, $A_2(1,t)$, $A_3 = (s,1)$. Тогда*

$$\|P\| = \frac{2}{1-st} \max\Big((1-s)(1-t), s\Big) + 1.$$

Доказательство леммы 1 является техническим, оно дано в работе автора [3]. Студентам может быть предложена проверка равенства (1) с помощью компьютера и разбор частных случаев. Доказательство леммы 2 использует схему вычисления нормы проектора, отмеченную в учебном пособии [1], и может быть предложено студентам в качестве полезного упражнения.

Теорема 1. *Для любого интерполяционного проектора $P : C(Q_2) \to \Pi_1(\mathrm{R}^2)$ по трём узлам, принадлежащим Q_2, выполнено точное неравенство*

$$\|P\| \geqslant 2\nu + 1 = \frac{2\sqrt{5}}{5} + 1 = 1.89442719\ldots. \tag{2}$$

Равенство в (2) имеет место лишь для проектора по узлам $(0,0)$, $(1,\tau)$, $(\tau,1)$ и для тех трёх проекторов, узлы которых получаются из указанных поворотами вокруг центра Q_2 на углы, равные $\pi/2$, π и $3\pi/2$. Для других P, кроме этих четырёх отмеченных, в (2) выполняется строгое неравенство.

Доказательство. Используя не очень сложную редукцию, задачу можно свести к оценке нормы проектора по узлам $(0,0)$, $(1,t)$, $(1,s)$. Затем надо использовать последовательно лемму 2 и лемму 1. Так получается неравенство (2). Точность константы $1.89442719\ldots$, а также возможные варианты равенства следуют из леммы 1 и того обстоятельства, что норма проектора не зависит от поворота системы узлов относительно центра квадрата на любой из указанных углов. $\square$

Описанные в теореме 1 экстремальные расположения узлов имеют красивые геометрические свойства. Каждый такой набор состоит из вершины квадрата A_1 и двух точек A_2 и A_3, принадлежащих сторонам квадрата, не содержащим A_1. Точки A_2 и A_3 осуществляют "золотое" деление сторон квадрата, на которых они находятся, причём меньший отрезок прилегает к ближайшей для A_1 вершине квадрата. Это приводит к тому, что оказываются равными площади всех трёх треугольников, которые отсекаются сторонами треугольника $A_1A_2A_3$ от углов квадрата.

Интересно заметить, что точное неравенство (2) даёт характеризацию "золотого сечения" через минимальный проектор.

3. Случай $n = 3$

Трёхмерная ситуация — самая трудная для анализа. Теорема, формулировку которой мы приведём в этом пункте, доказана автором в работе [3] с использованием достаточно тонких геометрических результатов интересной статьи М. Лассака [4].

Теорема 2. *Для любого* $P : C(Q_3) \to \Pi_1(\mathrm{R}^3)$ *имеет место точное неравенство*

$$\|P\| \geqslant 2. \tag{3}$$

Равенство в (3) имеет место лишь для проекторов по узлам $(1-t,0,0)$, $(t,1,0)$, $(1,1-t,1)$, $(0,t,1)$ *при* $t = 0$ *и при* $t = 1/2$ *и по тем узлам, которые сводятся к отмеченным с помощью замены переменных.*

Таким образом, в трёхмерной ситуации любой минимальный проектор P принадлежит одному из следующих двух классов: 1) узлы P расположены в вершинах куба и образуют правильный тетраэдр с рёбрами длины $\sqrt{2}$; 2) узлы P совпадают с серединами противоположных рёбер двух противоположных граней куба (и не имеют общей плоскости). Других минимальных проекторов при линейной интерполяции на кубе $[0,1]^3$ нет.

Студентам можно предложить явно вычислить норму $g(t)$ проектора с узлами

$$(1-t,0,0), (t,1,0), (1,1-t,1), (0,t,1), \qquad 0 \leqslant t \leqslant \frac{1}{2},$$

и затем доказать, что всегда $g(t) \geqslant 2$, причём $g(t) = 2$ лишь при $t = 0$ и $t = 1/2$. Кроме того, в трёхмерной ситуации есть много возможностей для применения компьютера. Некоторые из задач такого рода предлагались студентам в рамках курсовых работ.

Отметитим в заключение, что автором доказано замечательное характеристическое свойство экстремального симплекса S, справедливое для всех $n \in \mathrm{N}$. Это свойство формулируется в терминах гомотетии относительно центра тяжести S. Его варианты для $n = 1, 2$ и 3 также могут быть рассмотрены на занятиях со студентами.

Список литературы

[1] *Невский М. В., Иродова И. П.* Некоторые вопросы теории приближения функций. Ярославль, 1999. 92 с.

[2] *Невский М. В.* Оценки для минимальной нормы проектора при линейной интерполяции по вершинам n-мерного куба // Моделирование и анализ информационных систем. 2003. Т. 10, №1. С. 9 – 19.

[3] *Невский М. В.* О минимальных проекторах при линейной интерполяции на квадрате и на кубе // Ярославль, 2006. 23 с. Деп. в ВИНИТИ 13.06.2006, №786–В2006.

[4] *Lassak M.* Parallelotopes of maximum volume in a simplex // Discrete Comput. Geom. 1999. V. 21. P. 449 – 462.

О специальном курсе "Дополнительные главы топологии"

Е. В. Никулина

Рассматривается вопрос о возможности развития наглядно-образного мышления студентов-математиков 5-го курса в процессе изучения специальной дисциплины "Дополнительные главы топологии".

Библиография: 3 названия.

Рассмотрим особенности мышления математика, а именно то, когда он в своей мыслительной деятельности использует те или иные образные представления.

Прежде всего, внесем ясность в терминологию: правомерно или нет употреблять часто используемый термин "математическое мышление".

С одной стороны, специфика мышления математика, естествоиспытателя, гуманитария присутствует. На основании этого многие исследователи пользуются термином "математическое мышление". Например, Г. Вейль под математическим способом мышления понимает "во-первых, особую форму рассуждений, посредством которых математика проникает в науки о внешнем мире и даже в наши размышления о повседневных делах и заботах, и, во-вторых, ту форму рассуждений, к которой прибегает в своей собственной области математик, будучи предоставленным самому себе" [1].

С другой стороны, отсутствует точное определение понятия "математическое мышление", его психологические основания и отличительные черты. Поэтому употребление термина "математическое мышление", как подчеркивается в [2], чисто условно. Некоторые авторы заменяют термин "математическое мышление" термином "математический стиль мышления".

По мнению Г. Д. Глейзера, "успех на пути исследования структуры "математического мышления" заложен в сопоставлении общих закономерностей мышления с методами математики как объективированным воплощением специфически математических способов мышления" [2]. Следуя данному утверждению, остановимся на некоторых вопросах методологии математики.

Во-первых, математика одна из самых абстрактных наук. Абстрагирование в математике идет значительно дальше, чем, например, в естествознании. В понятии геометрической точки, линии, переменной величины, функции, как и всех других математических понятий вообще, мы отвлекаемся от конкретного содержания и качественных особенностей предметов и процессов. Многие абстракции современной математики возникают через ряд последовательных ступеней отвлечения и последующего обобщения. Именно таким путем образуются все математические структуры.

Во-вторых, математические теории носят строго логический дедуктивный характер.

Таким образом, особенности математического познания подтверждают тот факт, что в структуре мышления математика превалирует абстрактно-логическое мышление. Термин "абстрактно-логическое мышление", на наш взгляд, наиболее характеризует специфику мышления матеатика, учитывая сказанное выше, по сравнению с имеющимся термином "понятийное мышление", поскольку последний одинаково употребим для любого ученого: и для ученого-математика, и для ученого в области гуманитарных наук, который зачастую использует не только абстрактные понятия, но и конкретные.

Соответствуя вышесказанному, весь современный процесс обучения в на математическом факультете классического университета ориентирован на развитие абстрактно-логического мышления студентов.

Но с другой стороны, математические структуры и реальный мир взаимосвязаны друг с другом. С исторической точки зрения развития математики, можно выделить следующие основные моменты этой взаимосвязи.

- Любая математическая структура — это результат пути математического познания от изучения *интуитивно очевидных* математических объектов ко все более обобщенным и абстрактным.

- После того, как математическая структура оказывается окончательно сформированной, начинается поиск ее различных *интерпретаций.*

Таким образом, в первом случае процесс математического познания идет от частного к общему: от конкретных интерпретаций к абстрактной математической структуре, во втором случае — от общего к частному: от математической структуры к конкретным интерпретациям.

С педагогической точки зрения в процессе обучения математике в вузе особенно велика роль наглядных — как правило, геометрических или физических — интерпретаций. Объясняется это следующим.

1. Они помогают понять студентам методы обнаружения нового математического знания.

2. Образы придают математическому знанию наглядность, делая при этом его более "живым" и конкретным, что облегчает его усвоение. Например, геометрические интерпретации зачастую опираются на знания евклидовой геометрии, полученные в школе. Благодаря их использованию обеспечивается осознанность усвоения абстрактного математического материала, изучаемого в вузе.

Данный тезис подтверждают современные представления о применении принципа наглядности в процессе обучения (В. Е. Евдокимов, З. И. Калмыкова, Л. Н. Нуридинов, Е. И. Смирнов), а также современные исследования в области визуализации знаний (О. О. Князева, В. Паронджанов).

Придерживаясь того факта, что мышление — единое, качественно-неоднородное психическое образование, имеющее сложную структуру, выделим основные компоненты в структуре мышления математика: абстрактно-логический компонент и наглядно-образный. В процессе деятельности математика последний проявляется, по крайней мере, в следующих случаях: когда изучение геометрических свойств фигур проводится непосредственно (не применяется алгебраический метод), когда для осознания и лучшего понимания абстрактных математических выкладок используются образы объектов, которые в действительности представить невозможно или в случае, когда сознательно ищется конкретная интерпретация (геометрическая, физическая) некоторой математической структуры.

На современном этапе, когда объем новой информации растет быстрыми темпами, а полученные в вузе знания быстро устаревают и человеку приходится непрерывно учиться в течение всей жизни, актуальным стало развивающее обучение. Развивающее обучение предполагает знакомство студентов с математическими методами. Для каждого раздела математики существуют свои частные методы: методы интегрирования, методы решения дифференциальных уравнений и т. п., но если говорить об общих математических методах, то одним из них может являться использование потенциала наглядно-образного мышления в процессе исследований. Но для того, чтобы образные представления помогали студенту в процессе обучения и в дальнейшей исследовательской работе, необходимо пополнять имеющуюся базу образных представлений и развивать умение оперировать ими. Данную цель преследуют, в частности, читаемые автором для студентов математического факультета 5 курса специальные курсы "Теория изображений" и "Дополнительные главы топологии". Остановимся на содержании последнего.

Специальный курс "Дополнительные главы топологии" читается студентам-математикам в 9 семестре в объеме 60 часов лекций. После изучения сдается зачет. К 5-му курсу студенты уже знакомы с основными понятиями и идеями топологии. Цель данного курса состоит в том, чтобы у студентов были сформированы соответствующие наглядно-образные представления, которые необходимы для лучшего понимания такого раздела математики как топология. В связи с этим основное внимание при изучении данного курса уделяется графическому, наглядному изображению основных объектов топологии. Студенты много рисуют на занятиях, поскольку каждое вводимое даже абстрактное понятие сопровождается большим количеством наглядных интерпретаций, не говоря уже о том, когда предметом непосредственного изучения являются поверхности и кривые. Программа данного курса включает в себя следующие темы.

- История возникновения топологии, предмет ее изучения. Определение топологического пространства, опирающееся на понятие открытого множества. Гомеоморфизм.

- Простейшие топологические инварианты и их наглядные интерпретации.

- Основная теорема топологии поверхностей.

- Определение топологического пространства, опирающееся на понятие близости точки и множества.
- Различные подходы к определению понятия линии (простые дуги, пути, канторовы линии, урысоновские линии).
- Размерность (определение по Урысону, по Лебегу – Брауэру, по Пуанкаре).
- Топологическое произведение. Размерность топологического произведения.
- Симплициальные комплексы. Полиэдры. Многообразия. Наглядные свойства трехмерных многообразий.
- Гомология. Группы гомологий.
- Гомотопия. Фундаментальные группы.
- Различные направления топологии (общая топология, комбинаторная топология, алгебраическая топология, дифференциальная топология, геометрическая топология).

В процессе изучении курса "Дополнительные главы топологии" багаж пространственных представлений студентов значительно пополняется новыми представлениями, развивается пространственное воображение. Как показывает практика, если студент не может четко сформулировать определение некоторого понятия или теорему, но при этом соответствующие наглядно-образные представления у него имеются, то, используя наводящие вопросы преподавателя, он достаточно быстро справляется с поставленной задачей. Это подтверждает тот факт, что если для понимания глубоких математических закономерностей можно использовать образы (предметные, знаковые или внутренние визуальные), то это необходимо делать.

Список литературы

[1] *Вейль Г.* Математическое мышление: Сборник / сост. Ю. А. Данилов; под ред. Б. В. Бирюкова, А. Н. Паршина. М.: "Наука", 1989. 400 с.

[2] *Глейзер Г. Д.* Развитие пространственных представлений школьников при обучении геометрии: Науч.-исслед. ин-т общего образования взрослых Акад. пед. наук СССР. М.: Педагогика, 1978. 104 с.

[3] Закономерности развития современной математики: Методологические аспекты / Отв. ред. М. И. Панов М.: Наука, 1987. 336 с.

О методе Гаусса

Ф. И. Папоркова

Рассматривается модификация метода Гаусса, которая может быть использована при создании программ, решающих системы линейных уравнений.

Изложение курса "Геометрия и алгебра" для студентов специальности "Прикладная математика и информатика" начинается с метода Гаусса решения систем линейных уравнений. Суть метода проста и элегантна.

Пусть дана система линейных уравнений

$$\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{array}\right. \tag{1}$$

Матрица

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & . \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{array}\right),$$

называемая блочной или расширенной, соответствующая заданной системе (1), элементарными преобразованиями приводится к матрице, соответствующей системе линейных уравнений, эквивалентной данной (1), и имеет один из трёх видов:

$$\left(\begin{array}{cccc|c} 1 & 0 & \dots & 0 & c_1 \\ 0 & 1 & \dots & 0 & c_2 \\ \dots & \dots & \dots & \dots & . \\ 0 & 0 & \dots & 1 & c_n \end{array}\right) \tag{2}$$

$$\left(\begin{array}{cccccccc|c} 1 & 0 & 0 & \dots & 0 & t_{1k+1} & \dots & t_{1n} & s_1 \\ 0 & 1 & 0 & \dots & 0 & t_{2k+1} & \dots & t_{2n} & s_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & . \\ 0 & 0 & 0 & \dots & 1 & t_{kk+1} & \dots & t_{kn} & s_k \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & . \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \end{array}\right) \tag{3}$$

$$\left(\begin{array}{cccccccc|c}
1 & 0 & 0 & \dots & 0 & t_{1k+1} & \dots & t_{1n} & s_1 \\
0 & 1 & 0 & \dots & 0 & t_{2k+1} & \dots & t_{2n} & s_2 \\
\multicolumn{8}{c|}{\dots\dots\dots\dots\dots\dots\dots\dots} & . \\
0 & 0 & 0 & \dots & 1 & t_{kk+1} & \dots & t_{kn} & s_k \\
0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & s_{k+1} \\
\multicolumn{8}{c|}{\dots\dots\dots\dots\dots\dots\dots\dots} & . \\
0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0
\end{array}\right) \tag{4}$$

что соответствует:
(2) — единственности решения (1),
(3) — бесконечному множеству решений (1),
(4) — отсутствию решений (1).

Метод Гаусса — метод исключения переменных — поистине универсальный. Решение широкого спектра задач указанного курса, на первый взгляд, никак не связанных с линейными уравнениями, тем не менее легко сводится к решению систем линейных уравнений. Не перичесляя этот длинный список задач, остановимся на простом примере вычисления определителя методом Гаусса.

Пример. Вычислить определитель

$$\begin{vmatrix} 1 & 2 & 1 \\ 4 & 3 & -2 \\ 3 & 2 & 1 \end{vmatrix}.$$

Известно, что элементарными преобразованиями, не изменяющими величины определителя, его можно привести к определителю треугольного вида. Это позволяет легко вычислить определитель по определению:

$$\begin{vmatrix} 1 & 2 & 1 \\ 4 & 3 & -2 \\ 3 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 1 \\ 0 & -5 & -6 \\ 0 & -4 & -2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 1 \\ 0 & -1 & -4 \\ 0 & -4 & -2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 1 \\ 0 & -1 & -4 \\ 0 & 0 & 14 \end{vmatrix} =$$
$$= 1 \cdot (-1) \cdot 14 = -14$$

Однако элементарные преобразования можно заменить умножением матриц. Для перестановки двух строк (i-й и j-й) матрицы $\mathbf{C}$ размеров $m \times n$ достаточно умножать её слева на квадратную матрицу порядка m вида:

$$\begin{array}{c}\begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \\ \begin{array}{ccccccc} 1 & & i & & j & & m \end{array}\end{array},$$

которая получена из единичной матрицы порядка m перестановкой i-го и j-го столбцов. Для умножения всех элементов одной i-й строки матрицы $\mathbf{C}$ на одно и то же число r, отличное от нуля, достаточно умножить матрицу $\mathbf{C}$ слева на матрицу вида:

$$\begin{array}{c}\begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & r & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 \end{pmatrix} \\ \begin{array}{ccccc} 1 & & i & & m \end{array}\end{array},$$

которая получена из единичной матрицы порядка m умножением i-ого столбца на число r.

Чтобы прибавить к j-й строке соответствующие элементы другой i-й строки, умноженные на одно и то же число, достаточно умножить матрицу $\mathbf{C}$ слева на матрицу вида:

$$\begin{array}{c}\begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & r & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \\ \begin{array}{ccccccc} 1 & & i & & j & & m \end{array}\end{array},$$

которая получена из единичной матрицы порядка m прибавлением к элементам i-го столбца соответствующих элементов j-го столбца, умноженных на число r.

Заменяя элементарные преобразования умножением матриц, вышеизложенный пример вычисления определителя

$$\begin{vmatrix} 1 & 2 & 1 \\ 4 & 3 & -2 \\ 3 & 2 & 1 \end{vmatrix} = |\mathbf{C}|$$

можно осуществить следующим образом.

Матрицу $\mathbf{C}$ умножаем слева на матрицу $\mathbf{P_1} = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, что соответствует прибавлению ко второй строке первой строки, умноженной на (-4).

$$\mathbf{P_1} \cdot \mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 1 \\ 4 & 3 & -2 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -5 & -6 \\ 3 & 2 & 1 \end{pmatrix}.$$

Выполняя далее аналогичные действия, то есть, умножая последовательно полученные результаты на необходимые матрицы, получим:

$$\mathbf{B} = \mathbf{P_4} \cdot \mathbf{P_3} \cdot \mathbf{P_2} \cdot \mathbf{P_1} \cdot \mathbf{C} =$$
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 4 & 3 & -2 \\ 3 & 2 & 1 \end{pmatrix} =$$
$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -4 \\ 0 & 0 & 14 \end{pmatrix}$$

Отсюда получаем, что $|\mathbf{B}| = |\mathbf{C}|$.

Учитывая особенности специальности “Прикладная математика и информатика” можно надеяться, что такая модификация метода Гаусса позволит упростить процесс составления некоторых программ при решении определеннного рода задач.

Методика оценивания трудоемкости алгоритма и построения характеристик временной сложности программ

В. С. Рублев

Описывается методика программного прогнозирования времени выполнения кода.

В практической деятельности выпускника специальностей "Прикладная математика и информатика" и "Математическое обеспечение и администрирование информационных технологий" весьма важным является умение снабдить разрабатываемую программу выводом для пользователя информации о времени ее выполнения или о доле уже выполненной работы. Это особенно нужно делать в тех случаях, когда программа может долго выполняться. Существуют 2 подхода к достижению этой цели.

1. В первом случае программа содержит циклы в явном виде, ее можно разбить на непересекающиеся циклы и при этом можно выделить главный цикл – основной по трудоемкости алгоритма. Это особенно легко сделать в тех случаях, когда такой внешний цикл единственный, и потому можно и не оценивать его трудоемкость. Вставим перед этим главным циклом программный код, оценивающий число повторений цикла, скажем, в переменной All, а внутрь главного цикла, но не в его подциклы вставим увеличение счетчика Count главного цикла и вывод процентного отношения Count/All*100 всей работы в текстовом или графическом виде при помощи шкалы прогресса. Если при этом программный код внутри этого цикла выполняется долго (например, более секунды), то в этом коде также найдем вложенный цикл, являющийся главным по трудоемкости, вставим перед его выполнением программный код оценивающий его число повторений, скажем, в переменной All_1, а внутрь его вставим код увеличения счетчика Count, обнуленного перед главным циклом, и вывод процентного отношения Count/(All*All_1)*100 или связанного с этим значением шкалы прогресса. Поступая так же дальше при долгом выполнении кода внутри второго цикла, мы, в конце концов, найдем приемлемое решение. Преимущество такого подхода в его простоте. Однако ряд недостатков затрудняет его использование или делает совсем невозможным:

1) внешний цикл программы может быть не единственным, и в этом случае не всегда является простой оценка трудоемкости каждого внешнего цикла;

2) при одинаковой трудоемкости нескольких внешних циклов нужно уметь оценить долю времени выполнения каждого из этих циклов, что уже требует более тонкого анализа алгоритма и определения характеристик временной сложности;

3) наконец, циклов в явном виде может не быть: например, при выполнении SQL-операторов циклы скрыты в его конструкциях, которые не позволяют вставлять код (да и если бы такое было возможно, мы бы узнали о его результате только после полного выполнения SQL-оператора).

2. Во втором случае программа не содержит циклы в явном виде (см. выше). В этом случае необходимо уметь разработать программный код, который перед выполнением программы или ее части оценивает время выполнения и информирует при этом пользователя. Для этого выполняется следующая последовательность действий:

   1) оценивается трудоемкость алгоритма программы (максимальная, минимальная и средняя);

   2) на основе этих оценок строится вид функций временной сложности программы (максимального, минимального и среднего времени в зависимости от параметров программы);

   3) разрабатываются тесты с различными параметрами программы для определения неизвестных коэффициентов характеристик временной сложности программы;

   4) проводится тестирование программы с целью определения времени ее выполнения в зависимости от набора параметров программы;

   5) проводится численный расчет коэффициентов характеристик временной сложности с целью приближенного определения этих характеристик;

   6) проводится тест производительности компьютера, на котором определены характеристики временной сложности, с целью их коррекции на компьютере использования;

   7) строится код определения временных характеристик в зависимости от параметров программы и код теста производительности для определения коэффициента коррекции и этот код встраивается для выполнения перед вызовом программы (или ее части) с

целью вычисления и выдачи пользователю ожидаемого временного интервала выполнения программы и среднего времени этого ожидания.

Несомненно, является полезным использование такого подхода в различных курсах обучения указанным специальностям, связанным с разработкой и анализом алгоритмов.

1. Анализ алгоритмов и методика оценивания трудоемкости

В качестве основной характеристики сложности алгоритма принято рассматривать число шагов t исполнения алгоритма. Однако t зависит от исходных данных задачи. Зависимость эта в большинстве случаев является весьма сложной, и не всегда ее возможно анализировать непосредственно. В то же время, если мы непосредственно не можем предсказать, сколько будет выполняться алгоритм, то разумно потребовать меньшее: *каковы будут временные рамки выполнения алгоритма* (максимальное и минимальное время), *сколько в среднем будет выполняться алгоритм* (среднее время). Но для любых вариантов задачи время (число шагов) ничем не ограничено. Так, при сортировке массива время, как правило, зависит от длины массива и растет с ростом числа элементов массива.

Принято входные данные V алгоритма характеризовать одним параметром или несколькими параметрами. Одной из таких характеристик является *объем входных данных* – число элементов входных данных. Эта характеристика объективно характеризует входные данные так же, как и число шагов объективно характеризует исполнение алгоритма. В свою очередь, устанавливают зависимость объема входных данных от одного или нескольких параметров, характеризующих задачу. Так, в задаче сортировки массива таким параметром является длина n массива, в задаче нахождения кратчайшего пути в графе объем входных данных связан с числом n вершин графа, а в задаче нахождения кратчайшего остовного дерева объем данных зависит от числа m ребер графа. В других задачах оптимизации на графах в качестве параметров выбирается и число n вершин, и число m ребер.

Так как число шагов алгоритма зависит не только от выбранных в задаче параметров $P = (n, m, \dots)$, характеризующих объем входных данных $V = (P, X)$, но и от других характеристик входных данных $X = (x_1, x_2, \dots)$, то мы введем оценку по всем этим характеристикам. Оценка наибольшего числа шагов, необходимых для выполнения алгоритма, в зависимости от параметров P:

$$t = t(P, X) \leq \max_X t(P, X) \equiv T(P),$$

называется *максимальной трудоемкостью* алгоритма или просто *трудоемкостью* алгоритма. Максимальная трудоемкость дает нам возможность оценить максимальное время, необходимое для исполнения алгоритма. Эта оценка может быть очень завышенной в некоторых случаях. Поэтому важно иметь оценку наименьшего числа шагов, которую мы назовем *минимальной* трудоемкостью:

$$t = t(P, X) \le \min_X t(P, X) \equiv T_{min}(P),$$

и оценку среднего числа шагов, которую мы назовем *средней* трудоемкостью:

$$t = t(P, X) \le \frac{1}{k} \sum_X t(P, X) \equiv T_{cp}(P),$$

где k – число вариантов других характеристик входных данных.

Для простоты мы будем предполагать, что задача имеет 1 параметр n, и обращать внимание на отличия, когда это не так. Трудоемкость алгоритма позволяет оценить время выполнения алгоритма при решении той или иной задачи:

$$T_\phi(X) \le \bar{\tau} \cdot T(n).$$

При этом коэффициент $\bar{\tau}$ статистически определяется для исполнителя или оценивается (что чаще бывает) некоторой константой.

Однако точный вид зависимости $T(n)$ от аргумента n часто очень трудно (практически невозможно) установить. Поэтому вместо установления вида функции для трудоемкости оценивается быстрота роста этой функции при помощи некоторой простой функции $f(n)$. Говорят, что

$$T(n) = O(f(n)),$$

если $|T(n)| \le C|f(n)|$ для всех значений $n > n_0$. Такая оценка роста функции $T(n)$ является односторонней, так как функция $f(n)$ может расти быстрее. Лучше оценивать рост трудоемкости функцией $f(n)$, имеющей тот же порядок роста, т. е. также $|T(n)| \ge C_1|f(n)|$. В этом случае пишут

$$T(n) = \Theta(f(n))$$

и говорят, что рост $T(n)$ оценивается ростом $f(n)$. Наиболее простыми функциями, оценивающими рост трудоемкости, являются полиномы $p(n) = n^k + c_1 n^{k-1} + \cdots + c_k$. В случае $T(n) = \Theta(p(n))$, учитывая, что $p(n) = \Theta(n^k)$, получаем $T(n) = \Theta(n^k)$. Говорят, что в этом случае трудоемкость полиномиальна, или алгоритм имеет полиномиальную трудоемкость. При $k = 1$ $T(n) = \Theta(n)$ и алгоритмы принято называть алгоритмами с линейной трудоемкостью.

Если есть 2 алгоритма A_1 и A_2 решения некоторой задачи и оба имеют полиномиальную трудоемкость, причем $k_1 < k_2$, то говорят, что первый алгоритм имеет меньшую трудоемкость. Но меньшая трудоемкость не означает, что время решения задачи первым алгоритмом будет меньше, чем вторым. Так, пусть

$$90n \le T_1(n) \le 100n, \quad 9n^2 \le T_2(n) \le 10n^2.$$

Тогда при $n < 10$ оказывается, что время решения задачи для первого алгоритма больше, чем для второго. Однако из определения ясно, что найдется такое n_0 (в примере $n_0 = 10$), что время решения задачи при $n > n_0$ будет всегда меньше для первого алгоритма.

Трудоемкость алгоритма может иметь скорость роста меньшую, чем линейная. Например, $T(n) = \Theta(\sqrt{n})$ или $T(n) = \log_2 n$. Но и в этом случае принято говорить о полиномиальной трудоемкости. Алгоритмы, трудоемкость которых растет быстрее любого полинома, принято называть алгоритмами экспоненциальной трудоемкости, даже если скорость роста трудоемкости оценивается более медленной функцией, чем экспонента. Например, экспоненциальными являются все алгоритмы со следующими трудоемкостями:

$$T(n) = \Theta(2^n), \qquad T(n) = \Theta(2^{\sqrt{n}}), \qquad T(n) = \Theta(n^{n^2}).$$

Причина, по которой используются только эти 2 названия трудоемкости (полиномиальная и экспоненциальная), состоит в том, что алгоритмы полиномиальной трудоемкости, как правило, эффективны, если показатель степени у полинома не слишком большой. А алгоритмы экспоненциальной трудоемкости не являются эффективными, так как время вычисления по этим алгоритмам растет «космически» быстро.

Заметим теперь, что при нескольких параметрах входных данных трудоемкость полиномиального алгоритма растет как полином от нескольких аргументов. Например,

$$T(n,m) = \Theta(n \cdot m^2), \qquad T(n,m,k) = \Theta(n^2 k \log_2 m).$$

Процесс получения оценки трудоемкости называется *оцениванием трудоемкости.* Для этого следует анализировать алгоритм с точки зрения быстроты роста числа его шагов при изменении параметров задачи (параметров входных данных).

Прежде всего в алгоритме следует выделить циклы. Если циклов нет, то число шагов линейной структуры алгоритма не зависит от параметров задачи и, следовательно, трудоемкость является константной, т. е. оценивается как $\Theta(1)$.

Циклическая структура алгоритма ведет к повторению выполнения его частей, что влияет на общее число шагов выполнения, т. е. на трудоемкость. Следует оценить для каждого цикла, от каких параметров задачи зависит число повторений цикла и как оно растет с ростом этих параметров.

Если цикл B с числом повторений $n(B)$ вложен в цикл A с числом повторений $n(A)$ и циклы независимы (число повторений цикла B не зависит от выполнения цикла A), то общее число повторений цикла B с учетом повторений цикла A составляет $n(A) \cdot n(B)$. Отсюда правило:

> *для вложенных независимых циклов их трудоемкости перемножаются:* $\Theta(AB) = \Theta(A) \cdot \Theta(B)$.

Если вложенные циклы не являются независимыми, т. е. число повторений внутреннего цикла $n_i(B)$ зависит от номера i повторения при выполнении внешнего цикла, то нужно проанализировать, как зависит общее число повторений внутреннего цикла $\sum_{i=1}^{n(A)} n_i(B)$ от параметров задачи.

Если циклы не являются вложенными, то трудоемкость определяется наибольшей из трудоемкостей циклов

$$\Theta(A + B) = \Theta(A) + \Theta(B) = max\{\Theta(A), \Theta(B)\}.$$

Заметим теперь, что при оценке максимальной трудоемкости следует подбирать такие примеры входных данных для тех или иных параметров задачи, на которых реализуется максимальное число шагов алгоритма. При оценке минимальной трудоемкости следует подбирать примеры, на которых реализуется минимальное число шагов алгоритма. Ввиду сложности некоторых алгоритмов такие примеры не всегда удается построить, но в таких случаях для оценки трудоемкости бывает достаточно примеров и близких по числу операций к максимальному или соответственно к минимальному числу.

Рассмотрим ряд примеров оценивания трудоемкости.

<u>Пример 1</u>. Алгоритм Дейкстры нахождения кратчайшего пути между вершинами a и b связного графа $G(X, U)$. Пусть $n = |X|$ – число вершин графа и $l(x_i, x_j)$ – расстояние от x_i до x_j. Можно привести следующее пошаговое описание алгоритма:

1. Определить начальный список S помеченных вершин, включив в него $a(0)$ (в скобках – пометка), начальное значение последней помеченной вершины $last = a$ и начальный список непомеченных вершин $N = \{x_1(\lambda_1), \dots\}$, записав в скобках расстояние от вершины до $a : \lambda_i = l(a, x_i),$ если $(a, x_i) \in U,$ или $\lambda_i = \infty,$ если иначе.

2. Найти минимальное значение $\lambda_j = \min_i \lambda_i$.

3. Переместить вершину x_j из списка N в список S с пометкой значением $last$ и определить новое значение $last = x_j$.

4. Пересчитать значения λ_i для списка N, положив $\lambda_i = min\{\lambda_i, \lambda_j + l(x_j, x_i)\}$.

5. Если $x_j = b$, закончить выполнение алгоритма, иначе перейти к шагу 2.

В качестве параметра задачи естественно выбрать число n вершин графа. Анализ алгоритма показывает, что он содержит несколько циклов:

1) цикл шага 1, который выполняется $n-1$ раз, т. е. его трудоемкость $\Theta(n)$;

2) выполнение шагов 2 - 5 повторяется, т. е. идет циклически. Назовем этот цикл внешним. Его повторение зависит от входных данных и максимально, когда вершина b включается в список S последней, и минимально, когда она включается сразу же за вершиной a. Поэтому при определении максимальной трудоемкости нужно брать для этого цикла оценку $\Theta(n)$, а при определении минимальной трудоемкости – оценку $\Theta(1)$;

3) внутри внешнего цикла выполняются 4 шага, при этом шаги 3 и 5 не содержат циклов, и потому их трудоемкости оцениваются как $\Theta(1)$. Шаги 2 и 4 содержат внутренние циклы, среднее число повторений которых равно $n/2$. Поэтому трудоемкости каждого из этих циклов можно оценить как $\Theta(n)$. Суммарная трудоемкость тела внешнего цикла равна $\Theta(n) + \Theta(1) + \Theta(n) + \Theta(1) = \Theta(n)$.

Учитывая полученные трудоемкости циклов, можно оценить максимальную трудоемкость как

$$\Theta(n) + \Theta(n) \cdot \Theta(n) = \Theta(n^2),$$

а минимальную трудоемкость – как

$$\Theta(n) + \Theta(1) \cdot \Theta(n) = \Theta(n).$$

Замечание. Внешний и внутренние циклы не являются независимыми – мы обошли эту трудность, взяв среднее число повторений внутренних циклов. Более аккуратно при оценивании максимальной трудоемкости следует брать число повторений внутренних циклов равным $n-1, n-2, \ldots, 1$, и тогда повторение их тела во внешнем цикле составит $(n-1)+(n-2)+\cdots+1 = n(n-1)/2$, что также дает $\Theta(n^2)$. При оценивании минимальной трудоемкости внешний цикл выполняется 1 раз, а внутренний – $(n-1)$ раз, поэтому оценка трудоемкости будет $\Theta(n)$.

Пример 2. Сортировка массива методом пузырька задана процедурой на C++:

$$
\begin{array}{l}
void\ SortP\ (int\ n,\ float\ X[\,])\\
\{\ \ Bool\ b = true;\\
\quad for\ (int\ k = n-1;\ b\ \&\&\ k > 0;\ k--)\\
\qquad for\ (int\ i = 0,\ b = false;\ i < k;\ i++)\\
\qquad\quad if\ (X[i] > X[i+1])\\
\qquad\quad \{\ \ \ float\ r = X[i];\\
\qquad\qquad X[i] = X[i+1];\\
\qquad\qquad X[i+1] = r;\\
\qquad\qquad b = true;\\
\qquad\quad \}\\
\}
\end{array}
$$

Процедура содержит вложенные циклы. Внешний цикл в случае массива входных данных, упорядоченного по убыванию, будет выполняться максимальное число раз: $n-1$, а в случае входного массива, упорядоченного по возрастанию, будет выполняться только 1 раз. Внутренний цикл во втором случае выполняется $n-1$ раз, а в первом случае циклы зависимы, но, как и в предыдущем примере, внутренний цикл в среднем выполняется $n/2$ раз. Поэтому максимальная трудоемкость (входные данные первого случая) оценивается как $\Theta(n)\cdot\Theta(n) = \Theta(n^2)$, а минимальная трудоемкость (входные данные второго случая) – как $\Theta(1)\cdot\Theta(n) = \Theta(n)$.

Пример 3. Оценить трудоемкость следующей процедуры на C++ при $n > 0$:

$$
\begin{array}{l}
long\ f\ (int\ n)\\
\{\ long\ j = 1;\\
\quad while\ (n > 0)\ j <<= n--;\\
\quad while\ (j > 0)\ j >>= 1;\\
\quad return\ j;\\
\}
\end{array}
$$

Процедура имеет 2 цикла, не вложенных один в другой. Но начальное значение параметра j второго цикла определяется в первом цикле, так что некоторая зависимость циклов есть. Первый цикл будет выполняться ровно n раз, если $n > 0$. Поэтому его трудоемкость оценивается как $\Theta(n)$. В результате выполнения первого цикла параметр j примет значение $2^{n+(n-1)+\dots+1} = 2^{n(n+1)/2}$. Второй цикл будет выполняться $\log_2 j = n(n+1)/2$ раз, так как в теле цикла параметр j уменьшается ровно в 2 раза, и, значит, трудоемкость этого цикла оценивается как $\Theta(n^2)$. Поэтому трудоемкость процедуры оценивается как $\Theta(n) + \Theta(n^2) = \Theta(n^2)$.

Пример 4. Оценить трудоемкость следующего SQL-оператора:

$$
\begin{aligned}
&select\ e.EmpName,\ d.DeptName\\
&\quad from\ Emp\ e,\ Dept\ d\\
&\qquad where\ e.DeptNo = d.DeptNo
\end{aligned}
$$

Оператор имеет 2 цикла: по таблице *Emp*, длина которой n и по связанной с ней внешним ключом таблице *Dept*, длина которой m. Если таблицы не проиндексированы совместно, то трудоемкость данного оператора оценивается как $\Theta(nm)$. Если же они проиндексированы совместно по полю DeptNo, которое является первичным ключом для таблицы *Dept* и внешним ключом для таблицы *Emp*, то при эффективной организации выполнения SQL-оператора трудоемкость может оцениваться как $\Theta(n)$, т. е. быть на порядок меньше.

2. Характеристики временной сложности программы

Оценка трудоемкости еще не позволяет ничего сказать о времени выполнения программы, реализующей алгоритм, для тех или иных данных. Мы можем судить только о быстроте роста этого времени. Скажем, если трудоемкость оценивается как $\Theta(n^2)$, то с увеличением параметра n в 2 раза требующееся время будет увеличиваться примерно в 4 раза. Такая качественная оценка не всегда удовлетворяет нас. Допустим, что мы запустили программу и она уже выполняется 10 минут. Может, программа зациклилась и ее следует снять? А если нет, то сколько она еще будет выполняться? Таким образом, оценка времени выполнения программы может стать актуальной. Если программа в начале своей работы выдает, что интервал времени ожидания ее выполнения такой-то и в среднем придется ждать столько-то, то на такие данные можно опираться при принятии решения, следует ли продолжить счет.

Введем 3 статистические функции $t_{min}(n),\ t_{max}(n),\ t_{cp}(n)$, оценивающие время работы программы в зависимости от значения параметра n (при большем числе параметров задачи эти функции имеют большее число аргументов). Их вид зависит от оценки соответствующей трудоемкости.

Рассмотрим вначале общий случай оценки трудоемкости полиномом $T(n) = \Theta(n^k)$. В этом случае выбирается соответствующий вид функции $t(n) = c_k n^k + c_{k-1} n^{k-1} + \cdots + c_1 n + c_0$. Здесь c_i $(i \in \overline{1,k})$ – неизвестные коэффициенты, и мы рассматриваем не только коэффициент c_k, который определяет время при довольно больших значениях параметра n, но и коэффициенты при меньших степенях n, которые могут влиять на время при малых значениях параметра n. Для получения значения коэффициентов c_i $(i \in \overline{1,k})$ следует выбрать ряд различных значений параметра $n = n_1, n_2, \ldots, n_p$, где $p > k$, и для каждого значения параметра n_j выбрать данные, соответствующие этому параметру, и провести вычислительный эксперимент с программой, замерив время ее выполнения τ_j. Затем по методу наименьших квадратов можно найти коэффициенты c_i, которые минимизируют сумму квадратов отклонений полинома в точках n_j от значений τ_j : $\sum_{j=1}^{p}(t(n_j) - \tau_j)^2$. Для более грубой оценки можно взять $p = k+1$ и решить систему уравнений

$$\sum_{i=0}^{k} c_i n_j^i = \tau_j \qquad (j \in \overline{1, k+1})$$

относительно неизвестных c_i.

При оценке трудоемкости логарифмической функцией или произведением полинома на логарифмическую функцию следует и функцию времени $t(n)$ строить в подобном виде. Так, при оценке трудоемкости $T(n) = \Theta(n \cdot \log_2 n)$ следует выбрать вид $t(n) = c_1 n \log_2 n + c_2 n + c_3 \log_2 n + c_4$, а при $T(n) = \Theta(n^2 \log_2 n)$ следует выбрать вид $t(n) = c_1 n^2 \log_2 n + c_2 n^2 + c_3 n \log_2 n + c_4 n + c_5 \log_2 n + c_6$. Проведя достаточное количество вычислительных экспериментов, построив и решив соответствующую систему уравнений, можно найти коэффициенты функции времени $t(n)$.

При оценке трудоемкости $T(n) = \Theta(\sqrt{n})$ следует выбрать вид $t(n) = c_1\sqrt{n} + c_2$, а затем провести вычислительные эксперименты и найти неизвестные коэффициенты, решив соответствующую систему уравнений.

При экспоненциальной оценке трудоемкости $T(n) = \Theta(2^n)$ можно взять соответствующий вид функции $t(n) = c_1 \cdot c_2^n$ и, проведя 2 вычислительных эксперимента для n_1, n_2, найти τ_1, τ_2. Так как $\frac{\tau_1}{\tau_2} = c_2^{n_1 - n_2}$, то $c_2 = (\frac{\tau_1}{\tau_2})^{\frac{1}{n_1 - n_2}}$ и $c_1 = \tau_1 c_2^{-n_1} = \tau_2 c_2^{-n_2}$. Аналогичным образом можно поступать при другом виде экспоненциальной оценки трудоемкости.

В тех случаях, когда оценка трудоемкости алгоритма включает в себя несколько параметров задачи, вид функций времени следует выбирать соответствующим такому виду оценки трудоемкости. Например, при $T(n, m) = \Theta(nm^2)$ следует функцию времени искать в виде $t(n, m) = c_1 nm^2 + c_2 nm + c_3 n + c_4 m + c_5$. И в других случаях следует поступать подобным образом.

Теперь сделаем 3 замечания относительно выбора данных, выбора ряда параметров задачи для вычислительного эксперимента и организации проведения эксперимента.

1. При построении функции максимального времени следует выбирать данные, приводящие к максимальному числу шагов алгоритма. Если не представляется возможным построить такие данные, то следует для каждого значения параметра n_j провести не 1, а

много экспериментов (например, 10) и взять максимальное значение времени. Аналогичным образом следует поступать при построении функций минимального времени и среднего времени.

2. Значения параметров для вычислительного эксперимента следует выбирать такими, чтобы время эксперимента не было слишком малым, так как точность измерения времени имеет порядок 0.02 сек. Поэтому разумно выбирать значения параметров, дающих время эксперимента не менее 10 сек. Однако в ряде случаев данные эксперимента могут составлять значительный объем и возникает проблема ввода таких данных и проблема памяти для них. Первую из них можно решить программной генерацией данных, возможно, с использованием датчика случайных чисел, но вторую проблему так просто одолеть не удается. Поэтому вместо параметров, приводящих к значительным объемам данных, можно выбрать меньшие значения параметров, но вызов процедуры повторять k раз $(k = 10^2, \ldots, 10^6)$ в цикле. Для исключения времени этого цикла следует замерить время пустого цикла с таким же числом повторений и вычесть его из времени цикла вызова процедуры. Полученное значение времени следует разделить на k.

3. Ряд значений параметра (параметров) также нужно выбирать разумным способом. Если трудоемкость оценивается как $\Theta(\log n)$, то выбор значений параметра n, отличающихся друг от друга на несколько единиц, мало скажется на изменении времени эксперимента и потому приведет к большим погрешностям (система уравнений получится близкой к вырожденной). Поэтому в таком случае каждое следующее значение параметра можно выбирать в 2 раза больше предыдущего. Если же трудоемкость имеет экспоненциальную оценку, то, наоборот, не следует выбирать слишком далекие друг от друга значения параметра: так, при $T(n) = \Theta(2^n)$ увеличение параметра на 10 ведет к увеличению времени в 1 000 раз. Будет разумным в этом случае, чтобы следующее значение параметра было на 1–2 больше значения предыдущего.

Помимо полученных функций максимального, минимального и среднего времени выполнения параметров следует также указать производительность процессора, на котором производились вычислительные эксперименты. Для использования этих функций на компьютере с другой производительностью следует всегда вычисленное значение функций умножать на отношение производительностей процессоров, для чего макросом задавать производительность используемого процессора.

В заключение сделаем следующее замечание. При вычислении коэффициентов функций времени может оказаться, что коэффициент при самой старшей степени полинома ничтожно мал, в особенности по сравнению с коэффициентами других степеней. Это может говорить о том, что соответствующая трудоемкость оценена неверно и на самом деле ее оценка ниже. В таком случае следует повторить расчеты с более низкой оценкой трудоемкости и проконтролировать эти расчеты дополнительными экспериментами с более высокими значениями параметров, для чего сравнить время эксперимента и значение функции. Такой контроль необходим и в случаях, когда такой ситуации нет. Если при контроле оказывается, что время эксперимента значительно превышает значение функции, то это говорит о том, что оценка соответствующей трудоемкости занижена. Необходимо и в этом случае пересмотреть оценку трудоемкости и пересчитать коэффициенты функции.

Использование современных методов NLP в образовательном процессе

А. Г. Солопов

Обсуждается применение современных методов нейро-лингвистического программирования (NLP) в образовательном процессе.

Библиография: 5 названий.

NLP – это английская аббревиатура для модного в последнее время течения современной психологии – нейро-лингвистического программирования. По своему происхождению NLP – синтез наиболее эффективных техник современных психологов и психотерапевтов, плюс собственные наработки его создателей.

Авторство NLP приписывают двум людям: Ричарду Бэндлеру и Джону Гриндеру. В середине 70-х годов, когда NLP только зарождалось, оно представляло собой попытку моделирования успешных людей. Когда Бэндлер вел у своих студентов занятия по гештальт-терапии, он максимально точно подражал её основателю, Фрицу Перлзу. Бендлер отрастил бороду, начал курить и даже стал говорить на английском с немецким акцентом. Потом постепенно количество несущественных деталей в поведении и внешности Бендлер стал сокращать, выделяя самое главное.

После Перлза копированию, позднее получившему название метамоделирования, подверглись философ Грегори Бейтсон, детский психиатр Вирджиния Сатир, аниматор Уолт Дисней – известные и знаковые для того времени личности. Огромнейший вклад в современное NLP внесло метамоделирование психолога Милтона Эриксона. Большинство техник данной статьи основываются именно на смоделированных с него шаблонах поведения. Полученные в первый период становления новой науки Бэндлером и Гриндером знания они оформили в виде первой книги о NLP "Из лягушек в принцы" [1].

Основная идея NLP в том, что если есть в чем-то успешные люди, то можно стать в равной им мере успешным, копируя их стратегии и модели поведения в ситуациях, где эти люди преуспевают. NLP не стремится объяснить те или иные происходящие явления и факты, оно вбирает в себя только то, что работает, не стараясь объяснить как.

Применительно к образовательному процессу задачу метамоделирования можно сформулировать следующим образом: если имеется некий успешный преподаватель, можно "снять" с него успешную стратегию ведения учебного процесса и использовать для других дисциплин, другими преподавателями и у других студентов.

С другой стороны, совершенно не обязательно искать идеального преподавателя. В аппарат NLP уже входит большое количество эффективных методов, которые с некоторой адаптацией уже можно использовать в преподавании. Статья посвящена некоторым техникам NLP, которые были использованы мной в преподавании ряда дисциплин факультета ИВТ ЯрГУ.

ТРАНСЫ. МИЛТОН-МОДЕЛЬ ЯЗЫКА. Выражаясь терминами нейрофизиологии, транс – явление бодрствования участков мозга при общей заторможенности большей части мозга. Транс – состояние, близкое к сну. Когда человек засыпает, мозг всегда задерживается в трансе на какое-то незначительное время. Многие сталкиваются с состоянием транса в общественном транспорте, когда мышление как бы "отключается", и человек может просто смотреть на проносящиеся мимо образы окружающего мира.

В контексте обучения транс очень полезен, поскольку вся поступающая в мозг информация попадает напрямую в подсознание. Это должно существенно повышать усваивание студентами информации по сравнению с осознанным её восприятием. Трансы часто используются профессиональными коучерами в проведении тренингов. Тренинги-интесивы, проходящие с утра до вечера в течение нескольких дней, показали, что состояния транса можно поддерживать значительное время. Вполне реальной является цифра в 3-5 пар подряд с небольшими перерывами, что соответствует среднему учебному дню обычного студента.

Транс – естественное состояние мозга, и существуют естественные способы его вызова. Трансами очень плотно занимался уже упомянутый Милтон Эриксон, и один из наиболее часто используемых способов наведения транса при помощи голоса, милтон-модель, назван в его честь.

Милтон-модель подразумевает использование неконкретного языка для перегрузки сознания человека и углубления тем самым трансового состояния. Суть метода в правильном конструировании фраз и постановке речи. По моим собственным наблюдениям, достаточно в начале лекции навести транс, поддержание его по сравнению с другими случаями (имеется ввиду не связанными с преподаванием) дается сравнительно легко. Я связываю это со спецификой постановки общения между преподавателем и студентами, когда большую часть времени говорит лектор. Это позволяет использовать, например, специальный грудной голос, тема которого выходит за рамки статьи.

Используемые трансы являются довольно лёгкими, и поэтому совершенно безопасны с морально-этической точки зрения. В них нельзя делать прямых внушений, и основная их задача – способствовать усвоению студентами материала.

Тема трансов слишком обширна для того, чтобы описать ее в рамках короткой статьи. Подробно о них можно прочитать у основателей NLP в книге [2], или у одного из крупнейших в России специалистов по эриксоновскому гипнозу Сергея Горина в книге [3].

ЯКОРЯ. Якорь – некая условно-рефлекторная связь между двумя явлениями. Самый простой и яркий пример якоря: выделение слюны у собаки при включении света в экспериментах Павлова. В процессе эксперимента известному биологу удалось поставить якорь на вызов физиологического состояния в результате включения внешнего раздражителя – света.

NLP говорит о том, можно вызывать и любые другие состояния: интерес, внимание или какое-то другое. Причем совсем необязательно на свет. Триггерами (событиями, вызывающими нужное состояние) могут быть жесты, мимика, положение человека в пространстве.

Действие вызванных якорями состояний краткосрочно, поэтому лучше всего они подходят для заострения внимания студентов на каких-либо важных фактах.

Чтобы активизировать якорь, нужно сначала его создать. Допустим, для пространственных якорей нужно переместиться в какую-то точку аудитории, куда в обычных обстоятельствах в течение лекции преподаватель не заходит. Например, к двери. Затем рассказать анекдот, интересную историю или что-то еще, что может привлечь внимание группы.

Впоследствии, когда преподаватель будет переходить в эту точку повторно, например, при формулировании важных теорем или определений, это подсознательно будет вызывать повышенное внимание (а, значит, и интерес) у студентов. Подробнее про якори можно прочитать в [4].

РАБОТА С МОДАЛЬНОСТЯМИ. Модальность – система восприятия информации человеком. Различают 6 модальностей: визуальную (через органы зрения), аудиальную (органы слуха), кинестетическую (тактильные, т. е. воспринимаемые нервными окончаниями на поверхности кожи прикосновения), густаторную (вкусовая) и олфлакторную (запахи). Чем больше систем восприятия задействуется человеком при получении информации, тем более полно и глубоко она будет им воспринята. Поэтому чем больше модальностей использует преподаватель при подаче материла студентом, тем лучше материал усвоится.

Густаторная и олфлакторные модальности в силу своей специфичности использовть по понятным причинам очень сложно (в том числе и в учебном процессе), поэтому их часто просто игнорируют.

Из оставшихся трех, на мой взгляд, кинестетическая модальность также малополезна, поскольку людей, в достаточной мере воспринимающих информацию в данном виде немного. По исследованиям психологов, число их не превышает 3-4. Визуальную модальность и аудиальную при передаче информации хорошо использовать примерно в равных пропорциях.

На основе своего опыта, можно дать несколько простых рекомендаций. Во-первых, все важные моменты желательно не просто проговаривать вслух, но и выписывать на доске. Даже если это будет не формула, а текстовая теорема или определение. Таким образом информация будет передаваться в мозг сразу по двум каналам: через глаза и через слух. Во-вторых, полезно использовать и комбинировать в речи речевые шаблоны модальностей. Так, например, звуковая информация будет хорошо восприниматься визуалами (людями, предпочитающими воспринимать окружающий мир глазами), если сопровождать речь шаблонами "как вы можете увидеть, в этой формуле ...", "совершенно ясно, что..", и т.п. Другие визуальные, а также аудиальные и кинестетические шаблоны можно найти в любом учебнике NLP, например в [4] или [5].

В целом NLP позволяет преподавателю расширить свой арсенал средств подачи и представления информации. На мой взгляд, это ни в коем случае не должно использоваться в разрыве от других методов, в частности педагогических, а по сути дожно являться дополнением к ним. Как раздел психологии, NLP сравнительно молод, еще не прошел проверку временем, но активно развивается. С другой стороны, судя по тому, насколько обширно в настоящее время NLP используется в бизнесе, коммерции и маркетинге, оно по-настоящему эффективно. В частности, вопросы использования NLP в образовании, насколько мне известно, практически очень плохо изучены. Это не может не вызывать сожаления, поскольку, на мой взгляд, потенциал использования психотехнологий в высшей школе огромен.

Список литературы

[1] *Bandler R., Grinder J.* Frogs into Princes. Real People Press, 1979.

[2] *Бендлер Р., Гриндер Д.* Шаблоны гипнотических техник Милтона Эриксона с точки зрения НЛП, Симферополь: "Реноме", 1998.

[3] *Горин С.* А вы пробовали гипноз? Санкт-Петербург: "Лань", 1995.

[4] *О'Коннор Д., Сеймор Д.* Введение в НЛП. Челябинск: "Версия", 1998.

[5] *Пьюселик Ф., Льюис Б.* Магия НЛП без тайн. Санкт-Петербург: "Белый кролик", 1994.

О применении формулы суммирования Эйлера – Маклорена

Е. А. Тимофеев

Рассматривается применение формулы Эйлера – Маклорена для функций, производные которых не монотонны.

Библиография: 3 названия.

Связь сумм и интегралов появляется уже в начальном курсе математического анализа. Например, интегральный тест показывает, что для убывающей функции $f(x)$ сумма $\sum_{k=0}^{\infty} f(k)$ и интеграл $\int_0^\infty f(x)\,dx$ сходятся или расходятся одновременно. Формула Эйлера – Маклорена устанавливает более точную связь между суммой и интегралом

$$\sum_{k=0}^{N-1} f(k) = \int\limits_0^N f(x)\,dx + \sum_{k=0}^{m-1} \frac{B_{k+1}}{(k+1)!} \left(f^{(k)}(N) - f^{(k)}(0)\right) + R_m, \qquad (1)$$

где

$$R_m = \frac{(-1)^{m+1}}{m!} \int\limits_0^N B_m(\{x\}) f^{(m)}(x)\,dx, \qquad (2)$$

$B_k(x)$ – многочлены Бернулли, $B_k = B_k(0)$ – числа Бернулли.

Общая теория оценки остатка R_m существует только для функций $f(x)$ с монотонными производными. На основе этих оценок и получаются такие обертывающие ряды как, например, формула Стирлинга,

$$\ln n! = \left(n+\frac{1}{2}\right)\ln n - n + \ln\sqrt{2\pi} - \sum_{k=1}^{m-1}\frac{(-1)^k B_{k+1}}{k(k+1)n^k} + \mathcal{O}\left(n^{-m}\right), \qquad (3)$$

приближение для гармонических чисел

$$H_n = \sum_{k=1}^{n}\frac{1}{k} = \ln n + \gamma - \sum_{k=1}^{m}\frac{(-1)^k B_k}{kn^k} + \mathcal{O}\left(n^{-m-1}\right), \qquad (4)$$

где $\gamma = 0.577215...$ – постоянная Эйлера.

Обычно, как и в приведенных примерах, отбрасываемый остаток не только меньше последнего оставленного слагаемого, но противоположен ему по знаку (поэтому ряды и называют обертывающими).

Случай функции $f(x)$, производные которой не монотонны, обычно не рассматривается.

Цель настоящей заметки — показать, что

- для функций, производные которых не монотонны, приближение, полученное отбрасыванием остатка R_m в формуле Эйлера – Маклорена (1), может иметь погрешность такую же, как первое слагаемое (при любом m);
- с помощью формулы суммирования Пуассона [1, (13.8.4)] можно получить приближенное значение суммы.

1 Пример формального применения

Рассмотрим применение формулы Эйлера – Маклорена (1) для функции

$$f(x) = 1 - \left(1-2^{-x}\right)^s. \qquad (5)$$

Здесь s является параметром, который будем считать достаточно большим и (для простоты) целочисленным. Подчеркнем, что этот пример не является искусственным, а возник в работе автора [2].

Нетрудно видеть, что функция $f(x)$ является монотонно убывающей, $f(0) = 1$, а $f^{(k)}(0) = 0$, если $0 < k < s$. Функция $f(x)$ и ее производные в бесконечности экспоненциально убывают, поэтому $f^{(k)}(\infty) = 0$ для всех k.

Нетрудно проверить, что

$$\int_0^\infty f(x)\,dx = \sum_{k=0}^{m-1}\int_0^\infty 2^{-x}\left(1-2^{-x}\right)^k\,dx = \frac{H_s}{\ln 2}.$$

Формальное применение формулы Эйлера – Маклорена (1) для функции $f(x)$ на отрезке $[0, \infty)$, т. е. отбрасывание остатка R_m, для любого $m < s$ дает одинаковое приближение

$$\sum_{k=0}^{\infty} \left[1 - \left(1 - 2^{-k}\right)^s\right] \approx \frac{H_s}{\ln 2} + \frac{1}{2}. \tag{6}$$

Полученное приближение в этой формуле не является корректным, поскольку в следующем разделе будет показано, что остаток R_m хотя и небольшой (порядка 10^{-6}), но имеет бесконечное число примерно одинаковых максимумов при $s \to \infty$.

2 Формула Пуассона

Следуя Г. Харди [1, (13.8.4)], будем так назвать формулу Эйлера – Маклорена, в которой остаток R_m представлен в следующем виде

$$R_{2m} = \frac{(-1)^m}{2^{2m-1}\pi^{2m}} \int\limits_0^N f^{(2m)}(x) \sum_{k=1}^{\infty} \frac{\cos 2\pi kx}{k^{2m}}\, dx, \tag{7}$$

Отметим, что формула Пуассона получается при подстановке в формулу Эйлера – Маклорена (1) разложения в ряд Фурье многочлена Бернулли $B_{2m}(\{t\})$.

Отметим также, что четные значения m рассматриваются только для простоты.

Применим формулу Пуассона для оценки R_2 с функцией (5). Разложение по следующим производным для рассматриваемой функции не имеет смысла, поскольку $f^{(m)}(0) = 0$, при $1 \leqslant m < s$. Поэтому при замене второй производной на m-ую получим такое же значение остаточного члена.

$$R_2 = -\frac{1}{2\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \int\limits_0^{\infty} f''(x) \cos 2\pi kx\, dx.$$

Подставляя в полученную формулу $f'(0) = 0$, вычисляя первый интеграл и интегрируя по частям второй, получим

$$R_2 = -\sum_{k=1}^{\infty} \frac{1}{\pi k} \int\limits_0^{\infty} f'(x) \sin 2\pi kx\, dx.$$

Подставляя производную, получим

$$R_2 = \sum_{k=1}^{\infty} \frac{s \ln 2}{\pi k} \int\limits_0^{\infty} 2^{-x} \left(1 - 2^{-x}\right)^{s-1} \sin 2\pi kx\, dx.$$

Заменяя синус на разность мнимых экспонент, получим

$$R_2 = \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{is\ln 2}{2\pi k} \int_0^{\infty} 2^{-x}\left(1-2^{-x}\right)^{s-1} e^{2\pi ikx}\,dx.$$

При замене $y = 2^{-x}$ интегралы сводятся к бета-функции, поэтому имеем

$$R_2 = \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{is}{2\pi k} \frac{\Gamma\left(1+\dfrac{2\pi ki}{\ln 2}\right)\Gamma(s)}{\Gamma\left(s+1+\dfrac{2\pi ki}{\ln 2}\right)}.$$

Применяя формулу $\Gamma(x+1) = x\Gamma(x)$, перепишем полученное выражение как

$$R_2 = -\frac{1}{\ln 2} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{\Gamma\left(\dfrac{2\pi ki}{\ln 2}\right)\Gamma(s+1)}{\Gamma\left(s+1+\dfrac{2\pi ki}{\ln 2}\right)}. \tag{8}$$

Применяя формулу Стирлинга, нетрудно вывести

$$\frac{\Gamma(s+1)}{\Gamma\left(s+1+\alpha\right)} = s^{-\alpha} + \mathcal{O}(s^{-1}).$$

Отметим, что оценка является равномерной по α.

Подставляя эту оценку для $\alpha = \frac{2\pi ki}{\ln 2}$ в (8), получим

$$R_2 = -\frac{1}{\ln 2} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \Gamma\left(\frac{2\pi ki}{\ln 2}\right) s^{-2\pi ki/\ln 2} + \mathcal{O}(s^{-1}), \tag{9}$$

На рис.5 приведены графики функций R_2 и ее приближения по формуле (10) (пунктир) в зависимости от s

Гамма-функция убывает по мнимой оси довольно быстро ([3, 8.332.1]),

$$|\Gamma(iy)| = \frac{\pi}{y\,\mathrm{sh}\,\pi y}$$

поэтому для получения приближенной формулы достаточно оставить только первый член ряда ($|k| = 1$). Вычисление гамма-функции с применением пакета *Mathematica* дает

$$\Gamma\left(1+\frac{2\pi i}{\ln 2}\right) = (3.1766226452 - 3.7861079986i)10^{-6}.$$

Приближение полученное из (9) отбрасыванием слагаемых с $|k| > 1$ и подстановкой этих чисел имеет вид

$$R_2 \approx \frac{10^{-6}}{\pi}\left[3.78610799\cos\left(2\pi\log_2 s\right) + 3.17662264\sin\left(2\pi\log_2 s\right)\right]. \tag{10}$$

Точность этого приближения примерно 10^{-8}.

Рис. 5: Графики остатка R_2 и его приближения по формуле (10) (пунктир) в зависимости от s.

Список литературы

[1] *Харди Г.* Расходящиеся ряды. М.: ИЛ, 1951.

[2] *Тимофеев Е. А.* Состоятельная оценка энтропии мер и динамических систем // Мат. заметки. 2005. Т.77. №6. С. 903 – 916.

[3] *Градштейн И. С., Рыжик И. М.* Таблицы интегралов, сумм, рядов и произведений. М.: Наука, 1971.

Как предупредить формальное усвоение знаний

В. Ф. Чаплыгин

Предлагаются способы, препятствующие формальному усвоению знаний. Они предусматривают некоторые методические приемы при чтении лекций и написание учебных пособий.

"Уразумения существа математики, именно этого не может дать логический формализм, взятый сам по себе." (Н. Бурбаки "Архитектура математики")

Лекция, учебное пособие, учебник — три составные части, из которых студент получает математические знания. Каждая из этих трех частей выполняет свою функцию и они скорее дополняют, нежели дублируют друг друга. Если говорить о математическом анализе, то по этой дисциплине имеется достаточное количество учебников. Хотелось бы выделить два из них: "Курс дифференциального и интегрального исчисления" в трех томах Г. М. Фихтенгольца и "Краткий курс математического анализа" А. Я. Хинчина. Они существенно отличаются друг от друга по объему и содержанию, но их объединяют очевидные достоинства, главное из которых — их педагогическая выдержанность. Они написаны доступным языком на разумном уровне строгости. Кроме того, в первом из них содержатся сведения по истории развития математического анализа и большое количество примеров как иллюстративного характера, так и на приложения к механике и физике. К сожалению, некоторые учебники очень формально излагают теоретический материал, содержат мало примеров и контрпримеров, вопросов, которые могут помочь студенту понять, правильно ли он воспринимает материал.

Из учебных пособий заслуживает внимания учебное пособие под ред. В. Ф. Бутузова "Математический анализ в вопросах и задачах" в двух частях (для функций одной и нескольких переменных). Это учебное пособие сопровождает курс математического анализа, написанный В. А. Ильиным и Э. Г. Позняком, "Основы математического анализа" в двух томах, предназначенный для физических и физико-математических факультетов университетов. В пособии содержится большое число вопросов для самоконтроля, примеры решения задач, уделяется внимание не только чисто "техническим" задачам, но и обсуждаются различные теоретические "тонкости". В частности, выясняются возможности применения той или иной теоремы, существенность ее условий, комментируются определения.

Но, наверное, лекция остается важнейшим источником и средством, благодаря которому студент приобретает знания. Лекция ни в коем случае не должна быть формальной копией какой-либо части учебника или учебного пособия. Известно, какую важную роль в математике играют определения, от их правильного усвоения зависит дальнейшее понимание математических доказательств. Введение основных понятий необходимо предварять пропедевтической работой по формированию предпонятия. Сошлемся на мнение известного французского математика М. Фреше: "Если что-нибудь действительно необходимо, так это — уничтожение догматического метода: не давать никаких определений, не указав, как они возникли, для чего они нужны, как они применяются." С введенным формальным определением необходимо поработать, проанализировать, проиллюстрировать различными примерами, предложить студентам привести свои примеры.

Не менее важным является комментарий к теореме, который должен проводиться перед ее формулированием, помогая понять ее условия, их существенность, что легко сделать с помощью контрпримеров. После завершения доказательства теоремы следует описать область возможных приложений, следствия, предупредить о наиболее вероятных ошибках и проведении неоправданных аналогий. Заслуживает внимания пример, показывающий, что теорема Лагранжа для функций комплексного переменного не выполняется. Взяв функцию $f(x) = e^{ix}$ на отрезке $[a;b]$ и предположив, что формула Лагранжа $f(b) - f(a) = f'(c)(b-a)$ верна, приходим к равенству $|\sin\frac{b-a}{2}| = \frac{|b-a|}{2}$, которое справедливо лишь в случае, когда $a = b$.

Лекция позволяет установить непосредственный контакт со слушателями, увидеть их реакцию, обратить внимание слушателей на самые важные моменты изложения. Чрезвычайно полезно время от времени озадачивать студентов вопросом "почему" и приучить их самих ставить перед собой этот вопрос при чтении математической литературы. Это совершенно необходимо, так как иначе студент становится на путь механического, чисто формального запоминания материала. Задача лектора не дать возможности студенту скатиться к формализму, помочь ему понять суть.

Наверное, пришло время отказаться от традиционной формы чтения лекций. В принципе дело можно свести к чтению установочной лекции перед изучением той или иной темы, консультационной помощи при работе студента над этой темой и итоговой лекции после завершения изучения темы. Все остальное студент должен самостоятельно почерпнуть из учебников и учебных пособий.

Автор предполагает создать учебное пособие по математическому анализу для студентов-физиков, близкое по духу пособию под ред. В. Ф. Бутузова. В пособии не будут систематически излагаться основы математического анализа, доказательства теорем, а будут приведены задачи преимущественно теоретического характера, вопросы, способствующие сознательному, активному и глубокому усвоению студентами этой фундаментальной математической дисциплины. Хотя автор не придерживается при чтении лекций какого-то одного учебника по математическому анализу, он считает, что в пособии лучше использовать терминологию, определения, формулировки теорем, в основном, из учебника В. А. Ильина и Э. Г. Позняка. Часть вопросов и задач можно заимствовать из задачника Б. П. Демидовича, пособий под ред. В. Ф. Бутузова, В. А. Садовничего "Задачи и упражнения по математическому анализу" в двух частях, замечательной книги Б. Гелбаума и Дж. Олмстеда "Контрпримеры в анализе". В задуманном пособии будут представлены основные разделы курса: теория вещественного числа, предел числовой последовательности и функции, непрерывные функции и их свойства, теория числовых и функциональных рядов, интегрируемость функций. В этих разделах должны содержаться основные определения с комментариями и примерами, формулировки основных теорем с обсуждением существенности их условий, следствия, приложения, различные примеры и контрпримеры, опровергающие ошибочные утверждения. Рассмотрим в качестве примера краткое содержание некоторых разделов.

В разделе "Теория вещественного числа" рациональное число определяется как отношение целого числа m к натуральному числу n, причем дробь $\frac{m}{n}$ должна быть несократимой. Приводится алгоритм обращения обыкновенной дроби в десятичную и десятичной периодической дроби в обыкновенную. Между множеством рациональных чисел и десятичных периодических дробей устанавливается взаимно-однозначное соответствие (дроби, период которых состоит из нулей, не рассматриваются, их можно заменить дробями, период которых содержит лишь цифру 9). Иррациональным числом будет названо число, которое не представимо в виде отношения $\frac{m}{n}$, где m — целое, n — натуральное числа, и, следовательно, оно представимо в виде бесконечной десятичной непериодической дроби. Таким образом, доказательство иррациональности заданного вещественного числа сводится либо к невозможности его представления в виде отношения $\frac{m}{n}$, либо к доказательству непериодичности бесконечной десятичной дроби, которой оно задается. Напоминаются определения точных нижней (inf) и верхней (sup) граней числового множества и теоремы о существовании inf и sup. Приведем образцы задач и вопросов этого раздела:

1) Почему при обращении обыкновенной дроби в десятичную не может получиться бесконечная непериодическая дробь?

2) Докажите, что числа $\sqrt{2}$, $\sqrt[3]{5}$, $\log_2 3$, $\sqrt{2} \pm \sqrt{3}$ - являются иррациональными.

3) Докажите, что числа $\alpha = 0,123456\ldots$ (выписываются подряд все натуральные числа) и $\beta = 0,10110111011110\ldots$ являются иррациональными.

4) Приведите несколько примеров иррациональных чисел.

5) Найдите 131-й десятичный знак в представлении чисел $\frac{5}{6}$, $\frac{1}{35}$, $\frac{1}{7}$.

6) Докажите, что сумма, разность, произведение, отношение рациональных чисел являются рациональными числами.

7) Можно ли утверждать, что сумма, разность, произведение, отношение иррациональных чисел являются иррациональными числами?

8) Останется ли справедливой теорема о существовании точной верхней грани, если удалить из множества всех вещественных чисел хотя бы одно?

9) Докажите, что если $X = \{x\}$ — ограниченное множество, то $-sup\{x\} = inf\{-x\}$.

10) Докажите, что между любыми двумя неравными вещественными числами находится бесконечно много различных между собой как рациональных, так и иррациональных чисел.

11) Укажите конкретное а) иррациональное число, лежащее между числами 1,09 и 1,1; б) рациональное число, лежащее между числами π и $\sqrt{10}$;

12) Что значит: найти сумму чисел $\sqrt{2}$ и $\sqrt{3}$?

Приведем часть вопросов и задач из раздела "Предел числовой последовательности". В начале раздела приводятся два определения числовой последовательности:

1) $\lim_{n\to\infty} x_n = a \Leftrightarrow \forall\varepsilon > 0\ \exists n_\varepsilon \epsilon N\ \forall n > n_\varepsilon$ выполнено неравенство $|x_n - a| < \varepsilon$;

2) число a является пределом числовой последовательности x_n, если последовательность $\alpha_n = x_n - a$ является бесконечно малой.

Доказывается эквивалентность этих определений. Предлагаются вопросы, помогающие лучше их усвоить.

1) Можно ли утверждать, что число a является пределом числовой последовательности x_n, если в любой ε-окрестности этой точки $(a-\varepsilon, a+\varepsilon)$ содержится бесконечно много членов последовательности?

2) Сохранит ли первое определение смысл, если в нем заменить а) первый квантор $\forall$ на $\exists$; б) квантор $\exists$ на $\forall$; в) второй квантор $\forall$ на $\exists$. Приведите на каждый из этих случаев пример последовательности.

3) Приведите примеры ограниченных последовательностей, не имеющих предела.

4) Приведите пример последовательности, для которой числа $1; 0; -1$ являются частичными пределами (т. е. пределами некоторых подпоследовательностей).

5) Докажите, что последовательность x_n имеет пределом число a тогда и только тогда, когда любая ее подпоследовательность имеет пределом число a.

6) Имеет ли предел последовательность а)$\frac{1}{n}\sin\frac{n\pi}{2}$; б) $\cos\frac{2\pi n}{3}$?

7) Может ли сумма, разность, произведение последовательностей не имеющих предела, иметь предел? Тот же вопрос для случая, когда одна из них сходится, а другая — нет.

8) Приведите пример, когда предел последовательности y_n равен 0, но предел отношения последовательностей $\frac{x_n}{y_n}$ существует.

В разделе "Предел функции в точке" приводится определение предела по Коши:

$\lim_{x\to a} f(x) = l \Leftrightarrow \forall\varepsilon > 0 \quad \exists\delta > 0 : 0 < |x-a| < \delta \Rightarrow |f(x) - l| < \varepsilon$,
в котором следует особо остановиться на том, что $|x-a| > 0$, и показать на простых примерах нахождение числа δ по выбранному ε:

а) $f(x) = 2^{-\frac{1}{(x-1)^2}}$, если $x \to 1$;

б) $\phi(x) = x\sin\frac{1}{x}$, если $x \to 0$;

в) $\psi(x) = x^2$ при $x \to 0$.

При работе с определением по Гейне необходимо объяснить, что только в случае, когда для <u>любой</u> последовательности $x_n \to a$ ($x_n \neq = a$) последовательность $f(x_n)$ сходится к одному и тому же числу l, то $l = \lim_{x\to a} f(x)$, а если найдутся хотя бы две последовательности $x'_n \to a$ и $x''_n \to a$, для которых $f(x'_n) \to l_1$ и $f(x''_n) \to l_2$, причем $l_1 \neq l_2$, то в точке $x = a$ функция $f(x)$ предела не имеет. В пособии предполагается доказать равносильность определений Коши и Гейне (в программу экзамена это обычно не включается).

Часть вопросов и задач в этом разделе аналогична тем, что приведены в предыдущем. Отметим лишь два примера.

Пусть функция $f(x)$ в рациональной точке $\frac{m}{n}$ (дробь несократимая) равна $\frac{1}{n}$, т. е. $f(\frac{m}{n}) = \frac{1}{n}$, а если x — иррационально, то $f(x) = 0$. Доказать, что функция $f(x)$ имеет предел в любой иррациональной точке.

Второй пример весьма поучителен как в теоретическом смысле, так и с практической точки зрения. В общем случае он формулируется так:

Пусть $x(t) \to a$ при $t \to t_0$ и $f(x) \to l$ при $x \to a$. Следует ли отсюда, что $\lim_{t\to t_0} f(x(t)) = l$? Ответ отрицательный, что видно из следующего примера.

Пусть $f(x) = sgn^2 x = \left\{ \begin{array}{l} 1,\ x \neq 0 \\ 0,\ x = 0 \end{array} \right.$, и $x(t) = \left\{ \begin{array}{l} t,\ t \in (\frac{1}{2k}; \frac{1}{2k-1}), \\ -t,\ t \in (\frac{1}{2k+1}; \frac{1}{2k}), \\ 0,\ t = \frac{1}{k}, \end{array} \right.$

где $k \in Z$. Легко видеть, что $\lim_{t\to 0} x(t) = 0$, $\lim_{x\to 0} f(x) = 1$, но $\lim_{t\to 0} f(x(t))$ не существует. Этот пример придется кстати при изучении непрерывности композиции двух непрерывных функций, и он же предупреждает о возможных ошибках при вычислении пределов. Довольно часто при вычислении пределов некоторых функций студенты заменяют функцию ее пределом, что не всегда обоснованно. Например, в случае, когда предел оставшегося выражения является бесконечным. Такого рода примеры будут приведены в пособии.

Функциональные стороны и формы экзамена

В. Ф. Чаплыгин, Н. Б. Чаплыгина

Рассматриваются методические вопросы, связанные с проведением экзаменов по математическим дисциплинам, приводятся достоинства и недостатки различных существующих форм экзамена.

Учебный процесс состоит из нескольких составляющих частей: аудиторные занятия, самостоятельная работа студента, консультации, контрольные работы и завершается экзаменом или зачетом. Экзамен является важной составляющей учебного процесса, его назначение многофункционально, а формы разнообразны. Безусловно, экзамен — это форма контроля, промежуточного или итогового, в зависимости от его положения в учебном плане. На многих факультетах проводится промежуточная аттестация студентов младших курсов в течение семестра. Например, на физическом факультете университета на протяжении последних десяти лет в середине первого семестра проводится так называемый промежуточный экзамен по математическому анализу. Цель его — помочь студентам выяснить, насколько успешно они усваивают эту дисциплину, дать возможность преподавателю скорректировать в зависимости от результата экзамена темп изложения, доступность его и наглядность, обратить внимание студентов на правильность использования математического языка, определений, символики, проведения доказательств, логики рассуждений. Промежуточный экзамен и экзамен в конце первого семестра для студентов чрезвычайно важны. Студенты первого курса еще не освободились от школьных привычек и очень слабо представляют, что такое экзамены в университете. Они даже спрашивают преподавателя: "А Вы дадите нам билеты?" Конечно, студент должен иметь на руках программу экзамена, его содержание, знать, в каком объеме от него требуется знание того или иного вопроса, какие вопросы или темы имеют особо важное значение, какие задачи будут предлагаться на экзамене, какими литературными источниками можно пользоваться, и т. д.

Давая студентам программу экзамена (не билеты!), экзаменатор заботится о полном охвате курса, о сохранении его логики. При этом часть материала может и не войти в билеты, но в процессе подготовки должны быть изучены студентом. И совершенно недопустимо преследовать студента и тем более наказывать его оценкой за то, что они не следуют при ответе на экзамене конспекту лекций, но вопрос излагает правильно, глубоко, показывает его понимание. И как это, может быть, ни парадоксально звучит, экзамен помогает экзаменатору-лектору увидеть свои недоработки. Если он на занятиях неудачно провел доказательство какой-либо теоремы или неудачно изложил тему — это безусловно проявится на экзамене. Вряд ли подлежит сомнению тот факт, что "зазубривание" теорем и последующее их механическое воспроизведение на экзамене без понимания смысла никакой пользы для обучения не приносит. А вот предложить студенту привести пример, когда утверждение теоремы не верно, если не все условия теоремы выполнены, опровергнуть ложное высказывание — очень важно. Так, при доказательстве теоремы Ролля можно поставить вопрос, останется ли в силе заключение теоремы, если опустить одно из трех условий: $f(a) = f(b)$, непрерывность функции на отрезке $[a, b]$, ее дифференцируемость на интервале (a, b); или какие условия позволяют доказать непрерывность справа функции распределения вероятностей $F_\xi(x) = P(\xi \leq x)$ и почему у нее может не быть непрерывности слева. Уместно задать вопросы такого рода: 1) может ли сумма (произведение) разрывных функций быть непрерывной функцией; 2) может ли композиция непрерывной и разрывной функций быть непрерывной функцией. Аналогичные вопросы могут быть заданы для дифференцируемых функций.

Как известно, существуют различные формы экзамена: устная, более традиционная, письменная, модульная (экзамен сдается по темам или частям с рейтинговым накоплением оценки в течение семестра) и др. Наверное, каждая из указанных форм имеет как достоинства, так и недостатки. Письменная форма проведения экзамена, безусловно, имеет множество достоинств. В случае письменного экзамена можно поставить студентов в равные условия по объему, содержанию, трудности предлагаемых вопросов и задач, предоставить им одинаковое время на выполнение экзаменационного задания. Студент имеет возможность сосредоточиться, ему никто не мешает, его никто не перебивает и не исправляет, как это бывает на устном экзамене. Написанная работа — объективный документ. Работы студентов легко сравнить и правильно оценить. И что немаловажно, имея на руках экзаменационную работу студента, можно ее обсудить, объяснить, показать автору работы, в чем он прав, а в чем состоят его ошибки, сравнить его работу с работами других студентов, если это необходимо. Но существенным недостатком такой формы проведения экзамена является то, что задание в большей степени составляется из задач на нахождение решения, на вычисления и меньше задач на доказательство различных утверждений, в том числе теорем, поскольку их выполнение состоит из цепочек логических выкладок. Для того чтобы правильно провести доказательство, нужно иметь достаточный уровень логического мышления, хорошо представлять, что в условном предложении является предпосылкой и логическим следствием, отличать необходимые условия от достаточных. Последнее требование у большой части студентов вызывает затруднения. Поэтому в письменном изложении решения задач на доказательство зачастую бывают ошибки такого рода, которые не позволяют преподавателю, проверяющему экзаменационную работу, принять решение, где кроется ошибка: в логическом рассуждении или в его неверном изложении. На наш взгляд, студент, ориентированный на письменную форму экзамена, слабее готовит теоретический материал, а больше внимания уделяет решению вычислительных задач. Если речь идет о математических дисциплинах, то доказательства утверждений — это фундаментальная часть, на которой и основываются все методы, алгоритмы и приемы решения вычислительных задач. Кроме того, в настоящее время сложились объективные условия, препятствующие проведению письменного экзамена в большой аудитории — это электронные средства связи, широко используемые студентами и сводящие к минимуму все положительные стороны письменной формы проведения экзамена.

По одной и той же дисциплине, например, по математическому анализу, экзамен может иметь разные акценты для студентов различных специальностей: математиков, математиков-прикладников или физиков. При устной форме приема экзамена студента легко поправить в случае оговорки или механической ошибки в процессе теоретической выкладки или вычислении, в формулировке определения или теоремы. Здесь можно выяснить и поправить смысловые ошибки изложения студентом материала. Нередко одно и тоже математическое понятие поразному формулируется или обозначается. Самый простой и довольно часто встречающийся пример такой ситуации, когда некоторые авторы придают словам "или" и "либо" различный логический смысл. Письменный экзамен не позволит в таких ситуациях правильно понять работу студента, а на устном экзамене преподаватель с помощью дополнительных вопросов может такие ситуации выявить и правильно оценить рассуждения студента. Однако выдержать при этом один и тот же уровень требований к каждому студенту трудно. Преподаватель, принимая экзамен у целой группы студентов, неизбежно корректирует свои требования к знаниям студентов, даже не желая этого, в зависимости от того, как студенты отвечают, как долго продолжается экзамен (а по времени такой экзамен зачастую затягивается), преподавателю на протяжении всего экзамена очень трудно быть эмоционально ровным. Студенты, представ на экзамене перед преподавателем лицом к лицу, по-разному на это реагируют и многие испытывают психологические трудности, которые мешают сосредоточиться, преодолеть появляющийся перед экзаменом страх. Если любой экзамен для студента — причина стресса, то устный экзамен — в большей степени.

Третья форма приема экзамена — модульная, ее условно можно назвать ступенчатой. Такой экзамен может осуществляться через коллоквиумы и контрольные работы и требует от преподавателя дополнительных усилий и времени. По нашему мнению, дробление курса на сравнительно мелкие темы, отчеты студента по отдельным разделам (модулям), конечно, облегчает задачу студента, но, возможно, помешает ему охватить курс в целом, понять его глубину, проследить все важные взаимосвязи. Такая форма экзамена, возможно, единственный шанс отчитаться по дисциплине для весьма "слабых" студентов.

Это, пожалуй, основные формы экзамена, хотя могут существовать и иные, например, способ проведения экзамена в виде тестирования, который вызывает довольно много споров. Несколько слов об использовании студентами учебниками, конспектами и другими литературными источниками на экзамене. На письменном экзамене, включающем теоретические вопросы, изложенные в учебниках, конечно, нельзя допускать использование таких подсобных материалов. Но любой преподаватель понимает, что студенту, не усвоившему материал и не подготовившемуся к экзамену, не помогут никакие учебники. Для того чтобы оценить ответ студента, достаточно задать несколько дополнительных уточняющих вопросов по ходу изложения материала. И многие преподаватели именно так и поступают, не контролируя подготовку к устному ответу студента на экзамене, не тратя сил на выявление списывающих студентов. Но, насколько показывает практика, студент, ориентированный на лояльное отношение преподавателя к использованию учебного материала "в твердой копии", как правило, хуже готовится к экзамену. У него нет стимула выучить и понять теоретические вопросы, поскольку есть иллюзия, кажущаяся надеждой, сделать это прямо на экзамене. Если преподаватель не хочет жестко ограничивать использование студентом подсобного материала, может быть, стоит разрешить пользоваться лишь тем материалом, который студент подготовил сам, написал своею рукой. Такая подготовка к экзамену также приносит пользу.

Очень важным является отбор материала, включаемого в программу экзамена. Вряд ли целесообразно пытаться вложить в нее все, что было рассказано в учебном процессе в течение семестра или учебного года. Не рекомендуется включать материал, рассказанный накануне экзамена, который не проработан на практических, лабораторных или семинарских занятиях. Обязательно необходимо позаботиться о том, чтобы были усвоены те вопросы, которые будут использоваться в дальнейшем в этой дисциплине или в других. Достаточно дискуссионным является мнение: есть ли такие вопросы, за незнание которых следует ставить неудовлетворительную оценку. К примеру, если студент совершенно не показал понимания определений предела числовой последовательности и функции в точке, производной и т. п., которые лежат в основе всех теоретических выкладок в курсе математического анализ, имеет ли право экзаменатор поставить в таком случае студенту положительную оценку?

Если говорить о системе оценок вообще, то необходимо заметить, что она крайне негибкая, маловыразительная, недостаточно выполняет дифференцирующие функции. Конечно, работа студента в семестре и оценка результатов этой работы на экзамене связаны и взаимозависимы. Мы считаем, что если студент добросовестно занимался в семестре, выполнял домашние и лабораторные задания, успешно писал контрольные работы, то на экзамене сомнения преподавателя в выборе оценки должны быть разрешены в пользу студента, что особенно важно в случае выбора оценки между "неудовлетворительно" и "удовлетворительно". Однако, на наш взгляд, допустимо оценить студента высоко на экзамене ("хорошо" или "отлично"), если он при недостаточной работе в течение года, добросовестно подготовился к экзамену и достаточно хорошо овладел материалом.

Остановимся еще на одном моменте. Принято считать, что экзамен должен носить обучающий характер. По нашему мнению, стрессовое состояние подавляющего числа студентов на экзамене не дают им возможности воспринять замечания экзаменатора, с которыми они скорее всего согласятся, совершенно не вникнув в их суть. Экзамены на младших и старших курсах могут иметь различный характер. На первых курсах ему присущ больше обучающий характер, а на последних — он может приобретать форму диалога преподавателя со студентом. Здесь преподаватель задает вопросы более обобщающего характера, выясняет вместе со студентом суть проблем (решенных и нерешенных), перспективы дальнейших исследований, возможные приложения. Сказанное в особой степени относится к экзаменам по дисциплинам специализации.

Все формы проведения экзамена имеют и плюсы и минусы. Выбор формы зависит в большей степени от самого преподавателя, который учитывает все стороны той или иной формы экзамена и свои личные качества, и в меньшей степени — от желания самих студентов, так как их много. Одному преподавателю подходит, например, устная форма, другому — письменная. Экзамен, проводимый в устной форме, должен проходить в спокойной доброжелательной обстановке. Целесообразно последнюю переэкзаменовку проводить комиссией. Преподаватель сам может стать инициатором создания такой комиссии. Во время проведения устного экзамена не стоит торопиться выставлять "неуд" студенту после первого неверного ответа даже на "самый важный" вопрос курса: главная цель обучения — добиться понимания материала, возможно причиной неверного ответа являются какие-либо психологические факторы.

Все формы экзамена могут и должны существовать. И возможно, кто-то предложит иные способы проведения экзамена, заслуживающие внимания и которые можно обсуждать.

Узлы и косы. Алгебраический подход

В. К. Шалашов

Излагается алгебраический подход для решения проблемы узлов.

Узел K есть образ единичной окружности S^1 при непрерывном и взаимно-однозначном отображении $F:\ S^1 \to \mathbb{R}^3$. Чтобы избежать подвохов, связанных с топологическими трюками, мы сосредоточим внимание на полигональных узлах, являющихся объединением конечного числа отрезков. Будем говорить, что узел K_0 эквивалентен узлу K_1, если существует непрерывное отображение

$$F:\ S^1 \times [0,1] \to \mathbb{R}^3$$

такое, что $F(S^1, 0) = K_0$, $F(S^1, 1) = K_1$ и для каждого фиксированного значения t, $0 \leqslant t \leqslant 1$, отображение $s \mapsto F(s,t)$, $s \in S^1$, есть взаимно-однозначное и приводит к узлу $\{F(s,t),\ s \in S^1\}$, являющемся полигональным. Далее мы требуем, чтобы существовало вещественное число $\delta > 0$ такое, что расстояние между каждой парой вершин полигонального узла $\{F(s,t):\ s \in S^1\}$ было не меньше чем δ для каждого $t \in [0,1]$. Это требование позволяет избежать процесса стягивания узла в тривиальный, как показано на следующем рисунке.

$$\longrightarrow \qquad\qquad\qquad \longrightarrow$$

Знаменитая проблема узлов состоит в поиске алгоритма, который для двух узлов K_0 и K_1 определяет, эквивалентны они или нет. Исходя из этого ясно, что указанная проблема — задача топологии. Эмиль Артин в 1925 году ввел группу кос, инициируя алгебраический подход для решения проблемы узлов. Группа кос с n нитями может быть определена как абстрактная группа B_n, задаваемая образующими $\sigma_1, \ldots, \sigma_{n-1}$ и определяющими соотношениями $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$, $i = 1, \ldots, n - 2$, $\sigma_i\sigma_j = \sigma_j\sigma_i$, $i, j = 1, \ldots, n-1$, $|i-j| > 1$. Всякая коса с n нитями записывается как слово в алфавите $\sigma_1^{\pm 1}, \sigma_2^{\pm 1}, \ldots, \sigma_{n-1}^{\pm 1}$. При этом образующей σ_i соответствует элементарная коса, в которой только i-ая и $i+1$-я нити однажды меняются местами, а степень ± 1 образующей σ_i указывает способ прохождения i-й нити над или под $i+1$-й нитью. Если коса $\sigma \in B_n$, то соединив каждый верхний конец косы σ с ее нижним концом, получим зацепление $L(\sigma)$, т. е. объединение конечного числа непересекающихся узлов. при этом $L(\sigma)$ называется замыканием косы σ. Дж. Александер в 1925 году доказал, что произвольное зацепление (в частности узел) есть замыкание некоторой косы.

Следующий важный шаг на пути алгебраизации проблемы узлов сделал в 1935 году А. А. Марков. Согласно Маркову, все сводится к трем операциям над косами. Первая операция M_1 — сопряжение в группе кос B_n, т. е., если α произвольная коса из B_n, то $M_1(\alpha)$ есть коса вида $\beta^{-1}\alpha\beta$, где $\beta \in B_n$. Вторая операция M_2 — набрасывание петли, т. е. переход от косы $\alpha \in B_n$ к косе $\alpha\sigma_n \in B_{n+1}$. Третья операция — сбрасывание петли, т. е. переход от косы $w(\sigma_1, \ldots, \sigma_{n-2})\sigma_{n-1}^{\pm 1}$ к косе $w(\sigma_1, \ldots, \sigma_{n-2})$. Здесь $w(\sigma_1, \ldots, \sigma_{n-2})$ — произвольное слово в образующих $\sigma_1^{\pm 1}, \ldots, \sigma_{n-2}^{\pm 1}$. Марков доказал, что два зацепления $L(\alpha)$ и $L(\beta)$, где $\alpha \in B_n$, $\beta \in B_m$, эквивалентны тогда и только тогда, когда от косы α к косе β можно перейти конечным числом операций M_1, M_2, M_3. Таким образом, все сводится к следующей задаче (алгебраическая проблема узлов). Даны две косы $\alpha \in B_n$, $\beta \in B_m$. Найти алгоритм, который определяет, существует или нет конечная последовательность кос

$$\alpha = \alpha_0, \alpha_1, \ldots, \alpha_k = \beta,$$

такая, что для каждого i, $i = 1, \ldots, k$, α_{i+1} есть образ косы α_i при одном из отображений M_1, M_2, M_3.

До сих пор ответ на вопрос, поставленный в алгебраической проблеме узлов, не получен. Следующие примеры подчеркивают трудность в ее решении.

Пример 1. Коса

$$\beta = \sigma_3^{-2}\sigma_2\sigma_3^{-1}\sigma_2\sigma_1^3\sigma_2^{-1}\sigma_1\sigma_2^{-1}$$

не сопряжена косе вида $w(\sigma_1, \sigma_2)\sigma_3^{\pm 1}$, где $w(\sigma_1, \sigma_2)$ — слово в образующих $\sigma_1^{\pm 1}$, $\sigma_2^{\pm 1}$, тем не менее замыкание $L(\beta)$ косы β есть узел, эквивалентный тривиальному узлу. Другими словами, в этом случае (коса на четырех нитях) приходится сначала набрасывать петлю на одну из кос, сопряженных косе β, чтобы в конце концов придти к единичной косе на одной нити.

Говорят, что замыкание $L(\sigma)$ косы σ имеет *косовый индекс* n, если $\sigma \in B_n$, но $L(\sigma)$ не эквивалентно $L(\alpha)$ ни для какой косы $\alpha \in B_{n-1}$.

Из фактов, приведенных в примере 1, ясно, что косовый индекс для $L(\beta)$ равен 1. Следует подчеркнуть, что задача вычисления косового индекса для $L(\alpha)$, где α — произвольная коса, также не решена.

Пример 2. Пусть

$$\gamma_r = \sigma_1^{-1}\sigma_2^{2r}\sigma_1^{-2r}\sigma_2, \quad \gamma_r' = \sigma_1\sigma_2^{-2r}\sigma_1^{2r}\sigma_2^{-1},$$

где $r = 1, 2, 3, \ldots$.

Тогда обе косы γ_r и γ_r' имеют косовый индекс 3, они не являются сопряженными в B_3. При этом $L(\gamma_r)$ эквивалентно $L(\gamma_r')$.

О построении курса "Алгебраическая алгоритмика" для специальности "Компьютерная безопасность"

С. И. Яблокова

Обсуждаются некоторые особенности программы курса "Алгебраическая алгоритмика".

Библиография: 2 названия.

Программа курса "Алгебраическая алгоритмика" включает в себя широкий круг разнообразных вопросов прикладной алгебры таких, как расширенный алгоритм Евклида для чисел и многочленов с оценкой сложности этого алгоритма, теорию сравнений, включая китайскую теорему об остатках для чисел и многочленов, некоторые вопросы теории чисел, теории групп, поля Галуа, а также несколько быстрых алгоритмов цифровой обработки сигналов.

Курс насыщен новыми и важными понятиями и совсем не простыми утверждениями. Читается же он во втором и третьем семестрах, т. е. для студентов первого и второго курсов, еще только осваивающих курс алгебры, поэтому он требует постепенного погружения обучающихся в предмет. Для студентов первого набора этой специальности курс читался в 6 и 7 семестрах, т. е. на третьем и четвертом курсах, и тогда за основу была взята книга [2]. Даже для студентов старших курсов это не простой учебник. Книга написана для специалистов, в ней широко используются такие алгебраические структуры как факториальные, евклидовы и квазиевклидовы кольца, кольца главных идеалов, модули, фактор-кольца и т. п. Для студентов младших курсов эта книга сложна, у них еще недостаточно математической культуры для того, чтобы ее освоить, поэтому за основу была взята книга [1]. Предлагаемый студентам курс в основном основан на этом учебнике, хотя для изучения некоторых вопросов привлекается и другая литература.

Рассмотрим некоторые особенности построения программы курса. Курс начинается с вопросов, знакомых слушателям по курсу алгебры: наибольший общий делитель чисел, его представление и свойства, алгоритмы нахождения наибольшего общего делителя. Однако, в отличие от курса алгебры, здесь алгоритм Евклида рассматривается для различных правил деления (остаток от деления положителен – классический вариант; остаток от деления наименьший по модулю – центрированное деление; знак остатка совпадает со знаком делимого, со знаком делителя). Все эти способы деления сравниваются для выяснения, какой из них наиболее эффективен по числу требуемых операций деления. Затем выводятся теоремы о сложности алгоритма Евклида в классическом случае (теорема Ламе) и в случае центрированного деления. В ходе доказательства теоремы Ламе приходится вспомнить и о числах Фибоначчи, и найти решение бесконечной последовательности рекуррентных уравнений, и вывести формулы Бинэ. Аналогичная оценочная теорема в случае центрированного деления для своего доказательства требует сходных рассуждений, поэтому ее доказательство разбивается на части, которые даются студентам в качестве теоретических задач на дом. На следующей лекции тот, кто справился с задачами, докладывает свои результаты аудитории. Как правило, на каждом курсе находятся слушатели, способные решить такого рода задачи.

В дальнейшем курс вновь обращается к алгоритму Евклида, рассматривая расширенный алгоритм, который кроме наибольшего общего делителя позволяет вычислять также коэффициенты Безу. Доказывается оценочная теорема о величине этих коэффициентов, если они получены в ходе работы расширенного алгоритма Евклида.

Как полезное приложение алгоритма Евклида можно рассматривать следующую тему курса — "Цепные (или непрерывные) дроби". Цепные дроби потребуются студентам специальности "Компьютерная безопасность" в дальнейшем при изучении такого курса, как "Теоретико-числовые методы в криптографии", и хотя теорему о фундаментальном соответствии давать на первом курсе рано, но понятие цепной дроби, ее свойства, теоремы о представлении рациональных чисел в виде цепной дроби, о точности приближения подходящими дробями вполне могут быть усвоены слушателями.

Вопрос о разложении целого числа на простые множители позволяет вспомнить основную теорему арифметики и обобщить результат, вводя евклидовы кольца и доказав теорему о факториальности таких колец. В качестве полезного и наглядного примера рассматривается кольцо целых гауссовых чисел $\mathbb{Z}[i]$, доказывается его факториальность и выводится ряд полезных теорем о неприводимости элементов в этом кольце. Обсуждаются также самые простые алгоритмы разложения целых чисел на простые множители.

К вопросам из теории чисел относятся такие несложные, но важные теоремы, как малая теорема Ферма, теорема Эйлера, теорема Вильсона. Все они имеют сравнительно несложные доказательства и интересные и полезные применения. Естественно возникает необходимость ввести понятие сравнения и вспомнить свойства сравнений. Теория групп и теория чисел здесь тесно переплетаются. Понятие группы уже знакомо слушателям, но еще недостаточно ими освоено, поэтому повторяются основные определения и теоремы. Теория сравнений приводит к введению колец вычетов и групп их обратимых элементов. Введение в теорию сравнений заканчивается китайской теоремой об остатках. Эта теорема сначала доказывается в самом простом ее варианте для прямого произведения двух конечных абелевых групп $\mathbb{Z}_m$ и $\mathbb{Z}_n$ взаимно простых порядков, затем результат переносится на кольца вычетов и группы их обратимых элементов. После этого теорема формулируется в терминах решения системы линейных сравнений по взаимно простым модулям. Такой подход позволяет постепенно усложнять материал, кроме того, дает возможность получить два способа решения системы сравнений.

Обращаясь опять к теории групп курс постепенно подводит слушателей к одной из самых сложных теорем курса – теореме Гаусса, дающей необходимые и достаточные условия цикличности мультипликативных групп колец вычетов.

В качестве приложений изученной теории рассматриваются некоторые алгоритмы проверки чисел на простоту, разделенные на две группы: точные и вероятностные. При оценке вероятности успеха работы алгоритмов второй группы могут возникать некоторые трудности, поскольку студенты еще не знакомы с теорией вероятности. Приходится вводить понятие вероятности на интуитивном уровне. Но оставить курс совсем без таких приложений изученной теории для решения конкретных теоретико-числовых задач представляется неразумным. Обучаемый должен понимать, зачем ему нужны изученные ранее теоремы, как они могут быть использованы.

Алгоритмы факторизации чисел являются достаточно сложными, и их обоснование требует более широких математических знаний, нежели знания первокурсника, поэтому в курсе приводится в качестве примера всего один такой алгоритм.

В заключение первой части курса слушатели знакомятся с модулярной арифметикой, ее использованием в компьютерных вычислениях.

Следующий семестр начинается с возвращения к методам и понятиям, уже знакомым по первой части курса, но вместо кольца целых чисел студентам предлагается работать с кольцом многочленов над полем (и над факториальным кольцом). Алгоритм Евклида и расширенный алгоритм Евклида применяются к многочленам, оценивается сложность алгоритма. Вопрос об интерполяции многочлена по набору его значений в различных точках позволяет вспомнить и уже известные интерполяционные формулы, и получить новый алгоритм интерполяции, основанный на китайской теореме об остатках для многочленов в ее частном случае, когда взаимно простые модули есть линейные двучлены. Вводится понятие сравнимости по модулю многочлена и доказывается теорема о факторкольце $K[x]/(m(x))$, которая в случае неприводимости многочлена $m(x)$ в $K[x]$ дает способ построения новых полей. Китайская теорема об остатках формулируется в терминах многочленов.

Вопрос разложения многочлена на множители связан с понятием неприводимого многочлена над данным полем. Сначала слушателям предлагается вспомнить те теоремы о неприводимости многочленов над полями комплексных и действительных чисел, которые им известны из курса алгебры. Попутно устанавливается факториальность кольца многочленов над полем. Затем разбирается вопрос о неприводимости многочленов над полем рациональных и кольцом целых чисел. Показывается, что неприводимость в кольце $\mathbb{Z}[x]$ эквивалентна неприводимости в $\mathbb{Q}[x]$. Наконец, доказывается достаточный критерий неприводимости многочленов в $\mathbb{Z}[x]$. Далее рассматриваются неприводимые многочлены над $\mathbb{Z}_p[x]$. Вновь происходит обращение к теореме о факториальном кольце, но на этот раз упор делается на то, что $\mathbb{Z}_p[x]/(m(x))$ является линейным пространством над полем $\mathbb{Z}_p$ и состоит из конечного числа элементов. В это время слушатели уже познакомились с понятием линейного пространства в курсе линейной алгебры. Следуют теоремы, позволяющие найти количество неприводимых унитарных многочленов данной степени над полем $\mathbb{Z}_p$ и получить критерий неприводимости многочлена над конечным полем. В качестве полезного приложения рассматривается "решето Эратосфена" для многочленов над конечным полем.

Следующий важный раздел курса – поля Галуа, где опять происходит обращение к теореме о факториальном кольце $\mathbb{Z}_p[x]/(m(x))$ в случае, когда $m(x)$ – неприводимый многочлен над $\mathbb{Z}_p$. В этом разделе доказываются основные теоремы о конечных полях, вводятся понятия характеристики поля, примитивного элемента поля Галуа, минимального многочлена алгебраического элемента, дается способ построения такого многочлена. Слушатели знакомятся с рядом теорем, позволяющих разложить многочлен вида $x^{p^n} - x$ в произведение неприводимых в $\mathbb{Z}_p$. В качестве примера приводится способ построения поля Галуа из 2^n элементов и строится поле $GF(2^4)$.

Последняя часть курса посвящена быстрым алгоритмам цифровой обработки сигналов. В предлагаемых слушателям быстрых алгоритмах вычисления линейной и циклической сверток происходит обращение и к интерполяционным формулам (алгоритм Кука – Тоома), и к китайской теореме об остатках для многочленов (алгоритм Винограда). Быстрые алгоритмы вычисления дискретного преобразования Фурье, рассматриваемые далее, используют свойства мультипликативной группы поля Галуа и мультипликативных групп колец вычетов (алгоритм Рейдера, БПФ — алгоритм Винограда).

Список литературы

[1] *Акритас А.* Основы компьютерной алгебры с приложениями. М.: Мир, 1994. 544 с.

[2] *Ноден П., Китте К.* Алгебраическая алгоритмика. М.: Мир, 1999. 720 с.

Научное издание

Преподавание математики и компьютерных наук в классическом университете

Материалы 2-й научно-методической конференции преподавателей математического факультета и факультета информатики и вычислительной техники Ярославского государственного университета им. П. Г. Демидова

Компьютерный набор авторов.
Редактор, корректор А. А. Аладьева.
Компьютерная вёрстка М. В. Невского.

Подписано в печать 12.02.2007. Формат $60 \times 84^1/_8$. Печать офсетная.
Усл. печ. л. 20,46. Уч.-изд. л. 10,5. Тираж 100 экз. Заказ 069.

Оригинал-макет подготовлен в редакционно-издательском отделе Ярославского государственного университета.

Ярославский государственный университет им. П. Г. Демидова.
150000 Ярославль, ул. Советская, 14.

Отпечатано
ООО "Ремдер", ЛР ИД № 06151 от 26.10.2001.
г. Ярославль, пр. Октября, 94, офис 37
тел. (4852) 73-35-03, 58-03-48, факс 58-03-49.